Proposal: Generating Passive Income For Young Adults Through Value Investing Machine Learning

Problem Statement and Motivation

In the onset of a declining job market and tumultuous economy, many young adults do not have a stable source of income to rely upon for long term life goals; thus people begin to seek forms of passive income. A recent survey by *Stash* declares that over 90% of young adults express interest in learning how to invest but lack the knowledge to do so. Most people do not take the initiative to learn to invest because of the perception that it is 'complicated' and the fear of losing money because of the volatile nature of the market.

To increase financial literacy among young adults and encourage people to engage in the stock market, the team has created a machine learning model that is able to receive user input on paper money value, desired return value and when they would like to receive it. The machine learning model will conduct a financial analysis of S&P 500 stock by accessing an API of a publicly available mutual fund index. The goal is to ultimately target a stock at its lowest value and sell at its highest generating revenue for the client.

Related Works

An attractive quality of the stock market is its volatility - high risks yield high rewards yet it is the very thing that deters many prospective investors. Investors yearn for predictability, reassurance that their funds will not go to waste so they seek machine learning tools to increase confidence in their financial forecasts. Fortunately, there are many established bodies of work that utilize value investing methods to increase investor confidence.

Machine learning models specialize in making predictions based on historical data. One team explored a quantitative investment strategy that tracks stock volatility through a combination of time series and machine learning models. This approach tracks the stocks momentum including features such as on balance volume, average convergence and divergence to forecast upcoming changes to the stock value; this experiment achieves an annual return of 5 to 10% [1].

Another research group constructs a machine learning model to capture stock mispricings based on value investing principles. It calculates the likelihood of a mispricing based on a bond credit spread compensates for its risks. [2] It helps researchers calculate the stock's intrinsic value and inform users whether the stock is worth investing in.

Machine learning models can be programmed to make out of sample financial forecasts due to its ability to uncover new patterns from economical predictors. It is able to capture non linear relationships within historical data which gives it the ability to make accurate inferences.[3] Based on current research, our team has deduced the most profitable method to make financial forecasts is to combine traditional value investing techniques with state of the art machine learning models.

Hypothesis

- **Hypothesis 1:** A portfolio that uses the value investing metrics: Price to Earnings (P/E), Price to Book (P/B), Debt to Equity (D/E), and Free Cash Flow (FCF) will out preform both the S&P 500 and the Russell 1000 Value Index over the same period of time. Due to value investing focusing strictly on companies that are financially strong yet undervalued, based on these ratios, the portfolio should generate higher and less risky returns compared to the overall market.
- **Hypothesis 2:** By combining both traditional value based investing strategies with machine learning, the system will be able to determine relationships and patterns within financial data that a simple rule based approach could not. This will ensure that the model can make more accurate predictions/suggestions to users when it comes to finding undervalued stocks, which in turn, leads to better investments.

Goals

1. Create a data pipeline that collects (and cleans) both data used to calculate our metrics (P/E, P/B, D/E, FCF) and stock prices for the S&P 500 companies. Preferably primarily using APIs like Yahoo Finance and Alpha Vantage.
2. Create a baseline model that uses traditional value investing metrics to find and rank undervalued stocks. This model will then build a portfolio and it will be compared to the S&P 500 and Russell 1000 Value Index to figure out if it can outperform the market using these ratios.
3. Integrate a ML model that builds on the baseline approach by looking at historical data to find patterns/trends between metrics. The goal is for the model to find these patterns and make smarter, data driven predictions about which undervalued stocks are most likely going to outperform the S&P 500 and Russell 1000 Value Index over time.

Proposed Solutions

We propose a machine learning driven value investing framework that integrates financial data analysis with traditional investing metrics to suggest undervalued stocks to a user. The system will collect and analyze key metrics such as P/E, P/B, D/E, and FCF, in addition to balance sheet and market data from APIs like Yahoo Finance, Alpha Vantage, and FRED. These features are then used to train an ML model that is capable of finding hidden patterns and relationships that traditional methods might miss. A decision module will then evaluate the portfolio performance over time, comparing it against two benchmarks (S&P 500 and Russell Value 1000 index). This combined approach ensures that both financial logic and machine learning insights contribute to smarter, data driven investment recommendations for beginner investors.

Evaluations

1. **Accuracy:** Measure how often the model correctly predicts high-return investment opportunities and accurately identifies potential financial risks

2. **Time Efficiency:** Track how long it takes for the system to collect, process, and output investment recommendations
3. **Statistic Validation:** Compare the results of the machine learning model with the baseline traditional value-investing model using statistical tests to evaluate improvements in performance
4. **User Experience:** Ensure that users are given appropriate information that supports their journey in obtaining another avenue to passive income

References

- [1] Yan, K., & Li, Y. (2024). Machine learning-based analysis of Volatility Quantitative Investment Strategies for American Financial Stocks. *Quantitative Finance and Economics*, *8*(2), 364–386. https://doi.org/10.3934/qfe.2024014
- [2]Kirkpatrick, T., & LaGrange, C. (n.d.). *Robotic Surgery: Risks vs. Rewards*. PSNet. https://psnet.ahrq.gov/web-mm/robotic-surgery-risks-vs-rewards#:~:text=RAS%20shares%20the%20same%20risks,pelvis%20adequately%20to%20perform%20RALP.
- [3] Cao, K., & You, H. (2022, September 28). *Fundamental Analysis Via Machine Learning: Financial Analysts Journal: Vol 80, no 2*. Taylor & Francis Online. https://www.tandfonline.com/doi/abs/10.1080/0015198X.2024.2313692