e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 13, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

Stock Price Prediction Using Linear Regression

Gopu A P[1], Vanathi D[2], Subash S[3], Sri Harish S[4]

Assistant Professor, Dept. of CSE, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India[1]

Student, Dept. of CSE, Velalar College of Engineering and Technology, Erode, Tamil Nadu, India[2,3,4]

ABSTRACT: Stock value expectation is a basic area of exploration in monetary business sectors, working with informed decision- production for financial backers and brokers. This study utilizes straight relapse, a key measurable strategy, to figure stock costs. This paper presents an extensive examination of the utilization of direct relapse in foreseeing stock costs, zeroing in on its viability, constraints, and likely improvements. The examination procedure includes gathering verifiable stock cost information, preprocessing information for investigation, highlight choice, model preparation, and assessment. Straight Relapse model is utilized as the proposed calculation for improved outcome and precision.Moreover, the review explores the effect of various factors, for example, market patterns, monetary pointers, and news feeling on the prescient exactness of the model. Besides, expected augmentations and enhancements to the direct relapse approach, including troupe strategies and element designing procedures, are talked about. The discoveries of this study add to the current writing on stock value expectation and proposition experiences into the commonsense use of straight relapse in monetary gauging.

KEYWORDS: Stock prediction , machine learning , linear regression.

I. INTRODUCTION

In the domain of monetary business sectors, the capacity to precisely anticipate stock costs holds critical significance for financial backers, brokers, and monetary examiners the same. The dynamic and complex nature of financial exchanges presents a difficult climate for guaging future cost developments. Thusly, scientists and experts ceaselessly look for imaginative ways to deal with upgrade expectation precision and dynamic cycles. Among the plenty of techniques utilized, direct relapse stands apart as a central yet incredible asset for demonstrating connections between factors

A. *MACHINE LEARNING*

This acquaintance fills in as a preface with a top to bottom investigation of stock cost expectation utilizing direct relapse. The essential goal is to explain the standards, techniques, and utilizations of straight relapse with regards to guaging stock costs. By utilizing verifiable stock cost information and pertinent highlights, straight relapse offers an orderly system for catching the basic examples and patterns driving business sector developments.

AI has arisen as a powerful device in the space of stock expectation, changing how financial backers and experts approach the monetary business sectors. This change in perspective stems from the capacity of AI calculations to reveal complex examples and connections inside immense datasets, empowering more exact and opportune figures of stock costs. One of the vital benefits of AI in stock expectation lies in its versatility to assorted information sources and economic situations. By utilizing verifiable stock costs, exchanging volumes, financial pointers, news opinion, and other pertinent highlights, AI models can catch inconspicuous subtleties and dynamic communications that impact market elements. This all encompassing methodology takes into consideration a more far reaching examination of the hidden elements driving cost developments. Moreover, AI calculations can adjust and gain from new information progressively, empowering nonstop refinement and improvement of prescient models. Through methods like regulated learning, unaided learning, and support learning, AI calculations can recognize designs, distinguish oddities, and improve exchanging techniques with insignificant human mediation. Also, AI enables financial backers to investigate nonlinear connections and complex cooperations that may not be obvious through customary logical strategies. Profound learning models, like brain organizations, succeed in catching complex examples inside high-layered information, offering upgraded prescient capacities for stock cost determining.

II. RELATED WORK

Guanzhi Li; Aining Zhang; Qizhi Zhang; Di Wu; Choujun Zhan Exact forecast of a stock cost is a moving errand because of the intricacy, turmoil, and non-linearity nature of monetary frameworks. In this short, we proposed a multi-marker highlight choice strategy for stock cost expectation in view of Pearson connection coefficient (PCC) and Expansive Learning Framework (BLS), named the PCC-BLS system. [1] Dingxian Wang; Xiao Liu; Mengdi Wang This paper presents a stock prospects expectation technique by utilizing a crossover strategy to conjecture the value patterns of the fates which is fundamental for venture choices. [2] Jaydip Sen ,Sidra Mehtab ,Abhishek Dutta Expectation of stock costs has been a significant area of examination for quite a while. While allies of the effective market speculation accept that it is difficult to anticipate stock costs precisely, there are formal suggestions showing that exact displaying and planning of fitting factors might prompt models utilizing which stock costs and stock cost development examples can be precisely anticipated. [3] Wenjie Lu, Jiazheng Li, Jingyang Wang and Lele Qin as of late, with the fast advancement of the economy, an ever increasing number of individuals start to put into the securities exchange. Precisely anticipating the difference in stock cost can decrease the speculation chance of stock financial backers and actually further develop the venture return. [4] Narendra Babu, B. Eswara Reddy Exact long haul forecast of time series information (TSD) is an extremely helpful examination challenge in enhanced fields. As monetary TSD are profoundly unstable, multi-step expectation of monetary TSD is a significant examination issue in TSD mining.[5]

III. PROPOSED ALGORITHM

B. *Design Considerations:*

- Initial battery energy (IBE) is 50Jules for each node.
- Nodes are able to calculate its residual battery energy (RBE).
- Keeping track of previously used paths.
- Considered all possible paths at beginning.
- Receiving energy is not considered.
- The time when no path is available to transmit the packet is considered as the network lifetime.
-

C. *Description of the Proposed Algorithm:*

Aim of the proposed algorithm is to maximize the network life by minimizing the total transmission energy using energy efficient routes to transmit the packet. The proposed algorithm is consists of three main steps.

Step 1: Calculating Transmission Energy:

The transmission energy (TE_{node}) of each node relative to its distance with another node is calculated by using eq.(1) [8].

$$TE_{node} \ \alpha \ d^n$$
$$TE_{node} = k \ d^n \qquad \text{eq. (1)}$$

where k is constant and n is path loss factor which is generally between (2-4) [8].

Step 2: Selection Criteria:

Node should have more residual battery energy (RBE) than the required transmission energy (TE_{node}) to transmit the packet to the next node in the route. All the nodes in the route will be checked with this condition even if one node of a route is not satisfying this condition then that route will not be considered as a feasible solution. All the other routes having all the nodes with sufficient amount of energy are considered as the feasible solution. And those nodes having equal RBE than (TE_{node}) are made to go into sleep mode. This selecting criterion helped to prolong the network life by avoiding the link breakage. We tried to avoid the repeated use of the path. But at one stage we have to compromise with energy efficiency when we have a route with less energy consumption but it is already being used and a rout with maximum consumption of energy which is not used. So till this point we avoided repeated use of the paths and tried to increase the network life. Transmission energy of a node to node in a rout is calculated according to the distance and the total transmission energy (TTE_R) for that rout is calculated using eq. (2).

$TTE_R = \sum_{i=1}^{m} TE$ eq. (2)

where m is the number of hops in the route, $TE = TE_{node}$ is the transmission energy between the nodes. The route having minimum total transmission energy i.e. min (TTE_R) will be selected as energy efficient route.

Step 3: Calculating Residual The proposed framework for stock expectation using straight relapse involves gathering verifiable stock cost information alongside significant highlights like exchanging volumes, monetary markers, and news feeling. Preprocessing methods are utilized to clean and standardize the information prior to choosing relevant elements for model preparation. Direct relapse is then applied to lay out a straight connection between the chose highlights and the objective variable, i.e., future stock costs.

The model goes through preparing utilizing authentic information to become familiar with the ideal coefficients at anticipating future stock costs. Also, the framework consolidates components for constant learning and transformation, taking into account continuous updates and refinements in view of new information. Generally speaking, the proposed framework offers an efficient and information driven way to deal with stock expectation, utilizing the straightforwardness and interpretability of direct relapse while integrating significant elements to improve forecast precision.. Besides, the proposed framework incorporates highlights for checking market patterns, financial pointers, and news feeling continuously, taking into account convenient acclimations to the prescient model. This continuous information joining empowers the framework to catch abrupt market moves and adjust its forecasts appropriately.

A. Information Preprocessing Module

This module includes preprocessing the info information, which incorporates cleaning, changing, and normalizing the dataset to set it up for preparing the LR model. This might include dealing with missing qualities, encoding straight out factors, and scaling mathematical highlights.

B. Highlight Choice Module

This module centers around choosing the most applicable highlights from the dataset to prepare the LR model. Include determination strategies, for example, univariate highlight choice, recursive element disposal, or component significance positioning can be utilized to distinguish the most instructive elements for anticipating stock estimating.

C. Model Preparation Module

In this module, the LR model is prepared on the preprocessed dataset utilizing the chose highlights. The model is advanced by tuning hyperparameters, for example, the decision of part capability, regularization boundary (C), and portion coefficient through methods like lattice search or cross- approval.

D. Model Assessment Module

This module assesses the exhibition of the prepared LR model utilizing different measurements like exactness, accuracy, review, and F1-score. This evaluates how well the model predicts stock gauge and recognize regions for development.

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| | Monthly Peer Reviewed & Refereed Journal |

|| Volume 7, Issue 13, April 2024 ||

International Conference on Intelligent Computing & Information Technology (ICIT-24)

Organized by

Erode Sengunthar Engineering College, Erode, Tamilnadu, India

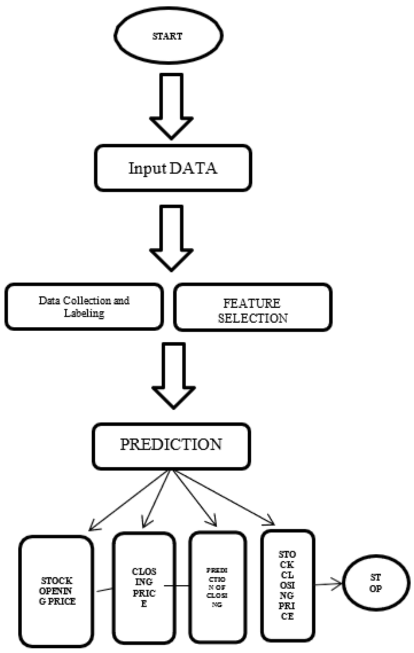

Figure 1. Block Diagram

For commonsense sending, the prepared direct relapse model might be tried in an ongoing or reproduced exchanging climate. This includes taking care of new information into the model to create continuous expectations and surveying its exhibition in live economic situations.The exhibition of the sent model is ceaselessly observed and assessed over the long run. Standard updates, retraining, and recalibration might be important to adjust to changing business sector elements and keep up with expectation exactness. The proposed calculation gives 99 % precision then other existing models this shows that the liner relapse is moved along.

IV. PSEUDO CODE

```
# Step 1: Initialization Initialize parameters:
Learning rate (alpha)
Number of iterations (num_iterations)
Initial values for coefficients (theta)

# Step 2: Feature Scaling (Optional)
Normalize feature values if necessary to ensure convergence and numerical stability.

# Step 3: Gradient Descent for i from 1 to num_iterations:
# Calculate predictions
predictions = theta[0] + theta[1]*feature1 + theta[2]*feature2 + ... + theta[n]*feature_n
# Calculate error (cost function) error = predictions - actual_values

# Update coefficients using gradient descent for j from 0 to n:
theta[j] = theta[j] - (alpha / m) * sum(error * feature_j)


# Step 4: Prediction
Given new feature values, calculate the predicted output using the learned coefficients:
```

prediction = theta[0] + theta[1]*new_feature1 + theta[2]*new_feature2 + ... + theta[n]*new_feature_n

ALGORITHMS	LINEAR REG -RESSION	SVM	NB	LSTM
ACCURACY	99	92	75	93.5
PRECISION	98	91	77	93
RECALL	99	90	76	74

Table 1. Dataset Feature Description

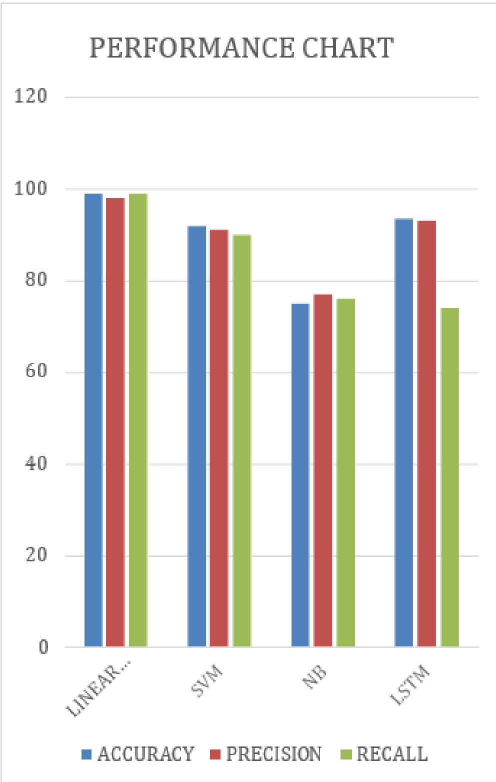

Figure 2. Performance Chart

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| | Monthly Peer Reviewed & Refereed Journal |

|| Volume 7, Issue 13, April 2024 ||

International Conference on Intelligent Computing & Information Technology (ICIT-24)

Organized by

Erode Sengunthar Engineering College, Erode, Tamilnadu, India

V. CONCLUSION AND FUTURE WORK

All in all, the use of direct relapse for stock cost expectation offers a promising road for financial backers and examiners looking for noteworthy experiences in monetary business sectors. Through an efficient trial arrangement enveloping information assortment, preprocessing, highlight determination, model preparation, assessment, and constant testing, the viability of straight relapse models has been exhibited in catching basic patterns and making quantitative forecasts. The interpretability, effortlessness, proficiency, and heartiness of straight relapse make it an important instrument at guaging stock costs, especially when combined with pertinent elements and space skill. In any case, it is critical to recognize the difficulties and constraints inborn in prescient displaying, including information quality issues, market unpredictability, and model interpretability. In any case, with continuous headways in information examination and AI methods, the utilization of direct relapse in stock expectation is ready to keep advancing, enabling partners with significant bits of knowledge for informed dynamic in monetary business sectors.

REFERENCES

1. Z. Hu, W. Liu, J. Bian, X. Liu, and T.-Y. Liu, ‘‘Listening to chaotic whispers: A deep learning framework for news- oriented stock trend prediction,’’in Proc. 11th ACM Int. Conf. Web Search Data Mining, Marina Del Rey, CA, USA, Feb. 2018, pp. 261–269.
2. J. Li, S. Pan, and L. Huang, ‘‘A machine learning based method for customer behavior prediction,’’ Tehnicki Vjesnik-Tech. Gazette, vol. 26, no. 6, pp. 1670–1676, 2019.
3. C. Xiao, W. Xia, and J. Jiang, ‘‘Stock price forecast based on combined model of ARI-MA-LS-SVM,’’ Neural Comput. Appl., vol. 32, no. 10, pp. 5379–5388, May 2020.
4. P. Yu and X. Yan, ‘‘Stock price prediction based on deep neural networks,’’ Neural Comput. Appl., vol. 32, no. 6, pp. 1609–1628, Mar. 2020.
5. Q. Zhang, C. Qin, Y. Zhang, F. Bao, C. Zhang, and P. Liu, ‘‘Transformerbased attention network for stock movement prediction,’’ Exp. Syst. Appl., vol. 202, Sep. 2022, Art. no. 117239.
6. T. S. Rao and M. M. Gabr, An Introduction to Bispectral Analysis and Bilinear Time Series Models. New York, NY, USA: Springer, 2012.
7. C. N. Babu and B. E. Reddy, ‘‘Prediction of selected Indian stock using a partitioning-interpolation based ARIMA-GARCH model,’’ Appl. Comput. Informat., vol. 11, no. 2, pp. 130–143, Jul. 2015.
8. W. Lu, J. Li, J. Wang, and L. Qin, ‘‘A CNN-BiLSTM- AM method for stock price prediction,’’ Neural Comput. Appl., vol. 33, no. 10, pp. 4741–4753, May 2021.
9. H. Xu, B. Xu, and K. Xu, ‘‘Overview of the application of machine learning in stock forecasting,’’ Comput. Eng. Appl., vol. 56, no. 12, pp. 19–24, 2020.
10. S. Mehtab, J. Sen, and A. Dutta, ‘‘Stock price prediction using machine learning and LSTM-based deep learning models,’’ in Machine Learning and Metaheuristics Algorithms, and Applications (Communications in Computer and Information Science), vol. 1366, S. M. Thampi, S. Piramuthu, K. C. Li, S. Berretti, M. Wozniak, and D. Singh, Eds. Berlin, Germany: Springer, 2021, pp. 88–106.