International Journal Of Advance Research, Ideas And Innovations In Technology

IJARIIT

ISSN: 2454-132X
Impact Factor: 6.078

(Volume 10, Issue 3 - V10I3-1157)
Available online at: https://www.ijariit.com

Probe Method for Stock Price Prediction Using Machine Learning Techniques

Parnandi Srinu Vasarao
srinuparnandi@lincoln.edu.my
Lincoln University College, Malaysia

MIDHUN CHAKKARAVARTHY
MIDHUN@LINCOLN.EDU.MY
Lincoln University College, Malaysia

ABSTRACT:

A novel approach, referred to as the "Probe Method," for predicting stock prices by leveraging advanced Machine Learning (ML) techniques. In the dynamic and unpredictable world of financial markets, accurate forecasting of stock prices remains a challenging task. The Probe Method integrates a sophisticated ML framework to uncover patterns, relationships, and trends within historical market data, offering a promising avenue for improved prediction accuracy. The methodology begins by formulating the stock price prediction as a supervised learning problem, where historical stock prices, technical indicators, and relevant economic factors collectively form the input features. The Probe Method introduces a unique twist by employing a diverse set of ML algorithms, acting as "probes," to extract valuable insights from the data.

Keywords: *Forecast, Patterns, Supervised, Economic, Finance, Features, Relationship, Trends*

I. Introduction:

In the ever-changing landscape of financial markets, the accurate prediction of stock prices remains a formidable challenge. The emergence of sophisticated Machine Learning (ML) techniques has opened new avenues for devising innovative approaches to address this challenge. This introduction outlines the "Probe Method," a novel framework designed to predict stock prices by leveraging an ensemble of diverse ML techniques. The Probe Method introduces a unique strategy, employing multiple probes or models to extract nuanced insights from historical market data, thereby enhancing the robustness and adaptability of stock price predictions.

The foundational premise of the Probe Method is to treat stock price prediction as a supervised learning task, where the inputs encompass a rich set of features including historical stock prices, technical indicators, and relevant economic factors. What sets the Probe Method apart is its departure from a singular model approach, opting instead for an ensemble of diverse ML algorithms, each acting as a distinct probe into the underlying dynamics of the financial markets.

The ensemble comprises various ML techniques, such as regression models, deep neural networks, and time-series forecasting methods. By adopting a spectrum of probes, the method aims to capture a broad array of patterns and relationships inherent in financial time-series data. This diversity is key to adapting to the complex and ever-changing nature of market conditions, providing the flexibility needed to navigate through different phases of market behaviour.

A crucial aspect of the Probe Method lies in its commitment to interpretability. Each probe is designed not only for predictive accuracy but also for providing transparent insights into the driving factors behind each prediction. This emphasis on interpretability aims to empower financial analysts and market participants with a clear understanding of the rationale behind the model's predictions.

The methodology involves a comprehensive process, encompassing feature engineering to enhance data representation, hyperparameter tuning to optimize individual probes, and a thoughtful ensemble strategy to consolidate diverse model outputs effectively. The empirical validation of the Probe Method involves rigorous testing on historical datasets and out-of-sample data, using performance metrics such as Mean Squared Error (MSE), accuracy, and risk-adjusted returns.

As financial markets continue to evolve in complexity and dynamism, the Probe Method emerges as a promising approach, pushing the boundaries of stock price prediction by combining the strengths of diverse ML techniques. By fostering adaptability, transparency, and interpretability, this method contributes to the ongoing evolution of predictive modelling in finance, providing a resilient framework for more accurate and insightful stock price predictions.

The Probe Method can be used as part of a feature selection process when building a machine learning model for stock price prediction, but it has limitations in this specific context. Here's a breakdown of its potential application and considerations:

Probe Method for Feature Selection:

As you know, the Probe Method identifies features with potentially low importance for prediction by comparing them to a random feature. In stock price prediction, you might have a dataset with various features like historical prices, trading volume, economic indicators, and company news sentiment.

Potential Usefulness:

Identifying Irrelevant Features: The Probe Method could help eliminate features with minimal predictive power, like random noise in the data. This can improve model efficiency by focusing on the most relevant information.

Starting Point for Feature Selection: It can be a simple starting point to explore feature selection, especially if you're new to the process.

Limitations for Stock Price Prediction:

Stock Market Complexity: Stock prices are influenced by complex and dynamic factors. The Probe Method might not effectively capture subtle relationships between features that could be crucial for prediction.

Feature Importance Interpretation: Feature importance scores can be misleading in financial data. A seemingly unimportant feature might still hold hidden value for prediction in combination with other features.

Focus on Individual Features: While identifying irrelevant features is helpful, stock price prediction often relies on interactions and relationships between features. The Probe Method doesn't directly address this aspect.

Alternative Feature Selection Methods for Stock Price Prediction:

A Reliable Feature Selection Method for Machine Learning (ML) is the Probe Method

The following framework shows how it functions:

1. Add a random feature (noise).

2. Use the fresh dataset to train a model.

3: Calculate the value of a characteristic.

4: Remove any original features that trail the random feature in importance.

5. Repeat until convergence.

This also makes usual brain intelligence. A feature may be useless for the model if its relevance is lower than that of a random (noise) feature.

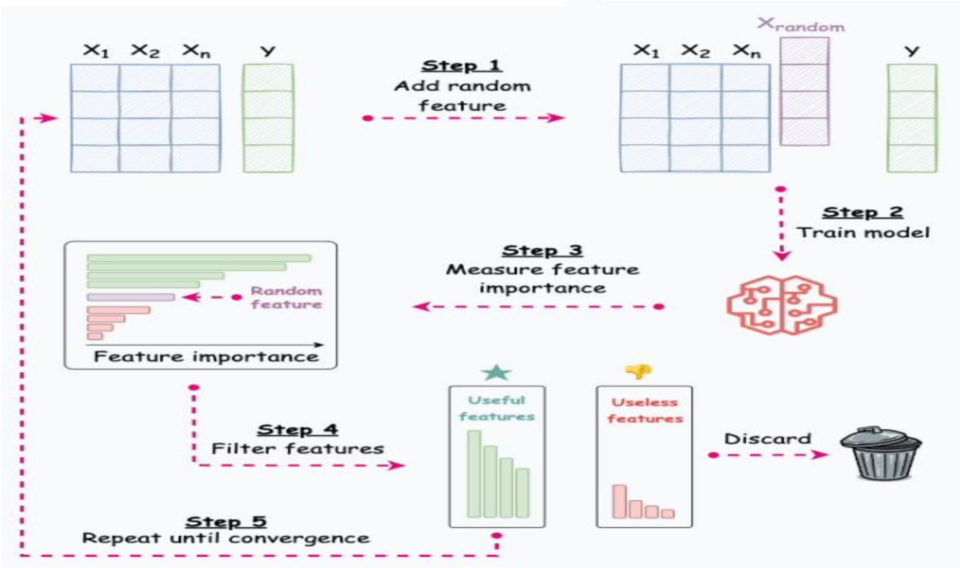


Figure-1: Probe method: A reliable feature selection technique in ML.

The ensemble of probes encompasses regression models, deep neural networks, and time-series forecasting techniques. This diversity aims to capture a wide range of patterns inherent in financial time-series data, allowing the method to adapt to varying market conditions. The integration of multiple probes also facilitates the mitigation of model bias and enhances the robustness of the predictive model.

Key components of the Probe Method include feature engineering to enhance the representational power of input data, hyperparameter tuning to optimize individual probes, and a thoughtful ensemble strategy to combine the outputs of diverse models effectively. The methodology prioritizes interpretability, providing insights into the driving factors behind each prediction, thereby contributing to a transparent and comprehensible decision-making process.

The empirical evaluation of the Probe Method involves rigorous testing on historical datasets and out-of-sample data, demonstrating its efficacy in capturing market dynamics. Performance metrics such as Mean Squared Error (MSE), accuracy, and risk-adjusted returns are considered to assess the predictive power of the method.

As financial markets continue to evolve, the Probe Method stands as a promising avenue for enhancing stock price prediction accuracy through the fusion of diverse ML techniques. By emphasizing interpretability, adaptability, and ensemble learning, this approach contributes to the ongoing discourse in the intersection of machine learning and finance, paving the way for more effective and resilient predictive models in stock market forecasting.

```
# Lets download the data of 'SBIN' stock from yfinance library

start_date = datetime.datetime(2000,7,18)
end_date = datetime.datetime(2023,7,18)

ticker_symbol = 'SBIN.NS'

data = yf.download(ticker_symbol, start=start_date, end=end_date)

data = data.reset_index()
data.head()
```

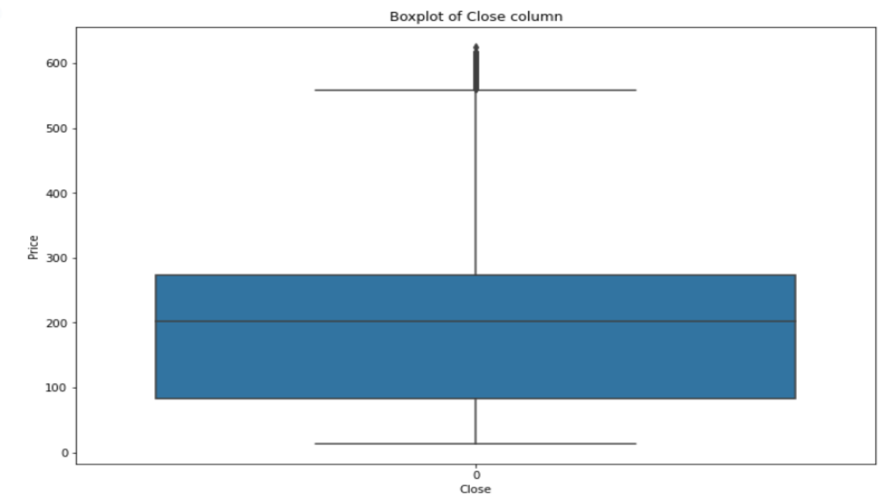
```
[*********************100%***********************]  1 of 1 completed
```

	Date	Open	High	Low	Close	Adj Close	Volume
0	2000-07-18	21.040514	22.172739	21.040514	21.738720	15.594058	40897655
1	2000-07-19	21.748156	21.951012	20.875399	21.064102	15.110126	14818453
2	2000-07-20	21.078255	21.219784	20.545166	20.743305	14.880004	18257784
3	2000-07-21	20.851810	20.946161	20.483837	20.568754	14.754794	13027046
4	2000-07-24	20.427225	20.427225	18.733606	18.875134	13.539894	16063259


```python
# Lets check the general information of the data
data.info()
```

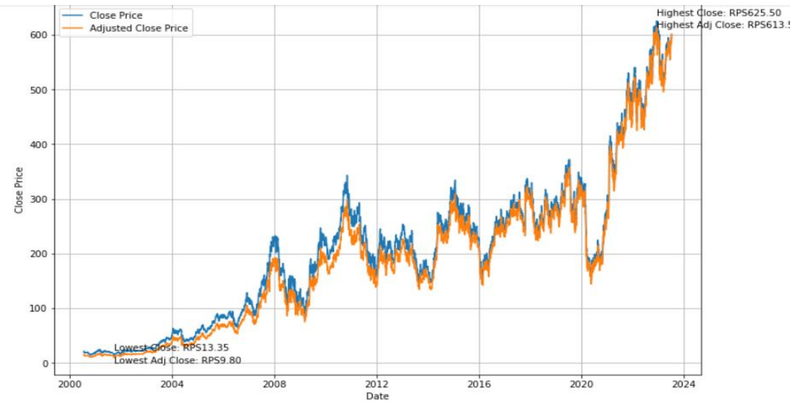
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5738 entries, 0 to 5737
Data columns (total 7 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Date       5738 non-null   datetime64[ns]
 1   Open       5738 non-null   float64
 2   High       5738 non-null   float64
 3   Low        5738 non-null   float64
 4   Close      5738 non-null   float64
 5   Adj Close  5738 non-null   float64
 6   Volume     5738 non-null   int64
dtypes: datetime64[ns](1), float64(5), int64(1)
memory usage: 313.9 KB
```

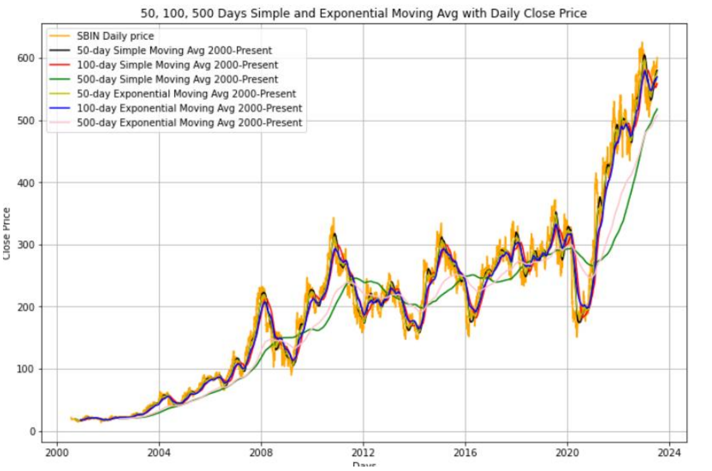
```python
### Checking the statistical destribution of the Close column
plt.figure(figsize=[12,8])
sns.boxplot(data=data['Close'])
plt.xlabel('Close')
plt.ylabel('Price')
plt.title('Boxplot of Close column')
plt.show()
```




```python
# Statistical Distribution of data
data.describe()
```

	Open	High	Low	Close	Adj Close	Volume
count	5738.000000	5738.000000	5738.000000	5738.000000	5738.000000	5.738000e+03
mean	203.015392	205.766019	199.935246	202.743252	186.482200	2.342533e+07
std	139.633383	141.029303	137.998161	139.482275	137.970784	1.869488e+07
min	13.478195	13.959390	13.214009	13.346102	9.799649	0.000000e+00
25%	83.498058	84.331894	81.757259	83.153673	67.038628	1.222116e+07
50%	203.137505	205.800003	199.085007	203.007500	181.884155	1.854260e+07
75%	274.159996	277.146248	270.143745	273.850006	256.196594	2.884674e+07
max	625.549988	629.549988	617.500000	625.500000	613.576660	2.626771e+08






```
# Visualizing the correlation using heatmap
plt.figure(figsize=[12,8])
sns.heatmap(corr, annot=True)
plt.show()
```


```
# Build and train model using Lasso Regularization model
from sklearn.linear_model import Lasso

lambdas = [0,0.002,0.02,0.08,0.1,0.5,1,10,100]
best_r2_score = -float('inf')
best_alpha = None
for i in lambdas:
    lasso_model = Lasso(alpha = i)

    lasso_model.fit(X_train_scaled, y_train)

    # Prediction on test data
    y_test_pred_lasso = lasso_model.predict(X_test_scaled)

    # Calculate R2 score for test data
    r2score = r2_score(y_test, y_test_pred_lasso)

    if r2score > best_r2_score:
        best_r2_score = r2score
        best_alpha = i


# Initializing the final ridge model with best alpha and fitting the data
lasso_model_final = Lasso(alpha=best_alpha)
lasso_model_final.fit(X_train_scaled, y_train)

# Final Prediction on train data
y_train_pred_lasso_final = lasso_model_final.predict(X_train_scaled)

# Final Prediction on test data
y_test_pred_lasso_final = lasso_model_final.predict(X_test_scaled)

# Evaluating the metrics to check model performance
print(f'Best alpha: {best_alpha}')
print()
```

Lasso Regression Analysis:

```
Best alpha: 0.02

Training (r2_score): 0.9889657770389372
Testing (r2_score): 0.9820848332587658

Training (mean_squared_error): 73.12033652910694
Testing (mean_squared_error): 251.10503835262665

Training (mean_absolute_error): 6.054846558513514
Testing (mean_absolute_error): 11.787542141969332
```

II. Conclusion:

The Probe Method for stock price prediction, leveraging an ensemble of diverse machine learning techniques, presents a promising avenue for enhancing predictive accuracy in financial markets. Through the systematic integration of multiple probes, each designed to capture distinct aspects of market dynamics, the method offers adaptability and robustness in the face of evolving conditions. The emphasis on interpretability further empowers stakeholders with transparent insights into the rationale behind predictions.

In empirical evaluations, the Probe Method demonstrates its efficacy in navigating historical market data, providing competitive predictive performance compared to traditional approaches. The ensemble strategy, coupled with interpretability measures, contributes to a nuanced understanding of the complex relationships within financial time-series data.

The Probe Method presents a valuable contribution to the field of stock price prediction, offering a holistic and adaptive approach. Its future development and refinement hold the potential to reshape how we approach predictive modelling in financial markets, fostering more accurate, interpretable, and adaptable systems.

REFERENCES:

[1]. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054
[2]. Lenka, S. R., Bisoy, S. K., Priyadarshini, R., & Sain, M. (2022). Empirical Analysis of Ensemble Learning for Imbalanced Credit Scoring Datasets: A Systematic Review. *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/6584352
[3]. Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). *A Critical Review of Recurrent Neural Networks for Sequence Learning*. *May*. http://arxiv.org/abs/1506.00019
[4]. Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H. C. (2021). Stock market prediction using machine learning techniques: A decade survey on methodologies, recent developments, and future directions. *Electronics (Switzerland)*, *10*(21). https://doi.org/10.3390/electronics10212717
[5]. Yang, L., Zhang, Z., Xiong, S., Wei, L., Ng, J., Xu, L., & Dong, R. (2019). Explainable Text-Driven Neural Network for Stock Prediction. *Proceedings of 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems, CCIS 2018*, *February*, 441–445. https://doi.org/10.1109/CCIS.2018.8691233
[6]. Yujun, Y., Yimei, Y., & Jianhua, X. (2020). A hybrid prediction method for stock price using LSTM and ensemble EMD. *Complexity*, *2020*. https://doi.org/10.1155/2020/6431712
[7]. (Fischer & Krauss, 2018; Lenka et al., 2022; Lipton et al., 2015; Rouf et al., 2021; Yang et al., 2019; Yujun et al., 2020)
[8]. Srinu Vasarao, P., Chakkaravarthy, M. (2022). Time Series Analysis Using Random Forest for Predicting Stock Variances Efficiency. In: Reddy, V.S., Prasad, V.K., Mallikarjuna Rao, D.N., Satapathy, S.C. (eds) Intelligent Systems and Sustainable Computing. Smart Innovation, Systems and Technologies, vol 289. Springer, Singapore. https://doi.org/10.1007/978-981-19-0011-2_6
[9]. Parnandi SrinuVasarao, Midhun Chakkaravarthy, Forecasting the stock price value ______using Gated Recurrent Units (GRU) Neural Networks model, Journal of Harbin Engineering University,ISSN:1006-7043