Predictive Modeling of Interest Rates Using Neural Networks

Technical Report

December 5th, 2023

Andrew Kroening, Jeremy Tan, Katherine Tian, Adler Viton

®

Executive Summary

In the dynamic landscape of financial markets, interest rates hold significant weight in monetary policies, investments, and various economic activities. Interest rates also measure economic health, reflected by their changes based on market sentiment, inflation, and monetary policies. From a commercial banking perspective, predicting interest rates helps institutions manage their positions and mitigate interest rate risk. In short, reliably forecasting interest rates could be a decisive advantage to any entity capable of doing so. This analysis investigates how a neural network regressor model performs when forecasting interest rates two decades into the future. Four experiments are conducted using various techniques, and the strength of each model is assessed by a combination of past metrics and observed viability of the predictions. While the models show some promise at making accurate predictions, further research is needed to enhance their forward-looking capabilities and provide a solid explanation for the performance range observed.

Table of Contents

Introduction

Forecasting interest rates is an integral part of a functioning economy, as policymakers use these rates as a lever to implement fiscal and monetary policy. Consequently, leveraging interest rate predictions helps these policymakers anticipate interconnected variables such as inflation and macroeconomic activities. Although the Efficient Market hypothesis suggests that forecasting interest rates may not be beneficial, the claim is disputed.[1] Forecasting has shown to be valuable, especially when the market has moments of inefficiency and mispricing. Interest rates also have temporal dependencies, where past interest rates and time play a role in their current values.[2]

This study analyzes the approach of predicting interest rates via neural networks. Prior literature has discussed methods of building linear models for forecasting, but they fail dramatically over time due to interest rate data often exhibiting non-linear patterns. Neural networks offer an alternative approach because they can capture these non-linear patterns, generalize patterns from historical data, and learn relevant features. However, there are concerns with the neural network's ability to capture the time-series nature of the data adequately, and predicting future rates mainly from past rates may not provide enough signal to the models to generate viable predictions.

We focus on how well an Artificial Neural Network (ANN), specifically a type of ANN called a Multi-Layered Perceptron (MLP) Regressor, can forecast interest rates. ANNs are known to be adaptable during their learning phase and can be suitable for time-series predictions.[3] Previous studies suggest its performance is a good start for using historical interest rate data to predict interest rates.[4] Projecting into the future, particularly when the model runs out of historical data to reference, is a key focus of this analysis. The study builds an MLP Regressor, tunes its hyperparameters, and trains the model on historical data. After assessing the model's performance, the model is then used to predict up to twenty years in the future. We then compare the forecasts and offer an interpretation of the outcomes and methods that might benefit this line of research in the future.

This research study is organized into a literature review, methodology, results, and conclusion sections. It aims to address two questions:

1. Is it possible to produce a realistic prediction for the interest rate yield curve into the future based on historical data?
2. Does the addition of other economic data positively influence the predictions?

[1] Malkiel, B. G. (2003). The efficient market hypothesis and its critics. Journal of Economic Perspectives, 17(1), 59–82. https://doi.org/10.1257/089533003321164958.
[2] Cajueiro, D. O., & Tabak, B. M. (2007). Time-varying long-range dependence in US interest rates. Chaos, Solitons & Fractals, 34(2), 360-367. https://doi.org/10.1016/j.chaos.2006.04.012.
[3] Sako, K., Mpinda, B. N., & Rodrigues, P. C. (2022, May 7). Neural Networks for Financial Time Series Forecasting. Entropy (Basel), 24(5), 657. https://doi.org/10.3390/e24050657.
[4] Yasir, M., Afzal, S., Latif, K., Chaudhary, G. M., Malik, N. Y., Shahzad, F., & Song, O-y. (2020). An Efficient Deep Learning Based Model to Predict Interest Rate Using Twitter Sentiment. Sustainability, 12(4), 1660. https://doi.org/10.3390/su12041660.

Literature Review

Prior work has investigated different approaches to predicting interest rates. To start, Yasir et al. compared the effectiveness of three modeling approaches.[5] They compared a fairly common linear regression modeling approach with two much more advanced techniques: a support vector regression and a convolutional neural network. Using only historical interest rate and exchange rate data, the authors demonstrated how the advanced convolutional neural network outperformed the other methods when training with data from January 2010 to October 2019 in four countries. Moreover, the convolutional neural network's ability to capture complex relationships is shown using social media for sentiment analysis as an unconventional indicator, boosting its prediction capabilities when included.[6] However, Yasir does not extend the analysis into the future, as doing so would require future social media sentiment data, which would have to be simulated.

In another paper, Politof and Ulmer discuss the efficacy of an ANN versus its simpler linear regression counterparts.[7] By examining the interest rates of Canadian treasury bills from January 1979 to December 1994, the ANN outperforms various multiple linear regression models with just a single hidden layer, a uniquely simple and shallow architecture for a potentially complex approach. The authors make several good observations from the results. Politof and Ulmer aptly describe the ANN as a black box, where we are given no insights into the relationships found during learning. This property is one of the trade-offs accepted when selecting that approach: the power of the model and its ability to detect hidden relationships is a consequence of the model's inability to provide an output that explains what those features are. The authors also note that the insights of their simpler multiple linear regression models are more interpretable but may be inaccurate due to the model's inability to handle non-linear relationships.[8] This tradeoff between predictability and interpretability is significant, particularly when understanding the rationale behind predictions is essential for decision-making in financial markets.

In the existing works, the interest rate data does not span more than 15 years. Yasir et al. compensates for this by adding other exchange rate data and social media sentiment. These additional features add depth to the dataset, but do not allow for the longitudinal study we seek to conduct here. Politof and Ulmer conducted their research before recent breakthroughs were made in the domain of neural networks, such as work in the interpretability of results and attention mechanisms. These limitations underscore the need for more comprehensive research, which our study aims to address by extending the temporal scope and incorporating newer advancements.

[5] Yasir, M., Afzal, S., Latif, K., Chaudhary, G. M., Malik, N. Y., Shahzad, F., & Song, O-y. (2020). An Efficient Deep Learning Based Model to Predict Interest Rate Using Twitter Sentiment. Sustainability, 12(4), 1660. https://doi.org/10.3390/su12041660.
[6] Ibid.
[7] Politof, T., & Ulmer, D. (1998). Predicting Interest Rates Using Artificial Neural Networks. In C. Zopounidis (Ed.), Operational Tools in the Management of Financial Risks (pp.291-305). Springer. https://doi.org/10.1007/978-1-4615-5495-0_17.
[8] Ibid.

Methodology

Data Collection

The data used in this research study was downloaded directly from the U.S. Department of the Treasury and includes the Daily Treasury Par Yield Curve Rates from January 3rd, 1990, through October 31st, 2023.[9] This three-decade range provides ample historical data from which the models can learn. Additionally, it permits the research to explore the optimal range for a lookback or reference window when training models. Since datasets with sentiment data before 2000 are sparse, and those that exist may not apply to our prediction aims, our study uses S&P 500 index data as another economic indicator.[10] This inclusion compensates for the lack of sentiment analysis data and introduces an additional layer of economic context to augment the MLP regressor's learning capabilities.

Cleaning Data

To prepare the data for experimentation, we clean the data to drop any unknown values and remove some time horizons where the Department of the Treasury data was not accounting for specific interest rates, such as the 1-month and 2-month yields, which are not consistently tracked in the earlier years of the dataset. With the interest rate data prepared, we then turned our attention to preparing the S&P 500 index data. Due to the nature of the MLP regressor, it is common practice to scale the S&P 500 index data by applying a logarithmic transformation to reduce the volatility in the data and ensure a consistent representation in the inputs for training.[11] We then merge the S&P 500 index data with our cleaned interest rate data for its inclusion as an economic indicator for our experiments. The table below shows the final dataset used in the experiments.

---

[9] U.S. Department of the Treasury. "Daily Treasury Par Yield Curve Rates." *Resource Center*, 2023, https://home.treasury.gov/resource-center/data-chart-center/interest-rates/TextView?type=daily_treasury_yield_curve&field_tdr_date_value=all. Accessed 31 October 2023.

[10] S&P Dow Jones Indices. (n.d.). S&P 500®. S&P Global. https://www.spglobal.com/spdji/en/indices/equity/sp-500.

[11] Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *International Conference on Artificial Intelligence and Statistics*.

Table 1. Merged Final Dataset

Date	3M	6M	1Y	2Y	3Y	5Y	7Y	10Y	open	high	low	close
Key	**Treasury features (values as float numbers, e.g. 4.28)**								**S&P 500 features**			

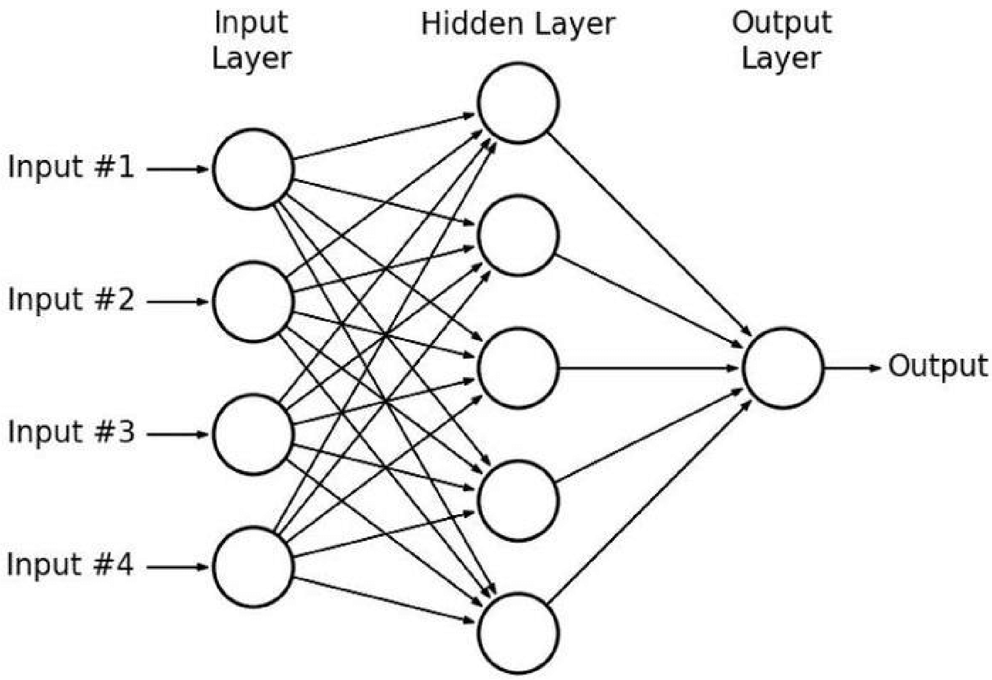
- The **Date** feature is the key value used to merge the datasets. It consists of one observation per trading day accounted for in the **Treasury** dataset.
- The **Treasury** features begin as raw values imported from the Department of the Treasury dataset. Experiments 1, 2, and 3 use only this data with transformations to investigate additional relationships. These are daily percent change (Experiment 2) and daily absolute change (Experiment 3).
- The **S&P 500** features are the daily open, high, low, and close for the S&P 500 with a logarithmic transformation applied. These values are only included in the dataset for Experiment 4.

Neural Network (MLP Regressor)

We use a Multi-Layer Perceptron Regressor from the Scikit-learn library as the workhorse for the experiments. We made this choice because the advanced ability of the neural network to learn non-linear relationships in the data may be helpful when learning from data that spans several days to several years in depth. The neural network starts with an input layer where we supply the historical values, transitions through hidden layers where mathematical functions are applied, and then produces results at an output layer. Each layer of the network has nodes and associated weights that dictate the flow of information through the layers. Each layer the inputs pass through helps the neural network expand its feature space through linear transformations. At the end of the training, or learning phase, the layers help the MLP Regressor capture complex relationships within the data to make informed predictions.

The figure below is a sample diagram of a simple neural network with four input nodes, one hidden layer of five nodes, and a single node in the output layer.

Fig 1. MLP Regression Visual Diagram[12]



The MLP Regressor used in the experiments conducted here was substantially more extensive in the number of hidden layers and the number of nodes in each layer used for the experiments. An additional noteworthy distinction is that the models used in our experiments also have eight or twelve nodes on the output layer, corresponding to whether the model attempts to predict interest rates using just previous interest rate data or if the S&P 500 data is included.

Modeling Data

The experimental design explores the application of the MLP Regressor in forecasting interest rates. Our study proposes a series of distinct experiments that vary based on the input data used. Because of the construct of the neural network, and as a consequence of our aim to predict interest rates up to 20 years into the future, the model will attempt to predict the daily output for the features supplied to it. In other words, if we provide only interest rates to the model, it will try to predict only the interest rates. When we add the S&P 500 data, it will also forecast the S&P 500 values. This is a consequence of our experimental design, a factor we discuss later in depth in the results section. The specific experiments are:

1. Predict daily interest rate absolute values from historical interest rate absolute values.

2. Predict daily interest rate percent change from historical daily interest rate percent changes.

[12] Data Science Central. (n.d.). The Mathematics of Data Science: Understanding the Foundations. Data Science Central. https://www.datasciencecentral.com/the-mathematics-of-data-science-understanding-the-foundations-of.

3. Predict interest rate daily change from historical interest rate daily changes.
4. Predict daily interest rate absolute value, and consequently S&P 500 index values, from historical interest rate absolute values with S&P 500 index data.

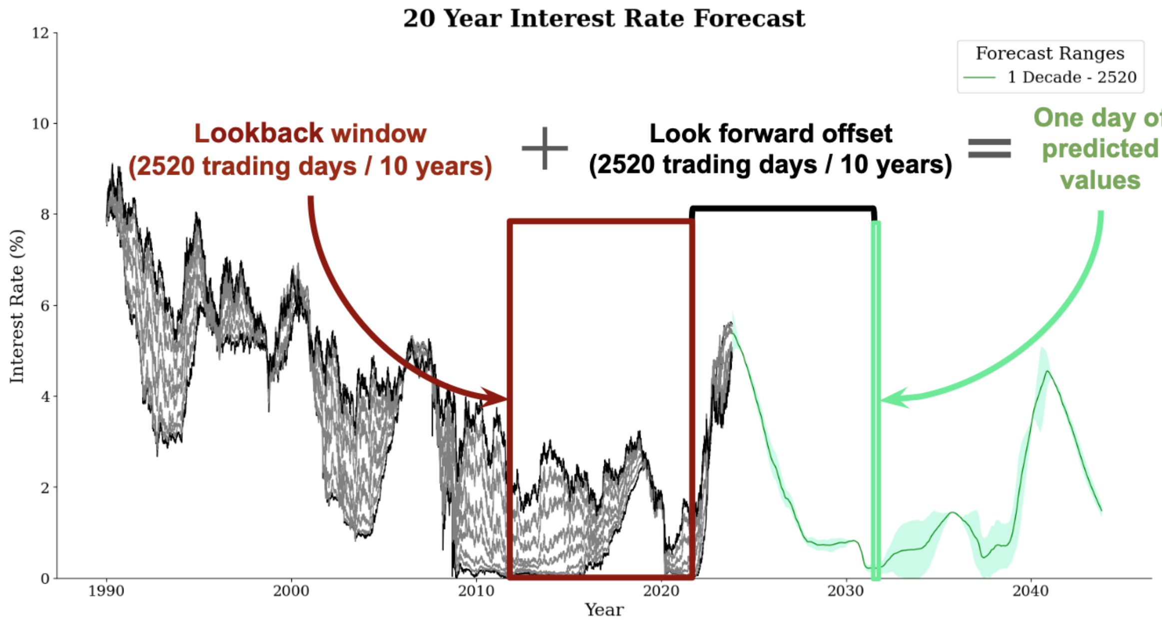
The initial experiment uses the raw absolute values of the interest rate data as inputs to understand what the MLP Regressor can learn without any applied transformations. In the second experiment, we calculate the daily percentage-wise variations in interest rates. By employing percent change, the aim is to capture the relative shifts in interest rates and understand how the MLP Regressor grasps these trends. The trade-off in this experiment is interpretability, which we cover in more detail when discussing this experiment's results below. In the third experiment, daily changes in interest rates are calculated to capture the short-term fluctuations in interest rate movements. Finally, we use the scaled S&P 500 index data as an additional economic indicator with raw interest rate values to evaluate if its predictive capabilities improve.

While performing the four experiments, we employ grid search, a systematic hyperparameter tuning technique, to explore an exhaustive set of combinations for the two values we seek to explore in this analysis:

- The *lookback range* is the number of trading days in the past that we use to predict the next day's interest rate yield curve
- The *hidden layer sizes* within the MLP Regressor are experimented with to determine if there are constructs that are more or less reactive to historical data and also more or less volatile in their predictions.

The final value we specify is the *look-forward range,* or the number of trading days into the future, which we offset our model to predict. The *look-forward range* is defined before each experiment. In each experiment, we utilize four horizons for this value: 1-day, 1-month, 1-year, and 1-decade. The figure below illustrates this concept, with the gray lines representing the historical data and the green line representing a sample prediction. The lookback and look-forward windows essentially "slide" in tandem from left to right through the dataset while training and then do the same when making predictions for the future values beginning on November 1st, 2023.

Fig 2. Experimental Design Visualization



Train-test backtesting, where the dataset is split into training and testing sets, was used to select the best combination of lookback range and hidden layer sizes for each prediction horizon in Experiments 1 and 4. In assessing the model, we examine two key performance metrics: Accuracy and Mean Squared Error (MSE). Accuracy provides an overall measure of the proportion of variance in the outcome interest rates that the historical data can explain. MSE calculates the average squared difference between predicted and actual interest rate values, providing insight into the precision of the model's predictions.

To provide a more granular analysis of the performance of the models across the various dataset transformations in Experiments 1, 2, and 3, we exhaustively searched the hyperparameter space in Experiment 1 for each of the four time horizons and then made predictions. After assessing those predictions, the team elected to use a consistent *hidden layer size* for future experiments to make comparison of relative performance easier to interpret. This allowed us to control for the neural network's performance somewhat while providing greater insight into the strengths or limitations of specific range combinations for *lookback* and *look-forward.*

After finalizing our models for the four experiments, the final phase involves using the models created to predict forward into the future by two decades. This forward prediction step evaluates each model's forecasting capabilities over an extended period. Because only a single month of "future" data has become available since the end of our dataset, we have limited ability to assess the performance quantifiably. However, we rate the performance of the models relative to each other and provide insight into which models are not feasible for future use and which offer potentially valuable insights into economic cycles, even if the fluctuations are probably out of the scope of realistic expectations.

Finally, to validate the interest data generated, we subject it to Augmented Dickey-Fuller (ADF) tests to assess non-stationary patterns in the data commonly observed in real-world interest rate data.[13]

[13] Rapach, D. E., & Weber, C. E. (2004). Are real interest rates really nonstationary? New evidence from tests with good size and power. Journal of Macroeconomics, 26(3), 409-430. https://doi.org/10.1016/j.jmacro.2003.03.001.

The ADF test, utilizing the Statsmodel TSA package, is designed to detect the presence of a unit root in a time series. A unit root suggests that the time series has a stochastic trend, indicating that the statistical properties of the data change over time and are non-stationary.

Results

Experiment 1 - Absolute Values

The first experiment conducted evaluated the ability of the neural network to learn from historical interest rate data and validated its ability to make predictions against known data. Once models with appropriate performance metrics were found, we extended that model twenty years into the future by conducting a simulation where the model predicted one day at a time and added its new prediction to the dataset.

This experiment's lookback ranges and hidden layer architecture are carried forward into the following two experiments. The complete grid search results are available in Appendix A for reference. The key hyperparameters chosen to take forward are:

- Hidden Layer Size: 10 Hidden Layers of 200 Nodes each
- 1-Day Forecast: 100-day Lookback
- 1-Month Forecast: 100-day Lookback
- 1-Month Forecast: 504-day Lookback (2 years)
- 1-Year Forecast: 756-day Lookback (3 years)
- 1-Decade Forecast: 1008-day Lookback (4 years)
- 1-Decade Forecast: 2520-day Lookback (10 years)

Below are figures depicting the 20-year forecast for two sets of predictions. The lines represent the average value of the interest rate curve for a given day, and the shaded area is the dispersion of the daily curve.

Fig 3. Experiment 1 Forecast A

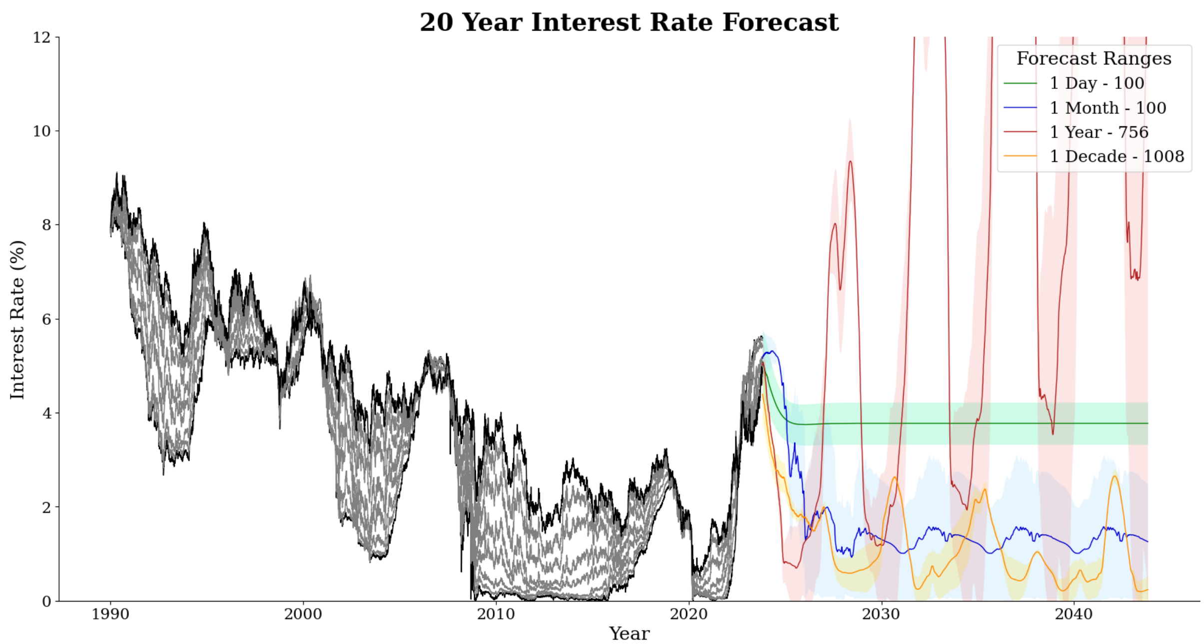
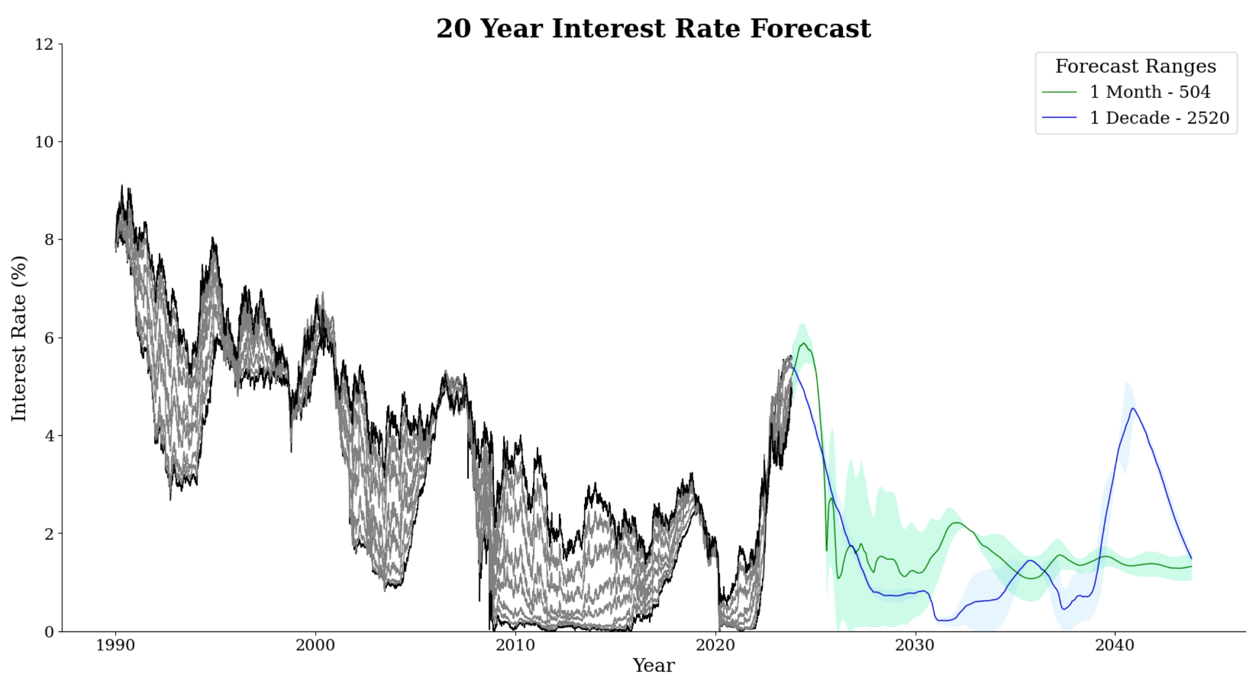


Fig 4. Experiment 1 Forecast B



Experiment 1 Forecast A was the first experiment conducted. The model fitting results were very promising (Appendix A), but the generated forecasts were anything but. Based on these results, the team

elected to continue searching the hyperparameter space for more suitable combinations of lookback and look-forward ranges. The team observed three takeaways from the forecasts generated in this forecast:

1. The 1-year and 1-decade models appear to have detected an underlying economic cycle influencing the rates. However, the 1-year model becomes incredibly unstable once detached from the historical data, and the 1-decade model needs some variation that makes it believable.

2. Some models, such as the 1-year and 1-month, do not perform well when moving past the historical data. One converges on a single value and goes completely flat; the other quickly falls towards the lower range and oscillates with little movement.

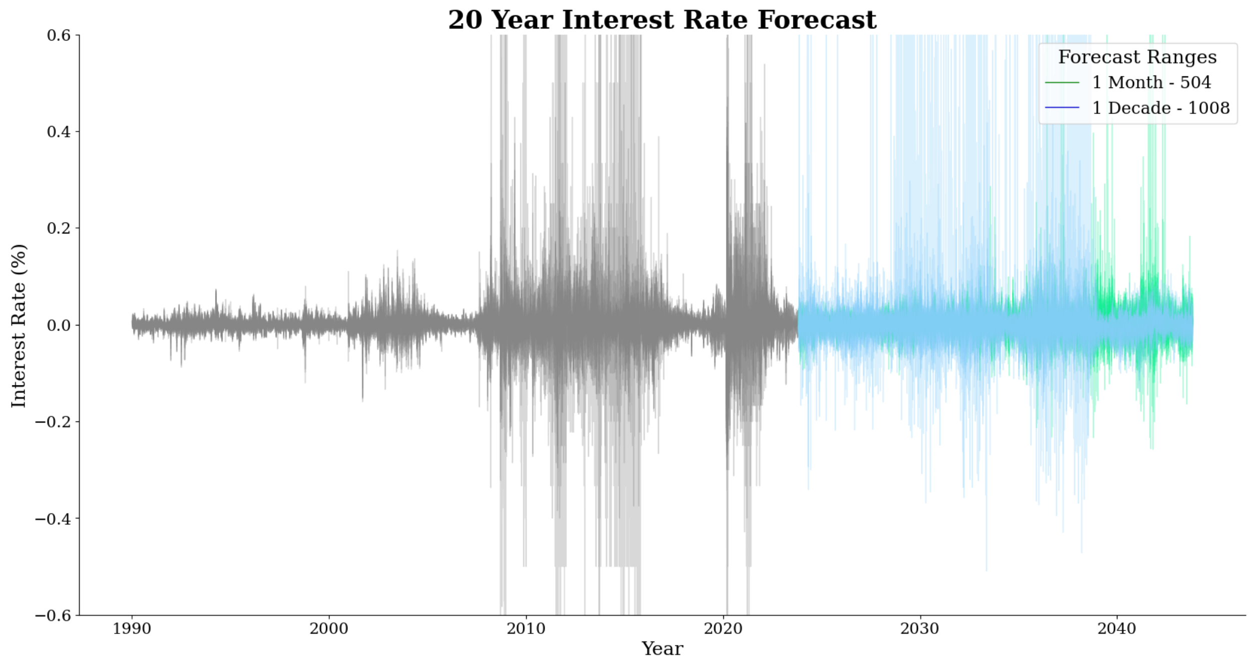
3. There does not appear to be a discernible, ideal ratio for the lookback-look-forward ranges.

Forecast B brought better results, although additional work must be done to fine-tune the predictions further. The selected 1-month model, with a four-year lookback range, offers an interesting prediction, at least into the early 2030s, before it appears to move towards a flat line. There is a continued rise into 2024-2025, followed by what appears to be a realistic approximation of rapid interest rate reductions with a broadening of the yield curve. The model struggles once pushed past the end of its historical data, but it does provide a potentially promising look into the near future. The 1-decade model with a ten-year lookback offers a similar promise. It predicts a slower reduction of rates followed by an uptick in the early 2040s. This model has less variation in the individual rates but still displays a believable future based on past observations.

Experiment 2 - Percent Change

Experiment 2 was far less insightful than Experiment 1, mainly due to the nature of the data and forecasts generated. There is potentially more to explore in this space; however, the models generally struggled far more in this experiment when learning from historical data (Appendix B) and failed to produce intuitive predictions. We expect that percent change may still be helpful in this prediction space but would need further research to implement meaningfully. The figure below displays the forecast for the two model configurations from Experiment 1 Forecast B above.

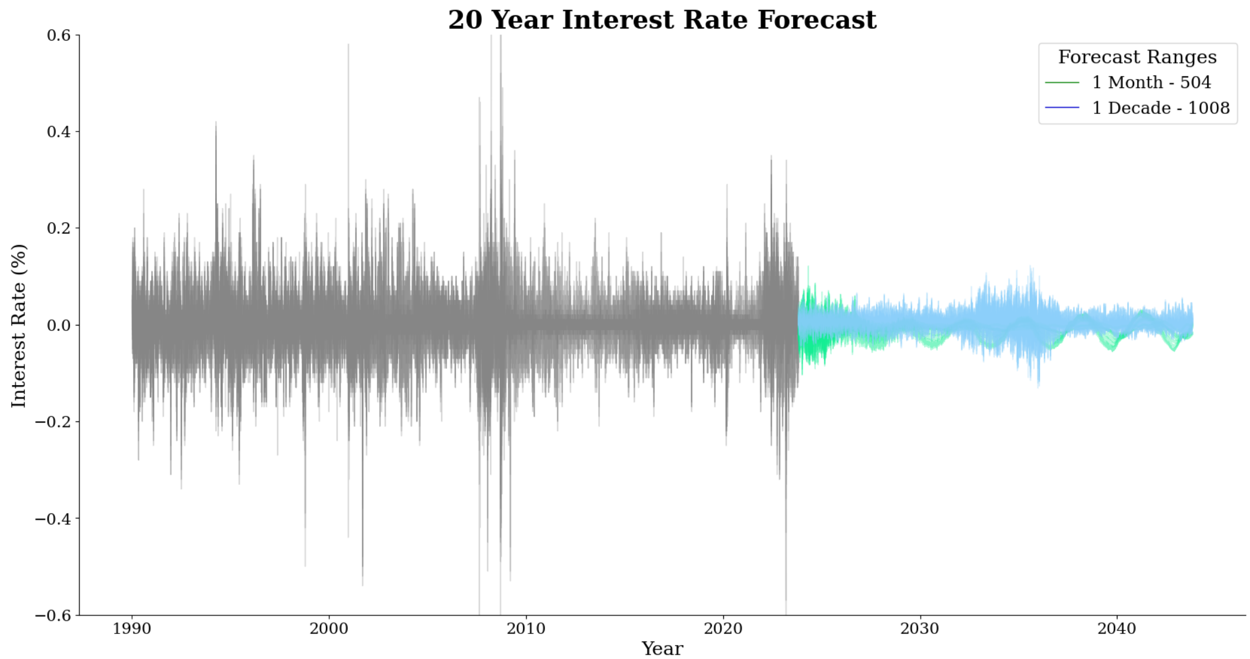
Fig 5. Experiment 2 Forecast



Experiment 3 - Daily Difference

The third experiment applied a transformation to the dataset where the interest rate values were the difference from the previous day. This technique reduced the absolute values to a much smaller range, and the models suffered only a minimal loss in their reported metrics when training (Appendix C). However, when the models were asked to generate predictions, the results were uninspiring. Below are the outputs from the same two models featured above, with the remaining four available in the appendix. The models simply regress toward zero and provide little indication that an economic cycle is present or learned from the data.

Fig 6. Experiment 3 Forecast

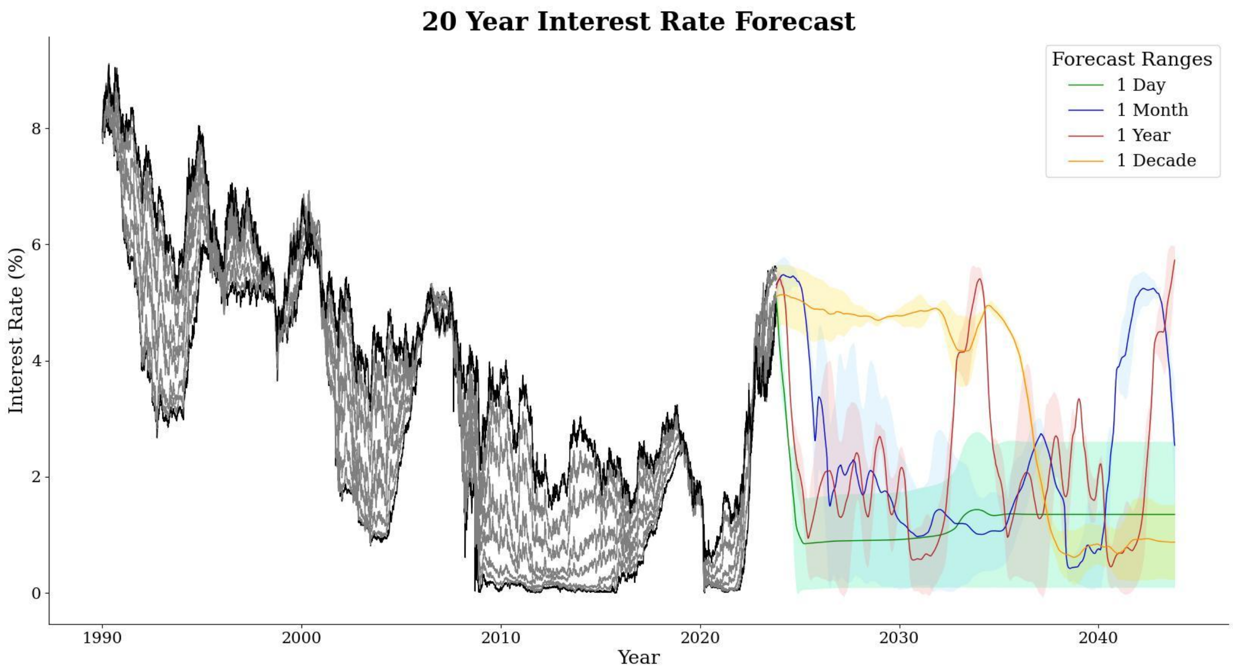


Experiment 4 - S&P 500 Data

Experiment 4 returned far more interesting results than Experiments 2 and 3, as visualized in the figure below. An essential factor to consider when evaluating this family of models is that the predictions include the daily interest rate yield curve and the daily performance of the S&P 500 index. As the model moves past the historical data range, it must also produce and store a day's worth of S&P 500 data to make future predictions.

The figure below captures the predictions from four different models, one for each time horizon, to display the variation in the projections. In summary, the S&P 500 data does not appear to boost confidence in the forecasts significantly. However, it offers an extra dimension to consider the models, and aspects of the outputs are worth exploring further.

Fig 7. Experiment 4 Forecast



The results from Experiment 4 show dramatic struggles from the 1-day model (which had issues in all experiments) and the 1-decade model. The 1-month and 1-year forecast models function more realistically, displaying some semblance of a variation over time and a periodic, cycle-like spike in the yield curve rates. The intermediate years seem quite choppy, which is unique to this experiment and may result from the dataset including the S&P 500 index data in its predictions.

Overall Results

The twenty-year forecasts for most models struggled significantly when the prediction range required the model to move past the available historical data and rely only on its own predictions to generate further values. This is not an unsurprising result; however, it was different from the outcome the team had initially aimed for. The following factors may contribute to the observed shortcomings of the models:

- The historical data used in this analysis only goes back to the early 1990s. A further range of data might provide more historical depth to the models and more variation to learn from.

- The models were trained from historical data that was "stacked" into a group of predictor values for each day. In some cases, these "predictor stacks" were several years long. This meant that when training, the models potentially saw some data points several times and may have overfit those points before making forecasts.

- The time-series nature of the data probably requires more nuance than an MLP Regressor can deliver. While the method is compelling and exciting to explore, it is not inherently designed to handle time-series data.

- Including other metrics, data, or index values would likely add complexity for the models to learn but might provide some additional variations for the models when making forecasts. To appropriately use this data in the forecasts, it could also be worth exploring alternate ways to generate those future values outside the predicted interest rates.

The ADF test results may also provide some insight into the issues faced with future simulations. The test found that the absolute values of the interest rate data are stationary, indicating that the overall properties of the data do not change significantly over time. The specific impact of this fact on the MLP Regressor models is not known, but this may provide a clue into the struggles observed in many of our models.

Conclusion

The actual outcome of the forecasts is disappointingly hard to determine, as only the passage of time can expose the true accuracy. To compensate for this, the team would recommend that the adopted modeling solution combines a series or family of models instead of just one. This ensemble approach might include several different time horizons for forecasts and weight each differently to emphasize the near-term and allow for more significant variation in the distant future where there are more unknowns.

We make this recommendation because the risks associated with inaccurate predictions could be catastrophic to an individual bank and, on a grand enough scale, the world economy. However, if a model or series of models could perform this task within a reasonable approximation of accuracy, it would be an enormous advantage in just about all markets.

To that end, future research should explore other methods for time-series modeling that might be more appropriate for the daily nature of this dataset. A variety of available Python packages can perform this task that were not leveraged here due to the scope of the experiments but could offer promise. Additionally, the team would recommend exploring other transformations to the dataset for the predictions. For instance, one might use a weekly average instead of daily rates. Or, instead of predicting a specific value, the rates might be rounded to the nearest tenth or whole point. These actions might make the data easier to leverage for the models and provide potentially brilliant predictions.

Appendix A - Experiment 1 Grid Search Results

Table 2. Experiment 1 Selected Model Performace

score	mse	model
0.993751	0.028849	{'lb': 100, 'lf': 2, 'hl': (200, 200, 200, 200...
0.996116	0.017721	{'lb': 100, 'lf': 30, 'hl': (200, 200, 200, 20...
0.993804	0.024596	{'lb': 756, 'lf': 252, 'hl': (200, 200, 200, 2...
0.946129	0.112234	{'lb': 1008, 'lf': 2520, 'hl': (200, 200, 200,...
0.982702	0.067162	{'lb': 504, 'lf': 30, 'hl': (200, 200, 200, 20...
0.975100	0.027223	{'lb': 2520, 'lf': 2520, 'hl': (200, 200, 200,...

Table 3. Experiment 1 Grid Search Results - MSE

Model MSE			
	504	756	1008
(100, 100, 100, 100)	0.032412	0.038337	0.048323
(100, 100, 100, 100, 100, 100, 100)	0.023644	0.037761	0.031064
(100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100)	0.024975	0.023557	0.031888
(200, 200, 200, 200)	0.553936	0.01865	0.031738
(200, 200, 200, 200, 200, 200, 200)	0.146001	0.021145	0.01879
(200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200)	0.038621	0.017016	0.014531
(300, 300, 300, 300, 300, 300, 300, 300, 300, 300, 300, 300, 300, 300)			0.033090
(500, 500, 500, 500)	0.02408	0.064427	0.032668
(500, 500, 500, 500, 500, 500, 500)	0.061218	0.027586	0.020573
(500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500)	0.027938	0.030202	0.039700
(750, 750, 750, 750)	0.060813	0.064164	0.040115
(750, 750, 750, 750, 750, 750, 750)	0.027886	0.041439	0.033888
(750, 750, 750, 750, 750, 750, 750, 750, 750, 750, 750, 750, 750, 750)	0.029353	1.290304	0.034296

Table 4. Experiment 1 Grid Search Results - Accuracy

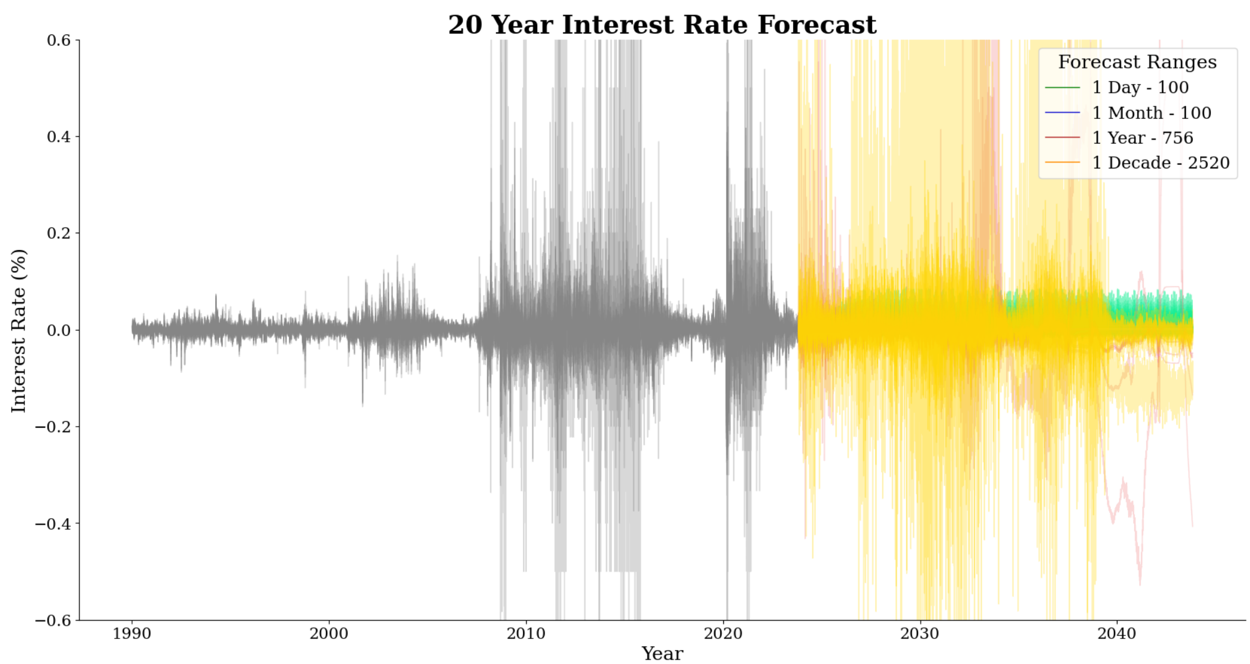
Model Score			
	504	756	1008
(100, 100, 100, 100)	0.982782	0.979123	0.976914
(100, 100, 100, 100, 100, 100, 100)	0.987007	0.980552	0.984223
(100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100)	0.986072	0.986686	0.982698
(200, 200, 200, 200)	0.752327	0.989786	0.981678
(200, 200, 200, 200, 200, 200, 200)	0.931822	0.988431	0.990426
(200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200)	0.979337	0.990712	0.992056
(300, 300, 300, 300, 300, 300, 300, 300, 300, 300, 300, 300, 300, 300)			0.981672
(500, 500, 500, 500)	0.986907	0.966376	0.984120
(500, 500, 500, 500, 500, 500, 500)	0.969922	0.984213	0.989287
(500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500)	0.984785	0.983897	0.979486
(750, 750, 750, 750)	0.970515	0.97005	0.977048
(750, 750, 750, 750, 750, 750, 750)	0.98495	0.975724	0.981518
(750, 750, 750, 750, 750, 750, 750, 750, 750, 750, 750, 750, 750, 750)	0.984052	0.415142	0.982182

Appendix B - Experiment 2 Model Fitting Results

Table 5. Experiment 2 Selected Model Performance

score	mse	model
0.643697	0.000603	{'lb': 100, 'lf': 2, 'hl': (200, 200, 200, 200...
0.249064	0.001975	{'lb': 100, 'lf': 30, 'hl': (200, 200, 200, 20...
0.576142	0.001502	{'lb': 756, 'lf': 252, 'hl': (200, 200, 200, 2...
0.816832	0.001163	{'lb': 1008, 'lf': 2520, 'hl': (200, 200, 200,...
0.758497	0.000597	{'lb': 504, 'lf': 30, 'hl': (200, 200, 200, 20...
0.780125	0.002031	{'lb': 2520, 'lf': 2520, 'hl': (200, 200, 200,...

Fig 8. Experiment 2 Alternate Model Forecast

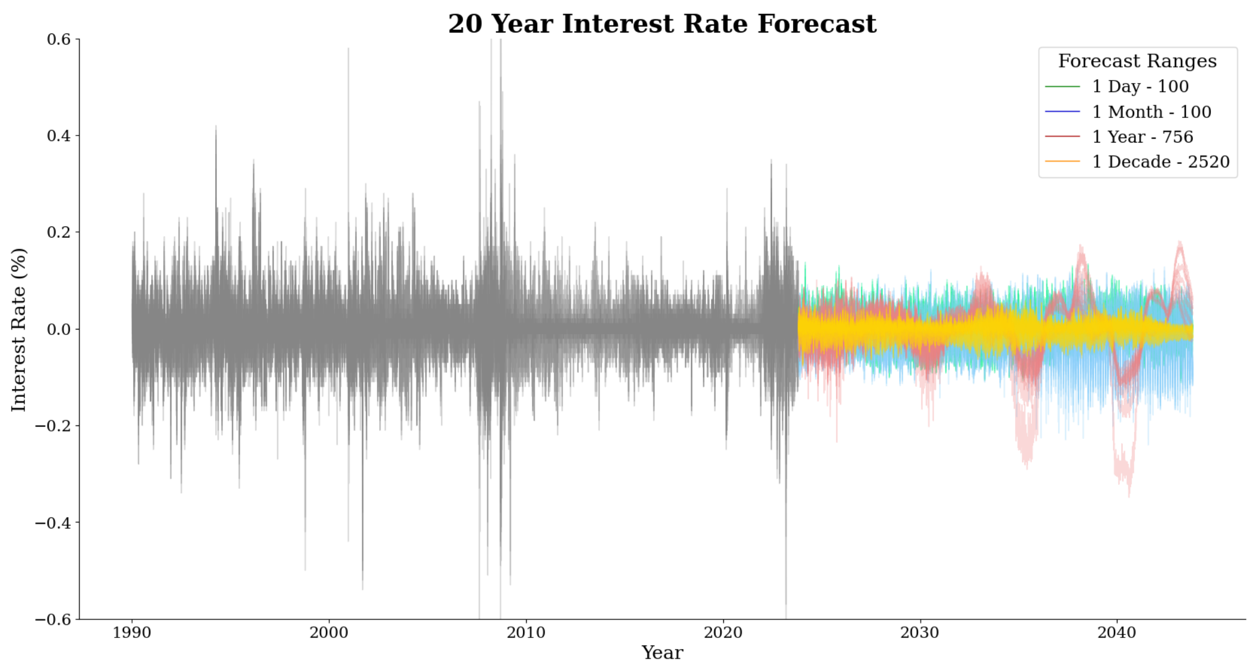


Appendix C - Experiment 3 Model Fitting Results

Table 6. Experiment 3 Selected Model Performance

score	mse	model
0.897482	0.000261	{'lb': 100, 'lf': 2, 'hl': (200, 200, 200, 200...
0.897385	0.000262	{'lb': 100, 'lf': 30, 'hl': (200, 200, 200, 20...
0.934950	0.000166	{'lb': 756, 'lf': 252, 'hl': (200, 200, 200, 2...
0.940441	0.000129	{'lb': 1008, 'lf': 2520, 'hl': (200, 200, 200,...
0.935574	0.000159	{'lb': 504, 'lf': 30, 'hl': (200, 200, 200, 20...
0.974183	0.000041	{'lb': 2520, 'lf': 2520, 'hl': (200, 200, 200,...

Fig 9. Experiment 3 Alternate Model Forecast



Appendix D - Experiment 4 Model Fitting Results

Table 7. Experiment 4 Selected Model Performance

	score	mse	model
0	0.973677	0.024012	{'lb': 100, 'lf': 2, 'hl': (200, 200, 200, 200...
0	0.987815	0.039854	{'lb': 100, 'lf': 30, 'hl': (200, 200, 200, 20...
0	0.966930	0.029522	{'lb': 756, 'lf': 252, 'hl': (200, 200, 200, 2...
0	0.981476	0.017261	{'lb': 1008, 'lf': 2520, 'hl': (200, 200, 200,...
0	0.990185	0.019141	{'lb': 504, 'lf': 30, 'hl': (200, 200, 200, 20...
0	0.970766	0.022095	{'lb': 2520, 'lf': 2520, 'hl': (200, 200, 200,...

References

Glorot, Xavier, and Yoshua Bengio. *Understanding the difficulty of training deep feedforward neural networks. International Conference on Artificial Intelligence and Statistics*, https://proceedings.mlr.press/v9/glorot10a/glorot10a.pdf.

"The Mathematics of Data Science: Understanding the foundations of Deep Learning through Linear Regression - DataScienceCentral.com." *Data Science Central*, 14 January 2019, https://www.datasciencecentral.com/the-mathematics-of-data-science-understanding-the-foundations-of. Accessed 4 December 2023.

Politof, Themistocles, and Dan Ulmer. "Predicting Interest Rates Using Artificial Neural Networks." *Operational Tools in the Management of Financial Risks*, edited by Constantin Zopounidis, Springer US, 2012, https://link.springer.com/chapter/10.1007/978-1-4615-5495-0_17#citeas. Accessed 4 December 2023.

Rapach, Daniel, and Christian Weber. "Are real interest rates really nonstationary? New evidence from tests with good size and power." *Journal of Macroeconomics*, vol. 26, no. 3, 2004, pp. 409-430. *ScienceDirect*, https://doi.org/10.1016/j.jmacro.2003.03.001.

S&P Global. "S&P 500® | S&P Dow Jones Indices." *S&P Global*, https://www.spglobal.com/spdji/en/indices/equity/sp-500. Accessed 4 December 2023.

U.S. Department of the Treasury. "Daily Treasury Par Yield Curve Rates." *Resource Center*, 2023, https://home.treasury.gov/resource-center/data-chart-center/interest-rates/TextView?type=daily_treasury_yield_curve&field_tdr_date_value=all. Accessed 31 October 2023.

Yasir, Muhammad, et al. "An Efficient Deep Learning Based Model to Predict Interest Rate Using Twitter Sentiment." *Sustainability*, vol. 12, no. 4, 2020, p. 1660, https://www.mdpi.com/2071-1050/12/4/1660.