Discussion

* Evaluate the work critically
* Answer questions raised in the introduction
* Limitations
* Further research

Conclusion

* Summarize thesis
* No new information
* Reinforce main points and why are they important
* Clear understanding of the research and its contribution to the field

It has been the main objective of this work to assess the uncertainty which is associated with the PNBD model’s prediction of the individual customer lifetime values. Previous research on this topic was found to be scarce. To fill this gap, several methods that produce prediction intervals were identified in the literature and applied and adapted to the specific context of CLV prediction. The resulting intervals were benchmarked with several metrics against an existing implementation in CLVTools, whereby a special focus was put on Conformal Prediction as a relatively young method that has recently gained importance in the statistical community.

The central contribution of this thesis is to successfully implement methods that deliver intervals with a high reliability on several real-world data sets and hence capture the underlying uncertainty, which no other literature has attempted before. It were especially the Bayesian Approach, Quantile Regression and Conformal Prediction which achieved this reliable coverage, which, however, comes at a cost. Capturing uncertainty adequately leads to wide intervals and little sharpness in the presence of high uncertainty, which is the case in the PNBD model. Two other methods, including the the bootstrap implementation from CLVTools, in contrast, create very short intervals, leading to low reliability but high sharpness. There is no method that provides a valid compromise between these two criteria.

A side objective was to figure out how the newly created uncertainly intervals allow to more accurately distinguish low-value customers from high-value customers. Unfortunately, the created intervals rather hold information about the uncertainty across a data set than for individual customers. Hence, distinguishing customers is not enhanced compared to solely using the point predictor, provided by the PNBD model.

Goal of this work:

* Benchmark bootstrap
* Alternative PIs
* Apply them in marketing

What methods were applied? (CP: Possible in this context)

Which results were achieved?

* First study to propose valid prediction intervals for CLV context
* Bootstrap was benchmarked
* Several methods applied and alternated
* Reliable PIs or sharp PIs
* No applicability for marketing

Weaknesses in the approach

* No method that combines everything or a compromise
* Only 4 data sets
* Past data
* No individual uncertainty

Future research

* Develop methods that internally incorporate individual customer uncertainty, not just their CET
* Use covariates that influence the individual uncertainty