

# Evaluation of Machine Learning Algorithms for the Classification of Process Data from the ERP System SAP S/4HANA



**Paul Oesterwitz, Business Information Systems**

## Description

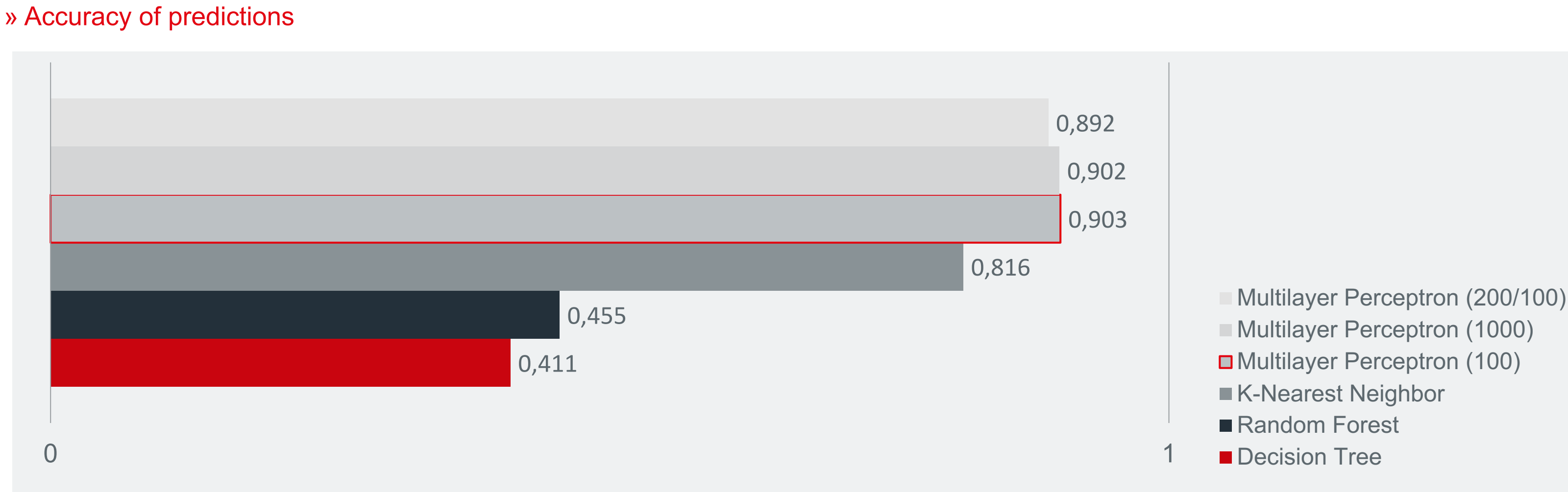
Machine Learning is one of the major technological topics of the present day and may be used in a large number of applications [1].

In particular, this technology has recently made inroads into the current ERP system »S/4HANA« by SAP SE [2]. Since SAP process data and the relational databases must adhere to pre-defined structures and rules, they are usually well-structured and of high quality. This allows to automatically find patterns in the data and thus to partially automatize a variety of business processes (such as the requisition of (indirect) material).

Preparing data proved to be surprisingly simple .

- » many of the otherwise necessary steps are already ensured by SAP application logic (correct formats, well-defined data types etc.)
- » Usage of standard embedding methods, such as OneHotEncoder or CountVectorizer proved to be effective [3].

» Accuracy of predictions



Surprisingly, the highest accuracy was reached by a rather small model comprising only 100 artificial neurons. Training more complex models to achieve higher accuracy is possible but certain care has to be taken to avoid model overfitting [4]. Without such regularization methods, larger models, such as an MLP with 1000 artificial neurons or deep models with two or more layers, showed decreasing prediction performance. Furthermore, re-using the model for other use cases is often almost trivial by just plugging in different data sets and defining other target goals.

With this, the corresponding business cases are likely to be economically attractive and may greatly automatize business processes related to the ERP-systems using Machine Learning, specifically, the »S/4HANA« by SAP SE. The insights gained from studying the use case, collecting the data set and training and evaluating the model is a great example of how theoretical knowledge can be applied to real-world problems and how the DHBW philosophy supports this endeavor.

## Results

Of the models tested, Multilayer Perceptron turned out to perform best with the prediction.

- » Accuracy of about 90% (30% of the training data set used for evaluation only)
- » This value is above the human prediction accuracy (in comparison to a group of SAP system users).
- » This accuracy was reached by a rather small model comprising only 100 artificial neurons.

## Future work

A rather small model trained in a standard fashion is already sufficiently powerful to guarantee good predictive accuracy for these specific SAP data sets. One may achieve even higher prediction accuracy with reasonable efforts by merely using a larger data set or adding further features.

## Cooperative partner

# MHP

A PORSCHE COMPANY

## Quellen

- » [1]E. Brynjolfsson and A. McAfee, "The Business of Artificial Intelligence", Harvard Business Review, 2022.
- » [2]B. Leukert, J. Müller and M. Noga, Das intelligente Unternehmen: Maschinelles Lernen mit SAP zielgerichtet einsetzen. Berlin, Heidelberg: Springer Gabler, 2018
- » [3]T. Pranckevičius and V. Marcinkevičius, "Comparison of Naive Bayes, Random Forest, Decision Tree, Support Vector Machines, and Logistic Regression Classifiers for Text Reviews Classification", Baltic Journal of Modern Computing, vol. 5, no. 2, 2017
- » [4]S. Marsland, "Machine Learning: An Algorithmic Perspective, Second Edition". 2014

## Kontakt

Duale Hochschule Baden-Württemberg

MHP Management- und IT-Beratung GmbH  
Edmund-Rumpler-Str. 3  
D-60549 Frankfurt am Main  
paul.oesterwitz@mhp.com