CLIENT DISENGAGEMENT

PURPOSE

To develop machine learning algorithms to model client disengagement and to assess strategies for leveraging the model by their potential impact on the profit of the company.

Data Used: Data from warehouse retrieved on 04-10-2018.
The following date ranges were used for the training, pre-holdout, holdout and test data.

- Training: 03-01-2015 to 10-05-2017
- Pre-holdout: 10-06-2017 to 12-05-2017
- Holdout: 12-06-2017 to 02-04-2018
- Test: 02-05-2018 to 03-06-2018

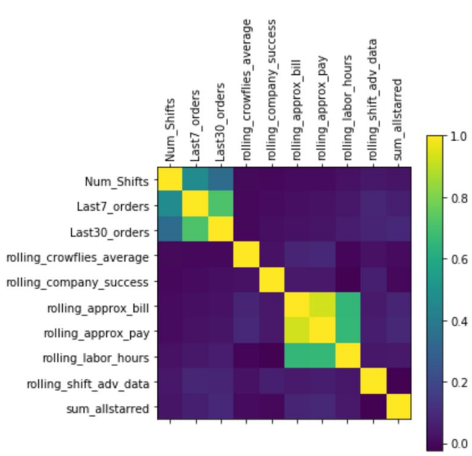
QUICK SUMMARY

Results validate the proof of concept. The best model removed 10% of uncertainty in whether or not a client is disengaging. This research suggests that up to $10,000 per disengaging client per month are left on the table due to loss of business. This report concludes with an outline of potential applications leveraging the model.

Ultimately, the model will need improvement. Potential additional strategies to improve its effectiveness include:

- Feature extraction
- Grid search and cross validation
- New features: e.g. demand and supply, shift demographics, client ratings, economics and geographic factors.

Correlation between predictor variables



There was significant interaction between past orders and the number of ordered shifts. There was also some amount of correlation between the labor hours and the pay variables.
These correlations can be used for future strategic approaches to reduce dimensionality by feature extraction which I feel could help in optimizing model performance.

Outline

A. **Methods**
B. **Results**
C. **Conclusion**
D. **Questions**
E. **Future Directions**
F. **Appendix**

A. Methods

To identify patterns of client disengagement, I began by coming up with a threshold on the inter-order delays to define client disengagement. After deciding on the features that would be used as inputs to model client disengagement, different classification algorithms were used to predict which client is likely to disengage in the near future. The impact of these clients on the profit of Shiftgig was also analyzed.

Feature Used for Predictive Modeling : (Feature Engineering)

Temporal Features

- **rolling_labor_hours:** Average number of hours a specialist has worked in previous 3 shifts.

Economic Features

- **rolling_approx_bill:** Average amount in dollars paid by the client to Shiftgig for previous 3 shifts.
- **rolling_approx_pay:** Average amount in dollars Shiftgig pays to specialists for previous 3 shifts.

Geographic Features

- **rolling_crowflies_average:** Average distance in miles from the specialist location to venue location for previous 3 shifts.

Demand and Supply Features

- **is_reconciled_needed_primary:** Boolean variable indicative of whether a client placed an order for a shift on a particular day
- **Num_Shifts:** Sum of Boolean True of **is_reconciled_needed_primary**

Miscellaneous Features

- **sum_allstarred:** Number of shifts all-starred in preceding 3 shifts.
- **rolling_shift_adv_data:** Amount of days from shift created to shift start for previous 3 shifts.
- **Last7_orders:** Num of shifts with a rolling window of 7 days
- **Last30_orders:** Num of shifts with a rolling window of 30 days
- **Rolling_company_success:** Average success rate for previous 3 shifts.

NOTE: Rolling = A sliding window of preceding 3 shifts was used on a per company basis.

Feature Used for Deciding Impact

- **Past60_profit**

NOTE: These above variables were based on an approximation from bill rates and pay rates.

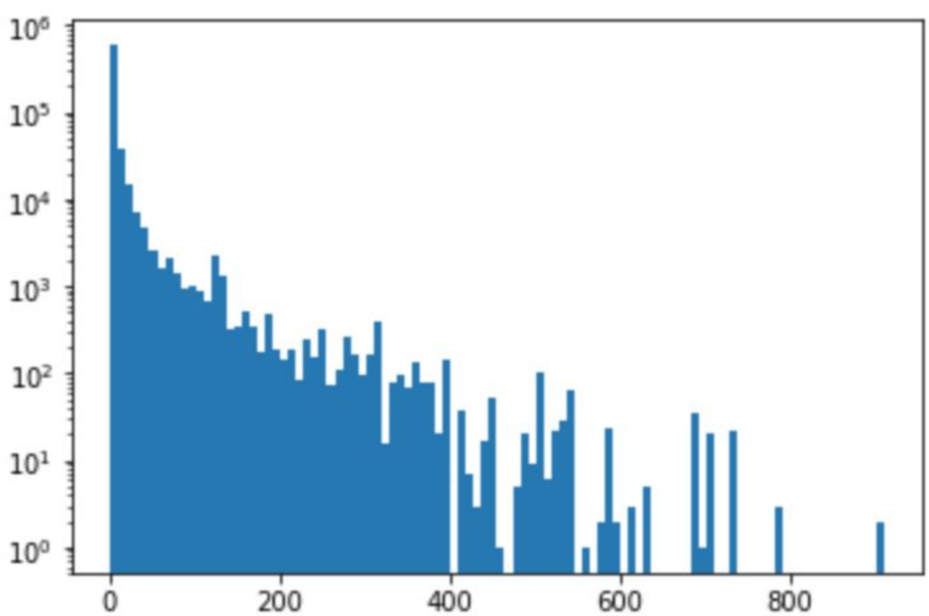
B. Results

I. Defining Churn

Churn and lifetime will be based off of a client's ordered shifts and the periods of inactivity between placement of shift orders by the client. The time between orders (creation) are denoted as Inter-Order Delays (IODs). Churn is determined based on a churn-threshold applied to the IOD's. The first IOD greater than the churn-threshold indicates that the worker has churned, and the client's lifetime is determined as the time from the start of their first order for a shift to the end of the last order of shift just prior to churn.

If all IOD's are under the churn-threshold, the worker is considered not churned and their lifetime is considered unknown.
Churning is have placed an order in the last 30 days and will not place in the next 30 days
Not Churning is have placed an order in the last 30 days and will place in the next 30 days
Churned is not placed an order in the last 30 days and will not place in the next 30 days

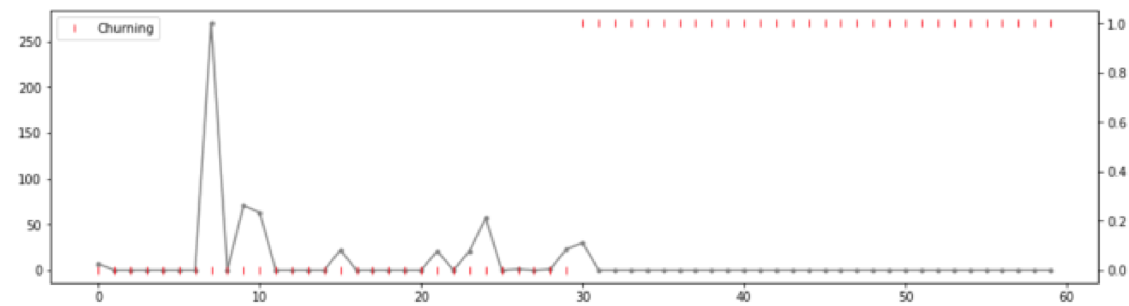
Distribution of IODs

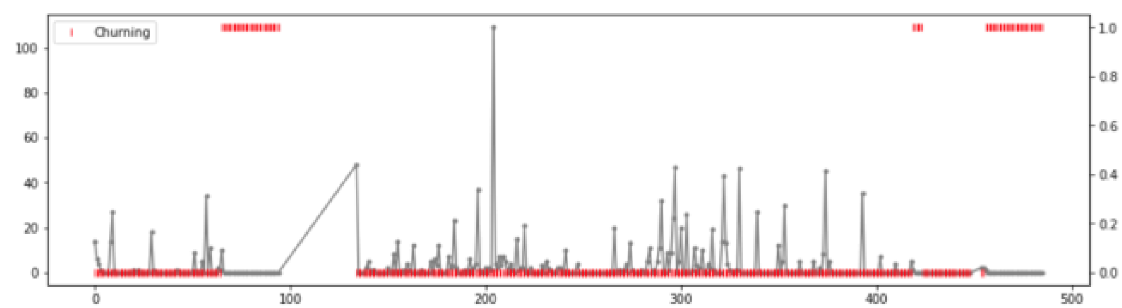


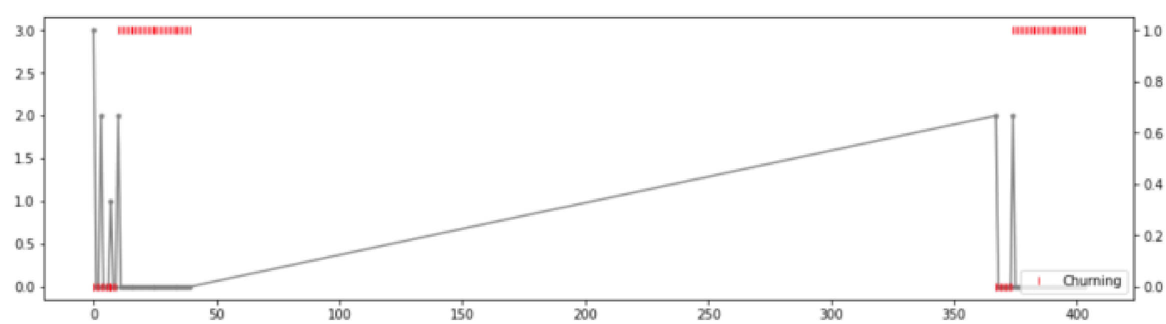

```
('95 percentile iod:', 28.0)
```

The sharp peak in the first bin of the histogram corresponds to the numerous clients that placed an order for only a single shift and whose lifetimes are thus the length of that shift.
Ultimately 95% cut off of the IODs will be considered for the churn-threshold which boiled down to 30 days a round numeric for convenience. 95 percentile indicates that 95% of the time clients have placed an order within or less than 28 days.

II. Client Behaviour







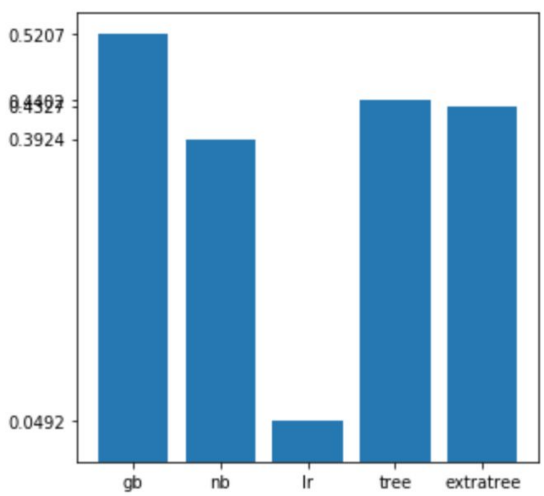
In the above figures, the grey points show the cumulative number of orders for shifts a client has placed (left Y axis) for certain number of days between orders (X axis). The red markers show whether or not a client is disengaging (right Y axis). 1st client shows a simple behaviour of a client. If a client is placing orders with us, he is not disengaging and vice versa. The second client is an active client while the third client is a seasonal client. The third client places orders literally after an year. Since the definition one churn is a difficult one to define because of variability in the placement of orders, the behaviour of clients in different epochs is an interesting attribute to look at with respect to disengagement.

III. Model Selection

The following 5 models were chosen to conduct a quantitative analysis of which clients are churning and which are not:

- Naive bayes
- Logistic Regression
- Decision Trees
- Extra Tree Classifier
- Gradient Boosted

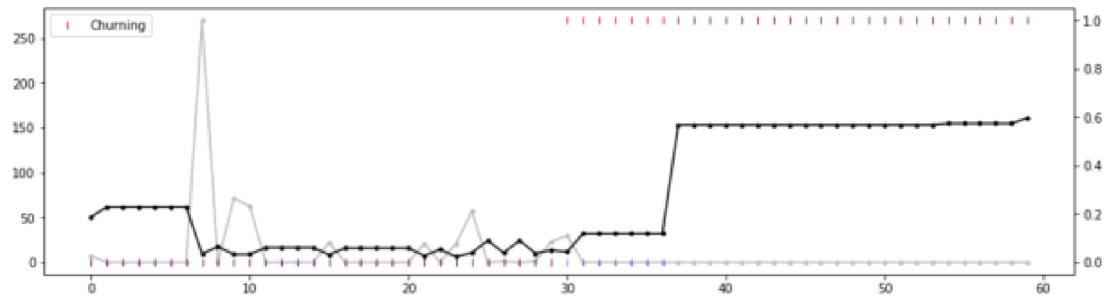
The best model was the Gradient Boosted. A data science performance metric called the F2 score was used to evaluate model performance.

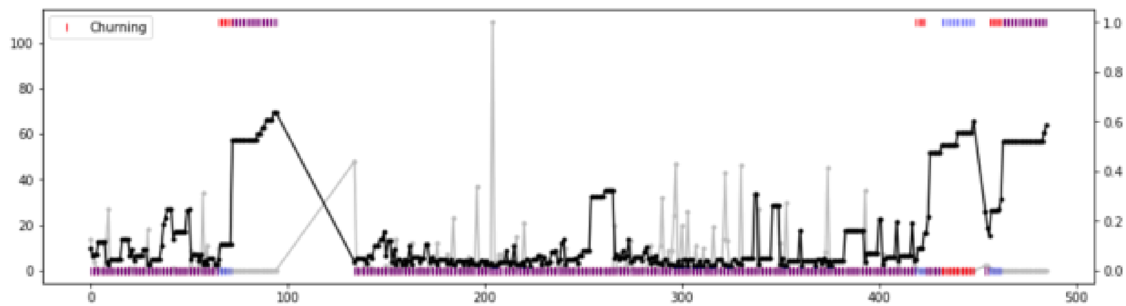


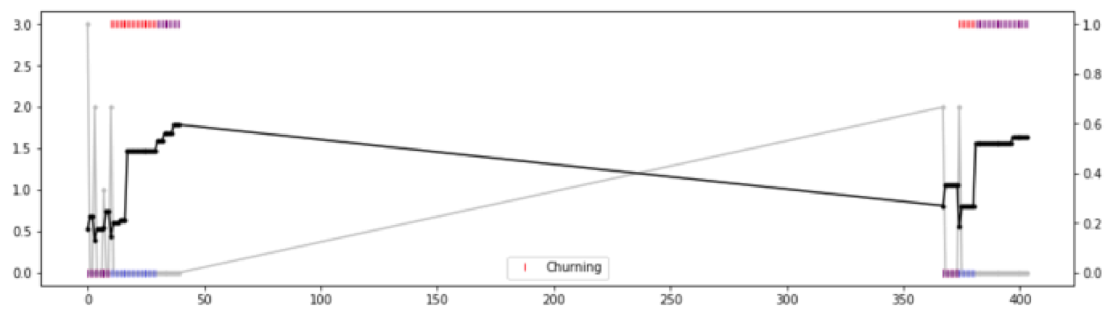
Considering Shiftgig is currently at a low-information state being a startup, it would be more fruitful in aiming to reduce the amount of false positives (clients that are wrongly predicted to be disengaging) . This will help to reduce costs incurred through irrelevant re-engagement campaigns. For this purpose, the F-score was chosen.

IV. Best Model Qualitative Evaluation

The model gives its estimate of the probability of disengagement which can be used to analyze the behaviour of the client.







From the first client and the second client it was seen that the model is able to predict clients who are not disengaging most of the time. In the case of the first client there was a delay of 7 days in predicting that the client is disengaging, but it was earlier in prediction by about 23 days considering the definition of churn in 30 days. In the case of the second client the delays in prediction during the 100th and 475th days were almost by the same amount. But during the 475th day's the client placed a very small order which confused the model and made it to err a bit.

In case of the 3rd client, the model delayed its prediction in the first epoch, but it caught up and reduced the delay in the second epoch. So on the whole, the model was able to make good predictions, but required further A/B testing to evaluate its performance.

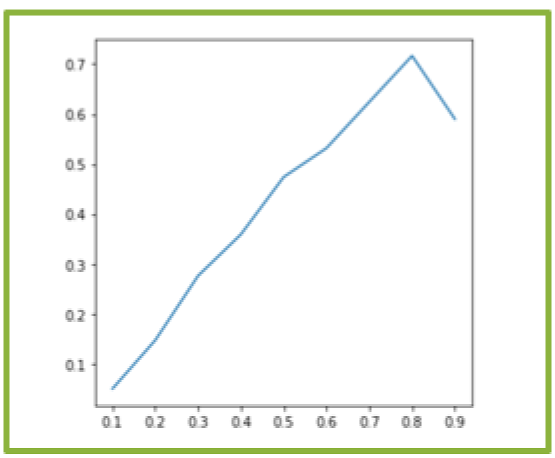
V. Best Model Quantitative Evaluation

```
[[36401  4836]
 [ 9093  6111]]
             precision    recall  f1-score   support

      False       0.80      0.88      0.84     41237
       True       0.56      0.40      0.47     15204

avg / total       0.73      0.75      0.74     56441

(array([0.80012749, 0.55823513]), array([0.88272668, 0.4019337 ]), array([0.8153871 , 0.51795159]), array([41237, 152
04]))
```

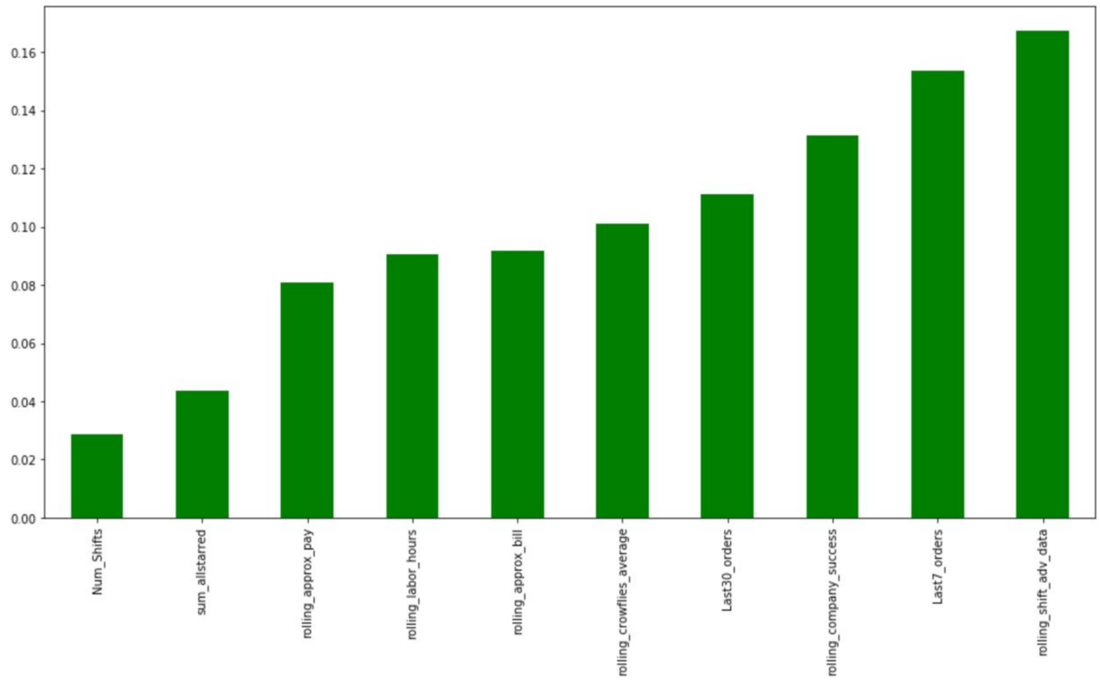
The gradient boosted model gave an accuracy if 73%. This number may be great and pleasing to the eyes, but the model is not fully implementable given the fact that 73% is just a little bit better than a naive prediction of saying most of the times the clients are not disengaging.

To get an assurance and trust the model for future purposes, a test on uncertainity removal was done through the concepts of mutual information and entropy between the actual and predicted values of pre-holdout data. It was found that the best model was able to remove 10% of the uncertainty.
Uncertainty = Mutual Information / Entropy

We must give credit that when a model predicts that 50% probability of the client churning, the model's estimate is on the same lines as that of the actual probabilities.



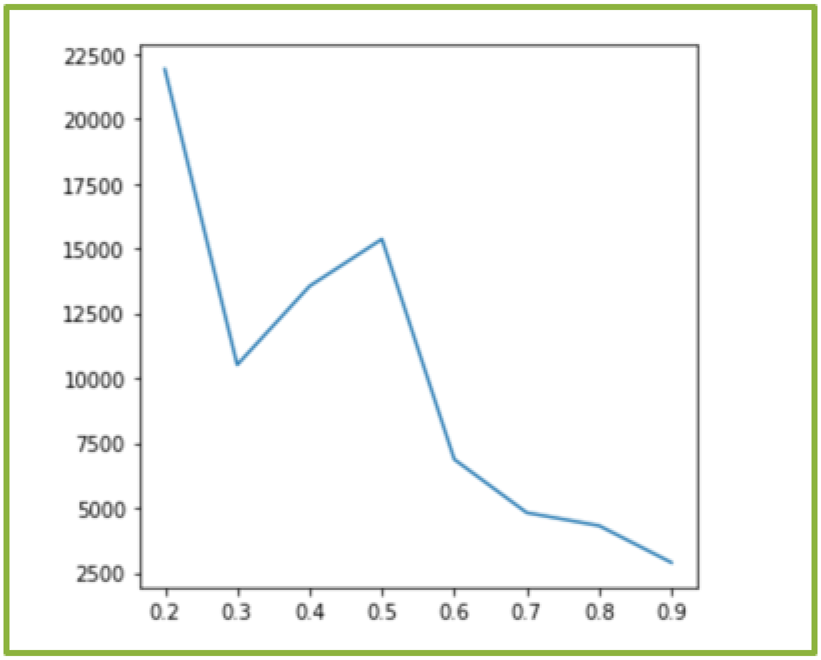
Following is the feature importance plot given by the best model. It shows that clients that have placed an order in the last 7 days and who place an order in advance are less likely to disengage.



C. Conclusion

Considering the observation that past orders do have an impact on client disengagement, it is a good way to say that if a client continues to place orders with us in the same way he has done in the past, the profit that we earn from the client will be the same as the profit that we earned from the client in the past. If the client disengaged, then the amount lost would be same as the amount earned,

Trusting the model, a further analysis of the disengaging clients on the past 60 profit of Shiftgig was conducted. It was observed that when the model's estimate of the probability of churn is 50% or 60%, we lose about 15000 and 7500 USD dollars respectively. It is a natural tendency to look that the clients who are likely to churn 80% or 90% of the time. But we must also look at the fact that the profit that we tend to lose from such clients is very minimal, almost in the range of 2500-5000 USD dollars. So it is better to target those clients from whom we are likely to lose more profit.



D. Future Directions

- Maintaining and optimizing client disengagement model
- Extend to specialist disengagement
- **Use the client disengagement model to power an automated retention messaging campaign**
- **Develop a model to directly assess how messaging impacts retention**
- Build a similar model that predicts profit directly
- Developing a client re-engagement model
- A RNN model that tracks behaviors in the app to predict a client placing orders
- A RNN model that tracks behaviors in the app to predict a specialist picking up a shift

We could use the client disengagement model to come up with who are the client we need to cater to with an automated messaging system that would trigger a tailored re engagement campaign. The results of this model could serve as an input to another develop model which would cater to analyzing the results of the re-engagement campaign and how the re-engagement directly impacted the retention of the clients. Two scenarios could be be used to test the performance of the model: One group of clients on whom we don't run the model and see their disengagement/re-engagement behaviour. Another group of clients on whom we run the model and analyze their behaviour towards re-engagement.