Improvement of Machine Learning Models for Time Series Forecasting in Radial-Axial Ring Rolling through Transfer Learning

SFU 2023

Johannes Seitz[1*], Qinwen Wang[2], Tonias Moser[1], Alexander Brosius[2], Bernd Kuhlenkötter[1]

1) Chair of Production Systems, Ruhr-University Bochum, Industriestraße 38c, 44894 Bochum
2) Institute of Manufacturing Technology, Chair of Forming and Machining Processes, Dresden University of Technology, Helmholtzstr. 10, 01062 Dresden
*Corresponding author: Johannes Seitz, seitz@lps.ruhr-uni-bochum.de

Abstract

Due to the increasing computing power and corresponding algorithms, the use of machine learning (ML) in production technology has risen sharply in the age of Industry 4.0 [1]. Data availability in particular is fundamental at this point and a prerequisite for the successful implementation of a ML application. If the quantity or quality of data is insufficient for a given problem, techniques such as data augmentation, the use of synthetic data and transfer learning of similar data sets can provide a remedy. In this paper, the concept of transfer learning is applied in the field of radial-axial ring rolling (rarr) and implemented using the example of time series prediction of the outer diameter over the process time. Radial-axial ring rolling is a hot forming process and is used for seamless ring production.

In this paper, the transfer between two different rolling mills with unequal data availability is analysed. The rolling mill with a comparatively large database (1300 rolling cycles) is from thyssenkrupp rothe erde GmbH (source space) and the rolling mill with the small database (60 rolling cycles) is affiliated to the Ruhr-University Bochum (target space). This rolling mill is smaller in size and can apply fewer forming forces. An already existing model according to SEITZ ET AL. [2], which represents the first 50% of the rolling process, is taken up and modified in such a way that a prediction of the outer diameter of the entire rolling process is possible. The underlying algorithm is based on recurrent neural networks and is implemented using an LSTM (Long Short-Term Memory) architecture. In order to achieve an extension to the entire rolling process, it is necessary to discard the time scaling used in [2], the time series sampling set to a fixed number of time steps, and to choose a new approach. The selected input dimension now runs each index sequentially step-by-step for all rolling ids and thus takes into account different rolling speeds and lengths. This makes it easier to compare the characteristics of the rolling processes. In addition, the concept of a hybrid database is taken into account, which transfers the training data of the target space into the training data of the source space, so the corresponding rolling strategies of the target are already taken into account in the source model. In addition to the classic source models, model generalization methods are also applied, on the one hand by adding dropout layers and on the other hand by modifying the underlying data split in the hybrid database. For each hyperparameter configuration, five models are trained and the mean value of the mean square deviation (mse) is used as the final evaluation. In addition,

the standard deviation of the five models between each other is determined. The base model, which is trained within the data of the target space, shows an mse of $1.109*10^{-3}$. Transfer learning results in an improvement of 61% of the mse and a reduction of the standard deviation of 45%. Through generalising the source model, an equal improvement is achieved, but consists of a reduced standard deviation of 81% compared to the base model. For this reason, generalisation is recommended. Using the hybrid database does not lead to the best transfer results, but a positive transfer can be observed for all hyperparameter configurations. When considering the source database for transfer learning isolated, this is not the case for all configurations.

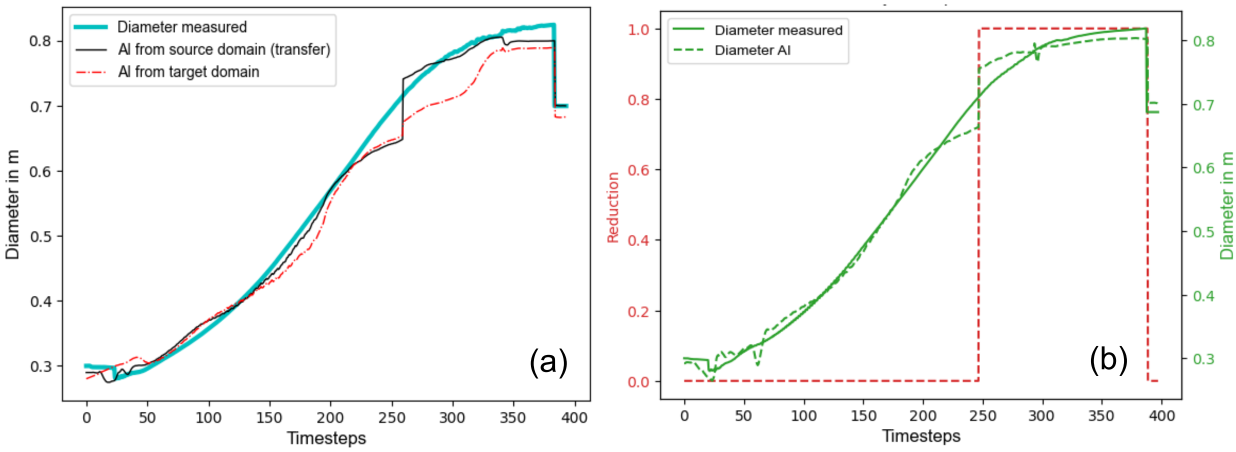

Figure 1: *Rolling process out of the test data (a) and anomaly in the prediction (b)*

The predictions conducted by the transfer learned model follow the course of the measured values best (see Figure 1a). However, the prediction quality decreases after 200 time steps and shows a clearly identifiable step in all test data trials. This step can be located at the moment when the binary value of the reduction increases from 0 to 1 (see Figure 1b). This step cannot be observed in the predictions of the source space. In the target space, this is caused by the decreasing correlations between the outer diameter and the process parameters with increasing process time, so that the algorithm is strongly orientated towards this binary step. This can be assigned to the data generation in the target space, where manual intervention in the process was targeted towards the end of the rolling operations as part of a design of experiments.

In summary, a positive transfer for the entire ring rolling process with an improvement of 61% can be achieved within the selected approach.

Keywords *Transfer Learning, Machine Learning, LSTM, Radial-Axial Ring Rolling*

References

[1] Wuest, T., Weimer, D.; Irgens, C., Thoben, K.: Machine learning in manufacturing: Advantages, challenges, and applications, 2016 Production & Manufacturing Research. 4, pp. 23-45.

[2] Seitz, J.; Moser, T.; Fahle, S.; Prinz, C.; Kuhlenkötter, B.: Transfer learning approaches in the domain of radial-axial ring rolling, 2023. In: Proceedings of the 5th Conference on Production Systems and Logistics (CPSL 2023).