DEEP INSIGHTS INTO MARKET EFFICIENCY: A NEURAL NETWORK BASED APPROACH.

Yeshwant S Nair
2239216

Dissertation Supervisor Dr. Vineeth Mohandas,
Department of Economics,
Christ (Deemed to be University); Bangalore

INTRODUCTION

The Efficient Market Hypothesis (EMH) asserts that financial markets are efficient, reflecting all relevant information in security prices.

Developed by Eugene Fama, and Paul Samuelson in the 1960s and 1970s.

EMH suggests it's difficult for investors to consistently outperform the market due to the randomness of stock prices.

Embracing the EMH, Andrew Lo (2004) propounded the Adaptive Market Hypothesis, which expands the idea of market efficiency by incorporating insights from behavioural finance and acknowledges the role of varying market conditions.

THEORETICAL LITERATURE REVIEW

Literature	Observations
Efficient capital markets: A review of theory and empirical work - Fama, E. F. (1970).	Defined Efficient market in all three forms : Weak, Semi-strong and Strong Form.
The adaptive markets hypothesis: Market efficiency from an evolutionary perspective - Lo, A. W. (2004).	Market's predictability varies over time, which can be attributed to the changing market conditions.
Efficient market hypothesis and forecasting.- Timmerman, A., & Granger, C. W. (2004).	The gains are short-lived for the first time users of new financial prediction methods
Efficient market hypothesis in European stock markets - Borges, M. R. (2010).	The transition of countries from emerging to developed markets in the early twenty-first century could explain these changes. As markets develop, increased participation, information flow, and regulatory changes might contribute to a decrease in return predictability and an increase in market efficiency
Are stock markets really efficient? Evidence of the adaptive market hypothesis - Urquhart, A., & McGroarty, F. (2016).	Return predictability fluctuated over time in each market, with each of the return series going through a period of significant predictability and periods associated with no predictability. Furthermore, different markets experience significant predictability at different points, suggesting that markets evolve differently with time

EMPIRICAL LITERATURE REVIEW

Literature	Observations
A comparison of artificial neural network and time series models for forecasting commodity prices - Kohzadi, N., Boyd, M. S., Kermanshahi, B., & Kaastra, I. (1996).	Compared the ANN with the ARIMA model for forecasting monthly data. The results showed that ANN forecasts provided a considerably more accurate prediction consistently and were able to capture more turning points compared to the ARIMA model.
Hansson, M., & Olsson, C. (2017). Feedforward neural networks with ReLU activation functions are linear splines. Bachelor's Theses in Mathematical Sciences	Used ReLU activation function to introduce non-linearity in their modelling process, they did so because the ReLU activation function helps prevent saturation of the gradient when the neural networks are deep thereby preventing the problems associated with the vanishing gradient problem.
Jais, I. K. M., Ismail, A. R., & Nisa, S. Q. (2019). Adam optimization algorithm for wide and deep neural network. Knowl. Eng. Data Sci., 2(1), 41-46.	Discussed the impact of Adam on wide and deep neural networks. They found that the use of Adam resulted in improved performance of the neural network which was backed by higher accuracy and better metrics as opposed to when Adam was not used. Their work suggested that the Adam optimization function was ideal for wide and deep neural networks.
Automatic neural network modelling for univariate time series - Balkin, S. D., & Ord, J. K. (2000).	Compared the ANN with the ARIMA model for forecasting monthly data. The results showed that ANN forecasts provided a considerably more accurate prediction consistently and were able to capture more turning points compared to the ARIMA model.
Balkin, S. D., & Ord, J. K. (2000). Automatic neural network modeling for univariate time series. *International Journal of Forecasting*, *16*(4), 509-515.	ANN usually performs better when we have a sufficiently long series to detect the non-linearity and provide reliable estimates of the parameters. When these conditions are satisfied, the ANN model outperforms more straightforward methods

- Most of the literature for assessing the validity of EMH is done using traditional econometric methods
- To use AI/ML methods to predict future movement in prices using historical prices and information.
- To check the validity of EMH using AI/ML methods.

RESEARCH QUESTIONS

1. Can deep learning models be used to predict future movement in prices of the NIFTY 50 index? If yes, how do they compare to traditional econometric models?
2. Does incorporating historical news information into the deep learning models serve to improve the predictability of future price movements?
3. Based on the findings from the above three questions, do we validate or invalidate the weak and semi-strong forms of EMH?

OBJECTIVES

1. To evaluate the scope of deep learning models in predicting the future movement of the NIFTY 50 index using past prices and to compare their performance against the traditional econometric models such as ARCH
2. To use Natural Language Processing, to investigate the role of past news formations in predicting the future prices of NIFTY 50.
3. To assess the validity or invalidity of EMH forms in the NIFTY 50 Index based on the above objectives

DATA & METHODOLOGY

VARIABLES

- NIFTY 50 Index
- News from public domain

METHODS

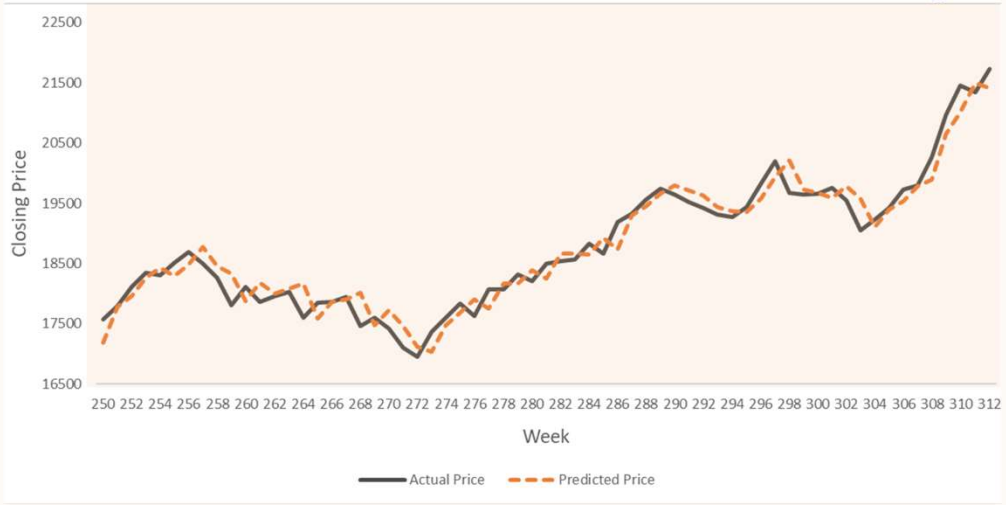
- Deep Neural Networks
- Natural Language Processing
- ARCH

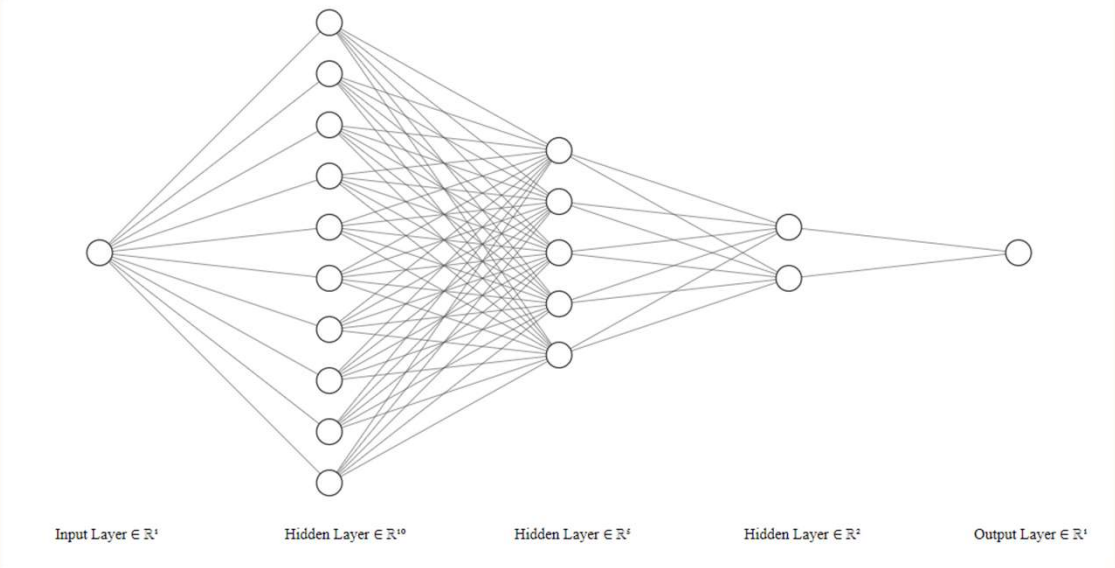
TIME PERIOD

- From 2018 to 2023

ANALYSIS

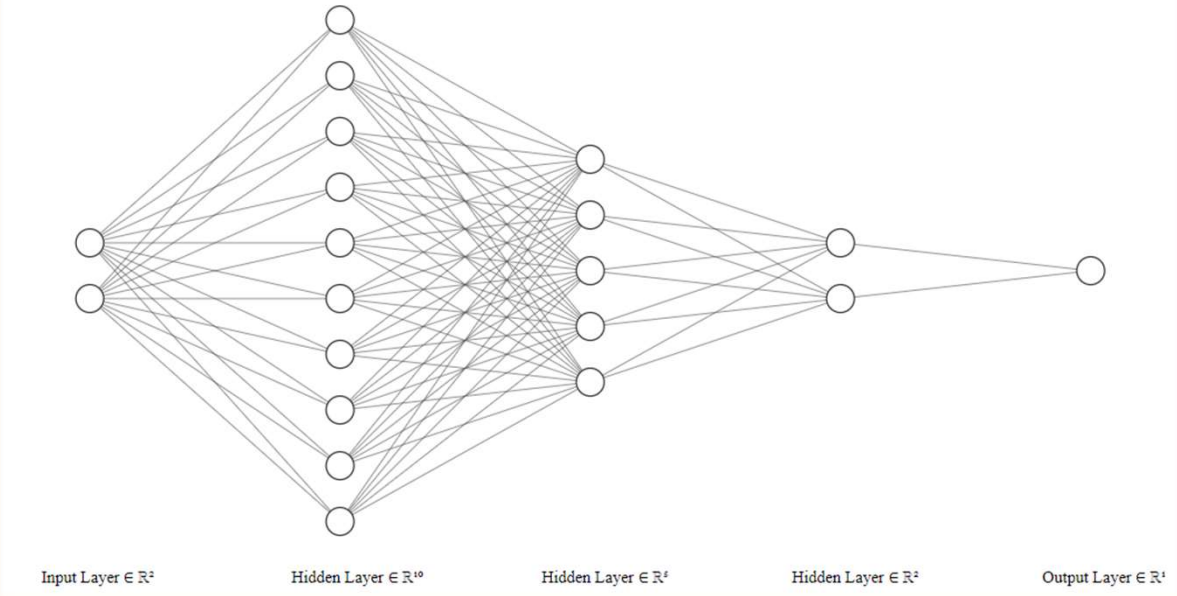
Weak Form





$$h1 = \mathrm{ReLU}(W1x + b1)$$
$$h2 = \mathrm{ReLU}(W2h2 + b1)$$
$$h3 = \mathrm{ReLU}(W3h2 + b1)$$
$$y = W4h3 + b4$$

Semi-Strong Form



$$h1 = \mathrm{ReLU}(W1x + b1)$$
$$h2 = \mathrm{ReLU}(W2h2 + b1)$$
$$h3 = \mathrm{ReLU}(W3h2 + b1)$$
$$y = W4h3 + b4$$

ARCH EQUATION

The estimated ARCH(1) equation will take the following form:

$$r_t = 39.48458 - 0.112107 r_{t-6} + \epsilon_t$$

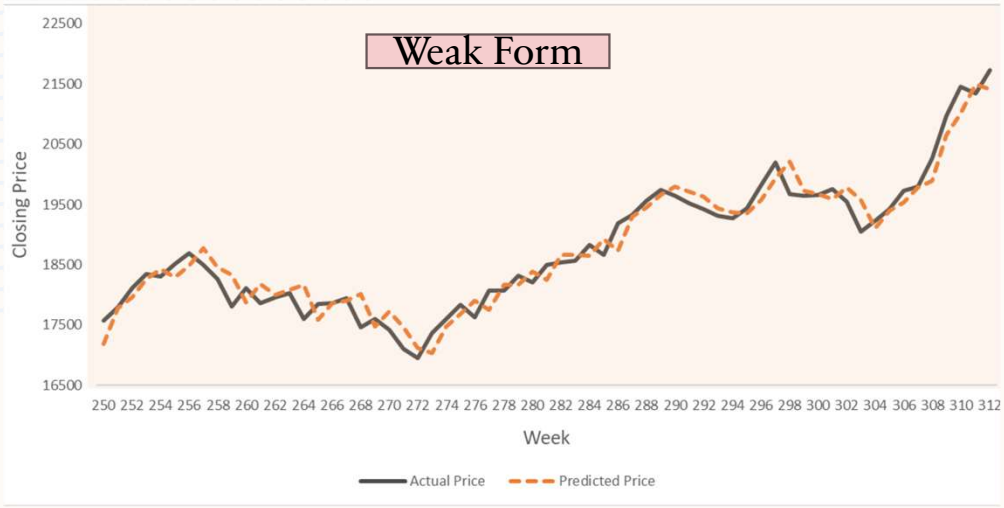
Variance

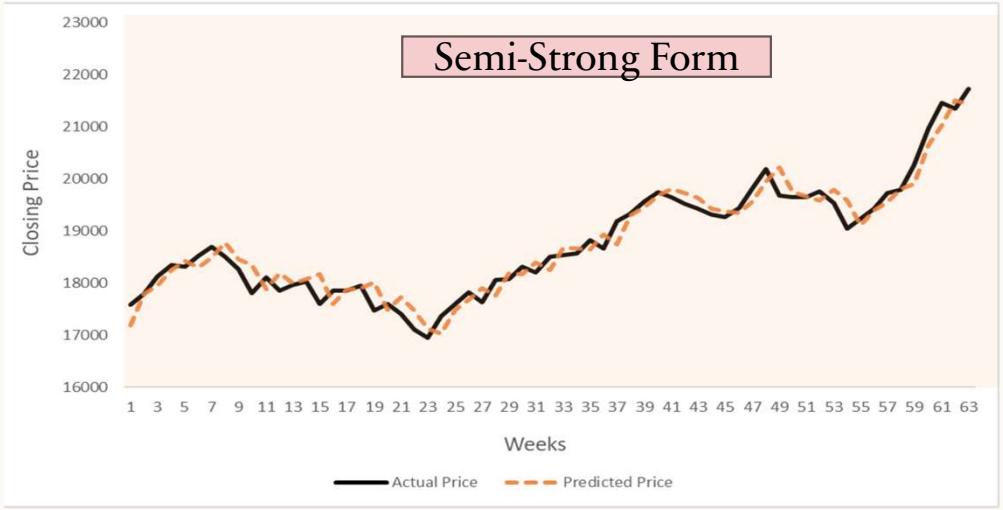
$$\sigma^2{}_t = 68485.08 + 0.171429 \alpha^2{}_{t-6}$$

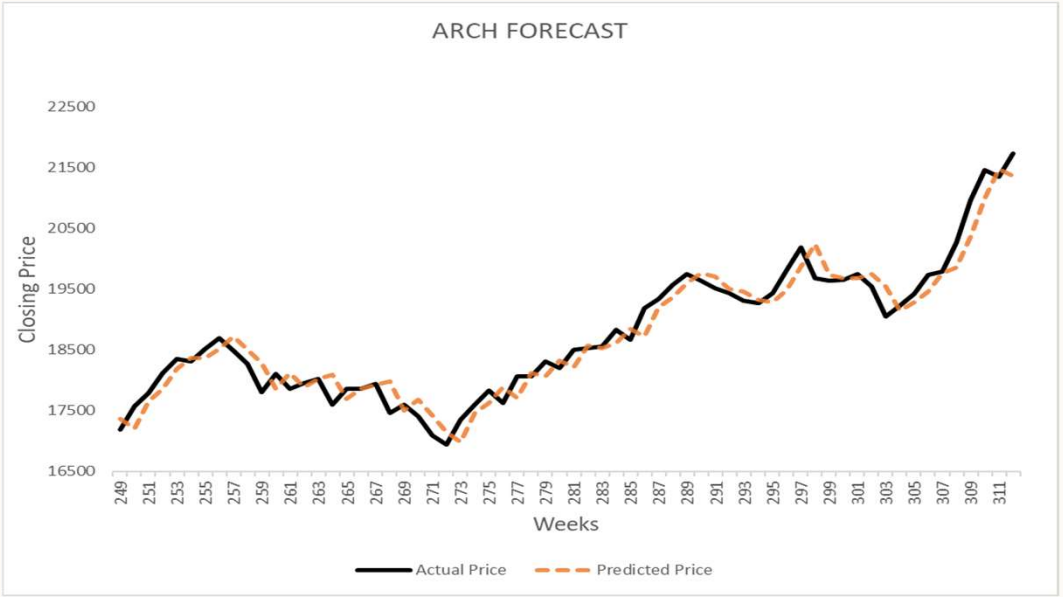
Where

r_t **is the returns at time t**

$\sigma^2{}_t$ **is the conditional variance at time t**


Weak Form
Closing Price
22500
21500
20500
19500
18500
17500
16500
250 252 254 256 258 260 262 264 266 268 270 272 274 276 278 280 282 284 286 288 290 292 294 296 298 300 302 304 306 308 310 312
Week
Actual Price
Predicted Price


Semi-Strong Form
Closing Price
23000
22000
21000
20000
19000
18000
17000
16000
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 61 63
Weeks
Actual Price
Predicted Price


ARCH FORECAST
Closing Price
22500
21500
20500
19500
18500
17500
16500
249 251 253 255 257 259 261 263 265 267 269 271 273 275 277 279 281 283 285 287 289 291 293 295 297 299 301 303 305 307 309 311
Weeks
Actual Price
Predicted Price

FINDINGS

Forecast Type	MAE	MAPE
FNN Weak Form	212.311	1.1353 %
FNN Semi-Strong Form	213.9576	1.1437 %
ARCH Forecast	217.468	1.15907%.

CONCLUSION

- **Leveraging the FNN it was found that based on the past prices it was possible for the neural network to outperform the random walk hypothesis**
- **The FNN model incorporating both factors exhibited superior performance compared to the EMH expectations, thus challenging the notion that all publicly available information is instantaneously**
- **The neural network model was successful in consistently outperforming the traditional ARCH model.**