

# Churn reduction

*Akshay Dinkar Patil*

# Contents

## 1 Introduction

- 1.1 Problem Statement .....
- 1.2 Data .....

## 2 Methodology

- 2.1 Pre Processing .....
- 2.1.1 Outlier Analysis .....
- 2.1.2 Feature Selection .....
- 2.2 Modelling .....
- 2.2.1 Model Selection .....
- 2.2.2 Multiple Linear Regression .....
- 2.2.3 Regression Trees .....
- 2.2.4 Classification .....

## 3 Conclusion

- 3.1 Model Evaluation .....
- 3.1.1 Mean Absolute Error (MAE) .....
- 3.1.2 Mean Squared Error (MSE) .....
- 3.2 Model Selection .....

## Appendix A - Extra Figures

## Appendix B - R Code

- Red Wine Multiple Histograms (Fig: 2.1) .....
- Red Wine Histograms with Means (Fig: 2.2) .....
- White Wine Multiple Histograms (Fig: 3.1) .....
- Red Wine Boxplots (Fig: 2.3) .....
- Effect of Outliers (Fig: 2.4) .....
- White Wine Boxplots (Fig: 3.2) .....
- Complete R File .....

## References

# Chapter 1

## Introduction

### 1.1 Problem Statement

---

Churn (loss of customers to competition) is a problem for companies because it is more expensive to acquire a new customer than to keep your existing one from leaving. This problem statement is targeted at enabling churn reduction using analytics concepts. The objective of this case is to predict customer behaviour.

### 1.2 Data

---

Task is to predict the churn score i.e. if the customer has moved (1=yes; 0 = no)  
The Predictors provided are as follows:

- account length
- international plan
- voicemail plan
- number of voicemail messages
- total day minutes used
- day calls made
- total day charge
- total evening minutes
- total evening calls
- total evening charge
- total night minutes
- total night calls
- total night charge
- total international minutes used
- total international calls made
- total international charge
- number of customer service calls made

The Target variable is 'move' if the customer has moved (1=yes; 0 = no)

