Report as of 22 Jun 2024

Introduction

Global financial markets are currently experiencing a period of transition. After a period of rising interest rates to combat inflation, some central banks are now starting to ease their policies. Earlier this week, the European Central Bank (ECB) took the significant step of cutting its benchmark interest rates by 25 basis points, signaling a potential shift towards a more accommodative monetary stance. In the US, the Federal Reserve is also facing calls for rate cuts, although they haven't announced any changes yet. This mixed picture, with some central banks easing and others remaining cautious, reflects the complex economic environment in discussion.
For Singapore investors who are highly responsive to global market conditions, these recent trends in interest rates will certainly alter their investment decisions on financial products. As interest rates fluctuate, the attractiveness of different investment options changes. Taking this into consideration, I want to take a deeper dive into understanding the implications these pose on one of the safest investment options in Singapore, the T-Bills.

Nature of T-Bills

Traditionally touted as an inflation hedge, the effectiveness of T-bills seems to depend heavily on the specific economic climate. Being backed by the full faith and credit of the government also makes them very low-risk and suitable for investors who prioritize safety over high returns.

Problem Statement

With that in mind, I set off to figure out whether the characteristics of the financial product could be accurately forecasted.

This will be executed through two means:
1. Tree-based models: These models are highly interpretable, and help users understand the factors influencing their predictions. Additionally, they handle complex relationships between variables, a key factor in financial markets. For this setup, a separate model is created to forecast each feature using information of every other feature, with a final output combining all the forecasted values together.
2. GPT-based models: These large language models are known for their prowess in pattern recognition and capturing intricate relationships within data. Using the same dataset developed for use in the tree-based models, combined with the power of prompt engineering, I seek to test if these models could prove valuable in analyzing financial product characteristics based on historical data and market trends.

Data Acquisition

A key breakthrough in developing this project stemmed from the ease of acquisition of the required data. As a matter of fact, the required data could be easily retrieved from the Monetary Authority of Singapore (MAS) website under the “Information for Individuals” section. From there, I was able to procure all T-bill auction statistics over the past 3 years, across the different forms, primarily 6-month and 12-month ones.

Data Cleaning

For the purpose of this project, I've chosen to focus solely on the outcome of 6-month T-bill auctions, and thus filtered the dataset as such. Given the standardised nature of the data and the absence of null values in the dataset, minimal data cleaning was required, and I was effectively able to use the dataset as-is.

Data Pre-Processing

The training data for the model was designed to be straightforward. We aimed to predict the target feature value for the upcoming fourth exercise (T-4) using data from the past three exercises (T-1 to T-3). To achieve this, a windowing function was used to extract consecutive sets of three exercises. These sets were then flattened into a single feature vector (X) and combined with the target value (y) for the following exercise (T-4). This process was repeated for every sequence of three consecutive periods found in the dataset.

	Auction Date	Issue Date	Maturity Date	Auction Tenor (Month)	Auction Amount (S$M)	Amount Applied (S$M)	Bid-to-Cover Ratio	Cut-off Yield (%)	Cut-off Price	Median Yield (%)	Median Price	Average Yield (%)	Average Price
0	24 Jun 2021	29 Jun 2021	28 Dec 2021	6	3900	7503.392	1.92	0.33	99.835	0.28	99.860	0.24	99.880
1	08 Jul 2021	13 Jul 2021	11 Jan 2022	6	4000	9095.272	2.27	0.32	99.840	0.26	99.870	0.22	99.890
3	22 Jul 2021	27 Jul 2021	26 Jul 2022	12	3700	7667.723	2.07	0.35	99.651	0.28	99.721	0.23	99.771

Figure 1: Initial Representation of Dataset

Median Yield (%) Growth1	Median Yield (%) Growth2	Median Yield (%) Growth3	Median Price Growth1	Median Price Growth2	Median Price Growth3	Average Yield (%) Growth1	Average Yield (%) Growth2	Average Yield (%) Growth3	Average Price Growth
0.000000	-0.071429	0.076923	0.000000	0.000100	-0.000100	0.000000	-0.083333	0.045455	-0.000100
-0.035714	0.000000	0.000000	0.000050	0.000000	0.000000	0.086957	-0.080000	-0.043478	-0.000100
0.000000	0.111111	0.066667	0.000000	-0.000150	-0.000100	0.090909	0.208333	0.000000	-0.000150
0.250000	0.075000	0.139535	-0.000391	-0.000150	-0.000301	0.103448	-0.062500	0.433333	0.000050
0.020408	-0.160000	0.190476	-0.000050	0.000401	-0.000401	-0.023256	-0.142857	0.222222	0.000301

Figure 2: State of processed dataset

Tree-Based Model

The tree model of choice for this project was the XGBoost (XGB) Regressor model. The model was trained on 18 features, and its performance was assessed using mean squared error (MSE) and mean absolute percentage error (MAPE). While errors were consistent across features, yield growth predictions showed a higher MAPE exceeding 2%.

Despite the limited dataset size due to the exclusion of overlapping data points, the model's predictions were reasonably close to the latest auction statistics. However, the model predicted a yield growth, whereas the actual outcome was a loss of about 4 basis points. This serves as a baseline for comparison with a GPT-based model discussed later.

Given the limitations of the dataset size, the XGBoost model's performance met expectations and serves as a valuable baseline for comparison with the forthcoming GPT-based model.

GPT-Based Model

The GPT model is the main model designed in this project, with the objective of testing its ability to comprehend current market conditions and make a sound judgment based off this set of information, in combination with the historical dataset of previous exercises.

With GPT's ability to comprehend textual information and perform sound data analysis on the back of it, I sought to tap on these capabilities in the model execution. Yet, GPT's 4-o engine, though relatively more current than previous models, lacked the ability to produce real-time information. i.e., It was unable to provide a statement detailing current market conditions and significant changes to the financial markets in the past week. As such, I turned to Gemini to retrieve this information, verifying the accuracy of its response, before subsequently feeding it into GPT.

With the gathered information, along with the dataset, I queried GPT to provide me with forecasted values for each of the 18 features predicted by the XGB Regressor earlier, as well as a rationale behind its predictions. In all fairness, the model was able to produce remarkably stellar results, unveiling impressive values and reasoning. While deviations were inherent, the degree of each deviation was far less than I had expected, against the XGB model.

Certainly, the inclusion of current market conditions provided GPT an advantage over XGB, but its ability to process this information and formulate its predictions and rationale under this context demonstrates the sheer power it possesses when used appropriately. Consistent with results and conclusions drawn from other projects, it has become increasingly imminent that GPT has the capabilities to harness information and perform in-depth analysis where necessary, but users must themselves be proficient in the knowledge of these models and prompt engineering, which is paramount in determining the quality of response returned from the model.

To conclude, I believe GPT's results have demonstrated a robust understanding of the problem and its interpretation has produced impressive results, which I'd be certain to incorporate for individual use. Moving forward, I will be looking to test its abilities on other larger-scale problems that would challenge its robustness when processing larger datasets.The GPT model is the main model designed in this project, with the objective of testing its ability to comprehend current market conditions and make a sound judgment based off this set of information, in combination with the historical dataset of previous exercises.

Conclusion

Certainly, the inclusion of current market conditions provided GPT an advantage over XGB, but its ability to process this information and formulate its predictions and rationale under this context demonstrates the sheer power it possesses when used appropriately. Consistent with results and conclusions drawn from other projects, it has become increasingly imminent that GPT has the capabilities to harness information and perform in-depth analysis where necessary, but users must themselves be proficient in the knowledge of these models and prompt engineering, which is paramount in determining the quality of response returned from the model.

To conclude, I believe GPT’s results have demonstrated a robust understanding of the problem and its interpretation has produced impressive results, which I'd be certain to incorporate for individual use. Moving forward, I will be looking to test its abilities on other larger-scale problems that would challenge its robustness when processing larger datasets.

Considerations

As auction amounts are posted a week prior to the auction date, this definitive statistic does not require prediction, and can instead be fed into the models as ground truth, potentially improving the accuracy of predicting other features.