VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELAGAVI – 590018,Karnataka

INTERNSHIP REPORT

ON

"Twitter Sentiment Analysis For Stocks"

Submitted in partial fulfillment for the award of degree(18CSI85)

BACHELOR OF ENGINEERING

IN

Computer Science and Engineering

Submitted by:

NAME:PRASANT POUDEL

USN:1BH19CS069

Varcons Technologies Pvt Ltd

Conducted at

Varcons Technologies Pvt Ltd

VTU Affiliated AICTE Approved BTI BANGALORE TECHNOLOGICAL INSTITUTE

BANGALORE TECHNOLOGICAL INSTITUTE

Department of Computer Science and Engineering

Approved By AICTE, New Delhi

Kodathi, Off Bangalore - Sarjapur Road, Bangalore East Taluk, Chikkakannalli, Bengaluru, Karnataka 560035

BANGALORE TECHNOLOGICAL INSTITUTE
Department of Computer Science and Engineering
Approved By AICTE, New Delhi

Kodathi, Off Bangalore - Sarjapur Road, Bangalore East Taluk, Chikkakannalli, Bengaluru, Karnataka 560035

CERTIFICATE

This is to certify that the Internship titled **"Twitter Sentiment Analysis For Stocks"** carried out by **Mr.Prasant Poudel,** a bonafide student of Bangalore Technological Institute, in partial fulfillment for the award of **Bachelor of Engineering**, in **Computer Science and Engineering** under Visvesvaraya Technological University, Belagavi, during the year 2022-2023. It is certified that all corrections/suggestions indicated have been incorporated in the report.

The project report has been approved as it satisfies the academic requirements in respect of Internship prescribed for the course Internship / Professional Practice (18CSI85)

Signature of Guide	**Signature of HOD**	**Signature of Principal**
……	……	……

External Viva:

Name of the Examiner	Signature with Date
……	……

D E C L A R A T I O N

I,**Prasant Poudel**, final year student of Computer Science and Engineering, **Bangalore Technological Institute -560035**, declare that the Internship has been successfully completed, in **VARACONS TECHNOLOGIES PVT LTD**. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Branch name, during the academic year 2022-2023.

Date :23/09/22 :

Place:Ramanashree Arcade,18 M G Road,Ulsoor,Bangalore-560001

NAME :Prasant Poudel
USN :1BH19CS069

OFFER LETTER

Internship Offer Letter

Date: **23rd August, 2022**

Name: **Prasant Poudel**
USN: **1BH19CS069**

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning With Python(Research Based)** Internship position with **Varcons Technologies Pvt Ltd**, effective Start Date **23rd August, 2022**, All of us are excited about this opportunity provided to you!

This internship is viewed as being an educational opportunity for you, rather than a part-time job. As such, your internship will include training/orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts of **Machine Learning With Python(Research Based)** through hands-on application of the knowledge you learn while you train with the senior developers. You will be bound to follow the rules and regulations of the company during your internship duration.

Again, congratulations and we look forward to working with you!

Sincerely,

Spoorthi H C
Director
Varcons Technologies Pvt Ltd
213, 2st Floor,
18 M G Road, Ulsoor,
Bangalore-560001

ACKNOWLEDGEMENT

Any achievement, be it scholastic or otherwise does not depend solely on the individual efforts but on the guidance, encouragement and cooperation of intellectuals and elders. We would like to take this opportunity to thank them all.

We heartily extend our words of gratitude to the technical guide of **Varacon Technologies Pvt. Ltd**. , for his valuable advice, encouragement and suggestion given to our team in the course of our Internship project. We convey our gratitude to her for having constantly monitored the development of the mini- project and setting up precise deadlines.

We would like to express our immense gratitude to Head of Department **Dr. Sohan Kumar Gupta**, for his unfailing encouragement and suggestions given to us in the course of our work.

We would like to take this opportunity to express our gratitude to the Principal,**Dr. H S Nanda**, for giving us this opportunity to enrich our knowledge.

We are grateful to the President **Dr. A Prabhakara Reddy** and Secretary, **Sri. C L Gowda** for having provided us with a great infrastructure and well-furnished labs.

Finally, a note of thanks to the Department of Computer Science and Engineering, both teaching and non-teaching staff for their cooperation extended to us.

Last but not the least, we acknowledge the support and feedback of our parents, guardians and friends, for their indispensable help always.

NAME: PRASANT POUDEL

USN: 1BH19CS069

ABSTRACT

The stock market can be influenced and moved by a variety of variables, including external and internal influences.Because of supply and demand fluctuations, stock prices fluctuate every second.Machine learning, on the other hand, will provide a more accurate, precise, and sensible process for resolving stock and market price concerns. ML algorithms have been employed to improve new ways of building simulation models that can estimate stock market movements and whether they will gain or lose. Support vector machines (SVM), Naive Bayes regression, and other approaches were used in several sentiment analysis research.

.

Table of Contents

CHAPTER 1

COMPANY PROFILE

1. COMPANY PROFILE

A Brief History of Varacons Technologies Pvt Ltd

Varacons Technologies, was incorporated with a goal ”To provide high quality and optimal Technological Solutions to business requirements of our clients”. Every business is different and has a unique business model and so are the technological requirements. They understand this and hence the solutions provided to these requirements are different as well. They focus on clients requirements and provide them with tailor made technological solutions. They also understand that Reach of their Product to its targeted market or the automation of the existing process into e-client and simple processes are the key features that our clients desire from the Technological Solution they are looking for and these are the features that we focus on while designing the solutions for their clients.

Sarvamoola Software Services. is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Sarvamoola Software Services. specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements.

Varacons Technologies, strive to be the front runner in creativity and innovation in software development through their well-researched expertise and establish it as an out of the box software development company in Bangalore, India. As a software development company, they translate this software development expertise into value for their customers through their professional solutions.

They understand that the best desired output can be achieved only by understanding the client's demand better. Varacons Technologies work with their clients and help them to define their exact solution requirements. Sometimes even they wonder if they have completely redefined their solution or new application requirement during the brainstorming session, and here they position themselves as an IT solutions consulting group composed of high caliber consultants.

They believe that Technology when used properly can help any business to scale and achieve new heights of success. It helps Improve its efficiency, profitability, reliability; to put it in one sentence ” Technology helps you to Delight your Customers” and that is what we want to achieve.

CHAPTER 2

ABOUT THE COMPANY

2. ABOUT THE COMPANY

Varacons Technologies is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Compsoft Technologies specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements. The organization where they have a right mix of professionals as stakeholders to help us serve our clients with best of our capability and at par industry standards. They have young, enthusiastic, passionate and creative Professionals to develop technological innovations in the field of Mobile technologies, Web applications as well as Business and Enterprise solutions. Motto of our organization is to "Collaborate with our clients to provide them with the best Technological solution hence creating a Good Present and Better Future for our client which will bring a cascading positive effect in their business shape as well". Providing a Complete suite of technical solutions is not just our tagline, it is Our Vision for Our Clients and for Us, We strive hard to achieve it.

Products of Varacons Technologies.

Android Apps

It is the process by which new applications are created for devices running the Android operating system. Applications are usually developed in Java (and/or Kotlin; or other such option) programming language using the Android software development kit (SDK).

The Android software development kit includes a comprehensive set of development tools. These include a debugger, libraries, a handset emulator based on QEMU, documentation, sample code, and tutorials. Currently supported development platforms include computers running Linux (any modern desktop Linux distribution), Mac OS X 10.5.8 or later, and Windows 7 or later. As of March 2015, the SDK is not available on Android itself, but software development is possible by using specialized Android applications.

Web Application

It is a client–server computer program in which the client (including the user interface and client- side logic) runs in a web browser. Common web applications include web mail, online

retail sales, online auctions, wikis, instant messaging services and many other functions. web applications use web documents written in a standard format such as HTML and JavaScript,which are supported by a variety of web browsers. Web applications can be considered as a specific variant of client–server software where the client software is downloaded to the client machine when visiting the relevant web page, using standard procedures such as HTTP. The Client web software updates may happen each time the web page is visited. During the session, the web browser interprets and displays the pages, and acts as the universal client for any web application. The use of web application frameworks can often reduce the number of errors in a program, both by making the code simpler, and by allowing one team to concentrate on the framework while another focuses on a specified use case. In applications which are exposed to constant hacking attempts on the Internet, security-related problems can be caused by errors in the program.

Frameworks can also promote the use of best practices such as GET after POST. There are some who view a web application as a two-tier architecture. This can be a “smart” client that performs all the work and queries a “dumb” server, or a “dumb” client that relies on a “smart” server. The client would handle the presentation tier, the server would have the database (storage tier), and the business logic (application tier) would be on one of them or on both. While this increases the scalability of the applications and separates the display and the database, it still doesn‟t allow for true specialization of layers, so most applications will outgrow this model. An emerging strategy for application software companies is to provide web access to software previously distributed as local applications. Depending on the type of application, it may require the development of an entirely different browser-based interface, or merely adapting an existing application to use different presentation technology. These programs allow the user to pay a monthly or yearly fee for use of a software application without having to install it on a local hard drive. A company which follows this strategy is known as an application service provider (ASP), and ASPs are currently receiving much attention in the software industry.

Security breaches on these kinds of applications are a major concern because it can involve both enterprise information and private customer data. Protecting these assets is an important part of any web application and there are some key operational areas that must be included in the development process. This includes processes for authentication, authorization, asset handling, input, and logging and auditing. Building security into the applications from the beginning can be more effective and less disruptive in the long run.

Web design

It encompasses many different skills and disciplines in the production and maintenance of websites. The different areas of web design include web graphic design; interface design; authoring, including standardized code and proprietary software; user experience design; and

search engine optimization. The term web design is normally used to describe the design process relating to the front-end (client side) design of a website including writing mark up. Web design partially overlaps web engineering in the broader scope of web development. Web designers are expected to have an awareness of usability and if their role involves creating markup then they are also expected to be up to date with web accessibility guidelines. Web design partially overlaps web engineering in the broader scope of web development.

Departments and services offered

Varacons Technologies plays an essential role as an institute, the level of education, development of student's skills are based on their trainers. If you do not have a good mentor then you may lag in many things from others and that is why we at Varacons Technologies gives you the facility of skilled employees so that you do not feel insecure about the academics. Personality development and academic status are some of those things which lie in the mentor's hands. If you are trained well then you can do well in your future and knowing its importance of Varacons Technologies always tries to give you the best.

They have a great team of skilled mentors who are always ready to direct their trainees in the best possible way they can and to ensure the skills of mentors we held many skill development programs as well so that each and every mentor can develop their own skills with the demands of the companies so that they can prepare a complete packaged trainee.

Services provided by Varacons Technologies.

• Core Java and Advanced Java

• Web services and development

• Dot Net Framework

• Python

• Selenium Testing

• Conference / Event Management Service

• Academic Project Guidance

• On The Job Training

• Software Training

CHAPTER 3

INTRODUCTION

3. INTRODUCTION

Introduction to ML

Machine learning is programming computers to optimize a performance criterion using example data or past experience . We have a model defined up to some parameters, and learning is the execution of a computer program to optimize the parameters of the model using the training data or past experience. The model may be predictive to make predictions in the future, or descriptive to gain knowledge from data.The field of study known as machine learning is concerned with the question of how to construct computer programs that automatically improve with experience.Currently, it is being used for various tasks such as image recognition, speech recognition, email filtering, Facebook auto-tagging, recommender system, and many more.Thus, machine learning algorithm is a set of rules and statistical techniques used to learn patterns from data and draw significant information from it. It is the logic behind a Machine Learning model.

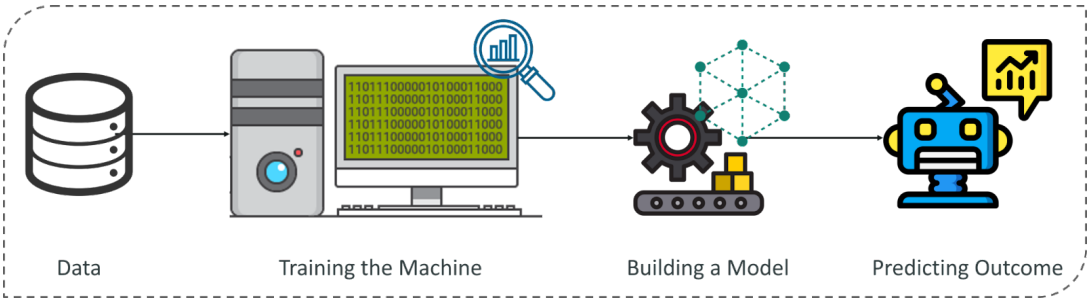


Figure: Overall machine Learning Phenomena

Problem Statement

Stock exchange is a subject that is highly affected by economic, social, and political factors. There are several factors e.g. external factors or internal factors which can affect and move the stock market. Stock prices rise and fall every second due to variations in supply and demand. But techniques using machine learning will give a more accurate, precise and simple way to solve such issues related to stock and market prices. The changes in stock prices of a company, the rises and falls, are correlated with the public opinions being expressed in tweets about that company. Understanding the author's opinion from a piece of text is the objective of sentiment analysis. Positive news and tweets in social media about a company would definitely encourage people to invest in the stocks of that company and as a result the stock price of that company would increase. A prediction model for finding and analyzing correlation between contents of tweets and stock prices and then making predictions for future prices can be developed by using machine learning.

CHAPTER 4
REQUIREMENT ANALYSIS

4. REQUIREMENT ANALYSIS

Hardware Requirement Specification

System	: Window OR Mac
RAM	: 8GB Min
Disk Space	: 100GB
CPU Architecture	: i5 or AMD5 Min
Generation	: 10 Gen

Software Requirement Specification

Coding Language	: Python
Tools	: Jupyter Notebook , Visual Studio Code
Environment	: Anaconda Environment
Library	: Scikit-learn, Tensorflow

CHAPTER 5

DESIGN ANALYSIS

5.DESIGN & ANALYSIS

5.1 PROBLEM STATEMENT & APPLICABILITY

Since the Advent of the Internet, humans have used it as a communication tool in the form of mostly text messages and nowadays video and audio streams and as we increase our dependence on technology it becomes increasingly important to better gauge human sentiments expressed with the help of technology. However, in this textual communication data, we lose the access to sentiments or the emotions conveyed behind a sentence, as we often use our hands and facial expressions to express our intent behind the statement. From this textual data, we can gain insights into the individual. Insights which can be used for multiple different uses such as content recommendation based on current mood, market segmentation analysis and psychological analysis in humans. In this project, we have attempted to classify human sentiment based on the tweet based on the stock prediction into two categories namely positive and negative. Which helps us better understand human thinking and gives us an insight which can be used in a variety of ways as stated above.

5.2 PROPOSED METHODOLOGY

In this paper we classify sentiments with the help of machine learning and natural language processing (NLP) algorithms, we use the datasets from Kaggle which was crawled from the internet and labeled positive/negative. The data provided comes with emoticons (emoji), usernames and hashtags which are required to be processed (so as to be readable) and converted into a standard form. We also need to extract useful features from the text such as unigrams and bigrams which is a form of representation of the “tweet”. We use various machine learning algorithms based on NLP (Natural Language Processing) to conduct sentiment analysis using the extracted features. Finally we report our experimental results and findings at the end.

5.3 DATA DESCRIPTION

The data given from the dataset is in the form of comma separated values files with “tweets” and their corresponding sentiments. The training dataset is a csv (comma separated value) file of type tweet_id, sentiment, tweet where the tweet_id is a unique integer identifying the tweet, sentiment is either 1 (positive) or 0 (negative), and tweet is the tweet enclosed in "". Similarly, the test dataset is a csv file of type tweet_id, tweet respectively. The dataset is a mixture of words, emoticons, symbols, URLs and references to people as seen usually on twitter. Words and emoticons contribute to predicting the sentiment, but URLs and references to people don’t. Therefore, URLs and references are being ignored. The words are also a mixture of misspelled words / incorrect, extra punctuations, and words with many repeated letters. The “tweets", therefore, must be preprocessed to standardize the dataset. The provided training and test dataset have 32000 and 2000 tweets respectively.

5.4 PREPROCESSING

Raw tweets scraped from twitter generally result in a noisy and obscure dataset. This is due to the casual and ingenious nature of people's usage of social media. Tweets have certain special characteristics such as retweets, emoticons, user mentions, etc. which should be suitably extracted. Therefore, raw twitter data must be normalized to create a dataset which can be easily learned by various classifiers. We have applied an extensive number of pre-processing steps to standardize the dataset and reduce its size. We first do some general preprocessing on tweets.which is as follows:

- Convert the tweet characters to lowercase letters.
- Taking only those characters which are alphanumeric.
- Removing the punctuation and stopword from the text sentence.
- Steaming the characters.

We handle special twitter features as follows:

5.4.1 Uniform Resource Locator (URL)

Users often share hyperlinks to other webpages in their tweets. Any particular given URL is not important for text classification as it would lead to very sparse features and incorrect classification. Therefore, we replace all the URLs in tweets with the word URL. The regular expression used to match URLs is ((www\.[\S]+)|(https?://[\S]+)).

5.4.2 User Mention

Every twitter user has a handle associated with them. Users often mention other users in their tweets by @handle. We replace all user mentions with the word USER_MENTION. The regular expression (regex) used to match user mention is @[\S]+.

5.5 Feature Extraction

We extract two types of features from our dataset, namely unigrams and bigrams. We create a frequency distribution of the unigrams and bigrams present in the dataset and choose top N unigrams and bigrams for our analysis.

5.5.1 Unigrams

Probably the simplest and the most commonly used features for text classification is the presence of single words or tokens in the text. We extract single words from the training dataset and create a frequency distribution of these words. A total of 181232 unique words are extracted from the dataset. Out of these words, most of the words at the end of frequency spectrum are noise and occur very few times to influence classification. We, therefore, only use top N words from these to create our vocabulary where N is 15000 for sparse vector classification.

5.5.2 Bigrams

Bigrams are word pairs in the dataset which occur in succession in the corpus. These features are a good way to model negation in natural language like the phrase– This is not good. A total of 1954953 unique bigrams were extracted from the dataset. Out of these, most of the bigrams at the end of the frequency spectrum are noise and occur very few times to influence classification. We therefore use only top 10000 bigrams from these to create our vocabulary.

5.6 MODEL

We will be using a model consisting of both test and training dataset models using various algorithms as due to the modular nature of the program we can add and remove the algorithms with ease. Let's understand the workflow of our system with the help of the above diagram. First, we have split the data into training and test sets. We also keep separate positive and negative pre-labelled datasets for training the model and checking their generalization classification in the test set. After this the training data is fed to some machine learning algorithms like Naive Bayes, RandomForestClassifier,DecisionTree, XGDBoost,SGDClassifier,KNeighborsClassifier,which learns to make predictions. To evaluate our system, we use Baseline Classification which is our evaluation metric in which test data is fed to the learned algorithm which in return generates recommended prediction ratings of words. With the help of a pre-classified golden set and evaluation metric we check the accuracy of our model.

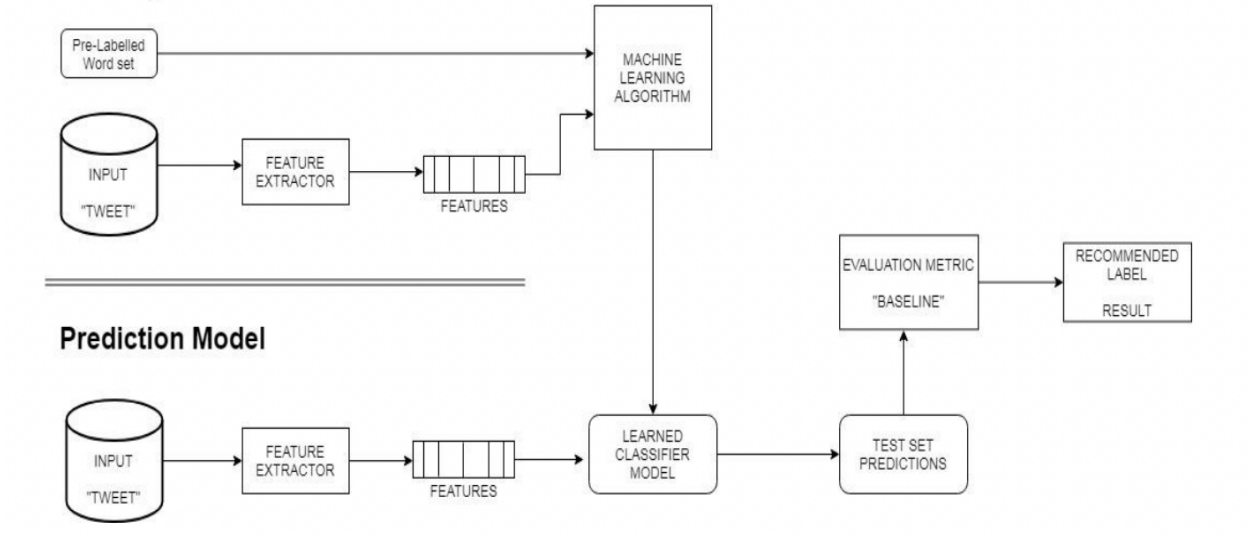


Fig 3: Machine Learning Model for Twitter System Analysis System

5.7 EVALUATION METRICS

For evaluation metrics, we use the baseline algorithm which uses simple positive and negative word counting methods to assign sentiment to a given tweet. We use the Golden Dataset of positive and negative words to classify tweets. In cases when the magnitude of positive and negative words is equal, we assign positive sentiment. A baseline is a method that uses heuristics, simple summary statistics, randomness, or machine learning to create predictions for a dataset. We can use these predictions to measure the baseline performance (e.g. accuracy) this metric will then become what we compare any other machine learning algorithm against.

5.8 Future Work

5.8.1 Handling Emotion Ranges

We can improve and train our models to handle a range of sentiments. Tweets don't always have positive or negative sentiment. At times they may have no sentiment i.e. neutral. Sentiment can also have gradations like the sentence, This is good, is positive but the sentence, This is extraordinary. Is somewhat more positive than the first. We can therefore classify the sentiment in ranges; say from -2 to +2.

5.8.2 Using symbols

During our pre-processing, we discard most of the symbols like commas, full-stops, and exclamation marks. These symbols may be helpful in assigning sentiment to a sentence.

CHAPTER 6

IMPLEMENTATION

IMPLEMENTATION

Implementation is the stage where the theoretical design is turned into a working system. The most crucial stage in achieving a new successful system and in giving confidence on the new system for the users that it will work efficiently and effectively.

The system can be implemented only after thorough testing is done and if it is found to work according to the specification. It involves careful planning, investigation of the current system and its constraints on implementation, design of methods to achieve the change over and an evaluation of change over methods apart from planning.

Two major tasks of preparing the implementation are education and training of the users and testing of the system. The more complex the system being implemented, the more involved will be the system analysis and design effort required just for implementation.

The implementation phase comprises several activities. The required hardware and software acquisition is carried out. The system may require some software to be developed. For this, programs are written and tested. The user then changes over to his new fully tested system and the old system is discontinued.

TESTING

The testing phase is an important part of software development. The Information zed system will help in automating the process of finding errors and missing operations and also a complete verification to determine whether the objectives are met and the user requirements are satisfied. Software testing is carried out in three steps:

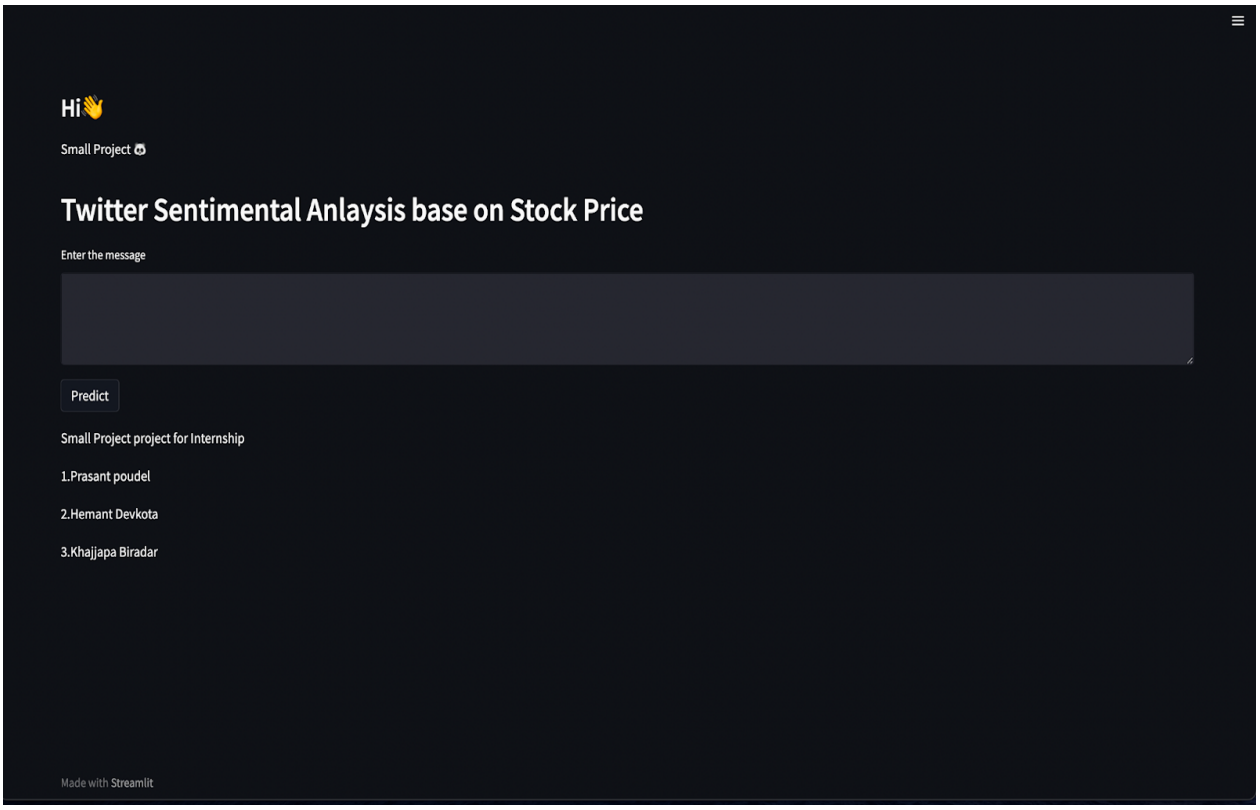
1. The first includes unit testing, where each module is tested to provide its correctness, validity and also determine any missing operations and to verify whether the objectives have been met. Errors are noted down and corrected immediately.
2. Unit testing is the important and major part of the project. So errors are rectified easily in particular modules and program clarity is increased. In this project the entire system is divided into several modules and is developed individually. So unit testing is conducted to individual modules.
3. The second step includes Integration testing. It need not be the case, the software whose modules when run individually and showing perfect results, will also show perfect results when run as a whole.

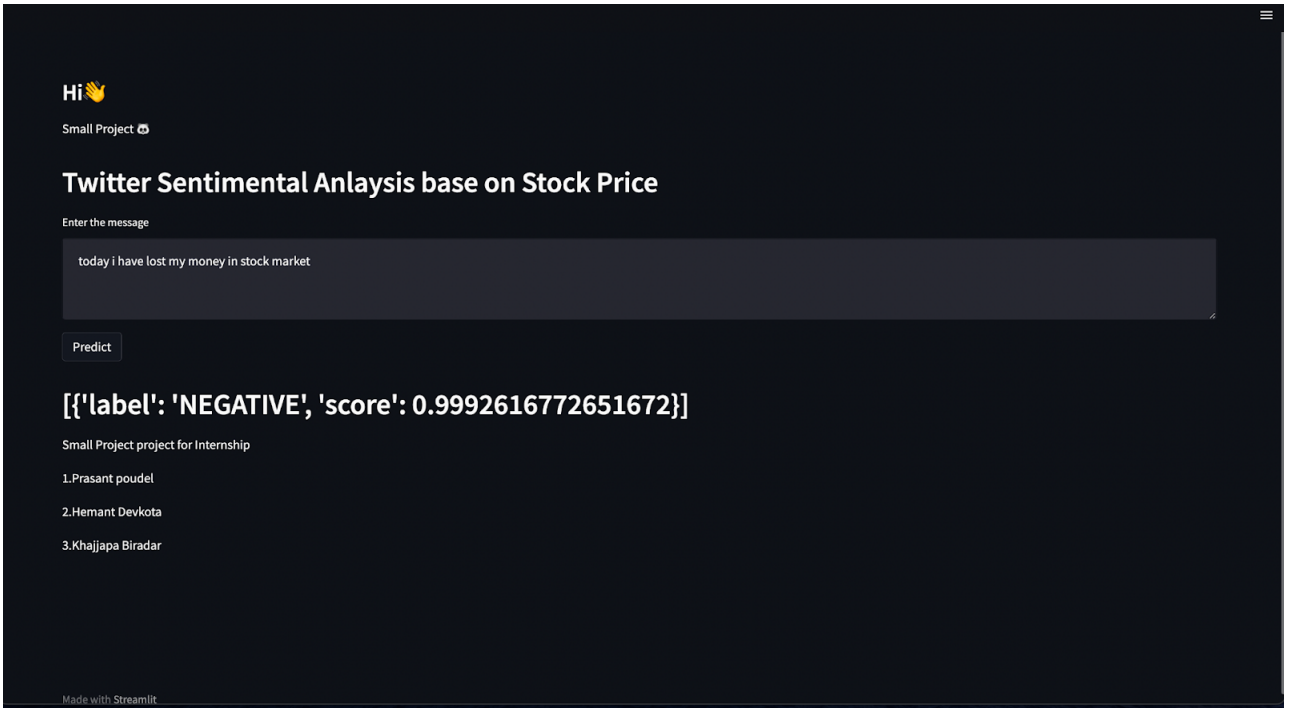
CHAPTER 7

SNAPSHOTS

7.SNAPSHOTS

Front End

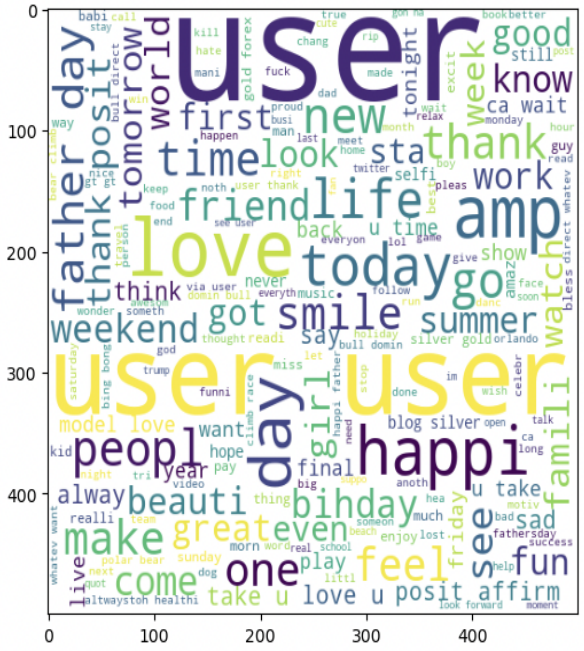




Code Snapshot

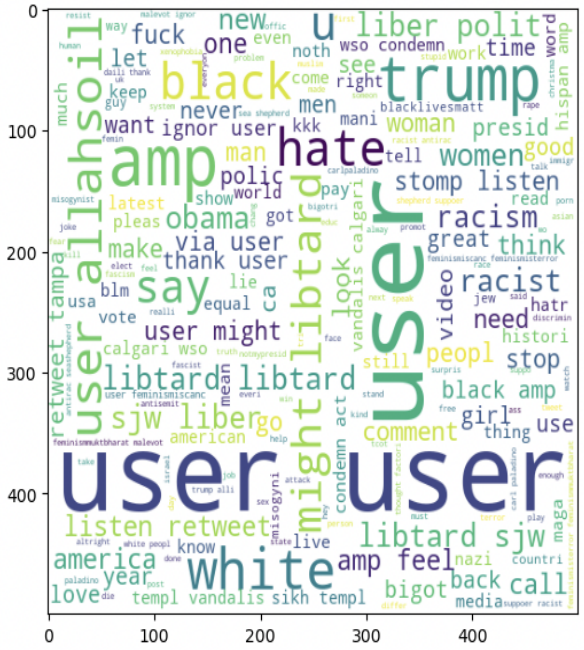
In [18]:
```
pos_wc=wc.generate(df[df['label']==0]['transform_tweet'].str.cat(sep=" "))
plt.figure(figsize=(18,6))
plt.imshow(pos_wc)
```

Out[18]: <matplotlib.figure.AxesImage at 0x15dd55e50>



In [19]:
```
neg_wc=wc.generate(df[df['label']==1]['transform_tweet'].str.cat(sep=" "))
plt.figure(figsize=(18,6))
plt.imshow(neg_wc)
```

Out[19]: <matplotlib.figure.AxesImage at 0x15f224250>



In [25]:
```
sgd = Pipeline([('vect', TfidfVectorizer()),
                ('tfidf', TfidfTransformer()),
                ('clf', SGDClassifier(loss='hinge', penalty='l2',alpha=1e-3, random_state=42, max_iter=5, tol=None)),
               ])
sgd.fit(x_train, y_train)
y_pred = sgd.predict(x_test)
print('accuracy %s' % accuracy_score(y_pred, y_test))
print(classification_report(y_test, y_pred,target_names=my_tags))
```

```
accuracy 0.9317968505579309
              precision    recall  f1-score   support

    negative       0.93      1.00      0.96      8910
    positive       1.00      0.04      0.07       679

    accuracy                           0.93      9589
   macro avg       0.97      0.52      0.52      9589
weighted avg       0.94      0.93      0.90      9589
```

In [26]:
```
randrf = Pipeline([('vect', TfidfVectorizer()),
                ('tfidf', TfidfTransformer()),
                ('clf', RandomForestClassifier()),
                   ])
randrf.fit(x_train, y_train)
y_pred =randrf.predict(x_test)

print('accuracy %s' % accuracy_score(y_pred, y_test))
print(classification_report(y_test, y_pred,target_names=my_tags))
```

```
accuracy 0.9583898216706643
              precision    recall  f1-score   support

    negative       0.96      1.00      0.98      8910
    positive       0.90      0.46      0.61       679

    accuracy                           0.96      9589
   macro avg       0.93      0.73      0.79      9589
weighted avg       0.96      0.96      0.95      9589
```

In [27]:
```
import pickle
pickle.dump(randrf,open('model.pkl','wb'))
```

In [28]:
```
rand = Pipeline([('vect', TfidfVectorizer()),
                ('tfidf', TfidfTransformer()),
                ('clf',DecisionTreeClassifier()),
               ])
rand.fit(x_train, y_train)
y_pred =rand.predict(x_test)

print('accuracy %s' % accuracy_score(y_pred, y_test))
print(classification_report(y_test, y_pred,target_names=my_tags))
```

```
accuracy 0.941599749713213
              precision    recall  f1-score   support

    negative       0.97      0.97      0.97      8910
    positive       0.60      0.54      0.57       679

    accuracy                           0.94      9589
   macro avg       0.78      0.76      0.77      9589
weighted avg       0.94      0.94      0.94      9589
```

In [29]:
```
rand = Pipeline([('vect', TfidfVectorizer()),
                ('tfidf', TfidfTransformer()),
                ('clf',KNeighborsClassifier()),
               ])
rand.fit(x_train, y_train)



y_pred =rand.predict(x_test)

print('accuracy %s' % accuracy_score(y_pred, y_test))
```

```
accuracy 0.9315882782354782
```

CHAPTER 8

CONCLUSION

CONCLUSION

The package was designed in such a way that future modifications can be done easily. The following conclusions can be deduced from the development of the project:

- Provides a better and higher perspective for analyzing the stocks through predictive modeling of the data of the sentiments.

- It can be implemented with techniques such as Wisdom of Crowd, which is basically incorporated with taking knowledge of tweets given by users to improve accuracy .

- It gives appropriate and better understanding for the stock related investments.

- It effectively analyzes whether the tweets are positive or negative.

- It provides flexibility with accuracy to get better sentiment prediction.

- The System has adequate scope for modification in future if it is necessary.

REFERENCE

1. Scikit-Learn https://scikit-learn.org/stable/
2. Tensorflow https://www.tensorflow.org/
3. Medium https://medium.com/
4. StackOverflow https://stackoverflow.com/
5. Kaggle https://www.kaggle.com/
6. Twitter Developer Portal https://developer.twitter.com/en/portal/dashboard