

# **Telco Feedback Management System using Sentiment Analysis**

Suhail Jamaldeen

149212 K

MSc in IT

University of Moratuwa

March 2018

## **Supervised by**

Mr. Saminda Premaratne

Senior Lecturer

University of Moratuwa

## **Declaration**

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Suhail Jamaldeen

Date:

Supervised by

Mr. Saminda Premaratne

Date:

## **Dedication**

This thesis is dedicated my father

Late Mr. HL. Jamaldeen  
(Senior Superintendent of Police)

who sacrificed his life to our mother nation Sri Lanka

## **Acknowledgements**

Foremost, I would like to express my sincere gratitude to my supervisor Mr. Saminda Premarathna, for the continuous support of my MSc study and research, for his patience, motivation, enthusiasm, and immense knowledge. I have been amazingly fortunate to have a supervisor who gave me the freedom to explore on my own and at the same time the guidance to recover when my steps faltered.

Besides my supervisor, I would also like to thank my family specially my mother, my wife and daughter for the support they provided me through my entire life and in particular.

I must also acknowledge my friends Chinthuja Varathalingam, Saranya Varathalingam and Arunnirthana Arulanantham for their support throughout my thesis submission.

## Abstract

Social Media such as Facebook, Twitter, blogs, etc. and social interactions are currently growing in an exploding speed. This leads Social Media in becoming a large repository of valuable opinions of a large scale of different people on numerous products and services. Twitter and Facebook plays a major contribution in social interactions. As such, people share their opinion on Telecommunication service providers using social media.

Usually any telecommunication provider in the world provide more than one services, such as Home Broadband, Mobile Devices, Prepaid, Postpaid, Mobile Broadband and Televisions.

When people share their opinions and feedbacks on the services which given by any telecommunication providers through social media. Telecommunication providers face a huge challenge in

- Categorizing the opinions as positive feedback, negative feedback, questions and statements
- Identifying for what services, the opinion is belong to
- Assigning the categorized opinion to a support engineer depending on the services which are provided by the telecommunication provider
- Replying to opinions, following them and managing them

When it's come to categorizing the opinion, we need a best approach to analyze the text and the data analysis approach should be able to detect only explicitly expressed opinions.

And then when it's come to assigning categorized opinion to a support engineer; it cannot be done from a social media itself. None of the social media allows us to use their platform to manage the individual's opinion.

Same as assigning; replying and managing individual's opinion cannot be done using any of the social media.

This research study will propose a best approach to analyze the text, which are mentioned a telecommunication provider and this study will provide a mechanism or a software solution to manage the opinion raised by the subscriber of a telecommunication.

In this thesis project, we focus on using Twitter, one of the most popular social media.

# Table of Contents

Declaration.....	i
Dedication.....	ii
Acknowledgements.....	iii
Abstract.....	iv
Table of Contents.....	vi
Chapter 1.....	1
Introduction.....	1
1.1 Prolegomena.....	1
1.2 Background and Motivation.....	1
1.3 Problem Statement.....	2
1.4 Hypothesis.....	3
1.5 Objectives.....	3
1.5 Proposed solution.....	4
1.6 Scope of the thesis project.....	4
1.6 Structure of the Thesis.....	5
Chapter 2.....	7
Literature Review and Background Study.....	7
2.1 Introduction.....	7
2.2 Sentiment and Sentiment Analysis.....	7
2.2.1 Sentiment.....	7
2.2.2 Sentiment Analysis.....	7
2.2.3 Early Study on Sentiment Analysis.....	8
2.2.4 Main Approaches to Sentiment Analysis.....	12
2.2.5 Sentiment Classification Techniques/Algorithms.....	13
2.2.6 Evaluation of Sentiment Classification.....	16
2.3 Machine Learning.....	16
2.3.1 Machine Learning Offerings.....	19
2.4 Future challenges of Sentiment Analysis.....	21
2.5 Twitter.....	22
2.5.1 Twitter API and Libraries.....	22
2.5.2 Authentication and Authorization on Twitter API.....	23
2.6 Related Works.....	24

2.6.1 Amolik et al., (2016) .....	24
2.6.2 Luo et al., (2013) .....	24
2.6.3 Pak & Paroubek (2010) .....	24
2.6.4 Davidov et al.,(2010) .....	25
2.6.5 Agarwal et al., (2011) .....	25
2.6.6 Parikh and Movassate (2009) .....	26
2.6.7 Kolchyna et al., (2015) .....	26
2.6.8 Overall Related Work Summery .....	26
2.7 Existing Sentiment Analysis Architecture for Twitter feed .....	27
2.7.1 Preprocessing .....	27
2.7.2 Feature Extraction.....	28
2.7.3 Training .....	28
2.7.4 Classification .....	28
2.8 Sentiment Analysis Approaches for Twitter Data .....	28
2.9 Problem Definition.....	30
2.10 Summery.....	30
Chapter 3 .....	31
Propose Approach and Technologies Adoption .....	31
3.1 Introduction .....	31
3.2 Selection of Twitter API .....	31
3.3 Authentication and Authorization on Twitter API .....	31
3.4 Data Storage Mechanism.....	32
3.5 Overall Sentiment Analysis Techniques.....	32
3.6 Selection of Sentiment Analysis Approach .....	33
3.7 Selection of Machine Learning Types .....	33
3.8 Selection of Machine Learning Offering .....	34
3.9 Deployment of Machine Learning Solutions.....	35
3. 10 Selection of Software Architecture.....	35
3.10.1 Presentation Layer .....	36
3.10.2 Business Logic Layer.....	36
3.10.3 Data Access Layer .....	37
3.11 Hosting and Cloud Offering.....	37
3.11 Software Used.....	37
Chapter 4.....	39



Design and Implementation .....	39
4.1 Introduction .....	39
4.2 Research Methodology .....	39
4.3 Design Assumptions .....	40
4.4 System Architecture .....	40
4.5 Design Components .....	41
4.5.3 Feedback Management Module .....	44
Chapter 5 .....	46
Evaluation .....	46
5.1 Evaluation .....	46
Chapter 6 .....	48
Discussion .....	48
6.1 Challenges Faced During the Sentiment Analysis .....	48
6.2 Future Work .....	49
Appendixes .....	53
1.0 Acronyms and their meaning used in preprocessing .....	53
2.0 Stopwords .....	54
3.0 Cycle of an Issue .....	55

# List of Figures

Figure 1 - Document Level Sentiment Analysis using Supervised Learning.....	9
Figure 2 - Artificial Intelligence vs Machine Learning vs Deep Learning.....	17
Figure 3 - Machine Learning Process .....	17
Figure 4 - Supervised Learning and Unsupervised Learning .....	18
Figure 5 - Sentiment Classification Based On Emoticons .....	25
Figure 6 - Lexicon-Based Model.....	29
Figure 7 - Overall Sentiment Analysis Techniques.....	32
Figure 8 - Compare Performance of Two Algorithms in Machine Learning Studio .....	34
Figure 9 - Use of Azure Machine Learning Studio for Machine Learning .....	35
Figure 10 - Three Layer Architecture .....	36
Figure 11 - Software Architecture.....	41
Figure 12 - Tweet crawling module .....	42
Figure 13 - Preprocessing Module .....	42
Figure 14 - Sentiment Analysis Module .....	43
Figure 15 - Support Vector Method.....	44
Figure 16 - Feedback Management Module.....	45

## **List of Table**

Table 1 - Confusion Matrix.....	16
Table 2 - Overall Related Work Summery.....	26
Table 3 - Evaluation Results for Sentiment Analysis.....	46
Table 4 - Evaluation Results for Service Category.....	46