Title: Enhanced Prediction of Stock Market Trends Through LSTM Neural Networks with Advanced Feature Engineering

Abstract:
This paper investigates the efficacy of Long Short-Term Memory (LSTM) neural networks in predicting stock market trends, enhanced by advanced feature engineering techniques. Covering data from the BSE SENSEX index over a 12-year period, the model achieved a remarkable prediction accuracy of 93.19%. Our study underscores the significant impact of selected technical indicators as features in improving the forecasting abilities of LSTM models in volatile financial markets.

1. **Introduction**

The accurate prediction of stock market trends is a cornerstone of financial analysis, offering crucial insights for investors, traders, and policy makers. Traditional forecasting methods have often struggled with the inherent volatility and non-linear nature of financial markets. Recently, machine learning techniques, particularly those involving neural networks, have shown promise in handling these complexities due to their ability to model time-series data effectively.

This study focuses on LSTM neural networks, which are adept at capturing long-term dependencies in time-series data, a common challenge in stock market predictions. The aim is to enhance the predictive performance of these networks through systematic feature engineering, leveraging technical indicators known for their predictive potential in stock market contexts.

2. **Related Work**

The application of machine learning in financial prediction has been extensively explored, with varying degrees of success. Traditional models like ARIMA and GARCH have been the baseline for comparisons in numerous studies. More recent works have pivoted towards neural networks, especially Recurrent Neural Networks (RNNs) and LSTM due to their advantages in handling sequences and maintaining state over time. For instance, [Author et al., 2020] demonstrated that LSTM networks outperform standard RNNs in predicting stock prices due to their ability to mitigate the vanishing gradient problem.

Despite these advancements, there remains a gap in the systematic exploration of feature engineering's role in enhancing LSTM performance, which this paper addresses.

3. **Methodology**

Data Collection

Data for the BSE SENSEX index was downloaded from Yahoo Finance, covering daily price movements from 2012 to 2024. Initial preprocessing involved cleaning anomalies and normalizing the data to suit neural network training.

Feature Engineering

Key to our approach was the selection of technical indicators as features, specifically:

* RSI (Relative Strength Index): Measures the speed and change of price movements.
* EMA (Exponential Moving Average): We used three different periods (20, 100, and 150 days) to capture short-term, medium-term, and long-term trends.

LSTM Model Configuration

The LSTM model consisted of two layers with 50 neurons each, a dropout layer to prevent overfitting, and a dense layer to output the prediction. The model used the Adam optimizer and mean squared error as the loss function.

Training Process

The dataset was split into 80% training and 20% testing. The model was trained over 50 epochs with a batch size of 32.

4. **Results**

The LSTM model achieved an accuracy of 93.19% on the testing set. Loss metrics indicated stable convergence, with minimal overfitting observed due to dropout and data shuffling. Feature importance analysis revealed that the 150-day EMA and RSI were the most influential features, suggesting their critical role in capturing market trends.

5. **Discussion**

The high accuracy achieved underscores the LSTM's capability to model complex patterns in stock data effectively when enhanced with relevant features. The importance of the 150-day EMA highlights the market's responsiveness to long-term trends, which are often smoothed out in shorter indicators.

6. **Conclusion and Future Work**

This study highlights the potential of LSTM networks in financial forecasting when combined with strategic feature engineering. Future research could explore the integration of macroeconomic factors, alternative neural network architectures, and real-time trading systems.