Machine Learning Engineer Nanodegree

Capstone Project

Minh Le
March 2nd, 2019

I. Definition

Project Overview

The federal funds rate is the interest rate at which depository institutions lend balances at the Federal Reserve (FED) to other depository institutions overnight. Changes in the federal funds rate trigger a chain of events that affect other short-term interest rates, foreign exchange rates, long-term interest rates, the amount of money and credit, and, ultimately, a range of economic variables, including employment, output, and prices of goods and services 1. This is one of the tools available to the FED to influence the availability and cost of money and credit to help promote national economic goals. The FED is supposed to arrive at interest rate decisions after considering many financial economic data and models. The decision is usually made with voting at quarterly FOMC (Federal Open Market Committee) meetings. Emergency decisions are rare but possible, which happened in 2008-2009 financial crisis.

Four financial economic data series are used in this research. These are Effective Federal Funds Rate, 5-Year Treasury Constant Maturity Rate, University of Michigan: Inflation Expectation, and Civilian Unemployment Rate. Data is downloaded from Federal Reserve of St. Louis.

Effective Federal Funds Rate (FEDFUNDS): https://fred.stlouisfed.org/series/fedfunds

5-Year Treasury Constant Maturity Rate (GS5): https://fred.stlouisfed.org/series/GS5

University of Michigan: Inflation Expectation (MICH): https://fred.stlouisfed.org/series/MICH

Civilian Unemployment Rate (UNRATE): https://fred.stlouisfed.org/series/UNRATE

All series have monthly frequency. From these series, 'CHANGE' is calculating as sequential period difference pf FEDFUNDS. Starting date is set at 1/1/1978 as the common earliest available for all.

The data series available start in 1971, near 'Nixon shock' event: wage-price controls, weakened gold standard, and temporary surcharge (tariff) on all imports. Fed lowered rate to boost growth[1]. The president followed up with devaluing dollar, stoking inflation. Fed raised rate to combat 3.5% YOY inflation. Economy entered Stagflation in the 1970s. The 1980s is marked with FED Chairman Paul Volker battling inflation by raising FED rates to unprecedented high double-digit level. This caused the economy to experience severe double-dip recession, but high inflation and Stagflation ended. The 1990s later periods marked consistently low rates and low inflation. The recent financial crisis witnessed zero-rate and unusually low inflation.

The goals of this research are: 1) Predict the FED interest rates, 2) Infer higher order logic IF-THEN the decision rules, 3) Asses the relative importance of input series on the rules.

---

[1] https://www.thebalance.com/fed-funds-rate-history-highs-lows-3306135

Problem Statement

Despite its paramount importance, it is not clear how these decisions are made. Voting members subjectivities may also play a role. The 'Taylor' rule2 is a conceptual of how interest rates should be, given unemployment and inflation. Output from Taylor rule, however, does not adequately reflect the FED's decision (Red and Green lines versus Blue line below)[2].

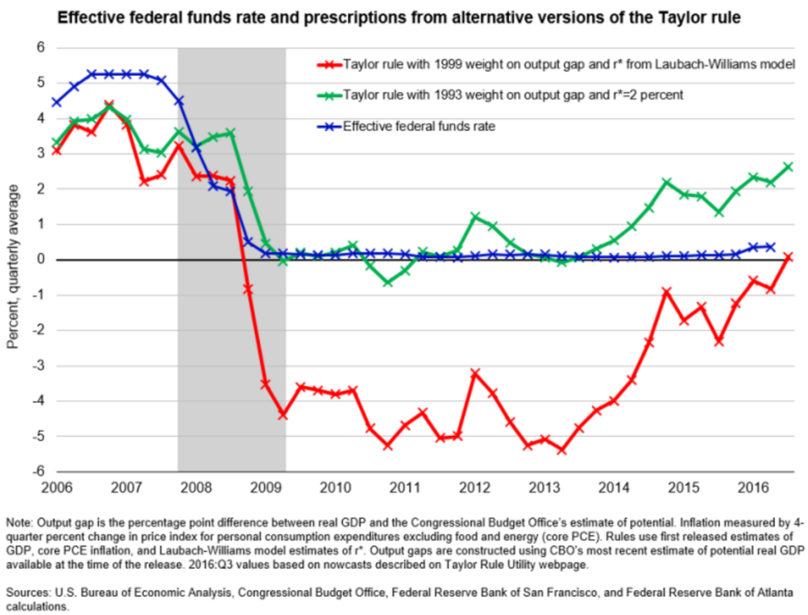


Figure 1 Effective Rate and Alternative Taylor Rule

Having rules closer to actual FED rates would be helpful. Furthermore, financial models have limited interpretation on the decision process. I propose using Adaptive fuzzy inference neural network as an alternative approach. Observable FED rates decision is modeled by changes in effective FED fund rates. Input data include common cited series, such as Inflation CPI/Income-change, Unemployment rate, GDP and SP500. Fuzzy logic membership functions will be Gaussian-shape.

The problems posed in previous section can be solved by:

- Building Fuzzy Membership functions from data patterns using Gaussian Mixture method.
- Extracting Fuzzy Rules from data and Membership functions.'
- Evaluate, experiment and compare the system's output with different rules modifications.

Metrics

This model has two deliveries:

- Predict the FOMC decision as reflected in changes in Effective Federal Fund Rates.
- Extract the decision rules in order to describe in high order logic.

---

[2] https://commons.wikimedia.org/wiki/File:Taylor_Rule_Prescriptions_for_Fed_Funds_Rate_2016.png

Model performance, therefore, is analyzed using quantitative and qualitative measures:

- Quantitative: model fitness, and root mean square errors for both in-sample and out-sample.
- Qualitative: soundness of extracted decision rules using economic theories.

Quantitative metrics are:

Root Mean Square Error (RMSE): This is the square root of sum of square of errors between prediction from ANFIS output and observed data. For benchmark model, this is the prediction of Decision Tree.

Maximum Absolute Error (MAE): This is maximum of absolute value of the differences between prediction output and observations. This metrics assesses maximum deviation while RMSE is about average deviation.

II. Analysis

Data Exploration

Federal Reserve's Federal Open Market Committee (FOMC) influences short-term interest rates by setting target interest rate. This rate might take the form of a point value or a tight range. The target rate is directly, and almost immediately, reflected by changes in Effective Federal Fund Rates. This is the interest rate at which depository institutions trade federal funds (balances held at Federal Reserve Banks) with each other overnight. The historical effective FED Funds Rates show how FED rate decision results in actual rates. These are the results of the decisions, with small variations.

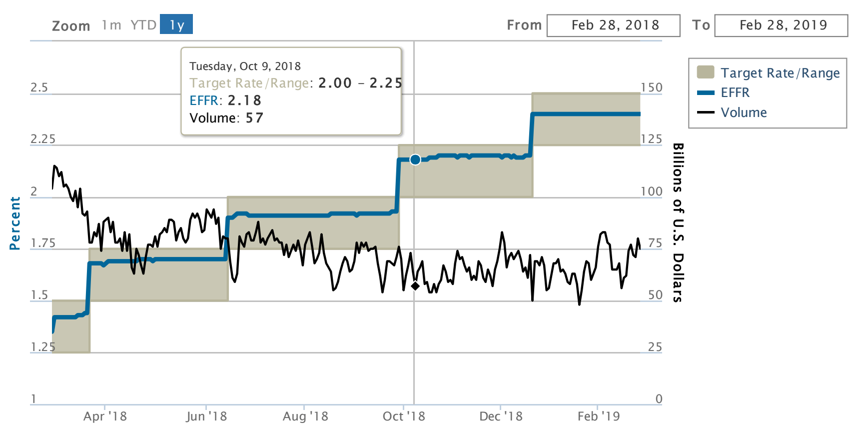


Figure 2 Target Rate and Effective Fed Fund Rate

The US economy has gone through multiple expansion contraction cycles in the last half centuries. The FED intervention actions had profound impacts on these cycles.

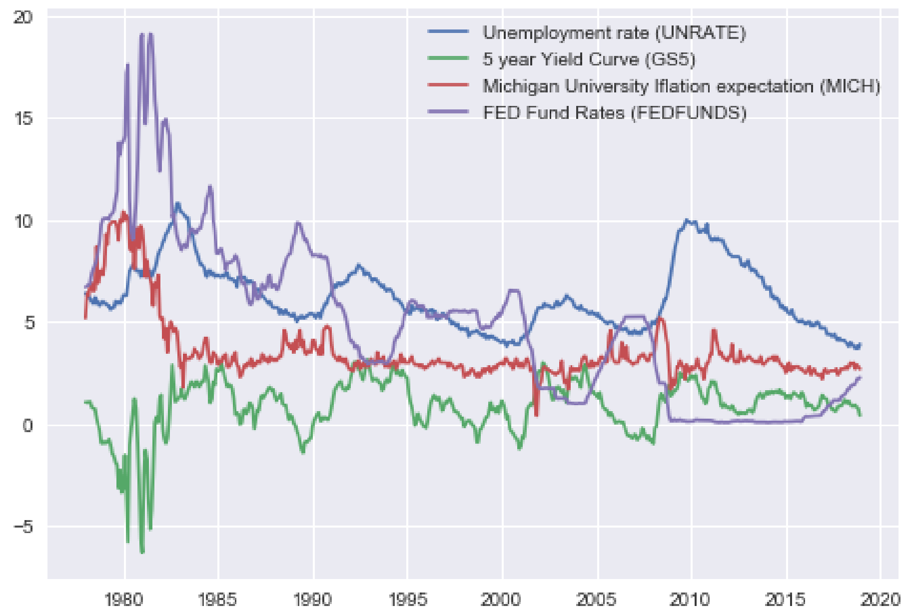


Figure 3 Date Series of the model

Following is the common narrative in monetary/business-cycle economic theory: When unemployment rate is high and inflation is low (recession), the FED would try stimulating the economy by cutting rates. As the economy recovers and expands, unemployment drops, and inflation increases. When the economy/inflation heats up, the FED increases rate (tightening) to cool it down. As results, the economic growth slows or even becomes negative (recession), and the cycle repeats. The FED faces a difficult task of devising subjective control policy to balance economic growth and inflation.

There is no basis to give definitive answer to the question on the 'optimal' rate. Lacking such basis, the Taylor-rule rate is used ad the 'neutral' benchmark to assess FED cut-tighten bias. The Taylor rule uses an objective formula:[3]

$$Target\,rate = \pi + r + 0.5(\pi - \pi_{desired}) + 0.5(y - y_{potential})$$

Equation 1 Taylor rule

In standard Taylor rule above, π is the rate of inflation, r is the equilibrium real interest rate, y is logarithm of GDP and $y_{potential}$ is the logarithm of potential output, as determined by a linear trend.

$\pi_{desired}$ is commonly accepted as 2%. Furthermore, data series have different report frequencies, with GDP only available on quarterly basis and often come in subsequent revisions. Thus, in this study, the GDP deviation is replaced with Unemployment deviation (UNRATE – U_longterm). U_longterm is commonly accepted at 5%. This modification allows all data series to be sampled at monthly frequency, and at the same time retains the objectivity of the rule. In addition, in order to ensure that GDP-tied economic growth is still represented, the model considers Yield-Curve spread between long-

[3] https://en.wikipedia.org/wiki/Taylor_rule

term 5-year and short-term rates. This spread is commonly used in financial economic literature as a proxy for economic growth expectation.

As described in the narrative, there are two distinctive regimes in the data. As result, the two versions of model implementation are developed. The main version is calibrated with data since January 1st, 1990, and the second with full sample. The first version is the main version, most recent and thus applicable. The second version is useful to provide deeper qualitative understanding.

Further Exploratory and Final Visualization

Financial economic data (refer to Figure 3) are highly volatile, often with revisions and annualizations. Often in financial economic researches, conversions are necessary to cope with noisy data and avoid spurious results. Following are appropriate steps:

Smoothing: Apply 12-period moving average to data series. Since original data frequency is monthly, this represent annualized smoothing.

Relative change: Calculate Relative-change instead of Level-change to Fed Fund interest rate. A 0.25% change of a 2% interest rate is huge (recent), while 0.25% of a 15% is barely noticeable (early 1980s). Changes in FED rates are numerator. Because the purpose of this work is about FED's deviation from neutral stand (Taylor rule), the denominator is the Taylor rule calculated at each month.

Normalization: With data series come in with very different scales and signs, they are normalized to be in [0,1] range. Normalization is important to learning efficiency.

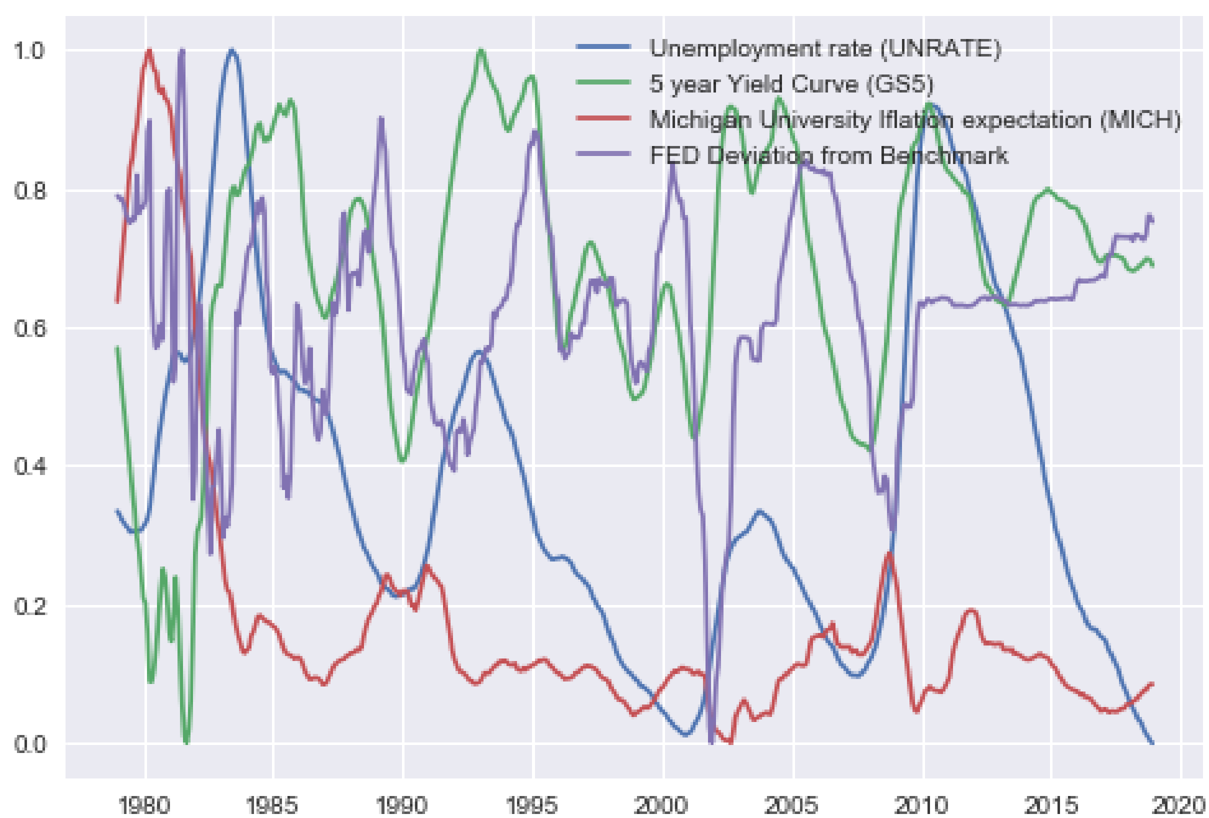


Figure 4 Final, and normalized, data series for modeling

Algorithms and Techniques

The model uses a combination of algorithms and techniques. I added to stat-of-art by modifications and incorporating techniques from other fields, as follows:

1. Fuzzy logic. Despite its wide uses in engineering control systems, Fuzzy logic is highly specialized and not as well known. I will give a short overview as needed background. In contrast to a crisp, single-value logic, fuzzy logic synonymous with the theory of fuzzy sets, a theory which relates to classes of objects with unsharp boundaries in which membership is a matter of degree[4]. The membership function of a fuzzy set is a generalization of the indicator function in classical sets. In fuzzy logic, it represents the degree of truth as an extension of valuation[5]. A high order logic fuzzy rule is applicable on a wide range of input, making it more stable and intuitive to human. Below is an illustration for linguistic terms[6].

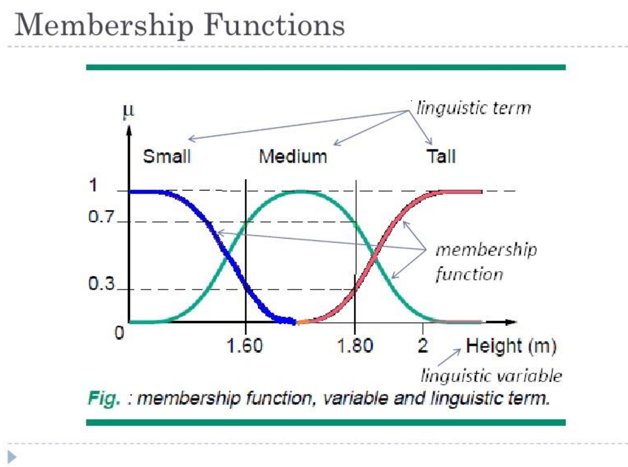


Figure 5 Example of membership functions

2. Modified Adaptive network-based fuzzy inference system (ANFIS): ANFIS is a kind of artificial neural network that is based on Takagi–Sagano fuzzy inference system. This inference system uses a set of fuzzy IF–THEN rules like human decision rules[7]. In ANFIS, the membership functions are estimated together with fuzzy rules. This estimation coupling may cause the end-result membership functions, while optimized to give better prediction, substantially differ from those observed with original input. Since one of the main objectives of this research is to understand FED decision in reacting to data, I decide to decouple this estimation.

[4] https://www.mathworks.com/help/fuzzy/what-is-fuzzy-logic.html

[5] https://en.wikipedia.org/wiki/Membership_function_(mathematics)

[6] https://www.slideshare.net/todkarmahesh/fuzzylogic

[7] http://liacs.leidenuniv.nl/~nijssensgr/CI/2011/anfis.pdf

The output of Modified ANFIS regression is mapped to higher order logic IF-THEN decision rules. Below are the first 4 rules of total 29 rules of FED rates decisions automatically extracted from data using Fuzzy Logic Modified ANFIS.

	UNRATE	GS5	MICH	RATE
0	UNRATE=High	GS5=Low	MICH=High	RATE=High
1	UNRATE=Low	GS5=Low	MICH=Low	RATE=Medium
2	UNRATE=Low	GS5=High	MICH=Low	RATE=High
3	UNRATE=High	GS5=Medium	MICH=Low	RATE=Low

3. Exponential-form density functions are widely used as Fuzzy membership functions. This fits naturally with unsupervised clustering using Gaussian Mixtures in Machine Learning. For certain data series, such as relative changes in interest rates, Gaussian Mixtures (GMM) can be unstable. Given the same data, even with the same mean points initialization, GMM can give different converged clustering structure. Because this is unsupervised learning, there is no 'best' structure. A fixed random initialization gives corresponding fixed clustering structure. This structure, however, may not be consistent with theoretical foundations. I devised a method combining mean-points initialization and repetitive warm-start GMM implementation to cope with this challenge. In Figure 6, the coordinates of each 3-D point are the mean values of a 3-component clustering structure for relative rate change data series. Each structure is a converged structure with the same mean-points initialization with different random seeds. Different structures result in different, sometime materially different, final inferences. The solution is to have large (500) random executions. *Best model based on Validation testing performance (RMSE) is selected and saved as binary object (pickle) as final version.*

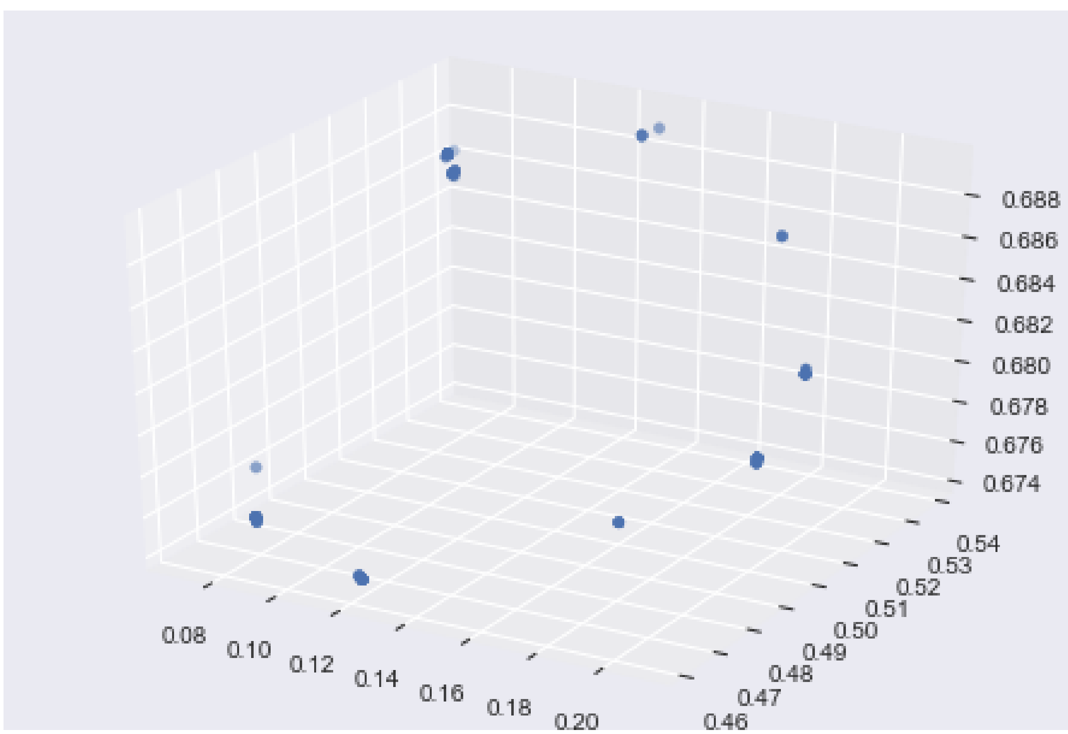


Figure 6 Multiple Gaussian Mixtures clustering structures

4. Conjoint analysis: Fuzzy inference systems, ANFIS included, stop at extracting IF-THEN rules. As explained in previous sections, the FED faces difficult choices and often contradicting data points. For the purpose of understanding FED decision, it is important to understand the relative importance of data series conditions. For example, all else equal, would 'Higher growth' relatively more important than 'High inflation'? This question is outside of Fuzzy inference system. To answer above question, I use a technique from marketing, call 'Conjoint analysis'. In marketing, 'Conjoint analysis' is a survey-based statistical technique used in market research that helps determine how people value different attributes (feature, function, benefits) that make up an individual product or service[8]. In this framework, I use data series conditions as features and assess their relative importance[9]. Figure 7 gives an example of the analytics used in conjoint analysis for smartphone. The relative calculation shows that 'Price' attribute has the highest relative importance. Conjoint analysis for this model will be discussed in detail in later section.

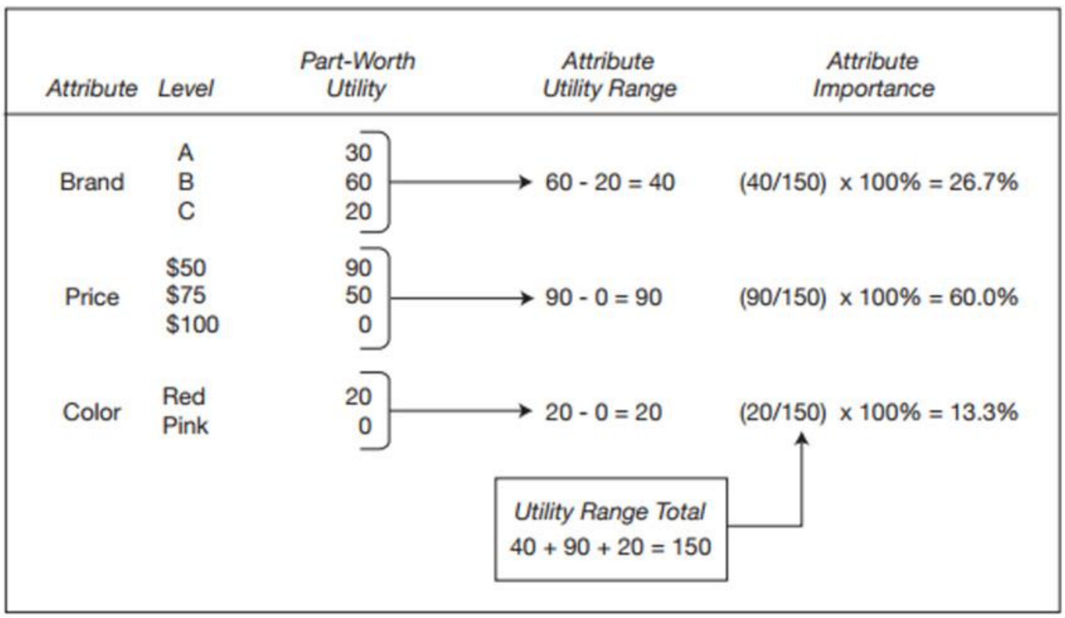


Figure 7 Example of conjoint analysis for smartphone

5. I incorporate both Fuzzy Logic Regression (Modified ANFIS) and Decision Tree Regression into Fuzzy Decision Rules mapping. This method allows a 'normal', crisp (non-Fuzzy), Decision Tree to be abstracted to higher order logic IF-THEN rules. This method strengthens the interpretation of Decision Tree learning mechanism. Comparing the two decision rules, which were derived from two regression engines, gives deeper insight on he quality of prediction in addition of standard measures such as Root-Mean-Square and Max Absolute Errors.

---

[8] https://en.wikipedia.org/wiki/Conjoint_analysis

[9] https://www.fieldboom.com/conjoint-analysis

Benchmark

Decision Tree Regression is chosen as benchmark model. A key strength of Decision Tree is to gives clear interpretation of a decision process, which aligns with this research. Decision Tree can deliver highly accurate prediction, albeit sometimes at the expense of overfitting and stability. With careful implementation, Decision Tree Regression can be a very good choice. In this example, Decision Tree Regression in deed gives highly accurate prediction, just slightly lower than the Fuzzy Logic model.

The key advantage of Fuzzy Logic model is higher logic, human like, interpretation of decision rules. As show below, the decision steps in a Decision Tree are number-based and mechanical.

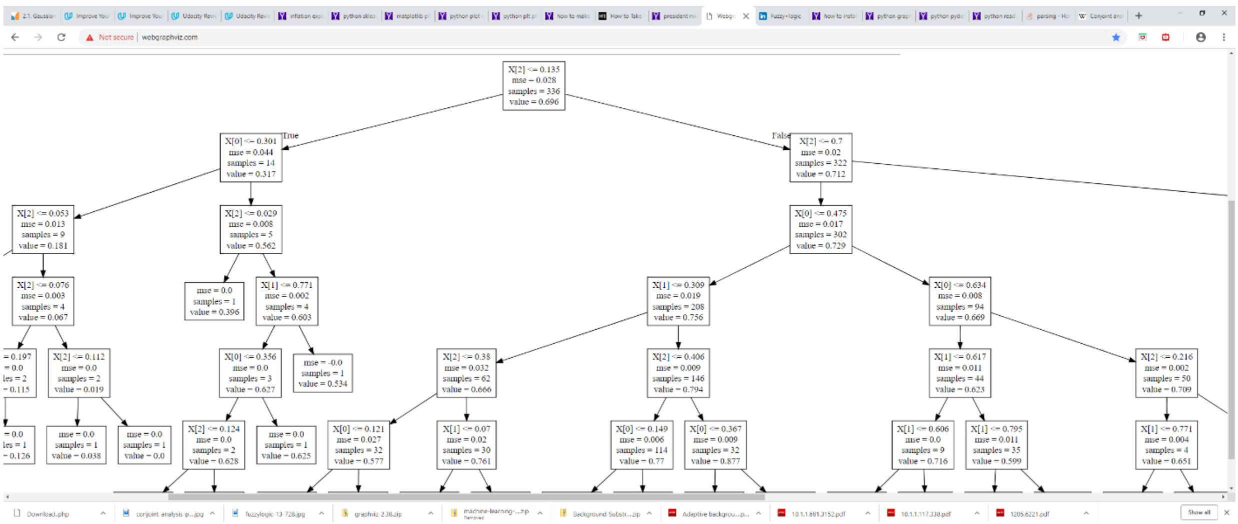


Figure 8 Decision Tree Regression

The results from this Decision Tree is mapped into Fuzzy Logic Rules through membership function of the observed output. Below are the first 4 rules of the total 23 FED rates rules derived automatically from data using Decision Tree Regression. Note the differences with those using ANFIS. The differences with those derived from Modified ANFIS will discussed in detail in the Results section.

	UNRATE	GS5	MICH	RATE
0	UNRATE=Low	GS5=Medium	MICH=Medium	RATE=High
1	UNRATE=Low	GS5=Low	MICH=Medium	RATE=Medium
2	UNRATE=Medium	GS5=Medium	MICH=Medium	RATE=High
3	UNRATE=High	GS5=Low	MICH=Medium	RATE=Medium

III. Methodology

Data Preprocessing

Data processing for Fuzzy Logic is different from crisp logic. It is divided into two parts: Raw data Processing and Membership Function Estimation.

Raw Data Processing

Data series are read from an external csv file. This data is collected from Federal Reserve Bank of St. Louis[10]. The start and end dates are standardized to be between 1/1/1978 and 12/1/2018, monthly frequency. This ensures all series to have the same length. The data series are:

1. Effective Fed Fund Rates (FEDFUNDS).
2. 5-Year Treasury Constant Maturity Rate (GS5). The difference of GS5 and FEDFUNDS is 5-year yield curve spread. From now on, GS5 denotes this spread.
3. Michigan University Inflation expectation (MICH).
4. Civilian Unemployment Rate (UNRATE).

Taylor rule objective rates are calculated using these series:

taylor = np.array(2 + MICH + 0.5*(MICH - 2) + 0.5*(UNRATE - 5))

Relative rate change is then calculated, adjusting one position for the 'diff' operator.

rates_diff = np.diff(rate)/taylor[1:]

Smoothing mechanism using 12-month moving average is applied to all data.
Normalization to be in [0,1], for example:

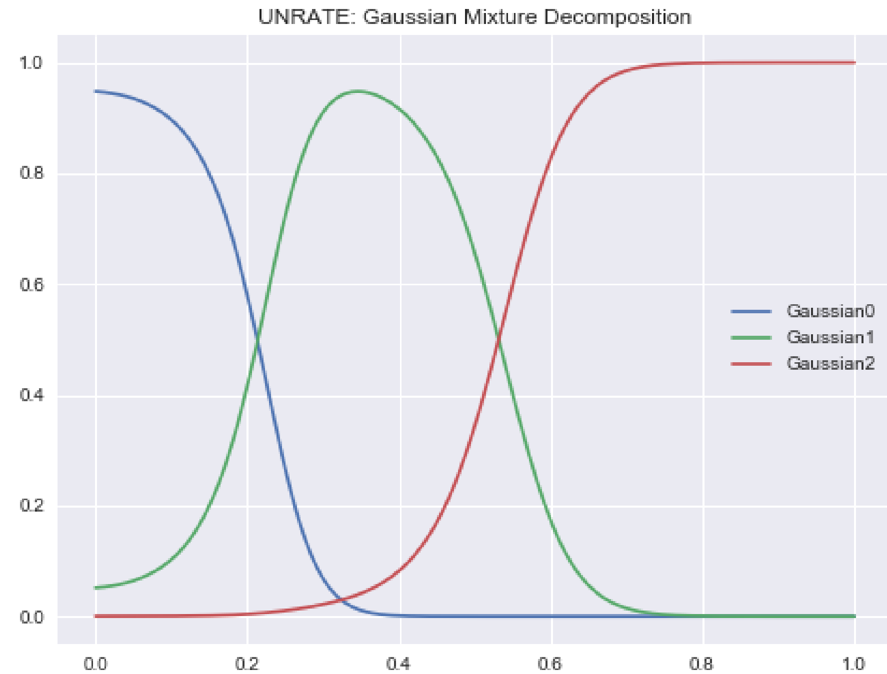
UNRATE_norm = (UNRATE - UNRATE.min(0))/UNRATE.ptp(0)

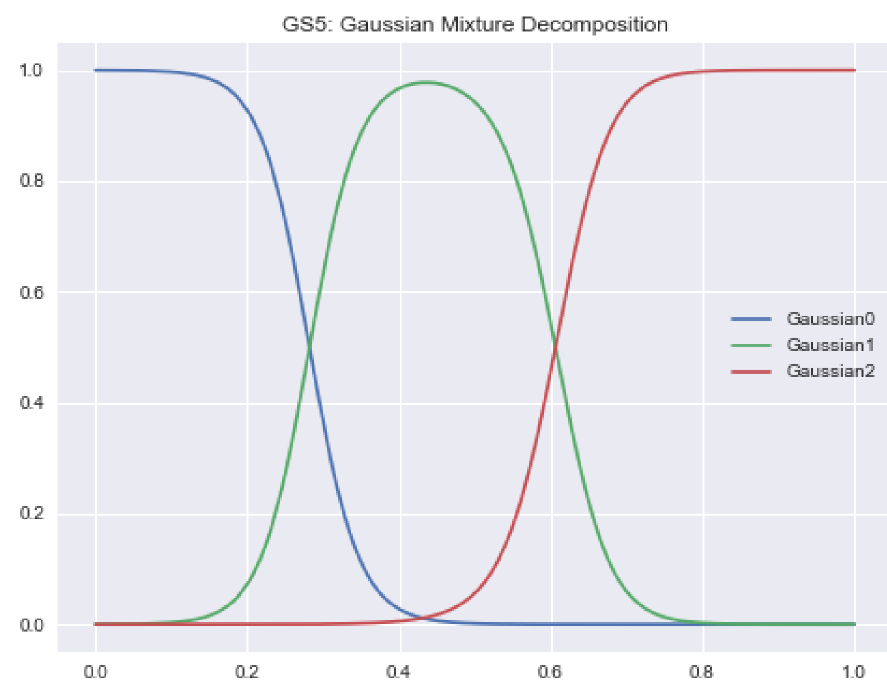
Membership Function Estimation

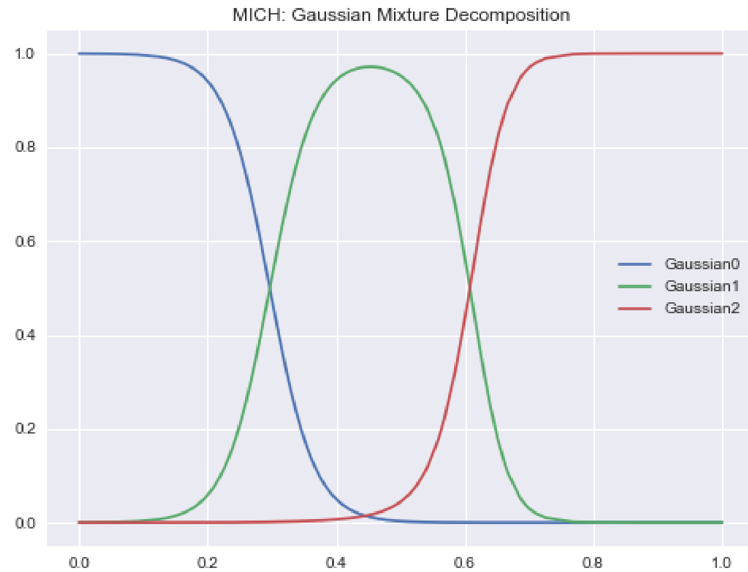
As explained in Algorithms and Techniques section, Machine Learning technique with Gaussian Mixtures Model for unsupervised clustering (GMM) is used to build Fuzzy Set membership functions. There are a few key decisions in this process:

Number of clusters: There is no hard rule for this choice. In this research, data size is not large (336 in main target period). There are three (3) feature series as independent variables. If 'n' is the number of clusters in each series, there are n^3 combinations of Fuzzy premises. The minimum value of 'n' to reflect economy conditions (Week, Average, Strong) is 3, with 27 combination. Increasing 'n' to 4 reduces average number of data points in each premise (336/64 ~= 5) without notable improvement in the shape of membership functions. Thus 'n' is fixed at 3. Below are GMM results for input series. The clustering results are stable with regards to random initialization. No mean-points initialization is needed.

[10] https://fred.stlouisfed.org/series/fedfunds


UNRATE: Gaussian Mixture Decomposition
Gaussian0
Gaussian1
Gaussian2
0.0
0.2
0.4
0.6
0.8
1.0


GS5: Gaussian Mixture Decomposition
Gaussian0
Gaussian1
Gaussian2
0.0
0.2
0.4
0.6
0.8
1.0



Random seed initialization and mean-points initialization: Building membership functions for RATE is a challenge. The number of clusters remain 3. Increasing it does not improve anything. Running with different random seeds, even with the same mean-points initialization gives different clustering structures (Figure 6). This is a known issue with GMM, which may get stuck at local optimization. It is not possible to derive 'global' optimization due to unsupervised learning. Hard coding a random seed is not stable with data changes. In order to estimate mean-points initialization, I use Python 'skfuzzy.cluster.cmeans' algorithm (below). The clusters centers give the initial means.

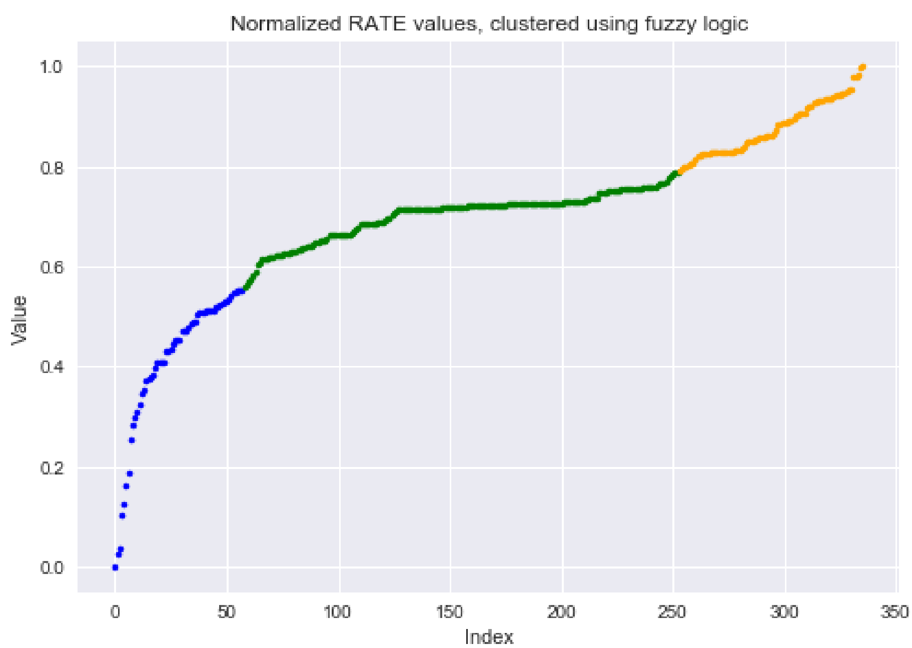


Figure 9 Fuzzy Mean Clustering using skfuzzy cmeans clustering

Implementation

The model has four (4) component stages.

Stage 1. Data processing and Membership Functions (MFs) Estimation.

This stage is described in detail in previous section, with emphasis on MFs

Stage 2. Regression (Fuzzy Logic for champion model and Decision Tree for Benchmark)

ANFIS's layers 1-3 are implemented to calculate Fuzzy probability weights based on input values and membership functions[11] (Jyh-shing and Roger Jang, 1993).

Layer 1: calculate weights w(i) of input X(i) using membership functions as densities.

Layer 2: cross-product of weights as probability estimation of combinations of inputs.

Layer 3: normalization to arrive at proper probability values of these combinations.

In the original paper, the authors use forward and backward learning procedures to estimate both membership function parameters and linear regression betas. In this model, the membership functions are estimated separately (see Algorithms and Techniques), and linear regression betas are estimated directly instead of multi-step learning.

Stage 3. Mapping regression fitted results to Fuzzy Rules.

Each of observed input, output and fitted/predicted values are mapped into Fuzzy regions using estimated membership functions, in two steps:

1. Getting Fuzzy range (region) names
2. Each data point is calculated and map to range name

Stage 4. Conjoint Analysis.

The high order logic IF-THEN rules extracted (29 via ANFIS and 23 via Decision Tree) contain different combinations of input conditions (premises). In order to understand their relative importance, conjoint analysis is used. Note that this analysis takes the IF-THEN Fuzzy Rules as input. Because these are categorical input, this analysis has two stages. First, the IF-THEN rules data is converted to dummies. Then, a linear regression is performed on these dummy variables.

The results of the final regression show the relative importance (Part Worth) of the rules features.

Refinement

ANFIS performance is quite high. Recent period (since 1990) gives Adjusted R^2 70.5%, RMSE 0.082, Maximum Absolute Error 0.47. Max Absolute Error is considerably larger than RMSE, which reflects average error, due to the spike noises in input data. Rule mapping accuracy 93.15%.

I applied careful calibrations to Decision Tree although it is a benchmark. A decision tree can give very high fit (overfit) by expanding tree depth. Maximum tree depth is a key parameter. Cross validation is a general helpful technique. In this case, there is an additional consideration. Because this is a benchmark model, the max tree depth is selected to deliver **average** regression performance comparable to Fuzzy Logic (RMSE). Even though RMSE is comparable, Decision Tree performance is notably worse in Maximum Absolute Error because Decision Tree predict flat-line output. In its final

[11] http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.42.2913

configuration, the Decision Tree has ‘max_depth’ of six (6). With three (3) input variables, the value range of each variable can be processed by the tree twice on average.

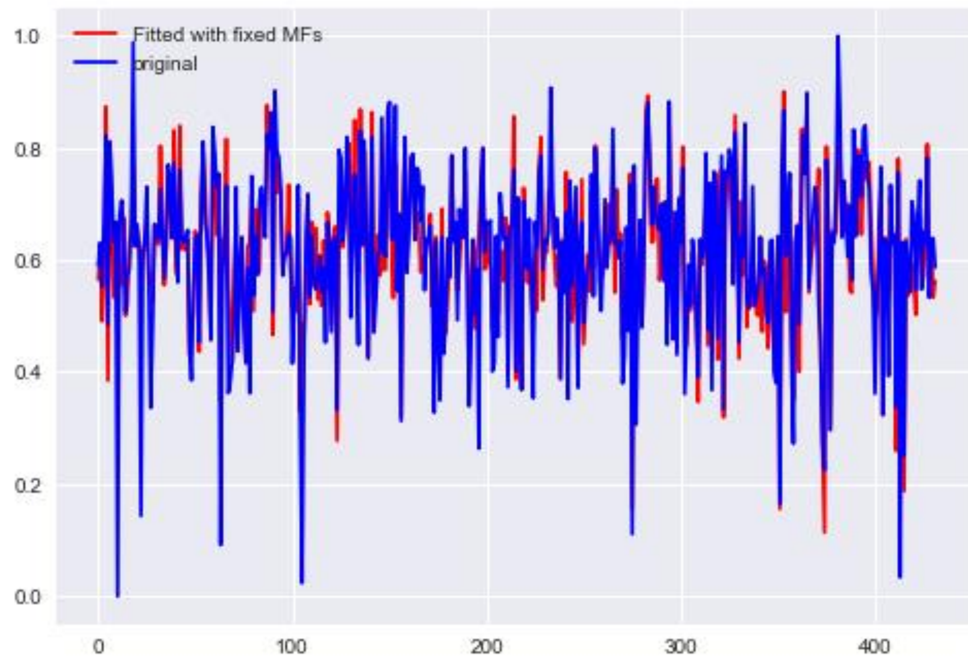


Figure 10 ANFIS fitted versus observed

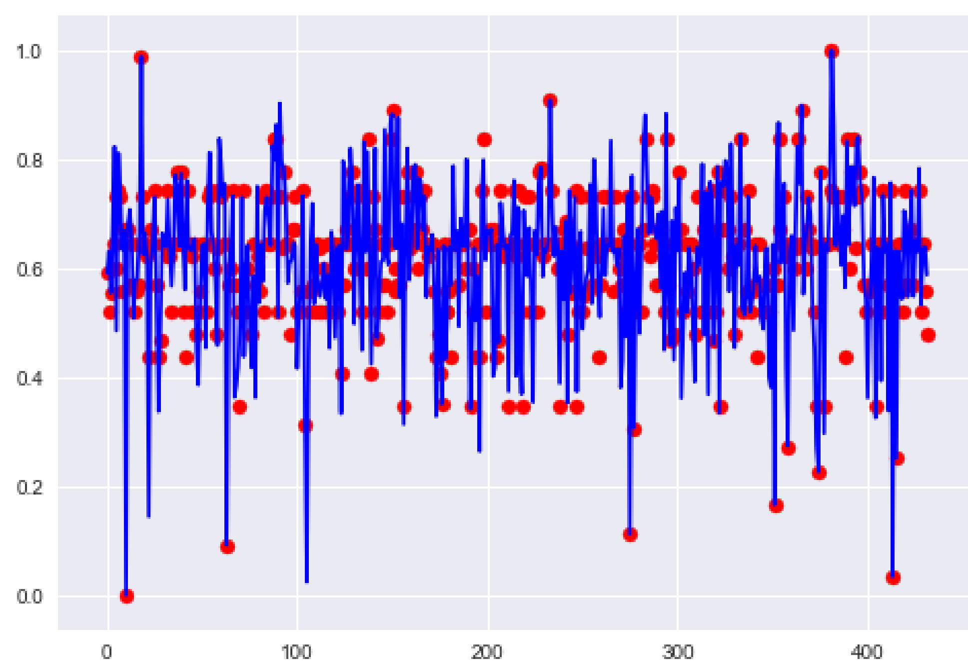


Figure 11 Decision Tree fitted versus observed

IV. Results

Model Evaluation and Validation

As discussed in Data Explanatory section, there are distinct regimes in the data. The total data points available are not as large (480). Thus, I randomly set aside 10% of the data for out-sample validation. This (48 months) is equivalent of four (4) calendar years, which is a sufficiently large testing set. The figures below show validation performance results for both champion and benchmark models. Note that while both models have performance in strong RMSE, champion model has much smaller Maximum Absolute Error (MAE). This measure ensures that in its worst case, the champion model does not deviate too far from ground truth. It is easy to see why Benchmark model has larger MAE, as the tree tends to give flat line predictions.

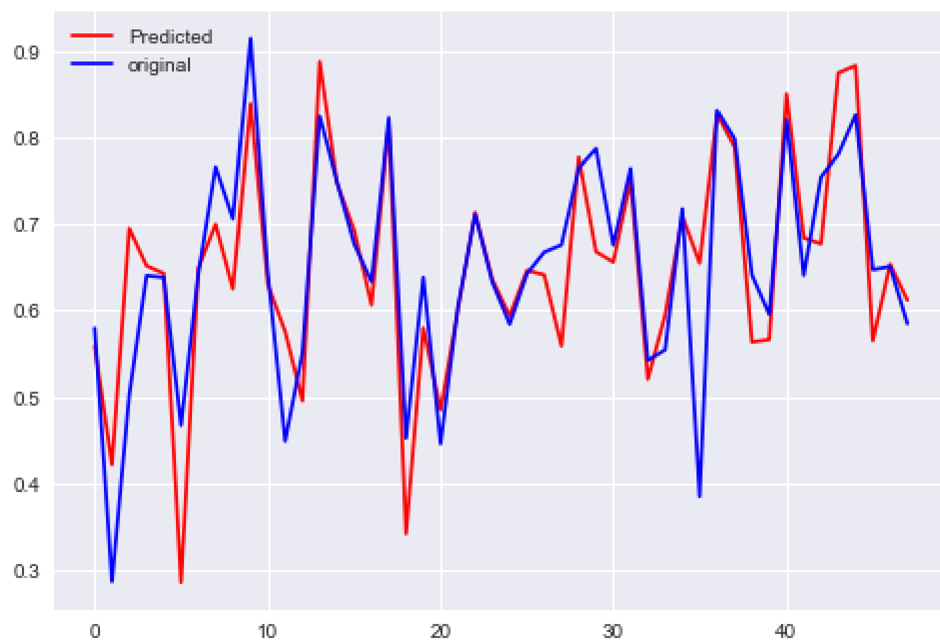


Figure 12 ANFIS validation performance (RMSE, MAE): 0.08 0.27

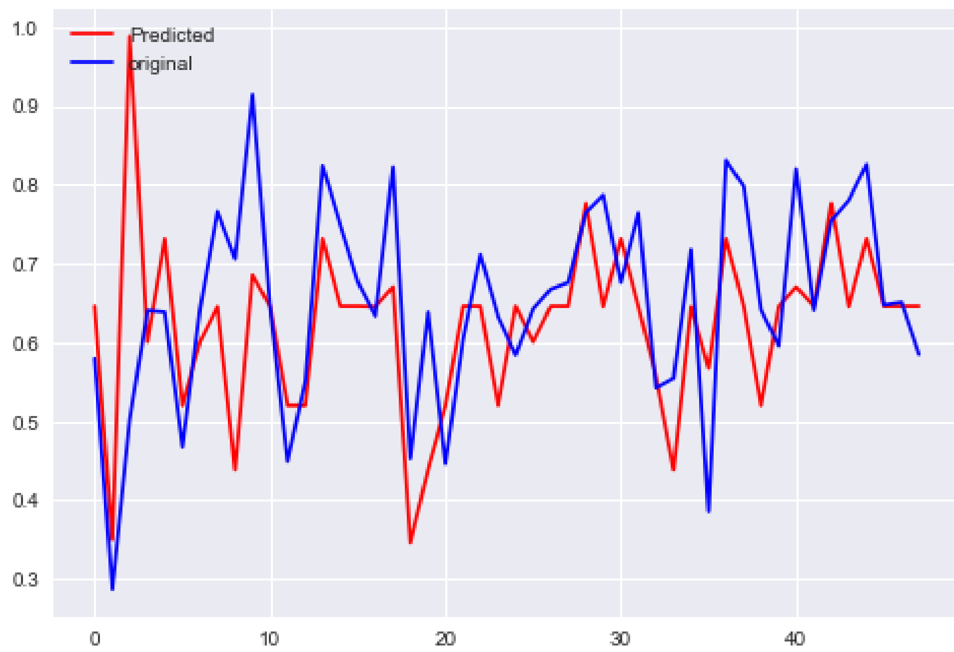


Figure 13 Decision Tree validation performance (RMSE, MA): 0.12 0.49

The Fuzzy model extracted the following high order rules. Recall that 'Loose', 'Neutral' and 'Tight' refer to the FED deviation from the 'fair' rule based on Taylor objective formula. The rules are consistent with empirical evidence shown on Figure 1, that the FED can systematically deviate from Taylor rule.

UNRATE	GS5	MICH	RATE
UNRATE=Low	GS5=Medium	MICH=Medium	RATE=Neutral
UNRATE=Medium	GS5=Medium	MICH=Medium	RATE=Neutral
UNRATE=High	GS5=Medium	MICH=Medium	RATE=Neutral
UNRATE=High	GS5=Medium	MICH=Low	RATE=Neutral
UNRATE=High	GS5=Medium	MICH=Medium	RATE=Tight
UNRATE=Medium	GS5=Medium	MICH=Medium	RATE=Tight
UNRATE=High	GS5=Low	MICH=High	RATE=Tight
UNRATE=Medium	GS5=Medium	MICH=High	RATE=Tight
UNRATE=Low	GS5=Medium	MICH=Medium	RATE=Tight
UNRATE=Medium	GS5=Medium	MICH=Low	RATE=Tight
UNRATE=Medium	GS5=Medium	MICH=Medium	RATE=Loose
UNRATE=High	GS5=High	MICH=Low	RATE=Tight
UNRATE=High	GS5=High	MICH=Low	RATE=Neutral
UNRATE=Medium	GS5=Low	MICH=High	RATE=Neutral
UNRATE=High	GS5=Low	MICH=Low	RATE=Neutral
UNRATE=High	GS5=High	MICH=High	RATE=Tight
UNRATE=Medium	GS5=High	MICH=High	RATE=Tight
UNRATE=Medium	GS5=Low	MICH=High	RATE=Tight

Table 1 Extract FED Rate Rules

The rules show that the FED must consider trade-off objectives. Based on these rules, conjoint analysis shows that 'High Inflation' regimes have the top relative importance. This is consistent with the FED's mandates and empirical events. The sample size is small (18) rules, just over 9 variables.

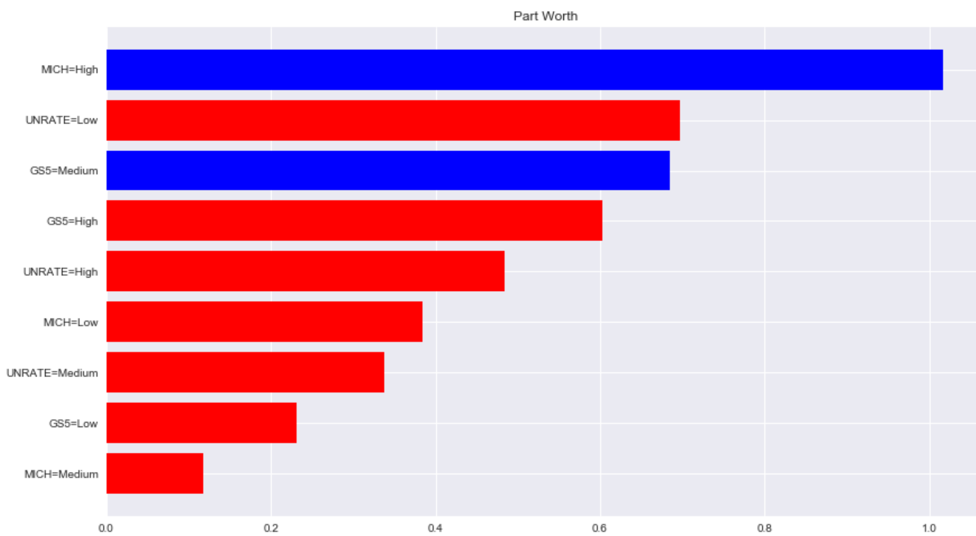


Figure 14 Relative feature importance, Conjoint Analysis

Justification

The champion model has stronger performance than benchmark, better interpretability. Its results are consistent with empirical and theoretical foundations. The model fulfills its objectives. It has strong crisp-value prediction capabilities and abilities to extract high order logic rules.

Solution Robustness

To evaluate the robustness of the solution. A set of one hundred (100) randomly selected validation is applied to the same model. Each test run generates its own performance measure in RMSE and MAE. The picture below shows the value distributions of these metrics.

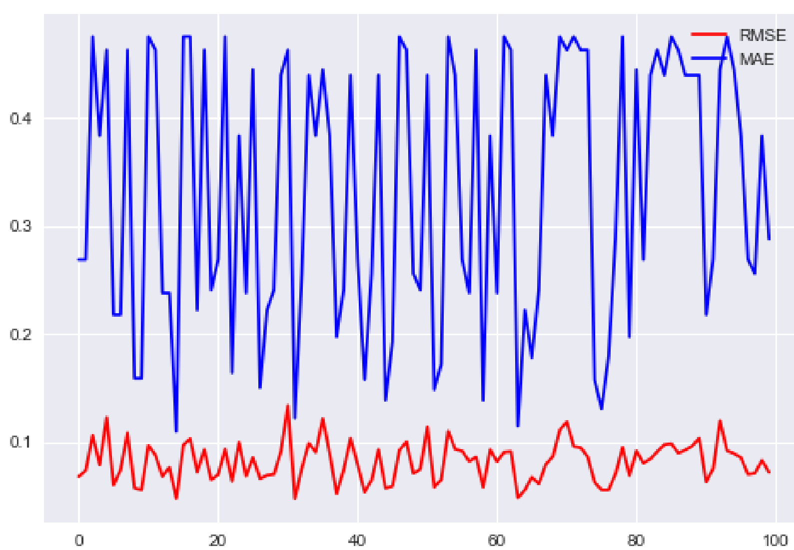


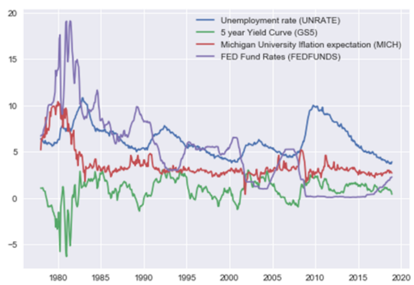
Figure 15 Performance metrics distribution

Performance result shows bounded, stable range, metrics. RMSE values are bounded to a very tight range (0.05, 0.13) and MAE in (0.11, 0.48). Result shows that model is stable.

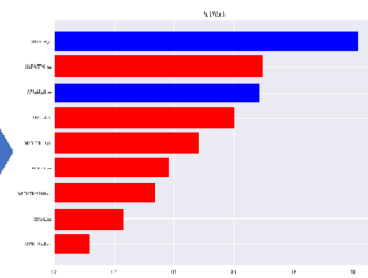
V. Conclusion

Free-Form Visualization

In my investigation, this is the first work to combine Machine Learning, Fuzzy Logic and Conjoint Analysis to derive models with higher order logic rules in financial economics. The model applies rigorous modeling techniques starting from set of noisy data, extracts high order logic Decision Rules, analyzes these rules to give order of decision variables importance.



UNRATE	GS5	MICH	RATE
UNRATE=High	GS5=Medium	MICH=Medium	RATE=Loose
UNRATE=High	GS5=Medium	MICH=Medium	RATE=Neutral
UNRATE=Low	GS5=High	MICH=Low	RATE=Neutral
UNRATE=Low	GS5=Low	MICH=Low	RATE=Neutral
UNRATE=Low	GS5=Medium	MICH=Low	RATE=Neutral
UNRATE=Medium	GS5=Medium	MICH=Medium	RATE=Neutral
UNRATE=Low	GS5=Medium	MICH=Medium	RATE=Neutral
UNRATE=High	GS5=Medium	MICH=High	RATE=Tight
UNRATE=Low	GS5=Medium	MICH=Medium	RATE=Tight
UNRATE=Medium	GS5=Medium	MICH=Medium	RATE=Tight
UNRATE=High	GS5=High	MICH=High	RATE=Tight
UNRATE=High	GS5=Medium	MICH=Low	RATE=Tight
UNRATE=Low	GS5=Low	MICH=High	RATE=Tight
UNRATE=Low	GS5=High	MICH=High	RATE=Tight
UNRATE=High	GS5=Low	MICH=High	RATE=Tight
UNRATE=High	GS5=Medium	MICH=Medium	RATE=Tight



One of the drawbacks of many Machine Learning models is weak interpretability. Applying Fuzzy Logic in addition to crisp value prediction is a promising approach[12]. The framework presented in this research can eb applied to other problems where it is important to understand not just what, but how the final decision is reached.

Reflection

Fuzzy Logic is not a mainstream approach. As result, there are few available ready to use packages. The skfuzzy library provides some functions, such as fuzzy clustering, but does not have the abilities to extract IF-THEN fuzzy rules from data.

Unsupervised Gaussian Mixtures Model is relatively simple in concept. To use GMM within Fuzzy Logic ANFIS framework comes with new challenges in clustering structure selection, result stability, and mean values initialization. I had to combine GMM/Fuzzy with other unsupervised clustering and cross validation to arrive at the optimized solution.

The dataset itself is another challenge. Multiple questions need to be solved. From simpler decision, such as over how many periods should moving average be, to how/if to deal with distinctive regimes in the data. The hardest part if to come up with a meaningful target output series. I arrived at the final from of deviation from Taylor Rule after a lot of experiments and thinking.

---

[12] https://arxiv.org/abs/1807.03215

Improvement

There are a few opportunities for improvement. The top three are:

1. Membership Function/ANFIS engine can be improved using TensorFlow/PyTorch. These engines have built-in optimization algorithms incorporating a lot of best practices.
2. More data series should eb considered than just the three above. Potential candidates include SP500 stock returns, VIX and financial stress indices. It is unlikely that the FED only consider three variables.
3. Experiment with different Fuzzy Inference approaches and algorithms, including those not yet in Python. I have tried the package 'frbs' in R[13] and found it quite comprehensive.

[13] https://cran.r-project.org/web/packages/frbs/index.html