

DECLARATION

-

This is to certify that Report titled “**SENTIMENT ANALYSIS**”, is submitted by us in partial fulfillment of the requirement for the award of degree B.Tech. in Computer Science & Engineering to BPIT, GGSIP University, Dwarka, Delhi. It comprises of our original work. The due acknowledgement has been made in the report for using others work.

Date: 7 Novemeber,2017

Priya Bhasin, 30920802714

Edwin John, 03620802714

Radhika, 09620802714

Certificate By Supervisor

This is to certify that Report titled “**SENTIMENT ANALYSIS**” is submitted by **Priya Bhasin(30920802714),Edwin John(03620802714), Radhika(09620802714)**in partial fulfillment of the requirement for the award of degree B.Tech in Computer Science & Engineering to BPIT, GGSIP University, Dwarka, Delhi. It is a record of the candidates own work carried out by them under my supervision. The matter embodied in this Report is original and has not been submitted for the award of any other degree.

Date: 7 Novemeber,2017

(Signature)

Supervisor

Certificate By HOD

This is to certify that Report titled“**SENTIMENT ANALYSIS**” is submitted by **Priya Bhasin(030920802714)**, **Radhika(09620802714)**, **Edwin John(03620802714)** under the guidance of **Dr. Deepali Virmani and Ms.Nikita** in partial fulfillment of the requirement for the award of degree B.Tech in Computer Science & Engineering to BPIT, GGSIP University, Dwarka, Delhi. The matter embodied in this Report is original and has been dully approved for the submission.

(signature)

Date: 7 November,2017

Dr. Deepali Virmani

ACKNOWLEDGEMENT

We would like to take this opportunity to express our gratitude towards our project guides **Dr. Deepali Virmani and Ms.Nikita** for thier valuable support and contribution towards this project. We also thankful to them for providing the direction essential for completing the project.

Expressing my sincere and heartfelt gratitude to **Priya Bhasin, Radhika, Edwin John**(Team Memebers)for their valuable support and encouragement, this helped us to successfully complete the project.

I would also like to thank all the other people who were directly or indirectly involved in the success of the project.

Finally I would like to thank my family for their love and continuous encouragement during the entire duration of the project.

Radhika, 09620802714

Priya Bhasin, 30920802714

Edwin John, 03620802714

ABSTRACT

The focuses on research work related to **Sentimental Analysis**. The said end-product is in form of an efficient algorithm with enhanced features and compare results with the pre existing one. Strict schedules are followed to ensure timely delivery.. The primary tools and technologies that are being used are Python, SQLite.

Python is a widely used high-level programming language for general-purpose programming, created by Guido van Rossum and first released in 1991. An interpreted language, Python has a design philosophy that emphasizes code readability and a syntax that allows programmers to express concepts in fewer lines of code than might be used in languages such as C++ or Java. The language provides constructs intended to enable writing clear programs on both a small and large scale.

Python features a dynamic type system and automatic memory management and supports multiple programming paradigms, including object-oriented, imperative, functional programming, and procedural styles. It has a large and comprehensive standard library.

SQLite is a relational database management system contained in a C programming library. In contrast to many other database management systems, SQLite is not a client–server database engine. Rather, it is embedded into the end program.

SQLite is ACID-compliant and implements most of the SQL standard, using a dynamically and weakly typed SQL syntax that does not guarantee the domain integrity.

SQLite is a popular choice as embedded database software for local/client storage in application software such as web browsers. It is arguably the most widely deployed database engine, as it is used today by several widespread browsers, operating systems, and embedded systems (such as mobile phones), among others. SQLite has bindings to many programming languages

Table of Contents

DECLARATION	I
CERTIFICATE BY SUPERVISOR	II
CERTIFICATE BY HOD.....	III
ACKNOWLEDGEMENT	IV
ABSTRACT.....	V
1 INTRODUCTION.....	1
1.1 SENTIMENT	1
1.1.1 TYPES.....	1
1.1.1.1 SUBJECTIVITY/OBJECTIVITYIDENTIFICATION.....	1
1.1.1.2 FEATURES/ASPECT BASED.....	2
1.1.2 APPROCHES AND FEATURES.....	2
1.1.3 EVALUATION.....	3
2 RELATED WORK	4
2.1 PROBLEM STATEMENT	4
2.2 PHASES.....	4
2.2.1 PHASE 1	4
2.2.2 PHASE 2.....	5
2.3 APPROACH USED.....	5
2.3.1 TWO MAIN PARTS OF LEXICON APPROACH	6
2.3.1.1 LEXICON.....	6
2.3.1.1.1 SIZE AND ORGANISATION.....	6
2.3.1.2 DATA DICTIONARY.....	6
2.4 NEED OF THE PROJECT	7
2.4.1 TO ADJUST MARKETING STRATEGIES.....	7
2.4.3 DEVELOP PRODUCT QUALITY.....	7
2.4.4 IMPROVE CUSTOMER.....	8
2.4.5 DETERMINE THE EMOTIIONAL TONE.....	8
3 SYSTEM ANALYSIS AND DESIGN	9
3.1 SYSTEM REQUIREMENT SPECIFICATION.....	9
3.1.1 SOFTWARE REQUIREMENT SPECIFICATION.....	9
3.1.2 HARDWARE REQUIREMENT SPECIFICATION.....	9
3.2 E-R DIAGRAM.....	10
3.3 0-LEVEL DFD.....	11
3.4 1-LEVEL DFD(PHASE 1)	12

3.5 1- LEVEL DFD (PHASE 2).....	13
4)_PROPOSED WORK	14
4.1 PROPOSED ALGORITHM.....	14
4.2 PSUEDO CODE (PHASE 1).....	14
4.3 PSUEDO CODE (PHASE 2).....	16
5 IMPLEMENTATION/RESULT	20
5.1 PHASE1.....	20
5.1.1 CODE.....	20
5.1.2 OUTPUT.....	26
5.1.2.1 DATA DICTIONARY.....	26
5.1.2.2 MAIN CODE.....	27
5.1.2.3 OUTPUTS.....	29
5.1.3 DATA SET.....	30
5.1.4 GRAPH ANALYSIS.....	31
5.1.5 CONCLUSION.....	32
5.2 PHASE2.....	33
5.2.1 CODE.....	33
5.2.2 OUTPUT.....	47
5.2.2.1 WORD DATABASE.....	47
5.2.2.2 REVIEW LINKS.....	48
5.2.2.4 UPDATING DICTIONARY.....	50
5.2.2.5 TWITTER ANALYSIS.....	52
5.2.3 GRAPGH ANALYSIS.....	53
5.2.4 CONCLUSION.....	54
5.3 COMBINE PHASE1 AND PHASE 2.....	55
5.3.1 DATA SET	55
5.3.2 PHASE1 RESULT	56
5.3.3 PHASE2 RESULT	57
6 CONCLUSION	59
6.1 CONCLUSION.....	59
6.2 DRAWBACKS ENCOUNTERED.....	59
7) FUTURE WORK.....	60
REFERENCES.....	61

LIST OF FIGURES

1. Figure 1: ER Diagram.....	10
2. Figure 2: Level-0 DFD.....	11
3. Figure 3: Phase I, Level-1 DFD.....	12
4. Figure 4: Phase II, Level-1 DFD.....	14
5. Figure 5: Data Dictionary.....	26
6. Figure 6: Main Code.....	27
7. Figure 7: Main Code.....	28
8. Figure 8: Output.....	29
9. Figure 9: Dataset.....	30
10. Figure 10:Phase I, Graph Analysis.....	31
11. Figure 11: Review Links.....	48
12. Figure 12: Review Rating.....	49
13. Figure 13: Dictionary Updating.....	50
14. Figure 14: Emoji Dataset.....	51
15. Figure 15: Twitter Analysis.....	52
16. Figure 16: Phase II, Graph Analysis.....	53
17. Figure 17: Test Dataset.....	55
18. Figure 18: Phase I Result.....	56
19. Figure 19: Phase II Result.....	57
20. Figure 20: Phase I & II, Graph Analysis.....	58

LIST OF TABLES

1. Table1: Comparison of different SA Approaches3

