CONTENTS

CHAPTERS	PAGE NO
1. INTRODUCTION	1
2. LITERATURE SURVEY	4
3. REQUIREMENT SPECIFICATION	8
4. SYSTEM ARCHITECTURE	9
4.1 Twitter Data	9
4.2 Mathematical Model	9
4.3 Software Environment	12
4.3.1 Hadoop Framework	12
4.3.2 Hadoop Distributed File System	13
4.3.3 Hadoop Map Reduced	14
4.3.4 Yarn	16
4.3.5 Flume	17
4.3.6 Hive	19
4.4 System Architecture	21
5. ALGORITHMS	22
6. IMPLEMENTATION	25
6.1 Problem Formulation	25
6.2 Objective	25
6.3 Creation of Twitter Account	26
6.4 Installing Hadoop	27
6.5 Flume Installation	33

6.6 Hive Installation 37
6.7 Installation & Workflow 40
6.8 Twitter Analysis Steps 41

7. RESULTS 44
7.1 Hash Tag Wise Tweet 46
7.2 Hash Tag Wise Tweet Count 46
7.3 Trending Topic of Elections 47
7.4 Sentiment Analysis of Tweet Dataset 48

8. CONCLUSION AND ENHANCEMENT 49

REFERENCES 51

APPENDIX:- Code 53