Hybrid CNN-LSTM model for stock price prediction

Deepak Singh, Hareesh P Nair, Mousina Barman, Snehal, Sonu Kumari

Indian Institute of Technology Madras
Chennai 600036, India

Motivation

Non-Linearity in the stock market data

The stock market is known to have an inherent non-linearity due to the complex interplay of various factors like investor sentiment, market psychology, geopolitical events, etc., which interact dynamically, leading to unpredictable price movements. Traditional linear models and statistical models fail to capture this non-linearity and other intricate relationships between the data, hence resulting in an unreliable price forecast.

The advancement in the field of computation and machine learning has opened up a wide range of applications for them in different industries, including finance. The inherent complexity of the stock market dynamics calls for the need for non-linear models like neural networks to predict the price changes in the stock market data accurately.

Objective: To develop a hybrid convolutional neural network and long short-term memory network model to predict the price of equity with high accuracy.

Mathematical Overview

Convolutional Neural Network

Convolutional neural networks are powerful algorithms that are used in variety of applications like image classification, text/speech recognition and a lot more. The algorithm can capture the short-term dependencies. It consists of three layers: Convolutional layer, Pooling layer and fully connected layer, which is the hidden layer of a feedforward neural network.

- Convolutional layer :

$$Y^{l+1} = [Y^l \circledast w^{l+1}](c,d) + b = \sum_{K_l}^{f=1}\sum_{f}^{e=1}\sum_{f}^{y=1}[Y_l^k(s_0c+x, s_0d+y)w_{l+1}^k(e,f)] + b$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & -3 & 0 & 1 \\ 2 & 1 & 1 & -1 & 0 \\ 0 & -1 & 1 & 2 & 1 \\ 1 & 2 & 1 & 1 & 1 \end{bmatrix} \otimes \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -2 & -1 \\ 2 & 2 & 4 \\ -1 & 0 & 0 \end{bmatrix}$$

- Pooling layer :

$$A_k^l(c,d) = \left[\sum_{x=1}^{f}\sum_{y=1}^{f} A_k^l(s_0c+e, s_0d+f)^p\right]^{\frac{1}{p}}$$

$$\begin{bmatrix} 2 & 2 & 7 & 3 \\ 9 & 4 & 6 & 1 \\ 8 & 5 & 2 & 4 \\ 3 & 1 & 2 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 9 & 7 \\ 8 & 6 \end{bmatrix}$$

Max pooling : Filter = (2x2), Stride = (2,2)

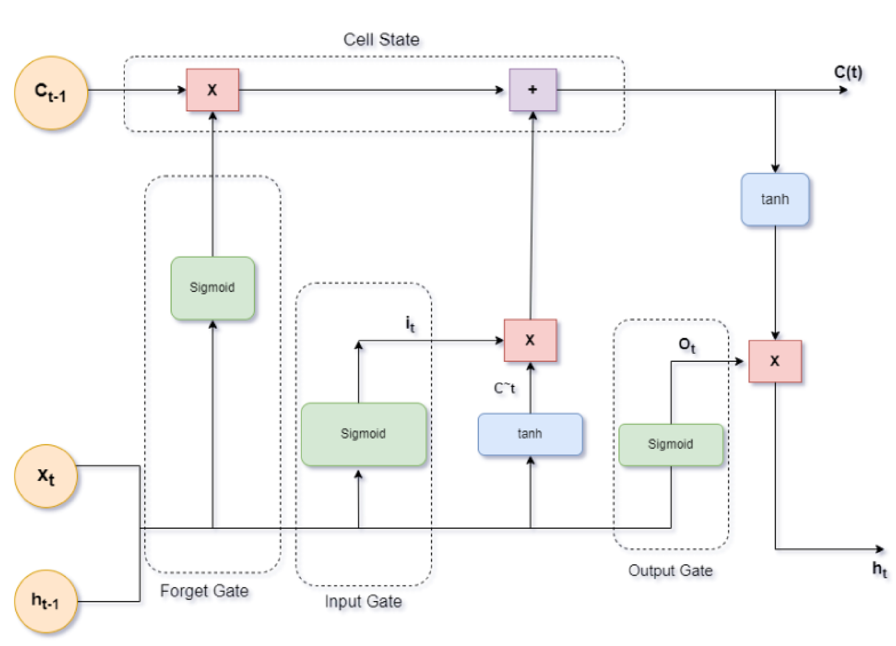
$$\begin{bmatrix} 2 & 2 & 7 & 3 \\ 9 & 4 & 6 & 1 \\ 8 & 5 & 2 & 4 \\ 3 & 1 & 2 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 4.25 & 4.25 \\ 4.25 & 3.5 \end{bmatrix}$$

Max pooling : Filter = (2x2), Stride = (2,2)

Long-Short Term Memory Network

Long-Short Term Memory networks are a special type of Recurrent Neural Network that can model sequential data, capture long-term dependencies, and make accurate predictions. It consists of multiple layers of cells which consist the following gates:

- Forget gate : $f_t = \sigma_g(W_f x_t + U_f c_{t-1} + b_f) \in (0,1)$
- Input gate : $i_t = \sigma_g(W_i x_t + U_i c_{t-1} + b_i) \in (0,1)$
- Output gate :$o_t = \sigma_g(W_o x_t + U_o c_{t-1} + b_o) \in (0,1)$

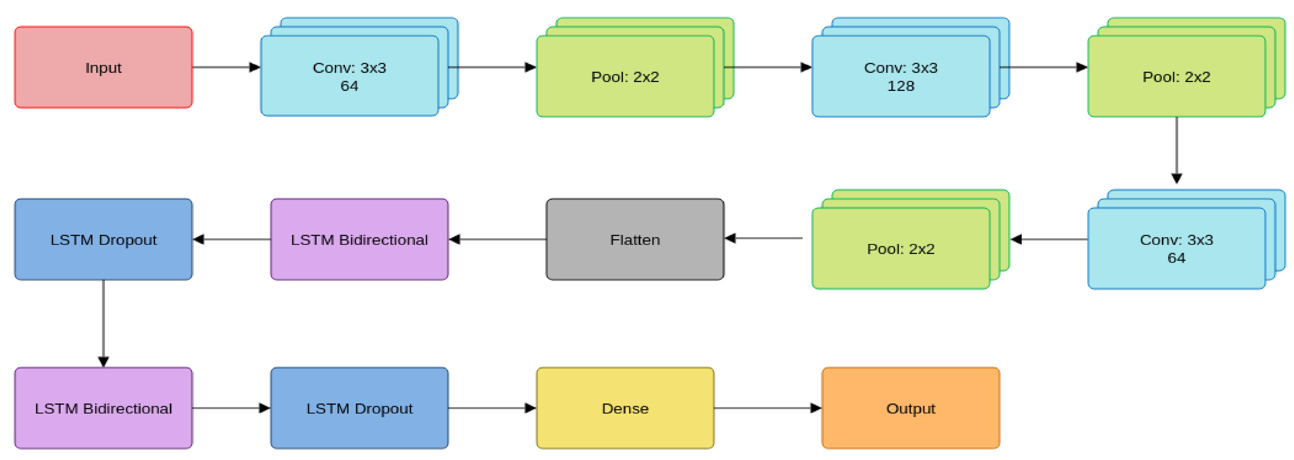


- $h_t = o_t \odot \sigma_h(c_t)$
- $c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$
- $\tilde{c}_t = \sigma_h(W_c x_c + U_c h_{t-1} + b_c)$

where σ_g is the sigmoid function.

Methodology

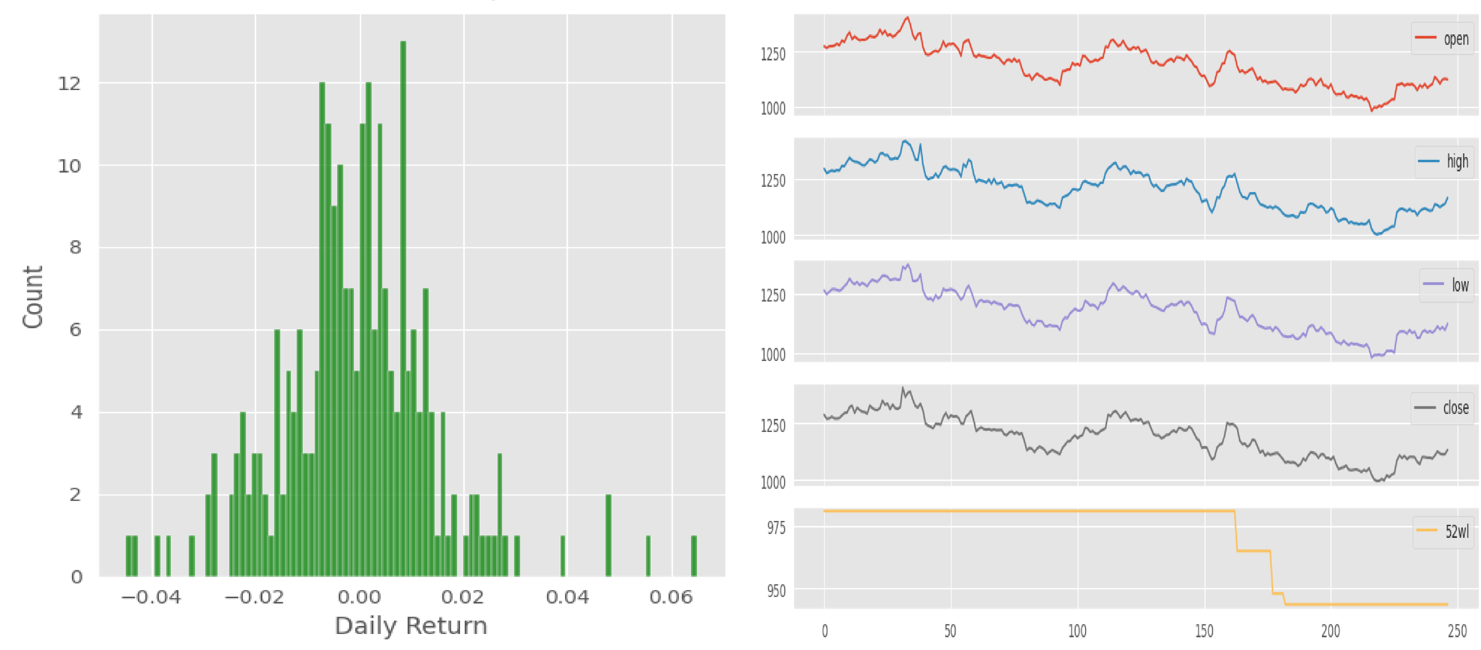
The CNN layers extract relevant features from input sequences, while the LSTM layers capture temporal dependencies and patterns, enabling accurate predictions of stock prices across various window lengths. The model's architecture facilitates effective information processing and learning, leading to high-performance metrics such as variance, R2 score, and maximum error evaluation.



The CNN-LSTM model is trained to predict the stocks for different window lengths and the accuracy of the model is compared with two different types of models.
Benchmark models : Random Forest, Support Vector Machine

Data

The market data used for training the model is obtained from NSE India. Various datasets of same format were used to benchmark the models. Certain data analysis techniques were utilized to obtain moving averages, calculate daily returns and visualize the distributions of closing price.



Results

Different stock price data were used to train and test the model. The below obtained are the accuracies of the CNN- LSTM model and the benchmark ML models

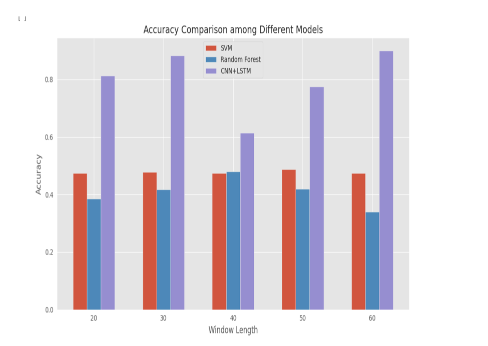


Fig. 2: Tech Mahindra

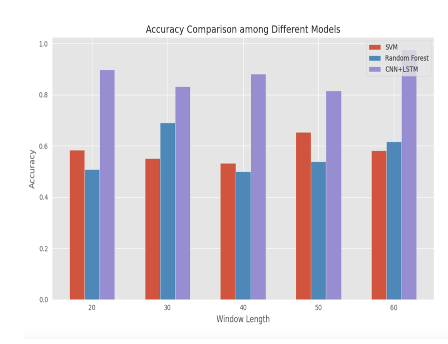


Fig. 3: SBI

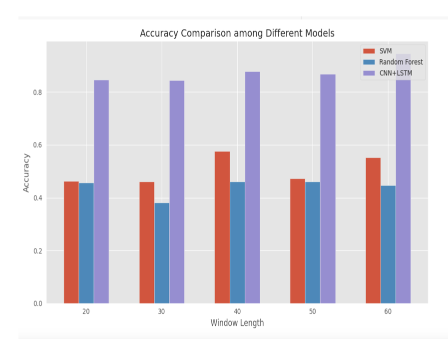


Fig. 4: TATA Motors

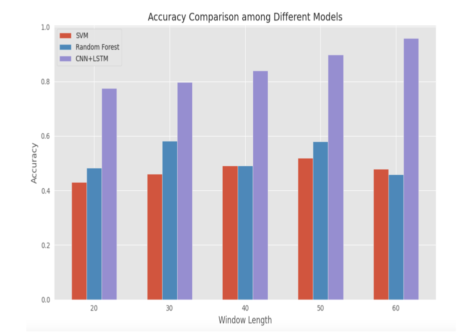


Fig. 5: Idea

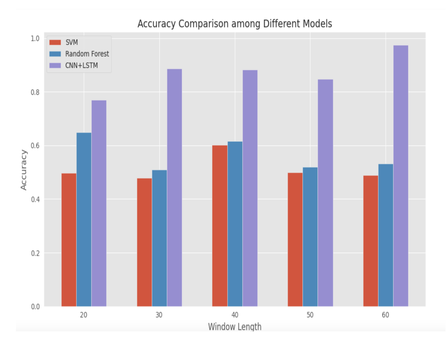


Fig. 6: Adani Port

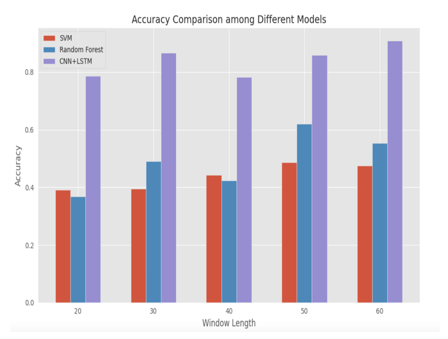


Fig. 7: TCS

Conclusions and future plan

Conclusions

- The model architecture for the CNN-LSTM is much more complex than benchmark models, hence resulting in a higher training time.
- The CNN-LSTM models have provided highly accurate predictions consistently across different stock market data and different prediction window lengths compared to the benchmark models
- Since the CNN-LSTM model can handle non-linear relationship and is robust to temporal patterns, it can be scaled and adapted to different market conditions, making it a versatile tool in the trading industry.

Future plans

- Develop adaptive learning algorithms to dynamically adjust the model parameters and architectures based on the market conditions.
- The model can be modified to make real-time probabilistic predictions by incorporating bayesian inference methods.