CAPSTONE PROJECT PUBLIC SUMMARY

Assignment Specification

The task was to analyse the performance of the customer care workflow and build methods for effectively predicting Service Level Agreement (SLA) breach events – based on any data from historical performance records.

The company handles service requests in a global ticketing system, which stores information about the requests themselves, monitors breach of contractual SLAs, and provides performance feedback across all dimensions of the company. In order to minimize costs from SLA breaches, the company would like to allocate extra efforts on the most likely worst offender service request. The sooner such a case can be detected during the workflow, the chances of minimizing contractual consequences are higher, hence the business case for the extra effort.

Business Background and Objective

The company hosting this capstone project is a big multinational firm, being one of the leader equipment and services supplier in the IT industry.

Among the offered services, Customer Support is one of the most paramount. Not only due to its high financial contribution to the company's results, but also because of its strategic importance: outstanding in-service-performance of products is the cornerstone of the company's brand, and Customer Support is fundamental in achieving this.

The company has a pretty wide portfolio of Customer Support services, tailored to different customer needs, spanning from basic best effort assistance to immediate and very deep recovery and resolution services, including predictive and preventive services. As of today, majority of Customer Support works are about resolving Trouble Tickets (TT) of various kinds: finding the reported faults and fixing those.

Customer Support contracts are centred around a concept called Service Level Agreement (SLA), which defines what percentage of the Trouble Tickets have to be resolved within what timeframe. Typically these durations vary by Customer, Product, Contract Type and Trouble Ticket Severity.

Customer Support revenues are highly dependent on how the company performs against these SLAs. A typical contract stipulates financial penalties at SLA breaches. For this reason, and also from an image perspective, it is imperative that the company meets its SLA obligations.

As many others, the company has gone through a rationalization and efficiency program the past years, where Customer Support organization have also been impacted. The consequence is that significantly less number of engineers are expected to cope with about the same amount of work as a year ago. As explained above, it is crucial that the workforce reduction must not imply noticeable service (SLA) degradation.

The idea is that if we could predict at the entry of this process section that a given Trouble Ticket would likely breach its SLA, then putting them on some "Fast Track" (i.e. short cutting the normal process and

CEU eTD Collection

sending those tickets directly to the experts), the "Ticket" could be "saved". The more are saved, the better the SLA improvement is, the higher the Customer Support revenues are.

This is the purpose of this capstone project.

Work Summary

Work has been carried out between the 15th of October and the 15th of December, 2018 at the company premises. The final dataset used for the latest analysis and prediction was extracted on the 7th of November, 2018. The data consisted of over 300k Trouble Tickets summary data, created between the 1st of January 2017 and the date of extraction. Each data entry consisted of 62 parameters. The assumption was that SLA failure follow some pattern, possibly to some extent impacted by those factors that are captured by these 62 variables, and this is what I was expected to find and leverage on.

The process followed was what is described in the high level CRISP data project model, except that deployment was not in the scope. The purpose of the capstone was just working out and evaluating the model ("Proof of Concept").

The work was done using R language on R Studio.

I faced a few dilemmas related to the dataset, most notably the followings:

- Definition of the available indicators was not always trivial, nor their relation to Trouble Ticket (TT) Service Level Agreement.
- In certain parameters the data was inconsistent.
- The short free text description of the fault, which I expected to reveal a lot about the problem, often was slang.
- Certain indicators might be changed any time during the life cycle of a ticket. The implication is that what the dataset contained might differed from how it looked like when the ticket was received. Using them would mean that the model is trained on something else than what is available at the time of the prediction.
- How old data should be considered valid for the model, what is the time period that has "good enough" stationarity to be used in a predicting?
- Most parameters are categorical types, some of those with thousands of possible values. Using them in Machine Learning require one-hot-encoding to binary variables, which could "explode" the size of the data.

I addressed these issues during the Data Engineering phase.

In the Data Modelling phase I built XGBoost and Deep Learning binary classifiers using the R "xgboost" and "keras" libraries. I spent substantial effort on building classifiers that handle both typical (i.e. numerical and categorical) features as well as free texts. In the XGBoost models I used shallow NLP methods, while in the Deep Learning models I used Recurrent Long Short Term Memory Deep Learning with GloVe embedding for the text processing and Simple Feed Forward Networks for processing rest of the features. In order to combine the best of both, I created an ensemble model as well, where an LSTM with Glove embedding Neural Net was responsible for the text parts, while XGBoost for the other features, and an XGBoost was put on the top.

CEU eTD Collection

I combined manual and grid-search based parameter tuning. I also tried AWS Sagemaker to tune XGBoost hyper-parameters (on a similar size dataset that hold no business information, whatsoever), but finally I abandoned this track: due to the size of the data the cost of this was exceeding my budget.

Evaluation

I evaluated all models on the latest Trouble Tickets, those that were created after the 30th of August, 2018. These observations were put aside from start in order to makes sure that the evaluation is done on data that the models have not "seen before". The result was promising: prediction on the test set showed that with such prediction the overall number of SLA failures can be reduced to an extent that positively impacts the business. I am convinced that there is even more potential in this approach, which I elaborated in the Technical Summary.

The evaluation consisted of two main parts:

From a Data Scientist perspective I compared the quantitative results of all the models, concluded that while LSTM does a good job on the free text based prediction, overall the performance of XGBoost was slightly superior than the concatenated Neural Networks and the Ensemble model.

Analysis from Business Perspective aimed at justifying the SLA predictor's delivered value. In order to quantify the business value, I did a simple "what if" analysis:

- With the best performing model I run a prediction on the Test Set,
- Quantified the result that matter from a business perspective, i.e. how many tickets are actually SLA breakers (Real Positives), how many of those were found by the model (True Positives), how many tickets the model classified as SLA breakers (Predicted Positives) and would put on "Fast Track". The latter represents the cost side of the business case.
- Weighted costs and business benefits.

The outcome of this simple analysis was that the model classified (and sent to "Fast Track") reasonably low number of tickets, and of those qualified correctly a number that could improve overall SLA levels to the extent that delivers more than decent financial benefits.

Possible Future Development Areas

This capstone was about a Proof of Concept, which demonstrated the potentials and limitations in an SLA predictor. It can deliver positive business impact even in its current shape, but there are some areas where it can be improved, such as:

- Further increasing the accuracy by including more features based on some other logs that were not available for this project,
- Time and industry dynamics should be more taken into account in the models,
- Allocating more resources (e.g. budget to use Sagemaker, time) in order to further tune the models

Finally a last note: it is one thing to predict SLA breakers early on, but the other part is to act upon those. Deployment would also require some changes in the current processes: introducing a Fast Track treatment for likely SLA breakers.

CEU eTD Collection

CEU eTD Collection

CEU eTD Collection

CEU eTD Collection

CEU eTD Collection

CEU eTD Collection

CEU eTD Collection

CEU eTD Collection