Group Contribution Report

Stock Market Sentiment Analysis

Group Members:

- Member 1: Akhil Karri
- Member 2: Ramita Deshmukh
- Member 3: Karthikeya Rao Bolamalla
- Member 4: Sai Leela Kuragayala

Group Contribution Split:

Name of Member	(%) Contribution	Contribution Area
Akhil Karri	25%	Data Collection and Preparation
Ramita Deshmukh	25%	Sentiment Analysis & Feature Engineering
Karthikeya Rao Bolamalla	25%	Model Building and Evaluation
Sai Leela Kuragayala	25%	Model Selection, Feature Importance, Correlation Analysis & Live Prediction

Member	Contribution Area & Detailed Task Breakdown
Akhil Karri	**Data Collection and Preparation** • Collected Tesla and Apple news datasets by loading CSV files. • Inspected datasets and selected essential columns: date, title, and content. • Added a company label to each record for clear identification. • Standardized and formatted the date fields, removing any timezone information. • Handled missing or invalid date entries by cleaning the datasets. • Filtered Apple news articles to match Tesla's 2022 time period for consistency. • Merged the cleaned Tesla and Apple datasets into a single unified dataset for analysis.
Ramita Deshmukh	**Sentiment Analysis and Feature Engineering** • Applied a pre-trained financial sentiment model (FinBERT) to analyze news article titles. • Extracted sentiment labels (Positive, Neutral, Negative) along with corresponding confidence scores, while ensuring token size limitations were handled appropriately. • Integrated sentiment predictions into the combined news dataset by adding sentiment label and sentiment score for each article. • Engineered additional features, including: ⇒ Previous day's sentiment score, ⇒ Previous day's closing stock price, ⇒ 5-day and 10-day moving averages of closing prices. • Merged the enhanced sentiment and stock features into a unified dataset for model training.
Karthikeya Rao Bolamalla	**Model Training and Building** • Selected key input features for model training, including price movements, trading volume, and sentiment scores. • Performed chronological data splitting, using 80% for training and 20% for testing to ensure no future data leakage. • Built and trained the following five machine learning models:

	⇒ Logistic Regression ⇒ Random Forest Classifier ⇒ XGBoost Classifier ⇒ Support Vector Machine (SVM) ⇒ K-Nearest Neighbors (KNN)
Sai Leela Kuragayala	**Model Evaluation, Correlation Analysis and Live Prediction** • Evaluated machine learning models using classification metrics and selected Random Forest as the best-performing model with approximately 70% test accuracy. • Conducted correlation analysis using Pearson and Spearman methods to study feature relationships and created heatmaps for visualization. • Analyzed and plotted feature importance for different machine learning models to identify key predictive features. • Set up livestock price fetching for Tesla using the Alpha Vantage API. • Retrieved latest Tesla news headlines from NewsAPI and Reddit. • Applied FinBERT sentiment analysis on live news titles for real-time sentiment scoring. • Predicted Tesla stock movement in real time based on live sentiment and stock data. • Built a Streamlit dashboard to display livestock movement predictions interactively.