Efficient Market Hypothesis in Contemporary Applications: A Systematic Review on Theoretical Models, Experimental Validation, And Practical Application

Chuer Xu
Hangzhou No.14 High School, Zhejiang, China
xuchuer2023@h14z.com

Abstract. Financial markets are fundamental to the stable development of the socioeconomic landscape. This study focuses on the efficient market hypothesis (EMH) positing that financial markets reflect all available information. The importance of studying EMH lies in assessing its feasibility in contemporary society and trying to make markets become "efficient". This paper reviews the theoretical models of EMH, explores statistical validation methods, and investigates leveraging machine learning for stock price prediction. Those findings suggest that while the feasibility of EMH in today's society may be reduced, its theoretical framework and derivative formulas still hold reference value. Furthermore, when utilizing machine learning for stock price prediction, it is essential to consider subjective human factors, such as irrational behavior. In conclusion, this research highlights the diminished feasibility of EMH in modern society, underscores the continued improvements of its theory and formulas, and emphasizes the significance of considering subjective human factors in utilizing machine learning for stock price prediction.

Keywords: Efficient Market Hypothesis, Statistical Validation, Machine Learning.

1. Introduction

Within the realms of economics and finance, the stock market holds significant importance, driving the attention and efforts of numerous economists and financial experts. It serves as a dynamic platform where buyers and sellers interact, leading to price discovery and allocation of resources. Understanding the workings of the market has been a central focus for researchers, as it provides insights into economic behavior, investment strategies, and overall financial stability. Efforts to study the market have resulted in various theories and hypotheses. One prominent theory is the efficient market hypothesis (EMH).

EMH is a fundamental concept in financial development that asserts it is difficult to consistently earn profit from the market by predicting the stock price. The importance of EMH lies in its implications for Investment Decision-making, Market Regulation, Risk Management, and Academic Research. In other words, studying market price trends is not only for profit-making, but also to enable issuers to better understand the overall situation, exert control over the long-term stability of the stock market, and ensure the security of the national economy.

This paper focuses on summarizing the theoretical models of EMH, the historical improvements, and criticisms put forth by economists; analyzing statistical validation methods of EMH; and providing examples of using machine learning for stock price prediction. The overall aim is to explore the feasibility of EMH theory in today's financial conditions.

2. Evolution of Theoretical Models

Going beyond a mere literature review, this paper classifies the foundation of EMH into two distinct categories and involves a comprehensive synthesis of the theoretical achievements of EMH, followed by an innovative exploration of subsequent improvements and challenges, as shown in Figure 1.
2.1. The Foundation of Efficient Market Hypothesis

In 1900, Louis Bachelier, a French mathematician, made a significant contribution to the development of the Efficient Market Hypothesis (EMH). Bachelier established the concept of "a random walk" [1] in stock prices and developed a mathematical framework to describe the behavior of stock prices based on a random walk process. Similar to Brownian motion, this theory proposed that stock prices exhibit a random walk pattern, characterized by the independence and unpredictability of future price changes from past price changes.

The formula is shown below, where $\mu$ is a drift constant, $\sigma$ is the standard deviation of the returns, $\Delta t$ is the change in time, and $Y_i$ is an i.i.d. random variable satisfying $Y_i \sim N(0,1)$.

$$S_{t+1} = S_t + \mu \Delta t S_t + \sigma \sqrt{\Delta t} S_t Y_i$$

Then, William Sharpe proposed the Capital Asset Pricing Model (CAPM) in his seminal paper [2], offering a framework to assess the suitable required return on an investment considering its risk level within a diversified portfolio. To wit, Shape introduced beta and systematic risk, the concept of market equilibrium, and Security Market Line (SML).

The formula is shown below, where $ER_i$ is expected return of investment, $R_f$ is risk-free rate, $\beta_i$ is beta oft he investment, and $(ER_m - R_f)$ is market risk premium.

$$ER_i = R_f + \beta_i(ER_m - R_f)$$

Bachelier's work on random walk laid the foundation for the understanding of stock price movements as a stochastic process, and Sharpe's CAPM model laid the foundation to analyze the correlation between risk and expected return in the context of market equilibrium, which can be summarized as informational efficiency.

2.2. Random Walk Model

After Bachelier's research on the random walk model, which was ahead of its time and provided a mathematical framework for understanding market behavior, was published, it gained agreement in Irving Fisher's 1930 book "The Theory of Interest"[3] and was further proved by Paul Samuelson's mathematical proof [4] in 1955. Moreover, on account of EMH and arbitrage opportunity unavailability, Fischer Black and Myron Scholes jointly developed the option pricing model known as the Black-Scholes model [5] in 1973. This model revolutionized the understanding and pricing of...
financial derivatives; provided a mathematical framework to calculate the option's fair valuation, taking into account variables like the current stock expense, strike price, expiration time, risk-free interest rate, and volatility; and allowed market participants to estimate the expected value of an option and make informed decisions about buying, selling, or hedging options.

However, the opposing views of random walk theory have emerged since the 1960s. Benoit Mandelbrot’s research on fractal geometry and its application to financial markets was brought out [6], suggesting that stock price movements exhibit fractal patterns and indicating there are repetitive and self-similar structures at different time scales. Additionally, the “Momentum Strategy”[7] concept proposed by Jegadeesh and Titman in 1993 presented empirical evidence supporting the existence of momentum in stock returns, implying that insights for forecasting future price movements can be offered by historical price patterns.

Later on, the "predictability" model was developed by Goyal and Welch. After examining the predictability of the equity premium, which investors expect to receive from investing in stocks compared to risk-free assets such as government bonds, researchers found that certain economic indicators have some predictive power for future equity premiums, contracting to the random walk theory.

The formulas are shown below.

\[ EPS = \frac{\text{Net Income} - \text{Dividend Payments}}{\text{Weighted Average Shares Outstanding}} \]  
\[ \text{Dividend Payout Ratio} = \frac{\text{Dividend Per Share (DPS)}}{\text{Earnings Per Share (EPS)}} \]

2.3. Market Efficiency and Informational Pricing

William Sharpe, John Lintner, and Jan Mossin boosted the development of CAPM during the 1960s. After Sharpe introduced a framework for pricing risky assets and determined the expected return on investment based on its systematic risk by beta, Lintner expanded on Sharpe’s work [9] and refined the CAPM by introducing the concept of dividend payout ratios and the relationship between dividends and stock prices. Jan Mossin contributed to the development of the CAPM with his 1966 paper "Equilibrium in a Capital Asset Market"[10], extending the CAPM to consider the impact of taxes and borrowing constraints on investment decisions. Together, Sharpe, Lintner, and Mossin laid the foundation for the CAPM, offering comprehension of risk-return association.

Nevertheless, several economists presented criticisms and improvements to CAPM, challenging certain assumptions of the CAPM and introducing alternative theories to explain asset pricing and market behavior. In 1968, for instance, Ray Ball and Philip Brown questioned the reliability of accounting measures in predicting stock returns [11].

Ten years later, Michael Jensen introduced "Jensen's alpha"[12], focusing on measuring the risk-adjusted performance of investment portfolios. The formula is shown below, where \( \bar{r}_p \) is expected total portfolio return, \( r_f \) is risk-free rate, \( \beta_p \) is Beta of the portfolio, and \( \bar{r}_m \) is expected market return

\[ \alpha_p = \bar{r}_p - [r_f + \beta_p(\bar{r}_m - r_f)] \]  

This measure aimed to assess the excess return earned by a portfolio beyond what would be expected based on its systematic risk. Unlike other criticisms, Jensen's alpha has a duality on EMH. On one hand, it supports the EMH by suggesting that excess returns should be close to zero, indicating that it is difficult to consistently outperform the market. On the other hand, it challenges the EMH by demonstrating that there are non-random factors or investment strategies that can generate excess returns. In essence, Jensen's alpha acknowledges the potential for market inefficiencies and the possibility of achieving superior performance through skillful investment strategies.
Since 2000, financial scholars have recognized the influence of investor sentiment and behavior on market prices. One notable contribution in this regard was Robert Shiller's work on "Behavioral Finance"[13], which introduced the concept of the "CAPE index."

$$CAPE\ Ratio = \frac{Current\ Market\ Price}{\text{Inflation\ Adjusted\ Earnings}_{10-Year\ Avg}}$$

This index measures the cyclically adjusted price-to-earnings ratio, taking into account psychological factors that impact asset pricing. By incorporating these psychological factors, Shiller's framework challenges the rationality assumptions of the CAPM. This shift towards behavioral finance has enriched the understanding of market phenomena and emphasized the need to consider psychological factors alongside traditional economic theories when analyzing asset prices and market behavior.

3. Statistical Validation and Empirical Testing

3.1. Historical Methods

Several influential models have been devoted to the understanding of asset pricing and market behavior. In 1984, Richard Roll introduced "Roll's Measure"[14], a systematic risk measure that captures the sensitivity of asset returns to common factors in the market. The following year, Rosenberg, Reid, and Lanstein proposed the "Arbitrage Pricing Theory (APT)"[15], suggested the determination role of multiple factors or risk sources in asset returns evaluation, and challenged the single variable in CAPM.

John Campbell's 1987 "Dividend-Price Ratio Model"[16] focused on the relationship between stock prices and dividends, indicating that higher dividend-price ratios can predict higher expected returns. Additionally, Chordia, Subrahmanyam, and Tong [17] explored the impact of "Investor Sentiment" and "Liquidity" on asset prices, investigating how factors such as investor sentiment and liquidity conditions can influence market prices.

As a far-reaching financial hypothesis, EMH's viability has been seen as cruciality. Scholars nowadays dwell on using statistical models to examine the applicability of EMH in light of changing market dynamics and evolving investor behavior.

3.2. Contemporary Validation of Viability

Brouty and Garcin [18] discuss the Hurst exponent, a commonly used statistical measure in finance, and state that predictions can be made using the fractional Brownian motion (fBm) model by determining the value of the Hurst exponent, leading to potential statistical arbitrage opportunities. However, the fBm model may not always accurately capture the changes in prices over time. Therefore, extensions of the fBm model have been proposed, such as eliminating the Gaussian distribution, introducing time-varying parameters deterministically or stochastically, and converting fBm into a stationary process. These extensions exhibit multifractal properties, quantified by the Hurst exponent, which may have an indication to the autocorrelation of the process. Consequently, the Hurst exponent may not be a suitable indicator of market efficiency in these cases, necessitating the use of alternative statistics with less reliance on specific models.

To address this issue, researchers have introduced the concept of Shannon entropy. As a measure of complexity in dynamical systems, Shannon entropy can be used to describe the complexity in the market, such as measuring liquidity. However, to make market efficiency more straightforward, the authors employ Risso's method, which studies the discrete distribution of successive price return phenomenon.

Due to the limited number of observations, estimates of this information might exist statistical errors. Instead, the researchers provide moments of the distribution of this estimate and a more comprehensive formula for its asymptotic distribution following a gamma law. They assess the
applicability of this asymptotic distribution through simulation studies and establish a statistical test for market efficiency.

In conclusion, the efficiency measure, referred to as market information, is determined by calculating the discrepancy between the entropy that aligns with the ideal EMH and the actual entropy observed in the market. Using the estimator conducted by the researchers, the market prices are evaluated and displayed. The notable result is that the experimenters cannot reject the null hypothesis of market efficiency for the Russell 2000 index whereas they can reject it for the CAC 40, Perkfect stock, and the cryptocurrency market with confidence levels lower than 99%. This indicates that there might be statistical arbitrage opportunities available for individual stocks within an index. However, when multiple stocks are combined in an index, these opportunities cease to exist at the index level.

3.3. Contemporary Challenges of Rationality in Certain Contexts

In the paper by Ammy-Driss and Garcin [19], they observe the disruptive impact of crisis events on financial markets. In this study, the memory parameter of an fLsm is employed as an efficiency indicator to measure the influence of COVID-19 on market efficiency. Similar to the previous paper, the researchers also discuss the Hurst exponent, and through empirical experiments, they find that the m of an fLsm is more effective in detecting market inefficiencies compared to the Hurst exponent H. Consequently, they provide a detailed explanation of dynamic alpha-stable distribution estimation and the discount factor select rule. They also describe the process of transitioning from fBm to fLsm, which allows for the filtration of m in H.

By utilizing the memory parameter m of an fLsm, the researchers observe the presence of an inefficiency period occurring predominantly at the onset of the crisis in various regions, including the USA (S&P 500, S&P 100), Europe (EURO STOXX 50, Euronext 100, DAX, CAC 40), Asia (Nikkei, KOSPI, SSE 180), and Australia (S&P/ASX 200). However, this inefficiency period is less pronounced for the Chinese and Australian indices.

4. Review of Computational Approaches

Machine learning (ML), a subset of Artificial Intelligence (AI), involves the process of learning from data to solve problems. To forecast stock prices, various machine-learning algorithms have been developed and utilized, showing promising results.

4.1. Prevalent Machine Learning Models

There are 11 prevalent ML models that are commonly used in practice. These models include Decision Tree, Random Forest, Adaptive Boosting (AdaBoost), eXtreme Gradient Boosting (XGBoost), Support Vector Classifier (SVC), Naïve Bayes, K-Nearest Neighbors (KNN), Logistic Regression, Artificial Neural Network (ANN), and another two deep learning models: Recurrent Neural Network (RNN), and Long short-term memory (LSTM)[20].

*Decision Tree* is a tree-like model that makes decisions based on input features, while *Random Forest* combines multiple decision trees to improve accuracy. *Adaboost* and *XGBoost* are ensemble learning algorithms that combine weak classifiers to create a strong classifier. *SVC* finds the best hyperplane to separate classes, *Naïve Bayes* is a probabilistic classifier, and *KNN* assigns labels based on the majority vote of its neighbors. *Logistic Regression* estimates the probability of binary classification. *ANN* is a computational model inspired by the brain, capable of learning complex patterns. These non-machine learning models rely on handcrafted features and perform well on tasks with relatively simpler patterns or smaller datasets and are generally interpretable and computationally efficient.

As for the two deep learning models, *RNN* processes sequential data using feedback connections, while *LSTM* addresses the vanishing gradient problem and captures long-term dependencies. These models learn hierarchical representations of data by progressively extracting more abstract features.
from raw inputs. Deep learning models can automatically learn features from raw data, eliminating the need for manual feature engineering.

ML offers a range of techniques for classification and prediction tasks, providing valuable understanding and performance in financial analysis.

4.2. Modern Technique: Chatgpt

In addition to traditional machine learning models used for predicting stock prices, some researchers have discovered the potential of large language models (LLMs) and tried to explore ChatGPT’s aptitude in evaluating stock market returns [21].

By utilizing the Center for Research in Security Prices (CRSP) daily returns, news headlines, and RavenPack datasets, this research explores the potential of LLMs in predicting stock market prices. The study establishes a strong foundation by examining the sentiment scores generated by ChatGPT and corresponding stock market returns. The integration with RavenPack ensures the use of relevant news for analysis. Prompts are employed to provide context and instructions. The empirical findings demonstrate the significant potential of ChatGPT as a tool to foretell stock market movements through sentiment analysis. In essence, this research elucidates the powerful potential of LLMs in stock price prediction and offers insights for further enhancing these models in finance.

4.3. Model Evaluation: Accuracy Matrices

To evaluate the accuracy of prediction models, various metrics are commonly used in the research. These include the mean absolute percentage error (MAPE), mean square error (MSE), mean absolute error (MAE), root mean squared error (RMSE), and correlation coefficient (R) [22,23,24,25].

R indicates the degree of association between predicted and actual values, with a higher value indicating better model performance.

$$ R = \frac{\sum_{i=1}^{n} (t_i - \bar{y})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (t_i - \bar{t})^2 \cdot \sum_{i=1}^{n} (y_i - \bar{y})^2}} $$  (7)

RMSE measures the residual between actual and predicted values.

$$ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (t_i - y_i)} $$  (8)

MAPE represents the average absolute percentage error, with lower values indicating better accuracy.

$$ MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{t_i - y_i}{t_i} \right| $$  (9)

MAE measures the average magnitude of errors without considering their direction.

$$ MAE = \frac{1}{n} \sum_{i=1}^{n} (t_i - y_i) $$  (10)

Apart from those specific metrics used to evaluate the accuracy and precision of prediction, volatility, momentum, accuracy, precision, recall, and F-score are the metrics that provide further insights into different aspects of analysis, including market behavior, trend strength, prediction accuracy, and the trade-off between precision and recall.

According to the results of the two research papers [20,21], the methods and evaluations are classified and listed below in Table 1.
Table 1. The Evaluation of Corresponding Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Decision Tree</th>
<th>Random Forest</th>
<th>AdaBoost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation</td>
<td>For Continuous Data: F-score: 0.681850 Accuracy: 0.663450</td>
<td>For Continuous Data: F-score: 0.731475 Accuracy: 0.722125</td>
<td>For Continuous Data: F-score: 0.731100 Accuracy: 0.714100</td>
</tr>
<tr>
<td></td>
<td>For Binary Data: F-score: 0.852425 Accuracy: 0.848725</td>
<td>For Binary Data: F-score: 0.859375 Accuracy: 0.854475</td>
<td>For Binary Data: F-score: 0.860425 Accuracy: 0.85545</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>XGBoost</td>
<td>SVC</td>
<td>Naïve Bayes</td>
</tr>
<tr>
<td>Evaluation</td>
<td>For Continuous Data: F-score: 0.723750 Accuracy: 0.708975</td>
<td>For Continuous Data: F-score: 0.739775 Accuracy: 0.718900</td>
<td>For Continuous Data: F-score: 0.731100 Accuracy: 0.714100</td>
</tr>
<tr>
<td></td>
<td>For Binary Data: F-score: 0.859650 Accuracy: 0.854800</td>
<td>For Binary Data: F-score: 0.867250 Accuracy: 0.865700</td>
<td>For Binary Data: F-score: 0.860425 Accuracy: 0.85545</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>KNN</td>
<td>Logistic Regression</td>
<td>ANN</td>
</tr>
<tr>
<td>Evaluation</td>
<td>For Continuous Data: F-score: 0.734625 Accuracy: 0.719250</td>
<td>For Continuous Data: F-score: 0.736775 Accuracy: 0.717325</td>
<td>For Continuous Data: F-score: 0.723750 Accuracy: 0.708975</td>
</tr>
<tr>
<td></td>
<td>For Binary Data: F-score: 0.865575 Accuracy: 0.859625</td>
<td>For Binary Data: F-score: 0.864375 Accuracy: 0.861875</td>
<td>For Binary Data: F-score: 0.731100 Accuracy: 0.714100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>RNN</td>
<td>LSTM</td>
<td>GPT</td>
</tr>
<tr>
<td>Evaluation</td>
<td>For Continuous Data: F-score: 0.857000 Accuracy: 0.846625</td>
<td>For Continuous Data: F-score: 0.859325 Accuracy: 0.849200</td>
<td>Accuracy: 0.511000 Precision: 0.509700 Recall: 0.940800 Specificity: 0.068600 F-Score: 0.6612</td>
</tr>
<tr>
<td></td>
<td>For Binary Data: F-score: 0.895750 Accuracy: 0.891600</td>
<td>For Binary Data: F-score: 0.893475 Accuracy: 0.889025</td>
<td></td>
</tr>
</tbody>
</table>

Naive-Bayes and Decision Tree models demonstrate lower accuracy (around 68%) compared to RNN and LSTM models, which perform exceptionally well (approximately 86%) with a significant difference from other models. However, the superior performance of RNN and LSTM comes at the cost of longer running times. Notably, deep learning methods (RNN and LSTM) exhibit strong predictive abilities for stock movement, particularly for continuous data, where traditional machine learning models struggle. While ChatGPT’s evaluation is not so high, it offers a novel approach: Future research can explore the integration of LLMs with diverse machine learning techniques and quantitative models, aiming to develop hybrid systems that capitalize on the respective strengths of these approaches.

5. Conclusion

EMH posits that financial markets are efficient in reflecting all available information, suggesting that it is impossible to consistently outperform the market through active trading or by exploiting mispriced assets. This hypothesis has sparked considerable debate and has been extensively examined by researchers seeking to understand the efficiency and predictability of markets. Under the continuous research of generations of economists, the original model of EMH has undergone
numerous derivatives. Financial behavioralists have also recognized the presence of irrational behavior in the market. The development of modern mathematics has facilitated the demonstration and validation of EMH theory in statistical analysis. Researchers assess the feasibility of EMH through hypothesis testing using confidence levels. With the advancement of computer science, ML has been increasingly applied to stock price prediction. However, simply inputting data into algorithms is insufficient. Current computer algorithm models are not yet perfect, and operators should consider the impact of socioeconomic events and news on investors. In further studies, financial experts should construct models that align with real-world conditions, integrating statistics, computer science, and subjective social factors.

References


