**ABSTRACT**

Nowadays, people are enjoying social media online web services like twitter for expressing their opinions and sentiment towards any person/political parties. Such people generated large volume of data or online reviews can be very useful for assessing the knowledge to political parties and even to public to know about the current scenario of politics within that region. In this project, we developed a workflow to integrate both large-scale data mining techniques as well as qualitative anlysis.We focused on public’s Twitter posts, where public post their reviews about political leaders to understand issues and problems that they have with them. Based on these results, we have Implemented a multi-label classification algorithm toclassify tweets. Reflecting public’s reviews about particular political leaders (Narendra Modi,Donald Trump and Rahul Gandhi).Hence ,sentiment analysis or opinion mining aims to use automated tools to detect positive/negative aspects of opinion.

**TABLE OF CONTENTS**

Topics Page No.

Abstract

List of Figures iv

List of Screens v

**1. Introduction**

1.1 Problem Definition 1

1.2 Motivation of the project 1

**2. Literature Survey**

**2.1 Data Mining**

**2.2 Twitter** 2

2.2.1 Why Twitter 2

**2.3 Sentiment analysis**  3

2.3.1 Definition 3

2.3.2 Types of Sentiment Analysis 3

2.3.3 Sentiment Classification techniques 4

**2.4 Algorithms used** 5

2.4.1 Naive bayes 5

2.4.2 Support vector machine 9

**3. Analysis**

**3.1 Introduction** 11

**3.2 Software Requirements**

3.2.1 Introduction 12

3.2.2 Functional Requirements

3.2.3 Non Functional Requirements  **3.3 System Requirements**

3.3.1 Introduction 15

3.3.2 Hardware Requirements 15

3.3.3 Software Requirements 16

**4. Design**

4.1 Introduction 17

4.2 Basic Building Block Of UML 18

4.2.1 Class Diagram 19

4.2.2 Use-case Diagram 21

4.2.3 Sequence Diagram 23

4.2.4 Collaboration Diagram 24

4.2.5 State-chart Diagram 26

4.3 UML Diagrams 36

**5. Implementation and Results**

**5.1 Introduction** 28

5.1.1 How to install R 5.1.2 Downloading RStudio 42

**5.2 Software Design**

5.2.1 Front End Design 29

**5.3 Sample Codes**

5.3.1 Naive Bayes

5.3.2 support vector machine(svm) 30

5.3.3 Web Page Coding

**5.4 Output Screenshots** 55

5.4.1 Twitter Creation

5.4.2 Output Screens of Narendra Modi

5.4.3 Output Screens of Donald Trump

**6. Testing and Validation**

**6.1 Introduction** 73

**6.2 Testing Objectives** 74

**6.3 Test levels**

**6.4 Types of testing**

6.4.1 Unit testing

6.4.2 Black box testing

6.4.3 White box testing

6.4.4 Integration testing

**6.5 Validation**  75

**7. Conclusion**  77

**8. References** 78

**LIST OF FIGURES**

**Figure no.**  **Figure Name Page No.**

1. Fig 2.1 Twitter 8

2. Fig 2.4.2 Support vector machine

3. Fig.4.2.1 Use Case Diagram 36

4. Fig.4.2.2 Class Diagram 37

5. Fig.4.2.3 Sequence Diagram 37

6. Fig.4.2.4 Collaboration Diagram 38

7. Fig.4.2.5 Activity Diagram 38

8. Fig.4.2.6 Deployment Diagram 39

9. Fig.4.2.7 Component Diagram 40

**LIST OF SCREENS**

**Figure no.**  **Figure Name Page No.**

1. Fig 5.2.1 Twitter Mobile No Verification 65

2. Fig 5.2.2 Twitter Email Verification 65

3. Fig 5.2.3 Signing Up Into Twitter 66

4. Fig 5.2.4 Hadoop Home Page 66

5. Fig 5.5 Hadoop Input Directories 67

6. Fig 5.6 Hadoop Output Directory 67

7. Fig 5.7 Hadoop Final Output 68

8. Fig 5.8 Login/Sign up Page 68

9. Fig 5.9 Registration Page 69

10. Fig 5.10 Login Page after registration 69

11. Fig 5.11 Login Page 70

12. Fig 5.12 Home Page 70

13. Fig 5.13 Get Result Button 71

14. Fig 5.14 Viewing Result 71

15. Fig 5.15 Logout Page 72