OPEN ACCESS

Volume: 11

Special Issue: 3

Month: July

Year: 2024

P-ISSN: 2321-788X

E-ISSN: 2582-0397

Received: 22.05.2024

Accepted: 20.06.2024

Published: 08.07.2024

Citation:
Hemamalini, D., et al. "Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data." *Shanlax International Journal of Arts Science and Humanities*, vol. 11, no. S3, 2024, pp. 26–33.

DOI:
https://doi.org/10.34293/sijash.v11iS3-July.7915

Predicting Stock Market Trends using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data

D. Hemamalini
Assistant Professor, Department of AI & DS, Arjun College Technology

C. Nagamani
Department of AI & DS, Arjun College 0f Technology

K. Deepesh
Department of AI & DS, Arjun College of Technology

P. Kamal
Department of AI & DS, Arjun College of Technology

Abstract

This investigation aimed to utilize machine learning algorithms for predicting stock market movements in Iran. The study centered on three specific sectors - diversified finance, information technology (IT), and metals - within the Tehran Stock Exchange. Ten years of historical data were analyzed . Incorporating ten technical indicators. To achieve this goal, six machine learning models were deployed. Support Vector Regression (Linear)Support Vector Regression (RBF)Linear Regression, Random -Forests ,K-Nearest Neighbours (KNN) Decision Trees.

Keywords: Machine Learning, Stock Market Prediction, Literature Review, Deep Learning, Support Vector Machine, KNN, LSTM, ANN, Investment Decision.

Introduction

The term "stock market" refers to various exchanges where shares of publicly held companies are bought and sold. These financial activities occur through formal exchanges and marketplaces that operate under a defined set of regulations. The stock market enables buyers and sellers of securities to meet, interact, and transact. It facilitates price discovery for corporate shares and serves as an economic barometer. Stock markets offer a secure and regulated environment where participants can confidently trade shares and other financial instruments with minimal operational risk. Operating under regulatory rules, stock markets function as both primary and secondary markets.

Machine-learning technology influences many aspects of modern society, from web searches and content filtering on social networks to recommendations on e-commerce websites. It is increasingly integrated into consumer products like cameras and smartphones.

Machine-learning systems are utilized to identify objects in images, transcribe speech to text, match news items, posts, or products.with user interests, and select relevant search results. Increasingly, these applications leverage deep learning techniques. Traditional machine-learning methods were limited in their ability to process natural data in its raw form, whereas deep learning offers enhanced capabilities.

Decision Tree

Decision Tree is a common supervised learning approach employed for both regression and classification problems. The goal of technique is forecasting a target by using easy decision rules shaped from the dataset and related features. Being easy to interpret or able to solve problems with different outputs aretwo advantages of using this model; on the contrary,constructing over-complex trees that cause overfitting is a typical disadvantage.

Random Forest Model

A random forest model consists of numerous decision trees and essentially averages their forecast results to create what is known as a forest. This algorithm incorporates three key random elements: randomly selecting training data when forming the trees, randomly choosing subsets of variables when splitting nodes, and considering only a subset of all variables for each node split in every individual decision tree.

K-Nearest Neighbour

KNN is known for its two main properties: lazy learning and being a non- parametric algorithm, as it does not assume any underlying data distribution. The method involves the following steps to find targets.

Support Vector Regression

This method is similar to Support Vector Regression with the parameter but uses lib linear instead of lib svm. The use of lib linear offers greater flexibility in selecting penalties and loss functions, and it provides better scalability for large datasets. Linear Regression: Linear Regression is a fundamental and widely-used machine learning

Motivation

Our motivation arises from the significant post- pandemic profits observed in the global economy, which have complicated the decision-making process for investors looking to buy or sell shares in highly profitable companies. To address this challenge, we have chosen this method to predict stock market trends, facilitating more informed and accurate investment decisions for buying and selling shares

Aim and Scope

Aim

The aim of this project is to predict the future value of stocks using a straightforward system. This approach ensures accessibility for administrators and stock brokers, enabling them to provide investors with essential information about potential stock investments. By leveraging a few machine learning algorithms and comparing their outputs, we aim to deliver the most accurate predictions.

Scope

With the rising number of investors and increasing share values, public interest in stock market investments is growing proportionally. Therefore, the use of a prediction system will become increasingly necessary and practical in the coming years.

Literature Review

We are going to represent literature overview of few papers that are studied for choosing the topicas follows:

1. "Forecasting stock market movement direction with support vector machine". This paper explains that Support Vector Machine (SVM) is a very specific type of learning algorithms characterized by the capacity control of the decision function, the use of the kernel functions and the sparsity of the solution. In this paper, we investigate the predictability of financial movementdirection with SVM by forecasting the weekly movement direction of NIKKEI 225 index.
2. "Predicting the direction of stock market prices using tree-based classifiers". This paper explains that the Predicting returns in the stock market is usually posed as a forecasting problem where prices are predicted. Intrinsic volatility in the stock market across the globe makes the task of prediction challenging. Consequently, forecasting and diffusion modelling undermines a diverse range of problems encountered in predicting trends in the stock market.

Methodology

System Architecture: Data Collection

This is the initial and crucial step towards developing a machine learning model, involving the gathering of relevant data.

Data Preprocessing

Data preprocessing involves preparing raw data and making it suitable for a machine learning model. This step ensures the data is clean and in the right format.

Feature Extraction

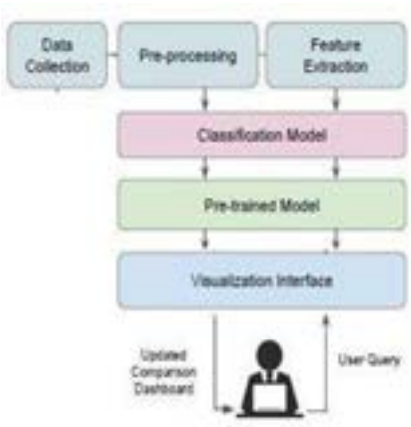
Feature extraction is a technique used to reduce the amount of resources required while preserving essential information. It plays a crucial role in enhancing the efficiency and accuracy of machine learning models.

Classification Model

Classification models are a type of supervised machine learning. They take input data and generate an output that categorizes the input into specific categories.

Pre-Trained Model

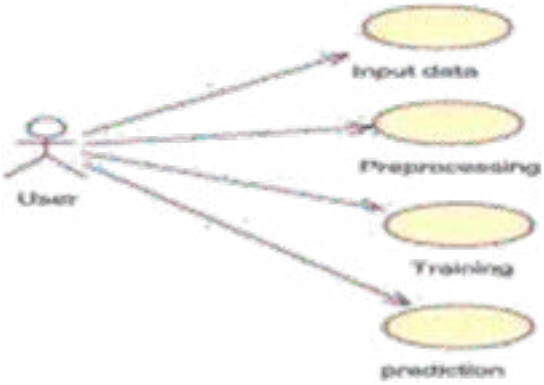
A pre-trained AI model is a deep learning model that mimics a brain-like neural network, identifying patterns or making predictions based on data. Trained on large datasets for a specific task, it can be used as is or further fine-tuned to meet the specific requirements of an application.



User Case Diagram

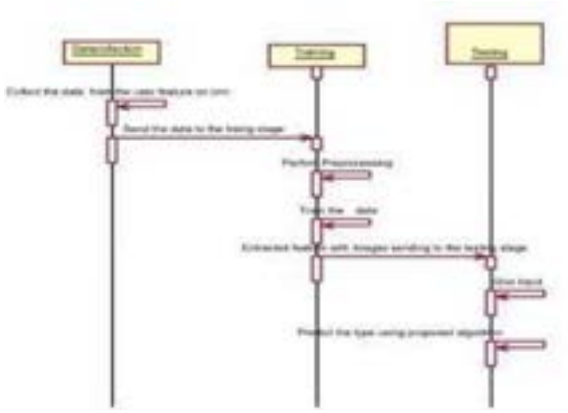
A use case diagram in the Unified Modelling Language (UML) is a type of behavioural diagram defined by and created from a Use- case analysis. Its purpose is to present a graphical overview of the functionality provided by a systemin terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is toshow what system functions are performed for which actor.

Sequence Diagram



Sequence Diagram

A sequence diagram in Unified Modelling Language (UML) is a kind of interaction diagram that shows how processes operate with one anotherand in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.



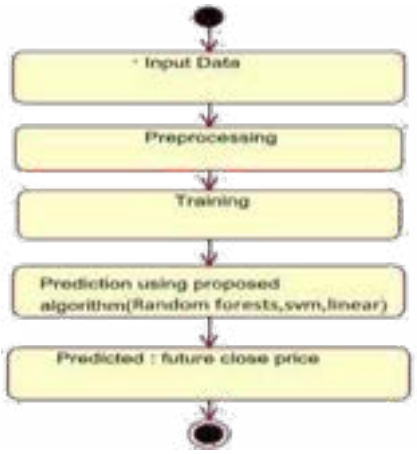
Activity Diagram

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modelling Language, activity diagrams canbe used to describe the business and operational step-by- step workflows of components in a system activity diagram shows the overall flow of control.

Proposed System

Modules

- Data Collections
- Dataset
- Data Preparation
- Model Selection
- Analyse and Prediction Accuracy



on test set

- Saving the Trained Model

Module Description

Data Collection

This is the first real step towards the real development of a machine learning model, collecting data. This is a critical step that will cascade in how good the model will be, the more and better data

that we get, the better our model will perform. There are several techniques to collect thedata, like web scraping, manual interventions and etc.

Dataset

The dataset consists of 100252 individual data. There are 8 columns in the dataset but we taken only two country data set one India and other USAtotally which are described below.

- Date - specifies trading date
- Open - opening price
- High - maximum price during the
- DayLow - minimum price during the
- Day Close - close price adjusted for splits
- Volume - the number of shares that changed handsduring a given day
- Turnover - It is the cost of goods sold divided byaverage inventory.

Data Preparation

We will transform the data. By getting rid of missing data and removing some columns. First Wewill create a list of column names that we want to keep or retain. Next we drop or remove all columnsexcept for the columns that we want to retain.Finally we drop or remove the rows that have missing values from the data set.

Model Selection

When developing a machine learning model, we require two datasets: one for training and one for testing. Currently, we have only one dataset, so we will split it into two parts with an 80:20 ratio. Additionally, we will separate the data frame into feature columns and label columns. For this purpose, we will use the function from Scikit-learn. We will also implement various algorithms, including Linear Regression, Random Forest Regression, K-Nearest Neighbors (KNN), Decision Trees, SVR (Linear), and SVR (RBF).

Mean Square Error (MSE)

When developing a machine learning model, two datasets are needed: one for training and one for testing. Since we currently have only one dataset, we will split it into two parts with an 80:20 ratio. We will also divide the data frame into feature columns and label columns. To accomplish this, we will use the function from Scikit-learn. Furthermore, we will implement several algorithms, including Linear Regression, Random Forest Regression, K-Nearest Neighbors (KNN), Decision Trees, SVR (Linear), and SVR (RBF).

MSE Can be Calculated as

MSE = 1 n Xn i=1 (yi − ˆyi)

Analyze and Prediction

1. Date - specifies trading date
2. Close - close price adjusted for splits

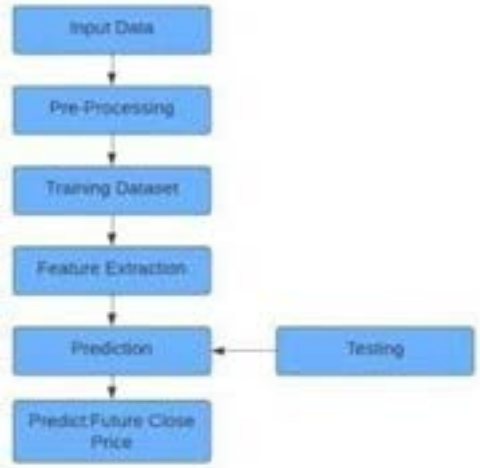
Accuracy on Test Set

We got a accuracy of 94.2% on test set.

Saving the Trained Model

Once you're confident enough to take your trained and tested model into the production-ready environment, the first step is to save it into a .p k l file using a library like pickle ..Make sure you have pickle installed in your environment. Next, let's import the module and dump the model into pkl file.

Flow Chart



Conclusion

Focus on the experiment: Researchers investigated the potential of machine learning to predict stock movements in Iran's Tehran Stock Exchange. They focused on three sectors: diversified finance, IT, and metals. The analysis used ten years of historical data with ten technical indicators, employing six machine learning models for comparison.

Highlight the data and models: This study explored the use of six machine learning models (Support Vector Regression with linear and RBF kernels, Linear Regression, Random Forests, K- Nearest Neighbors, and Decision Trees) to predict stock market movements in three sectors (diversified finance, IT, and metals) of the Tehran Stock Exchange. The analysis utilized ten years of historical data with ten technical features.

Simplify the technical details: To predict stock market trends in Iran, researchers used machine learning on ten years of data from three sectors (finance, IT, and metals) of the Tehran Stock Exchange. They compared six different machine learning approaches..

Future Scope

Predicting stock market trends using machine learning has a lot of potential for the future. Machine learning can help analyze large volumes of data, identify patterns, and make predictions based on historical data. Some potential future scope for predicting stock market trends using machine learning are:

As more data becomes available and machine learning algorithms become more advanced, the accuracy of predictions is likely to improve. This could lead to better investment decisions and increased profits. With the help of machine learning,it's possible to analyze large volumes of data in real-time. This could help investors make Better decisions and respond quickly to market changes. Machine learning can be used to develop automated trading algorithms that make trades based on real-time data and market trends. This could lead to increased efficiency and reduced costs for investors. Machine learning algorithms can learn from individual investors' preferences

and investment strategies. This could lead to more personalized investment recommendations that take into account an individual's risk tolerance,investment goals, and other factors.

References

1. W. Huang, Y. Nakamori, and "Forecasting stock market movement direction with support vector machine", vol. 32, no. 10, pp. 25132522, Oct. 2005.
2. S. Basak, S. Kar, S. Saha, L. S. R. Dey, "Predicting the direction of stock market prices using tree", vol. 47, pp. 552567, Jan. 2019.
3. Kelotra, A. and P. Pandey, "Stock market prediction using optimized deep- conv LSTM mode" vol. 80, Sep. 2017.
4. Lu, Ning, "A Machine Learning Approach to Automated Trading", Boston, MA,USA: Boston College Computer Science Senior, 2016.
5. J. Sun and H. Li, "Financial distress prediction using support vector machines:Ensemble vs. Individual", vol. 12, no. 8, pp. 22542265, Aug.2012.
6. J. Patel, S. Shah, P. Thakkar, "Predicting stock market index using fusion of machine learning techniques", vol. 42, no. 4, pp. 21622172, Mar.2015.
7. J. B. Duarte Duarte, L. H. Talero Sarmiento, and K. J. Sierra Juárez, "Evaluation of the effect of investor psychology on an artificial stock market through its degree of efciency", vol. 62, no. 4, pp.13611376, Oct. 2017.
8. Mojtaba Nabipor, pooyan nayyeri, Hame Jabani, Shahab.S, Amir Mosavi, "Predicting stock market using machine learning and Deep learning Algorithms Via Continuous and Binary Data; a Comparative Analysis", oct 2019.