Advanced Time Series Forecasting with Deep Learning and Explainability

Abstract:

This project implements advanced time series forecasting of stock prices using a multivariate LSTM model. Historical Open, High, Low, Close, and Volume features are used to predict future Close prices. SHAP-based explainability interprets the contribution of each feature across the temporal sequence, ensuring transparency and model understanding.

Introduction:

Time series forecasting is critical in finance for predicting future stock prices. Traditional models fail to capture nonlinear temporal dependencies. LSTM networks learn long-term dependencies effectively. Explainability is added using SHAP to interpret the contribution of each feature.

Methodology:

1. Download stock data (Reliance) using Yahoo Finance
2. Preprocess data and scale features
3. Create sequences for LSTM input (60-day window)
4. Train LSTM model with 64 units + Dense layer
5. Evaluate performance using RMSE and MAE
6. Explain predictions using SHAP with mean aggregation over timesteps

Model Architecture:

- Input: 60 timesteps × 5 features
- LSTM: 64 units
- Dense: 32 units, ReLU
- Output: 1 unit (Close price)
- Optimizer: Adam, Loss: MSE

Results:

- RMSE and MAE computed on test data
- Predicted vs Actual plot shows good trend capture
- SHAP summary plot highlights feature importance

Advantages

- Accurate Forecasting: LSTM captures nonlinear temporal dependencies, so it predicts stock prices more accurately than traditional statistical models like ARIMA.
- Multivariate Modeling: Using multiple features (Open, High, Low, Close, Volume) improves prediction reliability.
- Explainability: SHAP explains the influence of each feature over time, increasing transparency and trustworthiness.
- Adaptable Framework: The model can be extended to other stocks or financial time series easily.
- Handles Long-term Dependencies: LSTM remembers patterns over long sequences, useful for stock trends.

Disadvantages / Limitations

- Data Dependency: Requires large amounts of historical data to train effectively.
- Computationally Intensive: Training LSTM on long sequences can be slow, especially with large datasets.
- SHAP Complexity: SHAP explainability with LSTM is complex and can be unstable with large timesteps.
- No Guarantee in Stock Market: Stock prices are influenced by many unpredictable factors; model predictions are probabilistic, not certain.
- Limited Feature Scope: Only uses basic OHLCV features; adding technical indicators or news sentiment could improve performance.

Conclusion:

The LSTM model accurately forecasts stock prices using multivariate input. Explainability via SHAP improves transparency and supports understanding of temporal and feature-level influence.

Future Work:

- Multi-step prediction (future 30 days)
- Add technical indicators (RSI, Moving Average)
- Attention mechanism or Bidirectional LSTM