Machine Learning in the Real World

Presentation

Uday S Rachapudi
Danske Bank
uday.s.rachapudi@gmail.com
Mobile Number : 9573191475

Parnika Pancholi
Danske Bank
parnika.pancholi@gmail.com
Mobile Number : 7022036018

ABSTRACT

Despite analytics being a top investment priority for CXOs in the recent years, with significant effort spent to identify use cases to leverage data to make proactive decisions and drive efficiencies, organizations are yet to fully realize the value from developing and deploying analytical solutions.

In this presentation, we will talk in-depth about the various challenges and constraints organizations face in consuming analytical outcomes using a case study of customer churn in a leading Scandinavian bank. Specific focus will be on the role a data scientist has to play in overcoming these challenges in order to ensure financial impact of predictive models.

Churn is a pervasive problem across industries as developing a good predictive solution can be tricky especially if there is no clear cut definition of churn.
The presentation will focus both on complexities of churn in mortgages and the end to end aspects of solution design and deployment.

AUDIENCE

The presentation can be attended by

- **Beginners & intermediate practitioner**s in analytics. It can provide useful learnings from the case study, improve their solution design and provide guidelines to go beyond the algorithm and data science related tasks of their daily routines.
- **Business leaders & Product owners** who are working with analytics teams in order to improve their business processes. It can help them understand how to collaborate with analytics teams, how to ensure the solutions developed are implementable and scalable and how they can help maximize business value.

Time – 40-45 mins

INTRODUCTION

Objective

Portfolio management team of a leading Scandinavian bank experienced high outflow of customers. They identified that the existing churn prevention efforts were not working as expected and there was a need for a robust and data driven approach to retain customers.

Methodology

In analytics a lot of time and money can be wasted if a data scientist doesn't spend enough time defining and translating the business problem in analytical terms.
Identifying churn can also be a tough problem to solve for multiple reasons such as:

- Churn behavior might vary from product to product
- Churn in banking is typically passive at a product level
- Churn rate every month may be low but it still translates into a high revenue loss every year
- Lack of an integrated view of what drives a customer to churn

Analysis was done to understand the following instead of directly proceeding to data preparation and modeling:

- Is there any portfolio where churn is high?
- Is there any specific order or product pattern in which customers churn?
- What is the potential gain of implementing a predictive model?

As a result of this analysis, it was identified that 80% of churners first move their mortgages to other banks and subsequently move other products. Hence the scope was narrowed down to addressing churn in the home loan portfolio, based on which a conservative estimate was calculated to create a business case.

It is crucial to establish a ballpark ROI figure taking into account model and campaign performance especially in industries where AI and analytics are not the primary products but used to augment the primary products or services.

Multiple statistical modeling and machine learning algorithms like Random forest, XGboost, Memory based reasoning were explored and evaluated. The technique with the most stable performance was selected for the churn prevention campaign.

Implementation/Consumption-

It is not uncommon that even good models are rendered useless as there are multiple factors that contribute to its failure. As we know, creation of campaign involves multiple stages i.e. identifying-

- The target audience
- The channel
- The message

A model can only decide who should be contacted but if the message or channel is wrong, desired results cannot be achieved. Visibility into the implementation strategy and creating a feedback loop is critical to ensure consumption and right implementation of the models.

Close collaboration with the frontline helped in creating an optimal implementation strategy. It is also crucial to have a pilot phase, allowing enough time to test the model and fine tune it based on the feedback, prior to large scale rollout.

One of the key challenges identified was the messaging. Unlike sales campaigns, communication and rollout of a churn prevention initiative is not as straightforward as a single strategy cannot be used to target all customers. An iterative process was adopted to prepare "conversation starters" that can improve the conversation quality for the frontline.

It is also crucial for stakeholders and the data scientist involved to have a road map to measure performance and align on impact post execution. Hence, scientifically significant test and control groups were created with a strict cool off period to measure campaign performance. During measurement, it appeared that many business rules were applied to further filter the leads, reducing the potential impact. It was then decided that there is a need to closely align campaign execution with model development to maximize the potential benefit.

OUTCOMES/CONCLUSION

Data scientists understand what it takes to successfully build and implement machine learning models in the real world

Key takeaways for audience:

- Various challenges data scientists face in the practical world?
- How do we overcome them through design thinking (with an actual case study)?
- How should business and data scientist collaborate to avoid common pitfalls?

PARTICIPATION STATEMENT

I commit to attend the conference if my submission gets selected.

BIO

Parnika Pancholi is a Data Scientist who has spent 5 years in the banking and financial services domain enabling organizations to apply advanced analytics capabilities to create business impact. In Danske Bank, she is part of the flagship churn prevention team that developed statistical models to prevent revenue loss by identifying customers who are likely to move their business away from the bank. This required close collaboration with business stakeholders and frontline to ensure alignment between business priorities and analytical deliverables. Based on her experience, she is well positioned to ensure faster outcomes for customer retention across different industries.

Uday S Rachapudi is a Chief Data Scientist with over a decade of experience in enabling data driven decision making for multiple Fortune 500 companies across the globe. In Danske Bank, he leads the churn prevention initiative across all markets and business units, which is a key strategic priority for the bank. His role is to ensure validation of the business ask, alignment of the business problem to an analytical solution, a robust data foundation, application of the right statistical or machine learning techniques and recommendations on the right value proposition and Go-to-Market strategy to prevent customer attrition.

REFERENCES/BIBLIOGRAPGHY

1. 3.https:/www.mckinsey.com/industries/financial-services/our-insights/analytics-in-banking-time-to-realize-the-value
2. The Chartered Institute of Marketing, Cost of customer acquisition versus customer retention (2010)
3. https://hbr.org/2014/10/the-value-of-keeping-the-right-customers
4. https://bi-survey.com/challenges-big-data-analytics
5. https://thesai.org/Downloads/Volume9No2/Paper_38Machine_Learning_Techniques_for_Customer_Retention.pdf
6. https://aip.scitation.org/doi/abs/10.1063/1.4887705
7. Sadaf Nabavi, Shahram Jafar, Providing a Customer Churn Prediction Model Using Random Forest and Boosted Trees Techniques (Case Study: Solico Food Industries Group, Journal of Basic and Applied Scientific Research, volume 3, issue 6, pages:1018-1026, 2013
8. Keramati, A., Ghaneei, H. & Mirmohammadi, "Developing a prediction model for customer churn from electronic banking services using data mining", S.M. Financ Innov (2016) 2: 10.https://doi.org/10.1186/s40854-016-0029-6
9. Breiman,L. "Random Forests", Machine-learning Volume 45, Issue 1, pp 5–32, 2001