Abstract

Neural Network Approaches for Distributional Shifts in Environmental Sensors †

Tobias Sukianto [1,*]**, Sebastian A. Schober** [1,2]**, Cecilia Carbonelli** [1]**, Simon Mittermaier** [1] **and Robert Wille** [3,4]

[1] Infineon Technologies AG, 85579 Neubiberg, Germany; sebastian.schober@infineon.com (S.A.S.);
     cecilia.carbonelli@infineon.com (C.C.)
[2] Institute for Integrated Circuits, Johannes Kepler University Linz, 4040 Linz, Austria
[3] Chair for Design Automation, Technical University of Munich, 80330 Munich, Germany; robert.wille@tum.de
[4] Software Competence Center Hagenberg GmbH (SCCH), 4234 Hagenberg, Austria
[*] Correspondence: tobias.sukianto@infineon.com
[†] Presented at the XXXV EUROSENSORS Conference, Lecce, Italy, 10–13 September 2023.

Abstract: Harmful pollutants in the air have become a severe concern in our health-conscious society. Consequently, deploying low-cost environmental sensors and the application of machine learning algorithms to the sensor raw data are crucial to enabling an overall assessment of the air quality around us. Due to the distributional shift between the training and operational environment induced by sensor ageing and drift processes, the algorithms that predict air quality suffer from performance degradation during the products' lifetime. We propose a novel transformer-based model architecture inspired by the field of natural language processing, showing advantages compared to other architectures in the presence of distributional shifts.

Keywords: environmental sensors; neural networks; distributional shifts

1. Introduction

During sensor development, different effects, such as sensor-to-sensor variations and sensor aging, but also rapid and repeated optimization of the physical sensor structure throughout the design process, cause severe shifts in data distribution, requiring time-consuming and expensive data recollection. This is a bottleneck in the development of the sensor, where measurements and laboratory resources are often scarce. In order to tackle the problem of distributional shifts during development, we investigated a light transformer architecture (Figure 1a) inspired by the field of natural language processing [1] in terms of improved robustness to small sensor-specific distributional shifts. We compared it to a gated recurrent unit (GRU)-inspired model [2] and obtained very promising results.

(a)

(b)

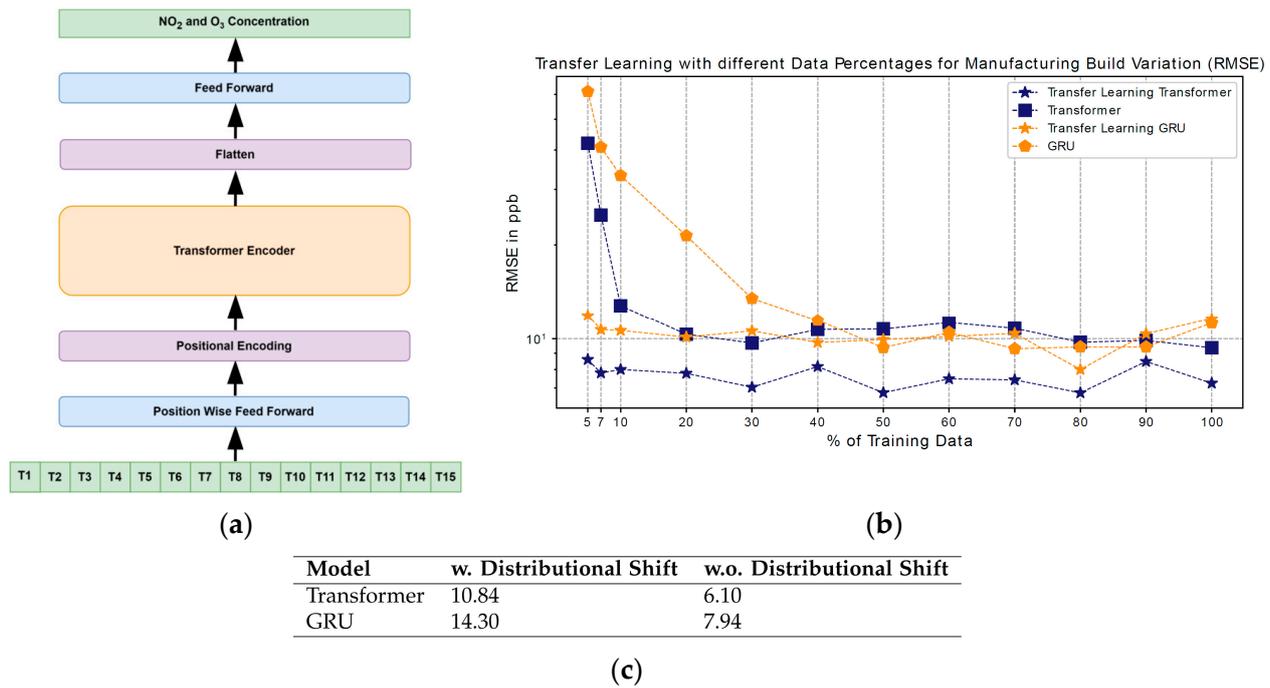| Model	w. Distributional Shift	w.o. Distributional Shift
Transformer	10.84	6.10
GRU	14.30	7.94

(c)

Figure 1. (**a**) Transformer algorithm, (**b**) transfer learning results, (**c**) architecture comparison.

2. Materials and Methods

The transformer-inspired architecture is a neural network-based regression architecture that can process the time-dependent sensor signals to predict the concentration of gases. For our specific application, we leverage only the transformer's encoder part to process an input sequence without exhibiting a recurrent structure, allowing the architecture to be highly parallelable. The encoder transforms each sensor's signal from the input sequence and forms a context-independent representation for multiple context-dependent representations. The fundamental purpose of this mechanism is to extract information about how relevant a sensor signal is to other sensor signals. Hence, the algorithm focuses on different key parts of the sequence for prediction tasks and increases its ability to learn long-term dependencies. The computational complexity of our architectures is seen to be computationally lightweight and is integrable into low-memory devices. We compare the performance of this architecture with a GRU architecture in the presence of distributional shifts, regarding data efficiency and the use of transfer learning (TL) [3].

3. Discussion

In the first experiment, we compared the performance and stability of the GRU and transformer architectures. The table in Figure 1c shows their prediction performances when the training and testing datasets are from sensors with different or the same physical structures. The results indicate that the transformer is more robust to distributional shifts than the GRU. In a second experiment, shown in Figure 1b, the applicability of TL for updating the regression models trained on old data stemming from a different sensor configuration with different quantities of new data is evaluated in comparison to completely train the models solely on the new data. It is observed that at low quantities of the new data, TL strongly enhances the sensor performance of both models' architectures, where the transformer is seen to be more data-efficient than the GRU. Our results also show that the combination of TL with transformers has a more significant performance gap to their newly trained counterpart than GRUs, which might indicate superior TL and prior knowledge-extraction capabilities. Overall, our experiments substantiate the capabilities of transformer architectures to increase environmental sensing algorithms' stability and data efficiency. Furthermore, TL techniques show promising benefits for reusing data from

previous sensor technologies, accelerating sensor development and time to market and saving measurements and laboratory resources. Further experiments regarding sensor-to-sensor variations and aging show equivalent results.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Dataset is available on request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, 4–9 December 2017.
2. Cho, K.; van Merrienboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
3. Karl, W.; Khoshgoftaar, T.M.; Wang, D.D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9.