

# **Sentiment User Review Analyzer for Android Applications**

Name of the Student : T I M P SILVA

Registration No : 169334E

Name of the Supervisor : Mr.SamindaPremaratne

M.Sc. in Information Technology  
Department of Information Technology  
University of Moratuwa

## DECLARATION

I do declare that this thesis is my own work and has not been submitted in any form for another diploma or degree at any university or other institution of tertiary education. This thesis was prepared for degree of Master of Science/ Information Technology under the guidance and supervision of Mr. Saminda Premaratne. All information taken from the published or unpublished work of others has been acknowledged in the text and a list of references is shown.

Name of Student

T I M P Silva

Signature of Student

.....

Date:

Name of Supervisor

Mr. S C Premaratne

Signature of Supervisor

.....

Date:

## ACKNOWLEDGMENTS

I would like to thank all who helped me to success this project. I would particularly like to single out my supervisor at the University of Moratuwa, Mr. SamindaPremaratne, Senior Lecture, I want to thank you for your excellent cooperation and for all of the opportunities I was given to conduct my project and further my thesis. When I was in trouble with the subject matters and personal matters, your guidance and encouragement make me to successfully complete this project.

I would also like to thank all the other lectures of the University of Moratuwa who have shared your knowledge throughout the whole period of study at university. And also giving my grateful thank to all my batch mates of 10<sup>th</sup> batch of MSc in IT. You encourage me every time when I was needed.

My colleague at the office, you gave me your support to evaluate the system taking your valuable time. Thank you.

## ABSTRACT

Usage of mobile phones rapidly increased for the last decade globally. Introducing smart phones has led to a renaissance in mobile application development. Not only for make calls, send SMS phone to become smart so that it can be used for email access, web browsing, e-commerce, bill payments and entertainments etc. Because of this popularity application markets became massive. For the same purpose there are many applications deployed by different developers. Some apps can be considered as good, and some are not working as the user expect. And also there are many harmful applications which are stealing sensitive user information.

As Android applications have been used by many people, lots of issues have been already identified and many type of research are done on this topic. Even though there are many android applications available in the market, it is difficult to select the most suitable one before use. To get an idea about the application which is going to be used, user needs to read many reviews and ratings. This manual process is time consuming and not efficient.

In this project main aim is to develop a User Review analyzer which is gives a good analytical picture of the application before install.

## Table of Contents

DECLARATION .....	i
ACKNOWLEDGMENTS .....	ii
ABSTRACT.....	iii
Table of Contents.....	iv
Table of Figures .....	vi
List of Tables .....	vii
CHAPTER 01 .....	1
1.1. Introduction.....	1
1.2. Background and Motivation .....	2
1.3. Statement of Research Problem.....	2
1.4. Aim and Objectives.....	2
1.4.1. Main objectives of the project.....	3
1.5. The scope of the Project.....	3
1.6. Overview of the Report.....	3
CHAPTER 02 .....	4
Review of Literature .....	4
2.1. Introduction.....	4
2.2. Android OS and Android Apps.....	4
2.3. Malware Detection.....	5
2.3.1. Static Analysis.....	5
2.3.2. Dynamic Analysis .....	7
2.3.3. Hybrid Analysis.....	8
2.4. User Reviews/Feedback Analysis .....	8
2.5. Current State .....	9
Chapter 3.....	10
Technologies used for developing the solution.....	10
3.1. Introduction.....	10
3.2. Back End/Server Side .....	10
3.3. Front End .....	10
3.4. Languages used.....	10
3.4.1. C# Language .....	10
3.4.2. Python .....	11
3.4.3. PHP Language: .....	11

3.4.4.	Java Programming Language .....	11
3.5.	Database.....	11
3.5.1.	MySQL:.....	11
3.6.	Tools and platforms used .....	12
3.6.1.	NLTK .....	12
3.6.2.	JSON .....	12
3.6.3.	Apache: .....	12
3.6.4.	Php Myadmin:.....	12
3.6.5.	Android Studio .....	12
3.7.	Summary .....	12
CHAPTER 04 .....		14
Approach.....		14
4.1.	Introduction.....	14
4.2.	Data collection .....	14
4.3.	Data Analysing.....	14
4.3.1	Data pre-processing .....	14
4.3.2.	Feature Extraction.....	15
4.3.3.	Classification .....	15
4.4.	Visualization .....	15
4.5.	Process .....	16
4.6.	Summary .....	16
CHAPTER 05 .....		17
Analyze and Design .....		17
5.1.	Introduction.....	17
5.2.	User Review extraction.....	18
5.3.	Sentiment Analysis .....	19
5.4.	Mobile Application .....	20
5.5.	Summary .....	21
CHAPTER 06 .....		22
Implementation .....		22
6.1.	Introduction.....	22
6.2.	Extracting user Reviews from Google Play Store.....	22
6.3.	Analysing the Reviews .....	24
6.3.1.	Stop word removal.....	25

6.3.2.	Stemming .....	25
6.3.3.	Tokenization.....	26
6.3.4.	Feature extraction.....	26
6.3.5.	Sentiment Classification and calculation the confident .....	26
6.3.6.	Calculating most discussed features.....	27
6.4.	Web Service.....	29
6.5.	Mobile Application.....	29
6.6.	Summary .....	32
CHAPTER 07 .....		33
Evaluation.....		33
7.1.	Introduction.....	33
7.2.	Evaluation of User Review Extractor .....	33
7.3.	Evaluation of Sentiment Analysis Module .....	34
7.4.	Evaluation of the Final Output.....	38
7.5.	Evaluation of Data Visualization Module.....	39
7.6.	Summary .....	39
CHAPTER 08 .....		40
Discussion & Future Works.....		40
8.1.	Introduction.....	40
8.2.	Conclusion .....	40
8.3.	Future Works .....	40
References.....		42
Appendix A.....		44
A.1.	Image of the User Review Extractor.....	44
A.2.	Image of Extracted data .....	45
A.3.	First screen of mobile application.....	46
A.4.	Result screen of mobile application.....	47

## Table of Figures

Figure 1: The Android component stack.....	4
Figure 2: Hybrid feature extraction.....	8
Figure 3: Sentiment Analysis Process.....	15
Figure 4: Top Level Design of the System .....	17

Figure 5: User Review Extraction Flow Chart.....	18
Figure 6: Sentiment Analysis Flow Chart.....	19
Figure 7: Mobile Application Flow Chart.....	20
Figure 8: Sample URL List.....	22
Figure 9: Code segment of expanding web page .....	23
Figure 10: Codes for Read Data from HTML.....	24
Figure 11: Information taken from CSV.....	25
Figure 12: Stop Words.....	25
Figure 13: Result of Tokenization .....	26
Figure 14: Feature Extraction .....	26
Figure 15: Codes for Classification .....	27
Figure 16: Result of Analysis .....	27
Figure 17: Codes for Classify to Selected Features .....	28
Figure 18: Loading App Names.....	30
Figure 19: Codes for Graphing .....	31
Figure 20: Codes for Bar Chart.....	31

## List of Tables

Table 1 Sentiment result of 25 games .....	34
Table 2 Feature-wise positive negative ratio.....	35
Table 3 Sample of result.....	36