

**DR. D. Y. PATIL SCHOOL OF SCIENCE AND TECHNOLOGY  
TATHAWADE, PUNE**

**A Foundation of Data Science-Report on  
S&P-500 Index Price Prediction.**

**SUBMITTED BY:**

<b>NAME OF STUDENT</b>	<b>ROLL NUMBER</b>
<b>1.TEJAS BHAVSAR</b>	<b>BTAI-204</b>
<b>2. ROHAN MASHERE</b>	<b>BTAI-228</b>
<b>3.PRATHEMESH NANGARE</b>	<b>BTAI-230</b>

**GUIDED BY:**

**Mrs. Mily Lal.**

## ABSTRACT

Predicting stock prices is a significant challenge in finance and a popular application of machine learning. This project aims to forecast the S&P 500 index using historical data and various machine learning models, focusing on technical and fundamental analysis. Stock prices are influenced by multiple factors, including market conditions, industry trends, leadership changes, and acquisitions. While sentiment analysis from news and social media can improve predictions, this study relies solely on historical price data and key technical indicators. Data from the last 20 years was extracted using the Yahoo Finance API, along with five major tech stocks (Apple, Amazon, Microsoft, Netflix, and Google) for comparison. Technical indicators such as future price change, simple moving averages, relative strength index, and exponential moving averages were incorporated as features. Several machine learning models, including Decision Trees, Random Forests, Gradient Boosting, Neural Networks, and Facebook's Prophet, were tested. Traditional models struggled with predictive accuracy, while Facebook's Prophet showed promising results, forecasting an upward trend in the S&P 500 over the next 2–3 years. This study highlights the potential of machine learning in stock price prediction while acknowledging its limitations, suggesting that time-series forecasting models like Prophet may be more effective. Future work could enhance predictions by integrating sentiment analysis and macroeconomic indicators.

**Keywords:** Stock Price Prediction, S&P 500, Machine Learning, Time-Series Forecasting, Technical Analysis, Fundamental Analysis, Decision Trees, Random Forests, Gradient Boosting, Neural Networks, Facebook's Prophet, Historical Data, Market Trends.

## **PROBLEM STATEMENT**

Stock prices are influenced by various factors, making their prediction challenging. This project aims to forecast S&P-500 stock prices using historical data and machine learning models, to identify potential trends and improve predictive accuracy.

## **OBJECTIVES**

1. Develop a predictive model using historical S&P 500 data to forecast stock price trends.
2. Analyse the impact of technical indicators such as moving averages, RSI, and exponential moving averages on stock price prediction.
3. Compare the performance of different machine learning models, including Decision Trees, Random Forests, Gradient Boosting, Neural Networks, and Facebook's Prophet.
4. Improve predictive accuracy by identifying the most effective model for forecasting stock prices.

## **DATASET**

The dataset for this project was obtained from the Yahoo Finance API, which provides historical stock market data dating back to the 1990s. Using the Yahoo Finance API, covering 20 years of S&P 500 stock price data were collected. Additionally, stock prices of five major technology companies (Apple, Amazon, Microsoft, Netflix, and Google) from the last five years were extracted for comparison. This dataset includes key stock market features such as Open, Close, High, Low, Volume, and Adjusted Close, which are essential for analysing price trends and market behaviour.

This data is highly relevant for stock price prediction as it captures long-term market trends and daily fluctuations. The inclusion of major tech stocks allows for correlation analysis with the S&P 500, helping to identify patterns and dependencies. By leveraging historical data, technical indicators, and machine learning models, this dataset serves as a strong foundation for predicting future stock prices.

**END!!**