

**Reliability of Mobile Application Analysis Based on
The User Reviews of Health and Fitness Category
Which are Available from Google Play Store**

Mr. Donbosco Prashanth
169326G



Dissertation submitted to the Faculty of Information Technology, University of Moratuwa,
Sri Lanka for the partial fulfillment of the requirements of Degree of Master of Science in
Information Technology

March 2019

Declaration

We declare that this thesis is our 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 Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Donbosco Prashanth

Signature of Student

Date:

Supervised by

Mr. Saminda Premaratne

Senior Lecturer

Faculty of Information Technology

University of Moratuwa

Signature of Supervisor

Date:

Acknowledgments

Completion of a project dissertation has a challenging task which requires an intense effort and dedication which must be endured for a considerable period of time, continually. The assistance, guidance, and familiarization of the author for unfamiliar concepts and knowledge by external but relevant parties would be imperative in addition to the aforementioned qualities have been imperative for successful completion of such a high-level assignment.

First and foremost, I would like to extend the sincere thanks and appreciation to my supervisor Mr. S. C. Premaratne, Department of Information Technology, Faculty of Information Technology, University of Moratuwa. His guidance and expertise had helped for the successful completion of this important assignment which has been the final frontier to conquer in the MSc Program. He has been the driving force for this author to select this topic and through the final project.

I wish to express my heartfelt gratitude to all academic staff of the Faculty of Information and technology, in particular for the opportunity given to be a student of such a valuable course of MSc program and for providing assistance in numerous ways by obliging many requests of the author for this successful completion.

The associates are remarkable in assisting me in my examinations in numerous ways. I would like to extend the heartiest gratitude for the great contribution and moral support from my sister and siblings and would like to mention the great bunch of batch mates who had been a wonderful motivational factor that had made the learning process of this MSc program highly fruitful, entertaining and knowledgeable.

Abstract

The current world everything embedded in the mobile phone and it has many uses in the field of life. Google apps store is a very important role to playing the mobile application development. This research proposed the reliable Health and fitness mobile application based on the Apps reviews which are given by end-user.

An end-user of mobile application review has to be analyzed in various methods because all the users do not have the same knowledge and same requirements but their review or feedback should depend on their knowledge, requirement and user experience.

In this research mainly we are targeting non-technical attributes which are gathered from Google play store to analyze the reliability of health and fitness mobile application based on user review, using the Natural Language Process techniques and classification models. In the current world has live technology improvements and day to day different need of user requirements so this will reflect the mobile application world trend, and it will make live changes in google play app store also.

Based on user review we can to find out proper path to analysis the google play store data to get an effective result in this research because that apps store has more complicated data in a different perspective of Health and Fitness area.

Research has tended to focus for Google play apps store customer has easy to find out the efficient application of Health and Fitness and comparing other mobile apps which are placed in the same apps store. Accordingly to that concept of our research final data will be produced in a structured database. In future idea from this updated database link with the mobile application to upload the google play store for the user access and user can easily predict the which application is efficient of their appropriate requirement.

Table of Content

DECLARATION.....	I
ACKNOWLEDGMENTS	II
ABSTRACT	III
TABLE OF CONTENT.....	IV
ABBREVIATIONS	VI
LIST OF FIGURE	VII
LIST OF TABLES	VIII
CHAPTER 1.....	1
Introduction	1
1.1 Prolegomena	1
1.2 Problem Statement.....	2
1.3 Natural Language Processing Approach.....	2
1.4 Goal of Research.....	3
1.5 Aim and Specific Objectives	3
1.6 Overview of the report.....	3
1.7 Summary	3
CHAPTER 2.....	4
Literature Review.....	4
2.1 Introduction.....	4
2.1 Summary	7
CHAPTER 3.....	8
Technology Adapted in Mobile Application Analysis.....	8
3.1 Introduction.....	8
3.2 Method for extract the features from user reviews text.....	8
3.3 Method of identifying sentiment polarity classification of user reviews.....	8
3.4 Method for identifying topics in a user review.....	9
3.5 Summary	10
CHAPTER 4.....	11
A Novel Approach for Mobile Application Analysis.....	11
4.1 Introduction.....	11
4.2 Proposed model.....	11
4.3 System overview	12
4.4 Data Collection from the Google play store.	12
4.5 Preprocessing data	13
4.6 Identify the suitable methods for NLP.....	13
4.7 Discovering Knowledge.....	13
4.8 Summary	14

CHAPTER 5.....	15
Analysis and Design of the Proposed Solution.....	15
5.1 Introduction.....	15
5.2 System design	15
5.2.1 Data collection module	15
5.2.2 NLP model for sentiment analysis.....	15
5.2.3 NLP model for topic analysis TFIDE and LSA.....	17
5.2.3.1 TFIDE	17
5.2.3.2 Latent Semantic Analysis	18
5.2.4 Data testing and visualizing.....	19
5.3 Summary.....	19
CHAPTER 6.....	20
Implementation of the Solution.....	20
6.1 Introduction.....	20
6.2 Implementation of data scraping tool	20
6.3 Implementation of the NLP model for sentiment analysis	22
6.4 Implementation of NLP model for topic analysis.....	24
6.5 Implementation of Mobile Application	25
6.6 Summary.....	25
CHAPTER 7.....	26
Evaluation	26
7.1 Introduction.....	26
7.2 Data collection module	26
7.3 NLP model for sentiment analysis.....	27
7.4 NLP model for topic analysis	31
7.5 Data visualizing and analyzing.....	33
CHAPTER 8.....	36
Discussion	36
8.1 Introduction.....	36
8.2 Limitations	36
8.3 Further Developments.....	36
8.4 Summary	36
REFERENCES:	37
APPENDIX A	39
APPENDIX B	40
APPENDIX C	41
APPENDIX D.....	42
APPENDIX E	43

ABBREVIATIONS

NLP	Natural Language Processing
WEKA	Waikato Environment for Knowledge Analysis
GUI	Graphical User Interface
TFIDE	Term Frequency Inverse Document Frequency
LSA	Latent Semantic Analysis
NLTK	Natural Language Tool Kit
API	Application Programming Interface

List of Figure

Figure 3.1 Sentence analysis	8
Figure 3.2 Topic model	9
Figure 4.1 Proposed model	11
Figure 4.2 Overview of the Google play store API structure	12
Figure 4.3 Clustering method of keyword analyzer	13
Figure 5.1 Solution TFIDE Model	18
Figure 5.2 Solution LSA Model	18
Figure 6.1 User review scarping tool	20
Figure 6.2 Google play Apps store structure change	21
Figure 6.3 Data pre-processing	21
Figure 6.4 Vader polarity value score for sentiment analysis	22
Figure 6.5 Tokenizing code for stop word and unwonted symbols	22
Figure 6.6 Polarity analysis score	23
Figure 6.7 Successful execution of sentiment analysis	23
Figure 6.8 Implementation of topic analysis module	24
Figure 6.9 Weka data visualization	25
Figure 7.1 Extracted user reviews of top 35 application from the google play	26
Figure 7.2 Sentiment analysis polarity score of user reviews	27
Figure 7.3 Output of accuracy using Varder sentiment analysis	29
Figure 7.4 Output of accuracy using traditional textblob sentiment analysis	30
Figure 7.5 Comparison of output traditional method and latest vadermethod	30
Figure 7.6 Output result of topic analysis	31
Figure 7.7 Percentage of total user reviews based on topics	32
Figure 7.8 Visually represent the summary of sentiment analysis	33
Figure 7.9 No of user reviews are contain no of positive & negative user	34
Figure 7.10 Mobile Application Interface	35

List of Tables

	Page
Table 2.1 The Summary of literature survey	7
Table 5.1 The Major python libraries used for solution	17
Table 7.1 Final summary of Sentiment Analysis from 35 Applications	28
Table 7.2 Topics and related word	31