International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Stock Price Prediction Using Machine Learning

ANUJ KUMAR GUPTA, GAURAV VERMA, KESHAV SHARMA

Under the Supervision of Prof. Badal Bhushan (Asstt. Professor),
Department of Computer Science and Engineering

Chapter : 1

Abstract:

Researchers have been studying different methods to effectively predict the stock market price. Useful prediction systems allow traders to get better insights about data such as: future trends. Also, investors have a major benefit since the analysis give future conditions of the market. One such method is to use machine learning algorithms for forecasting.
A number of researchers have come up with various ways to solve this problem, mainly there are traditional methods so far, such as artificial neural network is a way to get hidden patterns and classify the data which is used in predicting stock market. This project proposes a different method for prognosing stock market prices.

It does not fit the data to a specific model; rather we are identifying the latent dynamics existing in the data using machine learning architectures. In this work we use Machine learning architectures Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN) and Hybrid approach of LSTM + CNN for the price forecasting of NSE listed companies and differentiating their performance. On a long term basis, sling window approach has been applied and the performance was assessed by using root mean square error.

Chapter : 2

Motivation

The future price of a stock is the main motivation behind the stock price prediction. In various cases like business and industry, environmental science, finance and economics motivation can be useful. The future value of the company's stock can be determining

Chapter : 3
Literature Survey related to Topic of Project

SL No.	Paper Title	Authors	Year	Reference Link	Name of Publisher	Technology	Method
p1	Stock Closing Price Prediction using Machine Learning Techniques	Mehar Vijha , Deeksha Chandolab, Vinay Anand Tikkiwalb, Arun Kumarc	2019	https://www.sciencedirect.com/science/article/pii/S1877050920307924	Elsevier	Time series algorithms namely ARIMA, CNN, SVM and ANN.	Forecasting Using Data From Yahoo Finance and Analyzing Seasonal and Nonseasonal Trend
2	Stock Market Prediction Using Machine Learning:	Mehak Usman, Syed Hasan Adi	2018	https://sci-hub.se/10.1109/iccoins.2016.7783235	IEEE	Prediction Using Machine Learning	The researchers have used Artificial Neural Networks (ANN) and statistical technique ARIMA on almost 3 year's data to predict KSE-100 index.
3	STOCK PRICE PREDICTION USING REINFORCEMENT LEARNING	Jae Won idee	2001	https://scihub.se/10.1109/isie.2001.931880	IEEE	REINFORCEMENT LEARNING	The proposed approach provides a basis for applying reinforcement leaming to stock market prediction.

Sr.No.	Paper Title	Authors	Year	Name of Publisher	Reference Link	Technology	Method
4.	A Survey of Forex and Stock Price Prediction Using Deep Learning	Zexin Hu , Yiqi Zhao and Matloob Khushi	2021	MDPI	https://scihub.se/10.3390/asi4010009	Deep Learning	This paper provided a detailed review of 86 papers from 2015 to the present on predicting stock/Forex price movements through deep learning methods.
5	Stock price prediction using machine learning on least-squares linear regression basis	C. C. Emioma , S. O. Edeki 1	2020	IOP Publishing	https://sci-hub.se/10.1088/17426596/1734/1/012058	Machine Learning Techniques on least-squares linear regression basis	The prominent aim of this study is to [6] forecast the market performance of the Karachi Stock Exchange (KSE) on day closing using machine learning algorithms.

SL No.	Paper Title	Authors	Year	Name of Publisher	Reference Link	Technology
6.	Stock Price Prediction Using News Sentiment Analysis	Saloni Mohan, Sahitya Mullapudi, Sudheer Sammeta, Parag Vijayvergia and David C. Anastasiu	2019	IEEE	https://sci-hub.se/10.1109/bigdataservice.2019.00035	Deep Learning
7.	Stock Price Prediction: A Comparative Study between Traditional Statistical Approach and Machine Learning Approach	Indronil Bhattacharjee and Pryonti Bhattacharja	2019	IEEE	https://sci-hub.se/10.1109/eict48899.2019.9068850	Traditional Statistical Approach and Machine Learning Approach

SL No.	Paper Title	Authors	Year	Reference Link	Name of Publisher	Technology
8.	Stock Price Prediction Using Support Vector Regression on Daily and Up to the Minute Prices	Bruno Miranda Henrique, Vinicius Amorim Sobreiro, Herbert Kimura	2018	https://scihub.se/https://www.sciencedirect .com/science/article/pii/S2405918818300060	Elsevier	Support Vector Regression
9.	Stock market prediction using machine learning classifiers and social media, news	Wasiat Khan Mustansar Ali Ghazanfar Muhammad Awais Azam Amin Karami Khaled H. Alyoubi Ahmed S. Alfakeeh	2020	https://sci-hub.se/10.1007/ s12652-02001839-w	Springer	Prediction Using Machine Learning
10.	Mid-price Prediction Based on Machine Learning Methods with Technical and Quantitative Indicators	Adamantios Ntakarisa , Juho Kanniainenb , Moncef Gabbouja , Alexandros Iosifidisc	2018	https://scihub.se/10.2139/ssrn.3213389	SSRN	Machine Learning Methods with Technical and Quantitative Indicators

Chapter : 4

Literature review

Here, I have reviewed various approaches for

Stock price prediction. All approaches have their own advantages and disadvantages. CNN & LSTM is a most popular algorithm to prediction the stock price but there are some challenges in this method like use to need a lot of training data, High computational cost, without GPU data quite slow to train, depend on any previous information for prediction.

Advantages	Disadvantages
Easy Pattern Recognition	Needs lot of training data.
Adaptability	without GPU data quite slow to train
Hierarchical feature learning	High computational cost
Sequences modelling	depend on any previous information for prediction
Ability to learn patterns	More complex and expensive.
Feature Extraction	Difficulty in capturing sudden changes.

Chapter : 5

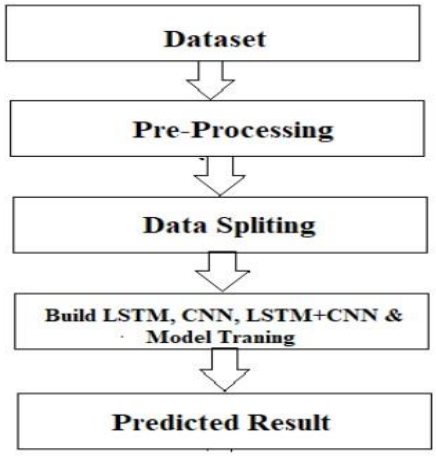
Problem formulation/Objectives

A stock market prediction is described as an action of attempting to classify the future value of the company stock or other financial investment traded on the stock exchange. The forthcoming price of a stock of the successful estimation is called the Yield significant profit.

This helps you to invest wisely for making good profits.

Chapter: 6

Methodology/ Planning of work



Proposed Workflow

Methodology/ Planning of work

The system presented here composes of five modules:-

1. Input as Dataset
2. Pre processing
3. Data splitting
4. Build & Model train Lstm, CNN and Hybrid approach of LSTM+CNN
5. Output as Predicted Result

Attribute such as: price of open, high, low, close, adjusted close price taken from huge dataset are fed as input to the models for training to pre-process the data techniques like normalization & one hot encoding in applied on dataset. After this data is divided in two sets namely training & testing which are ratio of 80:20 respectively. Then, this set are used to train a model using 3 different approaches: LSTM, CNN and Hybrid approach of LSTM+CNNS.
Finally, all these modules are evaluated using Root mean square error

Chapter: 7

Facilities required for proposed work

(Software/Hardware required for the development of the project.)

PYTHON

The language of select for this project was Python. This was a straightforward call for many reasons.

1. Python as a language has a vast community behind it. Any problems which may be faced is simply resolved with visit to Stack Overflow. Python is the foremost standard language on the positioning that makes it is very straight answer to any question.

2. Python is an abundance of powerful tools ready for scientific computing Packages. The packages like NumPy, Pandas and SciPy area unit freely available and well documented. These Packages will intensely scale back, and variation the code necessary to write a given program. This makes repetition fast.

3. Python is a language as forgiving and permits for the program that appear as if pseudo code. This can be helpful once pseudo code give in tutorial papers should be required and verified. Using python this step is sometimes fairly trivial. However, Python is not without its errors. The python is dynamically written language and packages are area unit infamous for Duck writing. This may be frustrating once a package technique returns one thing that, for instance, looks like an array instead of being an actual array. Plus the standard Python documentation did not clearly state the return type of a method, this can't lead without a lot of trials and error testing otherwise happen in a powerfully written language. This is a problem that produces learning to use a replacement Python package or library more difficult than it otherwise may be.

NUMPY

Numpy is python package which provide scientific and higher level mathematical abstractions wrapped in python. It is [20] the core library for scientific computing, that contains a provide tools for integrating C, strong n-dimensional array object, C++ etc.

It is also useful in random number capability, linear algebra etc. Numpy's array type augments the Python language with an efficient data structure used for numerical work, e.g., manipulating matrices. Numpy additionally provides basic numerical routines, like tools for locating Eigenvectors .

SCIKIT LEARN

Scikit-learn could be a free machine learning library for Python. It features numerous classification, clustering and regression algorithms like random forests, k-neighbours, support vector machine, and it furthermore supports Python scientific and numerical libraries like SciPy and NumPy.

In Python Scikit-learn is specifically written, with the core algorithms written in Cython to get the performance. Support vector machines are enforced by a Cython wrapper around LIBSVM .i.e., linear support vector machines and logistic regression by a similar wrapper around LIBLINEAR

TENSORFLOW

In the TensorFlow [22]has an open source software library for numerical computation using data flow graphs. Inside the graph nodes represent mathematical formulae, the edges of graph represent the multidimensional knowledge arrays (tensors) communicated between them. The versatile architecture permits to deploy the computation to at least one or many GPUs or CPUs in a desktop, mobile device, servers with a single API. TensorFlow was firstly developing by engineers and researchers acting on the Google Brain Team at intervals Google's Machine Intelligence analysis organization for the needs of conducting deep neural networks research and machine learning, but, the system is generally enough to be appropriate in a wide range of alternate domains as well.
Google Brain's second-generation system is TensorFlow. Whereas the reference implementation runs on single devices, TensorFlow can run on multiple GPUs and CPUs. TensorFlow is offered on Windows, macOS, 64-bit Linux and mobile computing platforms together with iOS and Android.

KERAS

Keras Is A High-level Neural Networks API, It Is Written In Python And Also Capable Of Running On Top Of The Theano, CNTK, Or.
Tensorflow. It Was Developed With Attention On Enabling Quick Experimentation. Having The Ability To Travel From Plan To Result With
The Smallest Amount Doable Delay Is Key To Doing Great Research.Keras Permits For Straightforward And Quick Prototyping (Through User-friendliness, Modularity, And Extensibility). Supports Each Recurrent Networks And Convolutional Networks, Also As Combinations Of The 2.
Runs Seamlessly On GPU And CPU. The Library Contains Numerous Implementations Of Generally Used Neural Network Building Blocks
Like Optimizers, Activation Functions, Layers, Objectives And A Number Of Tools To Create Operating With Text And Image Data Easier. The Code Is Hosted On Github, And Community Support Forums Embody The Github Issues Page, A Gitter Channel And A Slack Channel

COMPILER OPTION

Anaconda Is [19] Free Premium Open-source Distribution Of The R And Python Programming Languages For Scientific Computing, Predictive Analytics, And Large-scale Process That Aim Is To Modify Package Managing And Deployment. Package Versions Unit Managed By The Package Management System Conda.

JUPITER NOTEBOOK

The Jupyter Notebook Is An Open-source Web Application That Enables To Making And Sharing Documents That Contain Visualizations, Narrative Text, Live Code And Equations. Uses Include: Data , Data Visualization, Data Transformation, Statistical Modelling, Machine Learning, Numerical Simulation, Data Cleaning And Much More .

Chapter : 8

References

1. Masoud, Najeb MH. (2017) "The impact of stock market performance upon economic growth." International Journal of Economics and Financial Issues 3 (4) : 788–798.

2. J. Jagwani, M. Gupta, H. Sachdeva, and A. Singhal, "Stock Price Forecasting Using Data from Yahoo Finance and Analysing Seasonal and Nonseasonal Trend," in 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, Jun. 2018, pp. 462–467, doi: 10.1109/ICCONS.2018.8663035.

3. I. Parmar et al., "Stock Dec. 2018, pp. 574–576, doi: 10.1109/ICSCCC.2018.8703332.

4. Y. Lei, K. Zhou, and Y. Liu, "Multi-Category Events Driven Stock Price Trends Prediction," in 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), Nanjing, China, Nov. 2018, pp. 497–501, doi: 10.1109/CCIS.2018.8691392.

5. B. Jeevan, E. Naresh, B. P. V. kumar, and P. Kambli, "Share Price Prediction using Machine Learning Technique," in 2018 3rd International Conference on Circuits, Control, Communication and Computing (I4C), Bangalore, India, Oct. 2018, pp. 1–4, doi:
6. 10.1109/CIMCA.2018.8739647.

7. M. Usmani, S. H. Adil, K. Raza, and S. S. A. Ali, "Stock market prediction using machine learning techniques," in 2016 3rd International Conference on computer and Information Sciences (ICCOINS), 2016, pp. 322–327.

8. J. Du, Q. Liu, K. Chen, and J. Wang, "Forecasting stock prices in two ways based on LSTM neural network," in 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Mar. 2019, pp. 1083–1086, doi:
9. 10.1109/ITNEC.2019.8729026.

10. S. E. Gao, B. S. Lin, and C.-M. Wang, "Share Price Trend Prediction Using CRNN with LSTM Structure," in 2018 International Symposium on Computer, Consumer and Control (IS3C), Dec. 2018, pp. 10–13, doi: 10.1109/IS3C.2018.00012.

11. T. Gao, Y. Chai, and Y. Liu, "Applying long short term momory neural networks for predicting stock closing price," in 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, Nov. 2017, pp. 575–578, doi:
12. 10.1109/ICSESS.2017.8342981

13. R. Y. Nivetha and C. Dhaya, "Developing a Prediction Model for Stock Analysis," in 2017 International Conference on Technical Advancements in Computers and Communications (ICTACC), Melmaurvathur, India, Apr. 2017, pp. 1–3, doi:
14. 10.1109/ICTACC.2017.11.

15. Z. Yeze and W. Yiying, "Stock Price Prediction Based on Information Entropy and Artificial Neural Network," in 2019 5th International Conference on Information Management (ICIM), Cambridge, United Kingdom, Mar. 2019, pp. 248–251, doi:
16. 10.1109/INFOMAN.2019.8714662.

17. "Basic LSTM Unit Transfer Function Diagram from [10] | Download Scientific Diagram."
18. https://www.researchgate.net/figure/Basic-LSTM-Unit-Transfer-FunctionDiagram-from-10_fig8_308804546 (accessed Jun. 03, 2020