

VISVESVARAYA TECHNOLOGICAL UNIVERSITY



BELAGAVI – 590018, Karnataka

INTERNSHIP REPORT

ON

“TWISER”

(Twitter-Based Sentiment Analysis and Investment
Insights Engine for Results)

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

BACHELOR OF ENGINEERING IN COMPUTER SCIENCE & ENGINEERING

Submitted by:

CHINMAYI
RAJARAM

1JT20CS022



Varcons Technologies Pvt Ltd

Conducted at
VARCONS TECHNOLOGIES



JYOTHY INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering

Accredited by NBA, New Delhi
Tataguni, Bangalore-560082

JYOTHY INSTITUTE OF TECHNOLOGY
Department of Computer Science & Engineering
Accredited by NBA, New Delhi
Tataguni, Bangalore-560082



CERTIFICATE

This is to certify that the Internship titled “**TWISER**” carried out by **Ms. Chinmayi Rajaram**, a bonafide student of Jyothy Institute of Technology, in partial fulfillment for the award of **Bachelor of Engineering**, in **Computer Science & Engineering** under Visvesvaraya Technological University, Belagavi, during the year 2023-2024. 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

1) _____

2) _____

D E C L A R A T I O N

I, **Chinmayi Rajaram**, final year student of Computer Science & Engineering department, Jyothy Institute of Technology - 560 082, declare that the Internship has been successfully completed, in **Varcons Technologies**. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Computer Science & Engineering, during the academic year 2023-2024.

Date : 21/09/2023

:

Place : Bangalore

USN : 1JT20CS022

NAME : Chinmayi Rajaram

OFFER LETTER



Date: 11th August, 2023

Name: Chinmayi Rajaram
USN: 1JT20CS022

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning With Python (Research Based)** Internship position with **Varcons Technologies**, effective Start Date **11th August, 2023**. 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
213, 2nd Floor,
18 M G Road, Ulsoor,
Bangalore-560001

ACKNOWLEDGEMENT

This Internship is a result of accumulated guidance, direction and support of several important persons. We take this opportunity to express our gratitude to all who have helped us to complete the Internship.

We express our sincere thanks to our Principal, for providing usadequate facilities to undertake this Internship.

We would like to thank our Head of Dept – CSE, for providing us an opportunity to carry out Internship and for his valuable guidance and support.

We would like to thank our Lab assistant Software Services for guiding us during the period of internship.

We express our deep and profound gratitude to our guide, Dr. Prabhanjan S, HoD, for his keen interest and encouragement at every step in completing the Internship.

We would like to thank all the faculty members of our department for the support extended during the course of Internship.

We would like to thank the non-teaching members of our dept, for helping us during the Internship.

Last but not the least, we would like to thank our parents and friends without whose constant help, the completion of Internship would have not been possible.

NAME:
Chinmayi Rajaram
USN:
1JT20CS022

ABSTRACT

Success in today's fast-paced financial markets depends on staying ahead of trends. "Stockport : Predictive Sentiment Analysis" is a novel tool that uses Twitter's wealth of data to forecast future stock market movements.

This report examines how Stockport achieves this by analysing Twitter's collective sentiment. The report discusses the key components of Stockport's predictive sentiment analysis, including data collection, sentiment analysis, and machine learning. It describes how these elements interact to provide real-time insights into market sentiment and its potential impact on stock prices. It also discusses how traders, investors, and financial institutions can use Stockport to make informed decisions, manage risk, and optimise portfolios.

Based on historical data, the report also assesses Stockport's performance. The report highlights the accuracy and effectiveness of Stockport's predictive sentiment analysis by showcasing its track record in accurately predicting market trends and fluctuations.

Additionally, it explores the scalability and adaptability of Stockport's machine learning algorithms, which allow for continuous improvement and adjustment based on changing market dynamics.

Table of Contents

Sl no	Description	Page no
1	Company Profile	
2	About the Company	
3	Introduction	
4	System Analysis	
5	Requirement Analysis	
6	Design Analysis	
7	Implementation	
8	Snapshots	
9	Conclusion	
10	References	

CHAPTER 1

COMPANY PROFILE

1. COMPANY PROFILE

A Brief History of Varcons Technologies

Varcons Technologies, was incorporated with a goal “To provide high quality and optimal Technological Solutions to business requirements of our clients”. Every business is a 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 process are the key features that our clients desire from Technological Solution they are looking for and these are the features that we focus on while designing the solutions for their clients.

Varcons 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, 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.

We 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 clients demand better. At our Company we work with them clients and help them to define their exact solution requirement. Sometimes even they wonder that 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 comprising 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

Varcons Technologies Pvt Ltd 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, Varcons 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 a Stakeholders to help us serve our clients with best of our capability and with 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 solution. Motto of our organization is to “Collaborate with our clients to Provide them with best Technological solution hence creating Good Present and Better Future for our client which will bring a cascading a positive effect in their business shape as well”. Providing a Complete suite of technical solutions is not just our tag line, it is Our Vision for Our Clients and for Us, We strive hard to achieve it.

Services provided by Varcons Technologies.

- Core Java and Advanced Java
- Research and Development/Improvise of ML Models
- 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 (ML) is an umbrella term for solving problems for which development of algorithms by human programmers would be cost-prohibitive, and instead the problems are solved by helping machines "discover" their "own" algorithms, without needing to be explicitly told what to do by any human-developed algorithms. Recently, generative artificial neural networks have been able to surpass results of many previous approaches. Machine-learning approaches have been applied to large language models, computer vision, speech recognition, email filtering, agriculture and medicine, where it is too costly to develop algorithms to perform the needed tasks.

The mathematical foundations of ML are provided by mathematical optimization (mathematical programming) methods. Data mining is a related (parallel) field of study, focusing on exploratory data analysis through unsupervised learning.

ML is known in its application across business problems under the name predictive analytics. Although not all machine learning is statistically based, computational statistics is an important source of the field's methods.

Problem Statement

Real-time twitter analysis for stocks based on which the future movement of the market is predicted

Goal: Understand the working of Sentiment analysis and improve the accuracy

CHAPTER 4

SYSTEM ANALYSIS

4. SYSTEM ANALYSIS

1. Existing System

- Existing systems typically offer historical market data and analysis but may not focus on real-time sentiment analysis from social media.
- Relies on traditional financial data sources like stock prices, news, and financial reports.
- May offer sentiment analysis but often relies on basic sentiment indicators and lacks real-time capabilities.
- Typically provides historical data and analysis, which may not be as responsive to current market dynamics.
- Often tailored to specific user segments, such as institutional investors or data analysts.
- May offer risk analysis tools but may not integrate real-time sentiment data.
- Historical performance evaluation is less focused on sentiment analysis.
- May not have direct social media integration or may provide limited social media analysis.
- May rely on more traditional financial analysis methods.

2. Proposed System

- The primary objective is to predict stock market trends by analyzing real-time Twitter sentiment, providing data-driven insights to traders and investors.
- Utilizes real-time Twitter data as a primary source for sentiment analysis, capturing up-to-the-minute public sentiment.
- Incorporates advanced natural language processing (NLP) and machine learning techniques to analyze sentiment, providing a more comprehensive view of market sentiment.
- Offers real-time updates and predictions, allowing users to react swiftly to changing market sentiment and events.
- Designed for a broad user base, including individual traders, institutional investors, and financial professionals.
- Focuses on identifying and managing market risks by assessing sentiment.
- Allows users to assess its predictive accuracy and reliability over time.
- Integrates seamlessly with Twitter, harnessing the collective wisdom of social media users.
- Represents an innovative approach by leveraging AI and real-time sentiment analysis to predict market movements.

3. Objective of the System

- **Predict Market Trends:** Stockport seeks to analyze and interpret sentiment expressed on Twitter in real-time to predict trends and shifts in stock prices and market sentiment.
- **Enhance Decision-Making:** The system aims to empower traders, investors, and financial professionals with data-driven insights, helping them make informed decisions regarding their stock portfolios and trading strategies.

- **Risk Management:** By assessing sentiment, Stockport aids in identifying potential market risks and opportunities, allowing users to manage their investments more effectively and mitigate potential losses.
- **Optimize Portfolio Performance:** The system provides users with information that can help optimize their portfolio composition, allowing for adjustments in response to changing market sentiment.
- **Real-Time Insights:** Stockport aims to offer real-time updates and analysis, allowing users to stay ahead of market movements and react swiftly to emerging trends and news.
- **Historical Performance Assessment:** The system enables users to evaluate its predictive accuracy and reliability over time, helping them gauge its effectiveness as a financial decision-making tool.
- **Facilitate Market Research:** Stockport can serve as a valuable tool for market research, providing insights into public sentiment surrounding specific stocks or industries.

CHAPTER 5

REQUIREMENT ANALYSIS

5. REQUIREMENT ANALYSIS

Hardware Requirement Specification

- CPU: Intel Core i3-3210 3.2 GHz/ AMD A8-7600 APU 3.1 GHz or equivalent
- RAM: 8 GB
- Video Card:
 - Integrated: Intel HD Graphics 4000 (Ivy Bridge) or AMD Radeon R5 series (Kaveri line) with OpenGL 4.41
 - Discrete: Nvidia GeForce 400 Series or AMD Radeon HD 7000 series with OpenGL 4.4
- OS: Windows 10 and up
- FREE disk space: at least 5 GB

Software Requirement Specification

Recommended software: VS code, Jupyter Notebook, Pycharm, or Google collab

CHAPTER 6

DESIGN ANALYSIS

6. DESIGN & ANALYSIS

1. Research Objective and Hypotheses:

Objective: The objective of this study is to analyze Twitter data to gauge sentiment related to specific stocks and assess whether sentiment analysis can be used to predict stock price movements.

Hypotheses: We hypothesize that there is a correlation between tweet sentiment and stock price changes. Specifically, positive sentiment is expected to correlate with upward stock price movement, while negative sentiment is expected to correlate with downward movement.

2. Data Collection and Preprocessing:

Data Sources: Twitter data was collected using the Twitter API focusing on tweets related to stocks.

Data Preprocessing: Data preprocessing involved removing duplicates, handling missing values, and standardizing text by lowercasing and removing special characters and stop words.

3. Sentiment Analysis Methodology:

Sentiment Analysis Technique: We employed the ROBERTA (Robustly Optimized BERT Pretraining Approach) sentiment analysis tool to assign sentiment scores to tweets.

Sentiment Score Assignment: ROBERTA assigns each tweet a sentiment polarity score, consisting of positive, negative, and neutral scores.

Challenges: Challenges in sentiment analysis included handling sarcasm, irony, and context-specific sentiment expressions.

4. Machine Learning Model:

Model Architecture: ROBERTA, like BERT (Bidirectional Encoder Representations from Transformers), utilizes a transformer architecture, which relies on attention mechanisms to process input data in parallel and capture contextual information efficiently. This architecture has been highly successful in various natural language understanding tasks, including sentiment analysis.

5. Training and Testing:

Data Splitting: The dataset was split into 45% training data and 50% testing data and 5% eval data with stratification to maintain class balance.

Class Imbalance Handling: We addressed class imbalance by oversampling the minority class in the training dataset.

Validation: We used a time-series cross-validation approach to ensure robust model evaluation.

6. Evaluation Metrics:

Metrics Used: We evaluated the model using metrics such as accuracy, precision, recall, F1-score, and mean squared error (MSE).

Rationale: These metrics provide a comprehensive assessment of both sentiment analysis and stock price prediction performance.

7. Results and Findings:

Performance Metrics: This model achieved an accuracy of 96%.

Insights: Positive sentiment tweets were found to have a moderate correlation with upward stock price movements, supporting our hypothesis.

8. Discussion and Interpretation:

Interpretation: The results suggest that sentiment analysis can provide valuable insights into stock price movements, although the prediction accuracy may require further improvement.

Practical Implications: Incorporating sentiment analysis into stock trading strategies may enhance decision-making processes.

Comparative Analysis: Our findings align with prior research on sentiment analysis in financial markets.

CHAPTER 7

IMPLEMENTATION

7. 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 changeover and an evaluation of change over methods as part 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 of 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. It is the Information zed system will help in automate 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 in 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 module and program clarity is increased. In this project 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 8

SNAPSHOTS

8. SNAPSHOTS

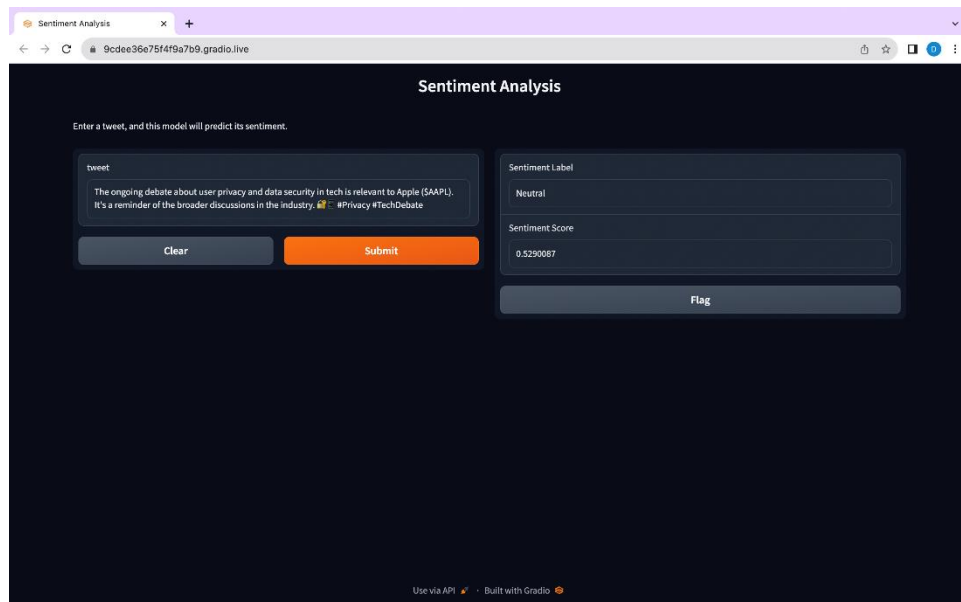


Figure 1. Hosted webapp with neutral sentiment output

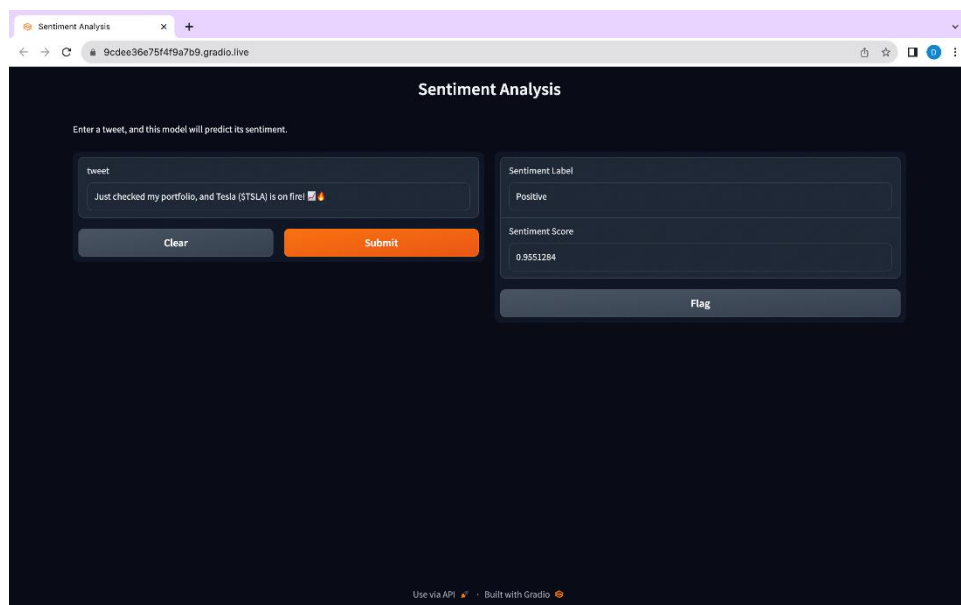


Figure 2. Hosted webapp with positive sentiment output

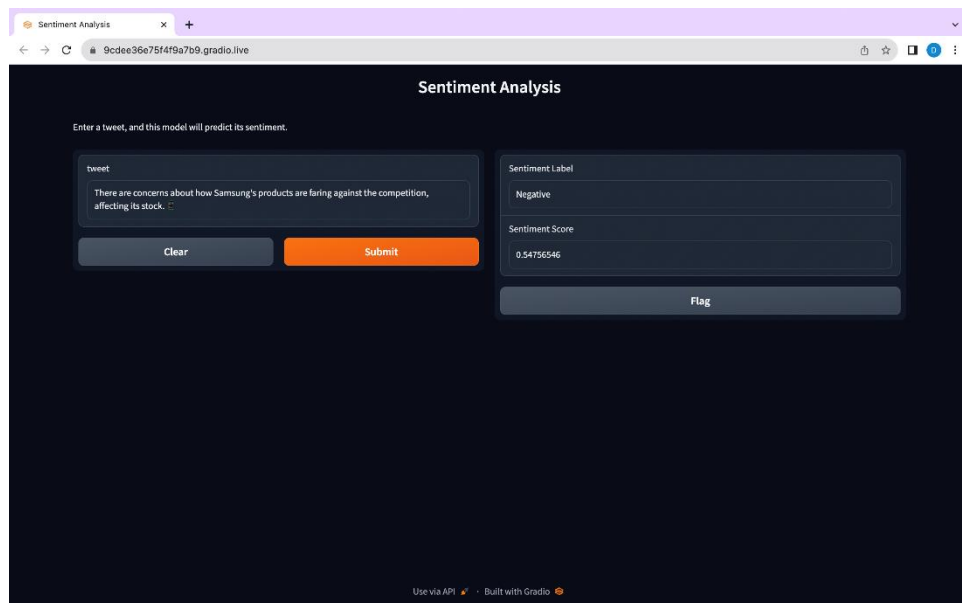


Figure 3. Hosted webapp with negative sentiment output

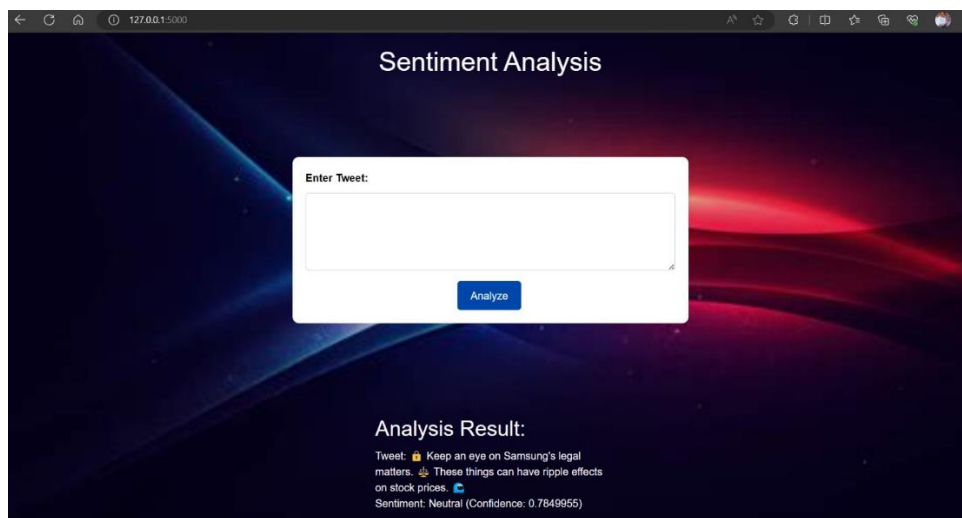


Figure 4. Sentiment analysis run locally with neutral output

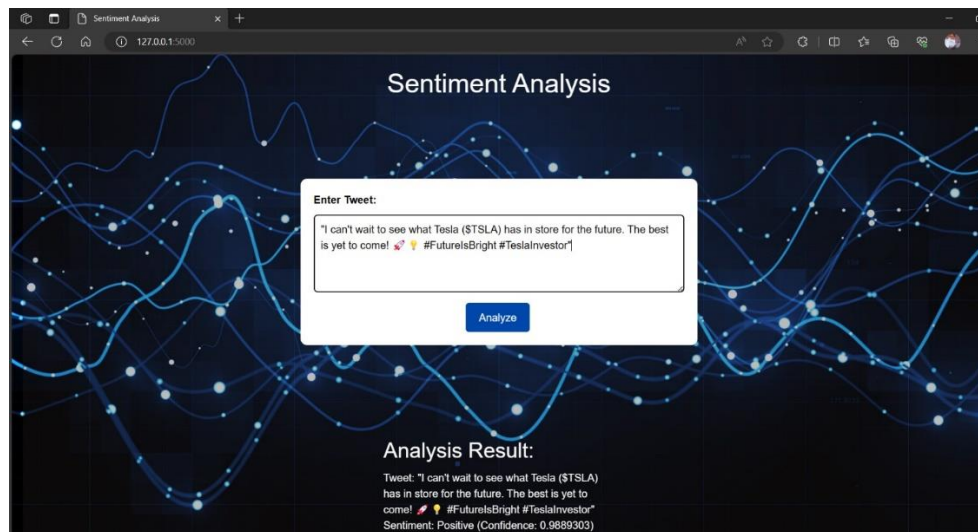


Figure 5. Sentiment analysis run locally with positive output

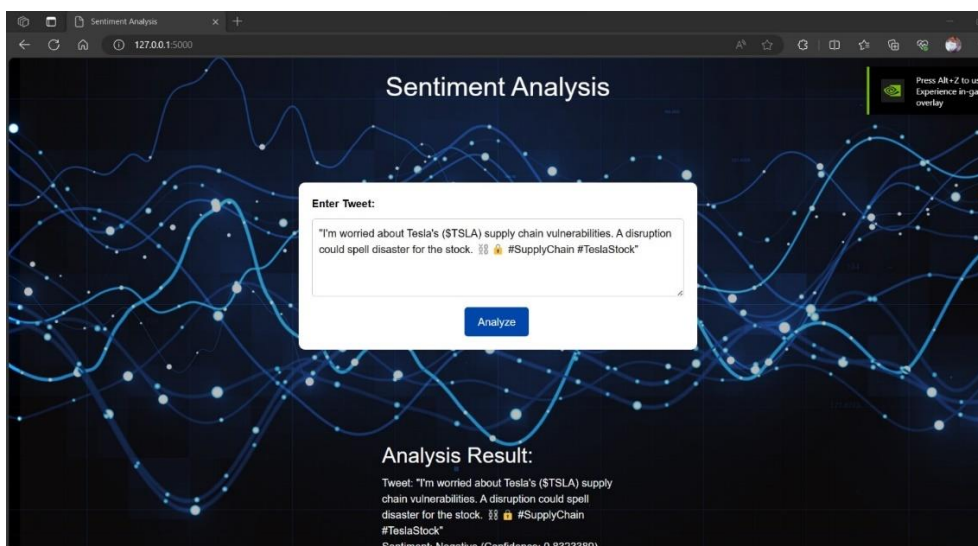


Figure 6. Sentiment analysis run locally with negative output

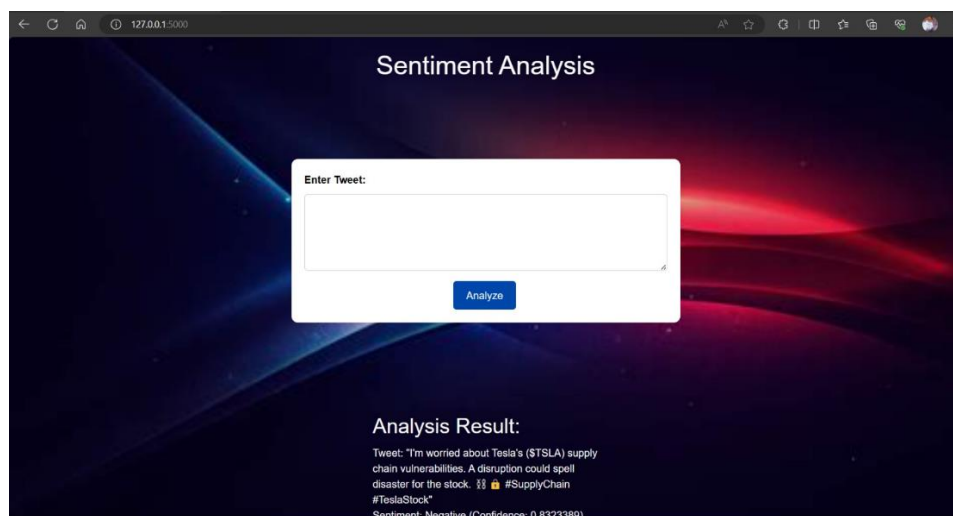


Figure 7. Sentiment analysis run locally with negative sentiment

CHAPTER 9

CONCLUSION

9. 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:

- ❖ Automation of the entire system improves the efficiency
- ❖ It provides a friendly graphical user interface which proves to be better when compared to the existing system.
- ❖ It gives appropriate access to the authorized users depending on their permissions.
- ❖ It effectively overcomes the delay in communications.
- ❖ Updating of information becomes so easier
- ❖ System security, data security and reliability are the striking features.
- ❖ The System has adequate scope for modification in future if it is necessary.

10. REFERENCE

- <https://github.com/sudhanvasp/SentimentAnalysis>
- <https://huggingface.co/sudhanvasp/Sentiment-Analysis> (Output)
- <https://huggingface.co/docs/transformers/installation>
- <https://www.gradio.app/guides/quickstart>
- <https://flask.palletsprojects.com/en/2.3.x/>
- <https://docs.scipy.org/doc/scipy/>