ACKNOWLEDGEMENT

We would like to take the privilege of the opportunity to express our gratitude in to the project work of **"TWITTER SENTIMENT ANALYSIS USING STREAMING API"** enabled us to express our special thanks to our honourable Chairman of the institution Sir **P.V .VISWAM** .

We are thankful to our honourable Principal **Prof .CH. SUBASH CHANDRA M.Tech (PhD)** who has shown keen interest in us and encouraged us by providing all the facilities to complete our project successfully.

We owe our gratitude to our Beloved head of the Department CSE **Mr. P.SUBBA RAO, M.Tech (PhD)** for assisting us in completing the project work.

We are extremely thankful to our project coordinator **Asst Prof. P.SRI MANIKANTA M.Tech,** who has been a source of inspiration for us throughout our project and for their valuable advice in making our project a success.

We express our sincere Thanks to our supervisor **Mr Asst Prof. P.SRI MANIKANTA M.Tech,** who has been a source of inspiration for us throughout our project and for his valuable advices in making our project successful.

We wish to express our sincere Thanks to all teaching staff of computer science and engineering .We wish to express our special Thanks to all the Faculty members of our college of their concern in subjects and their help throughout our course.

We are the Thankful to our parents and all our friends who had given us good co-operation and suggestions thought our project and helped us successful compilation.

P.PRAVALLIKA	**15B21A0583**
P.SATYANARAYANA	**15B21A0597**
P.J.CHANDANA	**15B21A0577**
K.SHALEM PRASAD	**15B21A05B8**
K.MANIKANTA	**15B21A0572**

ABSTRACT

Opinion mining is the computational study of people's opinions, sentiment, attitudes, and emotions expressed in written language. Sentiment analysis plays an important role in analyzing opinion mining and sentiment in texts. With rapid growth of social media, millions of users are sharing opinions on different aspects of life every day. To extract sentiment on particular topic or person from large data efficient techniques are required to collect data and extract meaningful information from them.

This project aims to provide an interactive automatic system which predicts the sentiment of real time tweets in twitter with particular hash tags using classification techniques in python. A precise method is used for predicting weighted sentiment polarity, which helps to improve marketing strategies.

INDEX

S.NO	CONTENTS	PAGENO
i.	LIST OF ABBREVIATIONS	-
ii.	LIST OF TABLES	-
iii.	LIST OF FIGURES	-
1.	INTRODUCTION	1
2.	LITERATURE SURVEY	3
	2.1. Survey1	3
	2.2. Survey2	3
	2.3. Survey3	4
	2.4. Survey4	5
	2.5. Survey5	6
3.	THEORETICAL BACKGROUND	8
	3.1. Sentiment Analysis	8
	3.2. Types of Sentiment Analysis	8
	3.3. Twitter	9
	3.4. Twitter Sentiment Analysis	9
	3.5. Twitter Streaming API	9
	3.6. Tweepy	10
4.	SYSTEM ANALYSIS	11
	4.1. Existing System	11
	4.2. Proposed system	11
5.	FEASIBILITY STUDY	12
	5.1. Economical feasibility	12
	5.2. Technical feasibility	12
	5.3. Social feasibility	12
6.	SYSTEM REQUIREMENTS	14
	6.1. Hardware requirements	14

	6.2. Software requirements	14
7.	SOFTWARE ENVIRONMENT	15
	7.1. About Python	15
	7.2. Features of Python	15
	7.3. What you can do with Python	18
8.	SYSTEM DESIGN	19
	8.1. Data flow diagram	19
	8.2. Flow Chart diagram	20
	8.3. UML diagrams	21
	8.3.1 Use case	21
	8.3.2 Sequence	23
	8.3.3 Activity	24
9.	IMPLEMENTATION	28
	9.1. Dataset Creation	28
	9.1.1. Positive Tokens	28
	9.1.2. Negative Tokens	30
	9.1.3. Neutral Tokens	32
	9.2. Natural Language Toolkit	32
	9.3. Bag Of Words	33
	9.4. Tokenizing Data	33
	9.5. Pre-Processing Data	34
	9.6. Classification	36
10.	SAMPLE CODING	39
	10.1. Training	39
	10.2. Streaming Tweets	40
	10.3. Classification	42
	10.4. Virtualization	45
11.	TEST CASES	47
	11.1. Test Case1	47

11.2. Test Case2 49
11.3. Test Case3 53
11.4. Test Case4 54
11.5. Test Case5 58
11.6. Test Case6 59
12. CONCLUSION 61
13. BIBILOGRAPHY 62

LIST OF ABBREVIATIONS

NLTK - Natural Language Tool Kit

ML - Machine Learning

NLP - Natural Language Processing

NB - Naïve Baye's

API - Application Programming Interface

BOW - Bag Of Words

IR - Information Retrieval

POS - Parts Of Speech

UML - Unified Modelling Language

AI - Artificial Intelligence

LIST OF TABLES

S.NO	TABLE.NO	CONTENTS	PAGENO
1.	9.1	Data set size	28
2.	9.2	Positive data se	28
3.	9.3	Negative data set	30
4.	9.4	Neutral data set	32

LIST OF FIGURES

S.NO	FIG.NO	CONTENTS	PAGENO
	8.1	Data Flow Diagram	19
	8.2	Flow Chart Diagram	20
	8.3	Use case notation	22
	8.4	Actor notation	22
	8.5	Use case diagram	23
	8.6	Sequence Diagram	24
	8.7	Initial state notation	25
	8.8	Activity state notation	25
	8.9	Control flow notation	25
	8.10	Branch notation	26
	8.11	Fork notation	26
	8.12	Join notation	26
	8.13	Activity Diagram	27
	9.1	Stemming word example	35
	11.1	Word polarity barchart for 10 tweets	53
	11.2	Classification pie chart for 10 tweets	54
	11.3	Word cloud for 10 tweets	54
	11.4	Word polarity barchart for 25 tweets	57
	11.5	Classification pie chart for 25 tweets	58
	11.6	Word cloud for 25 tweets	58
	11.7	Time analysis graph for different datasets	59
	11.8	Time analysis graph for same datasets with different sizes	60