FULL STACK PROJECT REPORT

On

"SENTIMENT ANALYSIS"

Submitted by

Name: DEEPAK SHARMA

Roll No: 171500094

Department of Computer Engineering & Applications

Institute of Engineering & Technology

GLA University

Mathura- 281406, INDIA

2019

ACKNOWLEDGEMENT

I thank the almighty for giving me the courage and perseverance in completing the project.

This project itself is acknowledgements for all those people who have give us their heartfelt co-operation in making this project a grand success.

I extend my sincere thanks to Pankaj Kapoor, Assistant Professor in C.S.E. Department for providing valuable guidance at every stage of this project work. I am profoundly grateful towards the unmatched services rendered by her.

Last but not least , we would like to express our deep sense of gratitude and earnest thanks giving to our dear parents for their moral support and heartfelt cooperation in doing the main project.

Thanks

DEEPAK SHARMA

(171500094)

ABSTRACT

Sentiment analysis has been predominantly used in data science for analysis of customer feedbacks on products and reviews. They are used to understand user ratings on different kinds of products, hospitality services like travel, hotel bookings.

It has also become popular to analyse user tweets — positive, negative or neutral by crawling twitter through APIs.

The objective is to design a website which have the data related to election prediction included graphs and facts.

DEEPAK SHARMA

(171500094)

CONTENTS

CHAPTER 1

Introduction

- **General Introduction**

Sentiment Analysis is the process of determining whether a piece of writing is positive, negative or neutral. A sentiment analysis system for text analysis combines natural language processing (NLP) and machine learning techniques to assign weighted sentiment scores to the entities, topics, themes and categories within a sentence or phrase. Sentiment analysis helps data analysts within large enterprises gauge public opinion, conduct nuanced market research, monitor brand and product reputation, and understand customer experiences. In addition, data analytics companies often integrate third-party sentiment analysis APIs into their own customer experience management, social media monitoring, or workforce analytics platform, in order to deliver useful insights to their own customers.

- **Objectives of the analysis**

In this study, we talk about sentiment analysis of the upcoming Loksabha Elections for the two major national parties in India which are Congress and BJP by crawling tweets from different hashtags of either parties, party leaders, as well as news hashtags

like NDTV. The sentiments analysed covers different user-reactions not only restricted to positive or negative sentiments but covers an in-depth analysis of various positive and negative moods along with the results of different ML models.

We categorise the analytics and research into 3 major milestones:

1. Crawling and scraping tweets in real time from Twitter using the Tweepy API and authentication using the OAuth Handler mechanism.
2. Applying Natural Language Processing techniques to clean the raw tweets from smileys, emojis, special characters and punctuations thereby extracting the important features.
3. Applying standard sentiment analysis techniques using the N-gram model, frequency distribution of words and identification of hate speech and popular sentiments.

CHAPTER 2

Hardware & Software Requirements

Hardware requirement:

Minimum:

-Computer with a 1.6 GHz or faster processor

-1 GB of RAM or more

-2.2 GB of available hard-disk space

-5400 RPM hard drive

-1366 × 768 or higher-resolution display

Recommended:

2.4 GHz or higher CPU, 1024 MB or more RAM, 1280x1024 display, 7200 RPM or higher hard disk.

Software requirement:

Windows 7 or higher

Microsoft Visual Studio 2008 or higher

Bracket

Sublime text

Web Browser

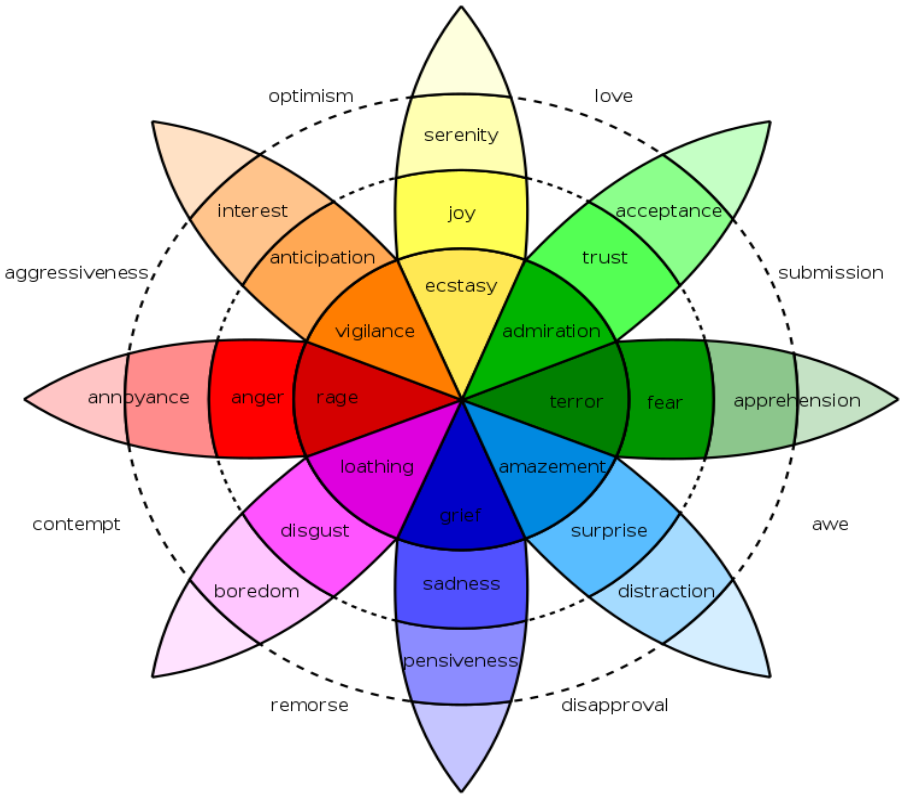
Technologies used:

- Front End : HTML5 , CSS , JAVASCRIPT
- Operating System : Window 10

CHAPTER 3

Identification of popular sentiment

There are many different types of emotions that have an influence on how we live and interact with others. At times, it may seem like we are ruled by these emotions. The choices we make, the actions we take, and the perceptions we have are all influenced by the emotions we are experiencing at any given moment.

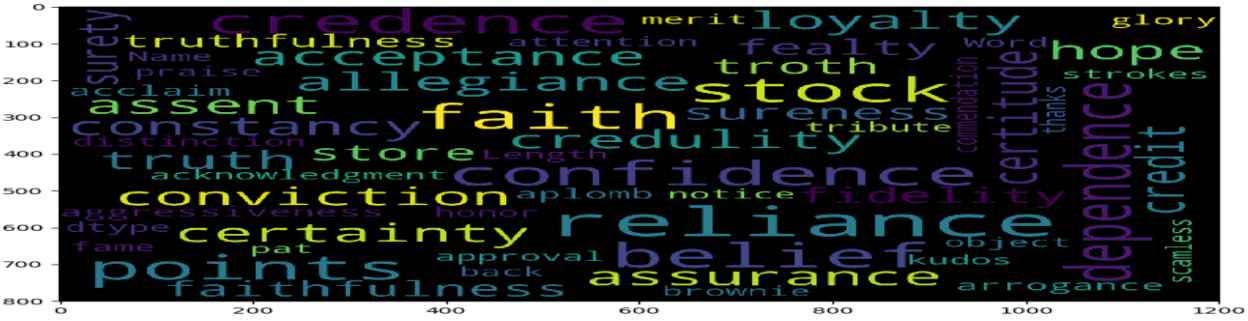


CHAPTER 4
Visualization of all the popular emotions

Anger

Dislike

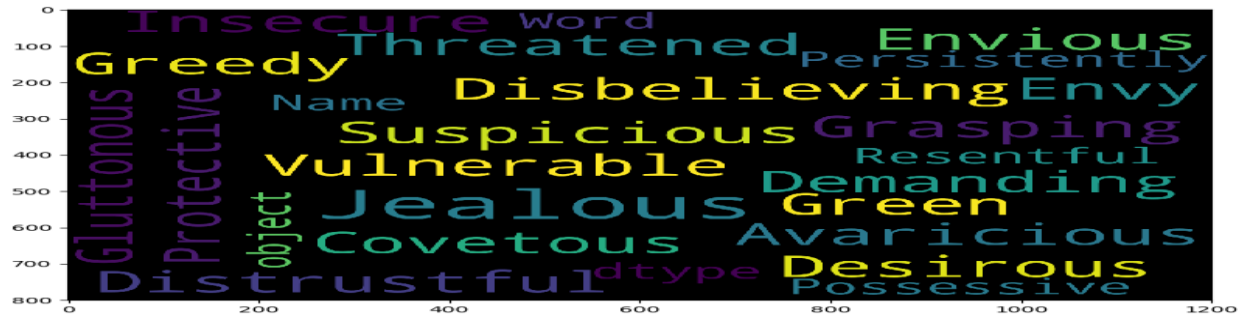
Faith



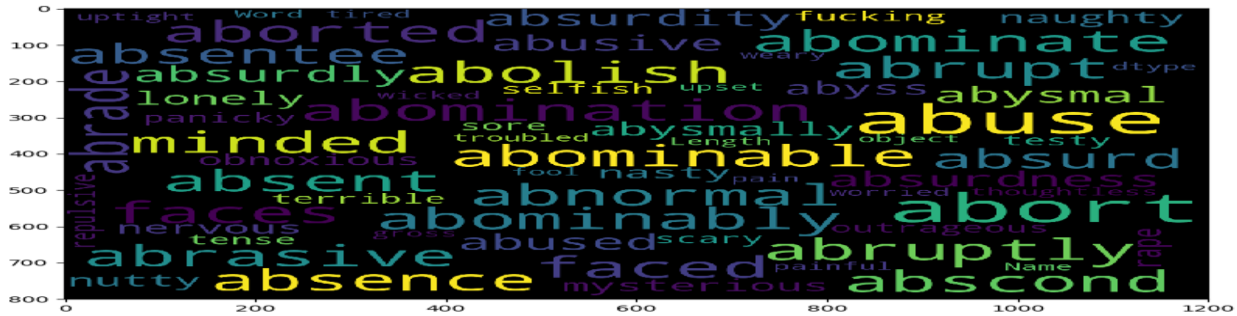
Fear

Happiness

Jealousy



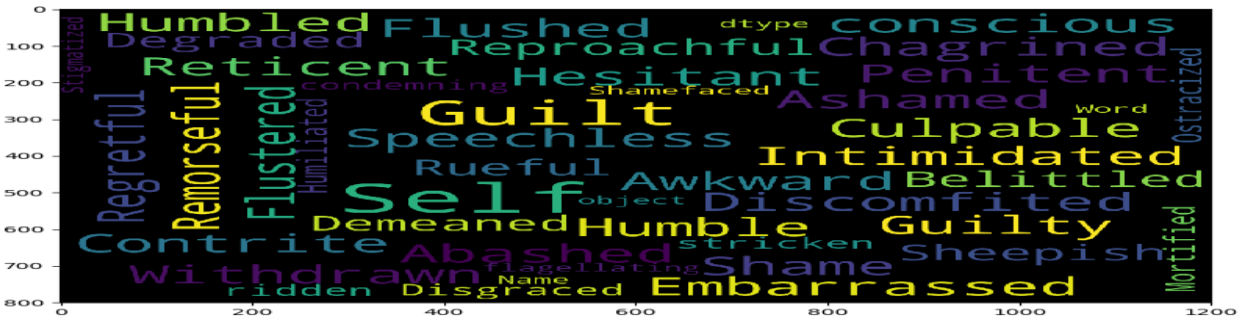
Negative Words



Positive Words

Sadness

Shame

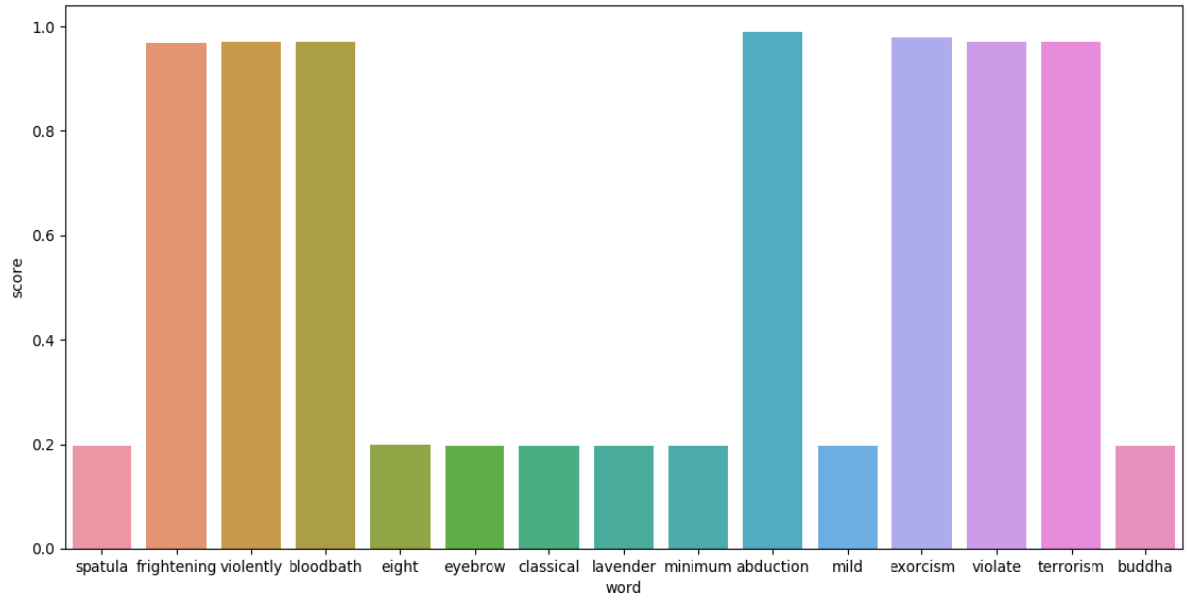


Suicidal

Surprise

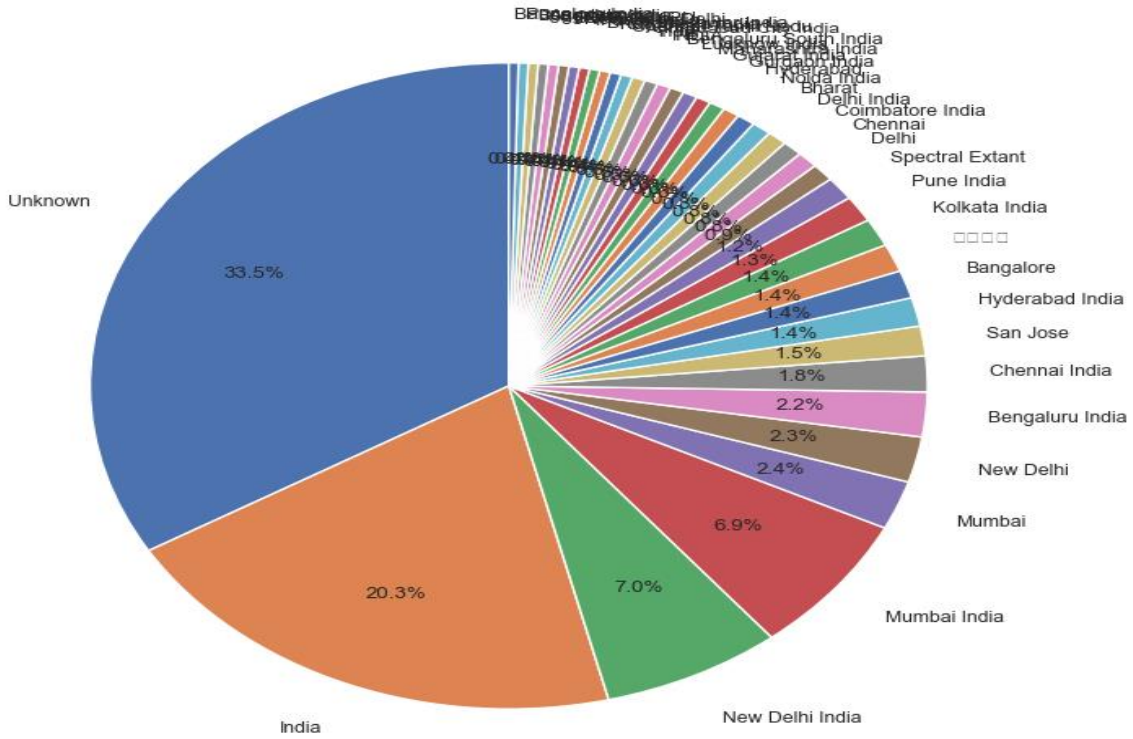
Understanding Twitter's word scoring algorithm

Twitter is automating some processes where suspicious account activity, like exceptionally high-volume tweeting with the same hashtag, or using the same @handle without a reply from the account you're mentioning, is flagged. The account owner has to complete a simple reCAPTCHA process or a password reset request to reclaim access to the account.



Identifying the location density of Twitter's users based on their tagged places

The primary conclusion that can be derived from the pie chart displayed is that most of twitter's user choose to be anonymous as far as the location is concerned. This anonymity might guarantee free speech to certain extent and also protects the first amendment right to speak anonymously against a person, organization or the government.



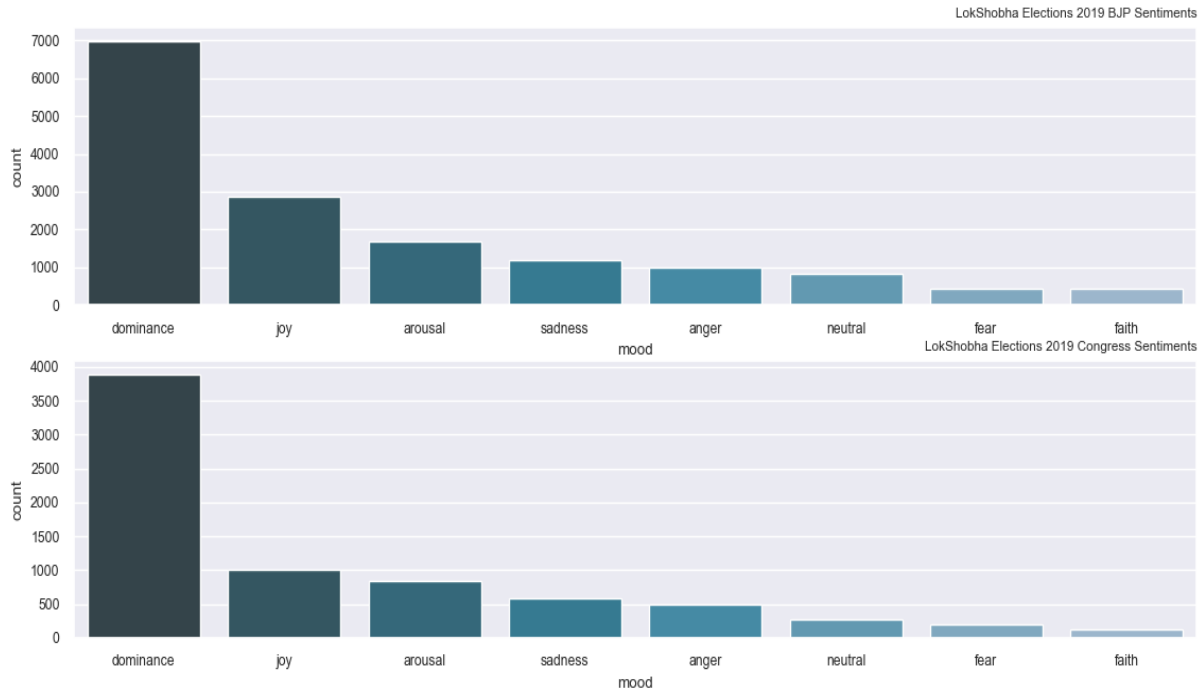
Extracting information from text and classification of sentiments

For any given question, it's likely that someone has written the answer down somewhere. The amount of natural language text that is available in electronic form is truly staggering, and is increasing every day. However, the complexity of natural language can make it very difficult to access the information in that text. The state of the art in NLP is still a long way from being able to build general-purpose representations of meaning from unrestricted text. If we instead focus our efforts on a limited set of questions or "entity relations," such as "where are different facilities located" or "who is employed by what company," we can make significant progress.

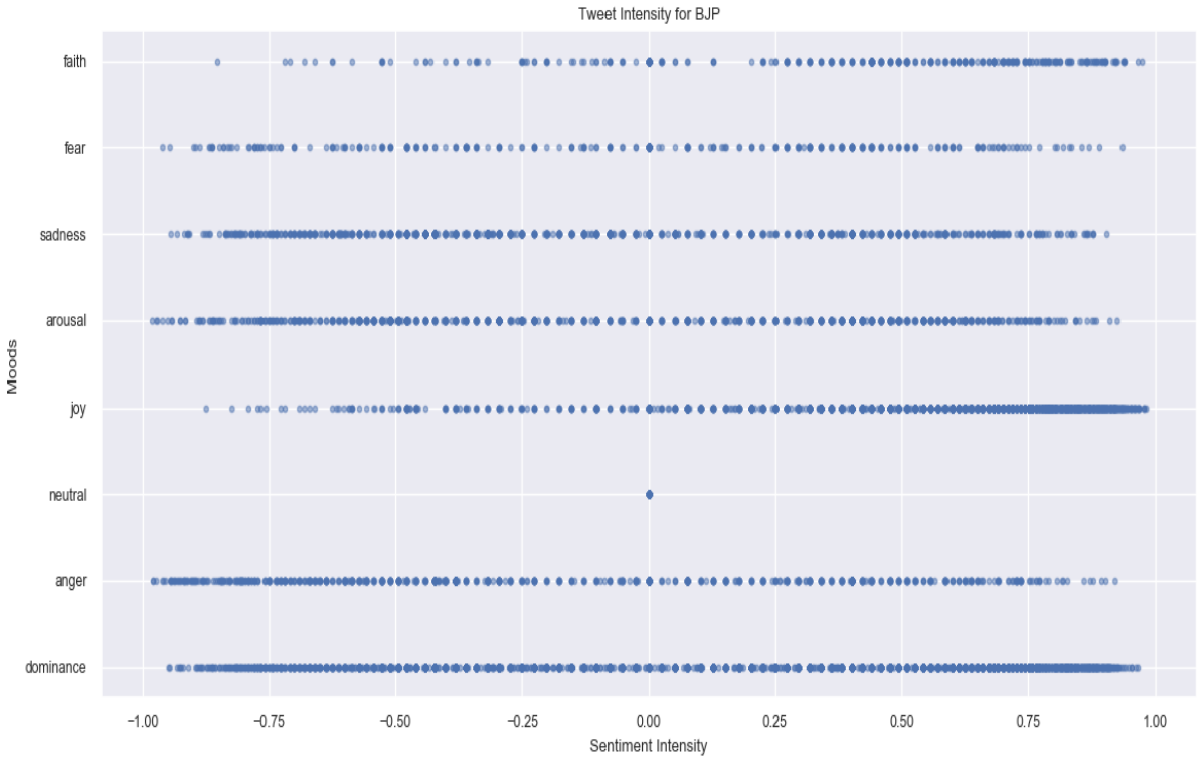
The goal of this chapter is to answer the following questions:

1. How can we identify the mood of a voter who is a hardcore supporter of a particular political party?
2. What are some robust methods for identifying the tweet intensity for a particular party over the years?
3. Which corpora, libraries, or tools are appropriate for this work, and how do we identify the retweet frequency?

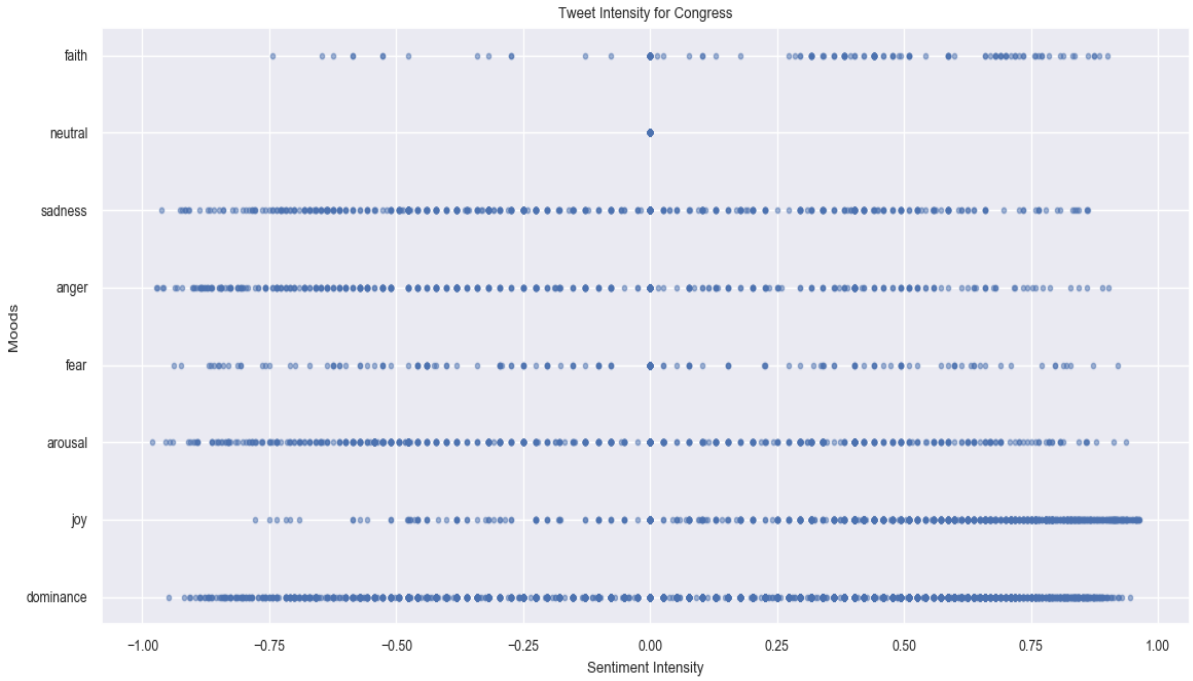
Histogram of sentiments for BJP and Congress



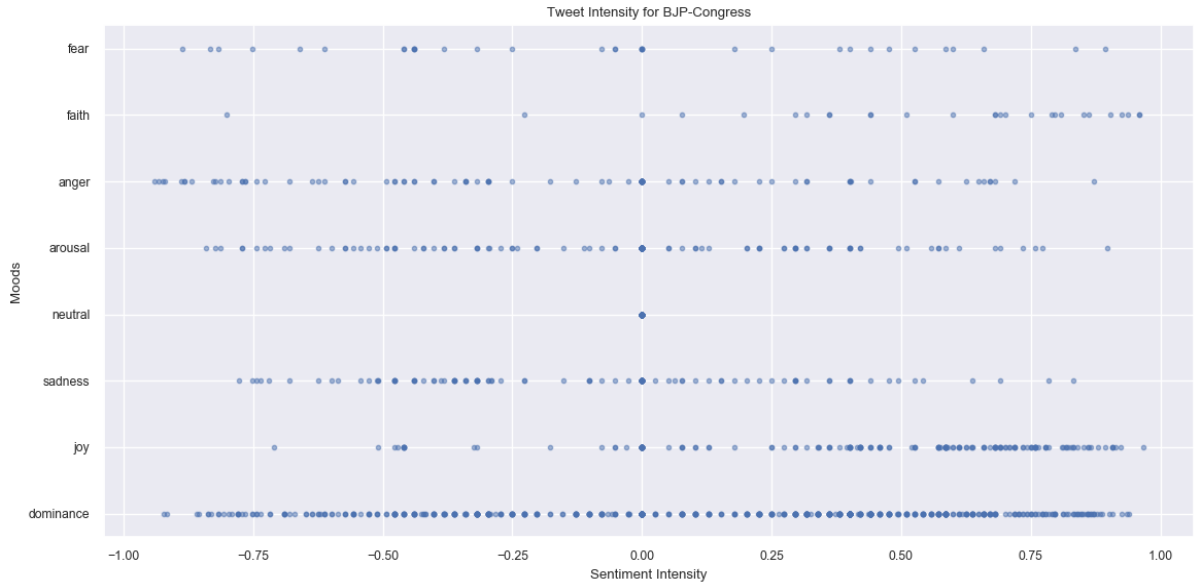
Dot plot representing the mood and tweet intensity for BJP



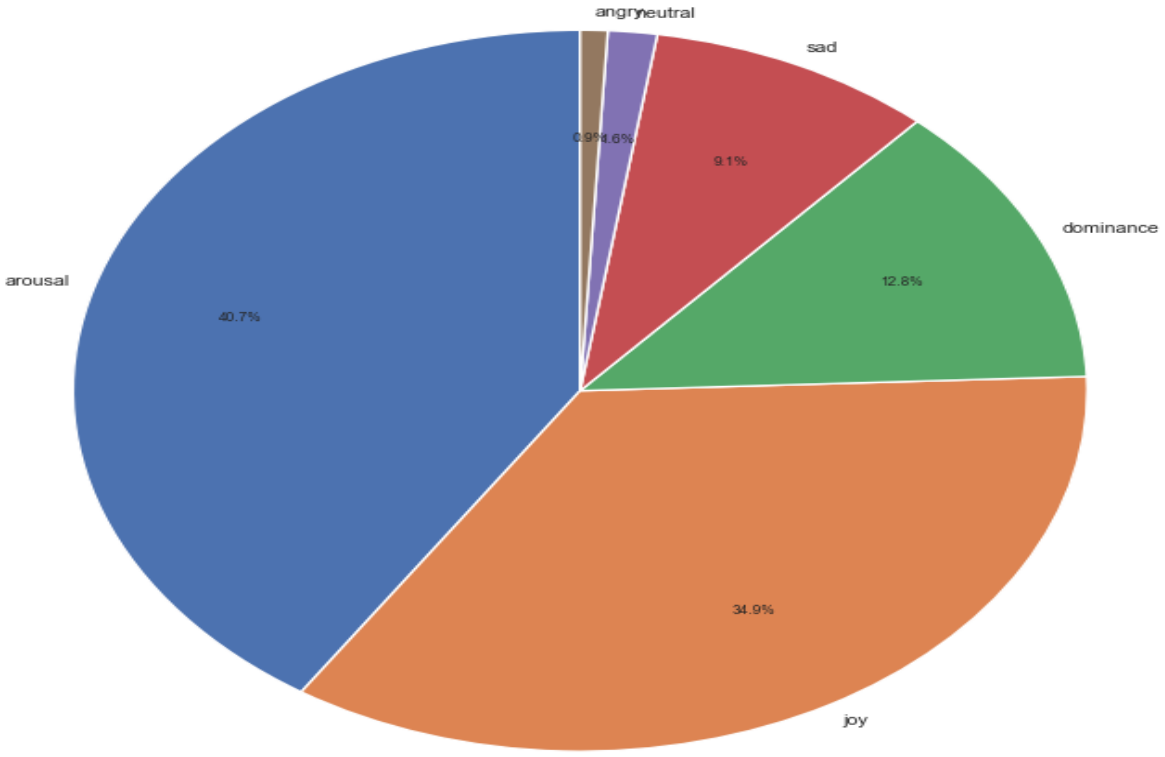
Dot plot representing the mood and tweet intensity for Congress



Dot plot representing the mood and tweet intensity for both the parties BJP and Congress



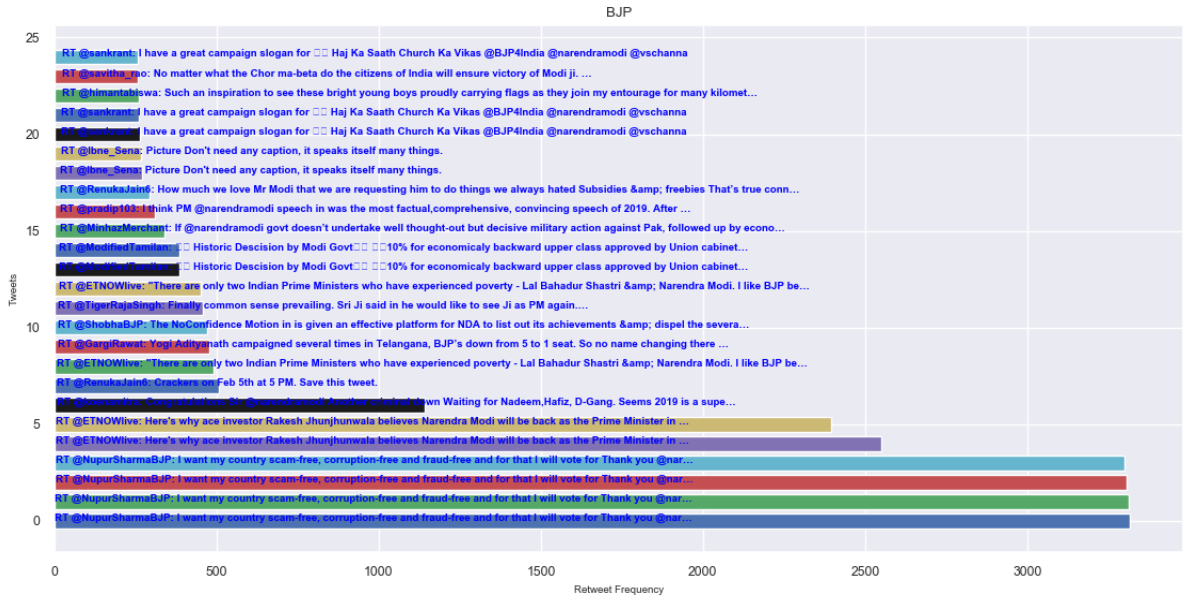
Pie chart representing the sentiment distribution of retweets given to the ruling party that is BJP



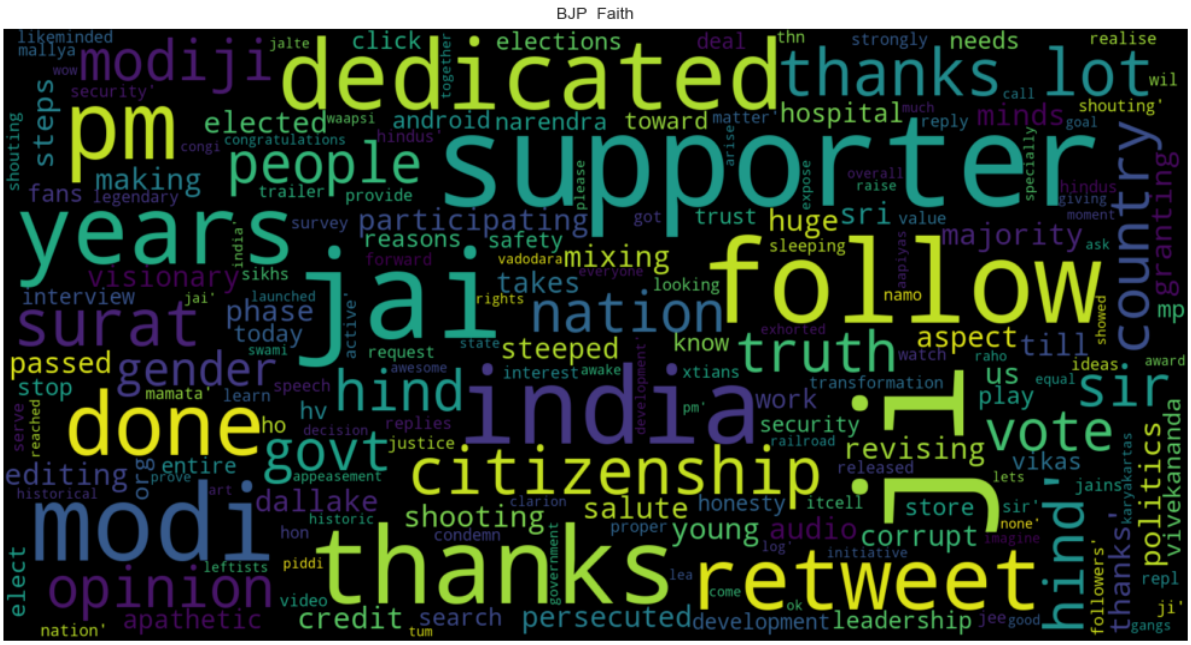
Understanding the online presence of the ruling party (BJP) and analysis of BJP's star campaigners and their tweets

Every political party now understands the importance of online volunteers who can tweet in the favor of a party and hijack the emotions of the normal users. Both BJP and Congress have dedicated IT cells that employ people to tweet in the favor of a particular political leader or policy.

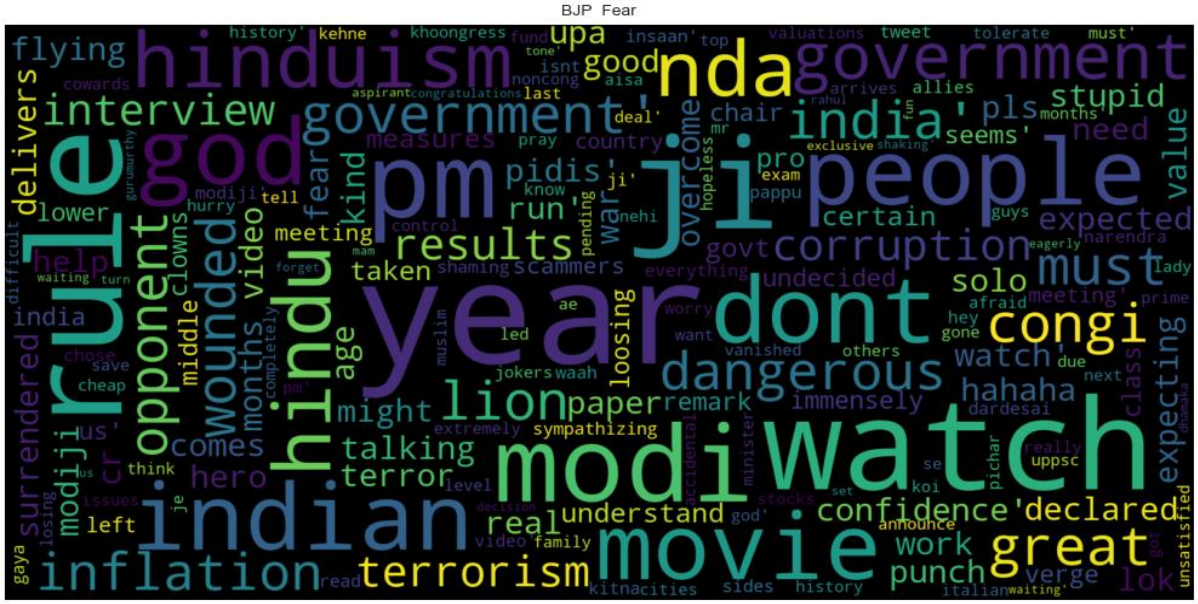
Twitter handle of popular BJP campaigners and their most retweeted tweet



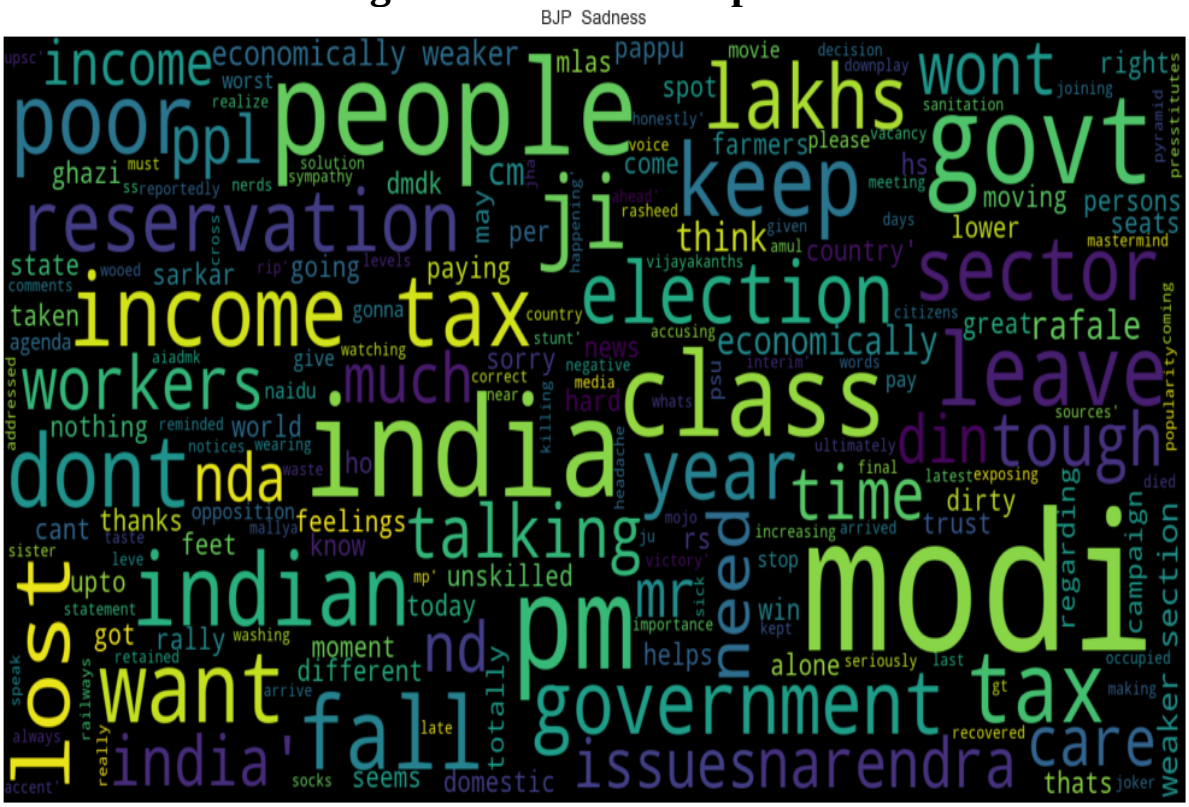
Wordcloud of the corpus that represents faith in BJP



Wordcloud of the corpus that represents fear if BJP is chosen again

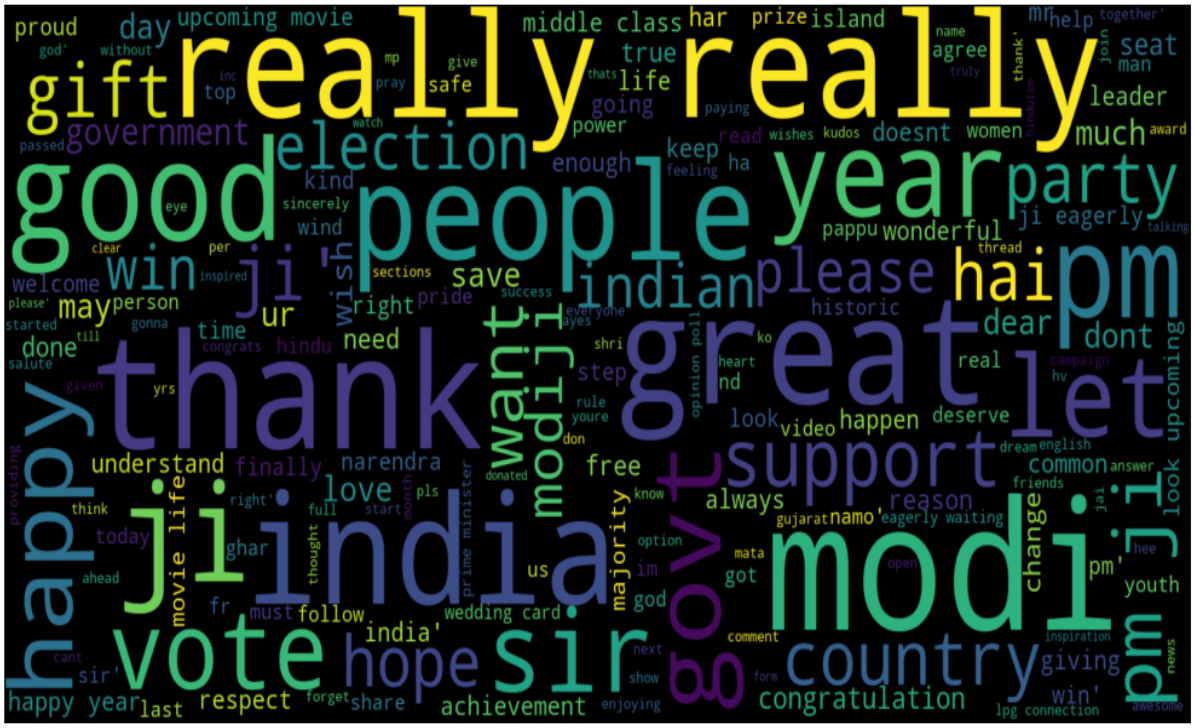


Wordcloud of the corpus that represents sadness with the current government and its policies



Wordcloud of the corpus that represents joy with the current tenure of BJP

BJP Joy


really
really
good
people
year
party
pm
thank
great
let
india
modi
vote
sir
govt
country
support
hope
happy
win
election
gift
proud
day
upcoming movie
middle class
har
prize
island
seat
leader
much
government
ji
please
hai
dont
upcoming
look
pm ji
wish
want
modiji
indian
save
pride
right
need
time
done
ur
may
person
welcome
understand
finally
narendra
love
free
step
dear
real
deserve
happen
video
historic
common
reason
change
always
majority
congratulation
achievement
respect
share
forget
last
happy year
follow
india'
wedding card
movie life
ghar
today
fr
must
ahead
us
got
god
im
option
mata
namo'
gujarat
eagerly
waiting
pm'
youth
giving
win'
lpg connection
awesome
inspiration
news
comment
next
show
form
enjoying
prime minister
opinion poll
thought
sir'
cant
respect
hindu
congrats
shri
ko
heart
nd
rule
youre
don
dream
english
friends
answer
campaign
hv
jai
hee
open
pappu
wonderful
ji eagerly
thread
talking
kudos
doesnt
women
wishes
read
keep
ha
feeling
enough
going
power
paying
life
true
thats
top
safe
give
pray
mp
watch
inc
without
god'
together'
join
thank'
truly
hinduism
award
mr
help
name
agree
man
passed
kind
sincerely
wind
eye
clear
per
inspired
sections
success
everyone
ayes
please'
started
gonna
till
salute
given
yrs
providing
donated
pls
start
full
right'
think
month

CHAPTER 5

Project Screenshots

Keep It Simple.

Put your awesome slogan here...

We All Love Good Typography.

Nov 11, 2019 • Wordpress, Ghost • Jane Doe

Search here...

CATEGORIES.

Wordpress (2)
Ghost (14)
Joomla (5)
Drupal (3)
Magento (2)
Uncategorized (9)

WIDGET TEXT.

Lorem ipsum Ullamco commodo laboris

ABOUT KEEP IT SIMPLE

This is Photoshop's version of Lorem Ipsum. Proin gravida nibh vel velit auctor aliquet. Aenean sollicitudin, lorem quis bibendum auctor, nisi elit consequat ipsum, nec sagittis sem nibh id elit.

Lorem ipsum Sed nulla deserunt voluptate elit occaecat culpa cupidatat sit irure sint sint incididunt cupidatat esse in Ut sed commodo tempor consequat culpa fugiat incididunt.

PHOTOSTREAM

NAVIGATE

- Home
- Blog
- Demo
- Archives
- About

CHAPTER 6

Applications

1. Tracking Your Employees' Feedback
2. Improving Your Customer Support
3. Providing Better Product Analytics
4. Keeping an Eye on Your Competition
5. Tracking User Generated Content
6. Social Media Monitoring
7. Managing a Crisis Better

CHAPTER 7

Conclusion

Sentiment analysis or opinion mining is a field of study that analyzes people's sentiments, attitudes, or emotions towards certain entities. This paper tackles a fundamental problem of sentiment analysis, sentiment polarity categorization. Online product reviews from Amazon.com are selected as data used for this study. A sentiment polarity categorization process has been proposed along with detailed descriptions of each step. Experiments for both sentence-level categorization and review-level categorization have been performed.

The age of getting meaningful insights from social media data has now arrived with the advance in technology. The Uber case study gives you a glimpse of the power of Contextual Semantic Search. It's time for your organization to move beyond overall sentiment and count based metrics. Companies have been leveraging the power of data lately, but to get the deepest of the information, you have to leverage the power of AI, Deep learning and intelligent classifiers like Contextual Semantic Search and Sentiment Analysis.

CHAPTER 8

References

Internet Sources:
[1] www.mysql.com
[2] www.php.net
[3]www.apache.org
[4] www.stackoverflow.com
GitHub:
https://github.com/iamdeepak55