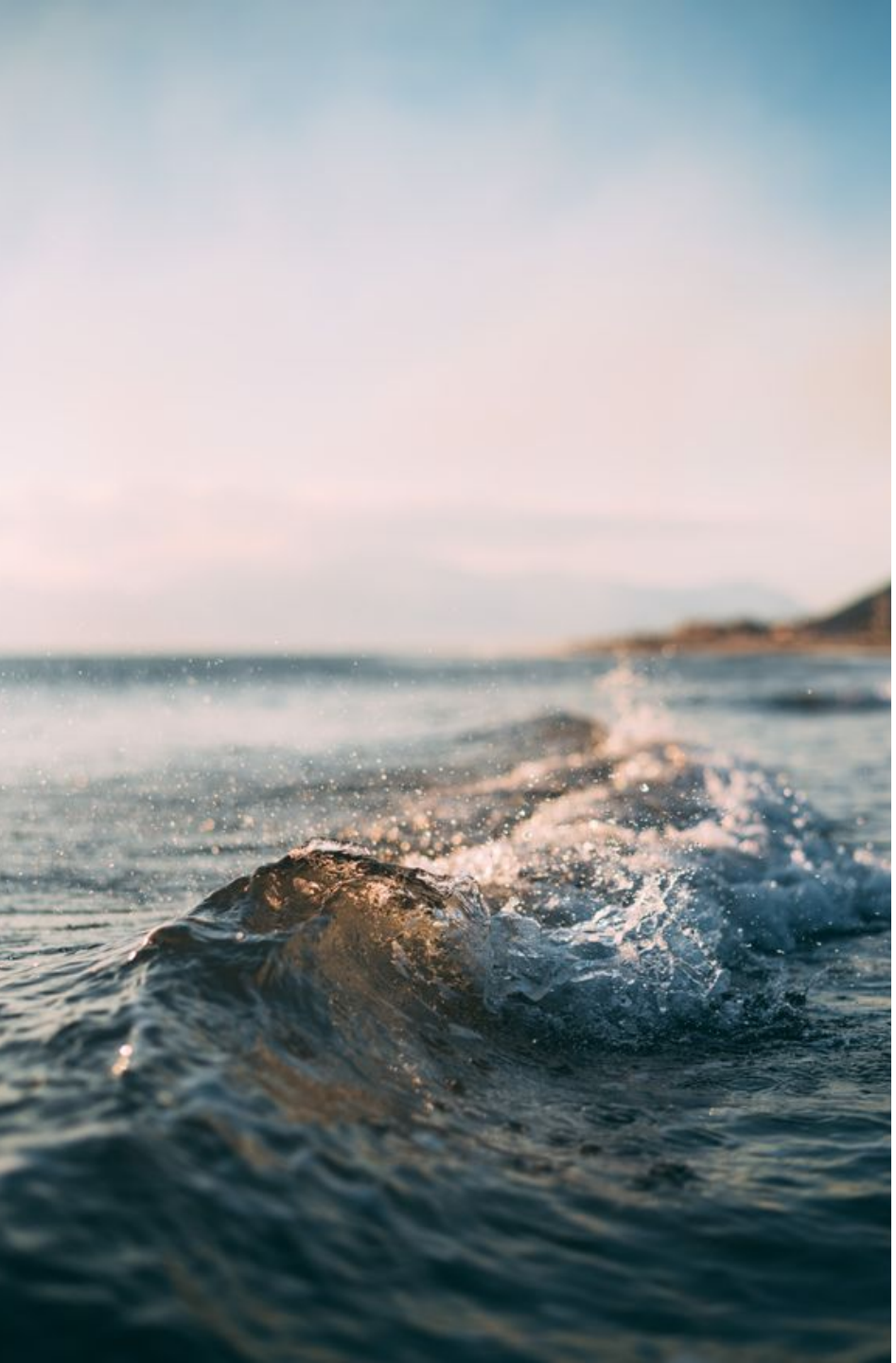
Churn Analysis for Telco Customers

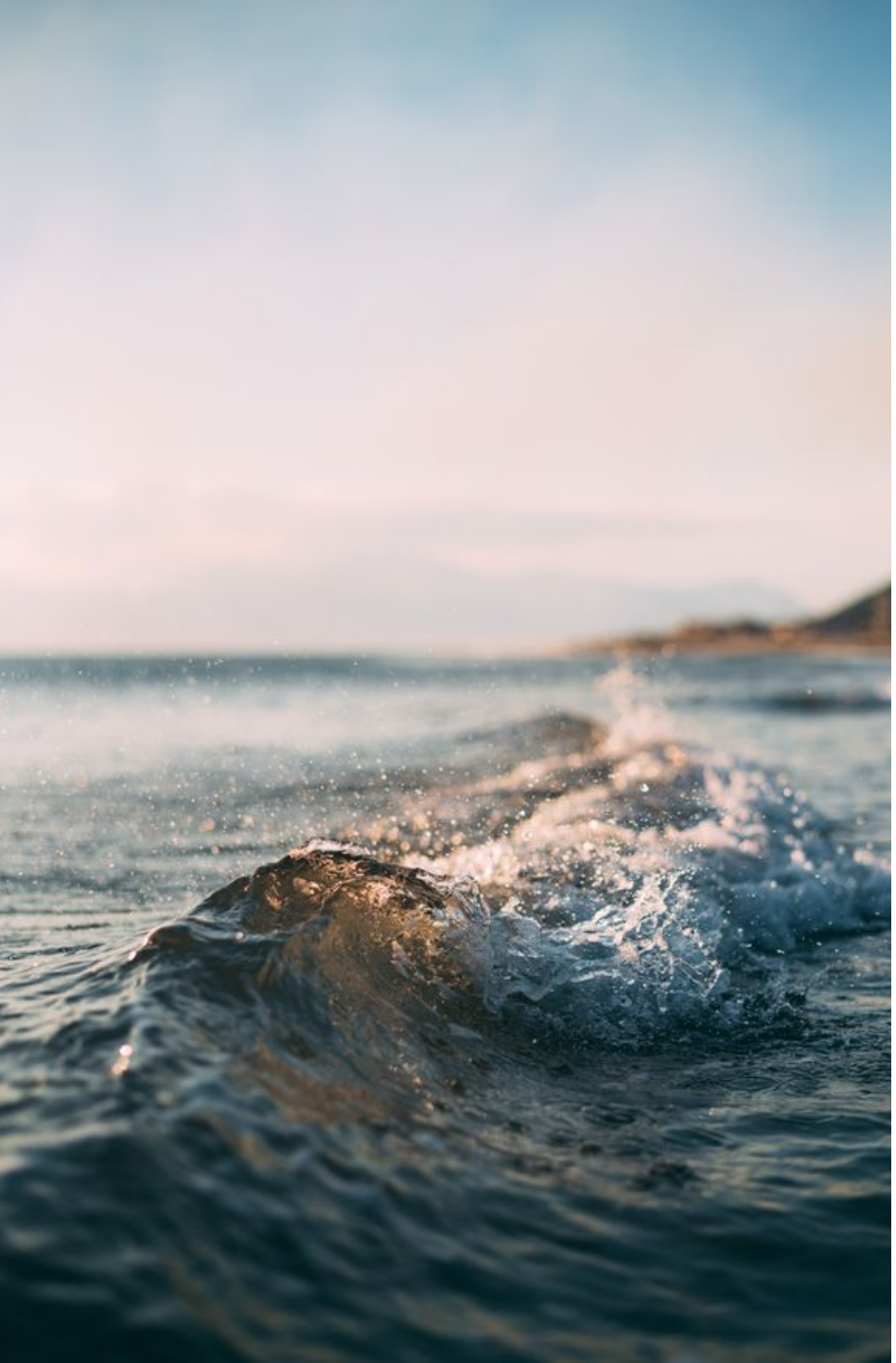


Outline

- Introduction and problem definition
- Exploratory data analysis
- Machine learning
- Conclusions and recommendations

Churn Analysis

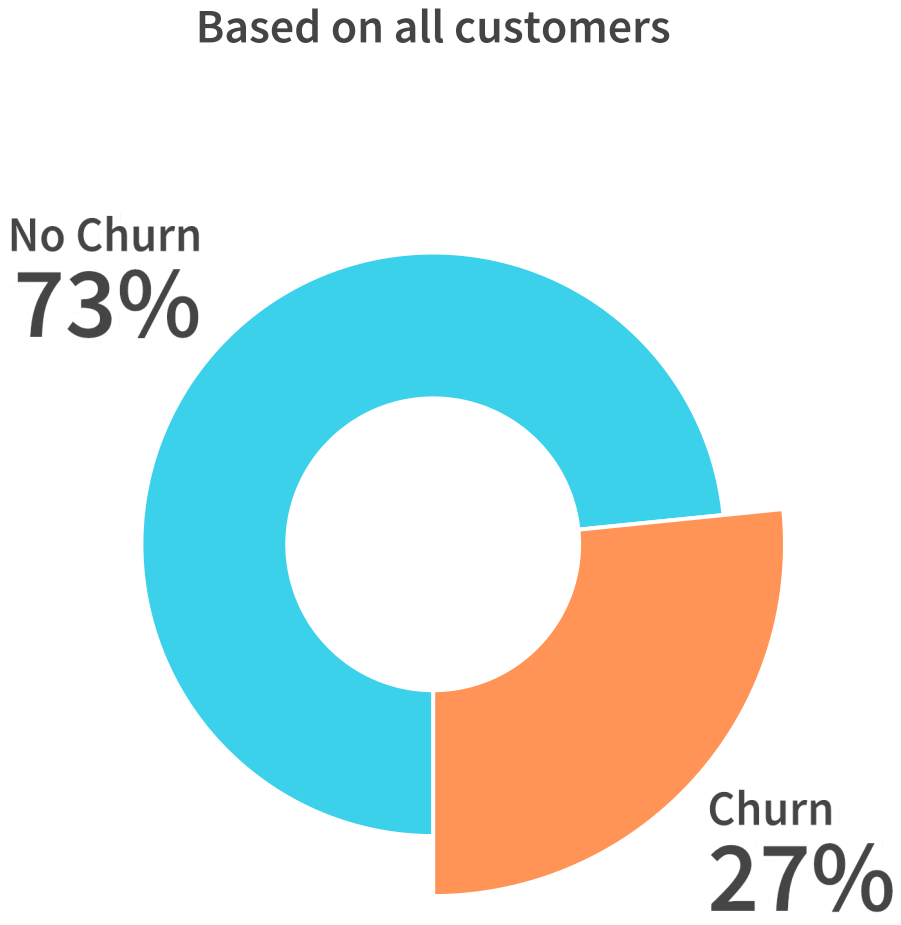
- Available dataset with information for over 7000 customers of telecom
- Data include subscribed services, add on products, some demographic information and service charges
- Can we use that information to determine
  1. Which customers are most likely to churn?
  2. What can we do to reduce churn?
- Machine learning used to determine important factors affecting churn



Outline

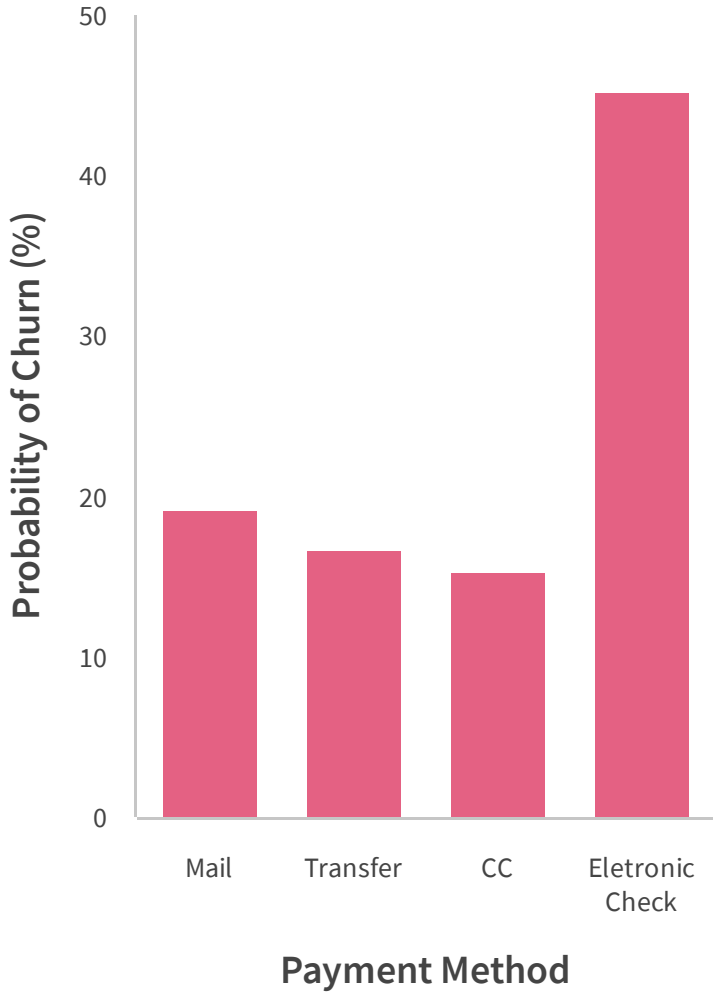
- [x] Introduction and problem definition
- [ ] Exploratory data analysis
- [ ] Machine learning
- [ ] Conclusions and recommendations

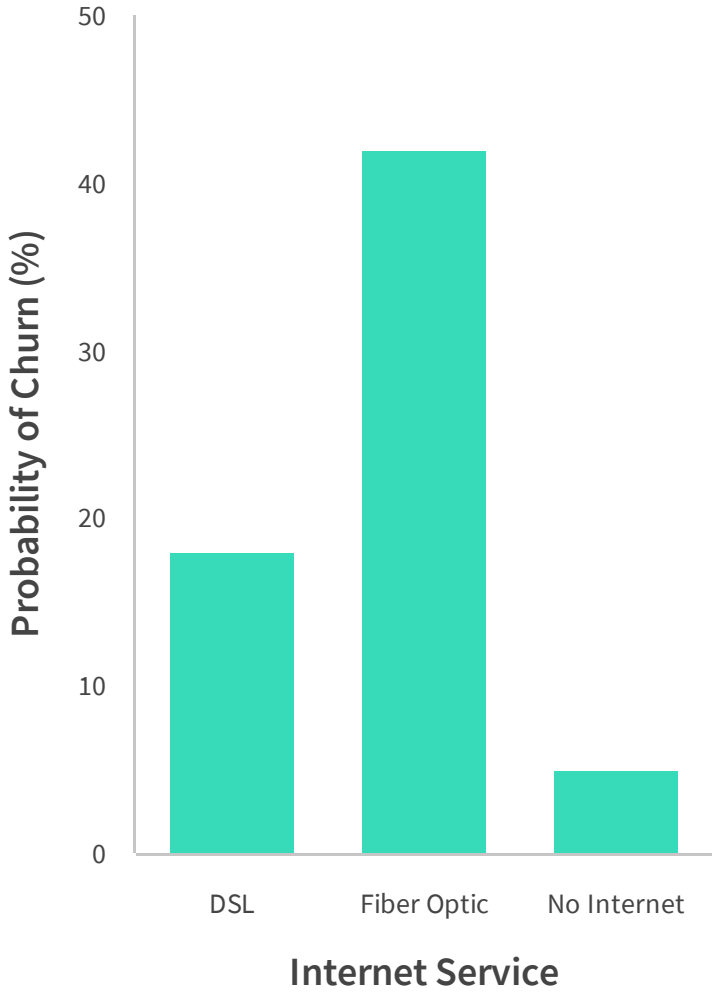
Exploratory Data Analysis

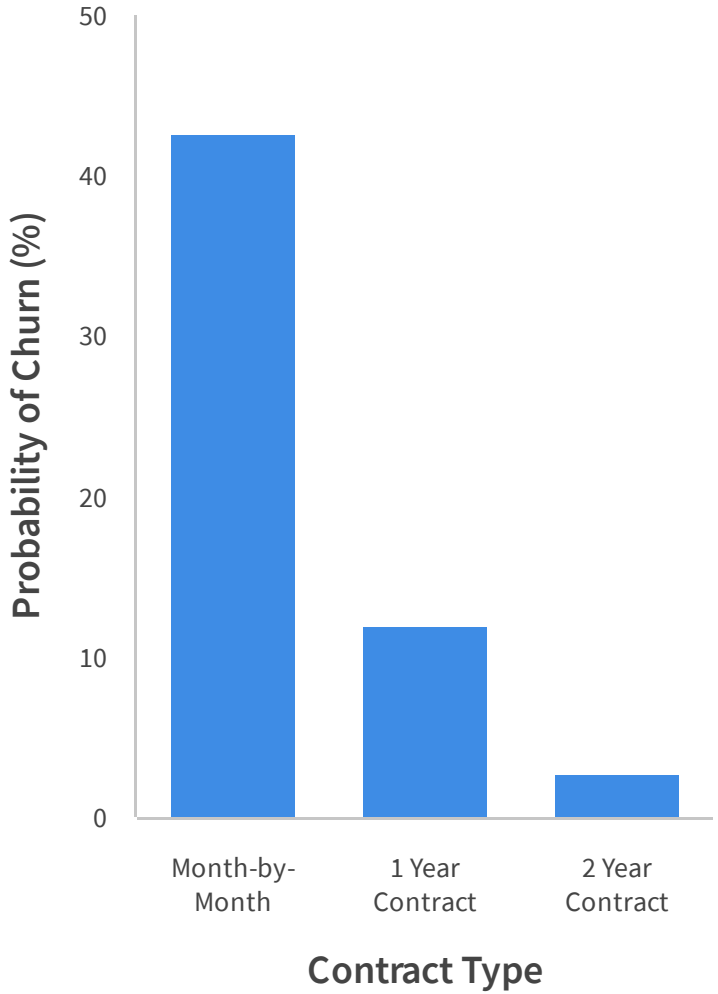


- Dataset was cleaned to remove any errors
- Overall churn is quite high at 27%
- Small churn for customers with telephone only services
- Little difference in churn for customers with or without streaming services

Overall Churn and Influence of Some Features

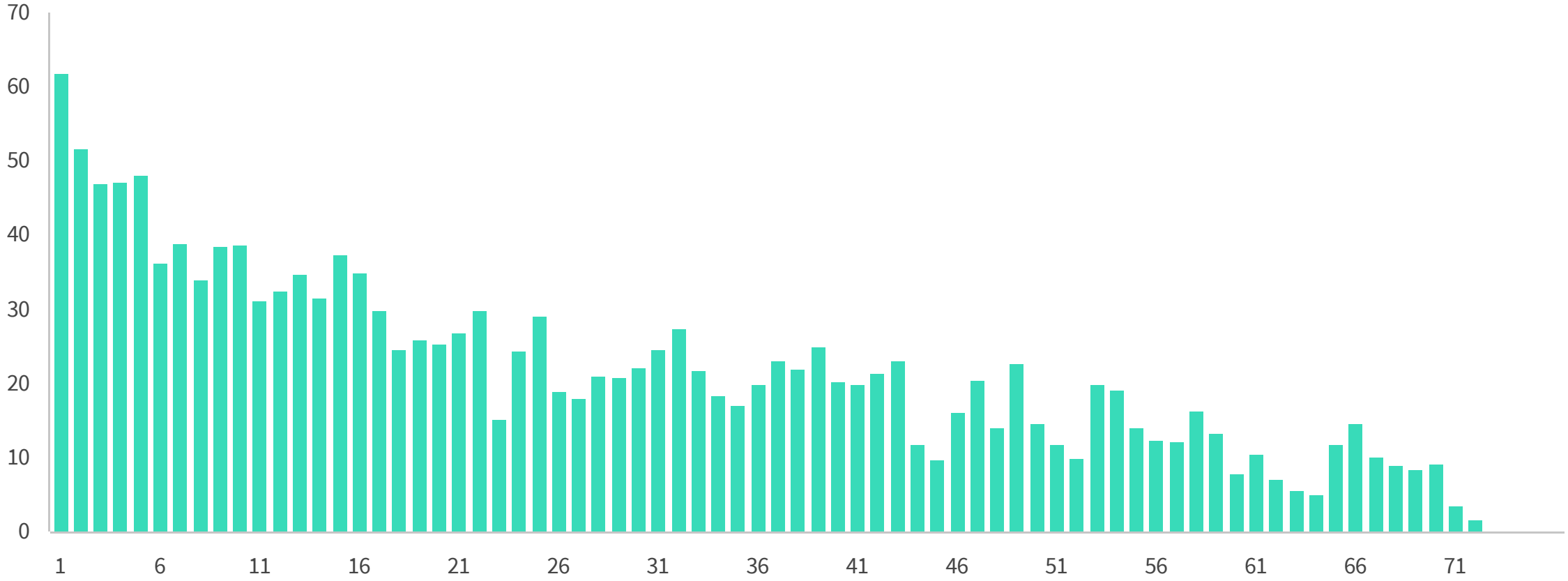


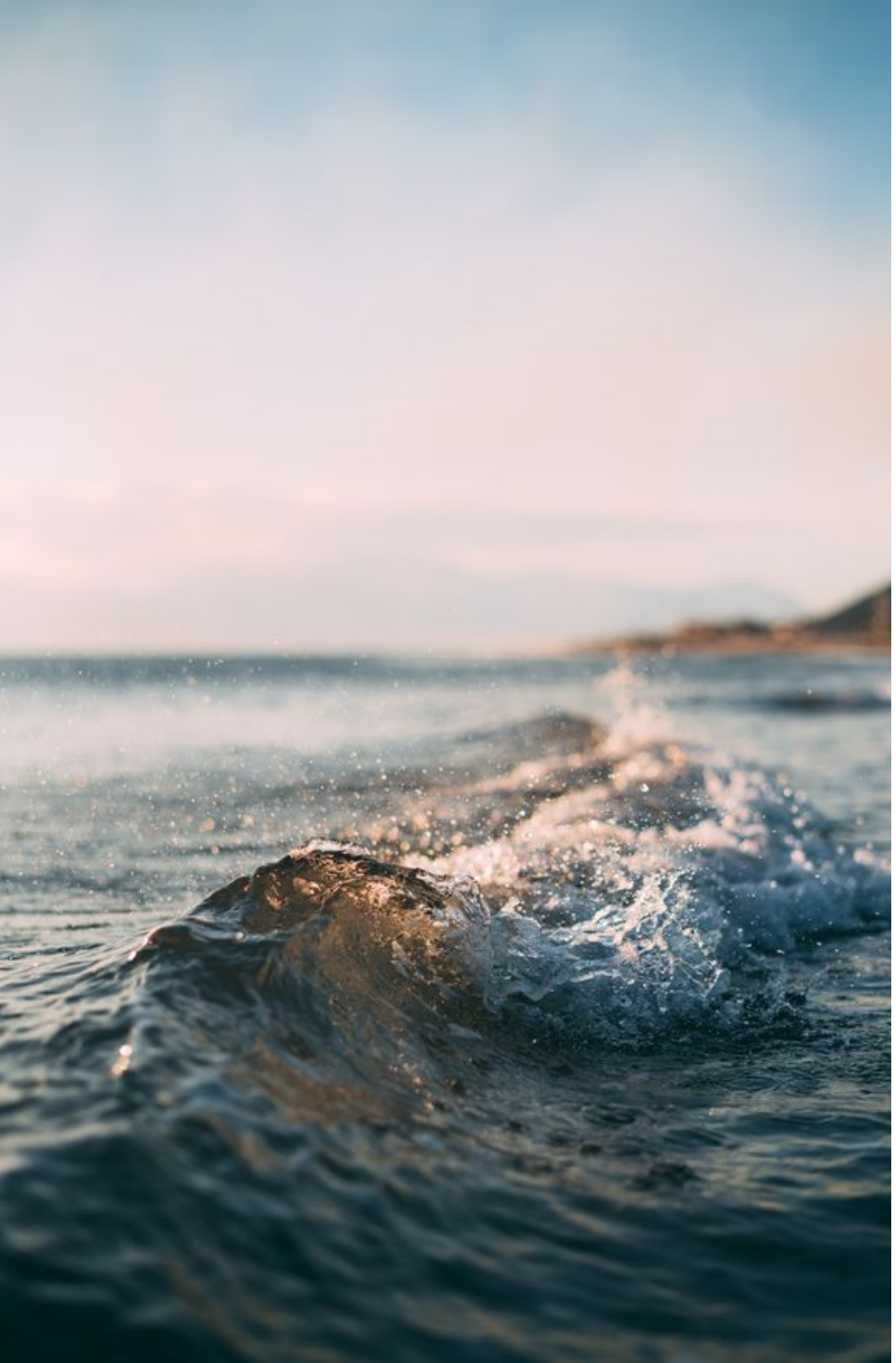




Churn by Tenure

The longer a customer has been with the company (tenure), the less likely the customer is to churn, with significant higher risk for tenures shorter than 18 months





Outline

- Introduction and problem definition
- Exploratory data analysis
- Machine learning
- Conclusions and recommendations

Machine Learning

- **Develop model to predict churn and identify relevant features**
- **Three models tested**

  XGBoost

  Random Forest Classifier

  Logistic Regression

- **All models have similar performance**

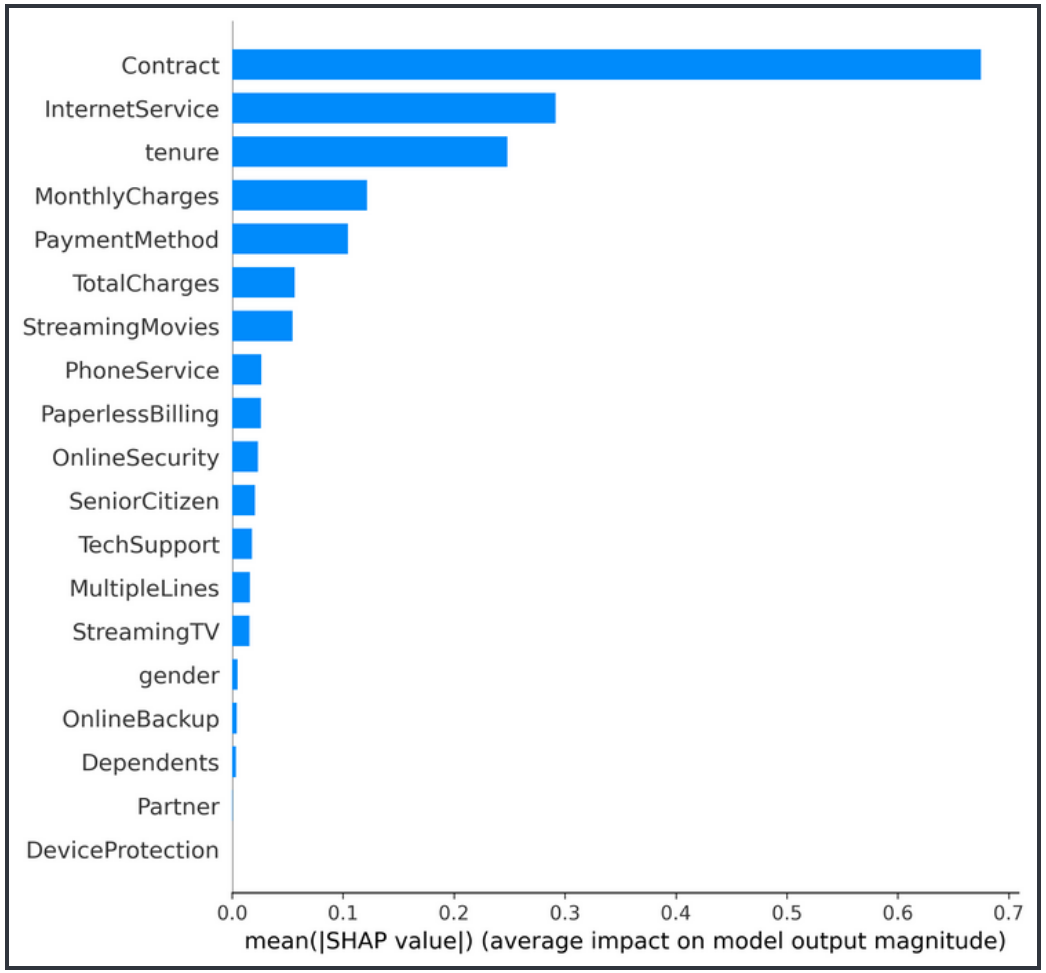
  Correctly predict ~80% customers who churn (recall)

  Incorrect predict churn for ~25% customers who don't churn (false positive)

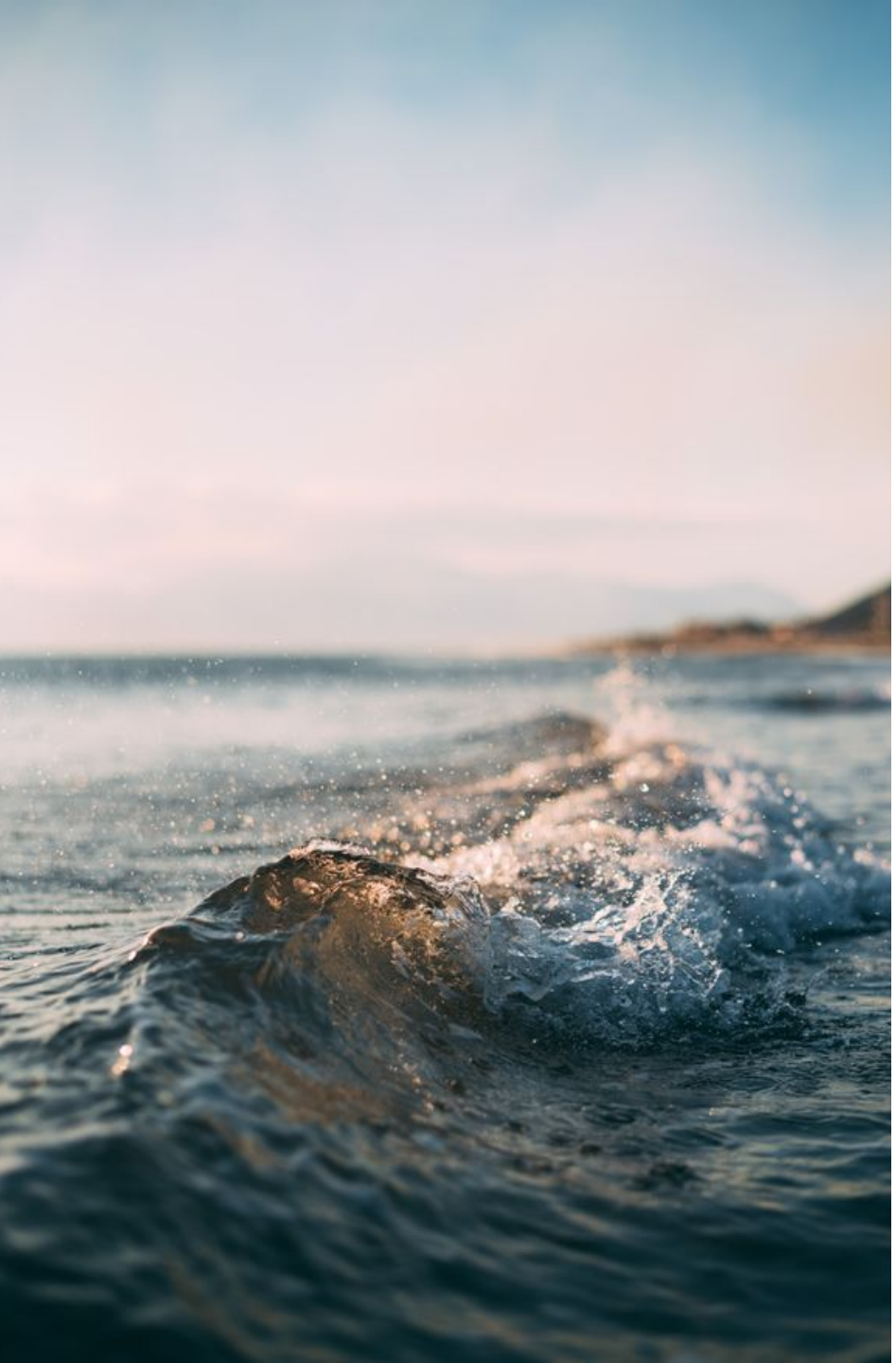
  Overall Precision ~52%

Feature Importance

The model identifies which are the most important features to predict customer churn



- Contract type
- Tenure
- Internet Service
- Monthly and Total Charges
- Payment Method



Outline

- Introduction and problem definition
- Exploratory data analysis
- Machine learning
- Conclusions and recommendations

Conclusions

Based on the results of the machine learning model, the following factors are most likely to influence a customer probability of churning

Contract type

Customers on month to month contracts are much more likely to churn than customers on longer contracts

Internet Service

Customers with Fiber Optic service are much more likely to churn than customers on DSL

Tenure

Recent customers with tenure < 18 months are more likely to churn than older customers

Internet Service and Charges

For Fiber Optic customer the risk of churn increases with a increase in monthly charges, but for DSL the risk is higher for cheaper plans

Payment method

Customer that use electronic checks are much more likely to churn than customers on other payment forms

Monthly Charges

For tenures shorter than 6 months, the risk of churn is independent of monthly charges, but after 6 months the risk is higher for customers with monthly charges higher than $80

Recommendations

Based on the features that influence churn, the following strategies could be used to reduce churn

Encourage long term contracts

Create incentives to move people from Month-to-Month to 1 or 2 year contracts. Incentives could include discounts for the first 6 months of service, bundle products such as modems or other promotions to encourage customers to sign up to longer plans

Encourage direct debit

Set up discounts for direct debit or a small fee for electronic checks to encourage uptake of direct debit. Customers on direct debit are much less inclined to leave than if they pay by electronic check

Improve Fiber Optic offering

As customers on Fiber Optic are much likely to churn than DSL customers, particularly at higher prices, there may be scope to reduce prices to increase competitiveness. This needs to balance the cost of churn vs a reduction in margin for the Fiber Optic plans

Benchmark Fiber Optic offering

A further recommended step is to benchmark the Fiber Optic offering. While these are the most expensive plans and some churn is driven by price, there also maybe other reasons for churn such as reliability of service. An analysis of customers service quality using outage/speed data could uncover technical issues influencing churn

Questions?

Luis Neumann
Data Scientist

luis.neumann@gmail.com

0408 187 365

https://www.linkedin.com/in/lneumannds/