Internship Project Report

Organization: Zidio Development
Internship Domain: Data Analytics & Analysis
Project Title: Stock Price Prediction Using Traditional and Deep Learning Models
Interns: Kushagra Sharma, Transhik Tyagi
Duration: 3 Months

Introduction

During our internship at Zidio Development, we worked on the project “Stock Price Prediction Using Traditional and Deep Learning Models” under the Data Analytics & Analysis domain. The objective was to apply data preprocessing techniques, explore multiple algorithms, and evaluate their performance in forecasting stock prices. This project provided us with hands-on exposure to time-series forecasting, regression modeling, boosting algorithms, and deep learning architectures.

Project Overview

The primary goal was to predict future stock prices based on historical stock data and compare traditional ML models with deep learning approaches. Traditional machine learning models such as SVR, KNN, Decision Tree, ElasticNet, Bayesian Ridge, Linear Regression, Random Forest, XGBoost, LightGBM were implemented. Time-series forecasting models like ARIMA and Facebook Prophet were tested. A GRU (Gated Recurrent Unit) deep learning model was built to capture sequential dependencies.

Dataset Description

• Source: Preprocessed CSV file inside a ZIP folder.
• Target Feature: Stock_1 column.
• Preprocessing Steps:
- Normalization using MinMaxScaler.
- Creation of sequences using a sliding window (60 days).
- Splitting into 80% training and 20% testing.

Methodology

Models Implemented:
• Traditional ML Models: SVR, KNN, Decision Tree, ElasticNet, Bayesian Ridge, Linear Regression, Random Forest.
• Boosting Models: XGBoost, LightGBM.
• Time-Series Models: ARIMA, Facebook Prophet.
• Deep Learning Model: GRU.

Evaluation Metrics

• Mean Squared Error (MSE): Lower is better.
• R^2 Score: Higher (closer to 1) is better.

Visualization

Predictions vs. actual values were plotted for each model. These helped to assess accuracy trends, visualize lag in predictions, and compare ML vs. Deep Learning performance.

Results Summary

Model	MSE	R^2 Score
SVR	0.0123	0.87
KNN	0.0148	0.84
Decision Tree	0.0205	0.79
ElasticNet	0.0189	0.81
Bayesian Ridge	0.0167	0.83
Linear Regression	0.0212	0.78
Random Forest	0.0109	0.89
XGBoost	0.0091	0.91
LightGBM	0.0094	0.90
ARIMA	0.0228	0.76
Prophet	0.0201	0.80
GRU	0.0076	0.93

Key Findings

• Among traditional models, Random Forest gave the best performance with an R^2 of 0.89.
• GRU (Deep Learning) achieved the highest accuracy ($R^2 = 0.93$).
• XGBoost and LightGBM also performed excellently.
• Time-series models like ARIMA and Prophet provided decent baseline forecasts but underperformed compared to ML/DL approaches.

Tools & Technologies Used

• Languages: Python
• Libraries: pandas, numpy, matplotlib, scikit-learn, TensorFlow, xgboost, lightgbm, statsmodels, prophet

Conclusion

This project successfully showcased how different machine learning and deep learning models perform in stock price prediction. GRU emerged as the most powerful model, while Random Forest, XGBoost, and LightGBM also performed strongly. The project enhanced our skills in data preprocessing, model building, evaluation, and visualization.

Future Work

• Implement advanced models like LSTM, Bi-LSTM, and Transformer-based architectures.
• Add technical indicators (RSI, MACD, moving averages).
• Apply transfer learning for multi-stock forecasting.
• Deploy predictions in a web-based interactive dashboard.