**ABSTRACT**

Predicting stock ups and downs has always been challenging, so here we are analyzing Stock Market Dataset to predict the future behavior of stocks. While it is impossible to predict the future movement of a specific stock, using technical analysis and research helps us make better predictions. We are using historical data, i.e., analyzing the previous trends to predict future trends. Here, we are considering factors like moving average, which is a stock indicator, to identify the trend direction of a stock or determine its support and resistance levels. It is a trend-following or lagging indicator based on past prices. Another factor we are considering is a stock correlation, i.e., how a stock performs w.r.t other stock, which helps us establish relationships over different stocks and improve our prediction. We are creating a stock portfolio, which helps us identify one’s goals and timeline and proceed accordingly. While making a portfolio, we try to follow the classical theory, i.e., minimizing risk and maximizing profitability. We will be generating random ratios of stocks and calculating each portfolio's risk and profitability according to the mathematical formulas. Since our dataset has a lot of random values, we use the Monte Carlo method to easier our task because it is a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results. We are using python libraries such as NumPy, pandas, matplotlib, and seaborn to help us complete our aim. We aim to make one of the most efficient stock predictors.

**PROBLEM DEFINITION**

We aim to analyze the stock market dataset to predict a specific stock, whether it’ll rise or fall, which will help us maximize profitability with minimum risk.

**STATE-OF-THE-ART**

We have been learning about stock market-related terms to deepen our understanding of various stock-market terminology. Also, we have decided on the python libraries we want to use and have a good deal of our dataset. We are close to finding the moving average of our dataset, which will be an essential factor to help us reach our aim. We have been analyzing different graphs and drawing conclusions from them.

**NOVELTY**

Over the years, different methods have been used to predict stocks. We have chosen a Monte Carlo Method technique to calculate the value at risk (VaR) of a portfolio; we will run a Monte Carlo simulation that attempts to predict the worst likely loss for a portfolio given a confidence interval over a specified period horizon. To make our program more efficient, we are not only dependent on Monte Carlo, but also on other factors like moving average and correlation, which will narrow down our search and probably give a better result than others.

**DATASET**

We are using a financial dataset. It consists of full historical daily price and volume data for all US-based stocks and ETFs trading on the NYSE, NASDAQ, and NYSE MKT. The size of the data set is approximately 1.5 GB. This Dataset consists of the following attributes: Date, Open, High, Low, Close, Volume, OpenInt. Open means the price at which a stock started trading when the opening bell rang. Close refers to the price of an individual stock when the stock exchange closed shop for the day. The high is the highest price at which a stock is traded during a period. The low is the lowest price of the period. A stock's high and low points for the day are often called its intraday high and low. Volume measures the number of shares traded in a stock or contracts traded in futures or options. We aim to analyze the behavior of this stock and predict its future, considering how it performed in the last couple of years and what factors affected it.