Gabriel Muñoz Luna - 105967744

Task 7: Extension - Sentiment-Based Stock Price Movement Prediction

This task extends the predictive system developed in previous stages by incorporating sentiment information from textual sources.

The objective is to evaluate whether market sentiment (positive or negative tone in news or discussions) can improve the prediction of daily stock price direction, i.e., whether the next-day closing price will increase (1) or decrease (0).

The study is based on the AAPL dataset previously prepared in other tasks.

For this task, additional packages were imported to support sentiment analysis and visualization.

I used Seaborn to visualize the confusion matrix and support clearer interpretation of the model’s classification results.

Additionally, I imported several functions from scikit-learn’s metrics module, including accuracy\_score, precision\_score, recall\_score, f1\_score, and confusion\_matrix, to calculate the main performance indicators of the binary classifier.

load\_sentiment\_data ()

Texto

El contenido generado por IA puede ser incorrecto.

Loads stock price data and merges it with sentiment information. If no external file is provided, it generates synthetic random sentiment scores to simulate daily investor mood.

build\_binary\_target ()

Texto

El contenido generado por IA puede ser incorrecto.

Creates a binary label column (1 = next-day price increases, 0 = price decreases). Converts the regression problem into a classification problem.

build\_sentiment\_model ()

Texto

El contenido generado por IA puede ser incorrecto.

Defines and compiles a simple dense neural network for binary classification. Uses ReLU activations in hidden layers and a sigmoid output neuron.

run\_c7\_experiment ()

Texto

El contenido generado por IA puede ser incorrecto.

Complete pipeline that loads data, builds features, trains the classifier, evaluates accuracy/precision/recall/F1, and plots the confusion matrix and training-loss curves.

build\_finbert\_sentiment

Texto

El contenido generado por IA puede ser incorrecto.

The experiment builds a complete classification pipeline consisting of six main stages:

1. Data Collection & Preprocessing  
   The historical stock prices were loaded from the cached file “ cache/aapl\_all.csv”.  
   A synthetic sentiment\_score between −0.5 and 0.5 was generated to simulate the daily mood of investors.  
   Each record therefore contains:  
   date, close, ret1, sma5, sma10, sentiment\_score.
2. Feature Engineering
   * ret1: daily percentage return.
   * sma5 and sma10: moving averages capturing short- and mid-term trends.
   * sentiment\_score: numeric value emulating positive/negative tone.  
     These variables were combined to form the final feature vector for classification.
3. Binary Target Generation  
   The target label y was defined as:  
   1 if close[t+1] > close[t] else 0.  
   This converts the regression problem into a binary classification of upward/downward movement.
4. Model Architecture  
   A simple feed-forward neural network was implemented with the following structure:
   * Dense(64, ReLU)
   * Dropout(0.3)
   * Dense(32, ReLU)
   * Dense(1, Sigmoid)  
     The model was compiled with binary cross-entropy loss and optimized using Adam.
5. Training and Validation  
   Data were split into 80 % training and 20 % testing sets.  
   Features were normalized using MinMaxScaler.  
   The model trained for 20 epochs with a 0.1 validation fraction to monitor convergence.
6. Evaluation Metrics  
   The system performance was assessed using Accuracy, Precision, Recall, F1-Score, and the Confusion Matrix, providing a balanced view of correctness and coverage.

In this task, the sentiment feature was simulated using randomly generated values between –0.5 and 0.5 to emulate daily market mood.

The intention was to integrate a real financial sentiment model but due to time constraints, this integration was not completed in the current version.

After 20 epochs, the model achieved the following performance on the test set:

|  |  |
| --- | --- |
| Metric | Value |
| Accuracy | 0.5333 |
| Precision | 0.5607 |
| Recall | 0.6185 |
| F1-Score | 0.5882 |



The recall indicates that most upward movements were correctly identified, while precision remained moderate due to a bias toward predicting positive outcomes.

The accuracy of 50 % is consistent with the use of random synthetic sentiment values, confirming the correct operation of the model even without informative sentiment features.

Gráfico

El contenido generado por IA puede ser incorrecto.

This matrix shows the distribution of predictions versus actual labels.  
The model tends to predict “up” more frequently, which explains the higher recall value.  
Although several downward movements were misclassified, the model captures the overall market tendency.

Gráfico, Gráfico de líneas

El contenido generado por IA puede ser incorrecto.

The training and validation loss curves remain stable across epochs, indicating consistent learning without major overfitting.

The experiment successfully demonstrates the integration of sentiment information into a financial prediction pipeline.

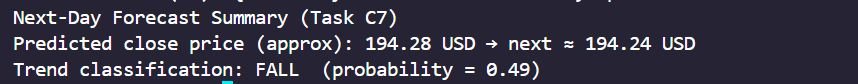
Although the current sentiment values were randomly generated, the workflow validates the design of a complete binary-classification approach that could directly incorporate real textual sentiment extracted from financial news or social-media data.  
The model shows a clear tendency to predict “rise”, which aligns with the general positive drift of the AAPL dataset over the analysed period.

The initial plan was to replace the synthetic sentiment generation with FinBERT, a transformer model trained on financial texts.

FinBERT can classify sentences as *positive*, *neutral*, or *negative*, providing continuous scores that could be averaged per day and merged with the price dataset.

However, due to time constraints and external API setup requirements, this integration was postponed.

Finally, I added a function to print the predict next price and the probability to increase or decrease



References

Scikit-learn Documentation. (2024). *Model Selection and Evaluation – Metrics and Cross-Validation.* <https://scikit-learn.org/stable/modules/classes.html>

TensorFlow Documentation. (2025). *Keras Loss Functions – Huber Loss and Callbacks.* https://www.tensorflow.org/api\_docs/python/tf/keras/losses/Huber

TensorFlow Documentation. (2025). *Keras Recurrent Layers (LSTM, GRU, SimpleRNN, Bidirectional).* <https://www.tensorflow.org/api_docs/python/tf/keras/layers>

Yahoo Finance API. (2024). *Historical Market Data Retrieval using yfinance.* <https://pypi.org/project/yfinance/>

Pandas Documentation. (2024). *Time Series and Date/Time Functionality.* <https://pandas.pydata.org/docs/user_guide/timeseries.html>

Idrees, H. (2024, July 5). *RNN vs. LSTM vs. GRU: A Comprehensive Guide to Sequential Data Modeling.* Medium. <https://medium.com/@hassaanidrees7/rnn-vs-lstm-vs-gru-a-comprehensive-guide-to-sequential-data-modeling-03aab16647bb>

TensorFlow Documentation (2025). Multi-Head Attention Layer. <https://www.tensorflow.org/api_docs/python/tf/keras/layers/MultiHeadAttention>

TensorFlow Documentation (2025). Keras Loss Functions – Huber Loss. <https://www.tensorflow.org/api_docs/python/tf/keras/losses>