

Data Science Project 7PAM2002-0509-2023
Semester C 2023
Logbook (Activities and GitHub submissions)

Student Name and ID: Tamilselvan Palanisamy and 22055049

Project Title: Predicting Stock Market Values Using Machine Learning and News Sentiment Analysis

Supervisor: Mykola Gordovskyy

Student GitHub URL: https://github.com/Tamil2095/Msc_Finalproject

Number of versions of the code submitted on GitHub: 6

User documentation has been submitted on GitHub: YES

Student GitHub URL has been shared with markers: YES

Log of Activities
Must record attendance at lectures and supervisions

Week	Date	Activity incl. lectures & supervisions	Reason if not attend lecture or supervision	Weekly project progress. How lecture/supervision was helpful to your project.
1	14/05/2024	Lecture 1		<ul style="list-style-type: none">○ Overview of the module○ How to choose a project○ Supervisor Allocation
2	21/05/2024	Lecture 2		<ul style="list-style-type: none">○ Project and Data Management Plan○ Data Selection○ Module Scoring
3	31/05/2024	Supervision 1		<ul style="list-style-type: none">○ Discussed about the project title○ Discussed about the Source of Dataset
4	03/06/2024	Lecture 3		<ul style="list-style-type: none">○ Data Ethics lecture (part-1)○ Knowledge on ethics○ Basic idea on PDM plan
5	14/06/2024	Supervision 2		<ul style="list-style-type: none">○ Presentation of PDM plan○ Q&A, then feedback○ Discussed about, how it will work
6	17/06/2024	Lecture 4		<ul style="list-style-type: none">○ Data ethics lecture(part-2)○ UK GDPR Law○ UH Ethical Policy
7	28/06/2024	Supervision 3		<ul style="list-style-type: none">○ Discussed about Ethics Law○ Discussed about the Project Status

8	12/07/2024	Supervision 4		<ul style="list-style-type: none"> ○ Got suggestion to make project even better ○ Changed the Stock value from Apple to Shell and BP
9	15/07/2024	Lecture 5		<ul style="list-style-type: none"> ○ Results and Analysis lecture ○ Got knowledge on report preparation
10	18/07/2024	Supervision 5		<ul style="list-style-type: none"> ○ Progress of the code ○ How the sentimental analysis worked
11	09/08/2024	Supervision 6		<ul style="list-style-type: none"> ○ Discussed about the code and report ○ Got suggestion to make Learning curve
12	21/08/2024	Supervision 7		<ul style="list-style-type: none"> ○ Mock viva and feedback ○ Suggestion in report

Log of GitHub Submissions

Record the versions of code and user documentation submitted on GitHub

Date	Filename and version submitted to GitHub	Description of code and/or documentation submitted (what has been added since the previous version).
30/07/2024	Stock_price_code.ipynb - Version 1.0	<ul style="list-style-type: none"> • This file is the initial implementation for analysing stock prices, specifically focusing on Apple Inc. • It includes code for loading historical stock price data, performing pre-processing steps such as handling missing values and formatting, and visualizing the stock trends over time. • Key features include the use of libraries like pandas for data manipulation and matplotlib for creating visualizations such as line charts representing stock price movements.
31/07/2024	news_fetching_apple.ipynb - Version 1.0	<ul style="list-style-type: none"> • This notebook is dedicated to fetching the latest news articles related to Apple Inc. from various online sources. • It employs web scraping techniques and APIs to gather news data, which is then cleaned and organized for further analysis. • The collected news articles are stored in a structured format, making them ready for

		<p>subsequent sentiment analysis.</p> <ul style="list-style-type: none"> • The file also includes a brief exploration of the data to ensure accuracy and completeness.
31/07/2024	Textblob_sentiment_apple.ipynb - Version 1.0	<ul style="list-style-type: none"> • This file applies sentiment analysis to the Apple news articles collected in the previous step using the TextBlob library. • The notebook processes each article's text to determine its sentiment polarity (positive, negative, or neutral). • It includes functions for text pre-processing, such as removing stop words and normalizing text, followed by applying TextBlob's sentiment analysis methods. • The output is a dataset of articles labelled with their respective sentiment scores.
09/08/2024	vader_sentimental_score.ipynb - Version 1.0	<ul style="list-style-type: none"> • This notebook leverages the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool to analyse the sentiment of Apple-related news articles. • VADER is particularly suited for analysing social media and news data. • The file includes code for applying VADER to each news article, generating a compound sentiment score that reflects the overall sentiment (positive, negative, or neutral). • The results are used to assess the general public sentiment towards Apple as depicted in the news.
28/08/2024	Final_project_fullcode_far.ipynb - Version 1.0	<ul style="list-style-type: none"> • This comprehensive notebook integrates all aspects of the project into a single workflow. • It combines the stock price analysis, news fetching, and sentiment analysis to create a full pipeline that analyses how news sentiment affects Apple's stock

		<p>prices.</p> <ul style="list-style-type: none"> • The file includes data integration, feature engineering, and exploratory data analysis (EDA). • It also includes some preliminary results and visualizations that correlate sentiment scores with stock price movements.
28/082024	22055049.ipynb - Version 1.0	<ul style="list-style-type: none"> • This file appears to be a specialized analysis related to the broader project. • It may include additional case studies, a focused analysis on a specific subset of data, or advanced techniques not covered in the previous notebooks. • The exact nature of the content would depend on the context, but it likely provides deeper insights or enhances the project with new findings or methods. • This notebook might also include comparisons of different sentiment analysis techniques or additional data sources.