# Effects of Time-Series Data Pre-processing on the Transformer-based Classification of Activities from Smart Glasses

Gabriela Ciortuz
Institute of Medical Informatics,
University of Lübeck
Lübeck, Germany
gabriela.ciortuz@uni-luebeck.de

Marcin Grzegorzek
Institute of Medical Informatics,
University of Lübeck
Lübeck, Germany
Department of Knowledge
Engineering, University of Economics
Katowice, Poland
marcin.grzegorzek@uni-luebeck.de

Sebastian Fudickar
Institute of Medical Informatics,
University of Lübeck
Lübeck, Germany
sebastian.fudickar@uni-luebeck.de

## ABSTRACT

Time-series classification is gaining significance in pattern recognition as time-series data becomes more abundant along with the increasing digitization of daily life and the rise of the Internet of Things (IoT). One of the biggest challenges lies in the ordered nature of time-series attributes, making traditional machine learning (ML) algorithms designed for static data unsuitable for processing temporal data. The Transformer architecture was introduced as a novel approach in natural language processing for machine translation tasks, relying solely on attention mechanisms without the need for convolution or recurrence. Since machine translation is similar to time-series data, where order is an important factor, it is also worth considering the Transformer for time-series classification. Pre-processing the data is a crucial step in the ML process and can influence the data and impact the effectiveness of the ML models. In this paper, we aim to address the effects of time-series pre-processing and data representation in combination with the Transformer model for Human Activity Recognition (HAR) using IMU data from smart glasses as input. We analyze the results based on established evaluation metrics such as the F1-score and the area under the curve (AUC).

## CCS CONCEPTS

• **Computing methodologies → Machine learning algorithms**; **Neural networks**; **Model verification and validation**.

## KEYWORDS

transformer, neural networks, human activity recognition, time-series data, pre-processing, data representation, wearables

**ACM Reference Format:**
Gabriela Ciortuz, Marcin Grzegorzek, and Sebastian Fudickar. 2023. Effects of Time-Series Data Pre-processing on the Transformer-based Classification of Activities from Smart Glasses. In *8th international Workshop on Sensor-Based Activity Recognition and Artificial Intelligence (iWOAR 2023), September 21–22, 2023, Lübeck, Germany*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3615834.3615858


*iWOAR 2023, September 21–22, 2023, Lübeck, Germany*

ACM ISBN 979-8-4007-0816-9/23/09.
https://doi.org/10.1145/3615834.3615858

## 1 INTRODUCTION

The Transformer architecture in deep learning (DL) has gained significant attention, due to its remarkable achievements in the field of natural language processing (NLP) [15]. Starting from NLP, there has been a surge in the development of numerous Transformer models, including in time-series classification, anomaly detection and forecasting [16].

[13] proposed an adaptation of the Transformer model for HAR. They fed their Transformer model with three seconds of normalized time-series data from smartphone motion sensors, specifically the accelerometer and gyroscope. Similarly, we aim at feeding a simple Transformer architecture that uses the Keras MultiHeadAttention Layer[14] with accelerometer and gyroscope data, not from a smartphone, but from smart glasses. Smart glasses have been used in various machine learning tasks, such as [9, 11], which used them in combinations with other wearables in order to determine the optimal sensor modalities for different classification tasks.

Classification of state activities from wearable data is a relevant step for HAR and corresponding health screening measures e.g. to predict frailty [6, 7]. State activities are activities characterising the pose or motion of a subject [11].

Data normalization is a crucial pre-processing step in Machine Learning (ML). Data representations can refer to different feature sets or data encodings used as input to the classifiers, and they can promote diversity and provide a more robust and accurate prediction [3].

Determining the proper pre-processing and data representation for a specific classification task typically requires an empirical comparison study of the available techniques [5]. Because datasets, labels, and models vary among classification tasks and because there are many factors that can influence the classification accuracy, there is no one for all solution [3]. In [2], we compared the performance of the Transformer to the performance of the LSTM and observed, that comparable results were achieved by both models, only when the input data representation differed.

Consequently, it is critical to consider and investigate the impact of data normalization approaches and data representation techniques on classification accuracy [1].

## 2 METHODOLOGY DESCRIPTION

In order to evaluate the effects of different time-series data pre-processing techniques on the Transformer-based model for HAR, we conducted the following experiments on the dataset from [11] made up of six state activities:

(1) using raw sensor data as input to the Transformer model
(2) applying standardisation
(3) applying normalisation

Each experiment was conducted separately for the following two common data representation techniques: A non-flattened condition with an LSTM-similar input, and flattened input data, instead. The non-flattened input means that the channels of each sensor are concatenated on the y-axis, while the flattened input means that the concatenation of all the channels forms a long row vector.

In this article, we applied these two data representation techniques in combination with the aforementioned three data preprocessing techniques for the same dataset.

Considering the type of sensor data, wearable-based HAR data of the JINS MEME smart glasses are used with a sampling rate of 20 Hz and a fixed segment length of four seconds during experiments. The smart glasses are equipped with an accelerometer and a gyroscope per each principal axis: pitch, roll and yaw. The following six state activities were considered as input to the transformer model: bending, lying, sitting, squatting, standing, and walking.

z-normalization and min-max normalization are conventional pre-processing steps for both univariate and multivariate data and can significantly impact the learning capabilities of deep neural networks [1]. Further research in this area could provide valuable insights into optimizing the performance of DNNs for time-series analysis [5].

We trained the Transformer architecture described in [2] and trained the model for 300 epochs using a batch size of eight. The robust average F1 Score (AF1) and Area Under the Curve (AUC) are used as main evaluation metrics. The AUC is the Area Under the Reciever Operating Characteristic (ROC) curve and reflects classifiers' overall ranking performance and was proven theoretically and empirically better than the accuracy metric [8]. F1-score calculates the harmonic mean between precision and recall [4]. The AF1-score is the average of the F1-scores calculated independently using a one-vs-all approach for each of the six classes as already commonly used in other works [11, 12]. We also plotted ROC curves and confusion matrices in order to visually analyze the results. ROC curves are commonly employed to visually represent the relationship between sensitivity and specificity for each feasible threshold in a test or a series of tests. In terms of multi-class, the micro-average ROC is preferred and it combines the contributions from all classes to compute the mean metric [10].

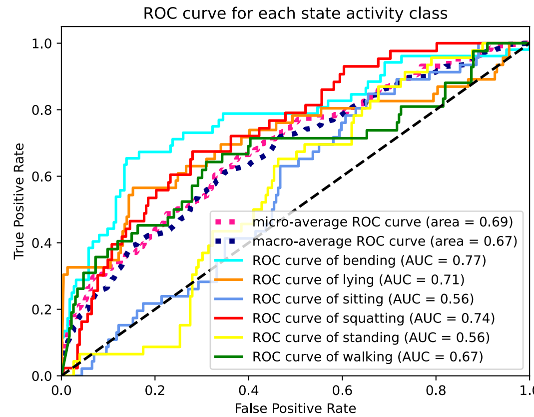


**Figure 1: ROC Curve of the activity classification with the Transformer using no preprocessing at all**

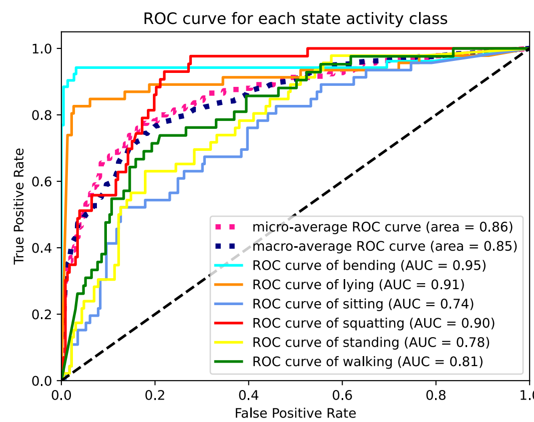


**Figure 2: ROC Curve of the activity classification with the Transformer using flattened and standardized data**

## 3 EXPERIMENTAL RESULTS

The experimental results are summarized in Table 1. 22% improvement can be observed between no pre-processing and using the best data representation for this case (40,82% and 62,42% AF1-score). Regarding AUC, the results improved greatly when applying any kind of pre-processing or better data representation, as seen in the ROC curves in the Figures 1 and 2.

We investigated the confusion matrices of both data representations with no pre-processing because they achieve the lowest and highest performance in terms of AF1 scores and accuracy (See figure 3, which is the confusion matrix of the classification with non-flattened data and figure 4 for the flattened data, respectively. These confusion matrices clearly indicate the benefit of adjusting the data representation While all classes were sharper separated in the second confusion matrix, indicating overall better performance, an enhancement was especially seen for the classes bending and lying.

## 4 CONCLUSION

Based on the experimental results, it is evident that any kind of pre-processing significantly improves the AUC in the Transformer model. One reason behind it might be that in general, the model is provided with a better feature representation, such as more informative and discriminative features, which in turn contribute to better performance.

Another noteworthy observation is that a well-designed data representation is essential for the Transformer's success. When the time-series data representation is more similar to the NLP case

**Table 1: Average F1-Scores, Accuracies, MAPs and AUC (in %) obtained by the Transformer model using the different pre-processing approaches**

| Data representation | Pre-processing | AF1 | Accuracy | MAP | AUC |
|---|---|---|---|---|---|
| Not-flattened | No pre-processing | 40.82 | 41.12 | 34.47 | 66.94 |
| | Min Max Normalisation | 47.52 | 51.71 | 51.64 | 81.29 |
| | Standardisation | 53.72 | 55.12 | 54.76 | 82.08 |
| Flattened input | No pre-processing | **62.42** | **62.09** | 57.91 | 84.70 |
| | Min max normalisation | 54.07 | 55.54 | 57.86 | 84.76 |
| | Standardisation | 54.95 | 55.53 | **59.52** | **86.00** |

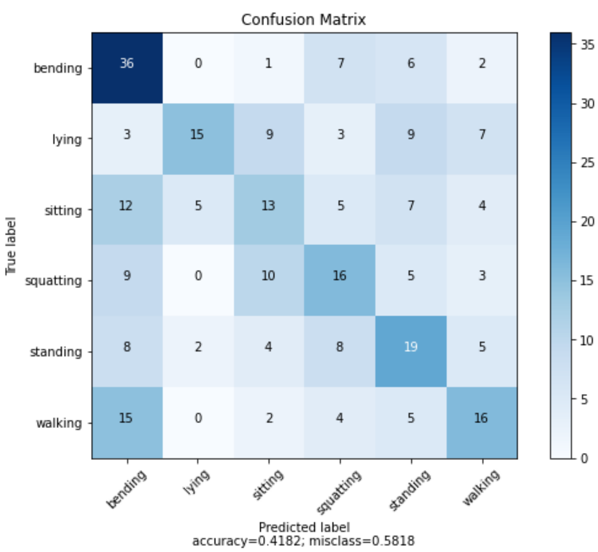


**Figure 3: Confusion matrix of the non-flattened, not-preprocessed data as input to the Transformer model**

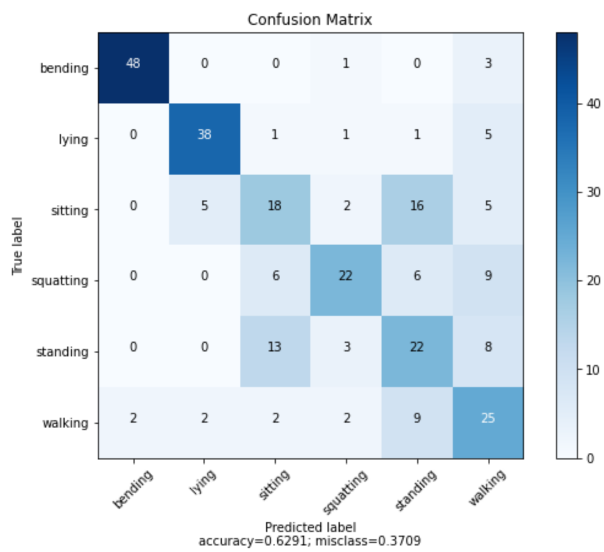


**Figure 4: Confusion matrix of the flattened, not-preprocessed data as input to the Transformer model**

presented as a long sequence of features, the Multi-Head Attention Layer exhibits improved performance.

In conclusion, the pre-processing steps applied to the data significantly impact the Transformer model's accuracy. By implementing these thoughtful pre-processing techniques, the Transformer model can achieve superior results, making it a powerful tool for various time-series classification tasks.

## ACKNOWLEDGMENTS

The study is funded by the German Federal Ministry of Education and Research (Project No. 01ZZ2007).

## REFERENCES


[1] Aishwarya Asesh. 2022. Normalization and Bias in Time Series Data. In *Digital Interaction and Machine Intelligence*, Cezary Biele, Janusz Kacprzyk, Wiesław Kopeć, Jan W. Owsiński, Andrzej Romanowski, and Marcin Sikorski (Eds.). Springer International Publishing, Cham, 88–97.

[2] Gabriela Augustinov, Muhammad Adeel Nisar, Frédéric Li, Amir Tabatabaei, Marcin Grzegorzek, Keywan Sohrabi, and Sebastian Fudickar. 2023. Transformer-Based Recognition of Activities of Daily Living from Wearable Sensor Data. In *Proceedings of the 7th International Workshop on Sensor-Based Activity Recognition and Artificial Intelligence (iWOAR '22)*. Association for Computing Machinery, Association for Computing Machinery, Rostock, Germany, 1–8. https://doi.org/10.1145/3558884.3558895

[3] Anthony Bagnall, Jason Lines, Aaron Bostrom, James Large, and Eamonn Keogh. 2017. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery* 31, 3 (01 May 2017), 606–660. https://doi.org/10.1007/s10618-016-0483-9

[4] Hercules Dalianis. 2018. *Evaluation Metrics and Evaluation.* Springer International Publishing, Cham, 45–53. https://doi.org/10.1007/978-3-319-78503-5_6

[5] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. 2019. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery* 33, 4 (mar 2019), 917–963. https://doi.org/10.1007/s10618-019-00619-1

[6] Sandra Hellmers, Sebastian Fudickar, Clemens Büse, Lena Dasenbrock, Andrea Heinks, Jürgen M. Bauer, and Andreas Hein. 2017. *Technology Supported Geriatric Assessment.* Springer International Publishing, Cham, 85–100. https://doi.org/10.1007/978-3-319-52322-4_6

[7] Sandra Hellmers., Tobias Kromke., Lena Dasenbrock., Andrea Heinks., Jürgen M. Bauer., Andreas Hein., and Sebastian Fudickar. 2018. Stair Climb Power Measurements via Inertial Measurement Units - Towards an Unsupervised Assessment of Strength in Domestic Environments. In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018) - HEALTHINF.* INSTICC, SciTePress, Funchal, 39–47. https://doi.org/10.5220/0006543900390047

[8] Mohammad Hossin and Sulaiman M.N. 2015. A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining and Knowledge Management Process* 5 (03 2015), 01–11. https://doi.org/10.5121/ijdkp.2015.5201

[9] Muhammad Tausif Irshad, Muhammad Adeel Nisar, Xinyu Huang, Jana Hartz, Olaf Flak, Frédéric Li, Philip Gouverneur, Artur Piet, Kerstin M. Oltmanns, and Marcin Grzegorzek. 2022. SenseHunger: Machine Learning Approach to Hunger Detection Using Wearable Sensors. *Sensors* 22, 20 (2022), 7711. https://doi.org/10.3390/s22207711

[10] Abdul Kader, Samiha Sharif, Pranta Bhowmick, Fahmida Mim, and Azmain Srizon. 2020. Effective Workflow for High-Performance Recognition of Fruits using Machine Learning Approaches. 7 (02 2020), 1516–1521.

[11] Frédéric Li, Kimiaki Shirahama, Muhammad Adeel Nisar, Xinyu Huang, and Marcin Grzegorzek. 2020. Deep Transfer Learning for Time Series Data Based on Sensor Modality Classification. *Sensors* 20, 15 (July 2020), 4271. https://doi.org/10.3390/s20154271

[12] Frédéric Li, Kimiaki Shirahama, Muhammad Adeel Nisar, Lukas Köping, and Marcin Grzegorzek. 2018. Comparison of Feature Learning Methods for Human Activity Recognition Using Wearable Sensors. *Sensors* 18, 2 (2018), 679. https://doi.org/10.3390/s18020679

[13] Iveta Dirgová Luptáková, Martin Kubovčík, and Jiří Pospíchal. 2022. Wearable Sensor-Based Human Activity Recognition with Transformer Model. *Sensors* 22, 5 (March 2022), 1911. https://doi.org/10.3390/s22051911

[14] Keras API Reference. 2021. *MultiHeadAttention Layer*. keras. Retrieved July 30, 2023 from https://keras.io/api/layers/attention_layers/multi_head_attention/

[15] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is All You Need. In *Attention is All You Need* (Long Beach, California, USA) *(NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, 6000–6010.

[16] Qingsong Wen, Tian Zhou, Chaoli Zhang, Weiqi Chen, Ziqing Ma, Junchi Yan, and Liang Sun. 2023. Transformers in Time Series: A Survey. arXiv:2202.07125 [cs.LG]


Received 1 August 2023