Dr. Felix Wick  
VP Data Science / Blue Yonder  
Ohiostr. 8  
76149 Karlsruhe

Computers & Operations Research  
Elsevier Publishing

Submission of

“Demand Forecasting of individual Probability Density Functions with Machine Learning”

Dear Professor Saldanha da Gama, dear Sir or Madam,

predicting future customer demand is central to optimizing the operational decisions retailers face. The prediction of demand is, however, complicated by the fact that each prediction is unique in the sense that it relates to a specific product sold at a specific store on a particular opening day and hence needs to be calculated at this granularity. Furthermore, various factors such as promotions, local weather, events, and many others influence demand. Additionally, since the sales of a given product influences the sales of other products, predicting future demand does not compose in separate newsvendor-type tasks but needs to be modeled holistically. In contrast to approaches to infer the operational decision, i.e. the determination of the order quantity, directly, modeling the demand and deriving the optimal order quantity in subsequent but separate steps has significant advantages, both from a practitioners as well as from a more theoretical perspective.

Machine learning is ideally suited to model the complex behavior of future demand. However, most machine learning approaches only provide simple point estimates, i.e. plain numbers, as output, what severely limits the usefulness of such predictions, because a full predicted probability density distribution for each individual prediction is required to fully optimize the downstream operational decisions. And also most evaluation metrics are geared towards the use of simple numbers. Furthermore, many popular machine learning methods, e.g. deep learning approaches, are “black-box” type algorithms, what can lead to trust issues, especially for operationally critical tasks.

In this work, we propose a method how to use the novel “Cyclic Boosting” machine learning algorithm to model demand on the finest granularity required operationally by predicting a full probability density distribution for each case, while being a “white-box” algorithm, meaning that each individual prediction is fully explainable. Additionally, we show how a range of quantitative and qualitative methods can be used to evaluate the accuracy of the predicted probability density distributions. We use a public dataset to demonstrate how these methods can be used in practice.

I hope that this article can be included in “Computer & Operations Research”.

Sincerely,

Felix Wick