



Technische Universität München
Fakultät für Sport- und Gesundheitswissenschaften

Informatische Methoden im Sport

Simulation, Data-Mining und Machine-Learning

Sebastian Georg Wenninger, M.Sc.

Vollständiger Abdruck der von der Fakultät für Sport- und
Gesundheitswissenschaften der Technischen Universität München zur
Erlangung des akademischen Grades eines

Doktors der Philosophie (Dr. Phil.)

genehmigten Dissertation.

Vorsitzender: Priv.-Doz. Dr. Felix Ehrlenspiel

Prüfer der 1. Prof. Dr. Martin Lames

Dissertation: 2. Prof. Dr. Arnold Baca

Die Dissertation wurde am 05.11.2021 bei der Technischen Universität München eingereicht
und durch die Fakultät für Sport und Gesundheitswissenschaften am 28.03.2022 angenommen.

Inhaltsverzeichnis

Abbildungsverzeichnis	II
Abkürzungsverzeichnis	IV
Tabellenverzeichnis	VI
Zusammenfassung	VII
Auflistung der eingebundenen Publikationen	VIII
1 Einleitung	1
1.1 Einordnung in die Informatik	2
1.2 Einordnung in die Trainingswissenschaft	7
1.2.1 Wettkampfdiagnostik in Sportspielen	9
1.2.2 Wettkampfdiagnostik in den behandelten Sportarten	11
1.3 Motivation	14
2 Methoden	17
2.1 Simulation	17
2.1.1 Markov-Ketten	18
2.1.2 Simulation mit Markov-Ketten	19
2.1.3 Numerische Differentiation	22
2.2 Data-Mining	23
2.2.1 Frequent Itemset Mining	24
2.2.2 Association Rule Mining	25
2.3 Machine-Learning	26
2.3.1 Neuronale Netze	27
2.3.2 Supervised-Learning	29
2.3.3 Convolutional Neural Networks	30
2.3.4 Recurrent Neural Networks	31
2.3.5 Entscheidungsbäume	32
3 Einzelbeiträge	35
3.1 Studie 1 - Performance Analysis in Table Tennis - Stochastic Simulation by Numerical Derivation	35
3.2 Studie 2 - Data Mining in Elite Beach Volleyball	36
3.3 Studie 3 - Performance of machine learning models in application to beach volleyball data	37
4 Diskussion	40
5 Fazit	45
Literatur	47
A Anhang	60
A.1 Genehmigung zur Veröffentlichung der Publikationen	60
A.2 Studie 1 - Originalveröffentlichung	61
A.3 Studie 2 - Originalveröffentlichung	76
A.4 Studie 3 - Originalveröffentlichung	95

Abbildungsverzeichnis

Abbildung 1	Strukturmodell der informatischen Disziplinen. Angelehnt an Claus et al. (1975).	3
Abbildung 2	Teilgebiete der Informatik und Anwendungsfelder in der Sportwissenschaft als Schnittstelle der Sportinformatik. Grafik entnommen aus Link und Lames (2009)	4
Abbildung 3	Pyramidenmodell der komplexen Sportspielleistung. Grafik entnommen aus Hohmann et al. (2020)	8
Abbildung 4	Markov-Kette mit 3 Zuständen und der Übergangsmatrix T . Grafik entnommen aus Grewal et al. (2019).	19
Abbildung 5	Betrag der Auslenkung in Abhängigkeit der Übergangswahrscheinlichkeit. Grafik entnommen aus Lames (1991).	21
Abbildung 6	Schritte des Data-Mining Prozesses nach Runkler (2010).	24
Abbildung 7	Schematische Darstellung eines einfachen neuronalen Netzwerks, auch Multilayer-Perceptron (MLP) oder fully connected network genannt.	28
Abbildung 8	Beispiel eines einzelnen Neurons mit den Eingangsdaten ($x_1 - x_n$), den korrespondierenden Gewichten ($w_1 - w_n$), dem Bias b und der Aktivierungsfunktion f , die auf die gewichtete Summe der Eingaben angewandt wird. („Activation functions and its Types in Artificial Neural network“, 2018).	28
Abbildung 9	Die Rectified Linear Unit (ReLU) Funktion. Grafik entnommen aus Goodfellow et al. (2016).	29
Abbildung 10	Beispielhafte Darstellung der Faltung der Eingabewerte $a - l$ mit dem Kernel $w - z$. Der Kernel wird sowohl auf der horizontalen wie auch auf der vertikalen Achse über die Eingabedaten angewandt. Grafik entnommen aus Goodfellow et al. (2016).	31
Abbildung 11	Architektur des LeNet-5, eines der ersten CNNs zur Bilderkennung. Grafik entnommen aus LeCun et al. (1998).	31
Abbildung 12	Links: Ein Neuronales Netz, das eine Schleife enthält. Rechts: Aufgerollte Form desselben neuronalen Netzes für eine Eingabesequenz der Länge t . Grafik entnommen aus Olah (2015).	32

Abbildung 13 Links: Partitionierte Daten. Rechts: Entscheidungsbaum für Klassifikationsmodell mit den Klassen 1, 2 und 3. An jedem Zwischenknoten wird der linke Kindknoten ausgewählt, falls die Bedingung des Knotens zutrifft. Andernfalls wird der rechte Kindknoten ausgewählt. Die vorhergesagte Klasse ist an den Blättern des Baumes dargestellt. Grafik entnommen aus Loh (2011).	33
Abbildung 14 Schematische Darstellung eines Tree-Ensembles. Die finale Klassifizierung erfolgt durch Votum der einzelnen Entscheidungsbäume. Grafik entnommen aus Shoaran et al. (2018).	34
Abbildung 15 Beispiel für einen durch <code>lr_find()</code> generierten Graphen. Grafik entnommen aus fast.ai - Dokumentation, 2020.....	43

Abkürzungsverzeichnis

A

ANOVA - Analysis of Variance	38
ARM - Association-Rule-Mining.....	25, 36, 37, 41

B

BPTT - Backpropagation through time.....	32
BSN - Body Sensor Networks.....	6

C

CNN - Convolutional Neural Network.....	30, 31, 43
---	------------

D

DL - Deep Learning.....	26
DSP - Digital Signal Processing.....	7

E

EPTS - Electronic Performance and Tracking Systems	2
ETTU - European Table Tennis Union.....	35

F

FIM - Frequent Itemset Mining.....	24, 25
------------------------------------	--------

G

GRU - Gated Recurrent Unit.....	32
---------------------------------	----

I

IMU - Inertial Measurement Unit	6
IOC - International Olympic Committee.....	35
ITTF - International Table Tennis Federation	35

K

KDD - Knowledge Discovery in Databases	23
KI - Künstliche Intelligenz.....	26

L

LSTM - Long-Short-Term-Memory.....	32
------------------------------------	----

M

MEMS - Microelectromechanical Systems.....	6
MLP - Multilayer-Perceptron.....	28, II
MPW - Mutual-Point-Winning-Probabilities	41

P

PLD - Praktische Leistungsdiagnostik..... 9, 45

R

ReLU - Rectified Linear Unit..... 28–30, II

RNN - Recurrent Neural Network..... 32

S

Std. - Standard..... 36–38

T

TLD - Theroretische Leistungsdiagnostik..... 8, 45

X

XAI - Explainable Artificial Intelligence..... 46

Tabellenverzeichnis

Tabelle 1 Beispiel für eine Datenbank von Kundentransaktionen mit der Transaction ID (TID) und dem Itemset der Transaktion. Die erste Transaktion repräsentiert einen Kunden, der die Items a , c und d gekauft hat. Entnommen aus Fournier-Viger et al. (2017).	25
--	----

Zusammenfassung

Die sportliche Leistung in Sportspielen ist geprägt von den komplexen dynamischen Interaktionen zwischen den Athleten. Um dieser Komplexität beizukommen, werden in der Sportwissenschaft zunehmend informatische Methoden eingesetzt. Begünstigt wird diese Entwicklung durch den technischen Fortschritt, der es ermöglicht, immer größere Datenmengen von Mannschaften und Spielern erfassen zu können, die schließlich die Grundlage für leistungsdiagnostische Analysen bilden. Eine Aufgabe der Sportinformatik ist dabei, bei der Erfassung, Verarbeitung und Auswertung dieser Daten zu unterstützen und dadurch sportwissenschaftlichen Mehrwert zu generieren.

Hier setzt diese Dissertation an, wobei ein besonderer Fokus auf das Problem der Mustererkennung innerhalb von Spielzügen in den bisher von der Sportinformatik wenig beachteten Sportarten Tischtennis und Beachvolleyball gelegt wird. Dazu kommen Methoden aus dem Feld der Simulation, des Data-Mining und des Machine-Learning zum Einsatz. So wird in der ersten Studie die Methode der Markov-Simulation im Tischtennis um die Simulation durch numerische Differentiation erweitert, wodurch effizientere leistungsdiagnostische Analysen ermöglicht werden. In der zweiten Studie werden mit Hilfe von Assoziationsanalysen sequenzielle Muster in Beachvolleyball-Spielzügen gesucht, die sowohl zur Trainingssteuerung als auch zur Gegneranalyse eingesetzt werden können. Dieser Ansatz wird in der dritten Studie mit mehreren Methoden des Machine-Learning weiterverfolgt, die zudem bezüglich ihrer Leistungsfähigkeit bei der Mustererkennung verglichen werden.

Zusammengefasst trägt diese Dissertation durch die Aufklärung innerer Zusammenhänge in Spielzügen zum Einen zu einem besseren Verständnis der Struktur der untersuchten Sportarten bei, wodurch Rückschlüsse für die praktische Leistungsdiagnostik gezogen werden können. Zum Anderen leistet sie einen sportinformatischen Beitrag durch die Untersuchung und den Vergleich informatischer Methoden in ihrer Anwendung im Sport.

Auflistung der eingebundenen Publikationen

Bei der vorliegenden Arbeit handelt es sich um eine kumulative Dissertation, wobei drei Full-Paper (peer-reviewed) aus englischsprachigen, international verbreiteten Publikationsorganen eingebunden wurden:

1. Wenninger, S., & Lames, M. (2016). Performance Analysis in Table Tennis - Stochastic Simulation by Numerical Derivation, *International Journal of Computer Science in Sport*, 15(1), 22-36. doi: <https://doi.org/10.1515/ijcss-2016-0002>
2. Wenninger, S., Link, D., & Lames, M. (2019). Data Mining in Elite Beach Volleyball – Detecting Tactical Patterns Using Market Basket Analysis, *International Journal of Computer Science in Sport*, 18(2), 1-19. doi: <https://doi.org/10.2478/ijcss-2019-0010>
3. Wenninger, S., Link, D., & Lames, M. (2019). Performance of machine learning models in application to beach volleyball data., *International Journal of Computer Science in Sport*, 19(1), 24-36. doi: <https://doi.org/10.2478/ijcss-2020-0002>

1. Einleitung

„Information is the oil of the 21st century and analytics is the combustion engine.“
- Peter Sondergaard (Gartner Research)

Daten treiben die Welt an. Laut dem Magazin Forbes sind allein in den Jahren 2016 bis 2018 mehr als 90 Prozent aller jemals erfasster Daten generiert worden (Marr, 2018). Diese Entwicklung zieht sich durch alle Lebensbereiche und spiegelt sich auch in der Sportwissenschaft wider. Hier können sowohl im Training als auch während des Wettkampfs Daten in einer Menge und Genauigkeit erfasst werden, die vor ein paar Jahren noch undenkbar waren. Zur Verarbeitung dieser Daten hat sich, ähnlich wie in anderen Disziplinen (z.B. Medizininformatik, Bioinformatik, Geoinformatik), auch in der Sportwissenschaft eine Unterdisziplin der Informatik, die Sportinformatik, gebildet. Hierbei handelt es sich um ein hochgradig interdisziplinäres Feld, das die vielfältigen Aspekte der Sport- und Informationswissenschaften vereint. Link und Lames (2009, S. 26) formulieren das Selbstverständnis dieser Disziplin folgendermaßen:

„Die Sportinformatik ist ein Verbund multi- und interdisziplinärer Forschungsprogramme im Schnittstellenbereich zwischen Sportwissenschaft und Informatik. Gegenstandsbereiche sind die Anwendung informatischer Werkzeuge, Methoden und Konzepte auf Fragestellungen des Sports und der Sportwissenschaft sowie die Integration sportwissenschaftlicher Erkenntnisse in Themenfelder der Informatik.“

Die Ursprünge der Sportinformatik gehen auf die Notwendigkeit zurück, den Menschen bei der Aufnahme, Verarbeitung, Speicherung, Analyse und Präsentation von Informationen zu unterstützen (Perl et al., 1997). Dabei ist die Informatik bereits seit Mitte der 60er-Jahre ein wichtiger interdisziplinärer Partner für die Sportwissenschaft (Lees, 1985). Während anfänglich mehrheitlich statistische Analysen sowie biomechanische Messungen durchgeführt wurden, ermöglichte der Fortschritt der Technik ab den 80ern den Einsatz von Computern auch zur Spielanalyse oder Modellierung von Spielsportarten (Hughes, 1985; Miethling & Perl, 1981).

Mit zunehmender Leistungsfähigkeit der Computer ergaben sich weitere Anwendungsbereiche, gleichzeitig aber auch zusätzliche Herausforderungen durch neu entstehende Technologien wie das Internet, digitale Videobilder oder Inertialsensoren (Perl, 2006). Der damit einhergehende stetige Anstieg an verfügbaren Daten trug dazu bei, dass die Sportinformatik eine zunehmend größer werdende Rolle in der Sportwissenschaft einnahm und heute einen essentiellen Bestandteil der sportwissenschaftlichen Gemeinschaft darstellt.

Gleichzeitig besteht seitens der Allgemeinheit (Fans, Verbände, Vereine) und damit natürlich auch durch die Medien ein gestiegenes Interesse am Sport. Dieses Interesse nutzen selbst international agierende Technologieunternehmen wie Amazon (Amazon, 2020), IBM

(Kumar, 2015) oder Google (Google, 2020), die den Sport als Demonstrationsfläche der Leistungsfähigkeit der eigenen Technologien verwenden. Dabei bietet der Sport die Möglichkeit, selbst komplexe informatische Methoden an verständlichen Anwendungsbeispielen zu demonstrieren und somit zugänglicher zu machen. Ein großer Treiber für diese Entwicklung ist die zunehmende Verfügbarkeit kontinuierlicher Positionsdaten, die durch den Einsatz sogenannter Electronic Performance and Tracking Systems (EPTS) erfasst werden können. Im Gegensatz zu populären Sportarten wie Fußball oder den amerikanischen Sportarten wie Basketball oder Baseball, in denen die Erfassung solcher Positionsdaten während Training und Wettkampf schon alltäglich ist, fehlt in den in dieser Arbeit behandelten Sportarten diese Datenbasis noch (Buchheit & Simpson, 2017; Sampaio et al., 2015). Die notwendigen Technologien dazu sind zwar vorhanden, deren Einführung hängt jedoch von einer Vielzahl komplexer Faktoren ab. So müssen neben finanziellen Hürden auch (sport-)politische Hindernisse überwunden werden. Gerade deswegen ist es die Aufgabe der Wissenschaft, empirisch zu validieren, dass sportinformatische Analysen einen entscheidenden Beitrag zur Leistungsdiagnostik leisten können (Ofoghi et al., 2013).

Dieses Ziel soll in der vorliegenden Arbeit verfolgt werden. Durch den Einsatz informatischer Methoden aus den Bereichen der Simulation, des Data-Mining und Machine-Learnings soll deren Anwendung geprüft und dadurch der Einsatz dieser Methoden in der Sportwissenschaft geschärft werden.

Das einleitende Kapitel ist weiter wie folgt strukturiert: In den Abschnitten 1.1 und 1.2 wird die vorliegende Arbeit systematisch in die Themengebiete der Informatik und Trainingswissenschaft eingebettet. Anschließend wird das in dieser Arbeit behandelte Anwendungsfeld der sportinformatischen Analysen - die Leistungsdiagnostik in den Spilsportarten - genauer beschrieben.

1.1. Einordnung in die Informatik

Wie bereits am Begriff zu erkennen ist, handelt es sich bei der „Informatik“ um diejenige Wissenschaft, die sich mit der maschinellen Verarbeitung von Informationen beschäftigt. Um dieses breit gefächerte Feld zu strukturieren, wird die Informatik im Allgemeinen in mehrere Teilgebiete gegliedert (siehe Abbildung 1).

Die Theoretische Informatik beschäftigt sich mit den abstrakten mathematischen und logischen Grundlagen der Informatik und bildet damit den Grundstein aller anderen Teilgebiete (Gumm & Sommer, 2010). Ihre Themengebiete umfassen unter anderem die Automatentheorie, formale Sprachen sowie Fragen nach der Berechenbarkeit und Komplexität von Problemen. Zum Einsatz kommen diese zum Beispiel beim Bau von Compilern - Programmen, die Quelltext einer meist höheren Programmiersprache in Zielcode (meistens Maschinencode) übersetzen.

Den formalen Rahmen der Theoretischen Informatik nutzen die Technische und Praktische Informatik.

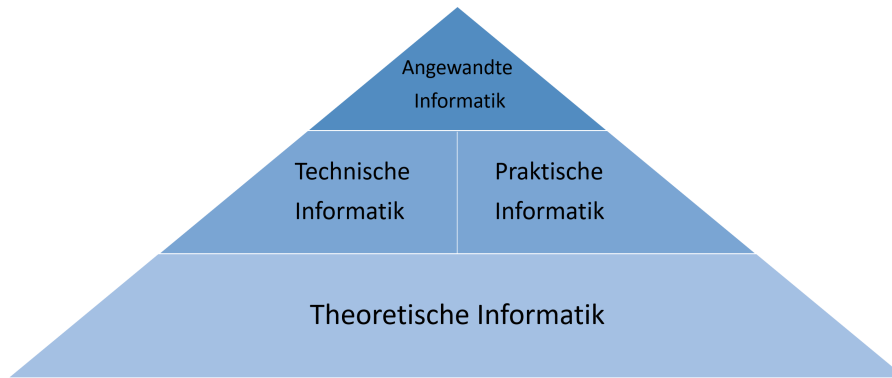


Abbildung 1 Strukturmodell der informatischen Disziplinen. Angelehnt an Claus et al. (1975).

Die Technische Informatik beschäftigt sich mit den Hardware-nahen Problemen der Informatik, also zum Beispiel der Konstruktion von Rechnern, Prozessoren und Speichermedien. Aber auch die systemnahe Programmierung gehört zu ihrem Kompetenzgebiet. Sie bildet die Schnittstelle zwischen den Fachbereichen der Informatik und der Elektrotechnik.

In der Praktischen Informatik zeigt sich der ingenieurwissenschaftliche Charakter der Informatik (Biundo et al., 2006). Sie entwickelt Lösungen für konkrete informatische Probleme, beispielsweise durch die Entwicklung von Algorithmen und Datenstrukturen. Ein prominentes Teilgebiet der Praktischen Informatik stellt die Softwaretechnik dar. Hier werden Konzepte und Methoden für die systematische Entwicklung und den Betrieb von Softwaresystemen entwickelt. Die drei Gebiete der Theoretischen, Technischen und Praktischen Informatik werden häufig unter dem Begriff „Kerninformatik“ zusammengefasst. Selbst wenn die hier gezeigte Strukturierung diese Disziplinen der Informatik strikt getrennt darstellt, muss angemerkt werden, dass sie natürlich nicht isoliert voneinander existieren. Die Grenzen zwischen den Teilgebieten verlaufen häufig fließend (Gumm & Sommer, 2010). Der bereits erwähnte Compilerbau existiert an der Grenze zwischen Theoretischer und Praktischer Informatik. Diese wiederum hat Schnittmengen mit der Technischen Informatik, da zum Beispiel die Entwicklung von Datenstrukturen selbstverständlich Abhängigkeiten zu der verwendeten Hardware hat.

Im Kontrast zur Kerninformatik steht die Angewandte Informatik, die sich mit der Anwendung von Methoden der Kerninformatik in anderen Wissenschaften befasst (Claus, 2006). Dies unterscheidet sie von der Praktischen Informatik, die sich mit informatischen Problemstellungen beschäftigt. Wie bereits erwähnt hat sich dazu eine Vielzahl von interdisziplinären Teilgebieten gebildet, in denen informatische Methoden mit fundierten Kenntnissen der Anwendungsgebiete verbunden werden, um fachspezifische Problemstellungen lösen zu können. Auch die Sportinformatik kann klar der Angewandten Informatik zugeordnet werden, da sie sich explizit mit Fragen der Anwendung informatischer Methoden in einem Kooperationsbereich der Informatik befasst.

Eine weitere Strukturierung der Angewandten Informatik kann nun anhand der behandelten Themen erfolgen. Dabei wird hier nur ein Ausschnitt des breiten Spektrums informatischer Anwendungsfelder gezeigt, der versucht, einen großen Teil der in der Sportinformatik relevanten Themen abzudecken, aber keinen Anspruch auf Vollständigkeit erhebt. Nach Baca

(2006) und Link und Lames (2009) lassen sich die informatischen Themenbereiche mit Anwendungspotenzial in der Sportwissenschaft wie folgt zusammenfassen (siehe auch Abbildung 2):

- Datenerfassung, -verarbeitung und -analyse
- Modellbildung und Simulation
- Präsentation und Visualisierung
- Netzwerke und Kommunikation

Die Studien dieser Arbeit fallen hierbei klar in die Themenfelder *Datenanalyse* und *Modellbildung und Simulation*, weshalb diese in den nächsten Abschnitten ausführlicher beleuchtet werden.

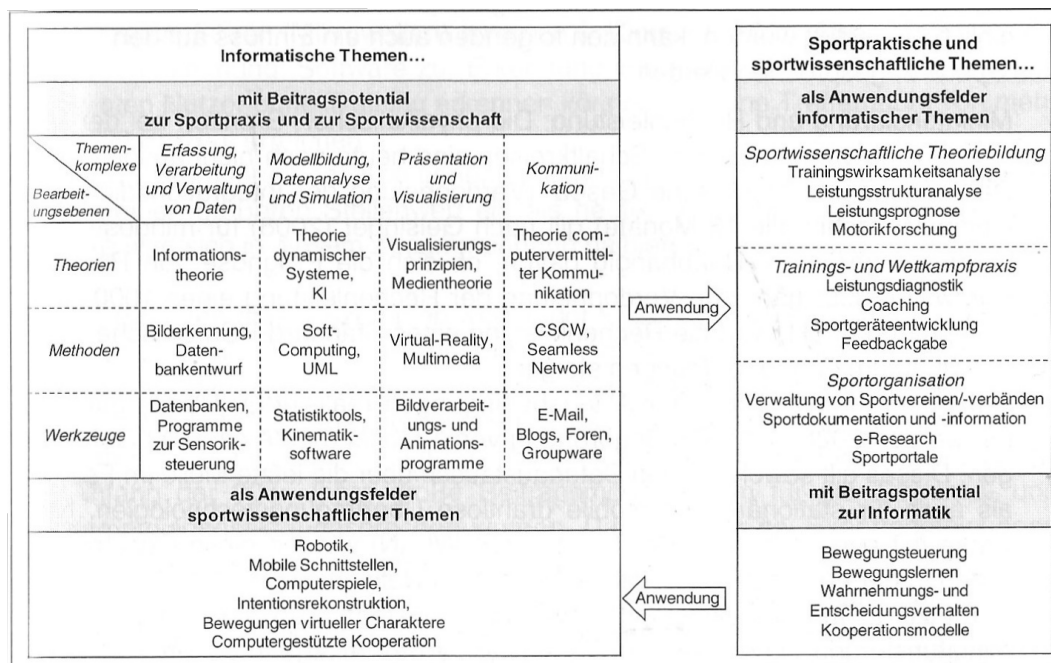


Abbildung 2 Teilgebiete der Informatik und Anwendungsfelder in der Sportwissenschaft als Schnittstelle der Sportinformatik. Grafik entnommen aus Link und Lames (2009)

Modellbildung und Simulation

Die allgemeine Modelltheorie nach Stachowiak (1973) identifiziert folgende wesentliche Merkmale des Begriffs „Modell“:

Das Abbildungsmerkmal besagt, dass Modelle immer Abbildungen eines Originals sind, die selbst wieder zum Modell werden können.

Das Verkürzungsmerkmal formuliert die Tatsache, dass Modelle normalerweise nicht alle Attribute des Originals abbilden.

Das pragmatische Merkmal besagt, dass Modelle den Originalen nicht eindeutig zuzuordnen sind. Sie sind immer Resultat bestimmter Individuen, für einen bestimmten Zeitraum

und zu bestimmten Zwecken.

Das Abundanzmerkmal drückt aus, dass Modelle auch überschüssige, das heißt nicht im Original vorkommende, Merkmale enthalten können.

Um die Abstraktion vom Original zu bewerkstelligen, werden Attribute des Originals entweder übergangen, also nicht im Modell abgebildet, oder neue Attribute dem Modell hinzugefügt (Stachowiak, 1973). Übrig bleiben die Attribute, die den Modellierern für den jeweiligen Zweck relevant erscheinen. Damit ist klar, dass die Modellbildung durch die Ziele und Perspektiven des Modellbilders beeinflusst wird (Steinmüller, 1993).

Ausgehend von einer weit gefassten Definition der allgemeinen Modelltheorie wird die Modellbildung in dieser Arbeit vor allem unter dem Aspekt der mathematischen Modellbildung betrachtet. Bei mathematischen Modellen wird die Abbildung von Original auf Modell durch mathematische Instrumente formalisiert. Beispiele für diese Instrumente enthalten unter anderem (Bungartz et al., 2009):

- algebraische (Un)Gleichungen
- gewöhnliche/partielle Differentialgleichungen
- Wahrscheinlichkeitsverteilungen
- algebraische Strukturen
- Automaten und Zustandsübergangsdigramme
- Graphen

Mit dieser Vielfalt an Modellierungswerkzeugen wird klar, dass die Schwierigkeit der Modellbildung unter anderem auch darin besteht, eine passende mathematische Formalisierung des Originals zu finden. Diese Auswahl kann zum Beispiel in Abhängigkeit von der Struktur der beschreibenden Variablen und der Zeitfunktion des Modells (diskret - stetig) getroffen werden (Lames, 1991).

Modelle selbst dienen nun nicht nur einem rein deskriptiven Zweck, sondern können mit Mitteln der Simulation auch dazu verwendet werden, Erkenntnisse über das Verhalten des abgebildeten Systems zu erlangen. Der zugehörige Prozess der Simulation wird in Kapitel 2 genauer dargestellt.

Datenerfassung, -verarbeitung und -analyse

Um sportliche Leistung beschreiben und analysieren zu können, müssen eine Vielzahl von Parametern erfasst werden (Baca, 2014): Biomechanische Untersuchungen benötigen beispielsweise Reaktionskräfte, kinematische oder elektromyographische Messungen. Positionsdaten von Ball und Spielern ermöglichen die Analyse von taktischen Mustern in Sportarten, während der Sauerstoffgehalt im Blut in leistungsphysiologischen Untersuchungen von Interesse ist. Diese Daten müssen gemessen, verarbeitet und schließlich analysiert werden, um Mehrwert in dem entsprechenden Anwendungsfeld zu erzeugen. Dazu kommen

weitere Anforderungen der Praxis, wie zum Beispiel geringe Latenzen für Echtzeitanwendungen, die wiederum die Anforderungen an das Übertragungsmedium bei der Datenerfassung beeinflussen.

Im weitesten Sinne stellt die Datenerfassung einen Prozess dar, in dem Informationen in maschinenlesbare Form gebracht und auf Datenträgern gespeichert werden (Claus, 1993). Dieser Prozess wird mit geeigneter Messtechnik durchgeführt, die abstrakt dargestellt analoge in digitale Signale umwandelt. Dazu sind im Allgemeinen Sensoren notwendig, die physikalische Signale wie Kräfte oder elektrische Ströme messen können. Im Kontext des Sports gibt es dabei eine Anzahl von Messgrößen, die von erhöhtem Interesse sind (Baca, 2014):

Positionen können durch Triangulierung anhand Kamera-basierter Systeme oder direkt z.B. mit GPS-Sensoren erfasst werden.

Distanzen und Geschwindigkeiten können direkt Z.B. mit Potentiometern und Radarmessungen erfasst oder algorithmisch aus Positionsdaten abgeleitet werden.

Winkel und Winkelgeschwindigkeiten können durch Drehpotentiometer oder Gyroskope erfasst werden.

Kräfte und Druck werden hauptsächlich von Spannungsprüfern, piezo-elektronischen oder kapazitiven Sensoren gemessen.

Beschleunigungen können mit piezo-elektronischen oder kapazitiven Sensoren gemessen werden. Moderne Sensoren können heutzutage als Microelectromechanical Systems (MEMS) gebaut werden, Systeme, die im Bereich von Mikrometern gebaut werden können.

Räumliche Orientierungen werden mit Inertialsensoren gemessen. Eine Inertial Measurement Unit (IMU) kombiniert Gyroskope und Beschleunigungssensoren, um mit Methoden der Sensor-Fusion Messungen der räumlichen Orientierung zu ermöglichen.

Um eine möglichst genaue Erfassung der realen Signale zu gewährleisten, müssen die Messungen mit diesen Sensoren gerade im Sport möglichst unbemerkt von den Akteuren durchgeführt werden. Hier bietet das sogenannte „Ubiquitous computing“ - oder auch „Pervasive computing“ - Lösungen an. Dieses Feld beschäftigt sich mit den Anwendungsmöglichkeiten von Netzen kleiner verbundener Recheneinheiten, die durch den technischen Fortschritt bei Sensoren, Mikroprozessoren und Netzwerken ermöglicht wurden (Hansmann et al., 2003). Im Sport bietet sich besonders der Einsatz sogenannter Body Sensor Networks (BSN) an, Netzwerke autonomer miniaturisierter Sensorknoten, die körpernah getragen werden können und damit minimale Einschränkungen der Träger versprechen (Yang & Yang, 2006).

Aber auch Software-gestützte Datenaufnahmen mittels der in der sportlichen Leistungsdiagnostik weit verbreiteten Beobachtungssysteme (vgl. Abschnitt 1.2.1) können als Messsysteme interpretiert werden, mit den von den Benutzern der Software eingetragenen Daten als digitalem Signal. Nicht zuletzt ist die Informatik heutzutage in der Lage, auch ohne den Einsatz körpernaher Sensoren Leistungsmerkmale zu messen, falls dies beispielsweise aus Kosten- oder Regelgründen nicht möglich ist. So können zum Beispiel mit Methoden der Bilderkennung automatisiert Positionsdaten von Spielern auf einem Fußballfeld aus einem Videosignal extrahiert werden (Liu et al., 2009).

Die Verarbeitung von digitalen Signalen beziehungsweise Daten wird häufig unter dem Begriff Digital Signal Processing (DSP) zusammengefasst. Das DSP transformiert Daten, um sie zweckmäßig verwertbar zu machen. So können diese beispielsweise für eine leichtere Übertragung komprimiert oder zur Verbesserung der Qualität des Signals manipuliert werden. Typische Operationen des DSP beinhalten unter anderem (Baca, 2014):

Skalierung Die gemessenen Daten müssen häufig skaliert werden, zum Beispiel um Umrechnungen der Einheit vorzunehmen. Dazu kann das Signal mit einem Faktor multipliziert und/oder Offsets addiert/subtrahiert werden.

Up/Downsampling Werden mehrere Signale mit unterschiedlichen Messfrequenzen aufgenommen, kann es zur Synchronisierung notwendig sein, die Sampling-Rate anzupassen. Dazu werden die Signale für das Upsampling häufig interpoliert, während für das Downsampling Datenpunkte im Signal fallengelassen werden.

Normalisierung Um Signalwerte besser vergleichen zu können, können diese auf festgelegte Intervalle normalisiert werden. Insbesondere beim Vergleich von Signalen mit unterschiedlichen Einheiten kann eine Normalisierung nützlich sein.

Filterung Häufig enthalten gemessene Daten Artefakte und Störsignale, die durch Ungenauigkeiten oder Fehler im Messequipment hervorgerufen werden können. Um störende Teile eines Signals zu entfernen, können digitale Filter eingesetzt werden, die als Software-Routinen auf die Daten angewandt werden. Bekannte Arten solcher Filter sind low-pass- oder high-pass-Filter, die jeweils niedrig- oder hochfrequente Teile eines Signals herausfiltern.

Diese Techniken können unabhängig von der eingesetzten Messtechnik zur Verbesserung der Qualität und Verwertbarkeit eingesetzt werden.

Wurden die entsprechenden Daten erfasst und für die jeweilige Fragestellung passend verarbeitet, kann die Analyse der Daten durchgeführt werden. Unter dem Begriff Datenanalyse wird umgangssprachlich gerne der komplette Prozess der Datenerfassung, Datenverarbeitung und Interpretation zusammengefasst. Streng betrachtet beinhaltet die Analyse jedoch lediglich den letzten Schritt der Interpretation. Hier bietet die Informatik neben statistischen Analysen mit den Themenfeldern des Data-Mining und Machine-Learning eine breite Auswahl an Methoden und Technologien, um Muster in den Daten für Analysen sichtbar zu machen. Nachdem die Datenanalyse einen großen Teil der Studien dieser Arbeit ausmacht, wird die ausführliche Beschreibung dieser Felder in den Kapiteln 2.2 und 2.3 vorgenommen.

1.2. Einordnung in die Trainingswissenschaft

Aus Sicht der Sportwissenschaft befasst sich die vorliegende Arbeit mit dem Themenfeld der Trainingswissenschaft, derjenigen „[...] sportwissenschaftlichen Disziplin, die sich aus einer ganzheitlichen und angewandten Perspektive mit der wissenschaftlichen Fundierung von Training und Wettkampf auf den Anwendungsfeldern des Sports beschäftigt“ (Hohmann et al., 2020, S. 19). Gleichzeitig versteht sich die Trainingswissenschaft als integrative, empirische

und angewandte Disziplin. Sie versucht also Erkenntnisse anderer sportwissenschaftlicher Disziplinen, wie z.B. der Sportmedizin, Sportpsychologie oder eben Sportinformatik einzubeziehen, Hypothesen aus der Sportpraxis zu generieren sowie ihrerseits als Unterstützung für die Sportpraxis zu agieren.

Als Gegenstandsbereiche der Trainingswissenschaft werden im Allgemeinen das Training, die Leistungsfähigkeit und der Wettkampf und insbesondere deren Wechselwirkungen untereinander angesehen. Um diese Wechselwirkungen analysieren zu können, wird die sportliche Leistung in der trainingswissenschaftlichen Leistungsdiagnostik in drei Schritten strukturiert (Hohmann et al., 2020). Im ersten Schritt, der Hierarchisierung, wird ein theoretisches Modell der Leistung mit aufeinander aufbauenden Ebenen gebildet (vgl. Abbildung 3). Die Leistungsmerkmale dieses Modells werden anschließend priorisiert und in einer internen Ordnung sortiert, um der Trainingspraxis einen Katalog von Trainingszielen zu liefern. Gleichzeitig generiert die Trainingswissenschaft Normwerte für relevante Leistungsmerkmale, anhand derer in der trainingspraktischen Leistungsdiagnostik Ist-Sollwert-Vergleiche, z.B. zur Identifikation von Stärken und Schwächen von Sportlern, durchgeführt werden können (Hohmann et al., 2020). Je nach untersuchter Fragestellung unterscheidet man auch zwischen theoretischer und praktischer Leistungsdiagnostik (Letzelter & Letzelter, 1982).



Abbildung 3 Pyramidenmodell der komplexen Sportspieleistung. Grafik entnommen aus Hohmann et al. (2020)

Theoretische Leistungsdiagnostik (TLD) Das Ziel der TLD ist die Erstellung allgemeiner Aussagen zur Leistungsstruktur einer Sportart. Typische Fragestellungen in der TLD beinhalten unter anderem die allgemeine Modellierung sportlicher Leistung, sportartspezifische Modellierung (z.B. die Bedeutung der Maximalkraft im Sprint), die Abhängigkeit der Leistungsstruktur vom Leistungsniveau (z.B. Einfluss der Ausdauer im Profifußball) oder die Erstellung statistischer Normen (z.B. die durchschnittliche Sprunghöhe von Volleyball-

spielen).

Praktische Leistungsdiagnostik (PLD) Im Gegensatz dazu beschäftigt sich die PLD mit der informationellen Unterstützung eines Trainingsprozesses. Hier wird also die Wettkampfleistung oder der aktuelle Leistungszustand einzelner Athleten/Mannschaften untersucht. Als Spezialfall der PLD existiert außerdem das Scouting, das sich mit der Diagnostik der Wettkampfleistung eines zukünftigen Gegners auseinandersetzt.

Ein besonders expandierender Bereich der Leistungsdiagnostik ist die Wettkampfdiagnostik (Hohmann et al., 2020), die zudem zu einem der wichtigsten Anwendungsfeldern der Sportinformatik zählt (Link & Lames, 2009). Ihr Ziel ist allgemein formuliert die Analyse des Wettkampfverhaltens, um daraus Rückschlüsse auf die Leistungsvoraussetzungen zu ziehen, aus denen wiederum Hinweise für die Trainingssteuerung generiert werden sollen. Die trainingswissenschaftlichen Anwendungen in den Studien dieser Arbeit fallen allesamt in die Domäne der Wettkampfdiagnostik, weswegen in den folgenden Abschnitten ein kurzer Überblick über die Wettkampfdiagnostik in Sportspielen sowohl im Allgemeinen als auch in den in dieser Arbeit behandelten Sportarten gegeben werden soll.

1.2.1. Wettkampfdiagnostik in Sportspielen

„Sportspiele sind Sportarten [...], bei denen zwei Parteien (Einzel, Doppel, Mannschaften) in einen Interaktionsprozess eintreten, der dadurch zustande kommt, dass beide Parteien gleichzeitig ihr eigenes Spielziel anstreben und behindern wollen, dass die gegnerische Partei ihr Spielziel erreicht; das Spielziel in den Sportspielen ist eine in den Regeln festgelegte, symbolische Handlung.“
(Lames, 1991, S. 33)

Im Gegensatz zu den Kraft- und Ausdauersportarten, bei denen die Zusammenhänge zwischen Leistungsvoraussetzungen und tatsächlich gezeigter Wettkampfleistung weitgehend regressionsanalytisch erfassbar sind, existiert bei Sportspielen eine Nicht-Linearität zwischen Leistungsfähigkeit, Wettkampfleistung und tatsächlichem Erfolg im Wettkampf (Hohmann et al., 2020, S. 246). Spielsportarten werden allgemein als komplexe, dynamische Systeme angesehen, bei denen die Interaktionen zwischen und unter den Parteien gewissermaßen chaotisches Verhalten auslöst, aus dem kaum noch auf die Leistungsvoraussetzungen geschlossen werden kann (Lames & McGarry, 2007; Mateus, 2005).

Diese Interaktionen werden dabei nicht nur durch die Bestrebung, die eigene Taktik durchzusetzen (Passos et al., 2016) beeinflusst, sondern auch durch bestimmte *constraints* [Einschränkungen] (Araujo et al., 2006). Nach Newell (1986) existieren drei Arten solcher *constraints*: *performer*, *environmental* und *task constraints* [etwa Ausführungs-, Umgebungs- und Aufgabeneinschränkungen]. Die *performer constraints* beziehen sich in diesem Zusammenhang auf die individuellen Charakteristiken einzelner Spieler, die Leistungsvoraussetzungen, die diese befähigen, spiel-relevante Aktionen auszuführen (Passos et al., 2016). *Environmental constraints* beschreiben hauptsächlich die Umgebungsbedingungen wie Wind, Temperatur oder die Beschaffenheit des Spielfeldes (z.B. Sand/Hallenboden im Volleyball). Die

task constraints schließlich beinhalten die Spielregeln sowie die räumlichen Einschränkungen durch das Spielfeld und die informationellen Einschränkungen, die die Spieler aufgrund ihrer Wahrnehmung einzelner Spielsituationen betreffen (Bastin et al., 2006). Diese simultan wirkenden Einschränkungen beeinflussen die Handlungen der Spieler im dynamischen Interaktionsprozess der Sportspiele (Araújo et al., 2013). Für die Wettkampfdiagnostik stellt sich durch die Nicht-Linearität die Frage nach geeigneten Modellen, die die sportliche Leistung in Sportspielen adäquat erfassen und beschreiben können.

Den traditionellen Ansatz zur Erfassung der Wettkampfleistung stellt die Wettkampfbeobachtung anhand von sportartspezifischen Beobachtungssystemen dar, die speziell zur Analyse taktischer Muster entworfen wurden (Hughes & Franks, 2007). Diese Beobachtungssysteme werden dazu genutzt, die Häufigkeit kritischer Ereignisse im Spiel zu dokumentieren und zu analysieren (Hughes & Bartlett, 2002). So werden im Sinne der trainingspraktischen Leistungsdiagnostik oft erfolgreiche und nicht erfolgreiche Leistungen miteinander verglichen (Hughes & Franks, 2007). Klassische Indikatoren, die von diesen Systemen erfasst werden, beinhalten häufig die Anzahl der Torschüsse, Ecken, Tore/erzielte Punkte und erfolgreiche Pässe (Castellano et al., 2012; Harrop & Nevill, 2014).

Obwohl diese Indikatoren nützliche Einblicke geben können und leicht zu erfassen sind, ist es auch damit schwer, die komplexen Leistungen in Sportarten zu erfassen (Carling et al., 2014; Mackenzie & Cushion, 2013). Resultierend daraus haben sich in den vergangenen Jahren neue Ansätze zur Wettkampfdiagnostik in Sportspielen entwickelt (vgl. z.B. Ofoghi et al. (2013) und Rein und Memmert (2016)). Ziel dieser Ansätze ist zum Großteil, den komplexen Interaktionsprozessen im Sportspiel gerecht zu werden (Passos et al., 2016). Begünstigt wird diese Entwicklung durch den technologischen Fortschritt, der immer genauere Daten aus der Spielbeobachtung liefern kann. Einer dieser Ansätze ist die Modellierung der komplexen Interaktionen im Sportspiel durch Konzepte aus der Theorie der dynamischen Systeme, wie zum Beispiel der Relativen Phase (Bourbousson et al., 2010; Folgado et al., 2014; Lames & McGarry, 2007; Sampaio & Maças, 2012).

Die Stärke dieser Methode besteht darin, dass synchrone Muster innerhalb oder zwischen Mannschaften aufgedeckt werden können. Es kann zum Beispiel die Dynamik der Bewegung verteidigender Spieler relativ zueinander erfasst werden, um Irregularitäten zu finden, die der angreifenden Mannschaft Vorteile verschaffen (Travassos et al., 2011). Analog zeigen Duarte et al. (2013) durch eine geclusterte Phasenanalyse, dass die Bewegungen einer Fußballmannschaft entlang des Spielfeldes eine höhere Synchronisation aufweisen, als Bewegungen in laterale Richtung.

Duarte et al. (2012) gehen dabei sogar noch einen Schritt weiter und beschreiben Sportmannschaften als „Superorganismen“. Dabei wird das Kollektiv zusammenarbeitender Spieler als eine einzige soziale Einheit gesehen, die durch Aufgabenteilung versucht, das vorgegebene Spielziel zu erreichen.

Aber auch probabilistische Ansätze, die die Unsicherheit durch Komplexität im Sportspiel durch Wahrscheinlichkeiten ausdrücken, haben Beachtung in der Forschung gefunden. So können Markov-Ketten zur Modellierung und Simulation taktischer Verhaltensweisen in Sportspielen verwendet werden (Lames, 1991). Besonders in Rückschlag-Sportarten wie Bad-

minton (Percy, 2009), Tischtennis (Pfeiffer et al., 2010) oder Volleyball (Lames & Hohmann, 1997) findet diese Methode der Modellierung Anwendungsmöglichkeiten. Aber auch im Fußball (Narizuka et al., 2014) oder Golf (Maher, 2013) finden sich Arbeiten, die Markov-Ketten als Modell verwenden.

Mit Bayesschen Modellen existiert ein weiterer Zweig probabilistischer Modelle zur Wettkampfdiagnostik. Diese Methoden bieten Vorteile, die zur Analyse komplexer Sportspielleistung hilfreich sein können. Sie erlauben es beispielsweise, Expertenwissen aus der Sportwissenschaft direkt in einem stochastischen Modell abzubilden, Unsicherheiten in Vorhersagen miteinzubeziehen und ihr Modell auch mittels im Nachhinein erfasster Daten anzupassen. Ein umfassender Überblick über die Forschung in diesem Feld findet sich in Santos-Fernandez et al. (2019).

Nicht zuletzt haben auch Methoden des Machine-Learning vermehrt ihren Weg in die Wettkampfdiagnostik gefunden. Machine-Learning-Anwendungen agieren Daten-getrieben, versuchen also, aus einem Datenbestand Muster zu lernen, um damit zumeist Vorhersagen für unbekannte Daten treffen zu können. Die verwendeten Modelle, wie z.B. neuronale Netze, versprechen somit, die vorhandene Nicht-Linearität implizit aus den Daten zu übernehmen. Solche Modelle wurden bereits in verschiedenen Sportarten eingesetzt, um die Gewinner von Spielen vorauszusagen (Bunker & Thabtah, 2019; Cai et al., 2019; Gu et al., 2019). Da zum Einsatz von Machine-Learning zumeist eine große Menge an Daten verfügbar sein muss, können mit diesen Methoden insbesondere in Sportarten, die kontinuierliche Positionsdaten ihrer Spieler als Datenbasis anbieten, Fortschritte erzielt werden. Dick und Brefeld (2019) und Decroos et al. (2019) lernen so Funktionen, um Spielsituationen im Fußball automatisch zu bewerten. Le et al. (2017) verwenden sogenannte Deep-Learning-Algorithmen, um das taktische Verhalten von Fußballspielern nachzuahmen und dadurch spielrelevante Strategien abzuschätzen. Im Basketball wurden ebenfalls mit Hilfe von Positionsdaten offensive (Wang & Zemel, 2016) sowie defensive (Tian et al., 2020) Spielzüge klassifiziert und spieler-spezifische Wurfprofile erstellt (Miller et al., 2014). Aber auch im Baseball (Koseler & Stephan, 2017), Ice-Hockey (Tora et al., 2017), American Football (Hsu, 2020) und Volleyball (Van Haaren et al., 2016) findet Machine-Learning Anwendung.

1.2.2. Wettkampfdiagnostik in den behandelten Sportarten

Nach der allgemeinen Betrachtung der Wettkampfdiagnostik in Sportspielen, soll hier noch einmal kurz genauer auf die in dieser Arbeit behandelten Sportarten Tischtennis und Beachvolleyball eingegangen werden.

Wettkampfdiagnostik im Tischtennis

Fuchs et al. (2018) identifizieren in ihrer Review fünf große Themenblöcke für Methoden in der wettkampfdiagnostischen Forschung im Tischtennis: Klassische Leistungsindizes, Momentum-Analysen, Footwork-Analysen, Qualitative und Simulative Ansätze.

Ein weit verbreiteter Ansatz, der Leistungsindizes verwendet, ist die sogenannte „Three-Phase“-Methode (Wu & Li, 1992). Bei dieser Methode wird ein Ballwechsel in drei Phasen aufgeteilt, die jeweils einzeln analysiert werden: Den ‚Angriff nach Aufschlag‘ beim ersten

und dritten Schlag, ‚Angriff nach Annahme‘ beim zweiten und vierten Schlag und ‚Rally‘ nach dem vierten Schlag im Ballwechsel. Abgeleitete Ansätze führen ähnliche Aufteilungen eines Ballwechsels basierend auf der Schlagreihenfolge durch und analysieren die resultierenden Spielphasen (Hao et al., 2010; Hsu, 2010; Hsu et al., 2014; Tamaki et al., 2017).

Momentum-Analysen, die das Konzept des Momentums (Adler, 1981) aus der Psychologie auf den Sport übertragen, versuchen der dynamischen Struktur der Sportspiele durch die Modellierung der momentanen Spielstärke oder Dominanz eines Spielers beziehungsweise Teams gerecht zu werden (Murray & Hughes, 2001). Dazu können zum Beispiel Moving-Averages eingesetzt werden (Lames, 2006; Liu et al., 2015), die nach einem Sliding-Window-Ansatz über den Matchverlauf gelegt werden und somit Kennzahlen für die momentane Spielstärke eines Spielers liefern.

Einen eher biomechanischen Ansatz verfolgen die Arbeiten der Footwork-Analysen [Beinarbeit-Analysen]. Hier werden die Schrittmuster, die Tischtennispieler zur Annäherung an den Ball verwenden, als Untersuchungsgegenstand verwendet. In diesem Zusammenhang konnten standardisierte Definitionen der am häufigsten verwendeten Schritttypen gefunden (Malagoli Lanzoni et al., 2007) und in Zusammenhang mit den verwendeten Schlagtechniken gebracht werden (Malagoli Lanzoni, Di Michele et al., 2013). Weitere Studien zeigen Unterschiede in Schrittmustern zwischen verschiedenen Altersklassen (Malagoli Lanzoni, Michele et al., 2013) sowie europäischen und asiatischen Spielern (Malagoli Lanzoni et al., 2014).

Ein Kritikpunkt der bisher erwähnten Methoden ist die Verwendung summarischer Statistiken als Leistungsindikatoren, da damit wichtige Aspekte der dynamischen Interaktionen verloren gehen können (Lames & McGarry, 2007). Wie bereits in Abschnitt 1.2.1 beschrieben, existiert mit der Simulation von Ballwechseln anhand von Markov-Ketten eine bereits im Tischtennis implementierte Möglichkeit, den dynamischen Aspekt des Sports zu berücksichtigen (Pfeiffer et al., 2010). Die Erweiterung dieser Methode ist Bestandteil dieser Dissertation, weswegen an dieser Stelle für eine genauere Beschreibung der Methode auf Abschnitt 2.1.1 beziehungsweise die Publikation in Kapitel A verwiesen wird.

Ebenso können qualitative Methoden helfen, den Abstraktionsgrad statistischer Analysen zu verringern. So wurden in einigen Studien Experten zur Effektivität und Häufigkeit technisch-taktischer Elemente (z.B. Schlagtechnik, Schlagkombinationen) im Tischtennis befragt, um die relevantesten Elemente der Spielstruktur zu identifizieren (Munivrana, Furjan-Mandić et al., 2015; Munivrana, Petrinović et al., 2015). Dazu wurden diese Elemente in Zusammenhang mit zusätzlichen, Tischtennis-relevanten Variablen wie zum Beispiel der Schlagposition über dem Tisch oder der Grifftechnik bewertet.

Zusätzlich zu den simulativen Ansätzen existieren weitere sportinformatische Arbeiten im Anwendungsfeld Tischtennis. Zuerst sind hier die Arbeiten zur automatisierten Datenerfassung zu nennen, die versuchen, den aufwändigen Prozess der Annotation von Wettkampfdaten zu vereinfachen. In diesen Studien wird versucht, mit Methoden des Machine-Learning sowohl die Position als auch Trajektorie des Balls aus Videoaufnahmen zu extrahieren (Ji et al., 2018; Tamaki & Saito, 2013; Voeikov et al., 2020). Ähnliche Ansätze zeigen die Arbeiten von Blank et al. (2015), Lim et al. (2018), Liu et al. (2019) und Tabrizi et al. (2020), die Bewegungsdaten tragbarer Sensoren zur Klassifikation von Schlägen im Tischtennis verwenden.

Machine-Learning wird weiterhin eingesetzt, um den Einfluss des Saisonverlaufs auf die Leistung von Tischtennisspielern zu untersuchen (Lai et al., 2018) und sowohl das Ergebnis von Spielen (Yang & Zhang, 2016) als auch von Ballwechseln vorherzusagen (Draschkowitz et al., 2014).

Nicht zuletzt trägt die Sportinformatik dazu bei, die Vermittlung wettkampfdiagnostischer Erkenntnisse durch neue Visualisierungen (Wu et al., 2018) und Analysekonzepte zu verbessern (Wang et al., 2019).

Wettkampfdiagnostik im Beachvolleyball

Für die leistungsdiagnostische Forschung im Beachvolleyball lassen sich verschiedene Themenblöcke identifizieren. Ein erstes Themenfeld stellen biomechanische Studien der im Beachvolleyball vorherrschenden Bewegungsmuster dar. Hier sind besonders die sportart-spezifischen Bewegungen im Angriffsschlag (Giatsis et al., 2019; Reeser et al., 2010; Serrien et al., 2016) und dem damit verbundenem Sprung (Giatsis et al., 2018; Tilp & Rindler, 2013; Tilp et al., 2008) von Interesse.

Weiterhin existieren zahlreiche Studien zur Leistungsstruktur im Beachvolleyball. Weltstandsanalysen (Hohmann et al., 2020, S. 265) in der männlichen (Giatsis et al., 2015; Giatsis & Zahariadis, 2008) und weiblichen (Häyrinen & Tampouratzis, 2012; Koch & Tilp, 2009a) Weltspitze sowie der Vergleich der wettkampfdiagnostischen Charakteristika zwischen den Geschlechtern (Chinchilla-Mira et al., 2012; Koch & Tilp, 2009b) belegen Tendenzen der erfolgreichsten Athleten und erlauben die Übertragung der Ergebnisse in die perspektivische Trainingsplanung.

Unter den sportinformatischen Arbeiten stechen die Bemühungen zur automatischen Erfassung von Positionsdaten der Spieler heraus (Gomez et al., 2014; Gomez et al., 2012; Mauthner et al., 2007). Ziel dieser Arbeiten ist die automatisierte Extraktion der Spielerpositionen und Ballkoordinaten aus einem Videobild. Der limitierende Faktor dieser Arbeiten ist der Einsatz einer einzelnen am Spielfeldrand platzierten Kamera mit eingeschränkter Auflösung, zusammen mit den Störfaktoren der Umgebung. Durch diese Umgebungsbedingungen weist besonders das Tracking des Balls und der im Hintergrund befindlichen Spieler noch zu geringe Genauigkeiten für den Praxiseinsatz auf.

Einen anderen Ansatz wählen Kautz et al. (2017), die tragbare Beschleunigungssensoren im Zusammenspiel mit einem Machine-Learning-Algorithmus einsetzen, um technische Aktionen der Spieler zu erkennen. Hier können unter den gegebenen Bedingungen Klassifikationsgenauigkeiten von ca. 80 Prozent erreicht werden. Haider et al. (2020) erweitern diesen Ansatz um Videobilder, die als zusätzliche Information in einen Klassifikationsalgorithmus eingespeist werden. Allerdings kann auch damit die Genauigkeit der Aktionserkennung nur um etwa vier Prozentpunkte erhöht werden.

Die Entwicklung und Evaluierung unterstützender Software zählt ebenfalls zum Themengebiet der Sportinformatik (Tilp et al., 2006). Um die Akzeptanz der Programme bei den Anwendern zu erhöhen, spielt dabei die Orientierung an den Anforderungen der Sportpraxis eine entscheidende Rolle (Link, 2014).

Hoffmeister (2019) liefert durch die Modellierung eines Beachvolleyball-Spiels als Markov-

Entscheidungsprozess (MDP, *Markov decision process*) eine interessante sportinformatische Anwendung. Anhand dieses Prozesses werden mathematisch fundierte Strategien zur Gewinnoptimierung hergeleitet, mit denen zusätzlich die Gewinnwahrscheinlichkeit eines Teams berechnet werden kann. So können zum Beispiel „riskante“ und „sichere“ Aufschlags- bzw. Angriffsstrategien bezüglich ihrer erwarteten Gewinnwahrscheinlichkeit verglichen werden.

1.3. Motivation

Die grundlegende Motivation dieser Arbeit, die sich aus dem Selbstverständnis der Sportinformatik ableitet, ist die Anwendung informatischer Methoden, um sportwissenschaftlichen Mehrwert zu generieren. Der Fokus liegt dabei gleichermaßen auf der Lösung sportwissenschaftlicher Problemstellungen wie auf der Evaluierung der Anwendung der Methoden selbst. Gleichzeitig hebt sich diese Arbeit durch folgende Gesichtspunkte von existierender sportinformatischer Forschung ab:

Ausweitung sportinformatischer Forschung in Randsportarten

Ein Merkmal dieser Dissertation ist die Ausweitung sportinformatischer Forschung auf die Sportarten Tischtennis und Beachvolleyball. Wie an den Arbeiten in Abschnitt 1.2.1 bereits angedeutet zu erkennen ist, konzentriert sich die Forschung im Bereich der Sportinformatik - insbesondere bei der Anwendung von Methoden des Data-Mining und Machine-Learning - bisher auf wenige, meist finanziell lukrativere und medienwirksamere Sportarten wie Fußball oder die amerikanischen Sportarten wie Basketball und American Football. Das ist sicherlich auch in der Verfügbarkeit leistungsdiagnostischer Daten begründet, was im nächsten Abschnitt noch genauer betrachtet wird.

Die dynamische Natur der Spielsportarten zeigt sich jedoch ebenso im Tischtennis und Beachvolleyball, auch wenn die geringere Zahl der Akteure zuerst eine verminderte Komplexität vermuten lässt. Das Grundproblem der nicht-linearen Interaktionen zwischen und unter den Parteien bleibt bestehen und bedingt somit den Einsatz von geeigneten Methoden, die unter anderem die Sportinformatik anbieten kann.

Anwendung aktueller informatischer Methoden auf diskreten Ereignisdaten

Natürlich spielen die verfügbaren Daten eine tragende Rolle beim Einsatz sportinformatischer Methoden in der Leistungsdiagnostik. Dies könnte einen Grund für den verstärkten Einsatz der Sportinformatik in den zuvor genannten Sportarten darstellen, die zeitkontinuierliche Positionsdaten durch den Einsatz entsprechender automatisierter Erfassungssysteme in Training und Wettkampf zur Verfügung stellen. Gerade mit dem wachsenden Interesse an Methoden des Data-Mining und Machine-Learning konzentrieren sich Studien vermehrt auf diejenigen Sportarten, die die dafür notwendigen Datenmengen zur Verfügung stellen können. Mackenzie und Cushion (2013) argumentieren sogar, dass die untersuchten Variablen vieler leistungsdiagnostischer Studien Resultat der Verfügbarkeit der Daten sind und nicht grundsätzlich nach ihrer Nützlichkeit ausgewählt wurden. Gerade in den in dieser Arbeit be-

handelten Sportarten, in denen noch keine automatisierte Datenerfassung etabliert ist, die Leistungsindikatoren also noch manuell erfasst werden, ist die Verfügbarkeit von Positionsdaten eingeschränkt. Hier werden die Daten noch größtenteils mit Hilfe von Beobachtungssystemen per Hand (jedoch Software-unterstützt, vgl. Link und Ahmann (2013)) durch eigens geschulte Scouts erfasst. Allerdings existieren im Tischtennis und Beachvolleyball schon allein auf Grund der Zahl der Akteure und der reduzierten Spielfeldgrößen (im Vergleich zu Sportarten wie z.B. Fußball) gewisse Einschränkungen, die diese Art der Datenerfassung erleichtern. Ziel der Arbeit ist es daher weiterhin, die Nützlichkeit informatisch gestützter Analysen auch auf diskreten Ereignisdaten aufzuzeigen. In allen drei Studien wurden Ereignisdaten als Basis der Analysen verwendet, wobei diese in der Machine-Learning-Studie zusätzlich mit zeit-diskreten Positionsdaten kombiniert wurden.

An dieser Stelle soll kurz auf eine Besonderheit der Datenbasis der Beach-Volleyball Studien hingewiesen werden. Selbst wenn hier noch keine vollautomatisiert erfassten Daten zur Verfügung stehen, können die Analysen in dieser Arbeit durch die seit Jahren andauernde Kooperation mit den deutschen Beach-Volleyball Nationalteams auf einen Datensatz zurückgreifen, der zum Verfassungszeitpunkt dieser Arbeit sowohl in seiner Größe als auch im Umfang der erfassten Merkmale weltweit einzigartig sein dürfte. Unterstützt wird dieser Prozess durch eine eigens zu diesem Zweck erstellte Scouting- und Analyse-Software (Link & Ahmann, 2013), an deren Entwicklung auch der Autor dieser Arbeit beteiligt war. Durch diese Kooperation wird sichergestellt, dass die erfassten Variablen einen leistungsdiagnostischen Zweck erfüllen, da sie von den Analysten der Nationalteams laufend zur Spielanalyse verwendet werden.

Leistungsdiagnostische Analysen auf Rally-Ebene

Das Ziel wettkampfdiagnostischer Untersuchungen ist es, die Wettkampfleistung und damit den Wettkampferfolg zu erklären. Und auch außerhalb der Wissenschaft existieren Akteure, wie z.B. die Wettindustrie, für deren Analysen der Gewinner einer Partie relevant ist. Die richtige Konsequenz daraus sind Studien, deren Primärziel es ist, Unterschiede zwischen Gewinnern und Verlierern von Partien herauszuarbeiten (Giatsis & Tzetzis, 2003; Katsikadelis et al., 2013; Lago-Peñas et al., 2010). Allerdings ist auch die Aufklärung der Leistungsstruktur einer Sportart, und damit die Aufklärung der Beziehungen der Leistungsindikatoren untereinander, erklärtes Ziel der Leistungsdiagnostik (Hohmann et al., 2020). So beeinflussen zum Beispiel Bewegungsmuster innerhalb von Ballwechseln den Entscheidungsprozess des defensiven Spielers im Beachvolleyball (Schläppi-Lienhard & Hossner, 2015). Die Optimierung dieses Prozesses führt wiederum zu einer Steigerung der Defensivleistung und folglich zu einer erhöhten Erfolgswahrscheinlichkeit. Eine Untersuchung solcher Muster, die nur indirekt auf den Erfolg einer Partie einwirken, kann also einen sinnvollen Beitrag zur Wettkampfdiagnostik leisten. Im Zuge dessen beschäftigen sich alle Studien dieser Dissertation mit Mustern innerhalb von Spielzügen, sei es anhand der Ballwechsellängen im Tischtennis (Simulations-Studie) oder anhand der Spielaktionen innerhalb von Beachvolleyball-Spielzügen (Data-Mining und Machine-Learning-Studie).

Der weitere Verlauf dieser Dissertation gliedert sich wie folgt: Kapitel 2 bietet einen tieferen Einblick in die informatischen Methoden, die in den Studien dieser Arbeit verwendet wurden. Diese Studien werden anschließend in Kapitel 3 genauer vorgestellt. Kapitel 4 diskutiert die Studien noch einmal in der Retrospektive und vergleicht sie mit aktuellen Forschungsansätzen. Abschließend wird in Kapitel 5 ein Fazit gezogen und ein Ausblick auf zukünftige Arbeiten im Bereich der Sportinformatik gegeben.

2. Methoden

Wie bereits im vorangegangenen Kapitel erwähnt, besteht das Ziel dieser Arbeit darin, sportwissenschaftlichen Mehrwert durch den Einsatz informatischer Methoden zu generieren. Aus dem breiten Feld informatischer Werkzeuge wurden dazu die Themengebiete der Simulation, des Data-Mining und des Machine-Learning ausgewählt. Die konkreten Methoden wurden jeweils unter folgenden Gesichtspunkten berücksichtigt:

1. Kann die Methode zur Lösung eines sportwissenschaftlichen Problems beitragen?
2. Bieten die vorhandenen Daten eine ausreichende Basis zur Anwendung der Methode?
3. Lassen sich die zu erwartenden Ergebnisse in der Sportpraxis nachvollziehen?

Am Beispiel der Data-Mining-Studie lässt sich dieser Prozess nachvollziehen: Das zu lösende Problem ist die Suche nach sequenziellen taktischen Mustern in Beachvolleyball-Spielzügen. Dazu wird eine Methode benötigt, die die Reihenfolge der Aktionen in den Spielzügen berücksichtigt und auf diskreten Event-Daten anwendbar ist. Beispiele dafür sind das Sequential-Pattern-Mining, eine Erweiterung des Association-Rule-Mining (siehe Abschnitt 2.2.2), oder auch rekurrente neuronale Netze (siehe Abschnitt 2.3.4). In diesem Fall fiel die Wahl auf das Sequential-Pattern-Mining, da neuronale Netze zum Einen eine deutlich größere Zahl an Daten benötigen, um Ergebnisse zu generieren. Zum Anderen sind die Algorithmen des Pattern-Mining leichter nachzuvollziehen, weshalb von einer größeren Akzeptanz der Ergebnisse in der Sportpraxis auszugehen ist.

Das restliche Kapitel widmet sich der ausführlicheren Beschreibung der in den Studien verwendeten Methoden der Simulation (2.1), des Data-Mining (2.2) und des Machine-Learning (2.3).

2.1. Simulation

„Simulation ist das Nachbilden von Prozessen realer Systeme in einem Modell und das anschließende Durchführen von Experimenten an diesem Modell.“ (Sauerbier, 1999, S. 5) Auf der Grundlage der durch die Simulation erzielten Ergebnisse werden anschließend Aussagen über das Originalsystem getroffen. Grundsätzlich bieten sich Simulationen als Analysewerkzeug in Situationen an, in denen das reale System nicht oder nur unter großem Aufwand untersucht werden kann. Dies kann der Fall sein, wenn das reale System (noch) nicht existiert, das Experiment in der Realität zu teuer wäre oder das System Eigenschaften aufweist, die eine analytische Lösung erschweren. Insbesondere Nichtlinearitäten, die bei der Leistungsdiagnostik in Sportspielen eine große Rolle spielen (siehe Abschnitt 1.2.1), verhindern den Einsatz analytischer Lösungsansätze (Sauerbier, 1999). Dabei haben Simulationen durch-

aus Vorteile gegenüber Experimenten am Realsystem: Verschiedene Systemkonfigurationen und Umgebungsbedingungen lassen sich relativ einfach durch Änderung der Modellparameter untersuchen. Der Prozess der Modellbildung erzwingt ein hohes Verständnis der Struktur und des Verhaltens des zu simulierenden Systems. Aber natürlich bringen Simulationen auch Nachteile mit sich. Schon die Modellierung bietet Potenzial für Fehler bei der Nachbildung des Systems, wodurch die Simulationsergebnisse beeinflusst werden können. Dementsprechend kommt der Modellbildung eine entsprechend wichtige Rolle im Simulationsprozess zu. Sie ist zudem notwendig, um den eingeschränkten (nicht unbegrenzten) Rechenkapazitäten der zur Simulation meistens verwendeten Rechner gerecht zu werden. Auch wenn die Rechenkapazitäten moderner Computer weiter steigen, schränkt dies die Komplexität der Modelle und damit die Genauigkeit der Simulation ein.

Das in dieser Studie verwendete mathematische Modell zur Simulation - die Markov-Ketten - und der dazugehörige Prozess der Simulation werden im folgenden Abschnitt genauer dargestellt.

2.1.1. Markov-Ketten

Bei Markov-Ketten handelt es sich um eine Form spezieller diskreter stochastischer Prozesse, die sich gut zur Modellierung von Systemen eignen, bei denen Grund zu der Annahme besteht, dass die Zustandsänderungen des Systems nur über eine gewisse Zeit hinweg Einfluss aufeinander ausüben. Dabei ist unter einem stochastischen Prozess ein System zeitlich geordneter Übergänge zwischen zuvor modellierten Zuständen zu verstehen. Diese Übergänge zwischen den Zuständen werden durch stochastische Wahrscheinlichkeiten ausgedrückt, die Tendenzen des Systems für Pfade durch seinen Zustandsraum beschreiben. Als ein finiter stochastischer Prozess wird ein stochastischer Prozess bezeichnet, der nur eine endliche Menge von Zuständen annehmen kann.

Die Markov-Eigenschaft eines stochastischen Prozesses formuliert allgemein Bedingungen daran, wie groß der Einfluss vergangener Zustände des Prozesses auf die Gegenwart ausfällt. Sie besagt, dass die bedingte Wahrscheinlichkeit für das Auftreten eines Zustandes im nächsten Schritt des Prozesses - gegeben aller vorheriger Ereignisse - gleich der bedingten Wahrscheinlichkeit für dieses Auftreten gegeben des aktuellen Zustandes ist. Formalisiert man diese Bedingung, so ist eine Markov-Kette eine Sequenz von Zufallsvariablen X_1, X_2, X_3, \dots mit Werten aus dem Zustandsraum $S = \{s_1, s_2, s_3, \dots\}$ für die gilt:

$$P(X_{t+1} = s_{j_{t+1}} | X_t = s_{j_t}, X_{t-1} = s_{j_{t-1}}, \dots, X_0 = s_{j_0}) = P(X_{t+1} = s_{j_{t+1}} | X_t = s_{j_t})$$

Die Übergangswahrscheinlichkeiten hängen also nur von dem augenblicklichen Zustand ab und bereits vergangene Ereignisse des Systems haben keine Auswirkungen auf zukünftige Zustände, womit die Markov-Kette auch als „gedächtnislos“ bezeichnet wird. Man kann zudem Markov-Ketten n -ter Ordnung beschreiben, bei denen der zukünftige Zustand von den n vorherigen Zuständen abhängig ist und somit ein Gedächtnis besitzt. Gedächtnislose Markov-Ketten entsprechen damit also Markov-Ketten erster Ordnung. Definiert man die

Übergangswahrscheinlichkeiten zwischen den einzelnen Zuständen als

$$p_{ij}(t) := P(X_{t+1} = s_j | X_t = s_i), i, j = 1, \dots, m,$$

lassen sich diese in einer quadratischen Übergangsmatrix $M(t) = (p_{ij}(t))_{s_i, s_j \in S}$ pro Zeitpunkt in der Kette zusammenfassen. Sind die Übergangswahrscheinlichkeiten in einer Markov-Kette unabhängig vom Zeitpunkt t , also von ihrer Position in der Ereigniskette, so spricht man von *homogenen* Markov-Ketten mit stationären Übergangswahrscheinlichkeiten. In diesem Fall können alle Übergangswahrscheinlichkeiten in einer einzelnen Übergangsmatrix festgehalten werden, die als Ausgangspunkt für die Simulation anhand von Markov-Ketten dient.

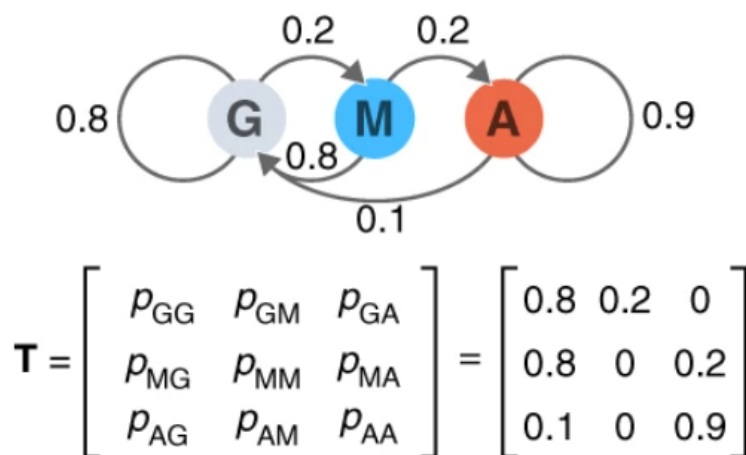


Abbildung 4 Markov-Kette mit 3 Zuständen und der Übergangsmatrix T . Grafik entnommen aus Grewal et al. (2019).

2.1.2. Simulation mit Markov-Ketten

Lames (1991, S. 92) beschreibt den Prozess der Simulation anhand von Markov-Ketten folgendermaßen:

“Die Leistungsrelevanz einer taktischen Verhaltensweise im Sportspiel wird quantifiziert durch die Differenz der Erfolgswahrscheinlichkeiten, die aus einer Ausgangsübergangsmatrix und einer im Sinne der taktische Verhaltensweise veränderten Übergangsmatrix berechnet werden.”

Zur Bestimmung der hier erwähnten Differenz der Erfolgswahrscheinlichkeiten sind dabei einige Zwischenschritte notwendig:

Ausgehend von einer bereits vorhandenen Übergangsmatrix, die entweder aus einer zuvor erfolgten Spielbeobachtung hervorgegangen ist oder wieder als Modell für einen bestimmten Leistungsbereich fungiert, können die Erfolgswahrscheinlichkeiten als die Wahrscheinlichkeiten für das Erreichen der absorbierenden Zustände berechnet werden.

Die Wahrscheinlichkeiten dieser Übergangsmatrix werden im folgenden Schritt so verändert, dass sie die gewünschten Ausschläge der taktischen Verhaltensweisen darstellen. So kann z.B. im Beachvolleyball eine erhöhte Ass-Rate untersucht werden, indem die Übergangswahrscheinlichkeit zwischen “Aufschlag” und “Punkt” erhöht wird. Im Gegenzug wird diese

Erhöhung durch Senkung der anderen Übergangswahrscheinlichkeiten des Zustandes "Aufschlag" kompensiert, um eine Zeilensumme von 1 in der Übergangsmatrix zu erhalten. Die resultierende Übergangsmatrix enthält nun Verhalten, das in dieser Form nicht direkt beobachtet wurde, gleichzeitig aber die gewünschten Auswirkungen auf die Häufigkeit der zu untersuchenden taktischen Muster darstellt.

Werden auf der veränderten Übergangsmatrix erneut die Erfolgswahrscheinlichkeiten ermittelt, kann anhand der Differenz zwischen der ursprünglichen und der veränderten Erfolgswahrscheinlichkeit ein Maß für die Leistungsrelevanz des untersuchten taktischen Verhaltens festgehalten werden.

Die Simulation durch Veränderung der Übergangswahrscheinlichkeiten eröffnet zwei Problemstellungen, die in den folgenden Abschnitten beschrieben werden.

Amplitude der Auslenkung

Das erste Problem stellt die Festlegung der Amplitude der Auslenkungen innerhalb der Übergangsmatrix dar. Geht man davon aus, dass das Ziel der leistungsdiagnostischen Analyse die Erstellung eines Prioritätenkataloges von taktischen Verhaltensweisen ist, so muss eine Rangfolge innerhalb dieses Kataloges anhand der Leistungsrelevanz der einzelnen Verhaltensweisen erstellt werden. Diese Leistungsrelevanz wird durch die Differenz der Erfolgswahrscheinlichkeiten zwischen Ausgangs- und ausgelenkter Übergangsmatrix beschrieben. Damit die resultierende Relevanz jedoch zwischen den verschiedenen Verhaltensweisen vergleichbar bleibt, müssen in einem eigenen Modellierungsschritt ähnlich "schwer" zu realisierende Auslenkungen pro Verhaltensweise abgebildet werden.

Wie jedoch definiert man eine solche Schwierigkeit, unabhängig von Sportart und Verhaltensweise? Eine konstante Auslenkung um beispielsweise .05 erscheint zuerst als einfachste Möglichkeit, bedeutet aber für eine ursprüngliche Übergangswahrscheinlichkeit von 5% eine Verdoppelung, während ein Übergang mit einer Wahrscheinlichkeit von 50% nur um 10% verändert würde. Zudem ist offensichtlich, dass zum Beispiel die Änderung der Übergangswahrscheinlichkeit von .50 auf .55 deutlich "leichter" in der Praxis durchzuführen ist, als zum Beispiel eine Änderung von .90 auf .99. Es muss also eine Modellierung auf Basis allgemeiner Beobachtungen über die Schwierigkeit von Verhaltensänderungen durchgeführt werden. Dabei definiert Lames (1991) folgende drei Bedingungen an das Modell:

1. Änderungen an den Wahrscheinlichkeiten im mittleren Wahrscheinlichkeitsbereich müssen leichter sein als an den Rändern (siehe oben genanntes Beispiel). Daraus ergibt sich eine größere Auslenkung in der Mitte des Intervalls $[0,1]$.
2. Je näher die ursprüngliche Übergangswahrscheinlichkeit an den Rändern des Wahrscheinlichkeitsbereichs liegt, desto schwieriger müssen Auslenkungen eben dieser Wahrscheinlichkeit sein. Die Modellfunktion sollte also symmetrisch um den Punkt 0.5 sein.
3. An den Randpunkten der Wahrscheinlichkeit kann ausgelenkt werden, wenn diese Punkte als Maxima/Minima der zugehörigen Verhaltensweisen angesehen werden. An den Grenzwahrscheinlichkeiten nimmt das Modell Auslenkungen mit Werten > 0 vor.

Als Lösung wird folgende Modellfunktion formuliert, die auch in Abbildung 5 dargestellt ist:

$$\delta\ddot{U}W(\ddot{U}W) = K + B * 4 * \ddot{U}W * (1 - \ddot{U}W)$$

Die Änderung der Übergangswahrscheinlichkeit besteht hier aus einem konstanten Betrag K für die Auslenkung in den Grenzbereichen sowie einer Konstante B , die die maximale Ausprägung festlegt. Dadurch lässt sich die Funktion an mögliche Vorkenntnisse in Bezug auf die Schwierigkeit der Verhaltensänderungen, z.B. in bestimmten Sportarten, anpassen, die Festlegung der Parameter muss allerdings auch für jedes Sportspiel neu ausgeführt werden.

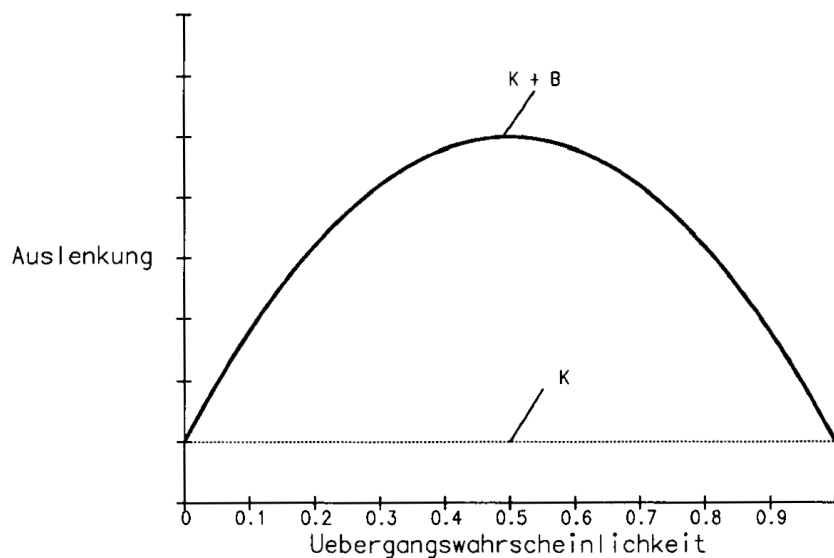


Abbildung 5 Betrag der Auslenkung in Abhängigkeit der Übergangswahrscheinlichkeit. Grafik entnommen aus Lames (1991)

Methodik der Kompensation

Neben der Problematik der Festlegung der Amplitude der möglichen Auslenkungen an den Übergangswahrscheinlichkeiten muss zudem die Art und Weise der Kompensation dieser Auslenkungen betrachtet werden. Wie bereits erwähnt muss die Veränderung einer Wahrscheinlichkeit $\delta\ddot{U}W_x$ durch eine kompensierende Veränderung anderer Übergänge für den betreffenden Zustand ausgeglichen werden, um eine Zeilensumme von 1 in der Übergangsmatrix zu erhalten. Dies bedeutet, dass die Summe aller einzelnen Veränderungen der Übergangswahrscheinlichkeiten eines Zustandes Null ergeben muss. Diese Randbedingung lässt sich auf mehrere Arten umsetzen:

1. *Kompensation durch eine einzige Übergangswahrscheinlichkeit:* Die Änderung wird durch eine komplementäre Auslenkung einer einzelnen Übergangswahrscheinlichkeit kompensiert, zum Beispiel durch die Senkung der Fehlerwahrscheinlichkeit beim Aufschlag im Beachvolleyball, wenn eine erhöhte Ass-Rate untersucht werden soll.
2. *Gleichmäßige Kompensation:* Die Änderung wird durch eine Auslenkung aller anderen Übergänge des Zustandes ausgeglichen. Die Höhe der Kompensation erfolgt dabei für jeden Übergang proportional zu dessen Anteil an der restlichen zur Verfügung stehen-

den Wahrscheinlichkeit $1 - \delta \dot{U}W_x$.

3. *Beliebige Kompensation*: Grundsätzlich kann jedes Modell, das die Rahmenbedingung der Nullsumme der Änderungen beachtet, zur Kompensation verwendet werden. So lassen sich a-priori Kenntnisse über eine Sportart abbilden, z.B. indem in unserem Beispiel des Aufschlags im Beachvolleyball die Änderung der Ass-Rate zu 80% durch die Senkung der Fehlerrate und zu 20% durch eine Senkung der Annahmerate kompensiert wird.

Welche Form der Kompensation schließlich angewandt wird, ist abhängig von der anwendungsspezifischen Fragestellung zu entscheiden.

2.1.3. Numerische Differentiation

Nachdem in der Markov-Studie die numerische Differentiation als Mittel zur Auslenkung der Übergangsmatrizen verwendet wird, soll hier kurz deren Methodik erläutert werden.

Die numerische Differentiation findet vor allem dann Anwendung, wenn keine analytische Form der abzuleitenden Funktion vorliegt. Dies ist besonders häufig in der Datenverarbeitung der Fall, in der Funktionen durch diskrete Messwerte charakterisiert werden. Diese Situation tritt auch bei der Simulation der Gewinnwahrscheinlichkeiten mit Hilfe von Markov-Ketten ein. Hier lässt sich die Berechnung der Gewinnwahrscheinlichkeit als eine Funktion ausdrücken, die eine komplette Übergangsmatrix als Eingangswert x entgegennimmt und die dazugehörige Gewinnwahrscheinlichkeit $f(x)$ ausgibt.

$f(x)$ beschreibt also die Gewinnwahrscheinlichkeit in Abhängigkeit aller möglicher Kombinationen von Wahrscheinlichkeiten in der Übergangsmatrix. Die Idee der Simulation durch Markov-Ketten, durch Anpassungen an der Übergangsmatrix taktische Verhaltensweisen zu analysieren, lässt sich nun auch auf die Funktion $f(x)$ übertragen. Dazu kann die numerische Differentiation verwendet werden, die im Grunde das Maß der Anpassung der Übergangsmatrizen vorgibt. Dabei wird die Definition der Ableitung über den Differenzialquotienten verwendet (Westermann, 2011):

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

Geometrisch wird also die Steigung der Tangente im Punkt $f(x)$ betrachtet. Nachdem der Grenzübergang $\lim_{h \rightarrow 0}$ numerisch allerdings nicht durchgeführt werden kann, wird die Ableitung einer Funktion im Punkt x durch die Sekantensteigung

$$f'(x) = \frac{f(x+h) - f(x)}{h}$$

mit $h > 0$ angenähert. Genauere Ergebnisse erhält man mit der Verwendung der sogenannten zentralen Differenzenformel:

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h}$$

Dabei repräsentiert x eine Konfiguration der Übergangsmatrix, während h die Amplitude der Auslenkung für diese Konfiguration beschreibt.

Anwendung im Sport

Im Gegensatz zu der in Abbildung 5 gezeigten Methode der Auslenkung wird mit der numerischen Differentiation kein zusätzliches Modell für die Auslenkung der Übergangswahrscheinlichkeiten benötigt. Dadurch wird der Aufwand für Analysen reduziert, da für verschiedene Fragestellungen nicht mehr einzelne Modelle der Auslenkung benötigt werden und schlicht die mathematisch vorgegebenen Schrittgrößen verwendet werden können. Zusätzlich dazu eröffnet dies die Möglichkeit automatisierter Analysen mit wechselnden Parametern, während sich gleichzeitig die Wahrscheinlichkeit für Modellierungsfehler verringert. Durch diese Vorteile gestaltet sich die Anwendung der Markov-Simulation effizienter und gleichzeitig robuster, was sie für den Einsatz in der Leistungsdiagnostik attraktiv macht.

2.2. Data-Mining

Wie im einleitenden Kapitel bereits geschildert, spielt die Erfassung, Speicherung und Analyse von Daten in der heutigen Zeit eine entscheidende Rolle in vielen Lebensbereichen. Im Data-Mining beschäftigt man sich deshalb mit der Extraktion von relevanten Informationen aus großen Datenmengen. Relevante Informationen sind solche, die neu, nützlich, nicht trivial und verständlich sind, was natürlich für jedes Anwendungsgebiet einzeln evaluiert werden muss und vom Vorwissen der Anwendungsexperten abhängt (Fayyad et al., 1996). Zum Einsatz kommen dazu neben konventionellen statistischen Verfahren wie Korrelation und Regression auch Methoden aus den Bereichen Mustererkennung, Klassifikation, Clusteranalyse und Machine-Learning (Runkler, 2010).

Data-Mining selbst ist ein Prozess, der häufig in den übergeordneten Prozess des Knowledge Discovery in Databases (KDD) eingeordnet wird (Hand, 2007). Dieser Prozess besteht aus mehreren Phasen (Runkler, 2010): Sollten noch keine Daten vorhanden sein, muss die Datenerfassung geplant und durchgeführt werden. Häufig wird Data-Mining aber auf einem bereits vorhandenen, zu anderen Zwecken gesammelten Datensatz durchgeführt (Hand, 2007). Diese Rohdaten sind häufig fehlerhaft, redundant oder unvollständig, weswegen vor der eigentlichen Datenanalyse der Schritt der Datenvorverarbeitung durchgeführt wird. In diesem Schritt werden die Daten je nach Notwendigkeit normalisiert, gefiltert oder transformiert und Fehler korrigiert. Der dadurch generierte „saubere“ Datensatz kann schließlich mit den bereits erwähnten Methoden der Datenanalyse verarbeitet werden. Im letzten Schritt werden die Ergebnisse der Analyse interpretiert und dokumentiert, um sie auswertbar zu machen. Der komplette Prozess ist in Abbildung 6 noch einmal dargestellt. Er muss jedoch nicht immer wie abgebildet linear durchlaufen werden, sondern kann je nach den Ergebnissen der einzelnen Teilschritte Phasen überspringen oder wiederholt durchlaufen.

Eines der bekanntesten Anwendungsgebiete des Data-Mining sind die sogenannten Warenkorbanalysen, die darauf abzielen, Präferenzen im Kaufverhalten der Kunden zu entdecken. So können im stationären Handel beispielsweise häufig zusammen gekaufte Produkte nah zusammen gestellt werden, um den Umsatz zu erhöhen. Im Onlinehandel werden die Ergebnisse von Warenkorbanalysen unter anderem dazu eingesetzt, Empfehlungen für interessante Produkte auf Basis des Kaufverhaltens auszusprechen.

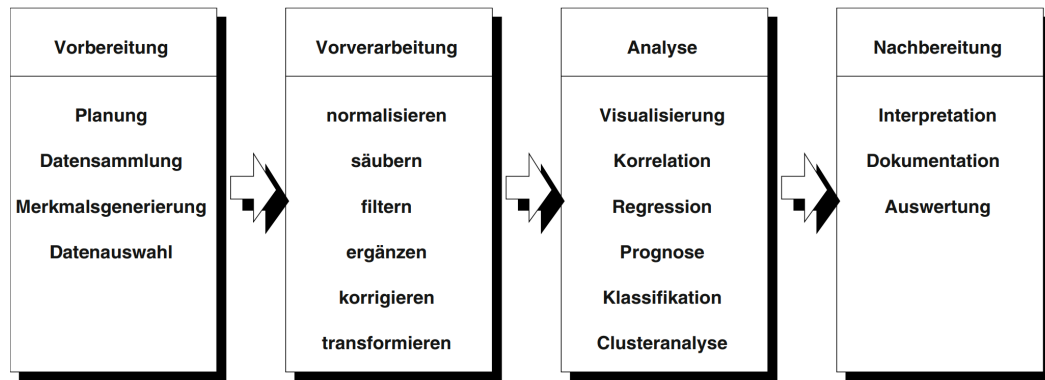


Abbildung 6 Schritte des Data-Mining Prozesses nach Runkler (2010)

Dazu sollten ursprünglich sogenannte *frequent itemsets*, also Gruppen von Gegenständen, die häufig zusammen gekauft werden, in einer Datenbank von Kundentransaktionen gefunden werden.

Ein berühmtes Beispiel für eine solche Assoziation ist die Regel von „Bier und Windeln“ (Power, 2002). Sinngemäß stellt sich die Geschichte dieser Regel folgendermaßen dar: Eine große amerikanische Supermarktkette analysierte die Einkäufe ihrer Kunden und fand heraus, dass junge männliche Käufer, die an Wochenenden Windeln einkaufen, häufig auch Bier mit nach Hause nahmen. Basierend auf dieser Erkenntnis entschloss sich der Supermarkt dazu, diese beiden Produkte nahe zusammen in den Regalen zu platzieren, wodurch der Umsatz an Bier deutlich gesteigert werden konnte. Auch wenn der Wahrheitsgehalt dieser Anekdote umstritten ist, steht sie heutzutage fast sinnbildlich für die Aufgabe des Data-Mining, zuvor unbekannte Zusammenhänge zwischen Variablen in einem Datensatz zum Vorschein zu bringen. Die zugehörigen Daten werden von sogenannten „Assoziationsanalysen“ ausgewertet, die im Folgenden genauer beschrieben werden.

2.2.1. Frequent Itemset Mining

Obwohl das Frequent Itemset Mining (FIM) im Anwendungsfeld der Warenkorb-Analysen bekannt wurde, wird es heute als eine universelle Methode des Data-Mining angesehen, die in vielen Domänen anwendbar ist. Da sich viele Arten von Daten analog zu Transaktionen bei Einkäufen modellieren lassen, hat FIM bereits Anwendungen in Feldern wie der Bioinformatik (Naulaerts et al., 2015), der Klassifizierung von Bildern (Fernando et al., 2012), Netzwerk-Analysen (Glatz et al., 2014), Malware-Detection (Duan et al., 2015) oder E-Learning (Mwamikazi et al., 2014) gefunden.

Formal wird das Problem des Frequent Itemset Mining folgendermaßen definiert (Agrawal, Srikant et al., 1994):

Es existiere eine Menge von Items $I = \{i_1, i_2, \dots, i_m\}$. Die Datenbank der Transaktionen $D = \{T_1, T_2, \dots, T_n\}$ ist dann eine Menge von Transaktionen $T_q \subseteq I (1 \leq q \leq n)$, so dass jede Transaktion eine Menge eindeutiger Items mit einer eindeutigen Bezeichnung (TID) darstellt. Tabelle 1 zeigt ein Beispiel für solch eine Datenbank mit den fünf Transaktionen, die die Items a, b, c, d, e enthalten können.

Ein Itemset X ist nun eine Menge von Items, so dass $X \subseteq I$. Das Ziel des Frequent Itemset Mining ist es, interessante Itemsets in einer Transaktionsdatenbank zu finden. Dabei gibt es im Feld des Itemset-Mining einige Methoden, um zu messen, wie interessant ein Itemset ist. Ein Standardmaß stellt der sogenannte *Support* $sup(X)$ eines Itemsets dar, der häufig entweder absolut als die Anzahl der Transaktionen, die das Itemset enthalten $sup(X) = |\{T|X \subseteq T \wedge T \in D\}|$, oder relativ als $sup(X) = sup(X)/|D|$ angegeben wird. In unserem Beispiel wäre also der Support des Itemsets $\{a, b\}$ 2 oder 0.4, da es in zwei von fünf Transaktionen auftritt (T_3 und T_5). Die meisten Algorithmen des FIM akzeptieren deshalb einen Parameter für den minimalen Support, den die gefundenen Itemsets aufweisen müssen. Würde ein Benutzer bei einem solchen Algorithmus den minimalen Support zum Beispiel auf 3 festlegen, würde das zuvor genannte Itemset $\{a, b\}$ nicht als interessantes Itemset zurückgegeben werden.

TID	Transaktion
T_1	$\{a, c, d\}$
T_2	$\{b, c, e\}$
T_3	$\{a, b, c, e\}$
T_4	$\{b, e\}$
T_5	$\{a, b, c, e\}$

Tabelle 1 Beispiel für eine Datenbank von Kundentransaktionen mit der Transaction ID (TID) und dem Itemset der Transaktion. Die erste Transaktion repräsentiert einen Kunden, der die Items a, c und d gekauft hat. Entnommen aus Fournier-Viger et al. (2017).

2.2.2. Association Rule Mining

Das Association-Rule-Mining (ARM) (Agrawal, Srikant et al., 1994) erweitert das Frequent Itemset Mining nun dadurch, dass nicht mehr Itemsets, sondern Regeln der Form $X \rightarrow Y$ gefunden werden sollen, wobei X und Y ihrerseits disjunkte Itemsets darstellen ($X \cap Y = \emptyset$). Die Interessantheit solcher Regeln wird dabei normalerweise durch den Support $sup(X \rightarrow Y) = sup(X \cup Y)$ und die Konfidenz $conf(X \rightarrow Y) = sup(X \cup Y)/sup(X)$ gemessen. Die Konfidenz stellt ein Maß für die Stärke der Assoziation zweier Itemsets dar, indem es die bedingte Wahrscheinlichkeit $P(Y|X)$ repräsentiert. Dadurch adressieren Association Rules eine Schwäche des Frequent Itemset Mining, das häufig viele Itemsets findet, die nur schwach assoziiert sind. Das Itemset $\{a, c\}$ beispielweise weist einen Support von 3 auf. Die Regel $\{a\} \rightarrow \{c\}$ besitzt jedoch eine Konfidenz von 100%, während die Regel $\{c\} \rightarrow \{a\}$ nur eine Konfidenz von 75% aufweist. Association Rules können also detailliertere Informationen über Itemsets darstellen, als reine Itemsets es vermögen. Neben Support und Konfidenz gibt es zudem mehr als 20 weitere Maße für die Interessantheit von Assoziationsregeln (Lencica et al., 2007), wodurch weitere Analysen der generierten Regeln möglich werden.

Anwendung im Sport

Für die Anwendung im Sport sind Association Rules besonders dann interessant, wenn Zusammenhänge im Auftreten diskreter Ereignisse untersucht werden sollen. Dabei kann es sich um technische oder taktische Merkmale, aber auch um Positionen oder Umweltvariablen handeln. Da die Algorithmen zumeist Eingabeparameter für den minimalen Support und die minimale Konfidenz, die die generierten Regeln aufweisen sollen, anbieten, lassen sie sich relativ gut auf die Anforderungen der jeweiligen sportwissenschaftlichen Fragestellung anpassen. Außerdem existieren weitere Erweiterungen des Association Rule Minings, wie zum Beispiel das Sequential Rule Mining, das die zeitliche Reihenfolge der Ereignisse berücksichtigt (Fournier-Viger et al., 2014). Damit lässt sich der zeitliche Kontext von Aktionen im Sport in die Analysen aufnehmen, der besonders in Rückschlagspielen wie im Beachvolleyball und Tischtennis eine Rolle spielt.

2.3. Machine-Learning

Heutzutage ist die Künstliche Intelligenz (KI) sowohl in der Forschung als auch in der praktischen Anwendung ein rasant wachsendes Themengebiet. Die Anwendungsfelder reichen dabei von der Automatisierung, autonomen Systemen, über die Sprach- und Bilderkennung bis zur Diagnostik in der Medizin (Goodfellow et al., 2016). Mit der kontinuierlichen Entwicklung der KI verändern sich auch die behandelten Themenfelder. Zu Beginn wurden hier Aufgaben bearbeitet, die nur schwer von Menschen, aber dafür relativ einfach von einer Maschine gelöst werden konnten - meistens Problemstellungen, die sich durch eine Reihe formaler mathematischer Regeln beschreiben lassen. Als die am schwersten zu lösenden Probleme gelten heute allerdings diejenigen, die ein Mensch intuitiv und zumeist automatisch lösen kann, wie zum Beispiel das Erkennen von Gesichtern und Sprachen - also gerade jene Klasse von Aufgaben, die sich nur schwer formalisieren lassen. Gerade hier gelangen in den letzten Jahren mit dem Aufkommen des Machine-Learning und insbesondere des sogenannten tiefen maschinellen Lernens (Deep Learning, DL) Fortschritte.

Das Machine-Learning stellt ein Teilgebiet der Künstlichen Intelligenz dar, dessen Ziel grob gesagt das Verständnis von Daten darstellt. Angelehnt an das menschliche Verhalten soll es Computern zum Einen ermöglicht werden, aus der Erfahrung zu lernen, und zum Anderen, ihr Problemverständnis hierarchisch aufzubauen. Durch diesen Ansatz wird es vermieden, dass ein menschlicher Experte das formale Wissen definieren muss, das eine Maschine zur Lösung eines spezifischen Problems benötigt. Der hierarchische Aufbau der gelernten Konzepte ermöglicht es Computern zudem, komplizierte Gebilde durch das Zusammensetzen einfacherer Lösungen zu verstehen. Hat dieser hierarchische Aufbau viele Ebenen, spricht man vom Deep Learning (Goodfellow et al., 2016).

Die meisten Probleme des maschinellen Lernens lassen sich in zwei Kategorien einteilen: Das überwachte (*Supervised-Learning*) und das unüberwachte (*Unsupervised-Learning*) Lernen. Beim Supervised-Learning wird versucht, ein Modell zu finden, das eine Antwortvariable mit den Eingabedaten in Verbindung bringt. Dies kann zum Ziel haben, die Antwortvariable möglichst genau vorherzusagen oder ein besseres Verständnis des Zusammenhangs zwi-

schen Eingabedaten und Antwortvariablen zu bekommen. Dies setzt natürlich voraus, dass die verfügbaren Daten eine solche Antwortvariable beinhalten. Ein einfaches Beispiel für ein Problem des Supervised-Learning ist die Vorhersage von Immobilienpreisen auf Basis von Daten wie zum Beispiel dem Standort der Immobilie, der Entfernung zu Freizeitangeboten oder Anstellungsmöglichkeiten durch Firmen in der Umgebung. Viele „klassische“ Lernverfahren, wie die lineare oder logistische Regression, gehören zur Kategorie des Supervised-Learning (James et al., 2013).

Im Gegensatz dazu beschreibt das Unsupervised-Learning die Situation, in der eine solche Antwortvariable nicht gegeben ist. Da in diesem Fall keine Variable vorhanden ist, die die Analyse „überwachen“ kann, wird meistens versucht, die Beziehung zwischen den Eingabevariablen oder verschiedener Beobachtungen innerhalb der Daten zu erforschen. Eine bekannte Methode aus der Domäne des Unsupervised-Learning ist die Cluster-Analyse, deren Ziel es ist, voneinander möglichst unterschiedliche Gruppen von Beobachtungen in einem Datensatz zu finden. Bleiben wir bei dem Beispiel der Immobilienpreise, könnte eine Cluster-Analyse Gruppen ähnlicher Immobilien basierend auf deren Preis, Lage, Alter, Größe und weiteren Merkmalen identifizieren.

Die in dieser Arbeit verwendeten Methoden fallen dabei allesamt in die Kategorie des Supervised-Learning, auf dessen Funktionsweise im Absatz 2.3.2 genauer eingegangen wird. Die meisten Modelle des Machine-Learning stellen sich als sogenannte Black-Box-Modelle dar, was bedeutet, dass nicht nachvollziehbar ist, wie die Eingabewerte der Modelle zu Ausgabewerten transformiert werden.

2.3.1. Neuronale Netze

Caudill (1987, S. 46) zitiert Dr. Robert Hecht-Nielson, den Erfinder eines der ersten Neuro-Computer, mit folgender recht einfachen Definition eines neuronalen Netzes als

“...a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs.”

[ein Computersystem, das aus einer Anzahl einfacher, untereinander hochgradig verbundener Rechenelemente besteht, die Informationen durch die Antwort auf externe Eingaben anhand ihres dynamischen Zustands verarbeiten.]

Die Struktur neuronaler Netze wurde dabei lose dem Aufbau des menschlichen Gehirns nachempfunden, zuerst allerdings in weitaus kleinerem Umfang. Durch den Fortschritt der verfügbaren Hardware erreichen neuronale Netze heute jedoch eine Komplexität, die der eines menschlichen Gehirns zumindest in der Anzahl der „Rechenelemente“ kaum noch nachsteht (Azevedo et al., 2009; Trask et al., 2015).

Neuronale Netze sind typischerweise in Ebenen (Layers) organisiert (siehe Abbildung 7), die aus einer Menge untereinander vernetzter Knoten, auch Neuronen genannt, bestehen.

Analog zu den elektrischen Signalen eines Gehirns können diese Neuronen basierend auf einer mathematischen Funktion, die für jedes Neuron gegeben ist, aktiviert werden. Diese

Funktion wird Aktivierungsfunktion genannt. Die Eingabedaten für die Neuronen werden also zuerst an die sogenannte Input-Layer übergeben und anschließend durch die Hidden-Layers über gewichtete Verbindungen zur Output-Layer weitergereicht.

Die eigentliche „Berechnung“ geschieht in den Hidden-Layers durch die Aktivierungsfunktion an den einzelnen Neuronen. Der Vorgang dieser Berechnung wird in Abbildung 8 dargestellt. Wie zu erkennen ist, wird die Aktivierungsfunktion auf die gewichtete Summe der Eingaben eines Neurons angewandt, um zu entscheiden, ob das Neuron aktiviert wird. Ohne diese zusätzliche Anwendung der Aktivierungsfunktion wären neuronale Netze ein lineares Modell, dessen Lösung zwar einfacher zu berechnen wäre, das aber die Möglichkeit verlieren würde, komplexe nicht-lineare Zusammenhänge zu erfassen.

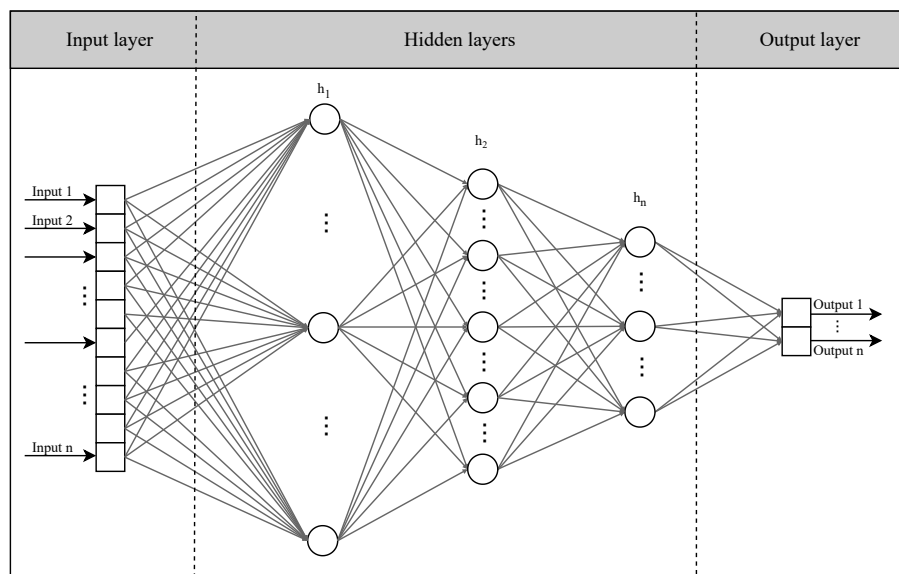


Abbildung 7 Schematische Darstellung eines einfachen neuronalen Netzwerks, auch Multilayer-Perceptron (MLP) oder fully connected network genannt.

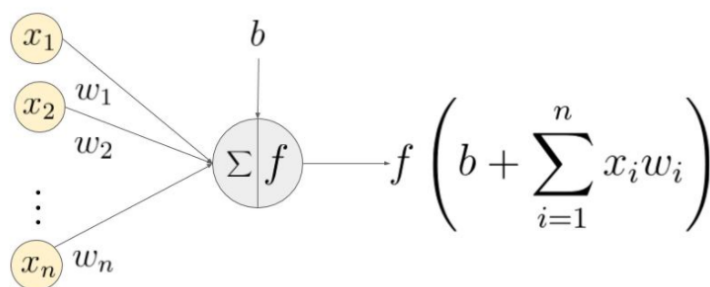


Abbildung 8 Beispiel eines einzelnen Neurons mit den Eingangsdaten ($x_1 - x_n$), den korrespondierenden Gewichten ($w_1 - w_n$), dem Bias b und der Aktivierungsfunktion f , die auf die gewichtete Summe der Eingaben angewandt wird. („Activation functions and its Types in Artificial Neural network“, 2018).

Kurz gesagt ist die Aufgabe der Aktivierungsfunktion, Neuronen aktiv zu schalten, sobald deren Eingabewerte ihren Schwellenwert überschreiten. Um dadurch nicht-Linearität im Modell des Netzwerks zu ermöglichen, sollte diese Funktion selbst also ebenfalls nicht-linear sein. Goodfellow et al. (2016) empfehlen heutzutage die in Abbildung 9 gezeigte Rectified Linear Unit (ReLU) als Standard-Aktivierungsfunktion für die meisten neuronalen Netze. Ih-

re Anwendung stellt zwar eine nicht-lineare Transformation dar, die Funktion selbst ist aber nahezu linear, da sie aus zwei linearen Teilfunktionen besteht. Dadurch bleiben dieser Funktion viele Eigenschaften erhalten, die sie wie lineare Funktionen leicht zu optimieren machen. Dies ist vorteilhaft für den eigentlichen Schritt des „Lernens“, welcher im nächsten Abschnitt genauer dargestellt wird.

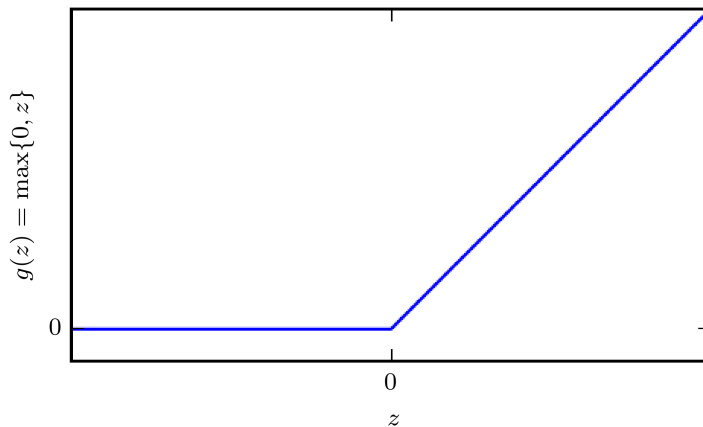


Abbildung 9 Die Rectified Linear Unit (ReLU) Funktion. Grafik entnommen aus Goodfellow et al. (2016).

2.3.2. Supervised-Learning

Im Fall des Supervised-Learnings, in dessen Bereich die Methoden dieser Arbeit fallen, lernen die meisten neuronalen Netze durch die Anpassung der Gewichte an den Verbindungen zwischen den Neuronen, ausgehend von den präsentierten Mustern der Input-Layer. Solche Netze werden „trainiert“, indem sie mit zufälligen Werten für die Gewichte \vec{w} initialisiert werden. Anschließend wird das neuronale Netz auf einem Datensatz mit den Input-Variablen X und den dazugehörigen Ausgabewerten \vec{y} angewandt, wodurch neue Ausgabewerte $\vec{\hat{y}}$ entstehen. Dieser Schritt wird auch *Forward-Pass* genannt. Die Ausgabewerte $\vec{\hat{y}}$ des Netzes sind also eine Funktion der Eingabewerte X , der zufällig initialisierten Gewichte \vec{w} und der Wahl der Aktivierungsfunktion T .

$$\vec{\hat{y}} = f(X, \vec{w}, T)$$

Die Differenz zwischen den durch das Netz generierten Werten $\vec{\hat{y}}$ und den im Datensatz vorhandenen „echten“ Ausgabewerten \vec{y} wird nun als Leistungsmerkmal für das Netz verwendet. Konnte eine gute Abbildung der Eingangswerte auf die gewünschten Ausgabewerte gelernt werden, so ist diese Differenz gering und das Netz weist einen geringen Fehler auf. Die Fehlerdifferenz wird dabei durch eine Kostenfunktion C berechnet, die durch den Anwender wählbar ist. Als Beispiele für solche Kostenfunktionen lassen sich Funktionen wie der quadratische Fehler oder die mittlere quadratische Abweichung

$$C = \frac{1}{n} \sum_{i=1}^n (y^i - \hat{y}^i)^2$$

anführen. Um eine möglichst gute Annäherung für die Funktion f zu finden, ist es also die Aufgabe des Netzes, diesen Fehler zu minimieren. Dazu kann nach dem Gradientenverfah-

ren f partiell nach den Gewichten \vec{w} abgeleitet werden. Dadurch ergibt sich ein Vektor, der die (lokal) optimale Änderung der Gewichte an der letzten Netzwerkebene (Output Layer) beschreibt. Diese Änderungen können mit Hilfe der Kettenregel durch das gesamte Netz bis an die Eingangsschicht zurückpropagiert werden, wodurch sich auch der Name dieses Verfahrens („Backpropagation“) herleitet. Der gesamte Prozess mit Forward-Pass, Fehlerberechnung und Backpropagation des Fehlers kann nun so lange wiederholt werden, bis entweder eine festgelegte Anzahl an Iterationen oder ein zuvor festgelegtes Leistungsniveau (z.B. minimale Fehlerrate) erreicht wurde.

Im Folgenden wird zusätzlich auf spezielle Architekturen von neuronalen Netzwerken, die in dieser Arbeit eine Rolle spielen, eingegangen. Unterschiede bestehen dabei im Aufbau der Ebenen von Neuronen innerhalb der Netze. Die Funktionsweise des Trainings und des Lernens ändert sich hingegen nicht.

2.3.3. Convolutional Neural Networks

Convolutional neural networks (CNN) (LeCun et al., 1989), zu deutsch etwa „faltende neuronale Netze“, wurden speziell dazu konstruiert, Daten zu verarbeiten, die in einer Gitterstruktur vorliegen. Insofern ist es wenig verwunderlich, dass diese Netzwerke heutzutage vorzugsweise in der Bildverarbeitung eingesetzt werden, mit den Pixeln der jeweiligen Farbkanaäle des Bildes als Eingabewerte. Aber auch Zeitserien, die als eindimensionales Gitter betrachtet werden können, kann diese Netzarchitektur effizient bearbeiten (Goodfellow et al., 2016).

Auch CNNs lernen durch die Anpassung ihrer Gewichte durch Backpropagation des Trainingsfehlers, die allerdings in sogenannten „Kernels“ organisiert sind (siehe Abbildung 10). Typischerweise besteht eine Ebene in einem CNN dann aus drei Stufen (Goodfellow et al., 2016): Zuerst wird der Kernel mit der mathematischen Operation der Faltung über das Gitter der Eingabewerte geschoben. Die resultierenden Gitter besitzen nun eine reduzierte Auflösung und stellen die Aktivierung des Netzwerkes in dem Bereich des Eingabe-Gitters dar, auf dem der Kernel angewandt wurde. Auf diesem reduzierten Gitter wird dann eine nicht-lineare Aktivierungsfunktion (wie z.B. ReLU) angewandt, bevor das sogenannte *Pooling* eine nicht-lineare Form des Downsamplings durchführt. Die Pooling-Funktion ersetzt die Ausgabe-Werte der Faltung durch eine übergreifende Statistik der benachbarten Ausgaben. So verwendet beispielsweise das *max pooling* (Zhou & Chellappa, 1988) den maximalen Wert innerhalb einer rechteckigen Nachbarschaft eines Ausgabewertes. Der Vorteil des Poolings ist zum Einen die Dimensionsreduktion der Eingaben, wodurch Rechenleistung gespart werden kann. Zum Anderen hilft es dabei, das Netzwerk robust gegenüber kleiner Variation der Eingabewerte zu machen (Goodfellow et al., 2016).

In vielen Fällen enthalten CNN-Architekturen nach den faltenden Ebenen noch zusätzliche MLP-Ebenen zur Klassifizierung der gewünschten Ausgabewerte, wodurch sich die Gesamtarchitektur meist ähnlich zu Abbildung 11 darstellt.

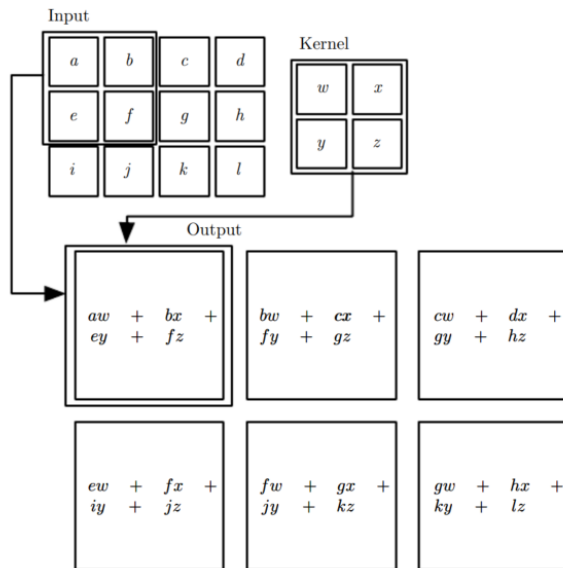


Abbildung 10 Beispielhafte Darstellung der Faltung der Eingabewerte $a - l$ mit dem Kernel $w - z$. Der Kernel wird sowohl auf der horizontalen als auch auf der vertikalen Achse über die Eingabedaten angewandt. Grafik entnommen aus Goodfellow et al. (2016).

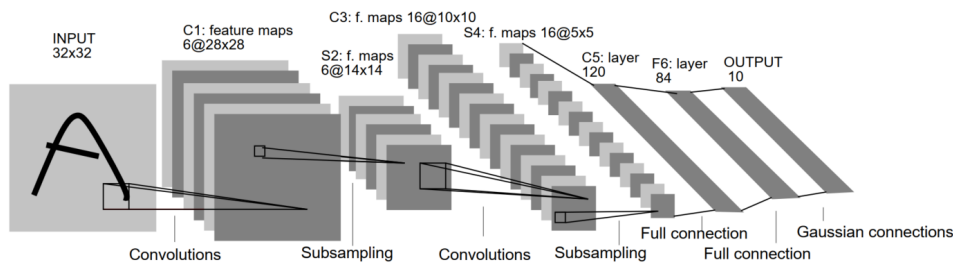


Abbildung 11 Architektur des LeNet-5, eines der ersten CNNs zur Bilderkennung. Grafik entnommen aus LeCun et al. (1998).

2.3.4. Recurrent Neural Networks

Ein Nachteil der bisher beschriebenen Netzwerkarchitekturen besteht darin, dass diese die Muster in den Trainingsdaten völlig ohne Kontext interpretieren. So betrachtet zum Beispiel ein CNN zur Bilderkennung in einem Trainingsdurchgang nur die Pixel eines Bildes, ohne zu beachten, welche Bilder es bereits in vorangegangenen Trainingsläufen bearbeitet hat. Diese Netze besitzen also kein „Gedächtnis“, keine Möglichkeit, während des Trainings gelernte Muster mit zuvor aufgetretenen Informationen in Verbindung zu setzen.

Recurrent Neural Networks (Rumelhart et al., 1986) (zu deutsch rekurrente neuronale Netze) behandeln diese Problematik, indem sie Schleifen, also rückläufige Verbindungen zu Neuronen anderer Schichten, sowie die persistente Speicherung von Zuständen ermöglichen. Diese Schleifen können einfach als Kopien desselben Netzwerkes gesehen werden, die Informationen an ihre Nachfolger weitergeben (siehe Abbildung 12). So wie CNNs prädestiniert für den Einsatz von in Gittern organisierten Daten sind, macht die kettenartige Struktur rekurrenter Netzwerke diese zu der Netzarchitektur, die auf die Verarbeitung von (temporalen) Sequenzen und Listen spezialisiert ist (Goodfellow et al., 2016). Insofern ist es wenig verwunderlich, dass diese Art von Netzen besonders bei der Text- und Sprachverarbeitung

erfolgreich eingesetzt wurde (Graves & Schmidhuber, 2008; Sak et al., 2014).

Trainiert werden rekurrente Neuronale Netze meistens mit einer abgewandelten Form der Backpropagation, der sogenannten Backpropagation through time (BPTT). Hierbei wird die Backpropagation auf die entrollte Form (Abb. 12) des neuronalen Netzes angewandt, indem der Gesamtfehler des Netzes als Summe der Fehler der einzelnen Zeitschritte des RNNs angesehen wird und die Ableitungen dieser Fehler mit der Kettenregel rückwärts durch die Zeit fortgesetzt werden.

Dabei kann man sich leicht vorstellen, dass die Werte dieser Ableitungen besonders bei langen Sequenzen, wie zum Beispiel Sätzen in einem Text, entweder schnell wachsen (*Exploding Gradient Problem*) oder schnell schrumpfen (*Vanishing Gradient Problem*) können.

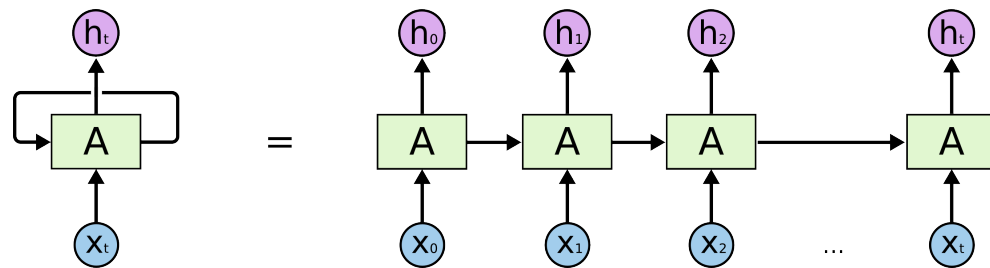


Abbildung 12 Links: Ein Neuronales Netz, das eine Schleife enthält. Rechts: Aufgerollte Form desselben neuronalen Netzes für eine Eingabesequenz der Länge t . Grafik entnommen aus Olah (2015).

Das bewirkt, dass diese Neuronen kaum noch Einfluss auf den Lernprozess des Netzwerks ausüben und damit ein „Langzeitgedächtnis“ nur schwer zu bewerkstelligen ist. Treten diese Probleme auf, können also nur ein paar wenige der vorangegangenen Zeitschritte das Lernergebnis beeinflussen.

Als Lösung wurden spezialisierte RNN-Architekturen erfunden, die das Problem des verschwindenden und explodierenden Gradienten beheben. Sowohl die Long-Short-Term-Memory (LSTM)-Netze (Hochreiter & Schmidhuber, 1997) als auch die jüngeren Gated Recurrent Unit (GRU) (Cho et al., 2014) lösen diese Probleme durch sogenannte Schleusen (Gates), die regulieren, welche Informationen in einem Zeitschritt gespeichert oder vergessen werden.

Interessanterweise können RNNs auch als Programme gesehen werden, die Eingabedaten mit ihrem internen Zustand und einer fixen (gelernten) Funktion in Ausgabedaten umwandeln. Tatsächlich konnte bereits nachgewiesen werden, dass rekurrente Netze Turing-vollständige, also universell programmierbare, Systeme darstellen (Siegelmann & Sontag, 1995).

2.3.5. Entscheidungsbäume

Genauso wie neuronale Netze können Entscheidungsbäume als Methode des Machine Learnings zur Klassifikation oder Regression einer Zielvariable auf Basis eines Trainingsdatensatzes verwendet werden. Die Grundidee der dazu verwendeten Algorithmen besteht darin, die Trainingsdaten rekursiv in rechteckige Partitionen aufzuteilen, so dass eine möglichst eindeutige Zuweisung von Partitionen zu den Klassen der Zielvariable entsteht (Abb. 13).

Die ersten veröffentlichten Algorithmen zur Erstellung solcher Entscheidungsbäume stammen aus den 1970er Jahren (Fielding & O'Muircheartaigh, 1977; Messenger & Mandell, 1972). Aufbauend darauf entstanden die heute häufig verwendeten Algorithmen C4.5 (Quin-

lan, 1993) und CART (Breiman et al., 1984).

Entscheidungsbäume haben einige Vorteile gegenüber neuronalen Netzen: Sie sind einfach zu visualisieren und damit einfach zu verstehen und zu interpretieren. Dabei sind sie sehr flexibel und operieren sowohl auf diskreten als auch kontinuierlichen Daten, ohne großen Aufwand bei der Vorbereitung der Daten zu verursachen. Allerdings neigen sie dazu, komplexe Bäume zu generieren, die nur schlecht auf Daten außerhalb des Trainingsdatensatzes anwendbar sind (*Overfitting*) (Strobl et al., 2009). Hinzu kommt, dass schon kleine Änderungen an den Trainingsdaten große Änderungen an den generierten Bäumen bewirken können (*Varianz*), wodurch die Ergebnisse nicht robust sind (James et al., 2013).

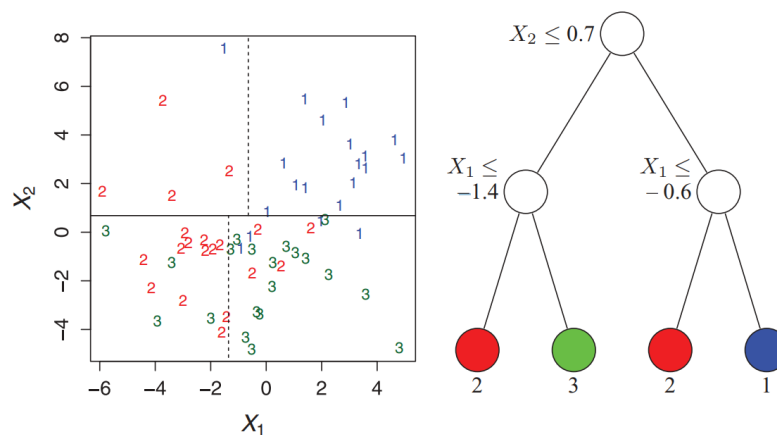


Abbildung 13 Links: Partitionierte Daten. Rechts: Entscheidungsbaum für Klassifikationsmodell mit den Klassen 1, 2 und 3. An jedem Zwischenknoten wird der linke Kindknoten ausgewählt, falls die Bedingung des Knotens zutrifft. Andernfalls wird der rechte Kindknoten ausgewählt. Die vorhergesagte Klasse ist an den Blättern des Baumes dargestellt. Grafik entnommen aus Loh (2011).

Um diese Schwächen auszugleichen existieren verschiedene Methoden, von denen hier insbesondere auf die *Random-Forests* und das *Boosting* eingegangen werden soll. Random-Forests generieren dazu nicht nur einen Baum, sondern eine große Anzahl unabhängig voneinander initialisierter Entscheidungsbäume, die das Ergebnis der Klassifizierung entweder als eine Art Mehrheitsentscheidung oder Durchschnitt der einzelnen Bäume bilden (James et al., 2013). Random-Forests zählen damit zur Klasse der *Ensemble-Methoden*, die eine finite Kombination von Machine-Learning-Modellen verwenden, um bessere Ergebnisse als ein einzelnes Modell zu erzielen. Auch das Boosting - im Fall von Entscheidungsbäumen spricht man häufig von *Boosted-Trees* - zählt zu den Ensemble-Methoden. Die Idee des Boostings ist dabei, aus mehreren „schwachen“, aber schnell zu berechnenden Lernmodellen ein starkes, leistungsfähiges Modell zusammenzusetzen (Freund et al., 1999).

Ähnlich zu den Random-Forests wird beim Boosting eine Mehrzahl von Entscheidungsbäumen generiert, die allerdings nicht unabhängig voneinander, sondern aufeinander aufbauend konstruiert werden. Das Boosting gewichtet dazu die Eingangsdaten nach der Schwierigkeit für den Algorithmus, diese Daten korrekt zu klassifizieren. Diese Gewichte werden analog zu den neuronalen Netzen während des Trainingsprozesses kontinuierlich angepasst. Nachdem der erste Entscheidungsbaum konstruiert wurde, optimiert das Boosting den nächsten Baum dermaßen, dass dieser gerade für die schwach gewichteten Fälle des ersten Baumes mög-

lichst gute Leistung zeigt (James et al., 2013). Dieser Prozess wird fortgesetzt, bis entweder eine gewisse Anzahl von Iterationen stattgefunden hat oder keine Verbesserung der Klassifikationsgenauigkeit mehr festgestellt werden kann. Das Ergebnis der Klassifikation ergibt sich nun wieder aus der Mehrheitsentscheidung oder dem Durchschnitt der Ergebnisse der einzelnen Bäume.

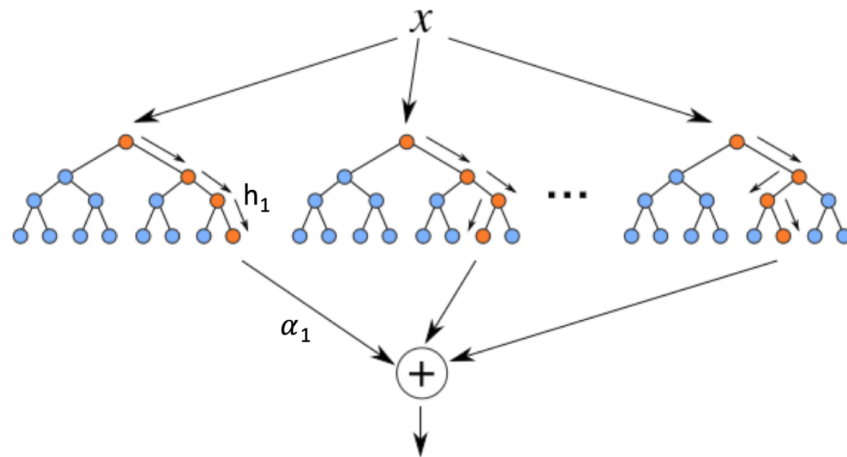


Abbildung 14 Schematische Darstellung eines Tree-Ensembles. Die finale Klassifizierung erfolgt durch Votum der einzelnen Entscheidungsbäume. Grafik entnommen aus Shoaran et al. (2018).

Anwendung im Sport

Wie in den vorangegangenen Abschnitten bereits beschrieben, haben die verschiedenen Methoden des Machine Learning unterschiedliche Stärken, die sie für den Einsatz in bestimmten Gebieten (z.B. Sprach- oder Bilderkennung) empfehlen. Allerdings sind sie damit natürlich nicht auf diese Einsatzgebiete beschränkt. Durch die richtige Architektur und eine passende Modellierung der Eingabedaten lassen sich die vorgestellten Methoden auf vielfältige sportwissenschaftliche Fragestellungen anwenden.

Das in dieser Arbeit relevante sportwissenschaftliche Problem ist die Suche nach Mustern von technischen und taktischen Merkmalen innerhalb von Ballwechseln im Beachvolleyball. Grundsätzlich sind die hier vorgestellten Methoden dazu in der Lage, dieses Problem zu bearbeiten. Dementsprechend stellt sich zusätzlich die Frage, ob eine der Techniken Vorteile bei der Analyse bietet. Dies soll durch einen Vergleich der Klassifikationsgenauigkeit der Methoden in dieser Arbeit beantwortet werden.

3. Einzelbeiträge

3.1. Studie 1 - Performance Analysis in Table Tennis - Stochastic Simulation by Numerical Derivation

Wenninger, S., & Lames, M. (2016). Performance Analysis in Table Tennis - Stochastic Simulation by Numerical Derivation, *International Journal of Computer Science in Sport*, 15(1), 22-36. doi: <https://doi.org/10.1515/ijcss-2016-0002>

Leistungsbeitrag des Autors

Der Promovend trug zur Entwicklung der Forschungsidee bei, speziell in Bezug auf die Erweiterung der Methode auf partielle Ableitung, Richtungsableitung und den Gradienten. Die Datenverarbeitung und Implementierung der notwendigen Algorithmen erfolgte ebenfalls durch den Promovenden. Die Datenauswertung und Aufbereitung der Ergebnisse wurde eigenständig durchgeführt und nur punktuell durch den Doktorvater unterstützt. Die Fertigstellung des Manuskripts erfolgte ausschließlich durch den Promovenden. Dies schließt die Literaturrecherche, die Darstellung der Methoden sowie Interpretation und Diskussion der Ergebnisse ein.

Zusammenfassung

Das Ziel dieser Studie war die Weiterentwicklung der Analyse-Möglichkeiten im Tischtennis durch Simulation anhand von Markov-Ketten. Dazu wurde die von Lames (1991) eingeführte Methodik der Anpassung und Auswertung von Übergangsmatrizen, die jeweils die taktischen Verhaltensweisen eines Spielers repräsentieren, um die Möglichkeit zur automatisierten Analyse durch numerische Ableitung erweitert. Während die ursprüngliche Methodik nach der Modellierung als Markov-System eine zweite Stufe der manuellen Modellierung beinhaltet, die notwendig ist um die Schwierigkeit einer taktischen Änderung einzubeziehen, wird dies durch die infinitesimal kleinen Änderungen der Übergangsmatrizen bei der numerischen Ableitung überflüssig.

Zur Validierung der Methode wurde ein Datensatz bestehend aus 259 Tischtennis-Spielen der Jahre 2008 bis 2012 verwendet, die alle bei offiziellen Turnieren internationaler Tischtennis-Verbände (International Table Tennis Federation (ITTF), European Table Tennis Union (ETTU), International Olympic Committee (IOC)) stattfanden. Erfasst wurden für jeden Ballwechsel dessen Schlagzahl, der Aufschlagspieler sowie der Spieler, der den Punktgewinn erzielen konnte. Darüber hinaus wurden Geschlecht der Spieler sowie das Ergebnis jedes Spiels pro Spieler ("Gewinner", "Verlierer") erfasst.

Auf Basis dieser Daten wurden schließlich beide Methoden verglichen und mathematische Konstrukte wie partielle Ableitung, Richtungsableitung und Gradient numerisch berechnet und sportwissenschaftlich interpretiert. Der Methodenvergleich erfolgte, indem für jedes Spiel die veränderten Gewinnwahrscheinlichkeiten der Originalmethode mit den partiellen Ablei-

tungen für dieselbe Übergangswahrscheinlichkeit korreliert wurden. Der resultierende durchschnittliche Korrelationskoeffizient von 0,457 ließ dabei nach Cohen (1988) auf einen moderaten Zusammenhang schließen. Um einen besseren Vergleich der Ergebnisse zu ermöglichen, wurden die Ergebnisse der Originalmethode zusätzlich in Steigungen umgerechnet und erneut mit den partiellen Ableitungen korreliert. Hierbei zeigte sich ein Korrelationskoeffizient von 0,995, was einen starken Zusammenhang der Methoden bestätigt und damit eine vergleichbare Wirkungsweise der numerischen Ableitung attestiert.

Für jedes der mathematischen Konstrukte konnten sportwissenschaftliche Anwendungen nachgewiesen werden. So ließ sich zum Beispiel der Gradient in Kombination mit einer Faktoranalyse dazu einsetzen, die einflussreichsten Übergänge der Markov-Matrizen zu identifizieren, während partielle Ableitung und Richtungsableitung Einsatzgebiete bei der Analyse spezieller taktischer Verhaltensweisen, wie z.B. riskantem Spiel, Einsatz fanden. Insbesondere lange Ballwechsel mit mehr als 6 Schlägen konnten hier durch hohe Ableitungswerte als relevante Spielsituationen identifiziert werden.

3.2. Studie 2 - Data Mining in Elite Beach Volleyball

Wenninger, S., Link, D., & Lames, M. (2019). Data Mining in Elite Beach Volleyball – Detecting Tactical Patterns Using Market Basket Analysis, *International Journal of Computer Science in Sport*, 18(2), 1-19. doi: <https://doi.org/10.2478/ijcss-2019-0010>

Leistungsbeitrag des Autors

Die Forschungsidee zu diesem Beitrag und insbesondere das Sequenzmodell von Beachvolleyball-Ballwechseln wurde vom Promovenden mit Unterstützung von Dr. Daniel Link im Rahmen eines vom Bundesinstitut für Sportwissenschaft (BISp) geförderten Projektes entwickelt. Datenverarbeitung, Auswertungen und Anfertigung des Artikels wurden vom Promovenden eigenständig durchgeführt. Der Doktorvater unterstützte bei der methodischen Entscheidungsfindung sowie der konzeptionellen Gestaltung des Manuskripts.

Zusammenfassung

Die Intention dieser Studie war die Suche nach sequenziellen taktischen Mustern innerhalb von Ballwechseln im Beachvolleyball. Dazu wurde eine Methode des Data Mining, das Association-Rule-Mining (ARM) eingesetzt. Im Einklang mit der Natur des Data Mining wurde ein explorativer Ansatz gewählt und die Methode sowohl zur theoretischen als auch zur praktischen Leistungsdiagnostik eingesetzt, ohne eine vordefinierte Hypothese zu verfolgen. Angewandt wurde der ERMiner-Algorithmus (Fournier-Viger et al., 2014) auf einem Datensatz von jeweils 413 Top-Level-Spielen der Männer sowie 552 Spielen der Frauen der Jahre 2013 bis 2016. Studien belegen dazu eine sehr gute Beobachterübereinstimmung zwischen den Analysten der deutschen Beachvolleyball-Nationalteams (Cohen's κ 0,93 bis 1,0). Analysiert wurden dabei nur die sogenannten Standard (Std.) Sideouts, diejenigen Ballwechsel in einem Beachvolleyball-Spiel, in denen das annehmende Team durch drei kontrolliert ausge-

fürte Berührungen einen strukturierten Spielaufbau betreiben kann. Dies wurde zum Einen dadurch begründet, dass sich taktische Verhaltensweisen durch die hohe Struktur gerade in diesen Ballwechseln manifestieren. Zum Anderen ist bekannt, dass diese Ballwechsel großen Einfluss auf den Ausgang eines Spiels haben (Zetou et al., 2007).

Die Datenauswertung setzte sich aus folgenden Schritten zusammen: Zuerst wurden die einzelnen Leistungsindikatoren innerhalb eines Std. Sideouts anhand eines selbst erstellten zeitlichen Modells geordnet und in eine Sequenz-Datenbank, wie durch den ERMiner-Algorithmus erwartet, konvertiert. Anschließend wurde der ARM-Algorithmus - je nach Fragestellung mit jeweils unterschiedlichen Parametern für Support und Konfidenz - auf die Daten angewandt. Um die große Zahl resultierender Regeln überschaubarer zu machen, wurde daraufhin eine hierarchische Clustering-Lösung entwickelt, die ähnliche Regeln anhand einer angepassten Form der Hamming-Distanz in Gruppen einordnet. Dabei wurde die optimale Anzahl an Clustern, ebenfalls spezifisch nach Fragestellung, durch eine Überprüfung der Silhouette-Metrik festgelegt. Für jeden Cluster wurde die repräsentative Regel als diejenige Regel mit dem geringsten durchschnittlichen Abstand zu allen anderen Regeln im Cluster definiert.

Die Auswertung der Ergebnisse zeigte, dass die Methode in der Lage ist, sowohl bereits bekannte taktische Muster zu reproduzieren als auch neue Muster aufzudecken. Dazu zählen zum Beispiel der Einsatz kurzer Aufschläge zur Behinderung des gegnerischen Anlaufs. Als großer Vorteil der Methode wurde dabei identifiziert, dass verschiedene Metriken für ungewöhnliche Kombinationen von Leistungsindikatoren bereitgestellt werden können, was den Analysten die Möglichkeit gibt, zusätzliches Wissen zu ihren erprobten Auswertungen zu generieren. Auf Grund der großen Anzahl generierter Regeln war es notwendig, Clustering als einen Post-Processing Schritt einzuführen. Vor diesem Hintergrund werden verschiedene Möglichkeiten diskutiert, die Einschränkungen der Methode entweder durch ein weiter verbessertes Post-Processing oder maßgeschneiderte und sport-spezifische Interessantheits-Metriken für Regeln zu umgehen.

3.3. Studie 3 - Performance of machine learning models in application to beach volleyball data

Wenninger, S., Link, D., & Lames, M. (2019). Performance of machine learning models in application to beach volleyball data., *International Journal of Computer Science in Sport*, 19(1), 24-36. doi: <https://doi.org/10.2478/ijcss-2020-0002>

Leistungsbeitrag des Autors

Die Forschungsidee zu dieser Studie wurde gemeinsam mit dem Doktorvater ausgearbeitet. Die Verarbeitung der Daten, Implementierung der Modelle sowie Datenauswertung und Verfassung des Artikels erfolgten durchgehend selbstständig durch den Promovenden. Unterstützung erfolgte durch den Doktorvater bei der Methodik zum Modellvergleich sowie punktuell bei der Manuskriptgestaltung.

Zusammenfassung

Die Zielsetzung der Studie bestand darin, den Einsatz von Methoden der künstlichen Intelligenz zur Auswertung der zum damaligen Zeitpunkt verfügbaren diskreten Positions- und Eventdaten im Beachvolleyball zu evaluieren. Als Motivation diente dabei vor allem die Einführung Sensor-basierter Metriken bei der medialen Ausstrahlung der Beachvolleyball-Weltmeisterschaften in den Jahren 2017 und 2019, die die Verfügbarkeit kontinuierlicher Positionsdaten in naher Zukunft in Aussicht stellt.

Im Detail wurde dafür die Genauigkeit unterschiedlicher Modelle aus dem Bereich des Machine Learning zur Vorhersage mehrerer Merkmale eines Ballwechsels verglichen. Eingesetzt wurden hierfür ein Convolutional Neural Network, ein Recurrent Neural Network, ein Multilayer-Perceptron sowie ein Gradient-Boosted-Tree. Um die Eigenschaften der verschiedenen Architekturen vergleichen zu können, wurde jedes Modell sowohl mit den reinen diskreten Positionsdaten als auch mit der Kombination von Positions- und Eventdaten als Input-Features trainiert. Zusätzlich dazu wurden die Modelle für jede dieser beiden Eingangsdatsätze jeweils zur Prädiktion dreier Zielvariablen (Angriffstechnik, Angriffsrichtung, Sideout-Erfolg) trainiert, wodurch sich insgesamt sechs Trainingsdurchläufe pro Modell ergaben.

Als Datensatz für das Training der Modelle dienten 1.356 Top-Level-Beachvolleyball-Spiele der Jahre 2013 bis 2018, die wiederum von professionellen Beachvolleyball-Analysten der deutschen Nationalteams erfasst wurden. Die Analyse beschränkte sich auch in dieser Studie wieder auf die Std. Sideouts, begründet auf denselben Argumenten wie bereits in der Studie zuvor. Für den Modellvergleich wurden die Modelle für jede der sechs Konfigurationen aus Trainingsdaten und Zielvariablen jeweils 50 mal für 10 Epochen trainiert und pro Iteration das trainierte Modell mit der besten Vorhersagegenauigkeit gespeichert. Die dadurch generierten 1200 Genauigkeitswerte dienten als Grundlage für eine drei-faktorielle Analysis of Variance (ANOVA) mit der Vorhersagegenauigkeit als abhängiger Variable und den Zielvariablen sowie den Input-Features als unabhängige Faktoren.

Alle Modelle konnten Vorhersagegenauigkeiten jenseits von 50 Prozent erreichen.

Die Ergebnisse der Varianzanalyse bestätigten signifikante Unterschiede in der Genauigkeit der Modelle, wobei auch die Interaktionseffekte zwischen den Faktoren signifikante Unterschiede vorweisen konnten. Durch die Überprüfung der Effektstärke ließ sich nachweisen, dass neben den Input-Features ($\eta^2 = .115$) insbesondere die Zielvariable ($\eta^2 = .735$) und die Interaktion dieser Faktoren ($\eta^2 = .146$) zur Aufklärung der Varianz beitragen, wohingegen die Varianzaufklärung des Faktors Modell nahezu vernachlässigbar klein war ($\eta^2 = .001$). Dies ließ darauf schließen, dass die signifikanten Unterschiede mit großer Wahrscheinlichkeit durch die geringe Varianz der Vorhersagegenauigkeit innerhalb der Gruppen hervorgerufen wurden, was ein Nebeneffekt der Methodik zum Modellvergleich war. Gleichzeitig gaben die kleinen absoluten Differenzen zwischen den Modellen Anlass zu der Annahme, dass zumindest mit den zur Zeit verfügbaren Daten eine faktische Obergrenze in der Genauigkeit der Modelle besteht. Begründet wurde dies durch den komplexen Interaktionsprozess in Sportarten generell und die Vielschichtigkeit taktischer Entscheidungen im Beachvolleyball im Speziellen, die insbesondere die Vorhersage des Erfolgs innerhalb eines Ballwechsels erschwert. Basierend auf diesen Erkenntnissen wurden schließlich die praktische Relevanz der

Methodik sowie weitere Möglichkeiten zur Anwendung in angrenzenden Gebieten der Leistungsdiagnostik diskutiert. Besonderer Fokus wurde dabei auf die in Zukunft möglicherweise verfügbaren kontinuierlichen Positionsdaten gelegt.

4. Diskussion

Die Publikationen dieser Arbeit befassen sich mit der Anwendung informatischer Methoden zur Leistungsdiagnostik in den Spilsportarten Tischtennis und Beach-Volleyball. Dazu wurde eine Reihe an Methoden verwendet, die im Kapitel 2 genauer erläutert werden. Da in den Publikationen auf Grund der beschränkten Länge der Artikel nicht alle Aspekte der entsprechenden Arbeiten diskutiert werden konnten, sollen im Folgenden diese Punkte detaillierter dargestellt werden. Zudem wird, falls möglich, ein Vergleich mit aktuellen Entwicklungen in der Sportinformatik gezogen.

Performance analysis in table tennis - Stochastic simulation by numerical derivation

In dieser Studie wurde die numerische Ableitung als Methode zur automatisierten Analyse eines Markov-Modells von Tischtennis-Spielen eingeführt und bezüglich ihrer leistungsdiagnostischen Aussagekraft untersucht. Der Fokus dieser leistungsdiagnostischen Untersuchung lag dabei auf Fragestellungen der theoretischen Leistungsdiagnostik. Natürlich lässt sich diese Methode auch in der praktischen Leistungsdiagnostik anwenden, indem die Übergangsmatrizen einzelner Spieler, zum Beispiel während eines Turniers, zum Gegenstand der Analysen gemacht werden. Somit könnten die Änderungen der Gewinnwahrscheinlichkeiten nach Anpassungen am taktischen Verhalten direkt zwischen Spielern verglichen werden. Durch die Automatisierung der Ableitung wäre es weiterhin möglich, auf der Übergangsmatrix eines Spielers eine Suche nach der effektivsten taktischen Änderung durchzuführen, ebenfalls mit informatischen Suchalgorithmen und dadurch ebenfalls automatisiert. Damit könnte eine Liste potenzieller taktischer „Gegenmaßnahmen“ gegen einen bestimmten Spieler gefunden werden, die einen entscheidenden Vorteil im Wettbewerb darstellen kann.

Ein anderer Punkt, der in der Diskussion dieser Publikation weniger ausführlich behandelt wurde, ist der Vergleich mit ähnlichen stochastischen Modellen. So existieren bereits seit den 70er-Jahren Modelle, die Tischtennis als einen alternierenden binomialen Prozess ansehen (Dominicy et al., 2013; Schulman & Hamdan, 1977). Indem jedem Spieler eine Gewinnwahrscheinlichkeit für einen Ballwechsel zugewiesen wird, kann über das Binomialmodell die Wahrscheinlichkeit des Matchgewinns als Funktion dieser beiden Wahrscheinlichkeiten determiniert werden. Diese Gewinnwahrscheinlichkeiten der Spieler können dazu beispielsweise empirisch aus einer Datenbank von Ballwechseln bestimmt werden. Im Vergleich zu dem in dieser Arbeit verwendeten Markov-Modell bietet dieser Ansatz allerdings keine Vorteile. Da das Modell der Ballwechsel nur den Ausgang desselben betrachtet und keinerlei Zustände innerhalb eines Ballwechsels beinhaltet, können keine Analysen zu taktischen Verhaltensweisen innerhalb eines Ballwechsels ausgeführt werden. Nun lässt sich argumentieren, dass auch das in unserer Studie verwendete Modell der Ballwechsel mit der reinen Modellierung nach Ballwechsellängen nicht merklich komplexer ist. Allerdings bieten Markov-Ketten die Möglichkeit, beliebig komplexe Modelle abzubilden und somit zum Beispiel auch technische Merkmale (Vor-/Rückhand, Schnitt, Platzierung, ...) mit in die Analyse aufzunehmen, voraus-

gesetzt diese Daten sind vorhanden. Die in dieser Arbeit entwickelte Methode der numerischen Ableitung lässt sich dazu auf jedes beliebige Markov-Modell anwenden.

Eine neuere Methode der stochastischen Analyse im Tischtennis wird von Ley et al. (2018) vorgestellt. In ihrer Arbeit verwenden sie für jede Paarung von Spielern sogenannte „Mutual-Point-Winning-Probabilities (MPW)“. MPW bestehen aus zwei Wahrscheinlichkeitswerten, die für jeden Spieler die Chance ausdrücken, einen Ballwechsel bei eigenem oder gegnerischem Aufschlag zu gewinnen. Da diese Werte jeweils für eine Spielerpaarung gebildet werden, stellen sie eine Kenngröße für die relative Spielstärke zweier Gegner dar und können nach jedem Match dieser Spieler aktualisiert werden. Ein besonderes Augenmerk legen die Autoren auf Gewichtungsfaktoren für bestimmte Spielsituationen (Satzende, Punktdifferenz, gespielte Sätze, etc.), die als Faktoren für die Bedeutung eines Ballwechsels in die Berechnung der MPW miteinbezogen werden. Dieses Modell erweitert also das von Dominicy et al. (2013) verwendete System um ein Maß der Wichtigkeit von Ballwechseln, um aussagekräftigere Ergebnisse für die Gewinnwahrscheinlichkeiten zu erhalten. Die Kritik an der ursprünglichen Methode, dass keine Aussagen über beliebige taktische Handlungen getätigt werden können, bleibt jedoch erhalten. Das zeigt gleichzeitig eine große Stärke der Simulation durch Markov-Ketten: Solange die gesuchten taktischen Muster in den Übergangsmatrizen abgebildet sind, lassen sie sich mit Hilfe der Markov-Ketten analysieren und anhand der Erweiterung durch numerische Ableitungen automatisiert auswerten. Denn auch die in der von Ley et al. (2018) verwendeten Gewichtungen der Ballwechsel stellen eine zusätzliche Modellbildung dar, die somit unweigerlich eine Quelle für Fehler in der Analyse einführen. Zur Demonstration des Nutzens ihrer Methode vergleichen die Autoren die Ergebnisse ihrer statistischen Vorhersagen mit den realen Ergebnissen selbst organisierter Turniere. Auch für die numerische Simulation wäre ein solcher Vergleich ein interessantes Mittel zur Veranschaulichung und Validierung der Ergebnisse gewesen. Den praktischen Nutzen der numerischen Simulation mit Markov-Ketten zeigt ihre Rezeption in Wu et al. (2018) oder Wang et al. (2019), die angepasste Versionen der hier vorgestellten Markov-Ketten in Software-Tools übernommen haben, die von Weltklasse-Athleten und -Trainern verwendet werden.

Data Mining in elite beach volleyball - Detecting tactical patterns using Market Basket Analysis

Eine Problematik der Studie war die hohe Anzahl an Regeln, die der Algorithmus mit den verwendeten Parametern lieferte. Die Wahl dieser Parameter begründet sich allerdings aus Anforderungen der Leistungsdiagnostik, die im Gegensatz zu den gängigen Einsatzgebieten des ARM eben nicht nur häufig vorkommende Muster als interessant ansieht. Um die Anwender der Methode nicht mit einer unübersehbaren Menge an Regeln zu überwältigen, wurde ein hierarchisches Clustering angewandt, das die Regelmenge strukturiert. Wie in der Studie bereits erwähnt, könnte ein eigens entwickeltes Interessantheitsmaß oder auch die Verwendung einer der vielen anderen Metriken für die Interessantheit von Regeln diese Problematik mindern (Hasler, 2015; Tan et al., 2004). Diese Metriken zu vergleichen und eine passende für den Einsatz im Beach-Volleyball auszuwählen beziehungsweise eine eigene Metrik zu entwickeln, hätte den Umfang dieser Studie jedoch überschritten und bleibt als Aufgabe für

zukünftige Forschung.

Die Methoden unterliegen weiterhin einer ständigen Entwicklung. Besonders Ansätze, die nicht auf der ursprünglichen Idee des Apriori-Algorithmus (Agrawal et al., 1993) basieren, finden in der Forschung neuerdings Beachtung. Diese Ansätze verzichten auf die Berechnung von *frequent itemsets*, um Leseoperationen in der Datenbank zu sparen und somit die Performance der Algorithmen insbesondere bei den immer größer werdenden Datensätzen zu erhöhen. Beispiele für diese Ansätze beinhalten den ebenfalls sehr bekannten FP-Growth Algorithmus (Han et al., 2004) und dessen Weiterentwicklungen (z.B. Narvekar und Syed (2015)) sowie heuristische Ansätze (siehe Ghafari und Tjortjis (2019)). Es ist weiterhin anzumerken, dass sich aktuelle Publikationen mit Methoden des Data-Mining im Sport auffällig häufig in Sportarten wie Cricket (Sahi, 2020; Singh et al., 2015), Baseball (Tolbert & Trafalis, 2016; Woodham et al., 2019) oder Tischtennis (Huiqun, 2006; Ma, 2020; Sun et al., 2010) finden. In diesen Sportarten sind die verfügbaren Daten immer noch mehrheitlich in diskreter Form vorhanden, was die Anwendung von Data-Mining-Algorithmen einfacher macht. In Sportarten, bei denen sich die gesammelten Daten immer mehr in Richtung zeitkontinuierlicher Positionsdaten verschieben, scheinen jedoch immer öfter Methoden des Machine Learning zur leistungsdiagnostischen Analyse herangezogen zu werden (Link et al., 2016; Theodoropoulos et al., 2020; Tian et al., 2020). Und selbst in den zuvor erwähnten Sportarten, die Data-Mining-Methoden begünstigen, findet man mittlerweile immer mehr Anwendungen des Machine-Learning (vgl. z.B. Balasundaram et al. (2020)).

Performance of machine learning methods in forecasting tactical behavior in beach volleyball

In dieser Studie wurden verschiedene Algorithmen des Machine-Learning dazu verwendet, nach taktischen Mustern in Positions- und Ereignisdaten zu suchen. Analog zur vorausgegangenen Studie wurden die jeweiligen Algorithmen mit bestimmten Parametern durchgeführt, die in der Studie dokumentiert wurden. Dabei wurde nicht genauer beschrieben, wie diese sogenannten Hyperparameter gefunden wurden, sondern nur dargestellt, dass diese aus einer experimentellen Phase hervorgegangen sind beziehungsweise auf die jeweilige Fragestellung bezogen ausgewählt wurden (z.B. die Zeitschritte bei den rekurrenten Netzwerken). Deswegen soll hier für die betreffenden Parameter noch einmal genauer auf diesen Prozess eingegangen werden.

Einer der wichtigsten Hyperparameter ist die sogenannte *learning rate*, die im Grunde den Grad der Anpassung eines neuronalen Netzes nach jedem Trainingsdurchlauf beschreibt. Dieser Parameter wurde mit Hilfe der Funktion `lr_find()` des verwendeten `fast.AI`-Frameworks gefunden, die den von Smith (2017) vorgestellten Algorithmus verwendet. Das definierte Modell wird dazu zyklisch mit exponentiell wachsenden *learning rates* trainiert und die resultierenden Verlustwerte¹ für die zugehörige *learning rate* in einem Graphen aufgetragen (siehe Abbildung 15).

Folgende Werte können dabei als Faustregel für die Wahl einer guten *learning rate* verwendet werden:

¹ Differenz zwischen Modell und wahren Trainingswert

- Ein Zehntel der minimalen *learning rate* bevor der Algorithmus beendet wurde.
- Die *learning rate* am Punkt der steilsten Steigung des Graphen.

In dieser Studie wurde jeweils der Punkt der steilsten Steigung als optimale *learning rate* für die Algorithmen identifiziert.

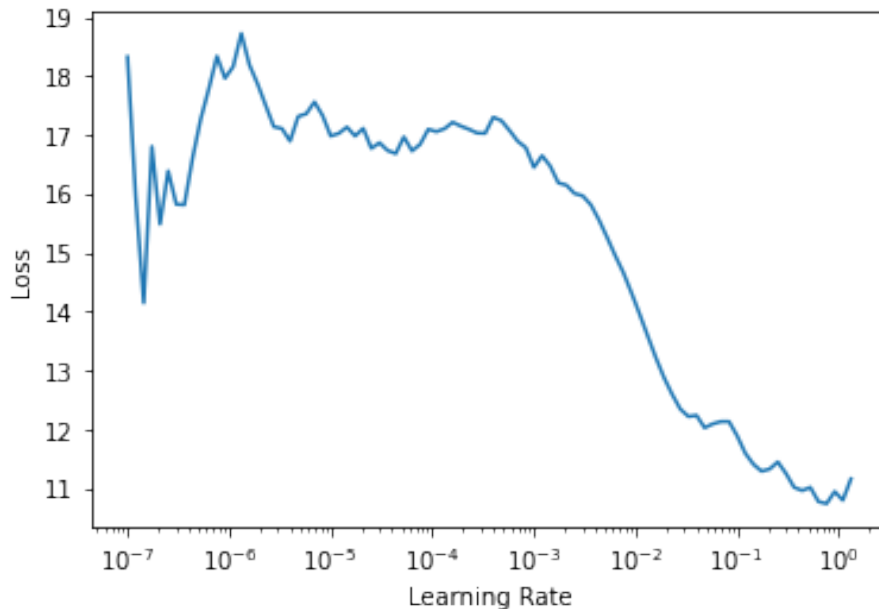


Abbildung 15 Beispiel für einen durch `lr_find()` generierten Graphen. Grafik entnommen aus fast.ai - Dokumentation, 2020

Auch die Architekturen der Modelle, also die Anzahl der Schichten und Neuronen pro Schicht im Fall der neuronalen Netze beziehungsweise die Anzahl der Bäume und die maximale Tiefe der generierten Bäume bei XGBOOST, entspringen einer experimentellen Phase vor der eigentlichen Studie, die die verwendeten Konfigurationen ergeben hat. Hier zeigten sich in den Experimenten mit den Architekturen keine gravierenden Unterschiede in der Genauigkeit der Algorithmen, weswegen auch unter dem Aspekt der verfügbaren Datenmenge eine möglichst einfache Struktur mit zwei linearen Ebenen gewählt wurde. Zudem wurde z.B. durch die Hinzunahme mehrerer 1D-Convolutions bei den CNN versucht, möglichst viele praktische Anwendungsfälle in den Modellen abzubilden. Hier wäre ein methodischer Vergleich der Auswirkungen dieser Architekturen auf die Klassifikationsgenauigkeiten sicherlich hilfreich gewesen, da nicht jede Kombination von Konfigurationen in der Testphase berücksichtigt werden konnte. Allerdings hätte dieser Architekturvergleich den Umfang einer eigenen Studie und mit der Zielstellung der hier vorgestellten Studie wenig gemein gehabt.

Natürlich existiert neben den hier verwendeten Architekturen und Methoden noch eine Vielzahl anderer Methoden des Machine-Learning, die nicht in diese Studie miteinbezogen wurden. Beispiele hierfür beinhalten beispielsweise U-Nets (Ronneberger et al., 2015), Generative adversarial networks (Goodfellow et al., 2014), Auto-Encoder (Goodfellow et al., 2016), Kohonen-Feature-Maps (Kohonen, 1982) oder sogar stochastische neuronale Netze wie zum Beispiel Boltzmann-Maschinen (Hinton et al., 1984). Jede dieser Methoden bietet Stärken und Schwächen (vgl. z.B. Serrien et al. (2017)), die im Kontext der Sportwissenschaft abge-

wägt und auf den jeweiligen Anwendungsfall geprüft werden müssen.

Betrachtet man die Studie aus leistungsdiagnostischer Sicht, ist anzumerken, dass die Auswahl des Test-Sets² aus den Daten angepasst werden könnte. Nachdem sich der Datensatz über mehrere Spielzeiten erstreckt, könnte man die Algorithmen beispielsweise mit den Daten aus den ersten 4 Jahren trainieren, während die letzte Saison als Test-Datensatz vorgehalten wird. Somit wäre ein direkter Vergleich der Klassifikationsergebnisse der Algorithmen mit realen Spieldaten der letzten Saison möglich gewesen, was im Vergleich zu der rein statistischen Analyse der Genauigkeit zusätzliche leistungsdiagnostische Erkenntnisse mit sich bringen könnte. Auch diese Methode hätte Einschränkungen (z.B. Wechsel der Spieler innerhalb der Teams), die durch die Auswahl der Daten aus dem Gesamtdatensatz wiederum umgangen werden. In Retrospektive wäre diese Art der Evaluation möglicherweise dennoch hilfreich gewesen, um den leistungsdiagnostischen Anteil der Arbeit herauszustellen und greifbare Beispiele für den Nutzen der Methoden zu generieren.

In der Studie wird weiterhin eine Problematik von Machine-Learning-Anwendungen in der praktischen Leistungsdiagnostik angesprochen: Häufig ist hier die verfügbare Datenmenge für entsprechende Analysen (z.B. Gegneranalysen) nicht groß genug, da sie nur in einem begrenzten Kontext (Turnier, Saison) erfasst werden kann. Mit Hilfe von Methoden des sogenannten Transfer-Learnings, das besonders in der Bilderkennung bereits sehr erfolgreich eingesetzt wird (Tan et al., 2018), würden sich die benötigten Daten für Analysen einzelner Spieler oder Teams deutlich reduzieren. Die Grundidee ist dabei, ein bereits zu einem anderen Zweck trainiertes Modell in einem neuen Kontext wiederzuverwenden. Nimmt man beispielsweise ein auf den Daten aller Spieler trainiertes neuronales Netz zur Klassifikation erfolgreicher Ballwechsel an, das eine hinreichende Klassifikationsgenauigkeit aufweist, könnte dieses als Grundlage für die Analyse einzelner Teams oder Spieler dienen. Dazu werden die untersten Schichten des Netzes „eingefroren“, wodurch dessen bereits gelernte Fähigkeit, grundlegende Muster in den Daten zu erkennen, bestehen bleibt. Nur die oberen Schichten könnten dann mit den Daten eines einzigen Teams/Spielers neu trainiert werden, um das Modell auf dieses/diesen zuzuschneiden und dadurch genauere Vorhersagen zu ermöglichen.

² randomisierte Auswahl aus allen Ballwechselln, 80% Trainingsdaten, 20% Testdaten

5. Fazit

Ein Fazit zu dieser Arbeit muss unter Berücksichtigung mehrerer Perspektiven gezogen werden. Betrachtet man diese Arbeit aus rein informatischer Sicht, ist zuerst festzuhalten, dass Sportdaten an sich ein hervorragendes Anwendungsfeld für informatische Analysen darstellen. Sie enthalten komplexe Muster in einem ebenso komplexen, dynamischen Kontext. Diese Komplexität wird aber gleichzeitig durch die Rahmenbedingungen im Sport (Regelwerk, Spielziel, räumliche Einschränkungen) so weit reduziert, dass sie informatisch beherrschbar werden (Link & Lames, 2009). Es konnte gezeigt werden, dass mit diskreten Event-Daten aussagekräftige Muster in verschiedenen Sportarten gefunden werden können. Insbesondere bei der Simulation und dem Data-Mining war die Anwendung der Methoden auf diesen Daten ohne erhöhten Aufwand zum Beispiel bei der Datenvorverarbeitung möglich. Machine-Learning erfordert für optimale Ergebnisse generell einen erhöhten Aufwand an Feature-Engineering. Dieser erhöhte Aufwand spiegelt sich mit den verfügbaren Daten allerdings noch nicht in den Ergebnissen wider. Hier könnte es lohnenswert sein, die Datenbasis in Zukunft auszuweiten und bevorzugt Positionsdaten angereichert mit diskreten Event-Daten zu verwenden. Aus sportwissenschaftlicher Sicht müssen die Ergebnisse aus zwei Perspektiven bewertet werden: Für die TLD, die Aussagen über die Leistungsstruktur von Sportarten trifft, hat sich die Methodik der Studien, Analysen auf dem Gesamtdatensatz aller Spieler beziehungsweise Teams auszuführen, bewährt. Die aus der theoretischen Leistungsdiagnostik gewonnenen Erkenntnisse und Modelle bilden die Grundlage für praktische Anwendungen und sind unerlässlich. Hierbei ist zu beachten, dass die direkte und unreflektierte Ableitung von praktischen Hinweisen aus der theoretischen Leistungsdiagnostik im Allgemeinen als problematisch zu erachten ist (Hohmann et al., 2020). In der praktische Leistungsdiagnostik, die Aussagen über einzelne Athleten oder Mannschaften treffen möchte, ist die direkte Anwendung solcher Methoden zudem schwieriger. Das liegt zum Einen an der bereits erwähnten geringeren Verfügbarkeit von Daten einzelner Akteure. Zum anderen müssen Anwendungen in der PLD direkt mit dem Expertenwissen von Trainern, Scouts oder Sportlern konkurrieren. Das setzt zum Einen hohe Ansprüche an die Ergebnisse, zum Anderen an die Methoden selbst. Die generierten Informationen sollten schließlich intuitiv verständlich für Trainer und Athleten sein (Glazier, 2010). Damit ist festzustellen, dass die Ergebnisse besser werden, je mehr Expertenwissen bereits in die Modellierung und das Feature Engineering einfließt. Es ist also immer ein tieferes Verständnis der Sportart erforderlich, um sinnvolle Analysen zu gestalten. Dies betont den interdisziplinären Charakter der Sportinformatik.

Somit bleibt die Leistungsdiagnostik in Spielsportarten wegen ihrer komplexen und dynamischen Struktur ein herausforderndes Forschungsgebiet. Die Aufgabe, aussagekräftige und vor allem menschlichen Experten bisher unbekannte Muster und Zusammenhänge in den Daten zu finden, bleibt nach wie vor schwierig, selbst mit den mittlerweile vorhandenen Datenmengen. Die Sportinformatik kann dabei zu einem besseren Verständnis der entstehenden Muster und Verhaltensweisen in diesen Sportarten beitragen. Dies wurde in dieser Arbeit am Beispiel des Tischtennis und Beachvolleyballs gezeigt.

Ausblick

Blickt man auf die Zukunft der Sportinformatik, ist klar, dass auch im Sport das stetige Wachstum der verfügbaren Daten weiter gehen wird. In den in dieser Arbeit behandelten Sportarten ist davon auszugehen, dass es nur eine Frage der Zeit ist, bis flächendeckend mit entsprechenden Sensoren erfasste kontinuierliche Positionsdaten zur Verfügung stehen. Im Tischtennis kann der Ballflug schon mit einer Genauigkeit verfolgt und vorhergesagt werden, dass Roboter damit trainiert werden können (Tebbe et al., 2018). Im (Beach)Volleyball hat die Erfassung dieser Daten in den letzten Jahren begonnen (comdirect, 2017). Gerade dabei bleibt die Aufgabe der Sportinformatik die Unterstützung der Experten im leistungsdiagnostischen Prozess, mit besonderem Augenmerk auf die Vorfilterung und vorbereitende Analyse der Daten. Das Ziel für die Zukunft der Sportinformatik muss sein, die Effizienz in diesem Bereich zu verbessern, und möglicherweise mit dem Einsatz neuer Methoden Anteile am Prozess der qualitativen Leistungsdiagnostik zu gewinnen. Das funktioniert nur dann, wenn erhöhter Wert auf den interdisziplinären Charakter der Sportinformatik gelegt wird. Dazu müssen die verwendeten Methoden geeignet sein, den nicht unbeachtlichen Graben zwischen der Sportpraxis und der Informatik zu überwinden. Betrachtet man zum Beispiel die sogenannten „Black-Box“-Modelle des Deep Learning, besteht ein besonderer Bedarf, den Entscheidungsprozess der Modelle erklären zu können. Zum Einen muss der Sportpraxis Vertrauen in die Funktionsweise der Modelle gegeben werden können, wenn Entscheidungen auf Basis ihrer Vorhersagen getroffen werden sollen. Zum Anderen müssen solche Entscheidungen gegenüber Dritten (Spieler, Trainer, Funktionären) begründet werden können. Das funktioniert nur, wenn Aussagen zur Nachvollziehbarkeit der Vorhersagen gegeben werden können.

Hier setzen beispielsweise neue Entwicklungen im Machine-Learning an, die mit Techniken der Explainable Artificial Intelligence (XAI) die Entscheidungsprozesse der Black-Box-Modelle erklärbar machen wollen (Geada et al., 2021; Ribeiro et al., 2016). Durch die Anreicherung des Entscheidungsprozesses mit Metadaten werden Nachvollziehbarkeit und Erklärbarkeit der Ergebnisse sichergestellt, die sich zudem durch das Modell verfolgen lassen, um einen weitreichenden Überblick über die getroffenen Entscheidungen und Vorhersagen zu erhalten. Dieses Beispiel veranschaulicht eine generelle Problematik, die momentan in der Sportinformatik besteht. Hier trifft ein Anwendungsfeld, in dem erst seit relativ kurzer Zeit eine große Menge an Daten zur Verfügung steht, auf ein breites Angebot immer ausdifferenzierterer informatischer Methoden. Kurz gesagt herrschen Idealbedingungen für die Erprobung dieser Methoden, die auch große Fortschritte in bisher nur schwer zu bearbeitenden Gebieten wie zum Beispiel der automatischen Erkennung von Events in Positionsdaten erzielen (Khaustov & Mozgovoy, 2020). Dabei besteht die Gefahr, dass die Anforderungen der Sportwissenschaft im Zuge immer komplizierterer Anwendungen auf der Strecke bleiben. Analog zu den informatischen Bemühungen muss also gleichzeitig der Aufwand erhöht werden, der betrieben wird, um das Verständnis der Sportpraxis sicherzustellen. Gerade hier wird zukünftig die große Aufgabe der Sportinformatik sein, weiter eine Brücke zwischen der Sportwissenschaft und der Informatik zu bilden.

Literatur

- Activation functions and its Types in Artificial Neural network.* (2018). Abgerufen am: 21. September 2021 unter <https://abhigoku10.medium.com/activation-functions-and-its-types-in-artificial-neural-network-14511f3080a8>
- Adler, P. (1981). *Momentum: A Theory of Social Action*. Sage Pubns.
- Agrawal, R., Imieliński, T. & Swami, A. (1993). Mining association rules between sets of items in large databases. In P. Bunemann & I. Shshil (Hrsg.), *Proceedings of the 1993 ACM SIGMOD international conference on Management of data* (S. 207–216).
- Agrawal, R., Srikant, R. et al. (1994). Fast algorithms for mining association rules. In J. Bocca, M. Jarke & C. Zaniolo (Hrsg.), *VLDB '94: Proceedings of the 20th International Conference on Very Large Data Bases* (S. 487–499). Morgan Kaufmann Publishers Inc.
- Amazon. (2020). *AWS powers NFL next Gen stats*. Abgerufen am: 30. Mai 2020 unter <https://aws.amazon.com/nfl/>
- Araujo, D., Davids, K. & Hristovski, R. (2006). The ecological dynamics of decision making in sport. *Psychology of sport and exercise*, 7(6), 653–676.
- Araújo, D., Davids, K. & Passos, P. (2013). The intending–perceiving–acting cycle in sports performance. *Routledge handbook of sports performance analysis* (S. 50–59). Routledge.
- Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E., Leite, R. E., Filho, W. J., Lent, R. & Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology*, 513(5), 532–541.
- Baca, A. (2006). Computer science in sport: an overview of history, present fields and future applications (part I). *International Journal of Computer Science in Sport*, 4(1), 25–35.
- Baca, A. (2014). *Computer science in sport: research and practice*. Routledge.
- Balasundaram, A., Ashokkumar, S., Jayashree, D. & Magesh Kumar, S. (2020). Data mining based Classification of Players in Game of Cricket. *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, 271–275.
- Bastin, J., Craig, C. & Montagne, G. (2006). Prospective strategies underlie the control of interceptive actions. *Human movement science*, 25(6), 718–732.
- Biundo, S., Claus, V. & Mayr, H. C. (2006). *Was ist Informatik? Positionspapier der Gesellschaft für Informatik*. Abgerufen am: 7. September 2021 unter <https://gi.de/fileadmin/GI/Hauptseite/Themen/was-ist-informatik-lang.pdf>
- Blank, P., Hoßbach, J., Schuldhuis, D. & Eskofier, B. M. (2015). Sensor-Based Stroke Detection and Stroke Type Classification in Table Tennis. In K. Mase, M. Langheinrich & D. Gatica-Perez (Hrsg.), *Proceedings of the 2015 ACM International Symposium on Wearable Computers* (S. 93–100). Association for Computing Machinery.

- Bourbousson, J., Sève, C. & McGarry, T. (2010). Space–time coordination dynamics in basketball: Part 1. Intra- and inter-couplings among player dyads. *Journal of sports sciences*, 28(3), 339–347.
- Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- Buchheit, M. & Simpson, B. M. (2017). Player-tracking technology: half-full or half-empty glass? *International journal of sports physiology and performance*, 12(s2), 35–41.
- Bungartz, H.-J., Zimmer, S., Buchholz, M. & Pflüger, D. (2009). *Modellbildung und Simulation: eine anwendungsorientierte Einführung*. Springer-Verlag.
- Bunker, R. P. & Thabtah, F. (2019). A machine learning framework for sport result prediction. *Applied Computing and Informatics*, 15(1), 27–33.
- Cai, W., Yu, D., Wu, Z., Du, X. & Zhou, T. (2019). A hybrid ensemble learning framework for basketball outcomes prediction. *Physica A: Statistical Mechanics and its Applications*, 528, 121461.
- Carling, C., Wright, C., Nelson, L. J. & Bradley, P. S. (2014). Comment on ‘Performance analysis in football: A critical review and implications for future research’. *Journal of sports sciences*, 32(1), 2–7.
- Castellano, J., Casamichana, D. & Lago, C. (2012). The use of match statistics that discriminate between successful and unsuccessful soccer teams. *Journal of human kinetics*, 31(1), 137–147.
- Caudill, M. (1987). Neural Networks Primer, Part I. *AI Expert*, 2(12), 46–52.
- Chinchilla-Mira, J. J., Pérez-Turpin, J. A., Martínez-Carbonell, J. A. & Jové-Tossi, M. A. (2012). Offensive zones in beach volleyball: differences by gender. *Journal of Human Sport and Exercise*, 7(3), 727–732.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Claus, V. (1993). *Duden-Informatik: ein Sachlexikon für Studium und Praxis* (2. Auflage). Dudenverlag.
- Claus, V. (2006). *Duden-Informatik A-Z: Fachlexikon für Studium, Ausbildung und Beruf* (4. Auflage). Dudenverlag.
- Claus, V., Schanz, P. D. M. & Avrutin, V. (1975). *Einführung in die Informatik*. Teubner Stuttgart.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates.
- comdirect. (2017). *comdirect ist offizieller Innovationspartner des Deutschen Volleyball-Verbandes (DVV)*. Abgerufen am: 21. September 2021 unter https://www.comdirect.de/cms/ueberuns/de/presse/innovationspartner_dvv.html
- Decroos, T., Bransen, L., Van Haaren, J. & Davis, J. (2019). Actions Speak Louder than Goals: Valuing Player Actions in Soccer. In A. Teredesai & V. Kumar (Hrsg.), *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (S. 1851–1861). Association for Computing Machinery.

- Dick, U. & Brefeld, U. (2019). Learning to Rate Player Positioning in Soccer. *Big Data*, 7(1), 71–82.
- Dominicy, Y., Ley, C., Swan, Y. et al. (2013). A stochastic analysis of table tennis. *Brazilian Journal of Probability and Statistics*, 27(4), 467–486.
- Draschkowitz, L., Draschkowitz, C. & Hlavacs, H. (2014). Predicting shot success for table tennis using video analysis and machine learning. In D. Reidsma, I. Choi & R. Bargar (Hrsg.), *International Conference on Intelligent Technologies for Interactive Entertainment* (S. 12–21).
- Duan, Y., Fu, X., Luo, B., Wang, Z., Shi, J. & Du, X. (2015). Detective: Automatically identify and analyze malware processes in forensic scenarios via DLLs. *2015 IEEE International Conference on Communications (ICC)*, 5691–5696.
- Duarte, R., Araújo, D., Correia, V. & Davids, K. (2012). Sports teams as superorganisms. *Sports medicine*, 42(8), 633–642.
- Duarte, R., Araújo, D., Correia, V., Davids, K., Marques, P. & Richardson, M. J. (2013). Competing together: Assessing the dynamics of team–team and player–team synchrony in professional association football. *Human movement science*, 32(4), 555–566.
- fast.ai - Dokumentation. (2020). *fast.ai Documentation*. Abgerufen am: 8. November 2020 unter <https://docs.fast.ai/callback.schedule>
- Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17(3), 37.
- Fernando, B., Fromont, E. & Tuytelaars, T. (2012). Effective use of frequent itemset mining for image classification. In A. Fitzgibbon, S. Lazebnik, P. Perona, Z. Sato & C. Schmid (Hrsg.), *12th European Conference on Computer Vision* (S. 214–227).
- Fielding, A. & O’Muircheartaigh, C. (1977). Binary segmentation in survey analysis with particular reference to AID. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 26(1), 17–28.
- Folgado, H., Duarte, R., Fernandes, O. & Sampaio, J. (2014). Competing with lower level opponents decreases intra-team movement synchronization and time-motion demands during pre-season soccer matches. *PloS one*, 9(5), 1–9.
- Fournier-Viger, P., Gueniche, T., Zida, S. & Tseng, V. S. (2014). ERMiner: Sequential Rule Mining Using Equivalence Classes. In H. Blockeel, M. van Leeuwen & V. Vinciotti (Hrsg.), *Advances in Intelligent Data Analysis XIII* (S. 108–119). Springer International Publishing.
- Fournier-Viger, P., Lin, J. C.-W., Vo, B., Chi, T. T., Zhang, J. & Le, H. B. (2017). A survey of itemset mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(4), e1207.
- Freund, Y., Schapire, R. & Abe, N. (1999). A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 14(771-780), 1612.
- Fuchs, M., Liu, R., Malagoli Lanzoni, I., Munivrana, G., Straub, G., Tamaki, S., Yoshida, K., Zhang, H. & Lames, M. (2018). Table tennis match analysis: a review. *Journal of sports sciences*, 36(23), 2653–2662.

- Geadia, R., Teofili, T., Vieira, R., Whitworth, R. & Zonca, D. (2021). TrustyAI Explainability Toolkit. *arXiv preprint arXiv:2104.12717*.
- Ghafari, S. M. & Tjortjijis, C. (2019). A survey on association rules mining using heuristics. *WIRES Data Mining and Knowledge Discovery*, 9(4), e1307.
- Giatsis, G., Lopez Martinez, A. B. & Gea García, G. M. (2015). The efficacy of the attack and block in game phases on male FIVB and CEV beach volleyball.
- Giatsis, G., Panoutsakopoulos, V. & Kollias, I. A. (2018). Biomechanical differences of arm swing countermovement jumps on sand and rigid surface performed by elite beach volleyball players. *Journal of sports sciences*, 36(9), 997–1008.
- Giatsis, G., Schrapf, N., Koraimann, T. & Tilp, M. (2019). Analysis of the arm swing technique during the spike attack in elite beach volleyball. *International Journal of Performance Analysis in Sport*, 19(3), 370–380.
- Giatsis, G. & Tzetzis, G. (2003). Comparison of performance for winning and losing beach volleyball teams on different court dimensions. *International Journal of Performance Analysis in Sport*, 3(1), 65–74.
- Giatsis, G. & Zahariadis, P. (2008). Statistical analysis of men's FIVB beach volleyball team performance. *International Journal of Performance Analysis in Sport*, 8(1), 31–43.
- Glatz, E., Mavromatidis, S., Ager, B. & Dimitropoulos, X. (2014). Visualizing big network traffic data using frequent pattern mining and hypergraphs. *Computing*, 96(1), 27–38.
- Glazier, P. S. (2010). Game, set and match? Substantive issues and future directions in performance analysis. *Sports medicine*, 40(8), 625–634.
- Gomez, G., Herrera López, P., Link, D. & Eskofier, B. (2014). Tracking of Ball and Players in Beach Volleyball Videos. *PLOS ONE*, 9(11), 1–19.
- Gomez, G., Linarth, A., Link, D. & Eskofier, B. (2012). Semi-automatic tracking of beach volleyball players. *Sportinformatik 2012*, 22.
- Goodfellow, I., Bengio, Y. & Courville, A. (2016). *Deep Learning* [<http://www.deeplearningbook.org>]. MIT Press.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. & Bengio, Y. (2014). Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence & K. Weinberger (Hrsg.), *Advances in Neural Information Processing Systems 27 (NIPS 2014)* (S. 2672–2680).
- Google. (2020). *Major League Baseball and Google Expand Relationship to Offer More Highly Personalized Experiences and Increase Engagement for Fans around the World*. Abgerufen am: 30. Mai 2020 unter <https://cloud.google.com/press-releases/2020/0303/mlb-and-google>
- Graves, A. & Schmidhuber, J. (2008). Offline handwriting recognition with multidimensional recurrent neural networks. *Advances in neural information processing systems*, 21, 545–552.
- Grewal, J. K., Krzywinski, M. & Altman, N. (2019). Markov models—Markov chains. *Nat. Methods*, 16, 663–664.
- Gu, W., Foster, K., Shang, J. & Wei, L. (2019). A game-predicting expert system using big data and machine learning. *Expert Systems with Applications*, 130, 293–305.

- Gumm, H. P. & Sommer, M. (2010). *Einführung in die Informatik* (Bd. 9). Oldenbourg Wissenschaftsverlag.
- Haider, F., Salim, F. A., Postma, D. B., van Delden, R., Reidsma, D., van Beijnum, B.-J. & Luz, S. (2020). A Super-Bagging Method for Volleyball Action Recognition Using Wearable Sensors. *Multimodal Technologies and Interaction*, 4(2).
- Han, J., Pei, J., Yin, Y. & Mao, R. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data mining and knowledge discovery*, 8(1), 53–87.
- Hand, D. J. (2007). Principles of data mining. *Drug safety*, 30(7), 621–622.
- Hansmann, U., Merk, L., Nicklous, M. S. & Stober, T. (2003). *Pervasive computing: The mobile world*. Springer Science & Business Media.
- Hao, Z., Tian, Z., Hao, Y. & Song, J. (2010). Analysis on technique and tactics of Lin Ma and Hao Wang in the men's single table tennis final in the 29th Olympic Games. *International Journal of Table Tennis Science*, 6, 74–78.
- Harrop, K. & Nevill, A. (2014). Performance indicators that predict success in an English professional League One soccer team. *International Journal of Performance Analysis in Sport*, 14(3), 907–920.
- Hasler, M. (2015). *A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules*. Abgerufen am: 12. November 2020 unter https://michael.hahsler.net/research/association_rules/measures.html
- Häyrinen, M. & Tampouratzis, K. (2012). Technical and tactical game analysis of elite female beach volleyball. *Jyväskylä, KIHU: Research Institute for Olympic Sports*.
- Hinton, G. E., Sejnowski, T. J. & Ackley, D. H. (1984). *Boltzmann machines: Constraint satisfaction networks that learn*. Carnegie-Mellon University, Department of Computer Science Pittsburgh, PA.
- Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.
- Hoffmeister, S. (2019). *Sport-Strategy Optimization with Markov Decision Processes* (Diss.).
- Hohmann, A., Lames, M., Letzelter, M. & Pfeiffer, M. (2020). Einführung in die Trainingswissenschaft.
- Hsu, M.-H. (2010). A study on the technical analysis and attack-defense performance of men's top four single players in 2008 Olympic Games. *International Journal of Table Tennis Science*, 6, 248–260.
- Hsu, M.-H., Chen, Y.-F. & Wang, S.-C. (2014). Offense-Defense Mode Analysis of the World Top Male Table Tennis Player-A Case Study by Chuang Chih-Yuan Who Participated in 2012 London Olympic Male Single Games. *Journal of Science and Innovation*, 4(2), 41–49.
- Hsu, Y.-C. (2020). Using Machine Learning and Candlestick Patterns to Predict the Outcomes of American Football Games. *Applied Sciences*, 10(13), 4484.
- Hughes, M. (1985). A comparison of the patterns of play of squash. *International ergonomics*, 85, 139–141.

- Hughes, M. & Bartlett, R. (2002). The use of performance indicators in performance analysis. *Journal of sports sciences*, 20(10), 739–754.
- Hughes, M. & Franks, I. (2007). *The essentials of performance analysis: an introduction*. Routledge.
- Huiqun, G. H. Z. (2006). Application of Association Rule Mining in Analyzing Techniques and Tactics of Table Tennis Match [J]. *Journal of North China University of Technology Beijing China*, 1(003).
- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
- Ji, Y.-F., Zhang, J.-W., Shi, Z.-h., Liu, M.-H. & Ren, J. (2018). Research on real – time tracking of table tennis ball based on machine learning with low-speed camera. *Systems Science & Control Engineering*, 6(1), 71–79.
- Katsikadelis, M., Pilianidis, T. & Mantzouranis, N. (2013). The interaction between serves and match winning in table tennis players in the London 2012 Olympic Games. In J.-F. Kahn, K. Yoshida & M. Kondric (Hrsg.), *Book of abstracts of the 8th International Table Tennis Federation Sports Science Congress* (S. 77–79).
- Kautz, T., Groh, B. H., Hannink, J., Jensen, U., Strubberg, H. & Eskofier, B. M. (2017). Activity recognition in beach volleyball using a Deep Convolutional Neural Network. *Data Mining and Knowledge Discovery*, 31(6), 1678–1705.
- Khaustov, V. & Mozgovoy, M. (2020). Recognizing Events in Spatiotemporal Soccer Data. *Applied Sciences*, 10(22), 8046.
- Koch, C. & Tilp, M. (2009a). Analysis of beach volleyball action sequences of female top athletes.
- Koch, C. & Tilp, M. (2009b). Beach volleyball techniques and tactics: A comparison of male and female playing characteristics. *Kinesiology: International journal of fundamental and applied kinesiology*, 41(1), 52–59.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1), 59–69.
- Koseler, K. & Stephan, M. (2017). Machine learning applications in baseball: A systematic literature review. *Applied Artificial Intelligence*, 31(9-10), 745–763.
- Kumar, P. (2015). *3 ways big data and analytics will change sports*. Abgerufen am: 30. Mai 2020 unter <https://www.ibmbigdatahub.com/blog/3-ways-big-data-and-analytics-will-change-sports>
- Lago-Peñas, C., Lago-Ballesteros, J., Dellal, A. & Gómez, M. (2010). Game-related statistics that discriminated winning, drawing and losing teams from the Spanish soccer league. *Journal of sports science & medicine*, 9(2), 288.
- Lai, M., Meo, R., Schifanella, R. & Sulis, E. (2018). The role of the network of matches on predicting success in table tennis. *Journal of Sports Sciences*, 36(23), 2691–2698. <https://doi.org/10.1080/02640414.2018.1482813>
- Lames, M. (1991). *Leistungsdiagnostik durch Computersimulation: Ein Beitrag zur Theorie der Sportspiele am Beispiel Tennis* (Bd. 17). Harri Deutsch.

- Lames, M. (2006). Modelling the interaction in game sports-Relative phase and moving correlations. *Journal of sports science & medicine*, 5(4), 556.
- Lames, M. & Hohmann, A. (1997). Zur Leistungsrelevanz von Spielhandlungen im Volleyball. *Integrative Aspekte in Theorie und Praxis der Rückschlagspiele*, 77(5), 121–128.
- Lames, M. & McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport*, 7(1), 62–79.
- Le, H. M., Carr, P., Yue, Y. & Lucey, P. (2017). Data-driven ghosting using deep imitation learning. *In Proceedings of the MIT Sloan Sports Analytics Conference 2017*.
- LeCun, Y. et al. (1989). Generalization and network design strategies. *Connectionism in perspective*, 19, 143–155.
- LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- Lees, A. (1985). Computers in sport. *Applied ergonomics*, 16(1), 3–10.
- Lenca, P., Vaillant, B., Meyer, P. & Lallich, S. (2007). Association rule interestingness measures: Experimental and theoretical studies. In F. Guillet & H. Hamilton (Hrsg.), *Quality Measures in Data Mining* (S. 51–76). Springer.
- Letzelter, H. & Letzelter, M. (1982). Die Struktur sportlicher Leistungen als Gegenstand der Leistungsdiagnostik in der Trainingswissenschaft. *Leistungssport*, 5(12), 351–361.
- Ley, C., Dominicy, Y. & Bruneel, W. (2018). Mutual point-winning probabilities (MPW): a new performance measure for table tennis. *Journal of sports sciences*, 36(23), 2684–2690.
- Lim, S.-M., Oh, H.-C., Kim, J., Lee, J. & Park, J. (2018). LSTM-Guided Coaching Assistant for Table Tennis Practice. *Sensors*, 18(12). <https://doi.org/10.3390/s18124112>
- Link, D. (2014). A Toolset for Beach Volleyball Game Analysis Based on Object Tracking. *Int. J. Comput. Sci. Sport*, 13.
- Link, D. & Ahmann, J. (2013). Moderne spielbeobachtung im beach-volleyball auf basis von positionsdaten. *Sportwissenschaft*, 43(1), 1–11.
- Link, D. & Lames, M. (2009). Matrix Reloaded - Struktur und Gegenstand der Sportinformatik. In M. Lames, C. Augste, O. Cordes, C. Dreckmann, K. Görsdorf & M. Siegle (Hrsg.), *Gegenstand und Anwendungsfelder der Sportinformatik*. Czwalina-Verlag.
- Link, D., Lang, S. & Seidenschwarz, P. (2016). Real time quantification of dangerousity in football using spatiotemporal tracking data. *PloS one*, 11(12), e0168768.
- Liu, J., Tong, X., Li, W., Wang, T., Zhang, Y. & Wang, H. (2009). Automatic player detection, labeling and tracking in broadcast soccer video [Video-based Object and Event Analysis]. *Pattern Recognition Letters*, 30(2), 103–113.
- Liu, R., Shi, Z. & Kong, L. (2015). The research on the application of double moving average in table tennis competition analysis. In M. Kondric, Z. Xiaopeng & X. Dandan (Hrsg.), *Book of abstracts of the 14th ITTF Sports Science Congress and 5th World Racquet Sports Congress* (S. 70).
- Liu, R., Wang, Z., Shi, X., Zhao, H., Qiu, S., Li, J. & Yang, N. (2019). Table tennis stroke recognition based on body sensor network. In R. Montella, A. Ciaramella, G. Fortina,

- A. Guerrieri & A. Liotta (Hrsg.), *International Conference on Internet and Distributed Computing Systems* (S. 1–10).
- Loh, W.-Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1), 14–23.
- Ma, H. (2020). Improvement of table tennis technology based on data mining in the environment of wireless sensor networks. *International Journal of Distributed Sensor Networks*, 16(10), 1550147720961343.
- Mackenzie, R. & Cushion, C. (2013). Performance analysis in football: A critical review and implications for future research. *Journal of sports sciences*, 31(6), 639–676.
- Maher, M. (2013). Predicting the outcome of the Ryder cup. *IMA Journal of Management Mathematics*, 24(3), 301–309.
- Malagoli Lanzoni, I., Lobiatti, R. & Merni, F. (2007). Footwork techniques used in table tennis: a qualitative analysis. In M. Kondric & J.-F. Kahn (Hrsg.), *Proceedings book of The 10th Anniversary ITTF Sport Science Congress* (S. 401–408).
- Malagoli Lanzoni, I., Di Michele, R. & Merni, F. (2013). Distribution of stroke and footwork types in top-level men's and women's table tennis. *Performance Analysis of Sport IX*, 168.
- Malagoli Lanzoni, I., Di Michele, R. & Merni, F. (2014). A notational analysis of shot characteristics in top-level table tennis players. *European journal of sport science*, 14(4), 309–317.
- Malagoli Lanzoni, I., Michele, R. & Merni, F. (2013). Technical and tactical performance of top-class senior, junior and cadet table tennis players. *International Journal of Table Tennis Sciences*, 8, 80–85.
- Marr, B. (2018). *How Much Data Do We Create Every Day?* Abgerufen am: 14. August 2020 unter <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#6c6a3db160ba>
- Mateus, J. (2005). In pursuit of an ecological and fractal approach to football coaching. *Science and Football V*, 571–583.
- Mauthner, T., Koch, C., Tilp, M. & Bischof, H. (2007). Visual tracking of athletes in beach volleyball using a single camera. *International Journal of Computer Science in Sport*, 6(2), 21–34.
- Messenger, R. & Mandell, L. (1972). A modal search technique for predictive nominal scale multivariate analysis. *Journal of the American statistical association*, 67(340), 768–772.
- Miethling, W. & Perl, J. (1981). *Computerunterstützte Sportspielanalyse: mathematische Ansätze und praktische Möglichkeiten*. Czwalina.
- Miller, A., Bornn, L., Adams, R. & Goldsberry, K. (2014). Factorized point process intensities: A spatial analysis of professional basketball. In E. P. Xing & T. Jebara (Hrsg.), *ICML'14: Proceedings of the 31st International Conference on International Conference on Machine Learning* (S. 235–243).

- Munivrana, G., Furjan-Mandić, G. & Kondrič, M. (2015). Determining the structure and evaluating the role of technical-tactical elements in basic table tennis playing systems. *International journal of sports science & coaching*, 10(1), 111–132.
- Munivrana, G., Petrinović, L. Z. & Kondrič, M. (2015). Structural analysis of technical-tactical elements in table tennis and their role in different playing zones. *Journal of Human Kinetics*, 47(1), 197–214.
- Murray, S. & Hughes, M. (2001). Tactical performance profiling in elite level senior squash. *pass. com, Cardiff: CPA, UWIC*, 194.
- Mwamikazi, E., Fournier-Viger, P., Moghrabi, C. & Baudouin, R. (2014). A dynamic questionnaire to further reduce questions in learning style assessment. In L. Iliadis, I. Maglogiannis & H. Papadopoulos (Hrsg.), *Proceedings of the 10th IFIP WG 12.5 International Conference on Artificial Intelligence Applications and Innovations (AIAI 2014)* (S. 224–235).
- Narizuka, T., Yamamoto, K. & Yamazaki, Y. (2014). Statistical properties of position-dependent ball-passing networks in football games. *Physica A: Statistical Mechanics and its Applications*, 412, 157–168.
- Narvekar, M. & Syed, S. F. (2015). An optimized algorithm for association rule mining using FP tree. *Procedia Computer Science*, 45, 101–110.
- Naulaerts, S., Meysman, P., Bittremieux, W., Vu, T. N., Vanden Berghe, W., Goethals, B. & Laukens, K. (2015). A primer to frequent itemset mining for bioinformatics. *Briefings in bioinformatics*, 16(2), 216–231.
- Newell, K. (1986). Constraints on the development of coordination. *Motor development in children: Aspects of coordination and control*.
- Ofoghi, B., Zeleznikow, J., MacMahon, C. & Raab, M. (2013). Data Mining in Elite Sports: A Review and a Framework. *Measurement in Physical Education and Exercise Science*, 17(3), 171–186. <https://doi.org/10.1080/1091367X.2013.805137>
- Olah, C. (2015). *Understanding LSTM Networks*. Abgerufen am: 29. Mai 2020 unter <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Passos, P., Araújo, D. & Volossovitch, A. (2016). *Performance analysis in team sports*. Taylor & Francis.
- Percy, D. F. (2009). A mathematical analysis of badminton scoring systems. *Journal of the Operational Research Society*, 60(1), 63–71.
- Perl, J. (2006). A computer science in sport: An overview of present fields and future applications (Part II). *Int. J. of Computer Science in Sport*, 4(1), 36–45.
- Perl, J., Lames, M. & Miethling, W.-D. (1997). *Informatik im Sport: ein Handbuch* (Bd. 117). Hofmann.
- Pfeiffer, M., Zhang, H. & Hohmann, A. (2010). A Markov chain model of elite table tennis competition. *International Journal of Sports Science & Coaching*, 5(2), 205–222.
- Power, D. J. (2002). *What is the "true story" about data mining, beer and diapers?* Abgerufen am: 29. Mai 2020 unter <http://www.dssresources.com/newsletters/66.php>
- Quinlan, J. R. (1993). *C4.5: programs for machine learning*. Elsevier.

- Reeser, J. C., Fleisig, G. S., Bolt, B. & Ruan, M. (2010). Upper limb biomechanics during the volleyball serve and spike. *Sports Health*, 2(5), 368–374.
- Rein, R. & Memmert, D. (2016). Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. *SpringerPlus*, 5(1), 1–13.
- Ribeiro, M. T., Singh, S. & Guestrin, C. (2016). "Why should i trust you? Explaining the predictions of any classifier. In B. Krisnapuram & M. Shah (Hrsg.), *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (S. 1135–1144).
- Ronneberger, O., Fischer, P. & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. Wells & A. Frangi (Hrsg.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (S. 234–241).
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533–536.
- Runkler, T. A. (2010). *Data mining*. Springer.
- Sahi, G. (2020). Applying Association Rule Mining to Analyze the Performance of Indian Cricket Team in T20 Format. Available at SSRN 3565907.
- Sak, H., Senior, A. W. & Beaufays, F. (2014). *Long short-term memory recurrent neural network architectures for large scale acoustic modeling*. Abgerufen am: 12. November 2020 unter <https://storage.googleapis.com/pub-tools-public-publication-data/pdf/43905.pdf>
- Sampaio, J. & Maçãs, V. (2012). Measuring tactical behaviour in football. *International journal of sports medicine*, 33(05), 395–401.
- Sampaio, J., McGarry, T., Calleja-González, J., Jiménez Sáiz, S., Schelling i del Alcázar, X. & Balciunas, M. (2015). Exploring game performance in the National Basketball Association using player tracking data. *PloS one*, 10(7), e0132894.
- Santos-Fernandez, E., Wu, P. & Mengersen, K. L. (2019). Bayesian statistics meets sports: a comprehensive review. *Journal of Quantitative Analysis in Sports*, 15(4), 289–312.
- Sauerbier, T. (1999). *Theorie und Praxis von Simulationssystemen*. Springer.
- Schläppi-Lienhard, O. & Hossner, E.-J. (2015). Decision making in beach volleyball defense: Crucial factors derived from interviews with top-level experts. *Psychology of Sport and Exercise*, 16, 60–73.
- Schulman, R. & Hamdan, M. (1977). A probabilistic model for table tennis. *Canadian Journal of Statistics*, 5(2), 179–186.
- Serrien, B., Goossens, M. & Baeyens, J. (2017). Issues in Using Self-Organizing Maps in Human Movement and Sport Science. *International Journal of Computer Science in Sport*, 16(1), 1–17.
- Serrien, B., Ooijen, J., Goossens, M. & Baeyens, J.-P. (2016). A motion analysis in the volleyball spike—part 1: three dimensional kinematics and performance. *Int J Hum Mov Sports Sci*, 4(4), 70–82.
- Shoaran, M., Haghi, B. A., Taghavi, M., Farivar, M. & Emami-Neyestanak, A. (2018). Energy-efficient classification for resource-constrained biomedical applications. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 8(4), 693–707.

- Siegelmann, H. & Sontag, E. D. (1995). On the computational power of neural nets. *J. Comp. Syst. Sci.*, 50, 132–150.
- Singh, T., Singla, V. & Bhatia, P. (2015). Score and winning prediction in cricket through data mining. *2015 International Conference on Soft Computing Techniques and Implementations (ICSCTI)*, 60–66.
- Smith, L. N. (2017). Cyclical learning rates for training neural networks. In G. Medioni, D. Michael & S. Sarkar (Hrsg.), *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)* (S. 464–472). IEEE Computer Society.
- Stachowiak, H. (1973). *Allgemeine Modelltheorie*. Springer.
- Steinmüller, W. (1993). *Informationstechnologie und Gesellschaft: Einführung in die angewandte Informatik*. Wissenschaftliche Buchgesellschaft.
- Strobl, C., Malley, J. & Tutz, G. (2009). An introduction to recursive partitioning: rationale, application and characteristics of classification and regression trees, bagging and random forests. *Psychological methods*, 14(4), 323.
- Sun, J., Yu, W. & Zhao, H. (2010). Study of association rule mining on technical action of ball games. In K. Lee & H.-X. Li (Hrsg.), *2010 International Conference on Measuring Technology and Mechatronics Automation* (S. 539–542).
- Tabrizi, S. S., Pashazadeh, S. & Javani, V. (2020). Comparative Study of Table Tennis Forehand Strokes Classification Using Deep Learning and SVM. *IEEE Sensors Journal*, 20(22), 13552–13561. <https://doi.org/10.1109/JSEN.2020.3005443>
- Tamaki, S. & Saito, H. (2013). Reconstruction of 3d trajectories for performance analysis in table tennis. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1019–1026.
- Tamaki, S., Yoshida, K. & Yamada, K. (2017). A shot number based approach to performance analysis in table tennis. *Journal of human kinetics*, 55(1), 7–18.
- Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C. & Liu, C. (2018). A survey on deep transfer learning. In V. Kurkova, Y. Manolopoulos, B. Hammer, L. Iliadis & I. Maglogiannis (Hrsg.), *Proceedings of the 27th International Conference on Artificial Neural Networks* (S. 270–279).
- Tan, P.-N., Kumar, V. & Srivastava, J. (2004). Selecting the right objective measure for association analysis. *Information Systems*, 29(4), 293–313.
- Tebbe, J., Gao, Y., Sastre-Rienietz, M. & Zell, A. (2018). A table tennis robot system using an industrial kuka robot arm. In T. Brox, A. Bruhn & M. Fritz (Hrsg.), *Proceedings of the 40th German Conference on Pattern Recognition (GCPR 2018)* (S. 33–45).
- Theodoropoulos, J. S., Bettle, J. & Kosy, J. D. (2020). The use of GPS and inertial devices for player monitoring in team sports: A review of current and future applications. *Orthopedic Reviews*, 12(1).
- Tian, C., De Silva, V., Caine, M. & Swanson, S. (2020). Use of Machine Learning to Automate the Identification of Basketball Strategies Using Whole Team Player Tracking Data. *Applied Sciences*, 10(1), 24.

- Tilp, M., Koch, C., Stifter, S. & Ruppert, S. G. (2006). Digital game analysis in beach volleyball. *International Journal of Performance Analysis in Sport*, 6(1), 140–148. <https://doi.org/10.1080/24748668.2006.11868362>
- Tilp, M. & Rindler, M. (2013). Landing techniques in beach volleyball. *Journal of Sports Science & Medicine*, 12(3), 447.
- Tilp, M., Wagner, H. & Müller, E. (2008). Differences in 3D kinematics between volleyball and beach volleyball spike movements. *Sports Biomechanics*, 7(3), 386–397.
- Tolbert, B. & Trafalis, T. (2016). Predicting major league baseball championship winners through data mining. *Athens Journal of Sports*, 3(4), 239–252.
- Tora, M. R., Chen, J. & Little, J. J. (2017). Classification of Puck Possession Events in Ice Hockey. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 91–98.
- Trask, A., Gilmore, D. & Russell, M. (2015). Modeling order in neural word embeddings at scale. *arXiv preprint arXiv:1506.02338*.
- Travassos, B., Araújo, D., Vilar, L. & McGarry, T. (2011). Interpersonal coordination and ball dynamics in futsal (indoor football). *Human movement science*, 30(6), 1245–1259.
- Van Haaren, J., Ben Shitrit, H., Davis, J. & Fua, P. (2016). Analyzing Volleyball Match Data from the 2014 World Championships Using Machine Learning Techniques. In B. Krishnapuram & M. Shah (Hrsg.), *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (S. 627–634). Association for Computing Machinery.
- Voeikov, R., Falaleev, N. & Baikulov, R. (2020). TNet: Real-time temporal and spatial video analysis of table tennis. In T. Boult, G. Medioni & R. Zabih (Hrsg.), *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (S. 884–885).
- Wang, J., Zhao, K., Deng, D., Cao, A., Xie, X., Zhou, Z., Zhang, H. & Wu, Y. (2019). Tac-Simur: Tactic-based simulative visual analytics of table tennis. *IEEE transactions on visualization and computer graphics*, 26(1), 407–417.
- Wang, K.-C. & Zemel, R. (2016). Classifying NBA offensive plays using neural networks. *Proceedings of the 2016 MIT Sloan Sports Analytics Conference*, 4.
- Westermann, T. (2011). *Mathematik für Ingenieure - Ein anwendungsorientiertes Lehrbuch*. Springer-Verlag Berlin Heidelberg. <http://www.home.hs-karlsruhe.de/~weth0002/buecher/mathe/start.htm>
- Woodham, M., Hawkins, J., Singh, A. & Chakraborty, S. (2019). When to Pull Starting Pitchers in Major League Baseball? A Data Mining Approach. In T. Khoshgoftaar, D. Wang & H. Wang (Hrsg.), *18th IEEE International Conference On Machine Learning And Applications (ICMLA 2019)* (S. 426–431).
- Wu, H. & Li, Z. (1992). Research on technical diagnosis method for table tennis players. *International Journal of Table Tennis Sciences*, 1, 99–103.
- Wu, Y., Lan, J., Shu, X., Ji, C., Zhao, K., Wang, J. & Zhang, H. (2018). iTTVis: Interactive Visualization of Table Tennis Data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1), 709–718.

- Yang, G.-Z. & Yang, G. (2006). *Body sensor networks* (Bd. 1). Springer.
- Yang, Q. & Zhang, H. (2016). Application of BP neural network and multiple regression in table tennis technical and tactical ability analysis. *Journal of Chengdu Sport University*, 42(1), 78–82.
- Zetou, E., Moustakidis, A., Tsigilis, N. & Komninakidou, A. (2007). Does Effectiveness of Skill in Complex I Predict Win in Men's Olympic Volleyball Games? *Journal of Quantitative Analysis in Sports*, 3(4). <https://doi.org/10.2202/1559-0410.1076>
- Zhou, Y.-T. & Chellappa, R. (1988). Computation of optical flow using a neural network. *IEEE International Conference on Neural Networks, 1998*, 71–78.

A. Anhang

A.1. Genehmigung zur Veröffentlichung der Publikationen

FW: Publishing permission for dissertation

Nowacka, Katarzyna <Katarzyna.Nowacka@sciendo.com>

Mo 13.09.2021 17:23

An: Wenninger, Sebastian <sebastian.wenninger@tum.de>

Dear Sebastian,

Thank you very much for your email.

Regarding your request we confirm that the journal International Journal of Computer Science in Sport is publishing on our SCIENDO platform in Open Access mode using the license CC BY-NC-ND-4.0., so you are free to share — copy and redistribute the material in any medium or format under some terms, that are described under following link: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Best Regards,

Katarzyna Nowacka
Customer Service Specialist
SCIENDO

T: +48 696 528 363

E: Katarzyna.Nowacka@sciendo.com

www.sciendo.com

Sciendo is De Gruyter's brand for publishing services
De Gruyter Poland Sp. z o.o., ul. Bogumiła Zuga 32A, 01-811 Warsaw, Poland
Walter de Gruyter GmbH, Genthiner Strasse 13, 10785 Berlin, Germany

Your personal data related to this mail is processed by De Gruyter Poland Sp. z o.o. with its registered office in Warsaw, ul. Zuga 32a, Poland, acting as the Data Controller. The data will be processed as necessary for this communication and corresponding purposes. Further information concerning the processing and protection of your personal data can be found in our Data Processing Policy.
an be found in our Data Processing Policy.

From: Wenninger, Sebastian <sebastian.wenninger@tum.de>

Sent: Sunday, September 12, 2021 4:03 PM

To: info | Sciendo <info@sciendo.com>

Subject: Publishing permission for dissertation

Dear Sir or Madam,

I kindly ask for your permission to include the following articles in full-text in my dissertation:

Performance Analysis in Table Tennis - Stochastic Simulation by Numerical Derivation (<https://doi.org/10.1515/ijcss-2016-0002>)

Data Mining in Elite Beach Volleyball – Detecting Tactical Patterns Using Market Basket Analysis (<https://doi.org/10.2478/ijcss-2019-0010>)

Performance of machine learning models in application to beach volleyball data. (<https://doi.org/10.2478/ijcss-2020-0002>)

Sincerely yours

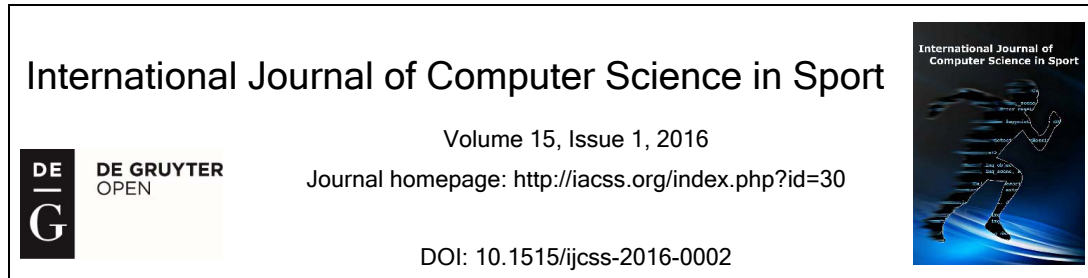
Sebastian Wenninger

Technische Universität München
Fakultät für Sport- und Gesundheitswissenschaft
Lehrstuhl für Trainingswissenschaft und Sportinformatik
Uptown München - Campus D – Raum M206
Georg-Brauchle-Ring 60/62
80992 München

E-mail: sebastian.wenninger@tum.de

Web: <http://www.trainingswissenschaft.se.tum.de/>

A.2. Studie 1 - Originalveröffentlichung



Performance Analysis in Table Tennis - Stochastic Simulation by Numerical Derivation

Wenninger, S., Lames, M.

Department of Training Science and Sport Informatics, Technische Universität München

Abstract

The aim of this study was to identify the impact of different tactical behaviors on the winning probability in table tennis. The performance analysis was done by mathematical simulation using a Markov chain model. 259 high-level table tennis games were evaluated by means of a new simulation approach using numerical derivation to remove the necessity to perform a second modeling step in order to determine the difficulty of tactical behaviors. Based on the derivation, several mathematical constructs like directional derivations and the gradient are examined for application in table tennis. Results reveal errors and long rallies as the most influencing game situations, together with the positive effect of risky play on the winning probability of losing players.

KEYWORDS: MARKOV CHAIN, PERFORMANCE ANALYSIS, TABLE TENNIS, SIMULATION, NUMERICAL ANALYSIS

Introduction

There has been growing interest in tactical and technical analysis in sports science (Lees, 2003). This process, also called performance analysis, aims to provide coaches with the information to optimize their decision-making during competitions as well as in training. Hughes and Bartlett (2002) analyzed the structure of different sports resulting in a classification of performance indicators into the three categories technical, tactical and biomechanical. Both O'Donoghue (2004, 2008) and Hughes (1998) examined match analysis in racket sports and suggested methods for the selection of indicators.

There exist several studies that aim to apply performance analysis to table tennis. Malagoli-Lanzoni, Di Michele and Merni (2011) summarize the different performance indicators commonly used in table tennis literature. The process of data collection was also analyzed and enhanced with the help of software tools by different authors (Baca, Baron, Leser & Kain, 2004; Juan, Zhang & Hu, 2008). Leser and Baca (2009) additionally developed a software for the qualitative analysis of performance in table tennis by collecting numerous types of data for

the stroke, ball positions, the players and the match. It is immediately obvious that performance analysis in table tennis is a large topic, with a wide range of research covering many aspects of performance to improve training and competition.

Most of the aforementioned studies try to assess performance by using notational analysis represented by statistics, indices and coefficients, which can be summarized as the absolute individual balance of successful and unsuccessful game actions (Lames et al., 1997). These indicators examine the individual performance outside of the context of the game actions, losing an important feature for evaluation. Hughes and Bartlett (2002) specifically emphasize that the use of match classification indicators, when observed in isolation, can be misleading.

To treat this, studies employed analysis which view the game as a system in which the behavior of the system is influenced by the interaction of the two playing parties, employing probabilistic models to calculate the outcome of matches (Strauss & Arnold, 1987; Newton & Aslam, 2009). Other research also included technical and tactical behavior into the stochastic models, first to predict playing patterns (McGarry & Franks, 1994) and second to determine the performance relevance of tactical behavior (Lames, 1991; Pfeiffer, 2005; Zhang, 2006; Pfeiffer & Hohmann 2008).

Lames (1991) proposes a performance analysis by mathematical simulation in tennis, which is based on transition matrices and tries to support the evaluation of game situations within their respective context. By assuming specific properties of the game actions, described by the Markov chain model, transition matrices can be treated as stochastic processes which then allow the application of different computations on the matrices. The impact of a tactical behavior hereby is described by the change in the winning probability induced by the changes to the transition matrices done through the simulation. Lames and McGarry (2007) use this simulative approach to determine the types of tactical behaviors relevant to athlete's performance in volleyball, while Pfeiffer, Zhang and Hohmann (2010) apply it to table tennis.

In addition to the first step of modeling, where the game actions are modeled by a Markov chain, these simulative approaches need a second step to determine the relative difficulty of tactical behaviors so the changes to the transition matrices fit into the semantics of the sport. The aim of this study was to improve the original simulation method by replacing the second modeling step through application of mathematical derivation. This substitutes the needed difficulty function with well-known mathematical principles.

Methods

Participants

259 matches of world class table tennis players, both male and female, were examined in this study. The evaluated games took place over the span of four years (2008-2012) in official tournaments of the ITTF, ETTU and IOC. The dataset consists of the length, server and winner per rally, together with the final result of each match.

Design

A table tennis rally was classified by the following discrete states: First to fifth stroke, more than five strokes, point and error, each for both players. Players are named "A" and "B", with player A to be known as the winner of the match. In context of the observation system, a single rally can be seen as a system that alternates between the stroke states of the players. This system starts with the serve (first stroke) and ends with a point/error for either player A or B.

Markov Simulation

Markov chains

A Markov chain is a special type of stochastic process, with a stochastic process defined as a system that moves stepwise through a defined set of states. The transitions between these states then occur with certain probabilities. A finite stochastic process is a stochastic process that can only reach a finite amount of states. The Markov property of a stochastic process defines the conditional probability for a state to be reached in the next step, given all previous steps of the process, as equal to the conditional probability for this event with only the current state given. Simply said, the Markov property assumes that each state of a rally only depends on the last former state and is independent of all earlier states. The system is assumed to have no memory. A Markov process is further called Markov chain, if the transition probability between two states is independent from the position of these states in the chain of events. Markov chains also know two types of states: Transient states, which cannot be returned to by the system, and absorbing states, which cannot be left (the transition probability to the own state equals 1). A Markov chain can then be expressed by transition matrices as seen in figure 1.

%	B2	B3	B4	B5	B>5	PA	PB
A1	82,0					5,0	13,0
A2		90,0				2,0	8,0
A3			88,4			2,3	9,3
A4				84,2		0,0	15,8
A5					76,7	10,0	13,3
A>5					83,3	0,9	15,7
	A2	A3	A4	A5	A>5	PB	PA
B1	100,0					0,0	0,0
B2		89,6				4,2	6,3
B3			84,4			4,4	11,1
B4				78,9		2,6	18,4
B5					75,0	0,0	25,0
B>5					74,3	1,8	23,9

Figure 1: Markov transition matrix for a table tennis game, with submatrices Q (yellow) and R (green)

These matrices consist of several submatrices, with Q and R being the most relevant. Q contains all transition probabilities of the transient (non-absorbing) states (the transitions between the strokes) among each other, while R contains the transitions from the transient states to the absorbing states (point/error), (Lames, 1991). The theory of Markov chains also allows the computation of several properties of the evaluated process. The normal matrix $N = (I - Q)^{-1}$, with I as the identity matrix, contains the mean frequency of occurrence for each transient state. From that, the matrix $B = N \cdot R$, an $s \times r$ matrix which contains the probabilities to be absorbed in the absorbing state r for each of the s transient states. The winning probability for a rally for each player can then be taken from the first two rows of the matrix B (Lames, 1991).

Simulation

After the observation system has been modeled as a Markov chain, the winning probabilities of the unmodified transition matrices can be calculated. In the second step the transition matrices have to be manipulated in such a way, that the new transition probabilities express the desired deviation in tactics. Additionally, in order to retain the stochastic nature of the transition

matrices, the tactical deviations have to be compensated by other transition probabilities, so that the probabilities per row still add up to one. Figure 2 shows an example manipulation on the transition probability of the serve of player A to the return of player B from the matrix in figure 1. In this case, the transition probabilities to the direct point and error states were reduced by 5%, while the transition from the serve of player A to the return of player B was increased by the resulting 10% in turn. This type of manipulation will later be defined as “safer play”, since it conveys the tendency of a player to trade a higher direct point rate with a lower direct error rate. In the last step, the winning probabilities have to be calculated again for the manipulated transition matrix. The resulting difference in the winning probability gives an estimate of the importance of the modeled behavior. However, the deflection of tactical behavior requires an additional modeling step, because the amount of the manipulation has to account for the relative difficulty of the behavioral change. As an example, if a player already achieves an extraordinary point rate of 90% at a certain stroke, increasing that point rate by 5% is very difficult in reality. Increasing the point rate from 50% to 60% is considerably easier though. Hence the simulation requires a “difficulty function”, which conducts small deviations for difficult behavioral changes, and bigger variations for easier tactical behaviors.

%	B2	B3	B4	B5	B>5	PA	PB
A1	↑ 92,0 (+10)					↓ 0,0 (-5)	↓ 8,0 (-5)

Figure 2: Transition matrix row after the manipulation of transitions, with the amount of manipulation in brackets

Numerical Derivation

The idea of this study was to replace the additional modeling step to determine the difficulty of tactical behavior with the evaluation of the derivative of the function for the winning probability at this point of the Markov chain. However, since there is no analytical definition for the derivative of the matrix operations involved in the computation of the winning probability, numerical derivation has to be used. This means that an appropriate numerical approximation for the derivative of markov winning probabilities has to be found.

Winning Probability as a Function

To compute the derivative of the winning probability of a player for a particular stroke number, we first have to define the function to derivate. The function $P_{Win} = f(\hat{x}) = f(x_1, \dots, x_{36})$ in the multi-dimensional state space of the Markov model takes the Markov transition matrices as input parameters and returns the winning probabilities for both players in a two-dimensional vector. In general, this function can be regarded as a blackbox around the matrix operations of the traditional simulation, so $f(x)$ essentially performs the computation of the matrix N and consequently matrix B . The wanted winning probability is then calculated from the matrix B under the assumption that the total number of services is distributed evenly among both players using the equation $f(\hat{x}) = 0.5 * (P_{ServeA} + P_{ServeB})$. Here, the point probabilities P_{ServeA} and P_{ServeB} denote the probability to end up in the corresponding absorbing state from either the serve of player A or the serve of player B. The overall winning probability for a rally is therefore composed of the probability to win the rally after ones own serve and the probability to win the rally after the opponent serves, both weighted equally. Although this does not entirely match the distribution of services between players, it gives a reasonable accurate estimation of service behavior over all games. The function $f = P_{Win}$ now has to be numerically derived to receive the impact of a tactical change on the winning probability.

Derivation

The original simulation used the absolute difference between the simulated winning probability $f(\hat{x} + \hat{h})$ and the observed winning probability $f(\hat{x})$ as an indicator for the impact of the simulated behavior. The vector \hat{x} describes the elements in the unmodified matrix, while the step size \hat{h} determines the amount of the manipulation to that vector. For the example shown in figure 2, \hat{h} would contain the value 10 for the transition to B2 and the value -5 for the transitions to PA and PB respectively, with all other values set to zero. The non-zero elements in \hat{h} therefore determine the behavior that is simulated by the derivation. The numerical derivation replaces this absolute difference with the slope of the winning probability. This slope can be interpreted as the ratio of vertical change (winning probability) to the horizontal change (manipulated transition probability).

There are several possibilities to compute the derivative of a function. The most common definition for the derivative is given by the forward difference

$$f' = \lim_{h \rightarrow 0} \frac{f(\hat{x} + \hat{h}) - f(\hat{x})}{|\hat{h}|}$$

which is already in numerically computable form. Another method is the backward difference

$$f' = \lim_{h \rightarrow 0} \frac{f(\hat{x}) - f(\hat{x} - \hat{h})}{|\hat{h}|}$$

that inverts the direction in which the slope is computed. However, in comparison to the forward difference and the backward difference, the central difference formula

$$f' = \lim_{h \rightarrow 0} \frac{f(\hat{x} + \hat{h}) - f(\hat{x} - h)}{2 * |\hat{h}|}$$

produces results with up to two orders of magnitude better precision (Press, Teukolsky, Vetterling & Flannery, 2007) and is hence used in most cases of the simulation. There exist two edge cases that require the application of the forward and backward difference, however. The forward difference is used to compute the derivative for matrix rows, where the transition probabilities have a value of zero, since it is impossible to apply the central difference at this point, because further subtraction of the step size would result in negative probabilities. Correspondingly, the backward difference is used for transitions that have the maximum probability value 1.

In order to achieve a high precision for the numerical computation, it is important to choose the step size \hat{h} so that it minimizes the round-off as well as the fractional error of the derivation computation. To this end, $\hat{h}(x) = \sqrt[4]{eps * x}$ has proven to be a suitable estimate for the central difference formula (Press et al., 2007). Here, x denotes the single transition probability for the currently simulated stroke, while eps describes the smallest machine-representable floating-point number. This is the point in the simulation, where the sports-specific difficulty function gets replaced by $\hat{h}(x)$, which only depends on numeric values present in the transition matrices and doesn't require further knowledge of the relative difficulty of tactical behaviors.

A special strategy to compute the step size $\hat{h}(x)$ is required in case the transition probability to simulate is zero. In that case the value of x is replaced by the mean value of all transition probabilities for the current dimension to provide a plausible step size for the central difference.

The value of the derivative now directly provides an indicator for the relevance of the simulated behavior.

Partial Derivation

In mathematics, the partial derivative of a multi-variable function is its derivative with respect to a single of those variables, while the rest is held constant. This means that only a single transition probability will be manipulated, or in other words, \hat{h} only contains one non-zero component for the respective transition. The goal of this analysis was to study the impact of point and error rates per stroke on the winning probability in table tennis. This is done by manipulating the transitions to the absorbing states “point” and “error” (PA or PB, depending on the player of the current stroke) respectively for each stroke. In order to reach a uniform distribution of the tactical change over the remaining variables, we chose to compensate it using the equation given by Lames (1991):

$$\delta P_{yi} = -\frac{P_{yi}}{1 - P_x} * \delta P_x, i = 1 \dots n$$

Here P_{yi} describes the n transition probabilities used to compensate for the change δP_x in the investigated probability P_x . The compensation also results in a directional derivative overall (\hat{h} contains > 1 non-zero elements) even though in theory the single change in the error / point probability represents a partial derivative.

Directional Derivation

The second tactical behavior examined in this study is the effect of safer/riskier play. The behavior “safer play” is defined as a decrease of both absorbing states at the current stroke. This roughly expresses the fact that a player is less likely to commit errors if he plays with more caution, but at the same time also scores less direct points. The term “riskier play” is defined accordingly by an increase of the error and point transition probabilities, which in turn expresses that the error rate is expected to grow at the same rate as the direct point rate if a player plays more aggressively. The necessary compensation for these manipulations was accounted for by the de- or increase of the total amount of change for the transition to the next state (e.g. from the third stroke of player A to the fourth stroke of player B) of that row, which is $2 * |\hat{h}|$ for the central difference formula and $|\hat{h}|$ for the forward and backward differences. Figure 2 shows an example of the manipulation of a matrix row for the simulation of safer play. To get the values for the simulation of risky play, the signs of the deflected values need to be switched.

Gradient

The mathematical gradient of a scalar function f is a vector operator, commonly denoted with the symbol ∇ . In a cartesian coordinate system it is a vector field whose components are the partial derivatives of f . Applied to the function for the winning probability the gradient results in the equation

$$\nabla f(x_1, \dots, x_{36}) = \frac{\partial f}{\partial x_1} e_1 + \dots + \frac{\partial f}{\partial x_{36}} e_{36}$$

where e_i are the orthogonal unit vectors of the coordinate system (Kaplan, 1992).

The orientation of ∇f is the direction in which the directional derivation of f has the largest value and $|\nabla f|$ (the length of the gradient vector) is the value of that directional derivative. When applied to the winning probability function $f(x_1, \dots, x_{36})$ it results in a vector of length

36 that contains the partial derivation of each possible transition in the original matrix. The main goal of the gradient analysis is to find the strokes with an outstanding influence on the outcome of a game. Since the gradient points in the direction of “steepest” change in the winning probability, it provides the possibility to automatically identify these transitions for each game and player. To that end, the simulation procedure was changed by a small step. In contrast to the aforementioned simulations, the service probability was not considered equally distributed over both players, but a game was simulated two times, one time for each player exclusively being at the serve. This changes the function for the winning probability to either $P_{Win} = P_{ServeA}$ or $P_{Win} = P_{ServeB}$, depending on which player is currently simulated. It also means that some transition probabilities can never be reached in a simulation. For example, if a player is serving, it can never happen that the same player hits the second or fourth ball in a rally. ∇f therefore contains the partial derivations of all transitions that can take place in an alternating state sequence, starting with the simulated player. As an example, for the simulation of player A as the serving player, this sequence would consist of the following states: (A1, B2, A3, B4, A5, B >5, A >5, ...). To obtain gradient vectors that contain only strokes of a single player (A1, A2, A3, ...), these alternations have to be taken into account. This is done by switching the corresponding derivation values between the gradient vectors. For example, to get the gradient values for the return of player A, the values from the simulation with player B as the serving player have to be considered. It also has to be taken into account that the states “point for player A” (PA) and “point for player B” (PB) have different semantics depending on which stroke is currently simulated. For example if player A is serving, PA stands for the service aces. However if player B is the serving player, PA represents the service errors.

This separated simulation of the players also causes the gradient vector to lose dimensions. Since not all 36 states are reachable from each configuration, the vector is reduced to the size of 21. Each stroke class adds three elements to the vector: The partial derivations of its transition probability to the next stroke and the two transitions to the absorbing point and error states. This adds up to 15 elements for the first five strokes of a player. The sixth stroke has to be split up in two classes, since it can be reached from both starting points of the simulation. This results in two subclasses for the sixth stroke: The sixth stroke after one’s own serve, and the sixth stroke after the serve of the opponent.

The separation of simulation for each player sounds like a contradiction to the original goal of the Markov model, to include the interaction process between players into the performance analysis. However, the interaction between the players is not omitted, but shifted so that only a subset of the original states is considered in the simulation.

The gradients for each game were analyzed using a factor analysis. The factor extraction was done by principal component analysis and Varimax rotation with Kaiser Normalization.

Results

Comparison between the methods

To validate the new simulation method, the partial derivation values for each game were compared with the values produced by the original method introduced by Lames (1991). Therefore the simulation was performed with the classical model (difficulty function) for the change in tactical behavior for \hat{h} .

$$\hat{h}(x) = K + B * 4 * x * (1 - x)$$

The constant K describes the amplitude of change in the edge case of the simulation, for example when the transition probability is zero, and was chosen with a value of 0.01. The other constant, B , determines the maximum amplitude of change and was chosen with the value of 0.05 in order to achieve the maximum change of 6% in transition probability, which was found to be stable in terms of its behavior by Lames (1991) and Pfeiffer (2005).

Since the original method works with the absolute change in the winning probability, in contrast to the slope generated by the Markov simulation, a direct comparison of both methods is difficult, however. The correlation between the derivation values and the absolute results showed a mean value of 0.457. A better way to compare both methods is to convert the absolute differences in the winning probability from the original method to slope values, which makes both methods easily comparable. This is done by dividing the results by the absolute amount of change $|\hat{h}|$ for each simulated value. The mean correlation between the original and the new method for the slope values then becomes 0.995, with a mean absolute difference of 0.022 between the values for each stroke.

Error and Winner rates

The impact of error and winner rates was simulated by the application of partial derivations to the winning probability function. A three-factor variance analysis with repeated measurements with the derivation direction (point/error) and the stroke number as within-subject effects and the gender as between-subject effect was used to analyze the resulting derivatives. Since the data did not comply with the sphericity criteria, the Greenhouse-Geisser corrected values had to be considered. The analysis shows significant results for both the stroke number and the interaction of stroke number with gender and point/error probability (table 1). The biggest share on the variance was found for the stroke number, with a partial eta-square value of .441, followed by the combinations of stroke number with gender and point. Between-subject tests indicate that the variance can also be explained by the gender with a significance of $p < .001$ and a partial eta square of .119 (table 2).

Table 1. 3-factor variance analysis, within-subject tests (Greenhouse-Geisser)

Source	Square sum	df	Mean squares	F	Sig	Partial eta ²
Point	.040	1.000	.040	1.275	.259	.002
Point * Gender	.000	1.000	.000	.006	.936	.000
Stroke Nr.	54.608	1.051	51.969	406.478	.000	.441
Stroke Nr. * Gender	9.355	1.051	8.903	69.638	.000	.119
Point * Stroke Nr.	.300	1.699	.176	15.608	.000	.029
Point * Stroke Nr. * Gender	.028	1.699	.017	1.467	.230	.003

Lastly, figure 3 compares the absolute mean values of the derivation of the point and error probability between stroke number and gender. The long rallies with more than five strokes show simulation values with an exceedingly higher absolute derivative than the rest of the strokes. With the absolute amount of change as important indicator for tactical behavior, the

long rallies can be considered as the most impacting rallies in a game.

Table 2. 3-factor variance analysis, between-subject tests

Source	Square sum	df	Mean squares	F	Sig	Partial eta ²
Constant term	399.568	1	399.568	11782.146	.000	.958
Gender	2.358	1	2.358	69.544	.000	.119
Error	17.499	516	.034			

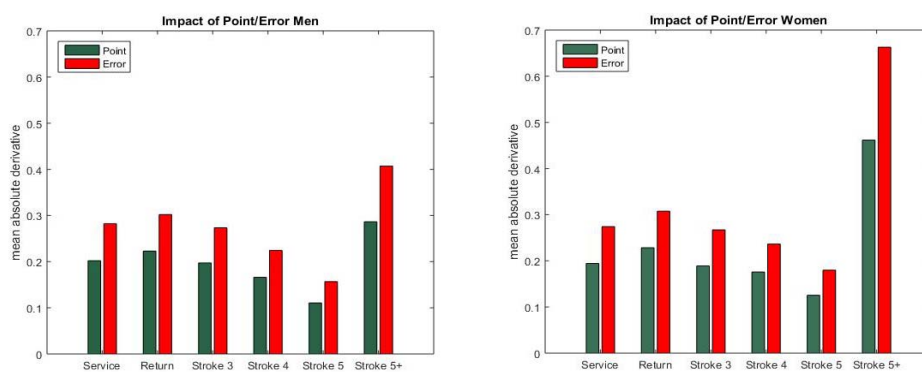


Figure 3. Impact of Point and Error probability per stroke and gender on the winning probability of a player. Absolute slope values are shown for better comparison of stroke types.

Gradient

This study used the gradient to identify the most important strokes for a player in terms of absolute impact on the winning probability. The factor analysis of the gradient clusters the elements of the transitions matrices into groups to find independent factors that explain the variance in stroke impact. The resulting groups of strokes with similar importance then represent the main clusters of performance in a table tennis game.

The scree plot for the computed factor analysis shows reason to differentiate four main components. However, these components are only able to explain 58.5% of the total variance, with the first component explaining 24% of the variance after rotation, the second component taking 16.8% and the last two components 9.6% and 8% respectively.

Table 3. Total Variance Explained

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6,148	29,277	29,277	5,050	24,048	24,048
2	2,641	12,575	41,852	3,528	16,800	40,849
3	1,973	9,397	51,248	2,018	9,608	50,457
4	1,533	7,299	58,547	1,699	8,090	58,547

Table 4 displays the rotated factor loadings for all factors with a correlation higher than .30. The first component features a high correlation with the long stroke classes, where the rallies last 6 strokes or longer. Component three is also clearly correlated to the service, while the high correlations of component four can be attributed to the return strokes. Finally, with the exception of the factor “Point 3.”, the second component exclusively contains noticeable correlations for error and transition probabilities.

Table 4. Rotated Component Matrix

	Component			
	1	2	3	4
Serve			,956	
Point Serve				
Error Serve			-,943	
Return				,895
Point Return				-,873
Error Return	,426	-,326		
3. Stroke		-,631		
Point 3.		,459		
Error 3.		,635		
4. Stroke		,680		
Point 4.				
Error 4.		-,747		
5. Stroke	-,321	-,510	-,302	
Point 5.				
Error 5.	,429	,605		
6 Opp. Serve	,818	,366		
Point 6 Opp. Serve	,873			
Error 6 Opp. Serve	-,837	-,411		
6 Own Serve	-,853	-,354		
Point 6 Own Serve	-,890			
Error 6 Own Serve	,860	,375		

Risk-taking

Directional derivatives were conducted to examine the importance of the tactical behaviors “risky play” and “safe play”. The results show a substantially higher amount of positive derivation values for the losing players, while winning players exhibit mostly negative values. Table 5 shows the results of the paired t-test for dependent samples for each stroke.

Table 5. T-Test between Winner and Loser for all strokes

Stroke	Mean	Std.dev	T	df	Sig(2-sided)	r	Cohen's d
Total	-.154	.201	-30.25	1553	.000	-.704	-1.417
1	-.159	.113	-22.531	258	.000	-.164	-2.136
2	-.154	.115	-21.566	258	.000	-.216	-2.090
3	-.137	.117	-18.747	258	.000	-.339	-1.906
4	-.111	.121	-14.927	258	.000	-.571	-1.645
5	-.085	.102	-13.482	258	.000	-.864	-1.617
5+	-.276	.395	-11.265	258	.000	-.940	-1.379

The data shows a significant difference between the means of the derivatives of the winning and losing player for each stroke. To compute the effect sizes for the dependent tests, the formula given by Dunlap, Cortina, Vaslow and Burke (1996) was used. All effect sizes indicate a large effect of the difference between winner and loser, although the effect strength diminishes with growing stroke number. Furthermore the sign of the correlation r validates the assumption of an inverse relation between the winning and the losing player.

Since only the impact of single behaviors was examined in previous sections, a multiple regression analysis was performed to expose any relation between the change in the winning probability (D5+) and risky/safe play (E5+, P5+) for rallies lasting longer than five strokes (tables 4 and 5). For the regression the point and error probabilities were chosen as predictors, and the derivative from the simulation as the dependent variable.

Table 6. Regression Summary

R	R ²	Adjusted R ²	Std.Error of the estimate
.444	.197	.194	.219
Predictors: (Constant), E5+, P5+			
Dependent Variable: D5+			

The regression summary reveals a rather small value with only 19.4% of the variance being defined by the predictors. The generated model can be trusted with values of $F = 63.230$ and $p < 0.01$ in an analysis of variance, however.

Table 7 reveals the impact of each predictor on the change in the winning probability.

Table 7. Regression coefficients

Model	B	Std.Error	Beta	T	Sig
(Constant)	-.278	.027		-10.268	.000
P5+	.197	-.758	-.121	-3.069	.002
E5+	1.021	.095	.425	10.756	.000

Both point and error probability have a significant effect on the winning probability. In comparison, the error probability $Beta = .425$ contributes almost four times as much to the explanation of the variance as the point probability $Beta = .121$.

Discussion

In general, the simulation using numerical derivation can be established as an effective alternative to the traditional simulation method. The criticism regarding the original method, especially the Markov-chain property, can still be cause for concern, since it demands stability of the transition probabilities regardless of the time the transition takes place within a rally or game. With the transitions set to the different stroke numbers, this point in time is fixed at least over the span of a rally for each transition in our Markov model.

As the comparison between both methods points out, the resulting values are highly correlating ($r = 0.995$) when brought onto the same scale, thus reinforcing the usability of the numerical method. This scale leads to one of the difficulties of the developed method. Unlike the traditional simulation, which used absolute differences in the winning probability as indicator for the performance, the derivation produces slopes that may be hard to interpret in the context of a table tennis game at first. In addition, the numerical simulation is susceptible to errors like rounding and machine imprecision. However, this is a well studied problem in numerics and can be minimized by the careful selection of the step size for the derivation according to Press et al. (2007). Ultimately, the stable properties of the simulation allow the application of several mathematical constructs, such as the partial derivation, directional derivation or the gradient.

The *partial derivation* analysis reveals significant differences between gender, stroke numbers, in regards to points and errors, with the higher influence of errors compared to the direct point rate especially apparent. The difference between the stroke numbers confirm that even though the model itself is rather simple, with only the length and result of rallies included and finer details as for example the technique of each stroke ignored, significant findings about the individual stroke lengths are possible. The high impact of long rallies with more than five strokes, as well as the higher impact of errors in comparison to direct points are notable findings of the analysis. While the partial derivation focuses solely on the influence of the point and error rates, the *gradient analysis* includes information about the relevance of all reachable transitions of the Markov simulation and hence provides an insight into the general structure of performance in a table tennis game. Four main components that contribute to the winning probability of a player were revealed by the factor analysis of the gradient. These components represent independent factors of performance and can give an indication, on which strokes to focus in more detailed analyses. In compliance with the findings of the partial derivation, the long rallies lasting more than five strokes were identified as the most important independent class of strokes during a match. The second component almost exclusively contains transient transitions (transitions to the next stroke), as well as transitions to the error states (direct errors). Because these behaviors are contradicting in the tactical context of a table

tennis game, the extraction of a general term that includes both of these stroke classes is quite difficult. One way of classifying these strokes could be to formulate a reverse definition and label them as the strokes, where no direct point is made. On the contrary, the third and fourth components are clearly related to the serve and the return strokes again.

The results of the *directional derivation* analysis show that the impact of the behavior “risky play” depends on the evaluated player. While risky play mostly results in a disadvantage for the winning player, losing players can benefit from a more risky playstyle. A possible explanation for this result could be that winning players already perform on a near optimal risk/reward level, while losing players still have room to improve in that regard. Again, although both error and direct point rates are found to have significant impact, the error rate contributes with an extraordinary share to the result.

Since all of the conducted analyses determine the *error rate* as highly impacting factor on the winning probability of a player and hence the result of a table tennis game, the prevention of errors, especially at highly impacting stroke types like the long rallies, appears to be a desirable goal for high-level players. Pfeiffer et al. (2010) find a similar effect in their study, where the influence of direct points is found smaller than the error rate. They explain this behavior by the outstanding defensive performance of world-class players, which induces a greater willingness to take risks. This outstanding defensive performance could also be the reason for the high impact of long rallies that we found in the partial derivation as well as in the gradient vectors. Since the possibility to score points early in a rally against world-class defensive players is relatively low, rallies tend to last longer and hence the performance in the long rallies becomes more important. This is further supported by the components of the gradient analysis, where the serve and return strokes showed the smallest contribution to the variance. Another explanation is given by Malagoli-Lanzoni, Di Michele and Merni (2014), who found that the stroke types that are mostly used at the beginning of a rally (push and flick) show a rather high proportion of returns. Again, this means that rallies tend to last longer and the performance in these long rallies has greater impact on the outcome of a game. They also discovered that aggressive stroke types like the top forehand and top-counter-top forehand, which are considered risky stroke techniques, are more associated to rally winners than the other strokes. This supports our assumption that the winners operate on a better risk/reward level than the losers of a game.

In general, the simulation using numerical derivation brings the advantage that a second modeling step to determine the difficulty of tactical behaviours can be skipped. This removes the need of game-specific knowledge to apply the simulation method on different sports. In fact, the simulation itself can be seen as a closed black-box, which gets the transition matrices as input, computes the desired numerical derivation on the matrices and returns the slopes of the winning probability as output, regardless of sport or behavior to examine. Naturally, this means that the resulting analyses are limited by the information contained in the input matrices, as their variables determine which behaviors can be examined and to what extent. Considering that the input data in this study is relatively simple, with only the rally lengths, server and winner known, it is possible to draw conclusions about the general structure of a table tennis game. Some interesting questions that are relevant in table tennis practice, like the impact of transitions between stroke techniques or stroke positions, as examined in other studies (Zhang, 2006; Pfeiffer et al., 2010; Malagoli-Lanzoni et al, 2014), can not be answered with the available data. However, since the simulation method itself is independent from the information contained in the input data, it would be possible to examine these features if it were included in the dataset.

Especially if you see the mathematical constructs we presented as a framework to automatically evaluate different sports on a mathematical basis without the need to adapt the simulation, this independence could prove to be advantageous. Given an arbitrary set of transition matrices, it would be possible to apply the gradient analysis on that set to automatically identify the most important transitions, then run consecutive simulation passes with more sophisticated partial or directional derivations on these transitions, resulting in a fully automatic analysis of the most important game behaviors. Future work on the method could also implement an improved way to compute the winning probability by including the actual distribution of services between both players. Additionally, the gradient vectors could be used for a cluster analyses on a per game basis to find similar game types, players or even match-ups. Finally, the differences between winning and losing players encourage further studies.

Conclusion

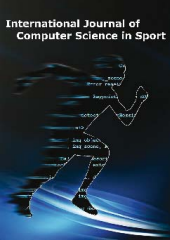
In the present study, stochastic simulation using a Markov chain model of table tennis and numerical derivation was used for the first time. In comparison with the traditional methods, it retains the advantages of the original Markov simulation, but skips the necessity to perform a second modeling step at the same time. This enables the application of the simulation in different fields without the need to adapt it to the difficulty of the sports-specific tactical behaviors, which makes it more approachable. The use of the numerical derivative also allows the employment of mathematical constructs that are based on derivation. The extension of the Markov simulation through numerical derivation hence is a profitable procedure in table tennis, allowing the computation and interpretation of the relevance of tactical behavior patterns.

References

- Baca, A., Baron, R., Leser, R., & Kain, H. (2004). A process oriented approach for match analysis in table tennis. In A. Lees, J. F. Kahn & I. W. Maynard (Eds.), *Science and racket sports III* (pp. 214-219). London: Routledge.
- Dunlap, W., Cortina, J., Vaslow, J., & Burke, M. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1(2), 170–177.
- Hughes, M. (1998). The application of notational analysis to racket sports. In A. Lees, I. Maynard, M. Hughes & T. Reilly (Eds.), *Science and Racket Sports II* (pp. 211-220). London: E & FN Spon.
- Hughes, M., & Bartlett, R. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences*, 20, 739-754.
- Juan, Y.L., Zhang, H., & Hu, J.J. (2008). Computer diagnostics for the analysis of table tennis matches. *International Journal of Sport Science and Engineering*, 2, 144-153.
- Kaplan, W. (1991). *Advanced Calculus* (4th ed). Reading, MA: Addison-Wesley.
- Lames, M. (1991). *Leistungsdiagnostik durch Computersimulation*. [Performance Diagnosis by Computer Simulation]. Frankfurt/Main: Harri Deutsch.
- Lames, M., Hohmann, A., Daum, M., Dierks, B., Fröhner, B., Seidel, I., & Wichmann, E. (1997). Top oder Flop: Die Erfassung der Spielleistung in den Mannschaftssportspielen. [Top or Flop: Measuring Performance in Team Games.] In E. Hossner & K. Roth (Eds.), *Sport-Spiel-Forschung Zwischen Trainerbank und Lehrstuhl* (pp. 101-117). Hamburg: Czwalina.
- Lames, M., & McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport*, 7(1), 62-79.

- Lees, A. (2003). Science and the major racket sports: a review. *Journal of Sports Sciences*, 21, 707-732.
- Leser, R., & Baca, A. (2009). Practice oriented match analyses in table tennis as a coaching aid. In A. Lees, D. Cabello & G. Torres (Eds.), *Science and Racket Sports IV* (pp. 214-219). London: Routledge.
- Malagoli-Lanzoni, I., Di Michele, R., & Merni, F. (2011). Performance indicators in table tennis: a review of the literature. *International Journal of Table Tennis Sciences*, 7, 71-75.
- Malagoli-Lanzoni, I., Di Michele, R., & Merni, F. (2014). A notational analysis of shot characteristics in top-level table tennis players, *European Journal of Sport Science*, 14(4), 309-317.
- McGarry, T., & Franks, I. M. (1994). A Stochastic Approach to Predicting Competition Squash Match-Play, *Journal of Sport Sciences*, 12, 573-584.
- Newton, P. K., & Aslam, K. (2009). Monte Carlo Tennis: A Stochastic Markov Chain Model, *Journal of Quantitative Analysis in Sports*, 5(3), 1-42.
- O'Donoghue, P. (2004). Match analysis in racket sports. In A. Lees, J. F. Kahn & I. W. Maynard (Eds.), *Science and Racket Sports III* (pp. 155-162). London: Routledge.
- O'Donoghue, P. (2008). Principal components analysis in the selection of key performance indicators in sport. *International Journal of Performance Analysis of Sport*, 8, 145-155.
- Pfeiffer, M. (2005). *Leistungsdiagnostik im Nachwuchstraining der Sportspiele* [Performance Diagnosis in the Game Sport Training of Young Athletes]. Köln: Sportbuch Strauß.
- Pfeiffer, M., & Hohmann, A. (2008). Leistungsdiagnostik mittels Systematischer Spielbeobachtung bei der FIFA WM 2006™ [Performance Diagnostic by Means of Systematical Game Observation at the FIFA World Championship™]. In M. Letzelter & H. Letzelter (Eds.), *Beiträge zur Theorie der Sportspiele*. Niedernhausen: Schors-Verlag, 65-98.
- Pfeiffer, M., Zhang H., & Hohmann A. (2010). A Markov Chain Model of Elite Table Tennis Competition. *International Journal of Sports Science and Coaching*, 5(2): 205-222.
- Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (2007). *Numerical Recipes: The Art of Scientific Computing* (3rd ed). Cambridge University Press.
- Strauss, D., & Arnold, B. C. (1987). The Rating of Players in Racquetball Tournaments, *Applied Statistics*, 36, 163-173.
- Zhang, H. (2006). *Leistungsdiagnostik im Tischtennis* [Performance Diagnosis in Table Tennis], Hamburg: Dr. Kovac.

A.3. Studie 2 - Originalveröffentlichung

<p>International Journal of Computer Science in Sport</p> <p>Volume 18, Issue 2 (Special Edition), 2019 Journal homepage: http://iacss.org/index.php?id=30</p> <p>DOI: 10.2478/ijcss-2019-0010</p>	
---	---

Data Mining in Elite Beach Volleyball – Detecting Tactical Patterns Using Market Basket Analysis

Sebastian Wenninger, Daniel Link, Martin Lames

Technical University of Munich

Abstract

Sports coaches today have access to a growing amount of information that describes the performance of their players. Methods such as data mining have become increasingly useful tools to deal with the analytical demands of these high volumes of data. In this paper, we present a sports data mining approach using a combination of sequential association rule mining and clustering to extract useful information from a database of more than 400 high level beach volleyball games gathered at FIVB events in the years from 2013 to 2016 for both men and women. We regard each rally as a sequence of transactions including the tactical behaviours of the players. Use cases of our approach are shown by its application on the aggregated data for both genders and by analyzing the sequential patterns of a single player. Results indicate that sequential rule mining in conjunction with clustering can be a useful tool to reveal interesting patterns in beach volleyball performance data.

KEYWORDS: DATA MINING, BEACH VOLLEYBALL, ASSOCIATION RULES, PERFORMANCE ANALYSIS

Introduction

In the last decade, performance analysis has gained significant importance as a means to support the decision making process of coaches and athletes of any level. At the same time, technological advances enabled the collection of vast amounts of performance data like video, match-events, or spatio-temporal data from electronic performance and tracking systems (EPTS) (Link, 2018). In order to assist humans in extracting useful information from this growing amount of digital data, the field of data mining and knowledge discovery is concerned with the development of methods and techniques for making sense of large amounts of data (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). By definition, data mining is an exploratory process to extract previously unknown information about patterns and regularities from archived or streaming data. It covers a wide arrange of methods from clustering to classification and relationship modeling (e.g. regression, neural networks) (Ofoghi, Zeleznikow, MacMahon, & Raab, 2013). Although traditionally, performance analysis in sports is based on expert knowledge and statistical analysis of performance indicators, data mining techniques have become increasingly popular in the field. A prime example is the software Advanced Scout (Bhandari et al., 1997), which uses a technique called attribute focusing to discover interesting events in basketball data. With this technique, events are tagged as interesting if the subset of attributes that define the event has a characteristically different distribution compared to an overall distribution for that event. Borrie, Jonsson and Magnusson (2002) use a method called T-Pattern detection to find similar sequences of passes from soccer games, while Bialkowski et al. (2014) cluster spatio-temporal data on players' positions during a game to find players' roles over time. In another approach, Cintia, Rinzivillo and Pappalardo (2015) model passing behaviour in soccer games as networks to predict the outcomes of soccer games.

There also exist applications in volleyball and beach volleyball. Sheng (2013) and Zhang, Zhao, and Wu (2006) employ Markov processes to analyse the tactical structure of volleyball games. Kang, Huh, and Choi (2015) on the other hand use algorithms from social network analysis and text mining to extract core key-words related with volleyball game performance. In another approach, Van Haaren, Ben Shitrit, Davis, and Fua (2016) utilize a relational learning technique called inductive logic programming to automatically detect play patterns in high level volleyball data. They present the top ranked, as well as the most distinguishing offensive patterns per team and compare playing patterns between men and women. Results show that attacks from outside positions are among the most successful offensive patterns. Furthermore, women's teams show more attack actions in the same number of rallies, hinting at a faster pace in men's volleyball, which makes it harder to gain control of the ball after an attack from the opponent. Because of the wide array of methods and applications that data mining encompasses, we would like to forward interested readers to Ofoghi et al. (2013), as well as Schumaker, Solieman, and Chen (2010) for an additional overview of data mining techniques applied to sports data.

A specific technique of data mining is the so called frequent itemset mining and its extension, association rule mining. Association rule mining attempts to find common patterns of items in large data sets (Agrawal & Srikant, 1994). While this method was first designed to help in the specific application of market basket analysis, where data mostly consists of groups of items that appear frequently together in transactions made by customers, this type of analysis has found applications in many other fields (Fournier-Viger et al., 2017). In fact, a customer transaction database can be generalized as a database of instances describing objects (the transactions), in which each object is described using nominal attribute values (the items). The task of frequent itemset mining then becomes the search for attribute values that frequently co-occur in a database. This allows the application of frequent itemset mining techniques in

multiple domains such as bioinformatics (Naulaerts et al., 2015), image classification (Fernando, Fromont, & Tuytelaars, 2012), network traffic analysis (Brauckhoff, Dimitropoulos, Wagner, & Salamatian, 2012), recommendation systems (Cakir & Aras, 2012) and even tourism (Bermingham & Lee, 2014), to name some examples. In order to address specific needs from different domains, the original algorithms were also improved and extended to discover correlated patterns (Fournier-Viger, Lin, Dinh, & Le, 2016a), patterns that generate a high profit (Liu, Liao, & Choudhary, 2005), or patterns in sequences (Mabroukeh & Ezeife, 2010) and graphs (Inokuchi, Washio, & Motoda, 2000).

Although the general nature of this data mining method makes it a good candidate for the application on sports data, it has not attracted much attention in the sports domain. Stöckl and Morgan (2013) use association rule mining in combination with the ISOPAR visualization method to investigate the characteristics of ball possession, passing, and attacking behavior of world-class hockey teams, revealing that teams tend to carry the ball on the right side of the pitch and that ball possession that resulted in goal shots were more likely to be neutral or left sided. Raj and Padma (2013) evaluate the influence of the variables venue (home/away), outcome of the coin toss, inning order (batting first/second) and opponent on the result of cricket games for the Indian national team over time from 1974 to 2010. In a similar work, Sun, Yu, and Zhao (2010) introduce a general framework to use association rules for the analysis of technical actions in ball games (e.g. volleyball, table tennis or basketball). They apply their framework on an example table tennis dataset and find actionable rules for individual players, defining the large amount of unreasonable association rules as one of the big remaining challenges.

Even though there are numerous complex ways to analyze data today, (clustering, regression, neural networks, random forests, support vector machines, etc.) the challenge with many of these approaches is that they can be hard to tune, difficult to interpret and require quite a bit of data preparation and feature engineering to get good results. On the other hand, association analysis is relatively light on the math concepts and easy to explain to non-technical people. In addition, it is an unsupervised learning tool that looks for hidden patterns so there is limited need for data preparation and feature engineering. This is important for the coaches and scouts, who most likely are not proficient in computer sciences and are expecting results that are easy to interpret and to communicate to the athletes. Secondly, it is widely used and easy to implement, which, in case the results prove worthwhile, may allow the inclusion of association rule algorithms in the existing scouting software. At last, the extension of association rules to cover sequential patterns fits the structure of beach volleyball performance quite well. We will describe the mapping from sequential transactions to beach volleyball rallies in more detail in the next section.

The aim of this study is to evaluate association rule mining as a tool to search for interesting, previously unknown tactical patterns in beach volleyball. In accordance with the nature of data mining, we take an explorative approach and apply our method in theoretical and practical performance analysis without any predetermined research questions. Lastly, the advantages and problems of this method are discussed together with possible enhancements and future research directions.

Methods

Sequential Rule Mining

A sequence database is a set of sequences in which each sequence is a list of itemsets (transactions). An itemset is an unordered set of items. A sequential rule $X \Rightarrow Y$ then is a

sequential relationship between two sets of items X and Y such that X and Y are disjoint and both X and Y are unordered. The left itemset of an association rule (X) is called antecedent, while the right one (Y) is called consequent. The support $sup(X \Rightarrow Y)$ of a rule $X \Rightarrow Y$ is the number of sequences in the database S that contain all items of X before all items from Y ($X < Y$) divided by the number of sequences in the database.

$$sup(X \Rightarrow Y) = \frac{|\{s \in S; X, Y \subseteq s, X < Y\}|}{|S|} \quad (1)$$

The confidence $conf(X \Rightarrow Y)$ of a rule is the number of sequences that contain all items of X before all items from Y , divided by the number of sequences that contains items in X .

$$conf(X \Rightarrow Y) = \frac{|\{s \in S; X, Y \subseteq s, X < Y\}|}{|\{z \in S; X \subseteq z\}|} \quad (2)$$

This means the confidence can be used as a measure for the predictive value of an association rule. As an example, the table shown below contains four sequences. The first sequence, named S1, contains 5 itemsets. It means that item 1 was followed by items 1 2 and 3 at the same time, which were followed by 1 and 3, followed by 4, and followed by 3 and 6. The rule $(1\ 4) \Rightarrow (3)$ then means that if 1 and 4 appear in any order, they will be followed by 3 with a confidence of 100%, since the item 3 appears after each occurrence of the itemset $\{1\ 4\}$. Moreover, this rule has a support of 75% because it appears in three (S1, S2 and S3) out of four sequences.

Table 1: Example sequence database

ID	Sequences
S1	(1), (1 2 3), (1 3), (4), (3 6)
S2	(1 4), (3), (2 3), (1 5)
S3	(5 6), (1 2), (4 6), (3), (2)
S4	(5), (7), (1 6), (3), (2), (3)

To measure the performance of rules, we employ the lift statistic, which compares the confidence of a rule $X \Rightarrow Y$ to the expected confidence of the rule in the database if X and Y were independent. The lift can be computed as follows (Tan, Kumar, & Srivastava, 2004):

$$lift(X \Rightarrow Y) = \frac{sup(X \Rightarrow Y)}{(sup(X) * sup(Y))} = \frac{conf(X \Rightarrow Y)}{sup(Y)} \quad (3)$$

In our example, the rule $(1\ 4) \Rightarrow 3$ has a lift value of 1, since the support of the item (3) is 100%. A lift value near 1 indicates that X and Y appear almost as often together as expected, this means that the occurrence of X has almost no effect on the occurrence Y . Hence no rule could be drawn involving those two events. A lift value greater than 1 indicates that X and Y appear more often together than expected, this means that the occurrence X has a positive effect on the occurrence of Y . A lift value smaller than 1 on the other hand means that the presence of X has a negative effect on the presence of Y .

Choice of Parameters

The sequential rule mining algorithm requires us to specify thresholds for the minimum support and confidence (values in $[0, 1]$) that the generated association rules need to satisfy. These thresholds influence the number and types of patterns that can be found by the algorithm. After an exploration phase we found this general relation: High support and

confidence thresholds provide a low number of higher quality (in term of the performance measures) rules, while low thresholds result in a high number of rules with lower performance measures. From a sports perspective, however, the rules with high support and confidence most of the time are the least interesting rules. They express the standard behaviours that happen frequently and are not new to the domain experts. Hence we choose these parameters according to the nature of patterns we want to discover for the different parts of our analysis:

- For the theoretical performance analysis (TPA), the analysis of the general structure of performance in beach volleyball, we choose a minimum support of 10% and a minimum confidence of 50%.
- For the analysis of individual players we choose a minimum support value of 5% and a minimum confidence of 50%.

Since TPA is concerned about the discovery of general relationships between performance indicators, we choose a higher support value to ensure that the patterns found by the algorithm appear in an adequate number of standard sideouts. This also reduces the time it takes to generate the sequential rules, because a large number of patterns that do not meet the support requirements can be filtered out beforehand. The individual player analysis, however, tries to find patterns that deviate from the standard behaviour, as obtained by the TPA. The lower support value for this type of analysis guarantees the inclusion of patterns that only happen in a small number of rallies, but may still be interesting from a performance analysis point of view. We also choose a relatively conservative confidence value of 50% in both cases to account for the variability top class beach volleyball players show in their tactical behaviour.

Modeling

The object of our analysis are the so called *standard sideouts*. Standard sideouts in beach volleyball are classified as the rallies, in which the receiving team has the chance to attack after a structured build-up (total 3 contacts of the ball). These sideouts are known to be one of the most influencing factors to the result of a game (Giatsis & Zahariadis, 2008). Since the athlete's actions, and hence the performance indicators in a beach volleyball rally, follow a natural temporal order (see Figure 1) in our observation system, sequential rules provide a well fitting model to the real world structure of beach volleyball. However, this also requires us to model rallies as a sequence of transactions. Given the sequential model as shown in figure 1 this can be done as follows:

- Generate a transaction for each unique timestep of a rally
- Add all indicators to the transaction of their corresponding time

This means that performance indicators that can appear at the same time according to our model will be added to the same transaction, and the sequences will contain exactly one transaction for each timestep in a rally. For example, the indicators *block position* and *block technique* will be added to the same transaction, while *attack zone* and *attack technique* will appear in separate itemsets. A beach volleyball game can then be modeled as a database of sequences, in which the sequences contain the actions for each player and rally.

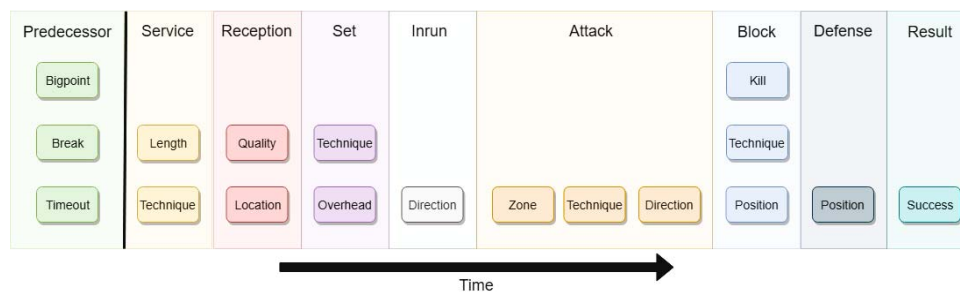


Figure 1. Sequential model of a beach volleyball rally, based on the variables employed. Performance indicators that can appear simultaneously in the temporal model are stacked on the vertical axis.

Sample

The dataset consists of 413 men's and 552 women's top-level games collected at FIVB world tour tournaments and championships in the years from 2013 to 2016. All data were annotated by professional beach volleyball analysts by using custom-made observation software (Link, 2014). Cohen's κ statistics showed substantial to perfect agreement between two observers based on a subset of 121 sequences ($\kappa = 0.93$ up to 1.0). For each rally in a game, the analysts collected more than 25 performance indicators. However, not all of those indicators are useful as input for sequential mining algorithms, since they only work with categorical data. This means that absolute positions (X and Y coordinates) for different actions (e.g. attack, defense) on the field have to be excluded from the analysis. However, these absolute positions are also included as categorical variables for some performance indicators (e.g. attack position and direction), which allows positional analysis nevertheless. In total, our database contains 25,918 standard sideouts (sequences) for the men, and 33,393 standard sideouts for the women. In order to evaluate our method in the context of practical performance analysis, we apply our analysis on a subset of our data for 47 games (1,451 rallies) of Alison Cerutti, the Brazilian 2016 olympic gold medalist.

Clustering

While association rules themselves already convey the information required to make educated decisions in the context of sports performance analysis, the typically large number of generated rules makes it necessary to perform some form of post processing. In order to find thematically connected groups of association rules, we apply hierarchical clustering on the resulting dataset of rules and then find representative rules per cluster to describe these groups (Jorge, 2004). We use the strategy of agglomerative clustering, a bottom-up approach which initially places each rule into its own cluster and then gradually merges pairs of clusters. In order to decide which clusters should be combined, a measure of dissimilarity between sets of rules is required. This is achieved by the use of a distance metric (between pairs of rules) in combination with a linkage criterion that defines the dissimilarity of sets as a function of the pairwise distances of rules in the sets. The choice of distance metric and linkage criterion has a strong influence on the shape of the clusters (Ashley, Kim, & Guo, 2005).

Commonly used distance metrics for clustering include the (squared) Euclidean distance, or Manhattan distance as some examples. However, since our association rules consist of non-numerical data (strings describing the performance indicators), we have to use a different metric. The Hamming distance describes the (dis-)similarity of two strings of equal length as the number of positions at which the corresponding symbols are different (Hamming, 1950). We define the hamming distance between two association rules $r1: X \Rightarrow Y$ and $r2: U \Rightarrow V$ to

be the average number of positions at which the performance indicator differ per rule side. This means we compute the Hamming distance between X and U , as well as between Y and V , by mapping each rule side to a vector that encodes the type of performance indicator (e.g. the attack direction) present in a rule side. We then compare the values of the performance indicators for these vectors to get the Hamming distances for the antecedent and the consequent of the two rules. The Hamming distances for both rule sides are then summed up and divided by two to get the average hamming distance over both sides (see formula 4). Note that we assign the right side of a rule a bigger weight, which clusters rules with similar consequents together. This emphasizes the effect of a rule and makes it easier for the analysts to browse rules by their outcome.

$$dist(r1,r2) = \frac{hamming(X,U) + 3 * hamming(Y,V)}{2} \quad (4)$$

Another possible way would be to map both rule sides onto one logical vector per rule and then compute the Hamming distance over those two vectors. However, our method produces thematically more consistent clusters, because the rule side a performance indicator appears in is taken into account. For the linkage criterion we choose the average linkage, which describes the distance between clusters as the mean pairwise distance of rules in those clusters as shown in formula 5.

$$\frac{1}{|A| \cdot |B|} \sum_{x \in A} \sum_{y \in B} dist(x,y) \quad (5)$$

Here A and B represent clusters of rules, while $dist(x,y)$ denotes our pairwise distance function.

Number of clusters

Unlike other popular clustering algorithms (e.g. k-means), hierarchical clustering has the advantage that the natural number of clusters in the data can be unknown beforehand. In order to determine the optimal number of clusters in a dataset, the hierarchy of clusters can be analysed using several methods. First, one can create a dendrogram of the cluster hierarchy and manually select the number of clusters to generate by setting the cut-off line at the appropriate level. This requires extensive knowledge about the data and the generated clusters, however.

As an alternative, the silhouette method (Rousseeuw, 1987) measures how closely the data in a cluster is matched within itself compared to how loosely it is matched to the data of the neighbouring cluster. Here, the neighbouring cluster is defined as the cluster whose average distance from the datum is the lowest. A silhouette value close to 1 implies the datum is in an appropriate cluster, while a silhouette close to -1 suggests that the datum is in the wrong cluster. The average silhouette value over all clusters then gives an estimation for the goodness of fit of the clustering solution. Figure 2 shows the average silhouette values for different cluster sizes for the men's data set. We chose the optimal number of clusters as the number where the highest gain in silhouette score appears (maximum first positive derivative with negative second derivative). According to those criteria, the number of clusters for the men's data was set to 13. The same analysis was done for the women's data (9 clusters), as well as for the individual players' data (10 clusters), and the number of clusters set accordingly.

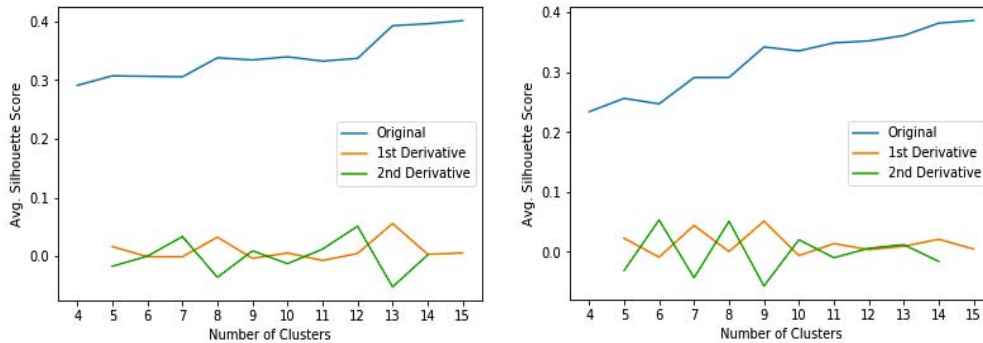


Figure 2: Average Silhouette values for different number of clusters in the men's (left) and women's (right) data set. The first and second derivative are used to determine the optimal cluster count, which is 13 for the men and 9 for the women.

Summarizing Clusters

To give users a rough description of the data inside each cluster, we try to find a subset of representative rules per cluster. This is done by choosing a representative rule $rep(R)$ of a rule set R as the rule r that has the minimal average distance to all other rules in that rule set (Jorge, 2004):

$$rep(R) = \underset{r \in R}{\operatorname{argmin}} [\operatorname{average}_{s \neq r \in R} \operatorname{dist}(r, s)] \quad (6)$$

If more than one rule minimizes the average distance, one of them is arbitrarily chosen. Of course, the representative rules can only give an overview and must not be mistaken for the most interesting rule in a cluster.

Tools

All analysis of association rules was done using the ERMiner algorithm (Fournier-Viger, Gueniche, Zida, & Tseng, 2014) implemented in the SPMF (Fournier-Viger et al., 2016b) software suite. The resulting rule files were then converted to the JSON (Bray, 2017) format and imported by the Python scikit-learn (Pedregosa et al., 2011) library for clustering and visualisation.

Abbreviations

Table 2: Abbreviations used in tables 3 to 6

Att	Attack	Loc	Location
Def	Defense	Pos	Position
Dir	Direction	Pred	Predecessor
JFloat	Jumpfloat	Qual	Quality
Len	Length	Tec	Technique

Results

Theoretical Performance Analysis

The ERMiner algorithm, with the parameters set as described in the previous section, generated 629 sequential association rules in 3.25 seconds for the men's data set. These association rules are clustered into 13 clusters with an average silhouette value of 0.393. For the women's data, 641 generated rules (6.64 s) are grouped into 9 clusters with an average silhouette score of 0.342.

Table 3: Description of rule clusters for men and women, including the number of contained rules (#), average support, confidence and lift per cluster. Some thematic groups (3 and 5) are also split over multiple clusters by our solution, producing multiple rows per group.

ID	Cluster Description	Men				Women			
		#	SUP	CONF	LIFT	#	SUP	CONF	LIFT
1	X => Att_Dir	8	.15	.51	2.12	6	.12	.51	2.11
2	X => Set_Tec_Forearm	28	.17	.61	1.59	54	.21	.76	1.04
3	X => Att_Pos_Left, Block_Pos_Line	69	.17	.77	2.73	136	.15	.79	3.27
		14	.12	.53	3.56	36	.14	.71	3.87
4	X => Att_Tec	77	.15	.58	1.56	67	.15	.53	1.70
5	Serve_Len_Short => Rec_Pos_Outside	20	.19	.71	3.18	32	.17	.76	3.46
		16	.16	.63	2.39				
6	X => Def_Pos_Line	127	.15	.85	3.59	164	.15	.86	4.03
7	X => Result	126	.14	.56	1.02	128	.15	.53	1.25
8	Serve_Len_Short => Rec_Qual_Good	24	.17	.63	2.56	18	.14	.55	3.38
9	Rec_Loc_Front => Inrun_Inside	5	.16	.56	1.33				
10	Pred_Break => Serve_Tec_JFloat	1	.19	.62	1.02				
11	X => Block_Pos_Line	114	.15	.76	3.15				

Clustering

Table 3 shows descriptions of the generated clusters per gender. These descriptions are generalized labels for the representative rule per cluster. According to our distance metric, the clusters mostly group association rules with similar consequents together. The clustering solution nicely captures the whole range of events present in beach volleyball rallies, starting from service and reception to setting technique, attack, defense and finally the result of a rally. Most of the clusters are present both for men and women, which is not surprising, since the general structure of play is the same for both genders.

For example, the rules in cluster 9 to 11, which are grouped into separate clusters for the men, are also present in the women's data. However, the clustering algorithm put them into cluster 2 for the women, since the consequents of these rules also contained the setting technique *Forearm Pass*. The opposite effect happened at clusters 3 and 5. The algorithm generated separate clusters for the same thematic group, because both the antecedents and the consequents of the contained rules were convoluted with additional performance indicators. To

discover the detailed differences in tactical behaviour between men and women, we have to look at the individual rules contained in these clusters, which we will do in the next section.

Rules

While the rule clusters already provide a rough overview over the performance patterns included in the data, they are mostly used to make the discovery of single interesting rules easier.

Table 4 depicts an excerpt of rules that the authors deemed interesting in the context of theoretical performance analysis. For both genders, the block position *Line* determines the defense position in the *Line* with a high confidence of over 90% and a lift value of 4.19. Next follow the rules involving the short services as a tactical measure as explained in the section before. Again, these rules show confidence values higher than 85% for the position of the reception, and more than 95% for the position of the attack in combination with high lift values.

Table 4: Selection of interesting rules contained in the clusters for men and women. Support and confidence are given as percentage values.

Rule	Men			Women		
	SUP	CONF	LIFT	SUP	CONF	LIFT
Block_Pos_Line => Def_Pos_Line	25.1	96.8	3.84	22.1	92.9	4.19
Serve_Len_Short => Rec_Pos_Outside	27.3	86.3	2.02	23.2	87.2	2.28
Serve_Len_Short, Rec_Pos_Outside => Att_Pos_Left	26.3	96.3	2.49	22.5	97.0	3.00
Serve_Len_Short => Rec_Qual_Good	23.1	73.0	1.02	17.5	65.9	1.02
Serve_Tec_JFloat => Rec_Qual_Good	45.1	74.2	1.03	47.1	63.4	0.98
Serve_Tec_Power => Rec_Qual_Good	19.8	64.5	0.90			
Pred_Break => Rec_Qual_Good	22.7	72.1	1.01	21.3	63.6	0.99
Rec_Qual_Bad => Result_NoSuccess	14.3	50.5	1.14	18.7	52.7	1.13
Rec_Qual_Good => Result_Success	41.9	58.4	1.04	36.6	56.8	1.06
Rec_Qual_Good, Att_Tec_Smash => Result_Success	25.9	61.6	1.10	18.0	58.3	1.09
Att_Tec_Smash => Result_Success	34.7	59.3	1.06	26.7	55.0	1.03
Att_Tec_Smash, Att_Dir_Diagonal => Result_Success	15.4	53.8	0.96	12.1	50.6	0.95
Att_Tec_Shot => Result_Success	20.9	51.5	0.92	26.5	52.2	0.98
Att_Tec_Shot => Att_Dir_Line	21.1	51.9	2.46			
Rec_Qual_Good => Set_Tec_Forearm	36.1	50.4	0.84	53.3	82.8	0.95
Rec_Qual_Bad => Set_Tec_Forearm	24.0	85.0	1.41	34.0	95.8	1.10

However, there exist noticeable differences in the effectiveness of different service techniques. The *Jumpfloat* serve, which is extensively used by both men (45.1% support) and women (47.1% support), leads to a good reception in 74.2% of the rallies for the men, while the women only achieve a good reception rate of 63.4%. It is also apparent that the preceeding rally has no influence on the reception performance of both male and female athletes. They achieve good receptions in 72.1% and 63.6% of the cases where rallies occurred after a break, which resembles the performance after *Jumpfloat* serves. The lift value of 1.02 confirms the weak relationship between those variables. Another recognizable pattern is the influence of the reception quality on the success of a sideout. The rallies in which the receiving player can

only produce a bad reception end with an unsuccessful result (no direct point in the sideout) in 50.5% of the cases for the men and 52.7% for the women. In turn, this means that they can achieve a successful result in less than 50% of the sideouts after a bad reception. After a good reception, on the other hand, both genders can achieve a much higher success rate of 58.4% and 56.8% respectively. Combined with the attacking technique *Smash* the success rate after a good reception further increases to 61.6% for the men and 58.3% for the women.

An 8 percentage points higher support for the rule *Attack_Technique_Smash => Result_Success* also indicates that men favor this attack technique compared to women, which is not surprising given that they are also more successful using it. Men perform a successful attack with the technique *Smash* with a confidence of 59.3%, while women are successful with this technique in only 55% of the cases. Interestingly, the success rate of both genders drops by around 5% if they use the technique *Smash* in combination with the attack direction *Diagonal*, which is also expressed by lift values slightly lower than 1. The success rate for the attack technique *Shot* shows to be on the same level for men (51.5%) and women (52.2%), but considerably lower compared to the technique *Smash*, which can be confirmed by the lift values. We can also see that men use the attack direction *Line* in more than 50% of the rallies when they use the technique *Shot*, while there is no such pattern found for the women. The last two patterns reveal the relationship between the reception quality and the setting technique *Forearm Pass*, which both genders preferably use after a bad reception, as seen by the higher confidence and lift values for this rule.

Player Profile

Compared to the theoretical performance analysis, practical performance analysis is more concerned with components and patterns of individual athlete's performance. Applied to the data of the player Alison, the ERMiner algorithm produced 2430 rules that were afterwards clustered into 10 clusters with an average silhouette value of 0.354. The higher amount of found rules can be attributed to the lower support threshold that was used for the analysis of players' data. Interestingly, the increased amount of rules does not result in an increased number of clusters.

Table 5: Description of rule clusters for Alison, including the number of contained rules (#), as well as average support, confidence and lift per cluster. The column "GID" contains the ID of the corresponding cluster in the general data set, if existent.

ID	Cluster Description	#	SUP	CONF	LIFT	GID
1	Att_Tec_Shot=> Att_Dir_Line	26	.09	.56	2.80	1
2	X => Rec_Qual_Good	51	.12	.63	2.85	8
3	Rec_Pos_Outside=> Attack_Pos_Left	252	.10	.71	3.42	3
4	Serve_Len_Short => Rec_Pos_Outside	100	.10	.67	3.72	5
5	X => Set_Tec_Bump	114	.10	.62	3.00	2
6	X => Att_Tec_Smash	271	.10	.66	2.32	4
7	X => Block_Pos_Line	453	.08	.78	4.42	11
8	X => Result_Success	587	.08	.59	0.99	7
9	X, Att_Tec_Smash => Att_Dir_Diagonal	66	.08	.51	2.32	1
10	X => Def_Pos_Line	510	.08	.85	4.87	6

In fact, the structure of clusters mostly stays the same, and the smaller clusters hardly gain an increase in rule numbers, with the exception of the cluster for the attack technique. The bigger

clusters see an extensive increase in rule numbers, however. This means that the lower support value results in the same core rules, but in a much higher number of combinations with different performance indicators that appear together. Table 5 shows an overview of the generated clusters. The rules generated for Alison are clustered in a similar way as the rules for the general men's data. Clusters 4, 5, 7 and 10 of Alison correspond to clusters 5, 2, 11 and 6 respectively. Clusters 1, 9, 2, 6 and 8 are more detailed versions of clusters 1, 8, 4 and 7 in the general data set, while cluster 3 of Alison seems to be included in cluster 3 of the men's data. Clusters 9 and 10 of the general data are not represented in the clusters for Alison, however.

The detailed tactical behaviours contained in the clusters are displayed in table 6. In agreement with the general pattern found for the men and women, short serves to the outside of the field are also used against Alison. These short serves do not influence the quality of the reception, as the rule *Serve_Len_Short => Rec_Qual_Good* with a confidence of 72.8 percent shows. Nonetheless, Alison achieves a higher success rate after the long serves (62.9% vs. 54.7%), even though his good reception rate is lower for that type of serve.

Compared to the general data, Alison has a considerably higher reception rate both for *Jumpfloat* (80.6%) and *Power* serves (72.9%). Of course this also influences his success rate, since a better reception results in a higher success rate as we have already seen in the men's data. After a good reception, Alison scores a point in 60.4% of the sideouts, which is slightly higher than the men's average (58.4%). However, even after a bad reception he achieves a success rate of 55.9% compared to a success rate of under 50% for the men's average.

The last group of rules reveals the success rate of different attack directions with the technique *Smash*. Even though the *Diagonal* direction has the highest support, which means is used most often by Alison, its success rate and lift is the lowest compared to the *Line* and the *Middle* direction. The lift even suggests that the diagonal direction is even detrimental to the success in a rally for Alison.

Discussion

Discussion of results

Given that the association rule algorithm was designed for an entirely different purpose, the results of its application are encouraging. The rules generated both for the aggregated data of both genders, as well as the individual player's data, are in accordance to the tactical patterns found by other studies. For example, Yianis (2008) confirms the significantly higher use of the smash attack by the men. Koch and Tilp (2009) also find that men prefer to use jump serves, the overhand pass as a setting technique and the smash in the attack, while women show a greater frequency of float serves and shots in the attack. Besides, our method finds a few additional patterns, that are not commonly explored in other work, like the use of short serves and their influence on the reception and attack position.

Clusters

It is already possible to make some deductions about the general structure of performance by only looking at the clusters for men and women. Cluster 2 apparently contains rules that determine the setting technique *forearm pass* both for men and women. The higher average support and confidence values show, however, that this technique is more often used by women, even though the lift indicates that the rules for this technique are more interesting for the men. This also indirectly explains the different number of generated clusters for men and women.

Cluster 5 also reveals another general pattern in beach volleyball. Teams try to impede the offensive options of the opponent by using short serves. For one, this limits the ability of the attacking player to use an extended inrun to generate jumping height for the attack. On the other hand, as seen in the description of cluster 5, short serves are also mostly played to the outside of the field, in order to limit the angle of attack for the receiving team. In this situation the defending teams seem use the boundaries of the field (the antenna) in combination with a block covering the line direction in order to cover as much area as possible with their defensive formation.

Table 6: Selection of interesting rules contained in the clusters for Alison. Support and confidence values are given in percent.

Rule	SUP	CONF	LIFT
Serve_Len_Short, Rec_Pos_Outside => Block_Pos_Line, Def_Pos_Line	19.4	92.2	4.74
Serve_Len_Short => Rec_Qual_Good	18.3	72.8	0.93
Serve_Len_Long => Rec_Qual_Good	5.4	67.2	0.86
Serve_Len_Short => Result_Success	13.7	54.7	0.92
Serve_Len_Long => Result_Success	5.0	62.9	1.05
Serve_Tec_JFloat => Rec_Qual_Good	49.7	80.6	1.03
Serve_Tec_Power => Rec_Qual_Good	24.9	72.9	0.93
Rec_Qual_Good => Result_Success	47.1	60.4	1.01
Rec_Qual_Bad => Result_Success	12.3	55.9	0.94
Att_Tec_Smash => Result_Success	41.8	62.4	1.05
Att_Tec_Shot => Result_Success	17.4	53.7	0.90
Rec_Qual_Good => Att_Tec_Smash	53.0	68.0	1.01
Rec_Qual_Bad => Att_Tec_Smash	13.9	63.1	0.94
Att_Tec_Smash, Att_Dir_Diagonal => Result_Success	18.4	55.1	0.92
Att_Tec_Smash, Att_Dir_Middle => Result_Success	8.6	70.2	1.18
Att_Tec_Smash, Att_Dir_Line => Result_Success	11.2	69.2	1.17

Looking at the clusters of the individual player's data, the more detailed clusters already give a hint of the behaviours Alison seems to favor compared to the men in general. These behaviours mostly concern the attack, where the technique *Smash* in combination with the direction diagonal seems to play a primary role. For the technique *Shot* on the other hand, the direction *Line* seems to be the favored behavior used by the player.

Finally it seems difficult to find relations between the performance indicators and the success of a rally for Alison, since the average lift value of cluster 8 lies close to 1.0, while the other clusters show lift values higher than 2.0. This difficulty to relate performance to success can also be seen in the individual rules in every dataset, since the lift values of these rules all lie between 0.9 and 1.2. This is not surprising and can be attributed to the variance of behaviours the high-level players use to score points in a rally.

Rules

The general pattern that the defensive player positions itself in the *Line* position when the block starts also in the *Line* conveys the widely used tactical pattern to “hide” the defensive player behind the blocking player, in order to give the attacking player as little information about the defensive formation as possible.

Additionally, we found patterns that employ short services as a tactical measure. In our opinion, they are used to hinder the inrun of the attacking player and, by playing the serve to the outside of the field, constrain the available attack angles from that position. The same pattern can be observed in Alison’s data, which is not surprising, since he is a player that relies on his physique and jumping height to achieve successful attacks, which is exactly what short serves are supposed to inhibit due to the limited inrun possibilities. It can also be seen that the short services are not used to directly cause bad receptions, since the rate of good receptions after short serves stands at 73% for the men and 65.9% for the women. This also holds for the rules of the individual player. In fact, long serves seem to more severely affect the reception of Alison, since both the lift and the proportion of good receptions in that case goes down to 0.86 and 67.2% respectively, even though this can still be considered a good reception rate. Surprisingly, Alison still achieves a higher success rate after the long serves, which could be caused by the fact that there is still enough room for the inrun to achieve his effective attack after a long serve. However, the lift values for the quality of the reception hint at a more complicated relation between service length and reception quality, even though these rules exhibit high confidence values.

Concerning the service technique, the fact that women are apparently much more effective with the use of the *Jumpfloat* serve than men may be explained by another pattern present in the data: Men use *Power* serves as an alternative to achieve higher effectiveness in the serve. *Power* serves are rarely used by women though and hence not present in the rule database, since they do not pass the support threshold of 10%. This means that women have to perfect the *Jumpfloat* serve in order to reach comparable levels of performance.

The missing influence of the predeceasing rally on the reception performance can be illustrated by looking at the raw sideout data. We can see that men choose the *Jumpfloat* serve in more than 60% of the rallies and women in almost 75% of all standard sideouts. The rule *Predecessor_Break => Reception_Quality_Good* is then actually caused by the intermediary variable serve technique and therefore also shows lift values close to 1.0. A good reception quality also plays a role in the success of a sideout. Obviously the structured build up that is possible after a good reception provides the best conditions for a successful attack. There also exists a connection between the reception quality and the setting technique *Forearm Pass*. Both genders preferably use this technique after a bad reception, most likely to avoid technical errors. Women also use this technique in a much higher proportion of cases both after good and bad receptions. This seems to be purely preference, since there is no apparent reason why they wouldn’t use other setting techniques.

The fact that men show higher success rates with the attack technique *Smash* is most likely due to physical advantages in height and strength, and can also be observed for the player Alison, who is highly effective with this technique due to his aforementioned physique. His rules with the antecedent attack technique and the consequent result success confirm this pattern. Alison’s success rate and lift with the smash is considerably higher compared to the shots (62.4% vs. 53.7%, 1.05 vs. 0.90), which suggests that shots actually have a negative effect on his success. And even after bad receptions Alison can use this technique in 63.1% of the rallies, which explains his higher success rate in this case. The drop in success for different attack directions with this technique in all datasets is most likely due to the standard defensive

formation used by most opponents. Especially the direction *Diagonal* shows lower success rates in this case. Possibly, this is the direction that the defensive formation primarily tries to cover. This needs further investigation, however. Surprisingly, Alison still achieves high success rates with a *Smash* to the *Line* and especially the *Middle* direction, since these are directions the block is supposed to cover in the defense.

On the other hand, the success of the attack technique *Shot* shows no differences between men and women. A possible explanation for this pattern is that shots have a higher chance to be defended against, since they are slower attacks that rely on the exact placement of the ball. While this technique is often used in combination with the direction *Line* by the men, women show no pattern regarding the attack direction. This hints at a greater variability of the women with that technique, which could be a consequence of their decreased success rate with the hard attacks (*Smash*).

Discussion of method

With the use of association rules, we can provide several measures for unusual combinations of performance indicators in the form of the confidence metric, which is especially interesting when a rule contains the success of a rally in its consequent. However, the large number of rules generated by the algorithm makes it necessary to employ some kind of post-processing in order to make the rule set approachable for scouts or coaches.

Clustering proves to be a worthwhile addition to association analysis for this case, since it breaks down the rule set into thematically consistent groups that can already give an indication about the broad structure of performance. We use a custom distance function to group rules mostly based on their consequent, which highlights the effect of tactical patterns. Of course, this can be changed based on the preferences of the scouts or coaches working with the tool. However, the addition of clustering adds another layer to the analysis toolchain that reduces the simplicity of the association analysis. Yet, the benefit of a structured result dataset surpasses this added complexity in our opinion. However, even the clusters can contain such a high number of rules that additional steps have to be taken to make analysis by a person feasible. These steps could be simple search algorithms like breadth-first search or depth-first search, after ordering the rules in a cluster by support or confidence, or more complex computations, like using the transitivity of rules to further compress the rule sets. There are many ways to post-process the rule set, as seen in Baesens, Viaene and Vanthienen (2000). The specific method to use ultimately depends on the structure of the rules and the question to be solved by the analysis. For example an interesting approach for future work could be to try and find clusters of players with similar emergent association rules.

Another option to make the rule set more approachable would be to develop a custom interestingness measure that rates rules based on their practical interestingness, since the standard measures of support and confidence hold no meaning in the context of the sport itself. A pattern is not necessarily interesting, just because it appears with a high frequency (high support). This would require expert knowledge, for example in form of an ontology, and a precise definition of “interestingness” though. For some analysis, the impact of a pattern on the success in the rally may be interesting, while for others the relative frequency of behaviours plays the deciding role. Clustering strikes a good balance here, as it groups rules without any kind of assessment regarding the interestingness. This imposes the decision of the interestingness of a rule on the user. In our case, this is not a problem, since the users - the scouts and coaches of the national teams - are of course experts in their field and hence able to decide which rules are interesting for the analysis at hand.

Judging the method from a performance analysis perspective, the association rules give a realistic picture of beach volleyball performance, even if the number of truly surprising rules is relatively low. This could be because of several reasons. First, beach volleyball is a highly structured and standardized sport, since there is no direct interaction between players as in hockey or soccer, for example. Interesting or surprising behavior then occurs when players have to deviate from their optimal tactics. However, because of the highly standardized structure of beach volleyball, players can freely choose their most efficient actions, which of course makes those the frequent patterns found by the algorithm. Additionally, the fact that we only analyze standard sideouts emphasizes this effect even more, since those are the most structured rallies present in this sport.

Second, it is possible that some of the interesting patterns were filtered out beforehand because of the parameters we used for the support and confidence thresholds. These thresholds were already set quite low, however, which makes this scenario unlikely. Since these values are parameters of the algorithm, they could be set to even lower values of course. This would further increase the number of rules found and in turn require to either improve the post-processing, or incur a higher manual workload for the users, because they would need to filter through bigger rule sets per cluster. Moreover, these parameters also influence the time the algorithm takes to generate the sequential rules. This means there exists a tradeoff between the number of rules that can be considered and the computing time, which is important given that this algorithm may be included in the scouting software used by the German national teams. For example, by halving the support parameter to 5% for the ERMiner algorithm on the women's data, the algorithm takes three times (from 6.64 to 18.11 seconds) longer than with a support of 10%.

In conclusion, sequential association rules, when used with the right post-processing techniques, provide a valuable tool to generate insights from high volumes of beach volleyball data. We use sequential association rules to find patterns within rallies of a beach volleyball game. Yet, our method could also be used with slight adjustments to find patterns between rallies instead. As we pointed out, the main challenge lies in finding the truly interesting rules in the high amount of generated patterns. We achieve this by breaking up the rule set into thematically ordered clusters, which are then manually evaluated. However, there is still room for improvement in the post-processing process, for example by eliminating redundant rules, in order to further automate the data mining process and to reduce work for the analysts. This is a non-trivial problem though, because one has to make sure that no interesting rules are eliminated. Adding more layers to the data mining process also removes its simplicity, which is one of the main advantages of the method. In case the association rules require a high amount of post-processing, other approaches like neural networks may provide better results with similar effort, given the database supports the use of such methods.

Acknowledgements

This study was supported by a grant from the German Federal Institute of Sports Science (BISp). We would also like to thank the scouts of the Olympic Training Center (OSP) Stuttgart, as well as the coaches of the national teams for their support in collecting the data.

References

- Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases, VLDB '94*, 487-499, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.
- Ashlock, D. A., Kim, E.Y., & Guo, L. (2005). Multi-clustering: avoiding the natural shape of underlying metrics. In C. H. Dagli et al. (Eds.), *Smart Engineering System Design: Vol.15. Neural Networks, Evolutionary Programming, and Artificial Life*, (pp. 453-461), ASME Press.
- Baesens, B., Viaene, S., & Vanthienen, J. (2000) Post-processing of association rules. At *The Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'2000)*. 20 - 23 Aug 2000.
- Birmingham, L., & Lee, I. (2014). Spatio-temporal sequential pattern mining for tourism sciences. *Procedia Computer Science*, 29, 379-389.
- Bhandari, I., Colet, E., Parker, J., Pines, Z., Pratap, R., & Ramanujam, K. K. (1997). Advanced scout: Data mining and knowledge discovery in NBA data. *Data Mining and Knowledge Discovery*, 1(1), 121-125.
- Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S., & Matthews, I. (2014). Large-Scale Analysis of Soccer Matches Using Spatiotemporal Tracking Data. In *2014 IEEE International Conference on Data Mining*, (pp. 725-730). IEEE.
- Borrie, A., Jonsson, G. K., & Magnusson, M. S. (2002). Temporal pattern analysis and its applicability in sport: An explanation and exemplar data. *Journal of Sports Sciences*, 10.
- Brauckhoff, D., Dimitropoulos, X., Wagner, A., & Salamatian, K. (2012). Anomaly extraction in backbone networks using association rules. *IEEE/ACM Transactions on Networking*, 20(6), 1788-1799.
- Bray, T. (2017). The JavaScript Object Notation (JSON) Data Interchange Format. RFC 8259, RFC Editor.
- Cakir, O., & Aras, M. E. (2012). A Recommendation Engine by Using Association Rules. *Procedia – Social and Behavioral Sciences*, 62, 452-456. World Conference on Business, Economics and Management (BEM-2012), May 4-6 2012, Antalya, Turkey.
- Cintia, P. U. d. P., Rinzivillo, S. I. N. R. C., & Pappalardo, L. U. d. P. (2015). A network-based approach to evaluate the performance of football teams. In *Machine Learning and Data Mining for Sports Analytics*.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37-54.
- Fernando, B., Fromont, E., & Tuytelaars, T. (2012). Effective use of frequent itemset mining for image classification. In Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., & Schmid, C. (Eds.), *Computer Vision - ECCV 2012*, (pp. 214-227), Berlin, Heidelberg. Springer Berlin Heidelberg.
- Fournier-Viger, P., & Tseng, V. S. (2011) Mining Top-K Sequential Rules. In *Proc. of the 7th Intern. Conf. on Advanced Data Mining and Applications (ADMA 2011)*, (pp. 180-194), Springer.
- Fournier-Viger P., Gueniche T., Zida S., & Tseng V.S. (2014) ERMiner: Sequential Rule Mining Using Equivalence Classes. In: Blockeel H., van Leeuwen M., Vinciotti V. (eds) *Advances in Intelligent Data Analysis XIII. IDA 2014. Lecture Notes in Computer Science, vol 8819*. Springer, Cham
- Fournier-Viger, P., Lin, J. C.-W., Dinh, T., & Le, H. B. (2016a). Mining correlated high-utility itemsets using the bond measure. In Martinez-Alvarez, F., Troncoso, A.,

- Quintian, H., & Corchado, E. (Eds.), *Hybrid Artificial Intelligent Systems*, (pp. 53-65)., Cham. Springer International Publishing.
- Fournier-Viger, P., Lin, J. C.-W., Gomariz, A., Gueniche, T., Soltani, A., Deng, Z., & Lam, H. T. (2016b). The spmf open-source data mining library version 2. In Berendt, B., Bringmann, B., Fromont, E., Garriga, G., Miettinen, P., Tatti, N., & Tresp, V. (Eds.), *Machine Learning and Knowledge Discovery in Databases*, (pp. 36-40)., Cham. Springer International Publishing.
- Fournier-Viger, P., Lin, J. C. W., Vo, B., Chi, T. T., Zhang, J., & Le, H. B. (2017). A survey of itemset mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(4), 1-41.
- Giatsis, G., & Zahariadis, P. (2008). Statistical analysis of men's fivb beach volleyball team performance. *International Journal of Performance Analysis in Sport*, 8(1), 31-43.
- Hamming, R. W. (1950). Error detecting and error correcting codes. *The Bell System Technical Journal*, 29(2), 147-160.
- Inokuchi, A., Washio, T., & Motoda, H. (2000). An Apriori-Based Algorithm for Mining Frequent Substructures from Graph Data. In Zighed, D. A., Komorowski, J., & Zytkow, J. (Eds.), *Principles of Data Mining and Knowledge Discovery*, (pp. 13-23)., Berlin, Heidelberg. Springer Berlin Heidelberg.
- Jorge, A. (2004). Hierarchical Clustering for thematic browsing and summarization of large sets of Association Rules. In *Proceedings of the 2004 SIAM International Conference on Data Mining*, (pp. 178-187).
- Kang, B., Huh, M., & Choi, S. (2015). Performance analysis of volleyball games using the social network and text mining techniques. *Journal of the Korean Data and Information Science Society*, 26(3), 619-630.
- Koch, C., & Tilp, M. (2009). Beach volleyball techniques and tactics: A comparison of male and female playing characteristics. *Kinesiology*, 41(1), 52–59.
- Link, D. (2014). A toolset for beach volleyball game analysis based on object tracking. *Int. J. Comp. Sci. Sport* 13, 24–35
- Link, D. (2018). *Data Analytics in Professional Soccer*. Springer Vieweg, Wiesbaden.
- Liu, Y., Liao, W.-k., & Choudhary, A. (2005). A two-phase algorithm for fast discovery of high utility itemsets. In Ho, T. B., Cheung, D., & Liu, H. (Eds.), *Advances in Knowledge Discovery and Data Mining*, (pp. 689-695)., Berlin, Heidelberg. Springer Berlin Heidelberg.
- Mabroukeh, N. R., & Ezeife, C. I. (2010). A taxonomy of sequential pattern mining algorithms. *ACM Computing Surveys*, 43(3), 1-41.
- Naulaerts, S., Meysman, P., Bittremieux, W., Vu, T. N., Berghe, W. V., Goethals, B., & Laukens, K. (2015). A primer to frequent itemset mining for bioinformatics. *Briefings in Bioinformatics*, 2, 216-231.
- Ofoghi, B., Zeleznikow, J., MacMahon, C., & Raab, M. (2013). Data Mining in Elite Sports: A Review and a Framework. *Measurement in Physical Education and Exercise Science*, 17(3), 171-186.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Raj, K. A. A. D., & Padma, P. (2013). Application of association rule mining: A case study on team india. In *2013 International Conference on Computer Communication and Informatics (ICCCI)*, (pp. 1-6). IEEE.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53 - 65.

- Schumaker, R. P., Solieman, O. K., & Chen, H. (2010). Sports knowledge management and data mining. *Annual Review of Information Science and Technology*, 44(1), 115-157.
- Sheng, L. (2013). Study of application of factors of volleyball game based on data mining. *Information Technology Journal*, 12(19), 5172-5176.
- Stöckl, M., & Morgan, S. (2013). Visualization and analysis of spatial characteristics of attacks in field hockey. *International Journal of Performance Analysis in Sport*, 13(1), 160-178.
- Sun, J., Yu, W., & Zhao, H. (2010). Study of association rule mining on technical action of ball games. *2010 International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2010*, 3, 539-542.
- Tan, P.-N., Kumar, V., & Srivastava, J. (2004). Selecting the right objective measure for association analysis. *Information Systems*, 29(4), 293-313.
- Van Haaren, J., Ben Shitrit, H., Davis, J., & Fua, P. (2016). Analyzing volleyball match data from the 2014 world championships using machine learning techniques. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, (pp. 627-634)., New York, NY, USA. ACM.
- Yiannis, L. (2008). Comparison of the basic characteristics of men's and women's beach volleyball from the Athens 2004 Olympics. *International Journal of Performance Analysis in Sport*, 8668, 8.
- Zhang, Y.-j., Zhao, H.-q., & Wu, J.-w. (2006). Research and application of data mining algorithm on technical-tactics analysis of volleyball matches. *Journal of Computer Applications*, 26(12), 3017-3029.

A.4. Studie 3 - Originalveröffentlichung



Performance of machine learning models in application to beach volleyball data.

Sebastian Wenninger, Daniel Link and Martin Lames

Technical University of Munich

Abstract

Driven by the increased availability of position and performance data, automated analyses are becoming the daily routine in many top-level sports. Methods from the domains of data mining and machine learning are more frequently used to generate new insights from massive amounts of data. This study evaluates the performance of four current models (multi-layer perceptron, convolutional network, recurrent network, gradient boosted tree) in classifying tactical behaviors on a beach volleyball dataset consisting of 1,356 top-level games. A three-way between-subjects analysis of variance was conducted to determine the effects of model, input features and target behavior on classification accuracy. Results show significant differences in classification accuracy between models as well as significant interaction effects between factors. Our models achieve classification performance similar to previous work in other sports. Nonetheless, they are not yet at the level to warrant practical application in day to day performance analysis in beach volleyball.

KEYWORDS: MACHINE LEARNING, SPORTS ANALYTICS, NEURAL NETWORKS, BEACH VOLLEYBALL

Introduction

Due to the large amount of available data and the resulting possibility to convert that data into useful knowledge, methods from artificial intelligence (AI) and data mining have become increasingly useful tools in many industry scenarios (Russel, & Norvig, 2010). This development also took place in the sports domain, where automated data analysis is a fast developing trend to assist the decision-making of practitioners (Link, 2018; Claudino et al, 2019).

However, predicting performance in beach volleyball is a subject that has not received much attention compared to other major sports. This may be due to the fact that teams in beach volleyball mostly operate for themselves, and data is not readily available. Unlike sports as football, hockey or basketball, where access to positional data from electronic position detection is common, beach volleyball only recently (at the world championships 2017 in Vienna) started to include sensor data to enhance media coverage. Additionally, the betting market, one of the biggest drivers for advances in performance prediction (Bunker, & Thabtah, 2019), mostly concentrates on other, financially stronger sports in order to optimize earnings.

The implementation and evaluation of different machine learning models on currently available performance data could serve as a first step to establish successful models from other application areas as viable methods of analysis. As Claudino et. al. (2019, p. 10) phrase it: Further evaluation research based on prospective methods is warranted to establish the predictive performance of specific AI techniques and methods. Additionally, the results of this study could inform future technological developments in the field of beach volleyball on the type of data and algorithms necessary for certain analyses.

Related Work

One of the earliest studies to consider AI in the analysis of sports performance was done by Lapham and Bartlett in 1995. They showed that the involvement of computers, with the help of methods from artificial intelligence, could be a rewarding future direction for the discipline. Since then, many different AI techniques were applied in a wide variety of sports, either to assess risk of injury (López-Valenciano et. Al., 2018), or to predict performance (Peterson, 2018).

In soccer, Perl & Memmert (2011) use DyCoN, a self-organizing neural network, to learn formation clusters from spatio-temporal data. These formation types are then fed to an analysis software to allow analysis of formation frequencies and interactions. Link and Hoernig (2017) employ a Bayesian network to classify ball control during individual ball possessions of soccer players with an accuracy of 96.7 percent. Bialkowski et al. (2015) learned the roles of players based on spatio-temporal data that describe their arrangement on the pitch. Using a minimum entropy model to partition the data into player roles, they showed that distinct formation classes can be discovered automatically. Spatio-temporal data was also used by Dick and Brefeld (2019) in combination with a deep convolutional network to automatically rate dangerous situations in soccer based on player positioning on the pitch.

The impact on spatio-temporal data on the possibilities of performance analysis can also be observed in basketball. Here it was used to compare self-organizing maps, a class of unsupervised neural networks, with a dynamical controlled network to classify three different pre-selected plays (Kempe, Grunz, & Memmert, 2015). Offensive play classification was also studied using deep recurrent neural networks by Wang and Zemel (2016), who state that the dynamic nature of team sports, especially the complexity of player interactions, makes this one of the hardest problems in sport analytics. Bianchi, Facchinetti and Zuccolotto (2017)

combine self-organizing maps with a fuzzy clustering algorithm to identify five new positional player roles in basketball based on seasonal player statistics data. On the other hand, Leicht, Gómez and Wood (2017) try to predict the match outcome in the Olympic basketball tournament using both a linear regression and a conditional inference classification tree. They conclude that even though the classification tree shows a slightly lower classification accuracy of 81.4% correctly classified matches, its ability to resolve non-linear phenomena offers greater practical utility.

The working group around Schrapf and Tilp (Hassan, Schrapf, Ramadan, & Tilp, 2017a; Hassan, Schrapf, & Tilp, 2017b; Schrapf, Alsaied, & Tilp, 2017) employed various types of neural networks (self-organizing maps, radial basis function network, dynamical controlled network) to analyze tactical interaction patterns, tactical training outcome and to predict shot positions on the field in team handball.

Of course, artificial intelligence methods also found applications in volleyball. Tümer and Koçer (2017) were able to successfully predict league standings of teams between seasons with 98 percent accuracy using a multi-layer perceptron and match results and match location (home/away) as input features to their model. In order to improve the training process, dynamical programming was used in conjunction with a k-nearest neighbour classifier to detect jumps in training data and choose appropriate intensity intervals in the training process (Vales-Alonso, Chaves-Dieguez, Lopez-Matencio, Alcaraz, Parrado-Garcia, & Gonzalez-Castano, 2015). To a similar purpose, Wang, Zhao, Chan and Li (2018) use inertial sensors placed on the wrists of volleyball players to assess their skill in the spiking technique. Their support vector machine was able to differentiate between elite, sub-elite and amateur players with an average accuracy of 94 percent. In another approach, Van Haaren, Ben Shitrit, Davis and Fua (2016) utilize a relational learning technique called inductive logic programming to automatically detect play patterns in high level volleyball data. They present the top ranked, as well as the most distinguishing offensive patterns per team and compare playing patterns between men and women. Results show that attacks from outside positions are among the most successful offensive patterns. Furthermore, women's teams show more attack actions in the same number of rallies, hinting at a faster pace in men's volleyball, which makes it harder to gain control of the ball after an attack from the opponent.

However, even though beach volleyball is at least as popular as indoor volleyball, with 425,000 cumulative visitors at the London Olympics 2012 ("Net gains - the evolution of beach volleyball", 2016), and arguably easier to analyze due to the lesser number of players on the court, AI methods have not attracted much interest. The only study we found used a deep convolutional neural network based on wearable sensors to automatically monitor player loads in beach volleyball (Kautz, Groh, Hannink, Jensen, Strubberg, & Eskofier, 2017).

The aim of this paper is to evaluate the application of AI methods on both positional as well as event data in beach volleyball. As far as the authors are aware, there exists no previous work to predict technical or tactical behaviors in beach volleyball. We will focus on three different neural network architectures together with a gradient boosted classification tree, some of the most used AI techniques employed to predict sporting performance (Claudino et. al., 2019), in order to cover a wider range of methods.

Methods

Data acquisition

The dataset consists of 569 men's and 787 women's top-level games collected at FIVB world tour tournaments and championships in the years from 2013 to 2018. All data was annotated

by professional beach volleyball analysts by using custom-made observation software (Link, 2014). Cohen's κ statistics showed substantial to perfect agreement between two observers based on a subset of 121 sequences ($\kappa = 0.93$ up to 1.0). For each rally in a game, the analysts collected more than 25 performance indicators (PIs) in addition to the discrete X and Y coordinates each for the serve, reception, set, approach and attack actions. These positions were tracked manually by the scouts through clicks on a calibrated projection of the beach volleyball field on the scouting video in the observation software. In total, our database contains 84,415 so-called standard sideouts. Standard sideouts in beach volleyball are classified as the rallies, in which the receiving team has the chance to attack after a structured build-up (total 3 contacts of the ball). These sideouts are known to be one of the most influencing factors to the result of a game (Zetou, Moustakidis, Tsigilis, & Komninakidou, 2007).

Data Preprocessing

Since not all games were analyzed in full detail by the analysts, we had to eliminate those rallies that did not contain valid X/Y coordinates. All X/Y coordinates were flipped so that the attacking position always appeared on the bottom half of the court and afterwards normalized to an area of 1 meter around the beach-volleyball court (see Fig.1) to account for positions outside the actual playing area (e.g. services). Additionally, we only used the PIs relevant to a single standard sideout from the collected data, leaving out general information such as tournament, match score and player information. We also did not include the defense position in the analysis, because for certain attacks, e.g. attack errors that were hit to the outside, this position also encoded the point where the ball hit the ground. Including this position would trivialize the prediction of success and attack direction. In total, this resulted in 14 PI and 10 positional (5 X/Y-coordinate pairs) input features for our models.

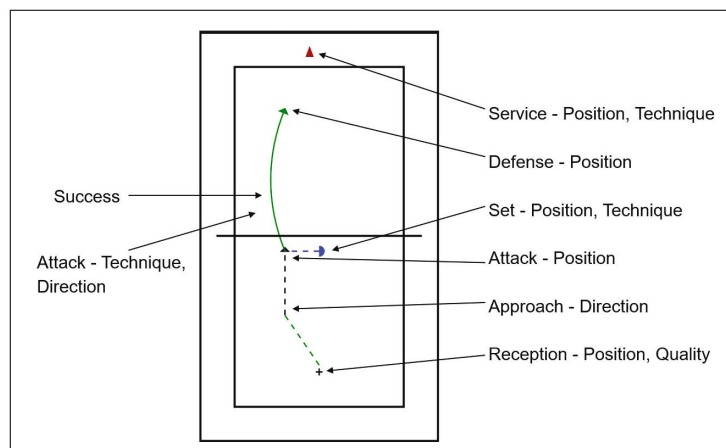


Figure 1. Positions and event features captured in a typical standard sideout. The figure sketches the 8x8m area of the beach volleyball court as well as the 1 meter wide area around it that accounts for positions outside the playing field. The black and green dotted lines depict player movement from reception to attack, while the blue dotted line represents the ball movement after the set.

Models

Except for the gradient boosted tree, all models were implemented using the PyTorch (Paszke et al., 2017) and fast.ai (Howard et al., 2018) python libraries, running on Google Compute Engine (GCE) virtual machines. The boosted tree was implemented leveraging the XGBOOST

(Chen, & Guestrin, 2016) python library and also ran on GCE.

Input features

We presented the models two different sets of input features: One with only the 5 positional X/Y tuples and one with the positional data combined with the performance indicators. In order to combine positional data and PIs for the neural networks, PIs were passed through automatically sized embedding layers (Kocmi, & Bojar, 2017) before concatenating the resulting output vector with the positional data. For the XGBOOST model, PIs were encoded using a One-Hot encoding scheme before passing them to the model along with the positional features.

Target variables

In addition to the different input features, we also evaluated the performance of our predictions for different target variables. We tried to predict the SUCCESS of a rally, the attack DIRECTION and attack TECHNIQUE based on the events and/or positions that occurred in the rally before. The SUCCESS was encoded as point/no point in the sideout, while the TECHNIQUE could take the values smash (strong attack) or shot (more precise, less strong attack). Finally, six values (diagonal, short diagonal, dink, cut, middle, line) were captured for the DIRECTION variable.

Training

Each model was trained six times, one time per combination of input features and target variable. We randomly split our data in training and test sets using a split size of 20 percent for the test set. The learning rate for the neural networks was set to $5e-3$, as determined by a learning rate finder (Smith, 2017). Training then consisted of 10 epochs using the 1-cycle policy (Smith, 2018) and the Adam optimizer (Kingma, & Ba, 2014), that were iterated 50 times. The model with the best accuracy on the test set per 10-epoch-iteration was saved for model evaluation, resulting in 50 accuracy values per model-input-target combination. The loss function used in all neural network training cases was the cross-entropy loss, which combines a logarithmic softmax-layer with the negative log likelihood loss.

XGBOOST

The gradient boosted tree model was set up to fit 50 trees with a maximum depth of 5 nodes. The objective function was set to logistic regression for binary classification for the SUCCESS and TECHNIQUE target variables, and to softmax for multiclass classification for the DIRECTION target.

MLP

In the multi-layer perceptron (Rosenblatt, 1961), the 10 positional input features are first passed through a 1-dimensional batch-normalization layer (Ioffe, & Szegedy, 2015), with momentum set to 0.1. We then use three blocks of linear layers followed by rectified linear unit (ReLU) activation functions (Maas, Hannun, & Ng, 2013), and batch normalization layers with sizes of 128, 64 and 32 neurons each. The final linear layer then maps to the class size of the respective target variable. In the combined input feature case, the embedding output is concatenated to the output of the first batch-normalization layer before passing it to the linear layers.

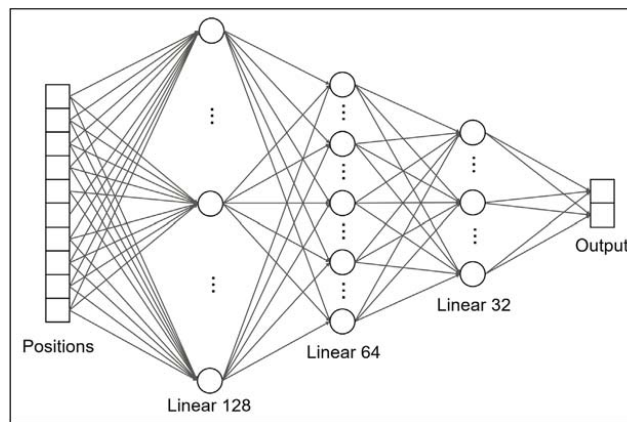


Figure 2. MLP architecture for positions only input, batch normalization and activation layers are omitted

CNN

Our convolutional network consisted of five 1-dimensional convolutional blocks with different kernel sizes, strides and dilations (Fig. 3) in order to capture different relative patterns between positions in the input. The concatenated output from those five blocks was then fed to three linear layers with batch-normalization and ReLU activations for classification. Again, we combined this architecture with embedding layers for the performance indicators, whose output was concatenated with the output of the convolutional blocks before it was passed to the linear layers.

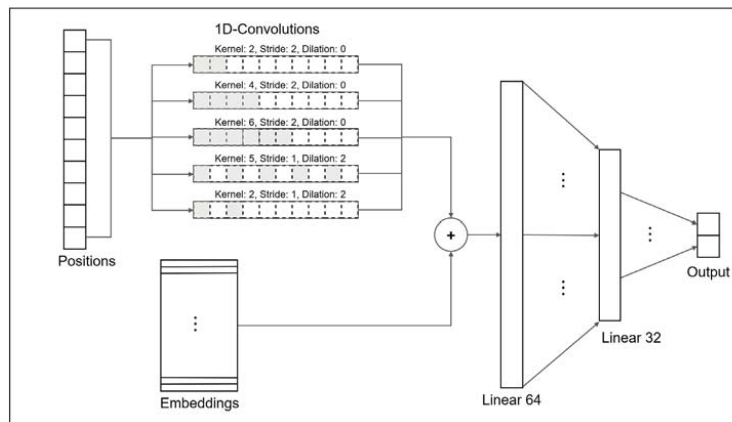


Figure 3. CNN architecture for the combined input case. BN-layers and activations are omitted for the linear layers.

RNN – GRU

Our choice of recurrent neural network was a so called Gated-Recurrent-Unit (GRU) (Cho, Van Merriënboer, Gulcehre, Bahdanau, Bougares, Schwenk, & Bengio, 2014), a kind of Long-Short-Term memory (LSTM) neural network that has proven to be easier to train on smaller datasets (Chung, Gulcehre, Cho, & Bengio, 2014). Our model consisted of a 5-layer GRU with a dropout rate of 20 percent, followed by two linear layers with batch-normalization and ReLU activations for classification. Since recurrent networks expect input data to be

sequenced, the positional data was reshaped to a sequence of five X/Y-coordinate tuples for the playing positions, ordered by occurrence in a beach volleyball rally. To combine positional and PI data, embedding layers were used for the PIs. The output of the embeddings was then concatenated with the output of the GRU before passing it to the final linear classification layer.

Evaluation

For the evaluation we performed a three-way ANOVA with classification accuracy as dependent variable and model, target variable and input features as factors. All statistical tests were performed using IBM SPSS Statistics for Windows, Version 23. G*Power analysis revealed that with a total sample size of 1200 accuracy values, small effect sizes ($\eta^2 = .018$) could be detected with a power of 0.96.

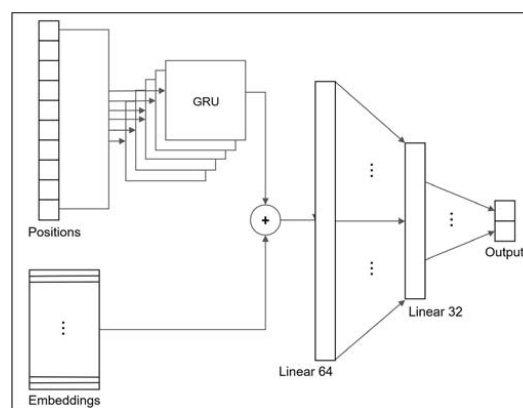


Figure 4. RNN architecture for the combined input case. Again, BN-layers and activations are omitted for the linear layers.

Results.

A three-way ANOVA was conducted to determine the effects of model, target variable and input features on the classification accuracy. There was a statistically significant three-way interaction between model, target variable and input features ($F(6, 1176) = 92.37, p < .001, \eta^2 = .001$). Statistical significance was accepted at the $p < .025$ level for simple two-way interactions and at $p < .016$ for simple main effects. There was a statistically significant simple two-way interaction between target and model for positional input ($F(6, 1176) = 122.5, p < .001$), as well as for the combined input ($F(6, 1176) = 148.75, p < .001$). There were statistically significant simple main effects of the model for the target variables Direction ($F(3, 1176) = 72.004, p < .001$), Success ($F(3, 1176) = 53.282, p < .001$) and Technique ($F(3, 1176) = 323.47, p < .001$), when training with the combined input features. The positional input also showed statistically significant ($p < .001$) simple main effects of model for all target variables (Direction: $F(3, 1176) = 93.41$, Success: $F(3, 1176) = 63.671$, Technique: $F(3, 1176) = 221.202$). All simple pairwise comparisons were run with a Bonferroni adjustment applied. Only two comparisons did not show significant differences in classification accuracy between models. The mean difference between the gradient boosted tree and the gated recurrent unit targeting the success variable, .001 percentage points (95% CI [-.001, .002]), was not statistically significant, $p = .783$. Additionally, the means between the convolutional neural net and the gated recurrent unit, .001 percentage points (95% CI [.000, .003]) showed no significant differences, $p = .119$.

Table 1: : 3-factor ANOVA-Results, sum of squares, F-value, significance p and effect size η^2 for each factor, three twofold interactions and the threefold interaction (*). Cell means are depicted in figure 5.

	<i>SS</i>	<i>F</i>	<i>p</i>	η^2
Factor Model	.0062	282.02	< .001	.001
Factor Features	.7254	99727.01	< .001	.115
Factor Target	4.6522	319803.89	< .001	.735
Model*Features	.0001	2.52	.056	< .001
Model*Target	.0078	178.89	< .001	.001
Features*Target	.9267	63702.91	< .001	.146
Model*Features*Target	.004	92.37	< .001	.001
Error	.0086			.001

Figure 5 shows the profile plot for the three-way interaction effect, together with the mean classification accuracies per combination of model, input data and target variable. Classification accuracies for the prediction of the attack direction showed values at roughly 37% for the positions-only input, and 51% for the combined input features. In contrast, the prediction of the attack technique showed relatively similar accuracies for both the combined (55%) and the positions-only input (54%). Finally, the difference in classification precision between the different input features for the prediction of the success showed even smaller discrepancies with roughly 59% correctly recognized rallies for both cases.

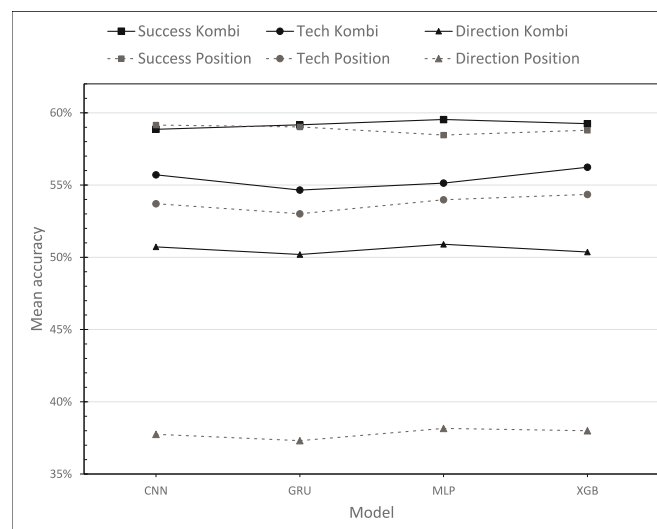


Figure 5. Profile plot for the three-way interaction effect. The values on the y-axis depict the mean classification accuracy, while the x-axis shows the employed models.

Discussion

Results revealed consistent significant differences between models in terms of classification accuracy. Our classification results achieve better performance than strict guessing in all cases

(see figure 5), with prediction accuracies ranging from around 37% for the forecasting of the attack direction (guessing: 17%) to almost 60% for the prediction of success and attack technique (guessing: 50%). The absolute differences in accuracy between models are in the range of .001 percentage points, however. The significant differences are most likely caused by the small within-group variances with coefficients of variation between .02% and 1.2%, which are the result of our evaluation method.

However, these small variations also give reason to believe that there exists an upper bound to accuracy. Similar studies have found these upper bounds in other sports (Weissbock & Inkpen, 2014) and gave difference in skill, as well as random chance as reasons for their existence. Given the available data and the innate complexity of the decision-making process, we believe that these upper bounds also exist in beach volleyball.

Effect sizes reveal, that most of the variation in classification accuracy is explained by the target of the prediction. This can also be observed in figure 5 by comparing the models in the interaction plot, where the differences between the models for the different target variables are more pronounced compared to the differences between models for varying input features. The strongest increase in classification accuracy for this case can be observed when using both position and PI data for the prediction of the attack direction compared to the prediction using only positional data. This is caused by the inclusion of the attack technique as an input feature in our opinion, which would also explain the high effect sizes of input features and the interaction between target variable and input features. Different attack techniques are used as a solution to certain block and defense constellations, which in turn imply specific shot directions due to standard defensive formations. One such example would be a line block in combination with a diagonal defense, which encourages a shot to the line as offensive solution. Interestingly, all effects involving the different models show eta-squared values lower or equal than .001, suggesting that it is not the choice of model, but the appropriate choice of input features and target variable in the sports context that affect classification accuracy the most. This also means that future studies should try to include the positions of the opponent team as model inputs, since top-level beach volleyball players can perceive the opponent's defensive formation during the approach and include this information in the decision-making process for the own attack. We did not have access to this type of data and it will most likely be only available if it becomes mandatory for all athletes to wear sensors providing positional data during tournaments in the future, however.

Considering the absolute classification performance, our models are not yet at the level to warrant practical application in day to day performance analysis in beach volleyball. However, even though there are no comparable studies in beach volleyball, work in other sports has shown that performance prediction in team sports is one of the hardest problems due to the dynamical interaction process of the playing parties. As examples, Wang and Zemel (2016) reach 77.9% top-3 accuracy in classifying NBA offensive plays, while Weissbock, Viktor and Inkpen (2014) were able to correctly classify success (Win/Loss) in Ice Hockey with an accuracy of 60.25 %. Parmar, James, Hughes, Jones and Hearne (2017) achieve a prediction accuracy of 85.5% in predicting team wins in rugby from PIs, even though they use data aggregated over whole games. The performance of our models is in a similar accuracy range, considering we are not able to use continuous position data, as Wang and Zemel for example. In contrast to previous work, we are also not trying to predict match outcomes, but actions and success in single rallies, using only the prior sequence of events from that rally as input. Under these circumstances, the results are encouraging.

Even though our model parameters were determined after a thorough experimentation phase, we are confident that more specialized models, with even more sophisticated feature

engineering can achieve higher accuracy values. Further studies could provide deeper comparisons between different network configurations per model (e.g. more linear layers or different hyper-parameter configurations) in order to increase classification performance. Especially the inclusion of continuous positional data, as provided by electronic tracking systems, may boost prediction performance considerably, since it inherently contains temporal information. Since we grouped the rallies for both men and women together, our models are not able to differentiate gender-specific behaviors. It could also make sense to train models for specific players or teams in order to gain more specific insights into individual tactics. This is especially important for practical performance analysis (Lames & McGarry, 2007), since it directly influences the possibilities to use our models e.g. for strategy generation against specific opponents as well as retrospective analysis of the own team for coaches and scouts. At the same time, this would greatly reduce the available training data and could lead to the inability to successfully train certain “deep” network models, however.

We did not compare models using only PI data for several reasons. Some models (e.g. RNNs and CNNs) expect data to be spatially or temporally structured in order to generate best results, which is not the case for categorical performance data. Moreover, the long-term goal using AI in performance analysis would be to rely only on positional data to automatically predict performance measures in a beach volleyball rally, without the help of (manually) collected PI data. Lastly, even though the absolute classification performance between models didn’t differ much, the gradient boosted tree model still has a particular advantage, since it allows easy interpretation of its decisions and can even assign importance to different input features to rank them for the experts. Yet, considering the recent trend in interpretability research for neural networks (Melis, & Jaakkola, 2018), it can be expected that at least some of these features will be available for neural networks in the near future too.

Future applications could also try to predict set/match outcome, for example employing recurrent neural networks to incorporate the course of the set/match. Another interesting approach would be to leverage the current advances in image recognition to classify successful rallies. Given the availability of continuous position data, this could be done by encoding player, and potentially ball positions together with additional information (e.g. velocity) in an image and using transfer learning to apply already trained models to the image classification task. This approach is already in use, for example to identify fraudulent website access via analysis of mouse movements (Esman, 2017). Another possible extension could try to group multiple models together using an ensemble method to increase classification accuracy for different target variables.

Regardless of which method will be used in the future, the prediction of tactical behavior could influence many applications and provide valuable insights in successful rally patterns in order to support both the scouting as well as the training process in beach volleyball.

Scouts could be able to rely on models to automatically detect and update behaviors in real-time, minimizing the required inputs to capture all tactical details per rally during live-scouting. Additionally, sufficiently accurate models could be able to fill out manually scouted rallies, again reducing the workload for the scouts, who currently have to analyze beach volleyball games in a two-stage process that takes roughly three hours per game. In turn, faster analysis of games could lead to a competitive advantage in tournaments, where athletes often have less than a day to prepare for their next match.

Coaches could use models trained on single player’s or opponent team’s data in conjunction with a simulative approach to find tactical behaviours that provide the highest chance at success against certain opponents. In the same way, they could simulate variances of the behavior of their own teams to find strategies for improvement in training.

In addition to the betting market, who could employ advanced models for win/loss-prediction to offer improved betting odds, the media could use such models to enhance their broadcasts by showing live analysis (e.g. overlays for the most probable attack technique/direction used) and statistics of beach volleyball rallies, which would increase viewer engagement and help the sport to grow in popularity.

Conclusion

In conclusion, since all models show relatively similar performance, our results suggest that for the moment the available data is the limiting factor for a implementation of rally classification in beach volleyball practice. Given that our dataset may be one of the biggest collected records of beach volleyball performance data to date, the impact of machine learning techniques on todays practical applications seems small. However, if we look at the evolution of data in other sports together with the recent development in beach volleyball, the use of continuous positional data for prediction may be possible in the near future. For this reason, and because we think the search for efficient solutions in performance analysis is one of the main tasks of computer science in sports, we encourage further studies in the AI domain to evaluate the performance of methods in different sports, using different data representations.

Acknowledgements

This study was supported by a grant from the German Federal Institute of Sports Science (BISp ZMVI4-070504/19-20). We would also like to thank the scouts of the Olympic Training Center (OSP) Stuttgart, as well as the coaches of the national teams for their support in collecting the data.

References

- Arndt, C., & Brefeld, U. (2016), Predicting the future performance of soccer players. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 9, 373-382.
- Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S., & Matthews, I. (2015). Large-Scale Analysis of Soccer Matches Using Spatiotemporal Tracking Data. *In Proceedings - IEEE International Conference on Data Mining, ICDM*.
- Bianchi, F., Facchinetti, T., & Zuccolotto, P. (2017). Role revolution: Towards a new meaning of positions in basketball. *Electronic Journal of Applied Statistical Analysis*, 10(3), 712–734.
- Bunker, R. P., & Thabtah, F. (2019). A machine learning framework for sport result prediction. *Applied Computing and Informatics*, 15(1), 27–33.
- Chen, T., & Guestrin, C. (2016). {XGBoost}: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- Claudino, J. G., Capanema, D. de O., de Souza, T. V., Serrão, J. C., Machado Pereira, A. C., & Nassis, G. P. (2019). Current Approaches to the Use of Artificial Intelligence for Injury Risk Assessment and Performance Prediction in Team Sports: A Systematic Review. *Sports Medicine - Open*, 5(1).

- Dick, U., & Brefeld, U. (2019). Learning to Rate Player Positioning in Soccer. *Big Data*, 7(1), 71–82.
- Esman, G. (2017, April 18). Splunk and Tensorflow for Security: Catching the Fraudster with Behavior Biometrics. Retrieved from <https://www.splunk.com/blog/2017/04/18/deep-learning-with-splunk-and-tensorflow-for-security-catching-the-fraudster-in-neural-networks-with-behavioral-biometrics.html>
- Hassan, A., Schrapf, N., Ramadan, W., & Tilp, M. (2017a). Evaluation of tactical training in team handball by means of artificial neural networks. *Journal of Sports Sciences*, 35(7).
- Hassan, A., Schrapf, N., & Tilp, M. (2017b). The prediction of action positions in team handball by non-linear hybrid neural networks. *International Journal of Performance Analysis in Sport*, 17(3), 293–302.
- Howard, J., et. al. (2018). fast.ai. GitHub. Retrieved from <https://github.com/fastai/fastai>
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.
- Kautz, T., Groh, B. H., Hannink, J., Jensen, U., Strubberg, H., & Eskofier, B. M. (2017). Activity recognition in beach volleyball using a Deep Convolutional Neural Network. *Data Mining and Knowledge Discovery*, 31(6), 1678–1705.
- Kempe, M., Grunz, A., & Memmert, D. (2015). Detecting tactical patterns in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks. *European Journal of Sport Science*, 15(4), 249–255.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kocmi, T., & Bojar, O. (2017). An exploration of word embedding initialization in deep-learning tasks. *arXiv preprint arXiv:1711.09160*.
- Lames, M., & McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport*, 7(1), 62-79.
- Lapham A.C. , & Bartlett, R.M. (1995) The use of artificial intelligence in the analysis of sports performance: A review of applications in human gait analysis and future directions for sports biomechanics, *Journal of Sports Sciences*, 13(3), 229-237.
- Leicht, A. S., Gómez, M. A., & Woods, C. T. (2017). Explaining match outcome during the men’s basketball tournament at the olympic games. *Journal of Sports Science and Medicine*, 16(4), 468–473.
- Link, D. (2014). A toolset for beach volleyball game analysis based on object tracking. *Int. J. Comp. Sci. Sport* 13, 24–35
- Link, D., & Hoernig, M. (2017). Individual ball possession in soccer. *PLoS ONE*. 12(7): e0179953. <https://doi.org/10.1371/journal.pone.0179953>
- Link, D. (2018). Data Analytics in Professional Soccer. Springer Fachmedien Wiesbaden.
- López-Valenciano A., Ayala F., Puerta J.M., De Ste Croix M., Vera-Garcia F.J., Hernández-Sánchez S., Ruiz-Pérez I., & Myer G.D. (2018). A preventive model for muscle injuries: a novel approach based on learning algorithms. *Med Sci Sports Exerc*. 50(5). 915–927
- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In *Proc. Icml*. 30(1), 3-9.
- Melis, D. A., & Jaakkola, T. (2018). Towards robust interpretability with self-explaining neural networks. In *Advances in Neural Information Processing Systems* (pp. 7775-7784).
- Net gains - the evolution of beach volleyball. (2016, July 21). Retrieved from <https://www.olympic.org/news/net-gains-the-evolution-of-beach-volleyball>

- Parmar, N., James, N., Hughes, M., Jones, H., & Hearne, G. (2017). Team performance indicators that predict match outcome and points difference in professional rugby league. *International Journal of Performance Analysis in Sport*, 17(6), 1044–1056.
- Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., ... & Lerer, A. (2017). Automatic differentiation in PyTorch. In Long Beach, California, USA: Autodiff Workshop.
- Perl, J., & Memmert, D. (2011). Net-Based Game Analysis by Means of the Software Tool SOCCER. *International Journal of Computer Science in Sport*, 10(2), 77–84.
- Peterson, K. D. (2018) Recurrent Neural Network to Forecast Sprint Performance, *Applied Artificial Intelligence*, 32(7-8), 692-706
- Rosenblatt, F. (1961). Principles of neurodynamics. perceptrons and the theory of brain mechanisms (No. VG-1196-G-8). Cornell Aeronautical Lab Inc Buffalo NY.
- Russell, S., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach - Third Edition. Pearson.
- Schrapf, N., Alsaied, S., & Tilp, M. (2017). Tactical interaction of offensive and defensive teams in team handball analysed by artificial neural networks. *Mathematical and Computer Modelling of Dynamical Systems*, 23(4), 363–371.
- Smith, L. N. (2017). Cyclical learning rates for training neural networks. In *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 464-472). IEEE.
- Smith, L. N. (2018). A disciplined approach to neural network hyper-parameters: Part 1--learning rate, batch size, momentum, and weight decay. *arXiv preprint arXiv:1803.09820*.
- Tümer, A. E., & Koçer, S. (2017). Prediction of team league's rankings in volleyball by artificial neural network method. *International Journal of Performance Analysis in Sport*, 17(3), 202–211.
- Vales-Alonso, J., Chaves-Dieiguez, D., Lopez-Matencio, P., Alcaraz, J. J., Parrado-Garcia, F. J., & Gonzalez-Castano, F. J. (2015). SAETA: A Smart Coaching Assistant for Professional Volleyball Training. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(8), 1138–1150.
- Van Haaren, J., Ben Shitrit, H., Davis, J., & Fua, P. (2016). Analyzing volleyball match data from the 2014 world championships using machine learning techniques. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, (pp. 627-634)., New York, NY, USA. ACM.
- Wang, K., & Zemel, R. (2016). Classifying NBA Offensive Plays Using Neural Networks. *MIT Sloan Sports Analytics Conference*, 1–9.
- Wang, Y., Zhao, Y., Chan, R. H. M., & Li, W. J. (2018). Volleyball Skill Assessment Using a Single Wearable Micro Inertial Measurement Unit at Wrist. *IEEE Access*, 6, 13758–13765.
- Weissbock, J., & Inkpen, D. (2014). Combining textual pre-game reports and statistical data for predicting success in the national hockey league. In *Canadian Conference on Artificial Intelligence* (pp. 251-262). Springer, Cham.
- Zetou, E., Moustakidis, A., Tsigilis, N., & Komninakidou, A. (2007). Does Effectiveness of Skill in Complex I Predict Win in Men's Olympic Volleyball Games? *Journal of Quantitative Analysis in Sports*, 3(4).