Predicting Customer Churn in Telecom   
using Machine Learning Techniques

Request for Proposal

Anish Mahapatra  
MSc in Data Science

Table of Contents

[Abstract 4](#_Toc61449128)

[1. Introduction 5](#_Toc61449129)

[1.1 Multi-touch Attribution 5](#_Toc61449130)

[1.2 Data for Multi-Touch Attribution 5](#_Toc61449131)

[1.3 Agent-Based Modelling 5](#_Toc61449132)

[2. Background and Related Research 5](#_Toc61449133)

[2.1 Background to Multi Touch Attribution 5](#_Toc61449134)

[2.2 Background to Agent-Based Modelling 6](#_Toc61449135)

[2.3 Related Study 6](#_Toc61449136)

[2.4 A Note on Validating Multi Touch Attribution Methods using Agent-Based Modelling 7](#_Toc61449137)

[3. Problem Statement 8](#_Toc61449138)

[4. Research Questions 8](#_Toc61449139)

[5. Aim and Objectives 8](#_Toc61449140)

[6. Significance of the study 9](#_Toc61449141)

[7. Scope of the study 9](#_Toc61449142)

[7.1 Limitations of the Study 9](#_Toc61449143)

[8. Research Methodology 11](#_Toc61449144)

[8.1 Simulated Data Generation 11](#_Toc61449145)

[8.1.1 Overview of Netlogo 11](#_Toc61449146)

[8.1.2 Understanding the basics of Netlogo 12](#_Toc61449147)

[8.1.3 Overview of the proposed design for Data Simulation 13](#_Toc61449148)

[8.2 Data Transformation 14](#_Toc61449149)

[8.3 Implementation of Multi Touch Attribution Techniques 14](#_Toc61449150)

[8.3.1 Heuristic Approaches to Multi Touch Attribution 15](#_Toc61449151)

[8.3.2 Algorithmic Approaches to Multi Touch Attribution 16](#_Toc61449152)

[8.4 Validation of Attribution Techniques 19](#_Toc61449153)

[9. Expected Outcomes 20](#_Toc61449154)

[9.1 Data Generation 20](#_Toc61449155)

[9.2 Data Transformation 20](#_Toc61449156)

[9.3 Implementation of Multi Touch Attribution Techniques 20](#_Toc61449157)

[9.4 Validation of Attribution Techniques 20](#_Toc61449158)

[10. Required Resources 21](#_Toc61449159)

[10.1 Hardware Requirements 21](#_Toc61449160)

[10.2 Software Requirements 21](#_Toc61449161)

[11. Research Plan 22](#_Toc61449162)

[References 23](#_Toc61449163)

# Abstract

With the advent of increasing competition in various segments of the market, it is crucial for companies to be able to retain customers. Customer retention policies can affect the annual turnover drastically depending on the rate of churn. In fact, the cost of customer churn to the Telecom industry is approximated to be around $10 billion per year. Studies show that the cost of customer acquisition is much higher than the cost of customer retention. This leads to the understanding that developing processes and efficient policies to reduce customer churn can reduce spend on customer relations.

Datasets for customer churn are quite large where many features can be found. Not all the features are significant for churn prediction. Hence, feature engineering requires not only computation, but a great amount of time as well. Through this we will find and the features that that will be significant for churn prediction. The aim is to be able to predict the churn accurately and show the performance difference of various algorithms.

# 1. Aim and Objectives

### 1.1 Multi-touch Attribution

### 1.2 Data for Multi-Touch Attribution

### 1.3 Agent-Based Modelling

#### 1.3.1 Using Agent-Based Modelling technique for Data Generation

# 2. Introduction

# 3. Background and Related Research

### 2.1 Background to Multi Touch Attribution

#### 2.1.1 Customer Journey

#### 2.1.2 Marketing Funnel

### 2.2 Background to Agent-Based Modelling

### 2.3 Related Study

#### 2.3.1 Data-driven multi-touch attribution models

#### 2.3.2 Multichannel Marketing Attribution Using Markov Chains

#### 2.3.3 Interpretable Deep Learning Model for Online Multi-touch Attribution

### 2.4 A Note on Validating Multi Touch Attribution Methods using Agent-Based Modelling

# 3. Problem Statement

# 4. Research Questions

# 5. Aim and Objectives

# 6. Significance of the study

# 7. Scope of the study

## 7.1 Limitations of the Study

# 8. Research Methodology

## 8.1 Simulated Data Generation

### 8.1.1 Overview of Netlogo

### 8.1.2 Understanding the basics of Netlogo

#### 8.1.2.1 Agent

#### 8.1.2.2 Command

#### 8.1.2.3 Reporter

### 8.1.3 Overview of the proposed design for Data Simulation

## 8.2 Data Transformation

## 8.3 Implementation of Multi Touch Attribution Techniques

### 8.3.1 Heuristic Approaches to Multi Touch Attribution

#### 8.3.1.1 Linear Attribution

#### 8.3.1.2 Time-decay Attribution

#### 8.3.1.3 Position based/U-Shaped Attribution

#### 8.3.1.4 W-Shaped Attribution

### 8.3.2 Algorithmic Approaches to Multi Touch Attribution

#### 8.3.2.1 Markov Model (Graph Theory Based Approach)

##### 8.3.2.1.1 Order of Markov Chains

##### 8.3.2.1.2 Removal Effect

#### 8.3.2.2 Hidden Markov Model (Graph Theory Based Approach)

#### 8.3.2.3 AIC and BIC (Supervised Approach)

#### 8.3.2.4 Bagged Logistic Regression (Supervised Approach)

#### 8.3.2.5 Shapley Value (Cooperative Game Theory Based Approach)

#### 8.3.2.5 Deep Neural Net with Attention (Deep Learning Approach)

## 8.4 Validation of Attribution Techniques

# 9. Expected Outcomes

## 9.1 Data Generation

## 9.2 Data Transformation

## 9.3 Implementation of Multi Touch Attribution Techniques

## 9.4 Validation of Attribution Techniques

# 10. Required Resources

## 10.1 Hardware Requirements

Based on the defined scope of the proposed thesis, the following are the required resources:

**NOTE:** Please ensure you have Administrator access in the machine (Windows/Ubuntu/macOS)

The minimum hardware requirements for this project are:

**RAM:** Minimum 4 GB (8 GB recommended for optimum performance)

**Disk space:** Minimum of 4GB free space needs to be allocated (Depends on the Simulation Size)

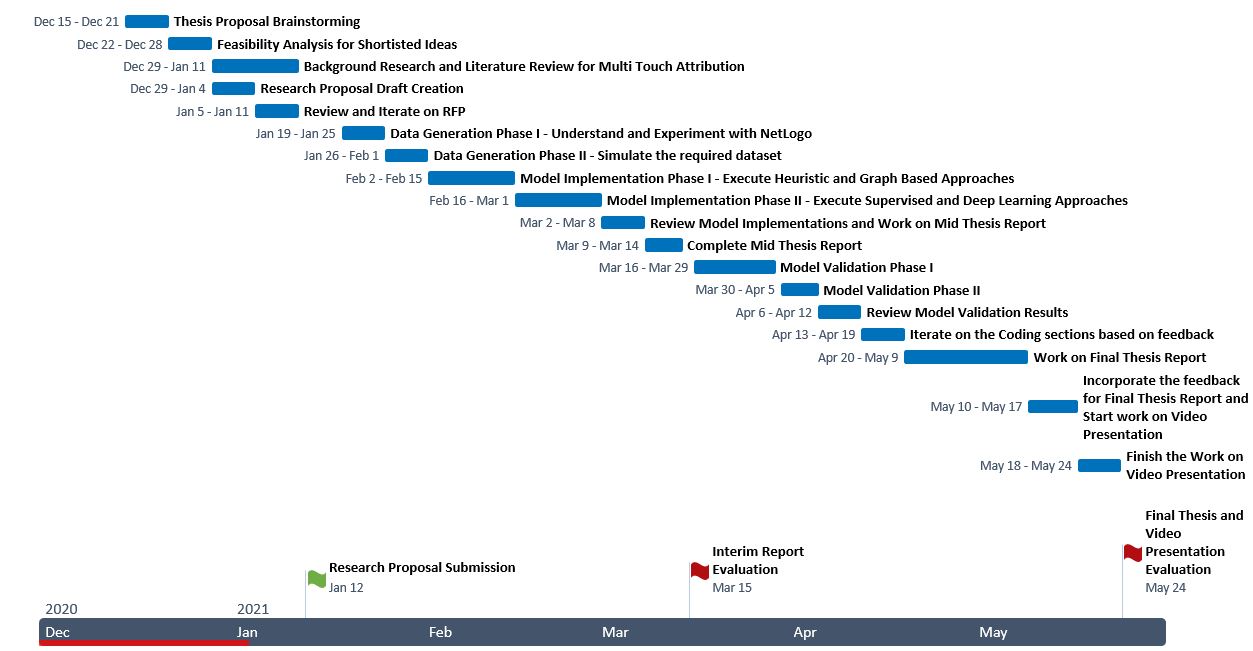
## 10.2 Software Requirements

|  |  |
| --- | --- |
| Software | Minimum Version |
| R | >= 3.6 |
| Python | >= 3.5 |
| NetLogo | >= 6.1.1 |
| R Studio | >= 1.2.0 |
| Java SDK | >= 8 |
| Excel | >= 2007 |

# 11. Research Plan

The following GANTT chart proposes the timeline for the research and implementation of the project.

Based on the complexity of the different phases, the timelines are subject to minor adjustments. Regardless, the candidate shall pledge to stick to the timeline as closely as possible.



# References