# PREDICTING THE STOCK



MARKET

#### MAG 7 STOCK PRICE PREDICTION THROUGH ML

- The MAG 7 Stocks
  - A combined market capitalization of about \$14 trillion
  - Exposure to high-growth technologies such as high-end software and hardware, cloud computing, and Artificial Intelligence
  - Each of the seven stocks has outperformed the S&P 500 by a huge margin in the past decade
- The ML Approach and Impact
  - Collect and preprocess stock price and Financial Indicators' data.
  - Combine Technical and Fundamental Factors to predict future stock prices.
  - Can be used by portfolio managers of investment funds or individual investors to maximize returns or to balance Risk and Rewards.





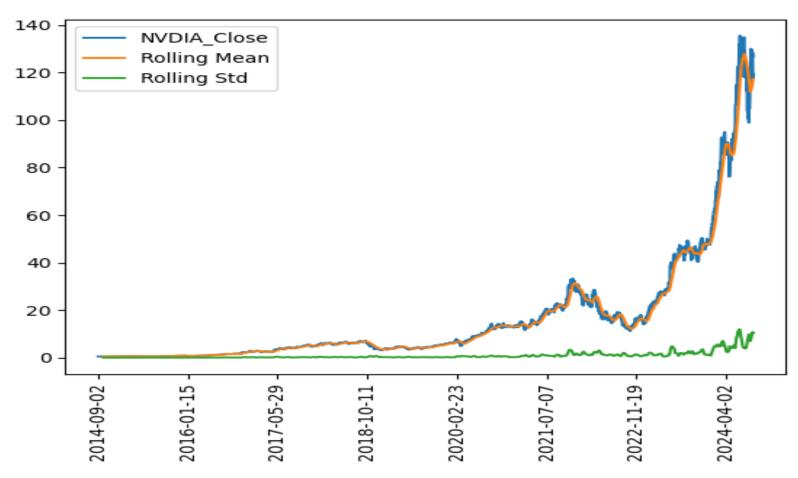
## COLLECTION OF DATA AND PREPROCESSING

- Gathered last 10 years' stock price data from "yfinance" APIs
- Gathered last 10 years' financial indicators' data from "Fred API"
- Financial indicators included "GDP Growth", "Inflation", "Unemployment Rate", "Retail Sales" and "Industrial Production".
- Fundamental Factors included "P/E Ratio" and "EPS"
- Calculated Technical Indicators like Rolling Average
- Forward filled and removed rows containing missing data
- Calculated lagged independent variables





# EXPLORATORY DATA ANALYSIS





GDP Growth -	1	0.086	0.13	0.056	-0.092	-0.011	-0.044	0.018	-0.0089	0.0069	0.57	0.092	0.036	0.065	-0.058	-0.033	0.027	-0.0069	0.0044
Inflation (CPI) -	0.086	1	-0.27	0.93	0.33	0.79	0.71	0.73	0.8	0.71	0.098	1	-0.27	0.93	0.31	0.72	0.72	0.8	0.7
Unemployment Rate -	0.13	-0.27	1	-0.28	-0.89	-0.17	-0.025	-0.18	-0.17	-0.16	0.0047	-0.27	0.83	-0.3	-0.8	-0.044	-0.18	-0.18	-0.16
Retail Sales -	0.056	0.93	-0.28	1	0.34						0.11	0.92	-0.21	0.89	0.25				
Industrial Production -	-0.092	0.33	-0.89	0.34	1	0.24	0.026	0.26	0.24	0.23	0.038	0.32	-0.73	0.35	0.87	0.052	0.26	0.25	0.22
NVIDIA_Close -	-0.011		-0.17		0.24	1	0.39	0.95	0.99	0.92	-0.0072		-0.17		0.22	0.41	0.95	0.98	0.87
PE Ratio -	-0.044		-0.025	0.7	0.026	0.39	1	0.24	0.39	0.3	-0.048		-0.0066		-0.012	0.95	0.22	0.4	0.32
Quarterly EPS -	0.018		-0.18		0.26	0.95	0.24		0.95	0.86	0.02		-0.19		0.25	0.29	0.98	0.95	0.82
Rolling Mean -	-0.0089		-0.17		0.24	0.99	0.39	0.95		0.89	-0.0028		-0.17		0.22	0.42	0.95	0.99	0.9
Rolling Std -	0.0069		-0.16		0.23	0.92	0.3	0.86	0.89		0.008		-0.17		0.21	0.31	0.86	0.87	0.73
Lagged_GDP -		0.098	0.0047	0.11	0.038	-0.0072	-0.048	0.02	-0.0028	0.008	1	0.095	0.12	0.064	-0.094	-0.038	0.03	0.001	0.017
Lagged_Inflation -	0.092		-0.27	0.92	0.32						0.095		-0.28	0.93	0.31			0.8	0.7
Lagged_Unemp_rate -	0.036	-0.27	0.83	-0.21		-0.17	-0.0066	-0.19	-0.17	-0.17	0.12	-0.28	1	-0.28	-0.89	-0.03	-0.19	-0.18	-0.17
Lagged_Retail_Sales -	0.065	0.93	-0.3	0.89	0.35						0.064	0.93	-0.28	1	0.32				
Lagged_Industrial_prod -	-0.058	0.31	-0.8	0.25	0.87	0.22	-0.012	0.25	0.22	0.21	-0.094	0.31	-0.89	0.32		0.017	0.25	0.23	0.22
Lagged_PE -	-0.033		-0.044	0.72	0.052	0.41	0.95	0.29	0.42	0.31	-0.038		-0.03		0.017	1	0.25	0.42	0.33
Lagged_EPS -	0.027		-0.18		0.26	0.95	0.22	0.98	0.95	0.86	0.03		-0.19		0.25	0.25		0.95	0.83
Lagged_Rolling_Mean -	-0.0069		-0.18		0.25	0.98	0.4	0.95	0.99	0.87	0.001		-0.18		0.23	0.42	0.95		0.87
Lagged_Rolling_Std -			-0.16		0.22	0.87	0.32	0.82	0.9		0.017		-0.17		0.22	0.33	0.83	0.87	1
Ву	rthweig degle de Roh	it Gro	Unergoyment Rate -	Retail Sales -	Industrial Production -	WIDIA_Close -	PE Ratio -	Quarterly EPS -	Rolling Mean -	Rolling Std -	Lagged_GDP -	Lagged_Inflation -	Lagged_Unemp_rate -	Lagged_Retail_Sales -	Lagged_Industrial_prod -	Lagged_PE -	Lagged_EPS -	Lagged_Rolling_Mean -	Lagged_Rolling_Std -





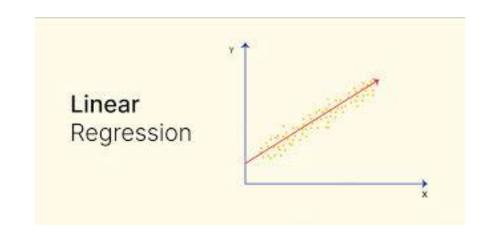
### ML MODELS

#### ➤ Baseline Model

- Train-Test Split and The Treachery of Data Leakage
  - Achieved an R-sq of 0.99 ©
- Split Data by Dates
  - Fit a Linear Regression Model
  - Got an R-sq of 0.77 after adjusting the model and feature elimination through Recursive Feature Elimination (RFE)

#### >Advanced Neural Network

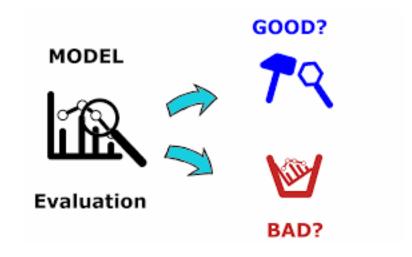
- Grid Search CV and Hyperparameter Tuning
- Improved the R-sq to 0.79





### MODEL EVALUATION

- Mean Absolute Error (MAE)
- Mean Squared Error (MSE) / Root Mean Squared Error (RMSE)
- R-squared (R^2)
- Directional Accuracy
- Mean Absolute Percentage Error (MAPE)





#### ML MODEL HOSTED ON STREAMLIT - DEMO

https://groverrohit-bs-capstone-predicting-mag7-stockprices.streamlit.app/





