Churn Prediction Model for Effective Gym Customer Retention

Jas Semrl
Student, Dept. of Computer Science,
University College London
Email: jas.semrl.15@ucl.ac.uk

Alexandru Matei
Digital Health Manager,
Nuffield Health
Email: alexandru.matei@nuffieldhealth.com

***Abstract*—In the fitness industry, rolling gym membership contracts allow customers to terminate a contract with little advanced notice. Customer churn prediction is a well known area in Machine Learning research. Many companies, however, face a data science skills gap when trying to translate this research onto their own datasets and IT infrastructure.**

In this paper we present a series of experiments that aim to predict customer behaviour, in order to increase gym utilisation and customer retention. We use two off-the-shelf machine learning platforms, so that we can evaluate whether these platforms, used by non ML experts, can help companies improve their services.

***Index Terms*—customer churn, gym attendance, machine learning**

I. Introduction

Predicting customer behaviour is of strategically important value in the lifestyle and fitness industry, and it is a topic that has been addressed by Machine Learning researchers [1], [2] and the industry [3].

However, many companies still face a shortage of data scientists available to recruit, able to translate ML research into practical solutions to the companies needs.

To address this gap, emerging off-the-shelf machine learning platforms claim to make basic predictive models easier to build, deploy and use, even by non-ML experts.

In this paper we evaluate two such platforms - Azure ML [4] and Big ML [5]. We use a standard question for many service industries - churn prediction - and a real data set that combines demographic and behavioural data of gym users.

We have chosen Azure ML as a representative example for the large cloud services providers. Microsoft's ML offering is well integrated with their other services (e.g. cloud storage, web hosting) and this may provide additional benefits.

We have chosen BigML as an example of a startup whose main mission is to make machine learning accessible to a wider public. Their focus is sharper, and we expect their platform to be more consumer centric.

II. Related Work

The work presented in this paper is based on the UCL, Nuffield Health and Microsoft project Transpire, which demonstrated a prototype dashboard with predictive models incorporated in its notification system [6].

Our evaluation of the platforms follows a structure used elsewhere in the literature [7]. We, however, approach the two platforms as non ML experts - a perspective that would be common in many companies, but not as often discussed in the research community.

III. Case Study: Gym Membership Churn Prediction

To investigate the two platforms, we have selected a fairly standard application domain for services industries: predicting customer churn in a contractual setting. Specifically, we consider gym membership contracts that may be renewed on a monthly basis.

We address churn from two complementary perspective. First, from a contractual perspective, we ask whether (in a given time period) a customer contract ends. Second, from a behavioural perspective, we ask whether (in a given time period), a user will not use the gym at all.

The available data was a sample of about 37,000 anonymous gym members' attendance record (membership card swipes). All members were on a monthly contract and we observe them from the start of their contract for the following 12 weeks. Each swipe was recorded as an individual row in the dataset - including date, time and location of the swipe. In addition to that, limited demographic data about the users was available, such as the age and the gender. Details about their contract included their start date, contract end date (if any), payment plan and monthly payments.

IV. Machine Learning Tasks

Observing members that have ended their contract within the first three months, most of these early cancelations happen weeks 4 to 8 - as shown in Fig. 1.

Hence, we propose a binary classification task, that based on user demographics and their gym attendance behaviour in the first month of the contract (weeks one to four), predicts whether a member's contract will end in month two (weeks five to eight).

The second machine learning task was a multi class clasification: predict the gym utilisations level for a member in the fifth week of their contract, based on their demographic data and their activity in the first four weeks.

978-1-5386-2366-4/17/$31.00 ©2017 IEEE

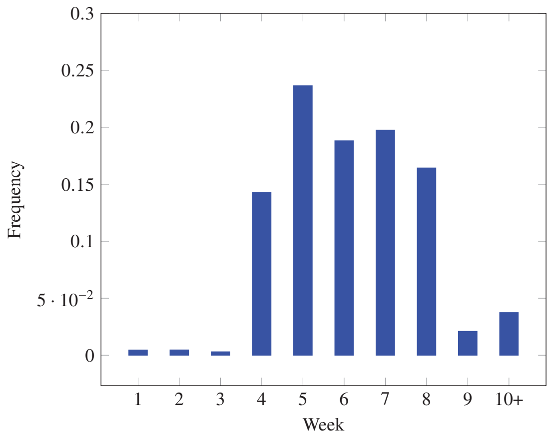

Fig. 1. Leaver Frequency per Week of Contract

In this case we define three classes to characterise gym utilisation: no swipes in the whole week (no attendance), one or two swipes (low attendance) and three or more swipes (high attendance).

V. Platform Evaluation Design

Based on previous literature [7], we evaluate the two platforms based on a number of criteria:

- data pre-processing capabilities
- whether the platforms offer algorithms suitable to our specific tasks and what is the performance of our models on each platform
- the learning and development time required to train and evaluate the models

VI. Results - Data Processing Capabilities

Before training the model, additional data manipulation was required. Firstly, those that left before the fifth week of their contract and those that have joined too recently to have a 12-week record. This reduced the number of users to about 30,000. The swipes of those users were then aggregated by week per user. This resulted in a table with one row per user, their demographical data and their per week attendance of the first four weeks.

Furthermore, two additional columns were added, to enable the classification tasks - one column to describe the status of the member at the end of the first three months (leaver or non leaver) and one column to describe their utilisation level in week 5 (none, low, high)

In this area, the two platforms follow different approaches. Azure ML allows data manipulation via SQL queries, R or Python scripts as well as a graphical interface for performing basic tasks such as filtering. BigML offers their own scripting language for data generation and filtering.

We consider AzureML to offer a broader range of capabilities, suitable to software engineers, but also to database administrators (who can write SQL transformations) and non-technical users (who can use the basic graphical interface).

VII. Results - Model Performance

One of the major factors in evaluating the platforms is the accuracy of the predictive models. To evaluate it, the data was randomly split into two parts. The first part (consisting of 80% of the records) was used to train the two models and the remaining part was used to score and evaluate the trained models.

The following algorithms were used in the two platforms:

AzureML	BigML
Decision Forests	Decision Forest
Decision Jungle	Boosted Trees
Logistic Regression	
Neural Network	

Since predicting customer churn was a two-class classification task, the measurement of the overall success was measured by the area under the ROC curve. We are interested in the false positive rate for leavers. This is the most meaningful measure from a business perspective, as it allows the company to focus its retention strategy.

Azure ML's best performing algorithm is the Logistic Regression with the AUC of 0.734. While Decision Jungle and Neural Network preform nearly as well (with AUCs of 0.732 and 0.726 respectively), the Decision Forest yields a significantly lower AUC of 0.675.

On the other hand, Predicting the customer activity was a multi-class task and the measurement of success was the accuracy. The Algorithms in AzureML performed similarly to the churn task. The best performing was the Neural Network with the Accuracy of 74.6% and the worst performing was the Decision Tree with the accuracy of 70.8%.

BigML offers a hyperparameter optimisation option for Decision Forests, called SMACdown [8].

For the first task, the best performing decision forest has an AUC of 0.748 . The false positive rate for leavers is 38.27% and the accuracy is 66.8%. For the second task, the accuracy we obtained for predicting the members with no swipes was 76%.

The initial results on both platforms have generated sufficient interest within the company to warrant additional investigation. As such, we consider the platforms have delivered value, by allowing the company to identify promising areas for further investigation.

VIII. Results - Platform Usability

Both platforms allow for easy prototyping of simple machine learning algorithms. With BigML, fewer steps are necessary to train and inspect the models. It also provides an easy way to compare the evaluations of different models, by plotting them on the same ROC graph.

AzureML, on the other hand, allows for a greater variety of algorithms, while also allowing for custom R and Python modules to be imported. It also provides more tools for data manipulation and cleansing.

IX. Conclusion

The models we developed on both platforms perform better than random or mean, so they can already provide a business benefit.

A future step is to add additional engineered features to our dataset. Inspired by behaviour change theory, we will look at whether consistency in behaviour, from one week to the next, is a predictor of engagement.

References

[1] E. Ascarza and B. G. Hardie, "Modeling churn and usage behavior in contractual settings," in *working paper, London Business School*, 2009.

[2] C.-F. Tsaia and Y.-H. Lub, "Customer churn prediction by hybrid neural networks," *Expert Systems with Applications*, p. 12547–12553, 2009.

[3] J. Burez and D. V. den Poel, "Handling class imbalance in customer churn prediction," *Expert Systems with Applications*, p. 4626–4636, 2009.

[4] "Azureml," https://studio.azureml.net/, accessed: 2017-06-04.

[5] "Bigml," https://bigml.com/, last Accessed: 2017-06-04.

[6] M. de Fontenay, M. Afsharmoqaddam, and J. Semrl, "Transpire," *UCL Computer Science 2017 Undergraduade and Postgraduate Abstract Proceedings*, p. 18, 2017.

[7] I. Almeida, "Machine learning services benchmark," https://www.slideshare.net/papisdotio/machine-learning-services-benchmark-ins-almeida-papis-connect, 2016, accessed: 2017-08-01.

[8] "Smac," http://www.cs.ubc.ca/labs/beta/Projects/SMAC/, accessed: 2017-08-01.