



NUS
National University
of Singapore



Final Year Project – KE 25

Smart Trader – Phase 2

Report

Team:

Gao Yong	(A0107658X)
Sathasivaiyer Nadarajasarma	(A0107664A)
Shriranga Kulkarni	(A0107666Y)
Tang Hui-Fen, Elizabeth	(A0107689N)
Weiyang Lim	(A0107648Y)

Table of Contents

1	Project Endorsement Certificate.....	3
2	Project Overview.....	4
2.1	Problem Description	4
2.2	Project Scope and Objectives.....	4
3	Project Planning	5
4	Domain Familiarization and Knowledge Acquisition	6
4.1	Domain Familiarization	6
4.1.1	Domain familiarization steps – Iterative method:	6
4.2	Knowledge Acquisition/Elicitation:	8
4.2.1	About the Interviewee:	8
5	User Requirement Specifications.....	10
5.1	Executive Summary.....	10
5.1.1	Project Description.....	10
5.1.2	Purpose and Scope of this Specification	10
5.1.3	Benefits and Costs.....	11
5.1.4	Solution Outline	11
5.2	Product/Service Description	12
5.2.1	Product Context	12
5.2.2	User Characteristics	12
5.2.3	Constraints	13
5.3	Dependencies.....	13
5.4	Functional requirements.....	14
5.5	User Interface requirements.....	14
5.6	Performance Measurement.....	15
5.7	Maintenance	15
5.8	Data requirements.....	15
6	System Design	17
6.1	Business Background	17
6.1.1	Model Factors/Variables.....	17
6.1.2	Model Process.....	22
6.1.3	Model Outcome	23
6.1.4	Data Exploration and Pre-processing.....	23
6.2	Detailed Modelling and Algorithms	28

6.2.1	Ensemble.....	28
6.2.2	Adaboost.....	28
6.2.3	Neural Network.....	33
6.3	System Architecture.....	35
6.3.2	Core Components	35
6.4	Solution Technology.....	37
6.4.1	Technology Candidacy	37
6.4.2	Technology Selection	37
6.5	Development Methodology.....	37
6.6	Development Platform, Framework and Tool	37
7	Progress Summary	38
7.1	Tasks accomplished as part of Phase 2:.....	38
7.2	Tasks planned in Pipeline for Phase 3:.....	38
8	Appendix A: Findings and Results (Comparison of various models – Sector by Sector).....	39
8.1	Sector – Consumer Discretionary	39
8.2	Consumer Staples	45
8.3	Sector - Energy	50
8.4	Energy – HealthCare	55
8.5	Sector – Industrials	60
8.6	Sector – Information Technology.....	65
8.7	Sector – Materials.....	70
8.8	Sector – Telecommunications.....	75
8.9	Sector – Utilities	80
9	Appendix B: System Design - US Stock Knowledge Model	85
10	Appendix B: Knowledge Acquisition Transcript	129
11	Appendix C: Project Schedule	133
12	References	134

1 Project Endorsement Certificate

Following is the letter of endorsement from PiQuant Capital Pte. Ltd
(<http://www.piquantcapital.com>).

We extend our thanks to PiQuant members and Weying Lim for the project endorsement.

To: Director (Master of Technology in Knowledge Engineering programme)
Institute Of Systems Science
Eusoff Hall, 25 Heng Mui Keng Terrace, 119615

Subject: Final Year Project for Master in Technology (Knowledge Engineering) program

The team of students from KE25 listed below will undertake a research and development project, possibly leading to an implementable system on behalf of PiQuant Capital Pte Ltd.

Lim Weiyang
Elizabeth Tang Huifen
Gao Yong
Shriranga Kulkarni
Nadarajasarma Sathasivaiyer

The topic of the project will relate to stock market trading in the U.S. market and seek to put into practice the skills, tools and techniques acquired by the students. The students will have the opportunity to work with the domain expert and understand the strategies in the organization and, suggest and apply knowledge-engineering techniques to provide identifiable benefits. The project may involve the experimental application and validation of several knowledge-engineering technologies such as random forest, clustering, neural net forecasting, ensemble methods etc to analyzing and predicting the stock market.

Piquant Capital Pte Ltd will endorse the objectives and strategies of the above Project and express its commitment to participate in this Project.

Sincerely,



Signed by



CEO of PiQuant Capital



CIO of PiQuant Capital
(GEORGE VARGHESE)

PIQUANT CAPITAL PTE. LTD.
(Company Registration No: 201105998C)

2 Project Overview

2.1 Problem Description

To develop an intelligent software system capable to

- Generate a portfolio of US stocks information on any particular day based on the S&P1500,
- Help the traders/portfolio managers to make informed decisions about selecting stocks to generate their desired absolute return.

2.2 Project Scope and Objectives

Overall build a dynamic and robust US stock trading system based on a multi-factor model that is systematic, rigorously and carefully tested across time and thus giving consistent and profitable returns in real life that can closely match the back test. Both long and short positions should be considered and combined wherever possible.

1. Customizable based on user's risk profile and style preference through a user interface. This can be done through the use of appropriate screens combining various factors/metrics, leverage, number of stocks chosen (though this has to be of a sufficient size to achieve diversification and model efficacy) etc.
2. Use of appropriate KE techniques to enhance performance if possible though not at the expense of sacrificing the intuition (through appropriate use of factors) behind what drives stock performance

3 Project Planning

Project plan is created and maintained via “Gantter” online which is powered by Google Docs, facilitating all team members to adhere to the schedule and progress collaboratively.

Please refer to Appendix C for the full project plan.

4 Domain Familiarization and Knowledge Acquisition

4.1 Domain Familiarization

The domain familiarization is an iterative and on-going process. The team members perform it both individually and collectively in order to get familiar with the domain and understand the various concepts involved in the domain.

4.1.1 Domain familiarization steps – Iterative method:

1. Formulation of an initial set of stock level data needed.
2. Downloading the stock data at regular interval of time, to get familiar with the data.
3. Exploratory data analysis to understand the data better.
4. Data cleaning/factor scaling and modeling.
5. Appropriate scoring/normalization and relative ranking of stocks by each factor score and grouping into respective portfolio quadrants. This factor scoring has to be sector specific/even capitalization specific (each sector has its own characteristics) and based on what we understand about each factor.
6. Generate daily factor returns for all portfolio quadrants for each factor. The best quadrant should preferably perform best over time.
7. Combine and apply appropriate weightings to individual factors, then form broad composite factors. A final factor can be determined and “factor returns” can again be generated for this. To form a final factor, different broad composite factors can be combined at different periods based on some clustering technique. However, this should be approached carefully as factor rotation is not an easy task to achieve. If possible, weightings of composite factors can change across time. Otherwise, equal weighting usually works quite well. There can be multiple combinations of broad composite factors that can be constructed that can perform well at different times. This has to be researched and tested.
8. Determine market favorability. Can be primarily based on stock/sector level aggregate fundamental/technical information. Economic/Monetary policies/Sentiment data can be used wherever appropriate. This can help determine the level of risk to take at any point in time.
9. Portfolio construction/implementation. This is where appropriate costs and actual rebalancing of the portfolio/execution of trades has to be considered. Can be a fixed duration such as monthly or dynamic rebalancing. Appropriate leverage, net long/net short

positions (if we also consider short positions) or quantity of stocks to recommend is also important based on appropriate risk to take.

Please refer to appendix A for the Domain Familiarization model details.

4.2 Knowledge Acquisition/Elicitation:

Knowledge acquisition and knowledge representation are the fundamental building blocks of knowledge-based systems (KBSs). How to efficiently elicit knowledge from experts and transform this elicited knowledge into a machine usable format is a significant and time consuming problem for KBS developers. Object-orientation provides several solutions to persistent knowledge acquisition and knowledge representation problems including transportability, knowledge reuse, and knowledge growth.

One of the initial and important processes in building the expert system is the Knowledge Acquisition for stock analysis and recommender in a certain area of interest. Since we had to get domain knowledge in the field of stocks, finance and related technology, we went through following procedures for accomplishing the knowledge elicitation process.

- 1) Primarily, interviewing Mr Yingwei Heng (Stanley Wang), a domain expert in the field of stocks trading and finance in general.
- 2) Discussions with Mr. Johnny Kwon, Investment manager for Woodsford Capital Management Ltd, Singapore.
- 3) Also do our own research in the stocks investment and financial domain using online/library resources.
- 4) Take guidance/assistance of our professor, Mr. Charles (for technical feasibility and clarification regarding implementation ways).

4.2.1 About the Interviewee:

Stanley Wang has a wealth of experience in the finance industry having worked in banks for a few years.

- He has considerable experience in trading equities, forex and warrants having traded for years.
- He has taken part and came in second in the Singapore round of an equity research competition organized by the CFA Institute.
- He believes that one can predict future market trend by analyzing historical data.

The interview was conducted online using Google hangouts air application and the further clarifications were discussed via mobile/email. The interview video was uploaded onto YouTube (<https://www.youtube.com/watch?v=UjAOcsTF1Aw>) with detailed transcript.

We followed mixture of unstructured interview/Semi-structured interview techniques for the domain knowledge acquisition.

During the early stage of the interview, elicitation followed **unstructured interview mode**, where session was free flowing which produced basics of knowledge domain [basically a broad discussion about the stocks and finance domain].

For later/final stages of the interview, **Semi-structured interview** mode was followed which is the main technique for elicitation, In this mode, Pre-defined questions were sent to expert prior to interview, supplementary questions asked at the interview. We followed this format since outcome of this kind of interview, can be used as part of validation.

Apart from the Interview with domain experts, we carried out our own research in finding the information about different aspects of stocks investment. Please refer to Reference section of the report for online references/resources used for the knowledge acquisition.

Please refer to appendix B for the summary of Knowledge acquisition transcripts.

5 User Requirement Specifications

5.1 Executive Summary

5.1.1 Project Description

Portfolio equity managers/traders are in the business of managing clients' assets and need to generate reasonable returns to satisfy their clients and generate fees for themselves. They perpetually face difficult decisions such as which stocks to buy/sell, when to buy/sell and how much capital to commit at any point in time. They need to understand the possible underlying factors that drive both the stock market as a whole and individual stock returns. The ability to determine which factors work and when they work can give the portfolio manager/trader an edge over his competitors.

5.1.2 Purpose and Scope of this Specification

The aim of this project is to develop an intelligent and dynamic software system which will generate a portfolio of US stock information from the S&P1500 based on a set of most relevant and important explanatory factors and accompanying models on any particular day, and help the traders/portfolio managers to make informed decisions about selecting the most ideal stocks to generate their desired absolute return. Below show the various scope of the project.

- Output will be a list of stocks which are expected to give high return relative to other stocks within universe on a particular day.
- User has the control over the number of maximum stocks in the list and the minimum possibility for the prediction.
- Factors selected by the models used in stock selection at any point in time and possible importance/explanatory ability.

5.1.2.1 Possible Extensions

- Integration with trading systems for real time trading of stocks, based on the historical models and results generated. Incorporate costs, use appropriate/realistic holding periods and risk control mechanisms and back-test strategy over historical period
- Incorporate market favorability/direction which can possibly be based on a predictive model and advice on stock allocation accordingly.

- User preference over the selection of factors.
- Other stock selection methods based on various styles such as value or dividend investing, growth investing and trend/momentum trading etc.

5.1.3 Benefits and Costs

Artificial intelligence and knowledge systems technology first surfaced in the latter part of the last century when they were used to predict market trends. Since then, there has been limited literature or widespread use of such technologies in the financial markets due to various reasons such as a lack of trust in its predictive accuracy or a greater belief in individual judgment rather than using the computer to make a trading decision.

However, the benefits of using these technologies are multifold. First, the technologies and models used have become increasingly comprehensive and therefore complicated, with the ability to synthesize a large amount of data particularly with comprehensive stock level information involving a large universe such as S&P1500 and multiple attributes that the human being will be unable to perform easily by himself.

Second, non-linear decision making computer programs that attempt to make the “best” decision are attempting to more closely model the type of decision making used by humans in an effort to better reflect our comprehensive problem solving capabilities. This may eventually result in the closest we have come to technologically reproducing human logic and intelligence. Third, these systems can be dynamic and evolve constantly with the markets while being able to learn new models and relationships between data with the arrival of new financial data.

There are also possible disadvantages of a system based on artificial intelligence and predictive modeling. There are limitations as to how far these technologies can go because of the inseparable human element in the fluctuations in the stock market and how well these fluctuations can be modeled accordingly. In addition, it can be difficult to use such technologies appropriately and effectively without a solid understanding.

5.1.4 Solution Outline

The ability to accurately determine the crucial factors affecting stock returns at any point in time, their importance and how to model them effectively can give the portfolio manager an edge. The portfolio manager/domain expert’s role is to determine, based on his understanding of the stock

market what such factors can be. Due to a multitude of explanatory factors and the difficulty in understanding precisely the mathematical relationship between factors and stock returns, artificial intelligence or knowledge systems technology can then be applied to better attempt to model or make sense of such relationships, possible in cases where there is non-linearity and some randomness.

KE techniques that will be explored include linear regression, decision trees, adaboost, support vector machines, neural networks. These methods will be used on the resulting set of factors (through a technique like random forest and other methods like correlation analysis etc) and used to predict individual stock returns to understand which method may be more effective.

Clustering is another method that can be used on the entire set of factors to analyze relationships between stock returns and factor attributes.

Allowing the change of weights to vary across time for different composite factor groups such as valuation, profitability, price related attributes can be explored as well.

5.2 Product/Service Description

As this is a prediction system, and the result of the prediction is based on a set of underlying explanatory factors, the output produced may differ from actual results depending on the unforeseen events in the market. System will be tested with the past data and validated for maximum reliability. The accuracy of the system is dependent on the real time data collected as trained data.

5.2.1 Product Context

The system is an independent system having no direct dependencies with other systems such as live stock data feed or brokerage systems. The analysis results produced by systems. This system may be used as input for trading strategies which are currently out of scope of the system.

5.2.2 User Characteristics

This system has two types of users.

- Admin

This mode is used to train the system and modify the core functionalities of the system.

Only the system developer and system administrator have the privileges of Admin user.

- End User

This mode is used by Traders/Portfolio Managers to generate the portfolio every day.

This is the target mode (end user mode) of this system.

5.2.3 Constraints

Because of the huge data to be analyzed to generate the relevant models and resulting portfolios, a significant processing time and processing power will be required.

5.3 Dependencies

- This system will require a daily download of data from the market data server.
- Downloading the data and transferring to appropriate place to be completed before starting this process.

5.4 Functional requirements

Requirement No.	Requirement	Comments	Priority	Reviewed/Approved
1	Data Availability	Downloading data from servers using appropriate tools/scripts	5	
2	Data Verification	Verifying the data for possible errors and missing values.	3	
3	Data Cleanup and preparation	Clean the data.	2	
4	Data Standardization	Rescale and Normalize.	2	
5	Data Analysis and Model Design and Implementation	Data mining and modeling	1	
6	Prediction	Predicted stocks with higher return.	1	

1 - Highest priority, 6 - Lowest priority

5.5 User Interface requirements

User has control over the following items:

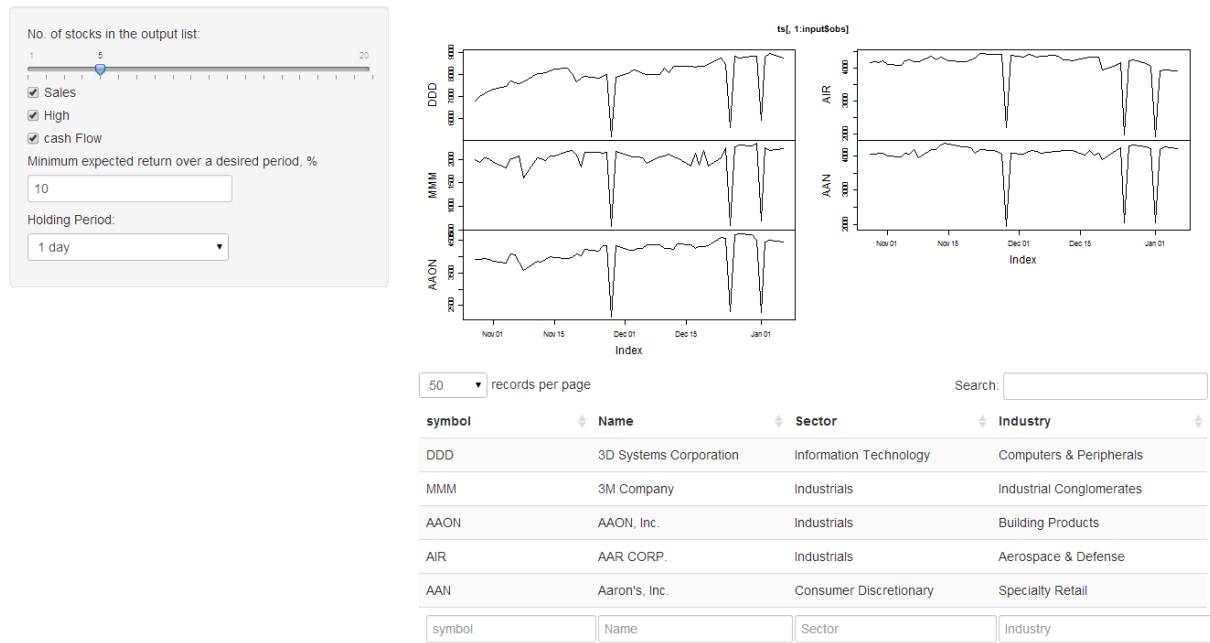
1. No. of stocks in the output list.
2. Minimum expected return over a desired period
3. Number and/or Type of variables to include/exclude for the prediction. (Possibly add on and for advanced users)

Possible Outputs:

1. List of stocks
2. Graphs
3. Variables analyzed, analyzed periods and details.

Schematic Representation of SmartTrader Output

SmartTrader



5.6 Performance Measurement

Performance will be measured as below:

1. Number of records processed and analyzed per unit time
2. System throughput

5.7 Maintenance

1. All the technical design and source code should be properly documented.
2. System should have interfaces to scale/extend in future.
3. Different levels of logging mechanisms are essential.

5.8 Data requirements

Requirement	Description
Types of information used by various functions	Market data ticks.
Frequency of use	Everyday

Data Security	Data should be protected from unauthorized access.
Data entities and relationships	Relationship and association between different data objects and variables should be designed
Integrity constraints	Data integrity refers to maintaining and assuring the accuracy and consistency of data over its entire life-cycle
Data retention	N/A
Valid range, accuracy, and/or tolerance	Range is different for different variables.
Units of measure	N/A
Data formats	CSV/XML, MSDB, or other applicable types
Default or initial values	N/A

6 System Design

6.1 Business Background

6.1.1 Model Factors/Variables

6.1.1.1 Objective Variable

Future stock returns over a desired holding period (a few possible selections such as 10, 20, 30 days) from the point where underlying predictive factor values are determined. Depending on which algorithm is used to train the model, this variable will be processed and labelled accordingly.

6.1.1.2 Predictive Variables

Following table illustrates the list of possible factors or variables which can be used to knowledge exploration, data discovery pattern and analysis and predictive modelling:

No	Factor	Category	Period
1	Book Yield FY1	Valuation	Forward Looking
2	Book Yield FY2	Valuation	Forward Looking
3	Book Yield FY3	Valuation	Forward Looking
4	Dividend Yield FY1	Valuation	Forward Looking
5	Dividend Yield FY2	Valuation	Forward Looking
6	Dividend Yield FY3	Valuation	Forward Looking
7	Earnings Yield FY1	Valuation	Forward Looking
8	Earnings Yield FY2	Valuation	Forward Looking
9	Earnings Yield FY3	Valuation	Forward Looking
10	EBITDA Yield FY1	Valuation	Forward Looking
11	EBITDA Yield FY2	Valuation	Forward Looking
12	EBITDA Yield FY3	Valuation	Forward

			Looking
13	Free Cash Flow Yield FY1	Valuation	Forward Looking
14	Free Cash Flow Yield FY2	Valuation	Forward Looking
15	Free Cash Flow Yield FY3	Valuation	Forward Looking
16	Sales Yield FY1	Valuation	Forward Looking
17	Sales Yield FY2	Valuation	Forward Looking
18	Sales Yield FY3	Valuation	Forward Looking
19	EV_EBIT Mean Forward 12M	Valuation	Forward Looking
20	EV_EBITDA Mean Forward 12M	Valuation	Forward Looking
21	EV_Free Cash Flow Mean Forward 12M	Valuation	Forward Looking
22	EV_Sales Mean Forward 12M	Valuation	Forward Looking
23	PEG Mean FY1	Valuation	Forward Looking
24	Earnings Yield (Past 1 Year)	Valuation	Historical
25	Free Cash Flow Yield (Past 1 Year)	Valuation	Historical
26	Sales Yield (Past 1 Year)	Valuation	Historical
27	EV_EBIT (Past 1 Year)	Valuation	Historical
28	EV_EBITDA (Past 1 Year)	Valuation	Historical
29	Book Value FY1 Revision % (Past 3M)	Analyst Revision	Forward Looking
30	Book Value FY1 Revision % (Past 6M)	Analyst Revision	Forward Looking
31	Dividend FY1 Revision % (Past 3M)	Analyst Revision	Forward Looking
32	Dividend FY1 Revision % (Past 6M)	Analyst Revision	Forward Looking
33	Earnings FY1 Revision % (Past 3M)	Analyst Revision	Forward

			Looking
34	Earnings FY1 Revision %(Past 6M)	Analyst Revision	Forward Looking
35	EBITDA FY1 Revision % (Past 3M)	Analyst Revision	Forward Looking
36	EBITDA FY1 Revision %(Past 6M)	Analyst Revision	Forward Looking
37	Free Cash Flow FY1 Revision % (Past 3M)	Analyst Revision	Forward Looking
38	Free Cash Flow FY1 Revision %(Past 6M)	Analyst Revision	Forward Looking
39	Sales FY1 Revision % (Past 3M)	Analyst Revision	Forward Looking
40	Sales FY1 Revision %(Past 6M)	Analyst Revision	Forward Looking
41	Earnings FY1 Estimate Variability	Analyst Variation	Forward Looking
42	Earnings FY2 Estimate Variability	Analyst Variation	Forward Looking
43	Earnings FY3 Estimate Variability	Analyst Variation	Forward Looking
44	EBITDA FY1 Estimate Variability	Analyst Variation	Forward Looking
45	EBITDA FY2 Estimate Variability	Analyst Variation	Forward Looking
46	EBITDA FY3 Estimate Variability	Analyst Variation	Forward Looking
47	Earnings FY1 UpDngrade(3M)	Analyst Grading	Forward Looking
48	Earnings FY2 UpDngrade(3M)	Analyst Grading	Forward Looking
49	Earnings FY3 UpDngrade(3M)	Analyst Grading	Forward Looking
50	Earnings FY1 UpDngrade(6M)	Analyst Grading	Forward Looking
51	Earnings FY2 UpDngrade(6M)	Analyst Grading	Forward Looking

52	Earnings FY3 UpDngrade(6M)	Analyst Grading	Forward Looking
53	EBITDA FY1 UpDngrade(3M)	Analyst Grading	Forward Looking
54	EBITDA FY2 UpDngrade(3M)	Analyst Grading	Forward Looking
55	EBITDA FY3 UpDngrade(3M)	Analyst Grading	Forward Looking
56	EBITDA FY1 UpDngrade(6M)	Analyst Grading	Forward Looking
57	EBITDA FY2 UpDngrade(6M)	Analyst Grading	Forward Looking
58	EBITDA FY3 UpDngrade(6M)	Analyst Grading	Forward Looking
59	Earnings FY1 Analyst Coverage	Analyst Coverage	Forward Looking
60	Gross Margin Forward 12M	Profitability	Forward Looking
61	EBITDA Margin MeanForward 12M	Profitability	Forward Looking
62	EBITMargin Mean Forward 12M	Profitability	Forward Looking
63	Net Margin Mean Forward 12M	Profitability	Forward Looking
64	ROIC Mean FY1	Profitability	Forward Looking
65	Asset Turnover Mean Forward 12M	Profitability	Forward Looking
66	Net Debt/Ebitda Mean Forward 12M	Profitability	Forward Looking
67	Capex_Sales Mean Forward 12M	Profitability	Forward Looking
68	Dividend Payout	Profitability	Historical
69	Return on invested capital	Profitability	Historical
70	Cash flow return on invested capital	Profitability	Historical
71	Debt/Capital	Quality	Historical
72	Debt/Total Assets	Quality	Historical

73	Debt/EBITDA	Quality	Historical
74	Operating Cash Flow/Financing Cash Flow (Past 1 Year)	Quality	Historical
75	Operating Cash Flow/Earnings (Past 1 Year)	Quality	Historical
76	Cash Dividend / Cash Flow from Operations	Quality	Historical
77	Net Cash/Total Debt (Past 1 Year)	Quality	Historical
78	Change in Shares Outstanding (Past 3 months)	Quality	Historical
79	Change in Shares Outstanding (Past 6 months)	Quality	Historical
80	Change in Shares Outstanding (Past 12 months)	Quality	Historical
81	EPS_Long Term Growth Mean	Growth	Forward Looking
82	EPS FY1 Slope (Past 6 Months)	Growth	Forward Looking
83	EPS Actual - slope (Past 6 Months)	Growth	Historical
84	EPS Actual - slope (Past 12 Months)	Growth	Historical
85	EPS Actual - slope (Past 36 Months)	Growth	Historical
86	EPS Actual - slope (Past 60 Months)	Growth	Historical
87	EPS Actual - slope over std (Past 36 Months)	Growth	Historical
88	EPS Actual - slope over std (Past 60 Months)	Growth	Historical
89	Sales Per Share Actual - slope (Past 6 Months)	Growth	Historical
90	Sales Per Share Actual - slope (Past 12 Months)	Growth	Historical
91	Sales Per Share Actual - slope (Past 36 Months)	Growth	Historical
92	Sales Per Share Actual - slope (Past 60 Months)	Growth	Historical
93	Sales Per Share Actual - slope over std (Past 36 Months)	Growth	Historical
94	Sales Per Share Actual - slope over std (Past 60 Months)	Growth	Historical
95	Net Margin Actual - slope over std (Past 12 Months)	Growth	Historical
96	Net Margin Actual - slope over std (Past 36 Months)	Growth	Historical
97	Price Slope (Past 20 Days)	Technical	Historical
98	Price Slope (Past 50 Days)	Technical	Historical
99	Price Slope (Past 100 Days)	Technical	Historical
100	Price Slope (Past 200 Days)	Technical	Historical
101	Price Slope Over std (Past 20 Days)	Technical	Historical
102	Price Slope Over std (Past 50 Days)	Technical	Historical
103	Price Slope Over std (Past 100 Days)	Technical	Historical
104	Price Slope Over std (Past 200 Days)	Technical	Historical
105	Money Flow (14 Days)	Technical	Historical
106	Money Flow (50 Days)	Technical	Historical
107	RSI (14 Days)	Technical	Historical

108	RSI (50 Days)	Technical	Historical
109	Price Momentum (20 Days)	Technical	Historical
110	Price Momentum (50 Days)	Technical	Historical
111	Price Momentum (100 Days)	Technical	Historical
112	Price Momentum (200 Days)	Technical	Historical
113	Price/200D Moving Average	Technical	Historical
114	Price/52 week high	Technical	Historical
115	Price/52 week low	Technical	Historical
116	Short Interest (%)	Technical	Historical

6.1.2 Model Process

6.1.2.1 Initial steps

Following steps are formulated to cover the Data Model and Analysis Process:

- As a first step, data files containing underlying factors will be downloaded at regular intervals of time (end of the day), historically for a period of 5-10 years.
- Factor values will evolve on a daily basis across the universe of stocks.
- As factors input values change every day, the model will dynamically evolve over time too. A model will be created either daily or at a fixed duration such as weekly or monthly.
- Data processing and the use of techniques in the model will however remain consistent over time so that the model can be evaluated at different stages over time. This is to ensure that the model keeps up to date (machine learning technique) with constantly evolving market information, always selecting the best set of factor inputs for modeling.

6.1.2.2 Possible data models

We believe that the data that goes into the algorithms matter as much if not more so than the algorithms being used. Hence, we have to decide the different aspects of the data that are used as input into the algorithms. Below are various components that determine which data to use as input for the algorithms. Each of the components can be used on its own or combined.

Moving window

A moving window of data since the start of the entire period should be used for training and testing to take into consideration the dynamic nature of the stock market over time. What length of window should we use? On one hand, using a longer rolling window will give us more training data and thus can increase the model accuracy, but on the other hand using longer a rolling window introduces more out-of-date data which might not be consistent with the current financial environment. We

choose a window of 12 months as a good cutoff point to train our model. For example, a first model to train will be to use stock input variables at the beginning of each month between January and December 2004 and use forward 20 day return just one day after the corresponding beginning of each month as the output variable.

Sector differentiation

A second variation of the model is to create various sector-based models, i.e. each sector will have its own model with the corresponding window used for training.

Seasonality differentiation

A third possible enhancement is to use seasonal based models. This is based on the idea that stocks can behave quite differently during different months of the year. Therefore, for each month across the training period (say between 2004-2009), we can aggregate cross sectional stock input factor data with their corresponding forward return and train a monthly model. The result is 12 monthly models for the available stock universe.

6.1.3 Model Outcome

“Future returns” of each stock for a specified duration will be predicted.

The duration used will be 20 days such that there will be sufficient time for stocks to perform and to allow for more realistic trading including costs. The stocks with the best predicted returns will be evaluated for how they perform in real life.

The accuracy rate will be calculated based on how well the selected stocks actually performed in returns terms compared to the average stock in the selected universe. If it beats the return of the average stock, it is considered as being successfully predicted. The accuracy rate can be determined by the percentage of successfully predictions.

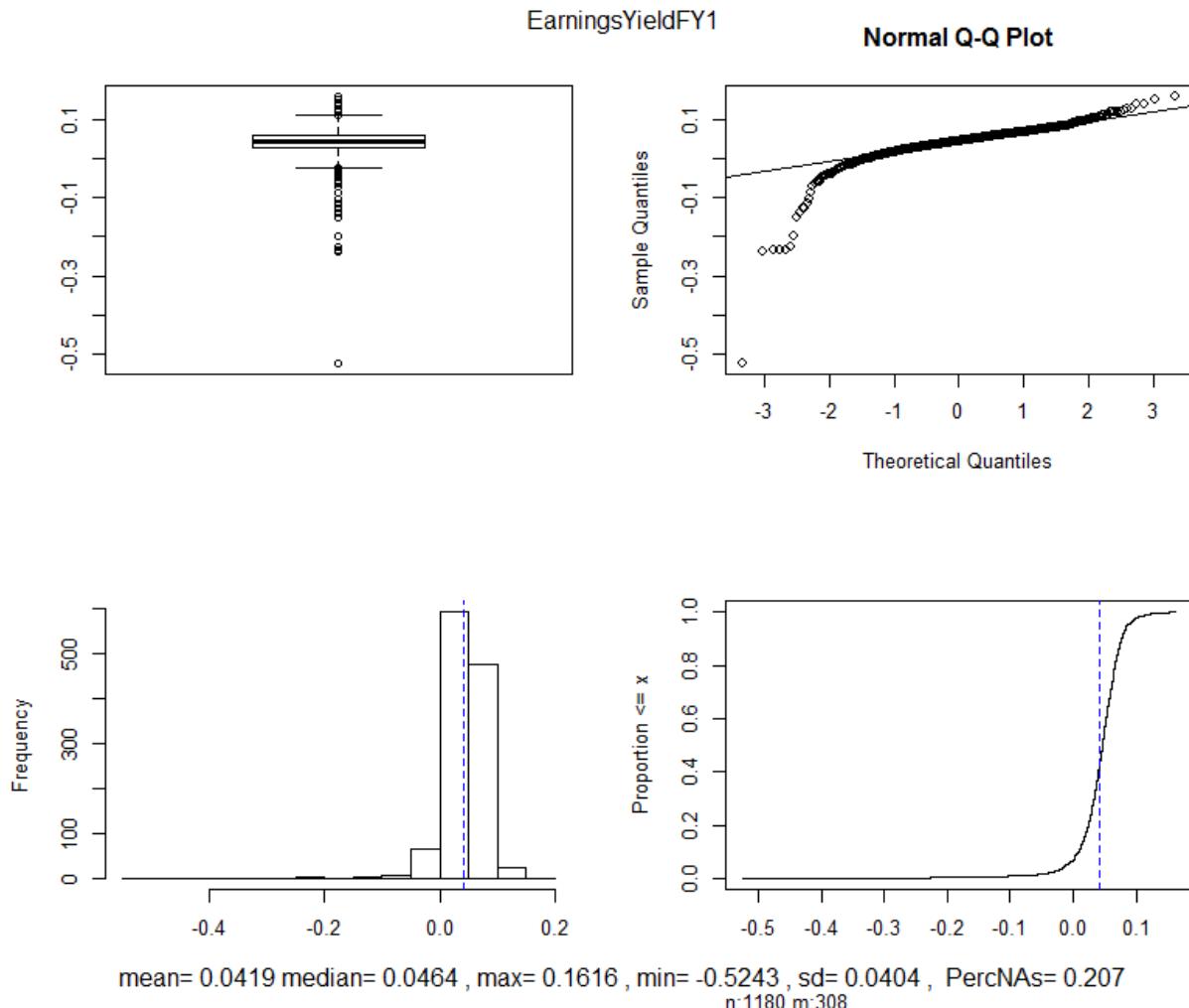
6.1.4 Data Exploration and Pre-processing

Two tables were created to store various summary statistics. First is for each factor for **various market capitalizations** (Large, Mid and Small Cap) on **each day** across time. Second is for each factor for **each sector** (for all 10 sectors) on **each day** across time.

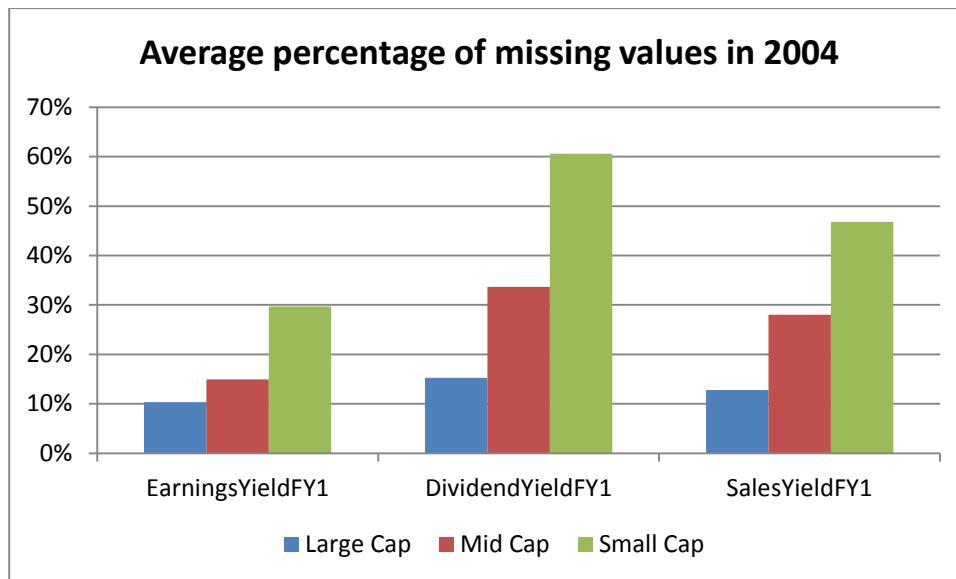
The summary statistics include the mean, the median, the maximum, the minimum, standard deviation, median absolute deviation and percentage with NAs/missing values, lower and upper limits with two mean absolute standard deviations.

This will allow us to understand the characteristics of stocks in various sectors and market capitalizations.

Various summary plots were also created on a single day as shown below. They include box plots, QQ-plots, histograms and cumulative density plots. Below shows the plots for the factor EarningsYieldFY1 for all stocks for Jan 15, 2004. This was done before any data pre-processing was carried out.



To determine whether we can use all the stocks in the S&P1500, we examine the number of missing values for various important factors. Below shows the average percentage of missing values for three of these factors in year 2004. The year 2004 was used as this is one of the earliest years of data where data might be less readily available. It can be observed that large cap stocks have the least number of missing values followed by mid cap and small cap stocks. As small cap stocks have a large number of missing values and they are less similar in characteristics compared to large and mid-cap stocks, we decide to omit these stocks for consideration. We include the mid cap names as well as we will like to have a sufficient number of observations for our modelling.



6.1.4.1 Handling of outliers/missing values

A uniform approach is applied to handle outliers and missing values for all input variables and all stocks across time. The reason is that the tremendous amount of data over time does not allow us to specifically examine different periods and have different rules for achieving this purpose at any point in time. Also, such an approach allows for better comparison of results over time.

Handling of outliers

An outlier is determined as being three median absolute deviations above or below the median. Such outliers will be restricted within a threshold of +/- three median absolute deviations. The threshold was set to be loose primarily only to curtail the extreme outliers and to give due consideration to values which deviate significantly from the median but within the outlier limits but are of predictive value.

We use the median absolute deviation instead of standard deviation as we do not want the mean to be unnecessarily skewed by outlier values.

Handling of missing values

This was done via a three step approach on any trading day as described below.

1. Remove stocks

First, we determine if any stock needs to be removed for prediction if it has too many missing values for its input variables on any particular trading day. To achieve this, a threshold for the percentage of missing values across all input variables for each stock on any trading day was set. This was arbitrarily set as 30% as a stock observation is only considered meaningful for prediction if it has at least a percentage of input variable values that are available. Otherwise it will be removed for

prediction for that trading day. Alternately, if a stock satisfies the threshold on other trading days, we will still use it for prediction for those days.

2. Avoid using and normalizing input variables

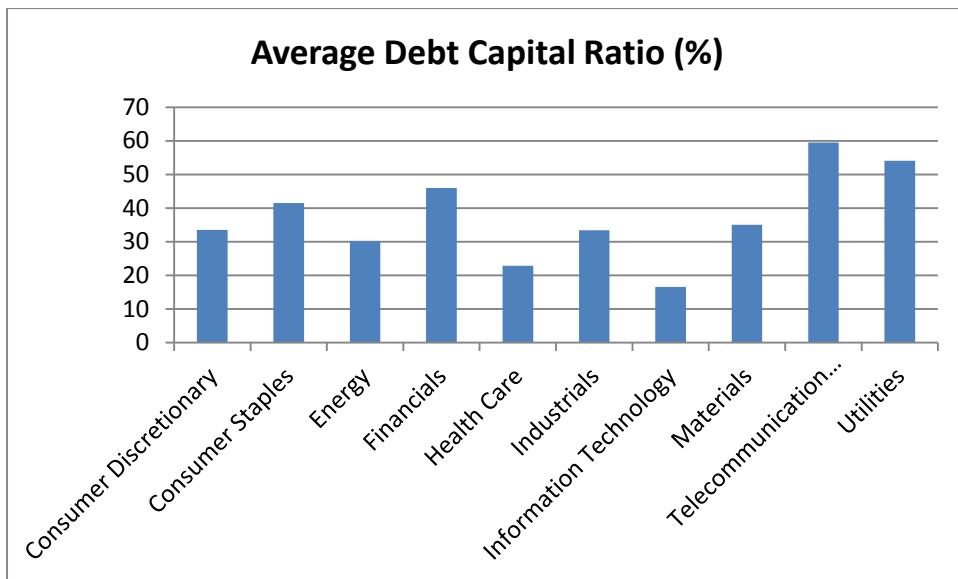
Next, we determine if any input variable should be removed for consideration on any trading day. The input variable will be considered meaningful for prediction if it has a minimum percentage of available observations. We arbitrarily determine the threshold for the percentage of missing values to be 30%. If the input variable has less than 70% of available values on any trading day, we will not use the input variable for prediction on that trading day. Alternately, if the input variable satisfies the threshold on other trading days, it will still be used for prediction for those days.

3. Set missing value

Finally, after the two steps above are done, missing values are set to be median values. A median value is considered to be neutral.

6.1.4.2 Normalizing input variables

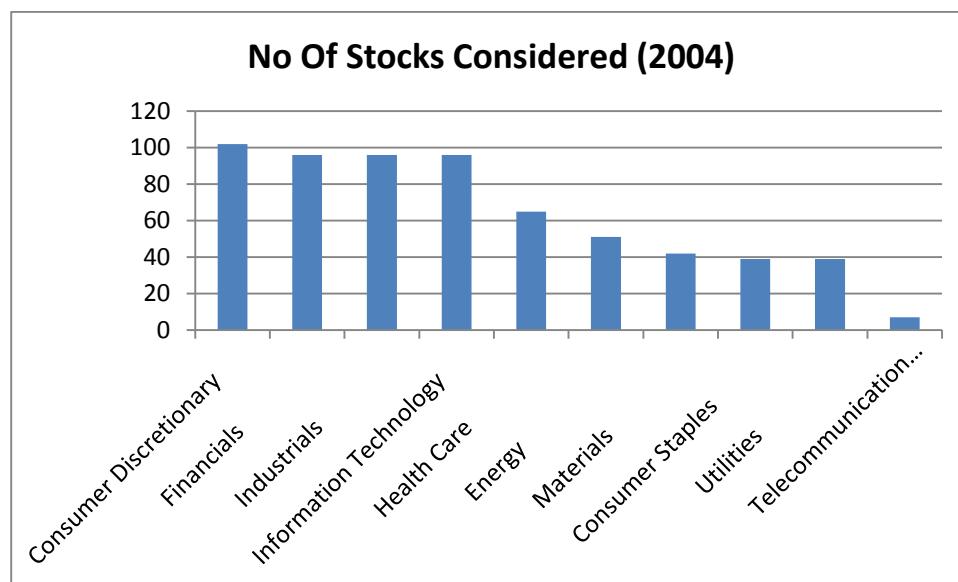
Stocks from different sectors usually have large differences in terms of factor distribution and performance. For example, for a factor variable such as debt capital ratio, the telecommunications/utilities sector will have a higher average value compared to other sectors as they are more capital intensive. The technology sector on the other hand is less capital-intensive and hence has less debt. Such factor variable value will naturally tilt towards certain sectors. Therefore, it is more ideal to neutralize the factor variable for sector bias. This means when normalizing factor variables, each sector of stocks should be considered on its own. Below shows a comparison of average debt capital ratio in year 2010 across sectors.



This step is done after the data has been handled for outliers and missing values on each trading day. To normalize each input variable for each sector, we simply use the z-score. The resulting values for each input variable will have a value between -3 and 3.

6.1.4.3 Identifying distribution of stocks across sectors

After the data has been pre-processed and normalized by sector, we examine the number of stocks in each sector that can be used for modelling. The figure below shows the number of stocks in each sector in year 2004. The average total number of stocks that can be used was 633. The telecommunications sector had the least average number of stocks at 7. Even if subsequent years where there were more available data, the number of stocks for this sector remained small. Given the small size of the telecommunications sector, we will omit it for modelling if we were to consider the use of sector models.



6.2 Detailed Modelling and Algorithms

6.2.1 Ensemble

The various modelling explored using various R package includes:

- Singular Method
 - Support Vector Machine
 - Works by finding the maximum margin hyper plane between classes in multi-dimensional space in order to classify. RBF kernel is used.
 - Neural Network
 - Train a single-hidden-layer neural network, currently using 100 hidden nodes to classify the classes.
 - Adaboost
 - Works by boosting weak variables. Increasing weight of error labels to re-train 100 times repeatedly.
 - Random Forest
 - Works by random sampling of data on random features for each tree then combining all 500 trees.
- Ensemble Methods
 - Voting - from above 4 models
 - Overall Probability

Several techniques are applied to analyse the data and 2 ensemble methods are used for choosing the best performing model. Details of findings of the various models sector by sector is attached in Appendix.

6.2.2 Adaboost

6.2.2.1 *Modelling approach - Classify the stocks using AdaBoost Algorithm*

Machine learning methods have become more and more popular in the quant world to predict stock returns. The main idea in using machine learning methods to predict future stock returns is to automatically learn to recognize complex data patterns and capture the hidden relationships among financial data, which are normally difficult to see with the human eye.

The AdaBoost model treats stock selection as a binary classification problem using supervised learning. We classify the stocks in our universe into outperformers and underperformers based on one-month forward stock returns. The construction of the

confidence score has two steps: the training step and the prediction step. In the training step, we use the end-of-month factor scores for each stock and one-month forward returns as training data to build the classifiers. In the prediction step, we use the current month factor score as the input for the classifiers we built in the training step, and the output is a confidence score.

6.2.2.2 Data preparation

We use the cross-sectional ranking of the factors, rather than the factor score itself as the input, and we calculate the factor ranking each month for all the available stocks. Then, we divide the factor ranking by the number of stocks to normalize the factor ranking to between (0, 1]. After normalizing the factors by rank and coverage, we then assemble the training data. We set the stocks in the top n%, measured by one-month forward return, as the outperformers. Similarly we set the bottom m% as the underperformers. Note that stocks not classified in the top or bottoms are disregarded. We perform this labelling exercise on a monthly basis, and pile the different months' data together as training data.

6.2.2.3 AdaBoost Algorithm

We use a machine learning algorithm called AdaBoost (Schapire [1998]) to build classifiers that can combine the best performing factors in previous months and get a confidence score. The higher the confidence score, the more likely the stock is going to outperform in the next month. The main idea of AdaBoost is that it adaptively builds a sequence of classifiers that are constantly being tweaked to emphasize misclassified stocks, thereby slowly improving the classification of stocks that would normally be incorrectly classified. Although certain classifiers can be weak, as long as their performance is not random, the performance of the final model will improve.

In our case, a weak classifier is simply defined by a factor. We divide the factor into quantiles, and calculate the weights of outperformers and underperformers in each quantile. Intuitively, the most effective factors are those which have the largest difference between the weights of outperformers and underperformers in each quantile. The higher the weight of outperformers relative to the weights of underperformers in that quantile, the higher value the output of that weak classifier will be. This way of defining the weak classifier can transform the non-linear factors into linear factors.

Initially, we equally-weighted each observation in the training data, and then the weights are updated in each round after a new weak classifier is found. The weight of each

incorrectly classified stock is increased and the weight of each correctly classified stock is decreased, so that the next classifier focuses on the stocks which have so far not been correctly classified. In each round we choose the most effective weak classifier, which can distinguish the outperformers and underperformers the most with the current set of weights. This means the current best performing factor is less correlated with the previously selected factors. Essentially, what our machine learning system does is select those factors that are complementary to each other. The output of the strong classifier is the sum of all the weak classifiers; it is a real value confidence score of how likely the stock is to be an outperformer.

After we train our machine learning model, we can predict month-ahead returns using the constructed classifiers with new data. First we transform factor scores into quantiles, and get the corresponding value of the weak classifier. Adding up all the weak classifiers, we get the final value of the strong classifier. We use the value of strong classifier as a new composite factor, which has better performance than any of the factors comprising it.

6.2.2.4 Flow Chart

Normalize the factor value to a real value [0,1] according to the rank of the factor in each month



Classify outperformers and underperformers.



Initially dually weight each stock.



Build weak classifiers for each factor.



Add the selected weak classifier to the strong classifier and updates the weights. Lower the weight for correctly classified samples and increase the weight for wrongly classified samples.



The final output of the strong classifier is the value of all the weak classifiers add up together.

AdaBoost Algorithm

1. Normalize all the factor value to a real value (0,1] according to the rank:

$$\text{Normalized_factor} = \text{factor_rank}/\text{total_number_of_factors_in_the_month}$$

2. Given stock performance set $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ and the factor pool F, where $y_i \{-1, 1\}$, N is the number of stocks in the training sample.

3. Initially equally weighted all the stocks $w_1(x) = 1/N$

4. For $l = 1, \dots, L$

- 1) For each factor f^k from the factor pool we build a weak classifier h^k .

- a) Divide the training data into Q quantiles, X_1, X_2, \dots, X_Q

- b) For each quantile j we calculate the total weight of outperformers and underperformers

$$W_{\pm}^j = \sum_{y_i = \pm 1, f(x_i) \in \text{quantile}_j} w(x_i)$$

- c) Calculate the discriminative objective function:

$$Z_l^k = \sum_{j=1}^n \sqrt{W_+^j W_-^j}$$

- d) Get a weak classifier:

$$h^k(x) = \frac{1}{2} \ln \left(\frac{W_+^j + \varepsilon}{W_-^j + \varepsilon} \right)$$

- 2) Find the best weak classifier according to the discriminative objective function:

$$h_l(x) = h^k(x) \text{ where } k = \arg \min_{f^k \in F} \{Z_l^k\}$$

- 3) Update the weights of each stock:

$$w_{(l+1)}(x_i) = w_l(x_i) \exp(-y_i h_l(x_i))$$

- 4) Normalize the weight $w_{l+1}(x_i)$ so that they add up to 1.

5. The final strong classifier: $H(x) = \sum_{l=1}^L (h_l(x))$

6.2.2.5 Implementation results (Partial)

Success rate - for each chosen stock (1 if beat the average stock return, 0 otherwise)

We achieved 72% accuracy with the following parameter settings:

- Iterations = 20
- Lower cut = 60%
- Upper cut = 90%

Average return for stocks chosen versus average return for universe (sector only/all sectors) for each month and aggregate over the months.

```
[snadaraj@FLXSGL64EQ02 ~/AdaBoost] $ g++ Ada.cpp
[snadaraj@FLXSGL64EQ02 ~/AdaBoost] $ ./a.out
Sucessfully Loaded the data! Rows=18999Factors=98
Upper cut=12.284 Lower cut=0.503707
After cleaning nRows =13300
Sucessfully Loaded the Test data! Rows=1498
Accuracy = 71.6288%
[snadaraj@FLXSGL64EQ02 ~/AdaBoost] $ █
```

6.2.3 Neural Network

6.2.3.1 Modelling approach

As a computational model, NN features as follows:

- Strength
 - High tolerance to noisy data
 - Ability to classify untrained patterns
 - Well-suited for continuous-valued inputs and outputs
 - Algorithms are inherently parallel
- Weakness
 - Long training time and large training data
 - Poor interpretability

In our stock trading scenario, 2 tentative NN models are investigated:

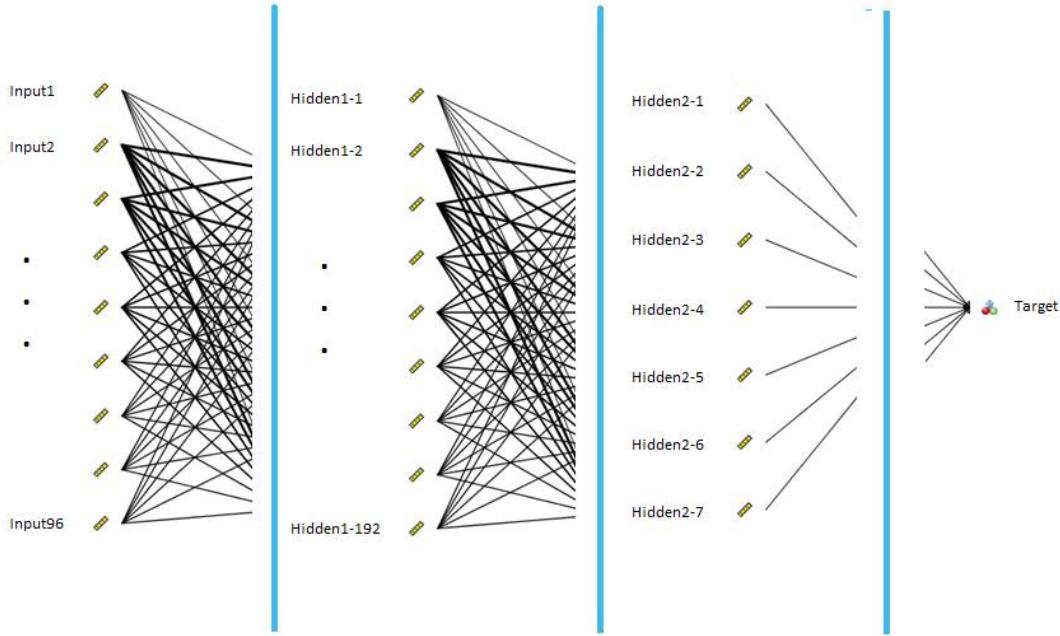
- Classification model: to predict stock price fluctuation trend, either up or down,
- Regression model: to predict the stock price fluctuation amplitude, the quantity

6.2.3.2 Regression Modelling

The network is designed as Two-hidden-layer Feed-forward Network with back-propagation learning with the architecture below:

- Input layer: 96 nodes, representing the collective input data

- Hidden layer 1: $92 * 2 = 192$ nodes, representing the compositional factors
- Hidden layer 2: $6 + 1 = 7$ nodes, representing the 6 aggregative factors and bias respectively
- Output layer: 1 node, representing the numerical target
- Activation function: Bipolar Sigmoid Function



6.2.3.3 Classification Modelling

The network is designed Probabilistic Neural Network with the architecture below:

- Input layer: 96 nodes, representing the collective input data
- Output layer: 2 node, representing the categorical trend, up or down
- Activation function: Gaussian function

6.2.3.4 Sample Results of Regression

```
Validation MSE: 41.8763281140175
Optimal Neural Network Parameters:
Input nodes - 96, Output nodes - 1, Hidden Layers - 2[192-7], Least Training MSE: 100.168317923734
Layer-Neuro[01,01] weights - 1.210,1.037,1.214,-0.530,1.562,1.289,-1.609,0.220,-1.026,-2.775,-3.657,-0.580,0.347,0.620,0.
Layer-Neuro[01,02] weights - 1.607,1.475,1.214,0.524,1.131,1.287,-0.032,0.201,1.273,1.495,1.173,2.156,0.785,1.660,1.392,0.
Layer-Neuro[01,03] weights - 1.263,0.450,0.673,1.084,0.626,1.113,0.832,0.487,0.437,0.580,0.967,0.268,0.357,0.208,0.561,0.
Layer-Neuro[01,04] weights - 1.077,0.813,0.383,-0.525,1.106,1.657,-2.074,-0.486,-0.788,-3.326,-3.448,-0.780,0.467,0.097,0.
Layer-Neuro[01,05] weights - -0.107,-0.091,0.156,0.273,0.556,0.696,-0.257,-0.387,0.121,0.786,0.942,0.493,0.489,0.413,0.58
Layer-Neuro[01,06] weights - 0.609,0.895,0.375,0.543,0.160,0.155,0.700,0.974,0.868,0.613,1.117,0.892,0.450,0.883,0.814,0.
Layer-Neuro[01,07] weights - 0.422,0.534,0.689,0.581,0.944,0.511,0.655,0.230,0.024,-0.122,0.381,0.505,0.736,0.349,0.823,0.
Layer-Neuro[01,08] weights - 0.906,0.831,1.082,0.026,-0.247,0.514,1.231,0.632,1.062,0.948,0.947,0.299,0.839,1.036,0.819,0.
Layer-Neuro[01,09] weights - 0.832,1.574,1.394,1.548,1.515,1.393,1.057,0.710,0.860,1.043,0.682,1.490,0.718,0.548,1.200,1.
Layer-Neuro[01,10] weights - 1.195,0.323,0.516,0.585,0.772,0.264,0.841,0.329,0.730,0.919,0.036,0.330,0.242,0.199,0.131,0.
Layer-Neuro[01,11] weights - 0.930,1.167,0.748,0.193,0.946,0.098,0.410,0.726,0.146,0.776,0.528,0.785,0.430,0.293,1.164,0.
Layer-Neuro[01,12] weights - 0.204,0.214,1.778,-0.845,-1.206,-0.037,1.912,1.913,1.634,0.373,0.534,0.472,-0.109,0.023,1.10
Layer-Neuro[01,13] weights - 0.928,0.676,0.845,0.068,-0.357,-0.402,0.439,1.150,0.556,1.179,0.445,0.859,1.309,0.643,1.309,
Layer-Neuro[01,14] weights - 0.867,0.232,0.366,0.032,0.546,0.248,0.177,0.417,-0.095,0.072,0.262,0.639,0.022,0.805,0.332,0.
Layer-Neuro[01,15] weights - -0.551,-1.509,0.735,-0.754,-2.617,1.047,2.195,1.591,1.397,1.582,1.586,1.625,-0.959,-1.263,1.
Layer-Neuro[01,16] weights - 0.345,0.295,1.047,0.048,0.914,0.415,0.163,0.759,0.756,0.902,0.675,0.301,0.605,0.768,0.506,0.
Layer-Neuro[01,17] weights - 0.729,0.951,0.944,-0.340,-1.480,-1.706,-2.383,-1.180,-1.037,-1.316,-0.116,-0.491,0.241,0.075
```

6.3 System Architecture

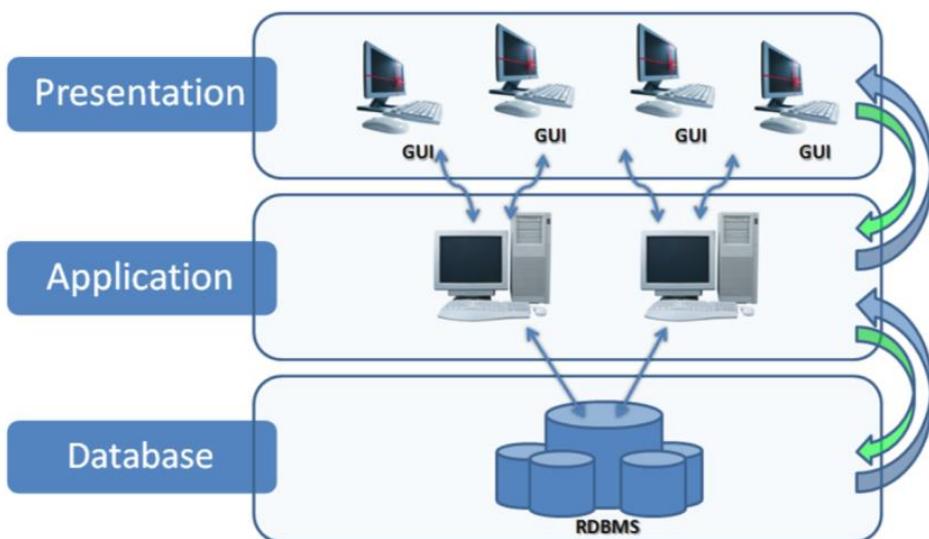
6.3.1.1 Architecture Overview

The system will be based on typical 3-tier Client/Server architecture for scalability, modularity, robustness and maintainability. The 3 layers are as shown below.

- Presentation layer
- Application layer
- Repository layer

The overall architecture can be represented as follows:

3-Tier Client/Server Architecture



6.3.2 Core Components

There will be a couple of core functional components within each layer and on top of them façade will be defined to expose functionality and service for the adjacent upper layer to consume.

6.3.2.1 Repository Layer

This layer will mainly manage incoming/outgoing data flow on the repository level to ensure data integrity, availability, consistency and maintainability:

- Data importing component: import stock data update on a periodic or ad-hoc basis
- Data accessing component: provide data access for upper and other components

6.3.2.2 Application Layer

This layer will encapsulate business inferential rule/logic, perform the various processes on request, and generate sensible and concrete outcomes as expected

- Inferential engine component
- Data intelligence enablement component
- User investment profiling component
- Factoring analytical and computational component

6.3.2.3 Presentation Layer

This layer is what administrator and end user will liaise with to initiate informative queries and perform predictive tasks:

- Interactive graphical user interface
- User input collection, validation, representation and transformation component
- Reporting/exporting component

6.4 Solution Technology

6.4.1 Technology Candidacy

To align with project objectives and fulfil the architectural requirements with the context of limited manpower and time, the technology candidates should possess the following merits to outstand:

- Maturity adoptable
- RAD (rapid application development) capable
- Rich open source resource available
- Flatter learning curve

6.4.2 Technology Selection

It consists of selection of technologies for the fundamental aspects/parts of the system design.

Aspects/Parts	Selection	Assessment
Database	MS SQL Server 2008	Prerequisite of sponsor for compatibility and integration
Presentation	ASP.NET Web Form or MVC	<ul style="list-style-type: none">• Mature• RAD well supported
Application	R R.NET NMath C#.NET AForge.NET Encog.NET	<ul style="list-style-type: none">• .NET background and foundation of team• Open source resource richness

6.5 Development Methodology

- Agile adherent
- CRISP
- Agile modeling, processing, and delivering
- Scrum planning and scheduling

6.6 Development Platform, Framework and Tool

There are 2 heterogeneous technologies proposed to take advantage of: R and .NET. In-between there is bridging framework/library like R.NET and NMath.

The major tools used include but not limited to:

- Microsoft Visual Studio
- R/RStudio IDE
- Microsoft SQL Server Management Studio

7 Progress Summary

7.1 Tasks accomplished as part of Phase 2:

- (1) Data Preparation and cleaning: Collection of periodic stock data, data cleaning and transformation operations.
- (2) Finalization of the various stock data factors.
- (3) Performing of sector by sector classification of factors.
- (4) Modelling on both full data as a whole and individual sector by sector
- (5) Implementation of customized Adaboost and Neural network algorithms

7.2 Tasks planned in Pipeline for Phase 3:

- (1) To Combine the different models designed/implemented to form a hybrid-solution model for implementing smart trader system
- (2) User Interface Design and Implementation
- (3) To perform Verification and Validation activities as part of Quality Assurance
- (4) To perform User Acceptance Test/Deployment Activities
- (5) Documentation of the overall project
- (6) Research paper preparation

8 Appendix A: Findings and Results (Comparison of various models – Sector by Sector)

8.1 Sector – Consumer Discretionary

sectors

Accuracy Precision by Sector

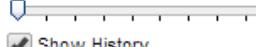
Overall Accuracy of Model

Overall Precision of Models

Overall Returns of Models

Sector

Consumer Dis... ▾

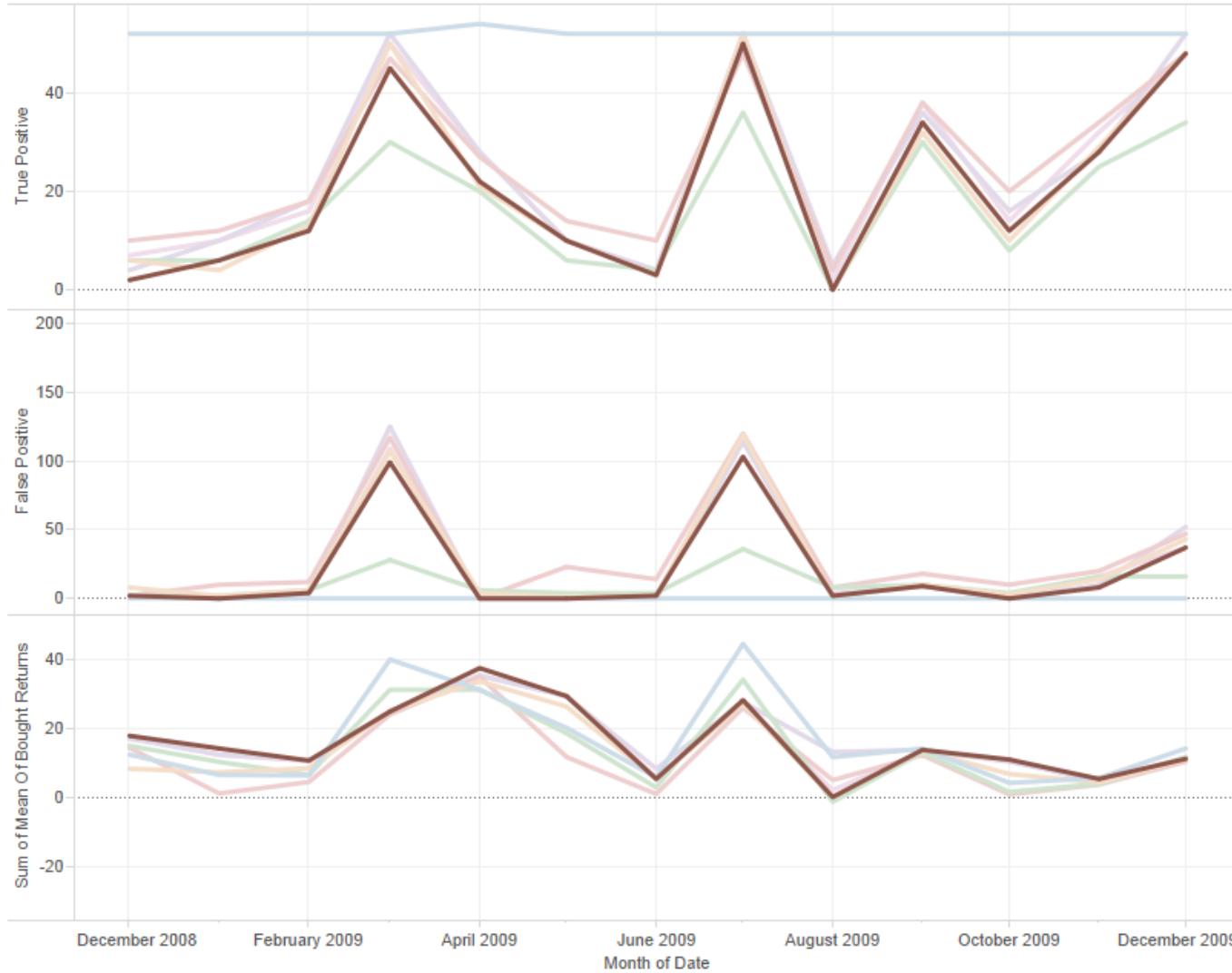


Show History

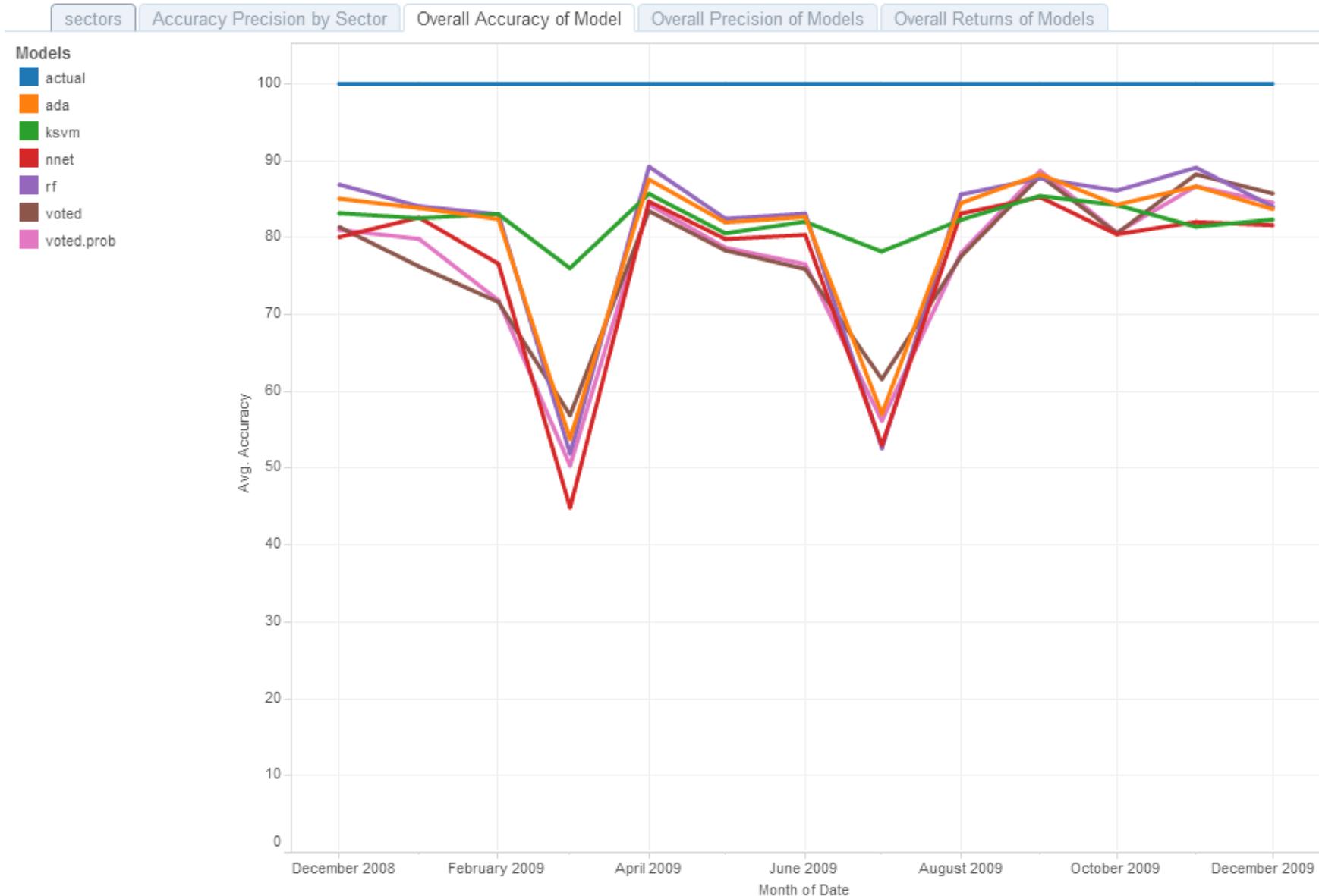
Models

nnet

 voted
 voted.prob

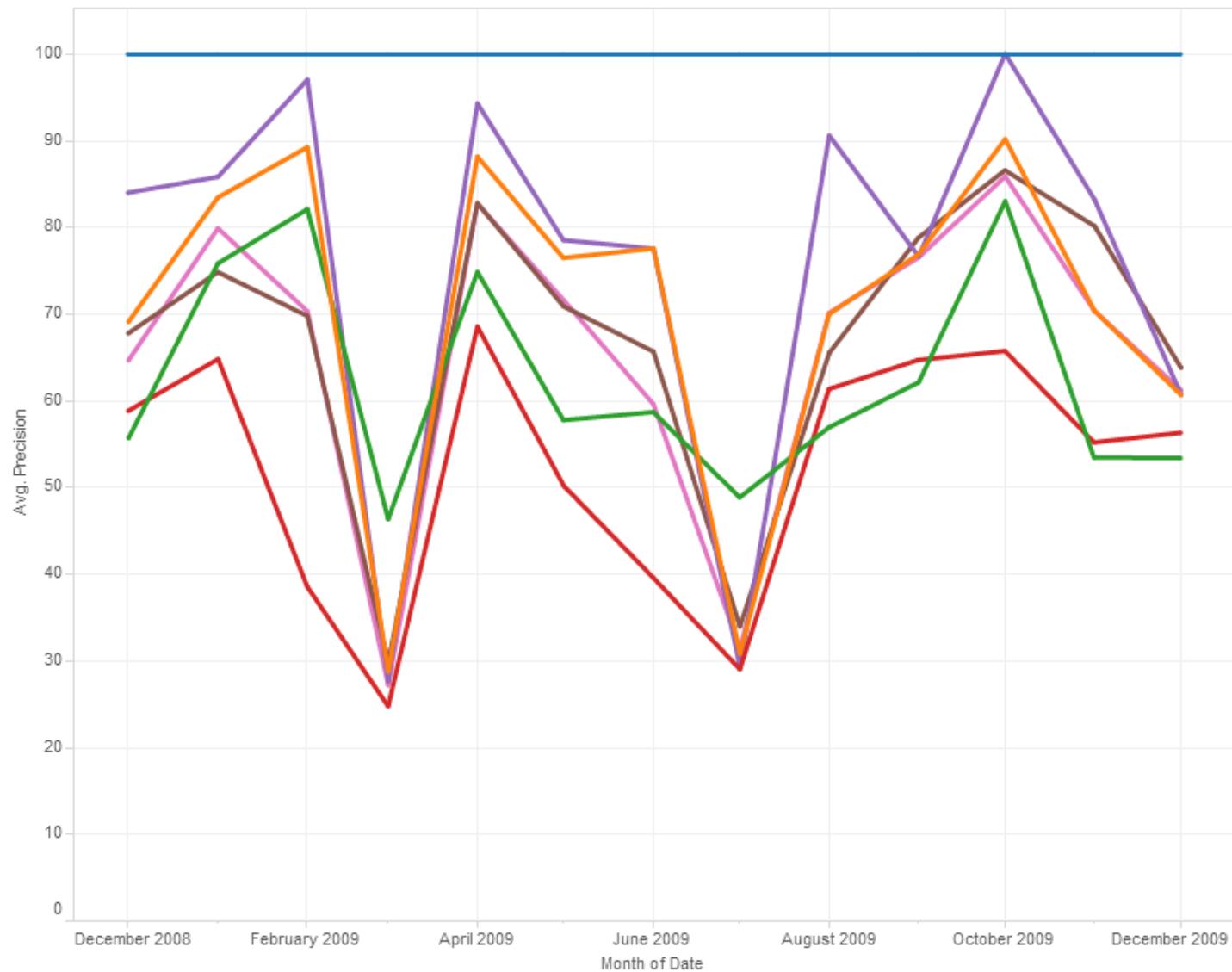


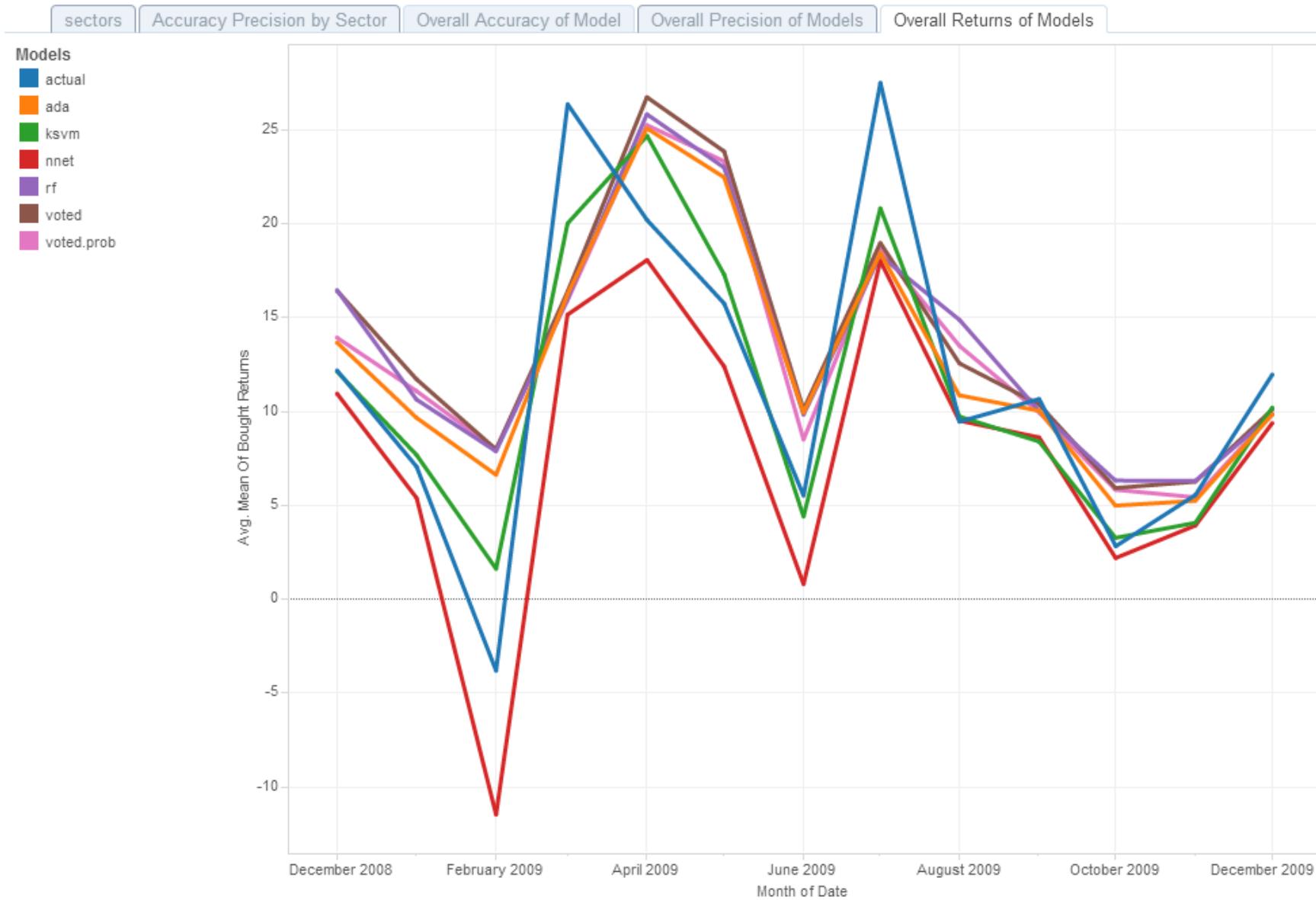




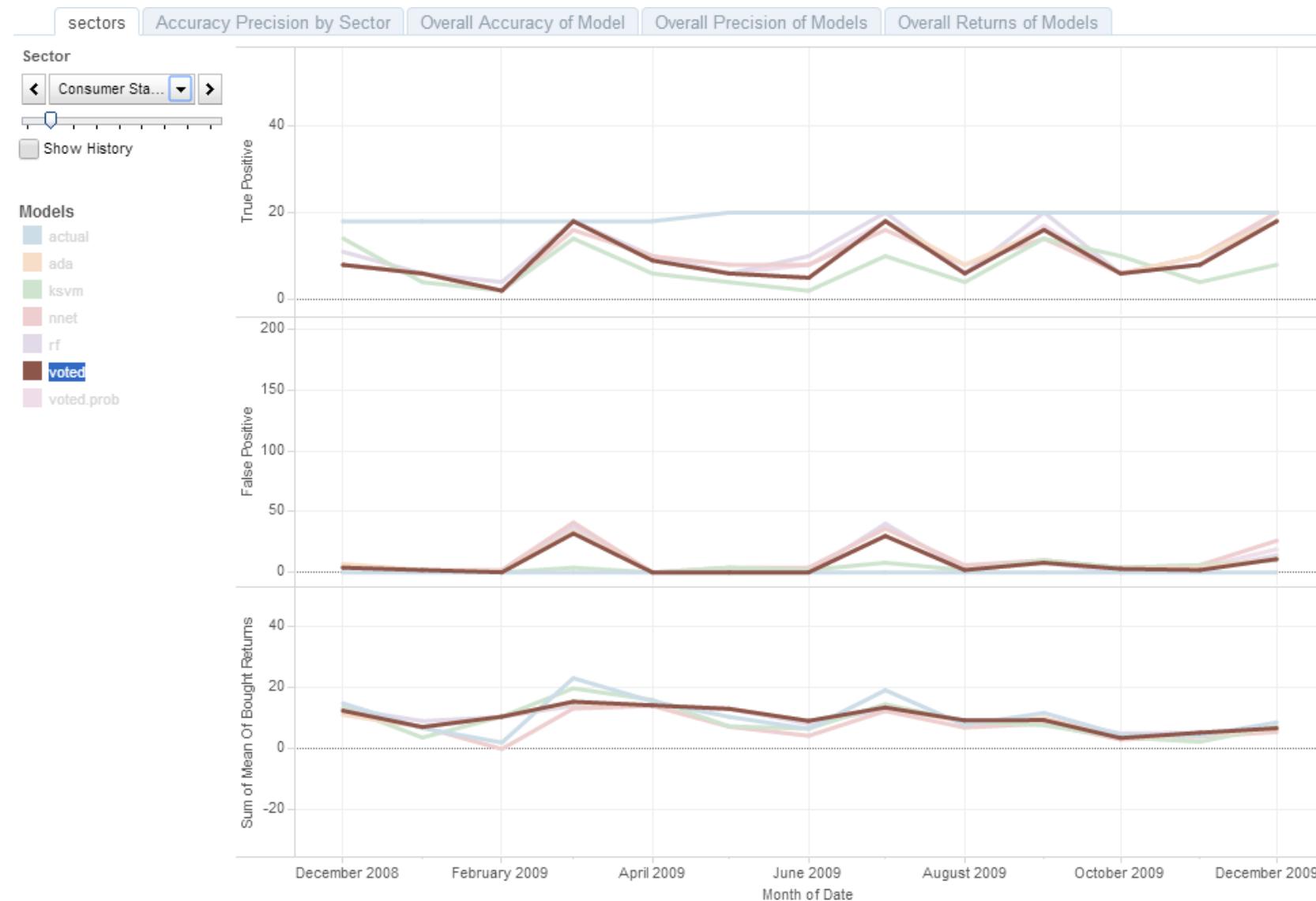
[sectors](#)[Accuracy Precision by Sector](#)[Overall Accuracy of Model](#)[Overall Precision of Models](#)[Overall Returns of Models](#)**Models**

- █ actual
- █ ada
- █ ksvm
- █ nnet
- █ rf
- █ voted
- █ voted.prob

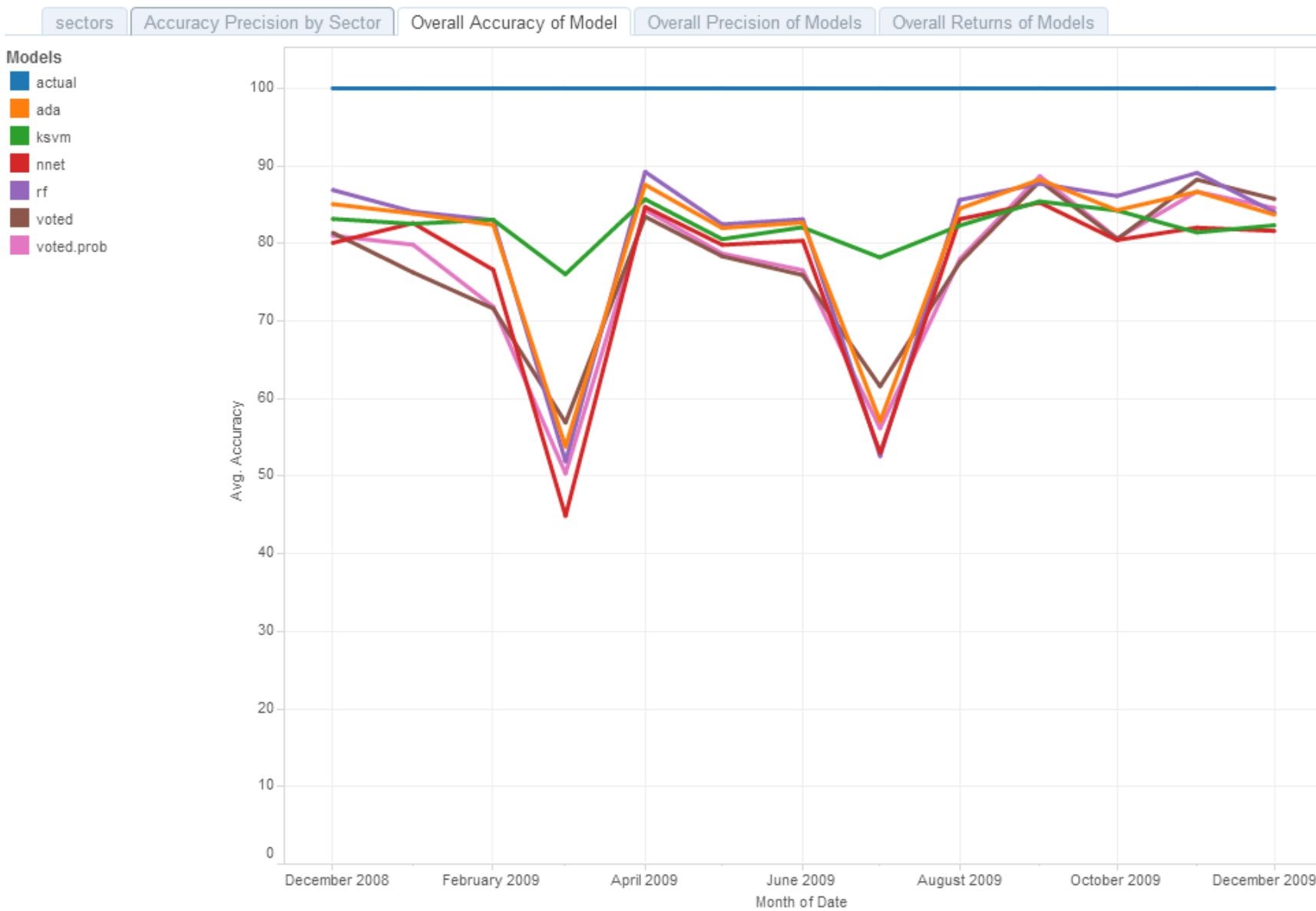




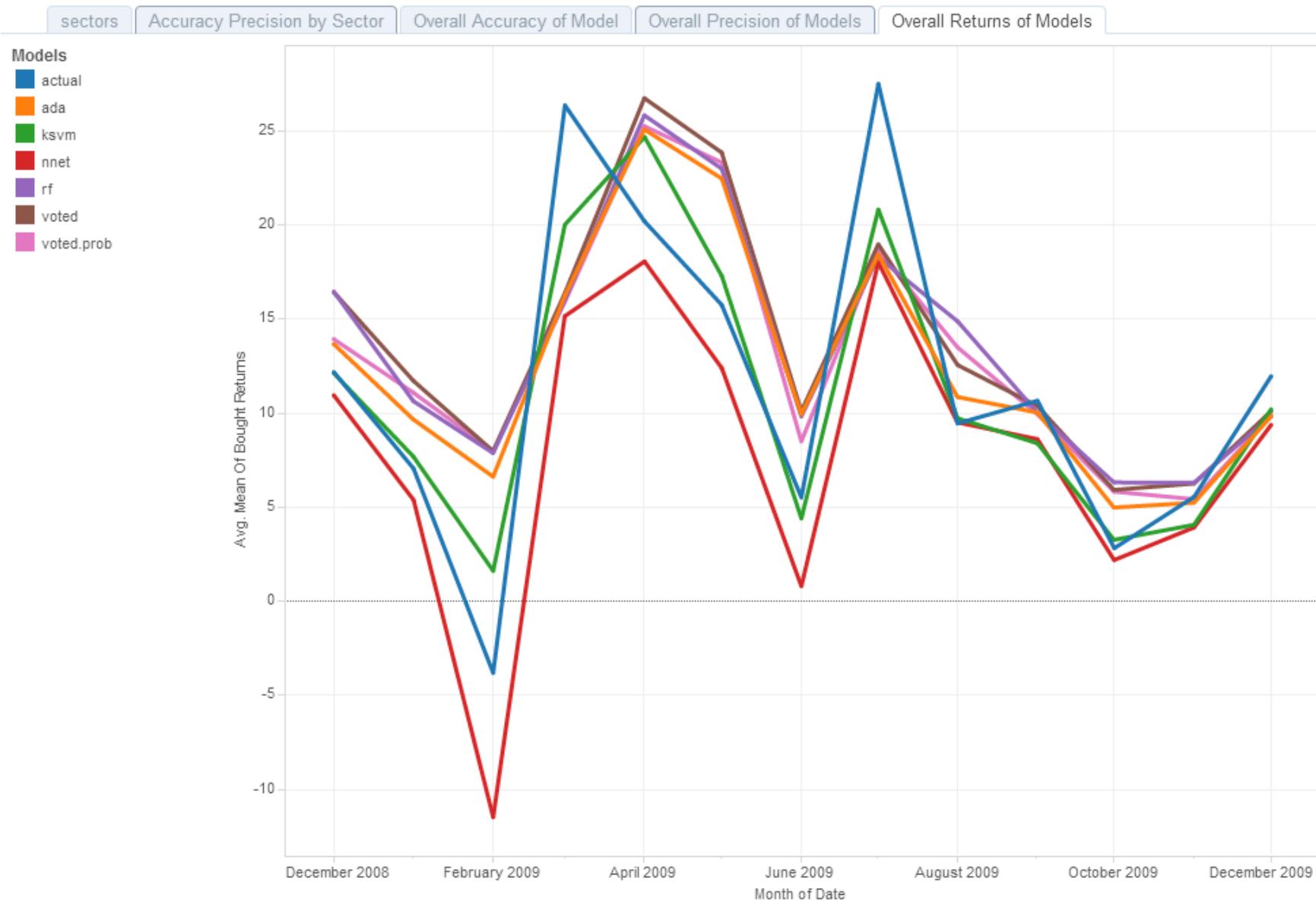
8.2 Consumer Staples



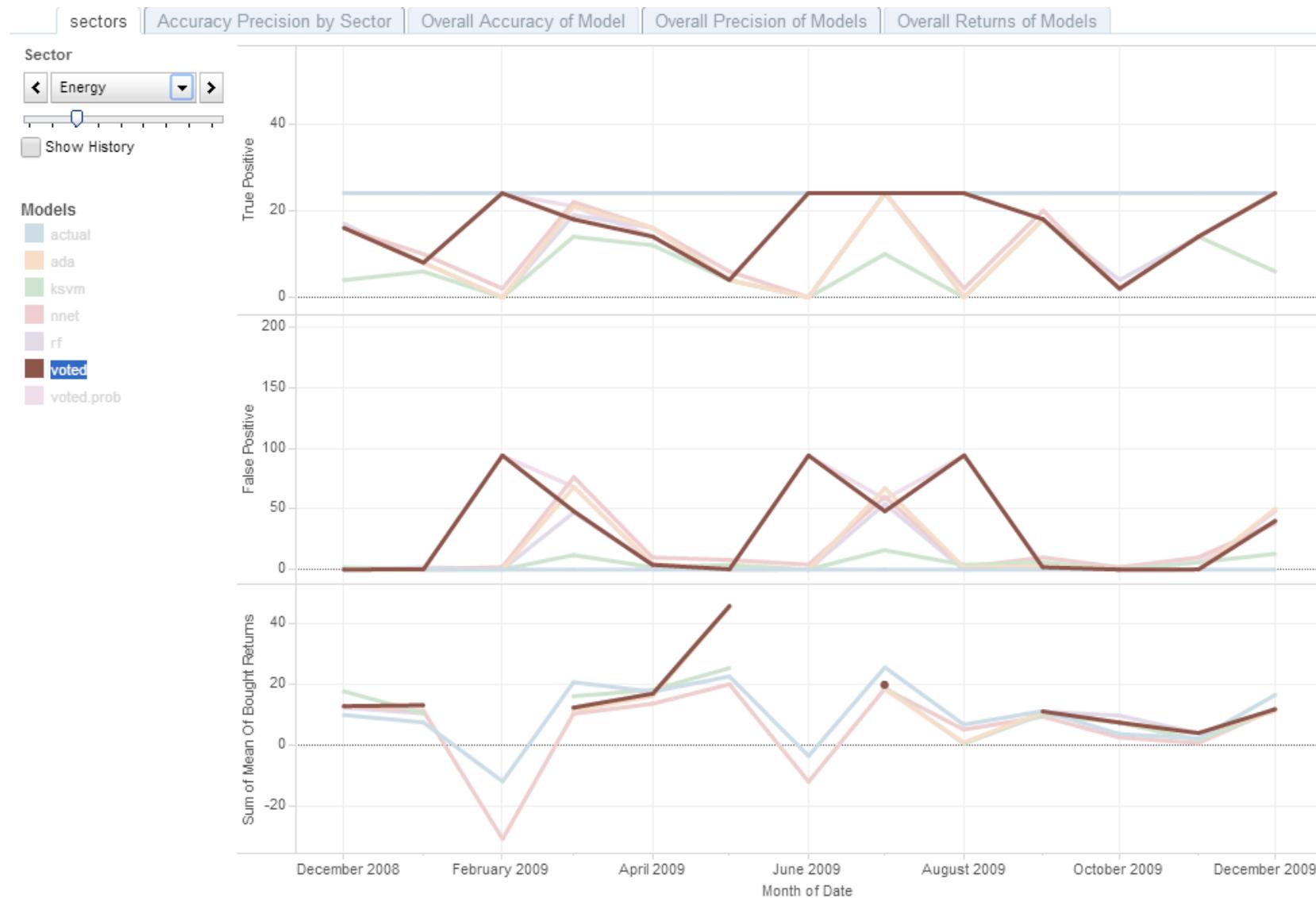


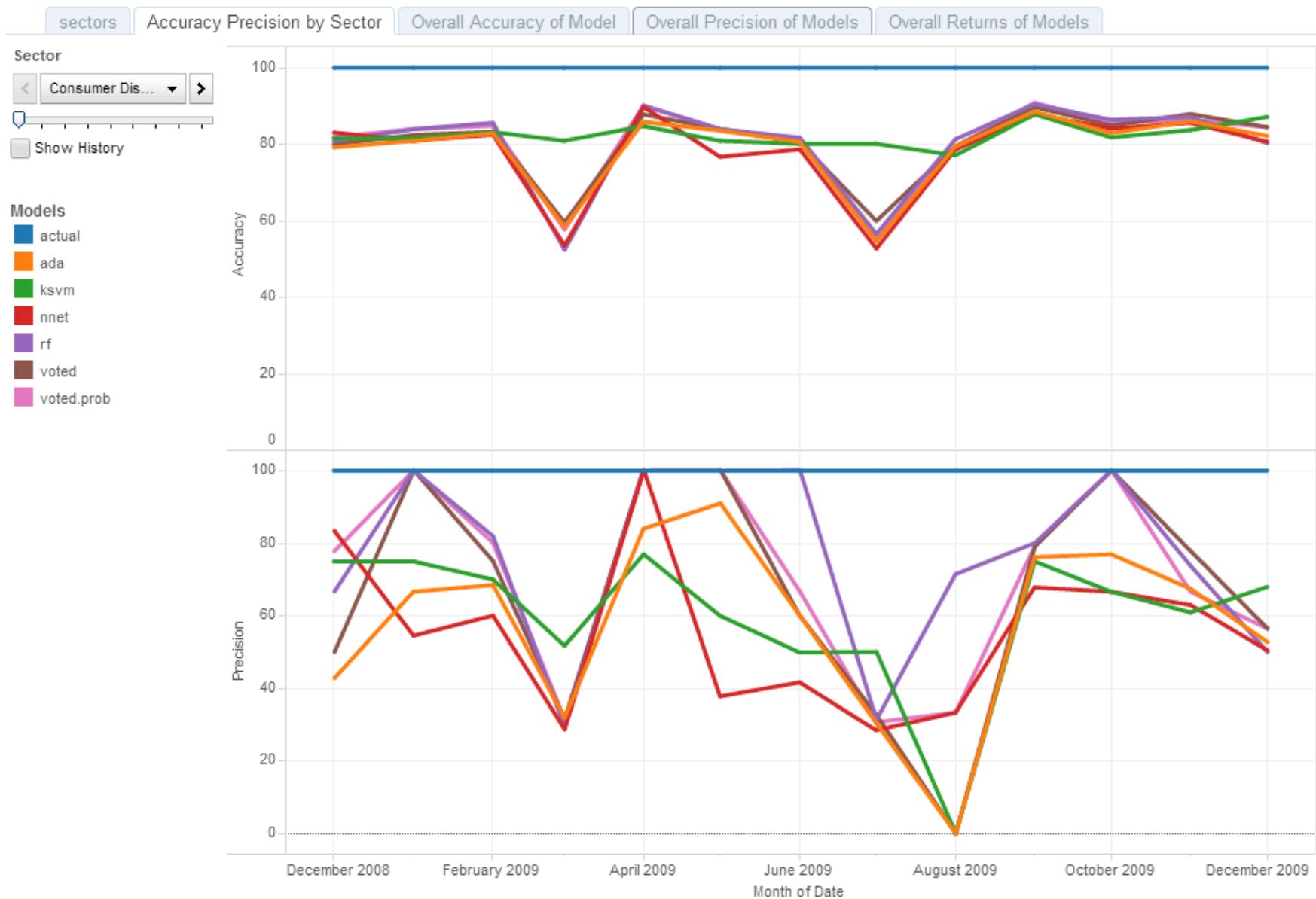


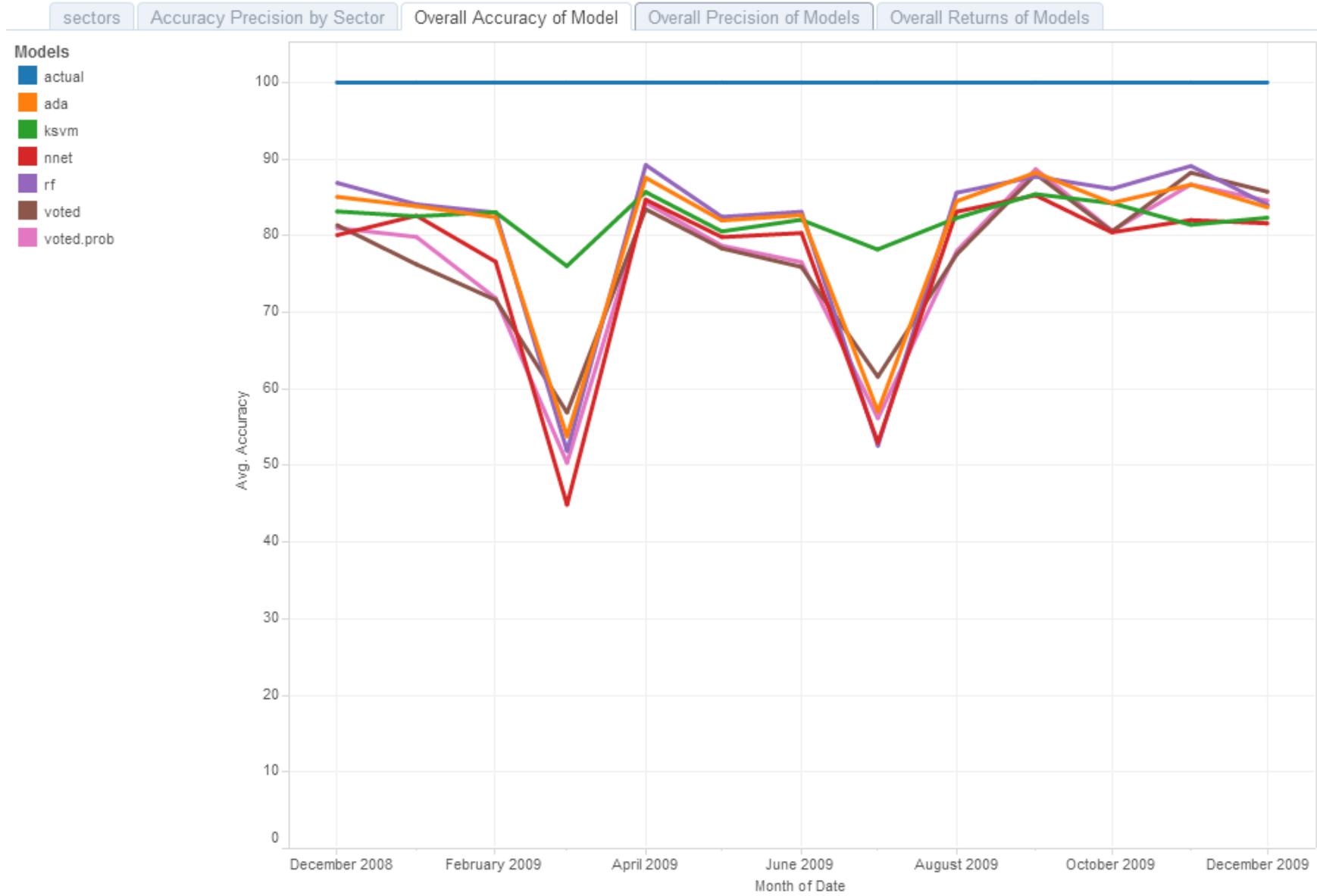




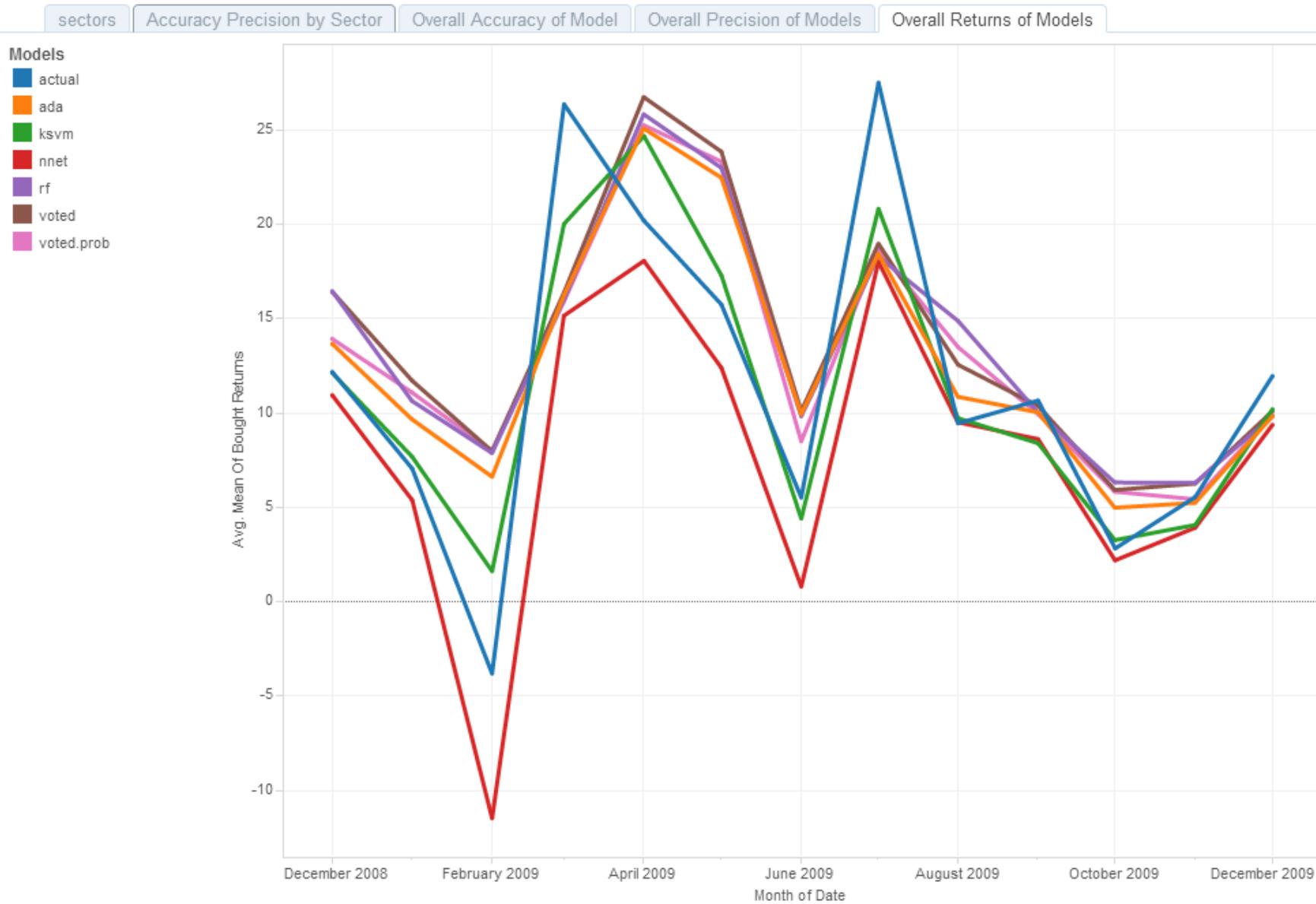
8.3 Sector - Energy



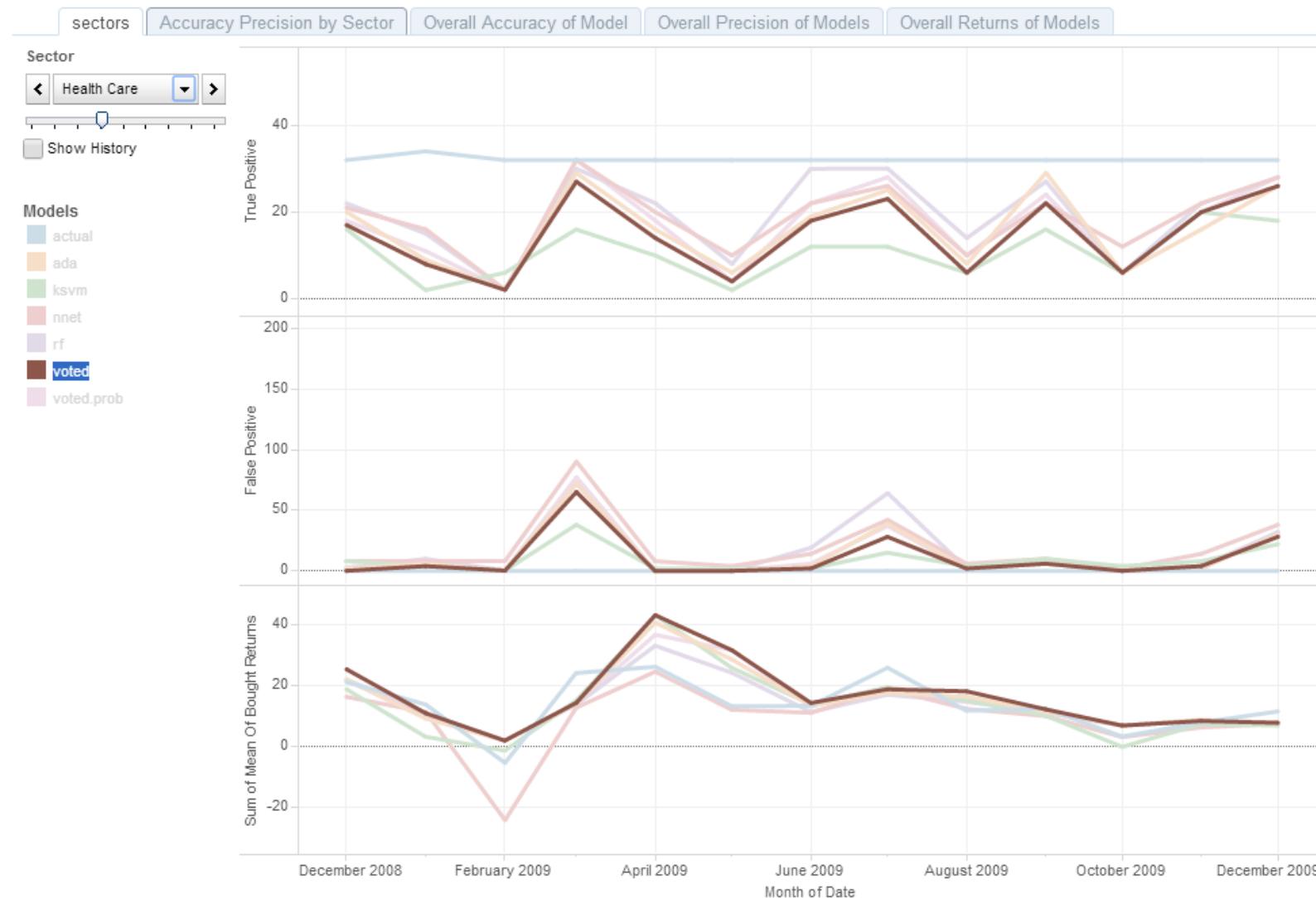








8.4 Energy - HealthCare



sectors

Accuracy Precision by Sector

Overall Accuracy of Model

Overall Precision of Models

Overall Returns of Models

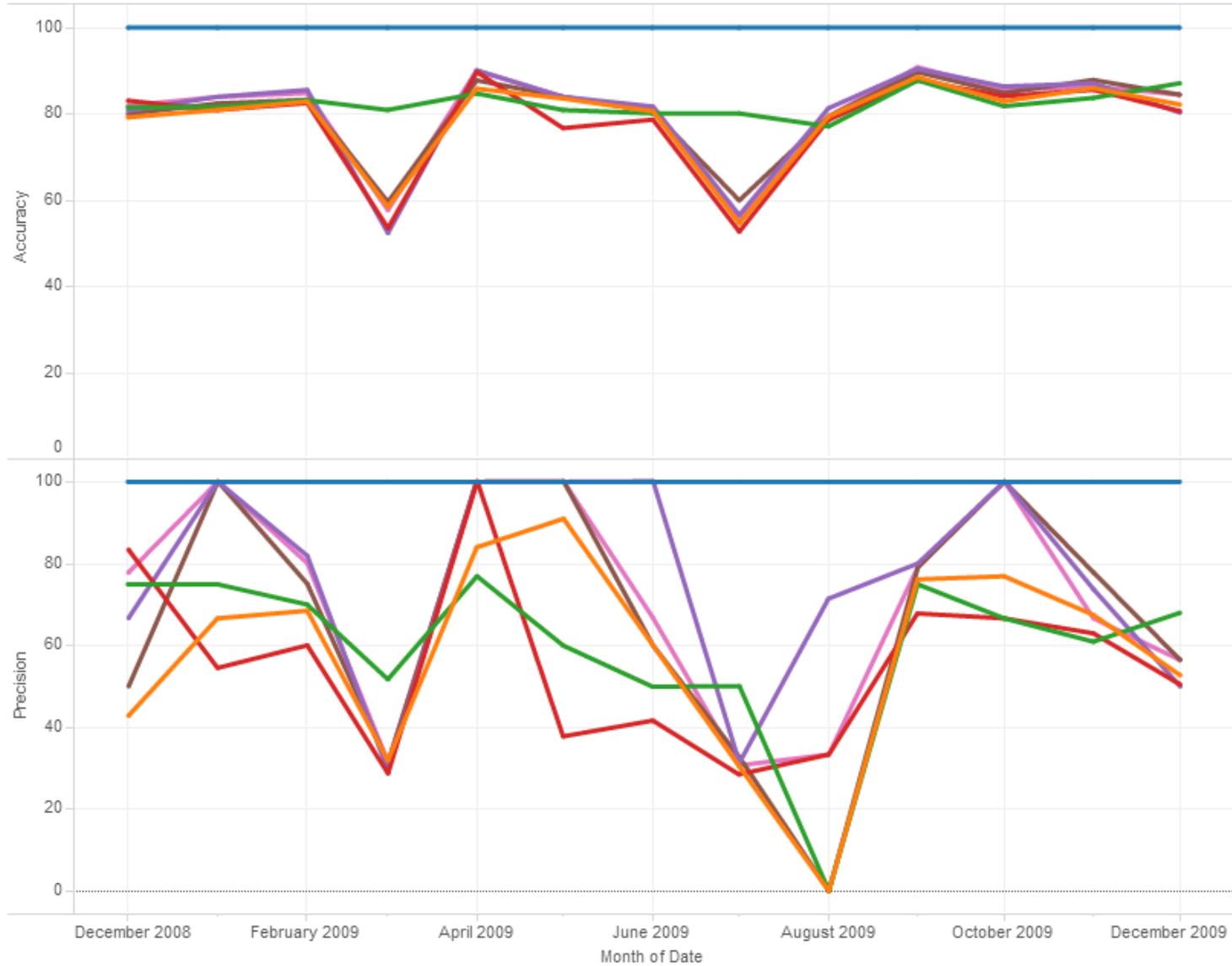
Sector

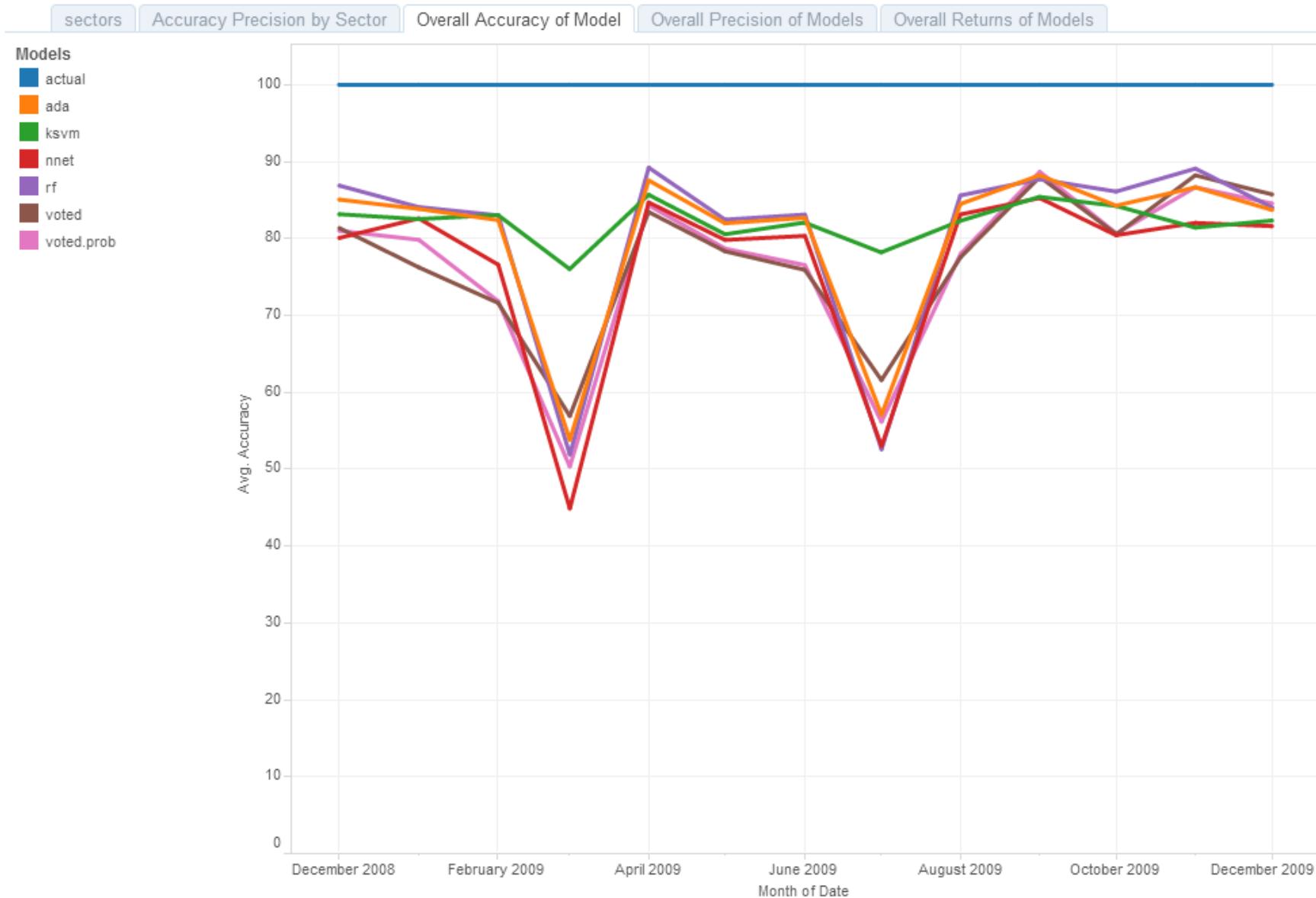
Consumer Dis... ▾

Show History

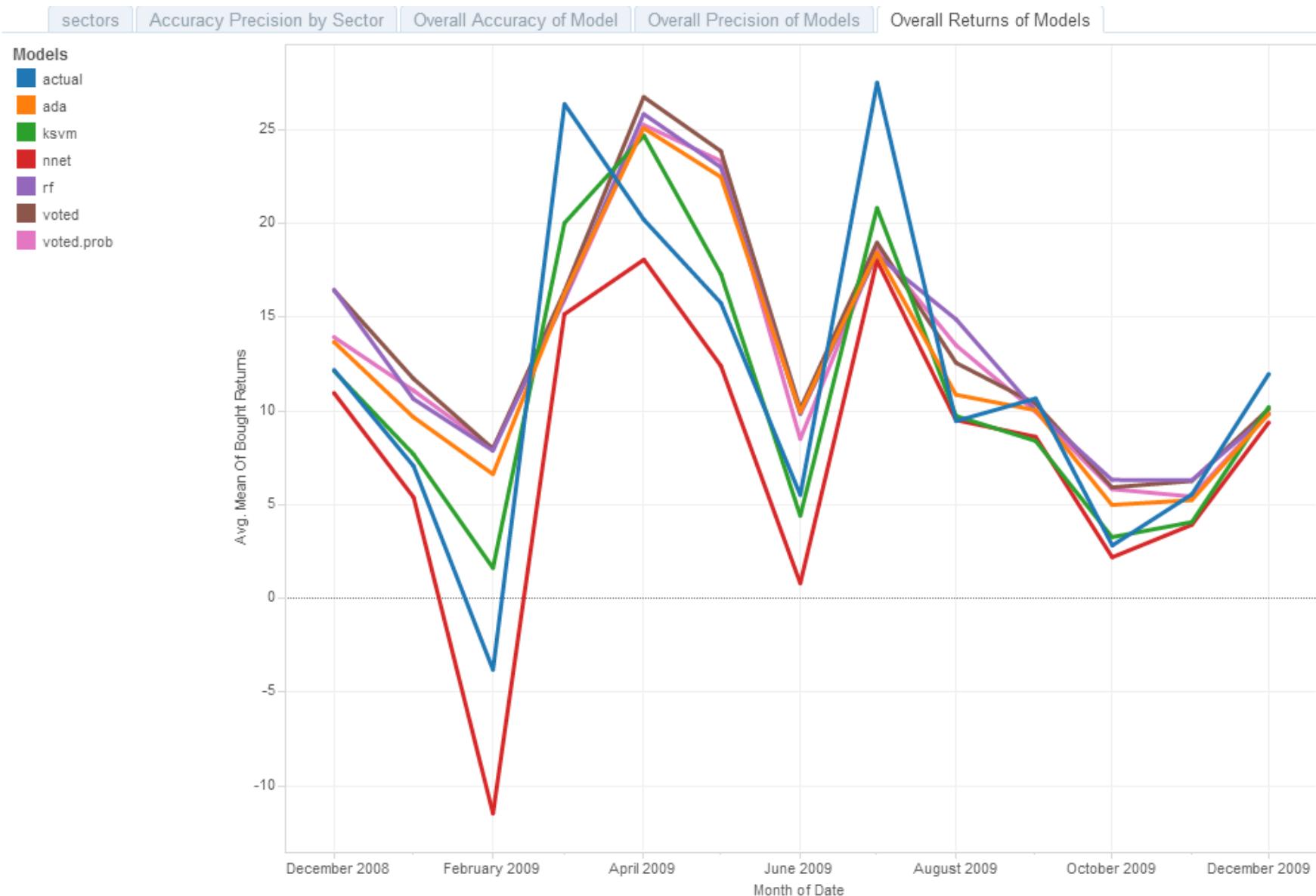
Models

- actual
 - ada
 - ksvm
 - nnet
 - rf
 - voted
 - voted.prob

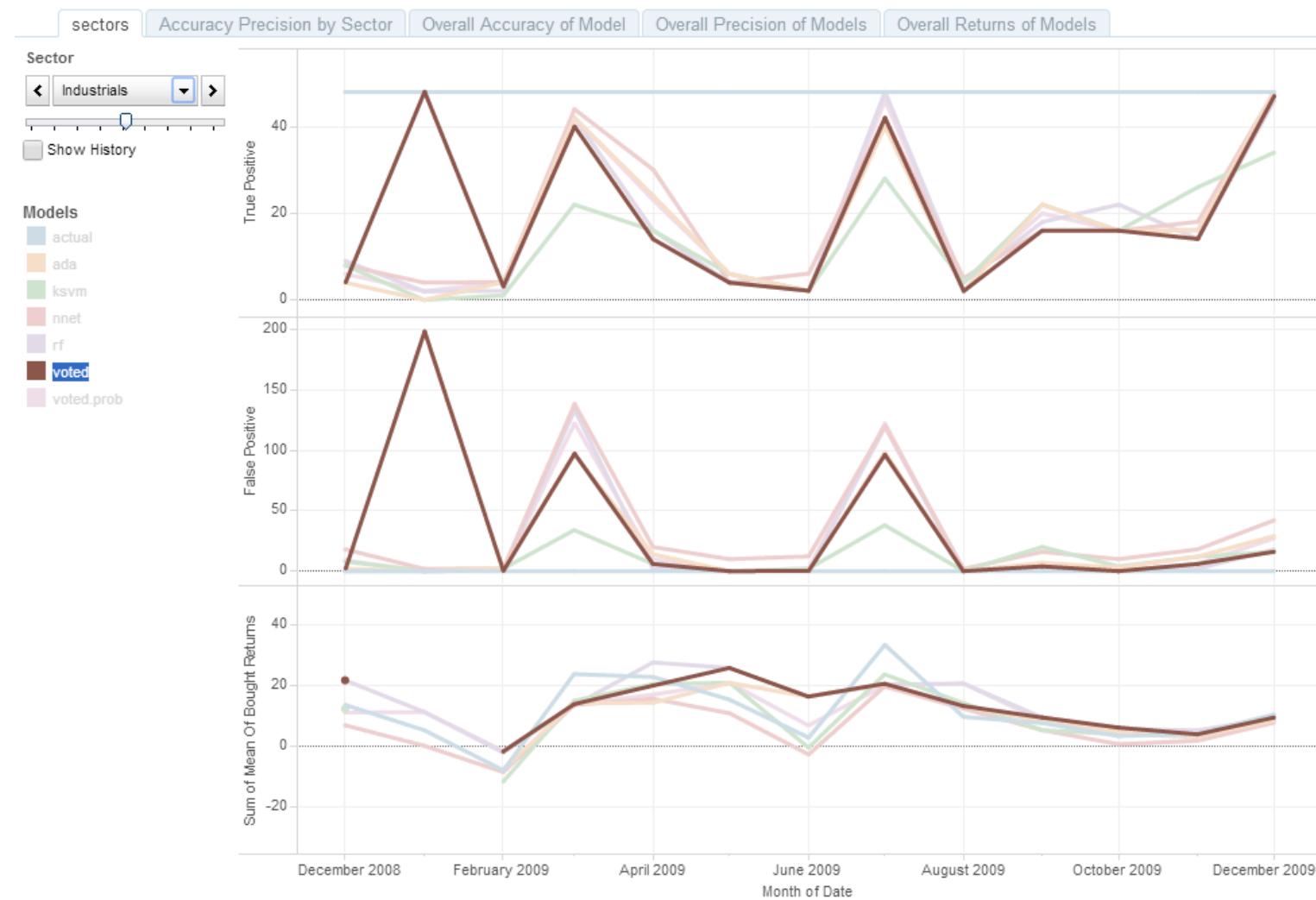


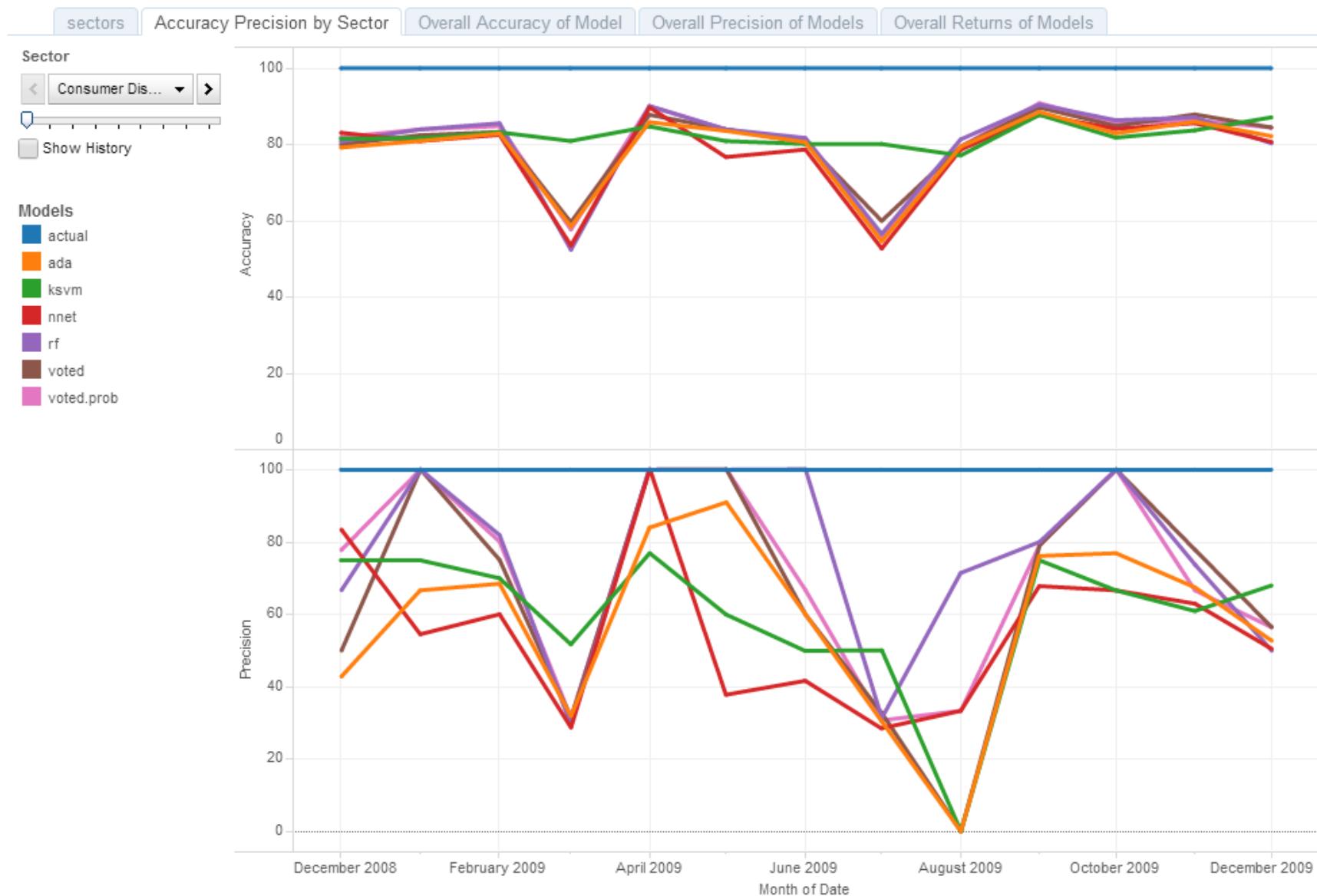


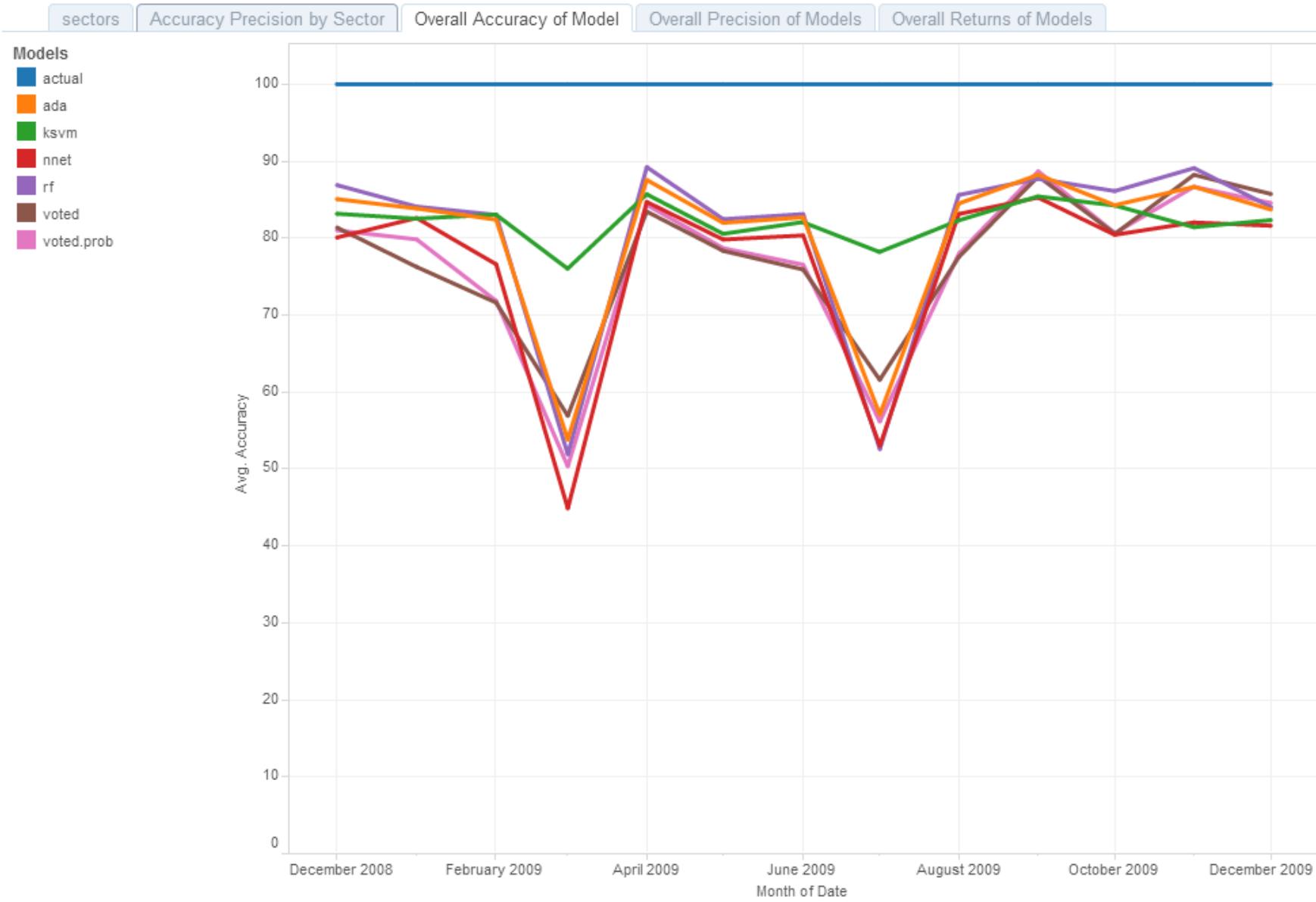




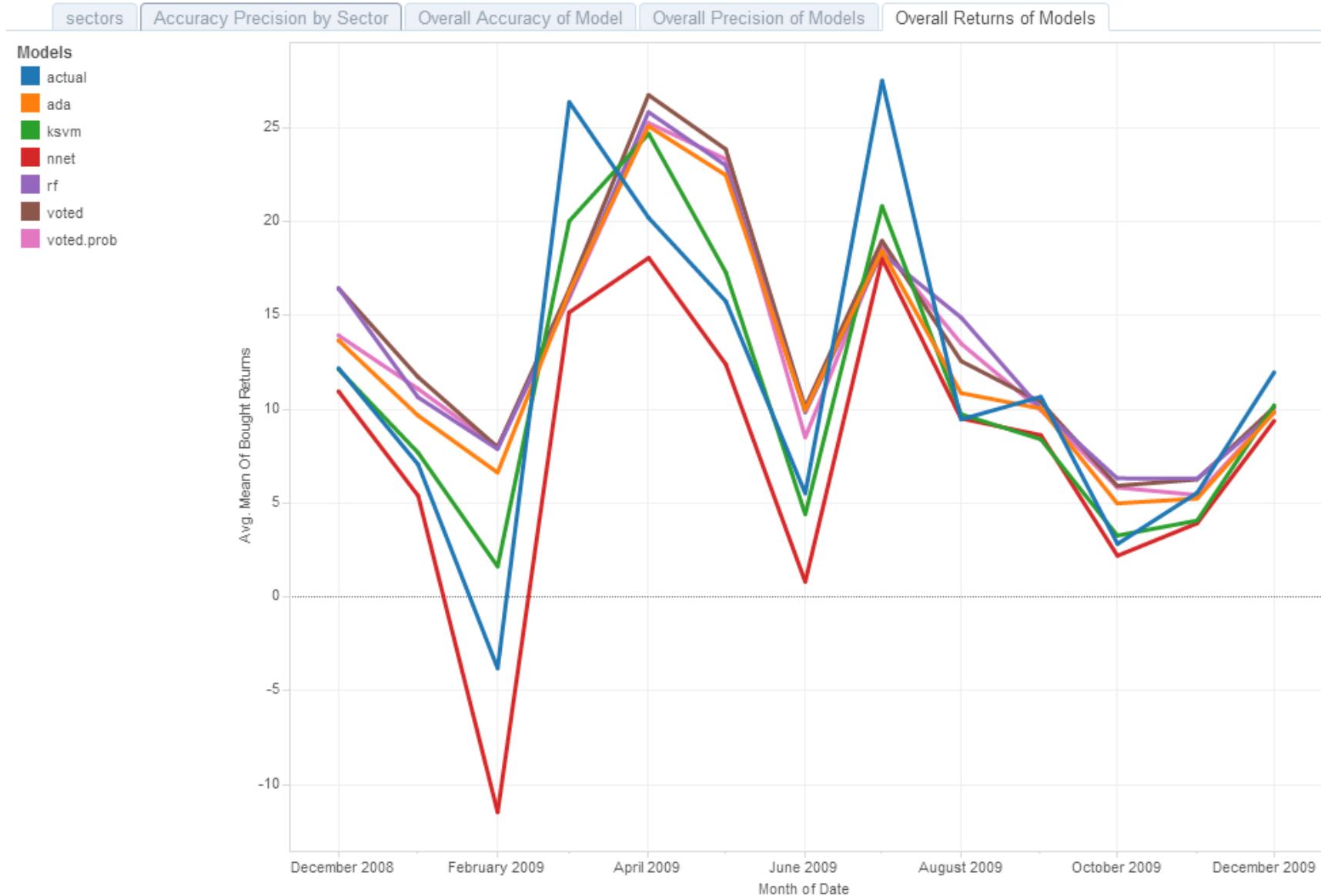
8.5 Sector - Industrials







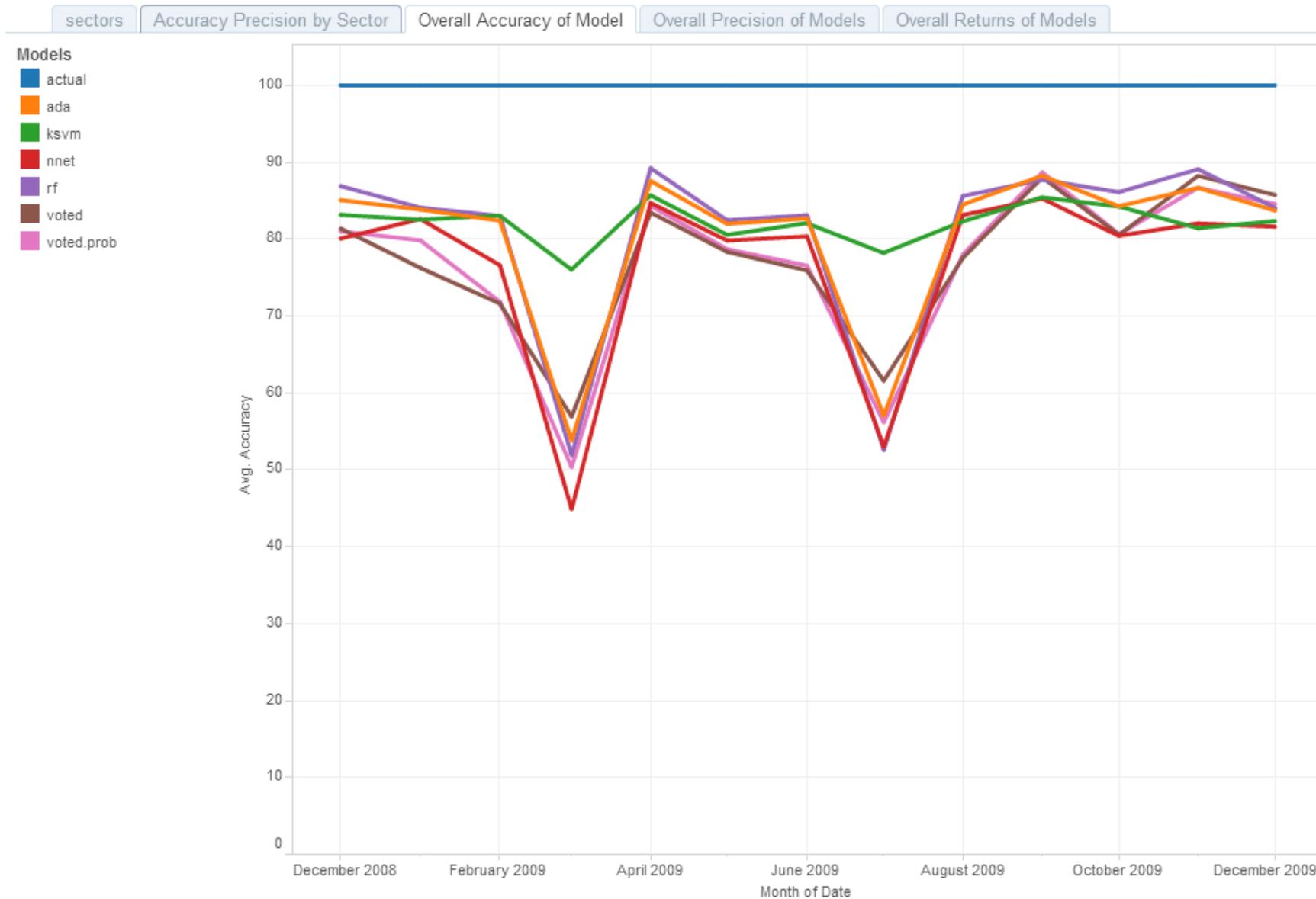




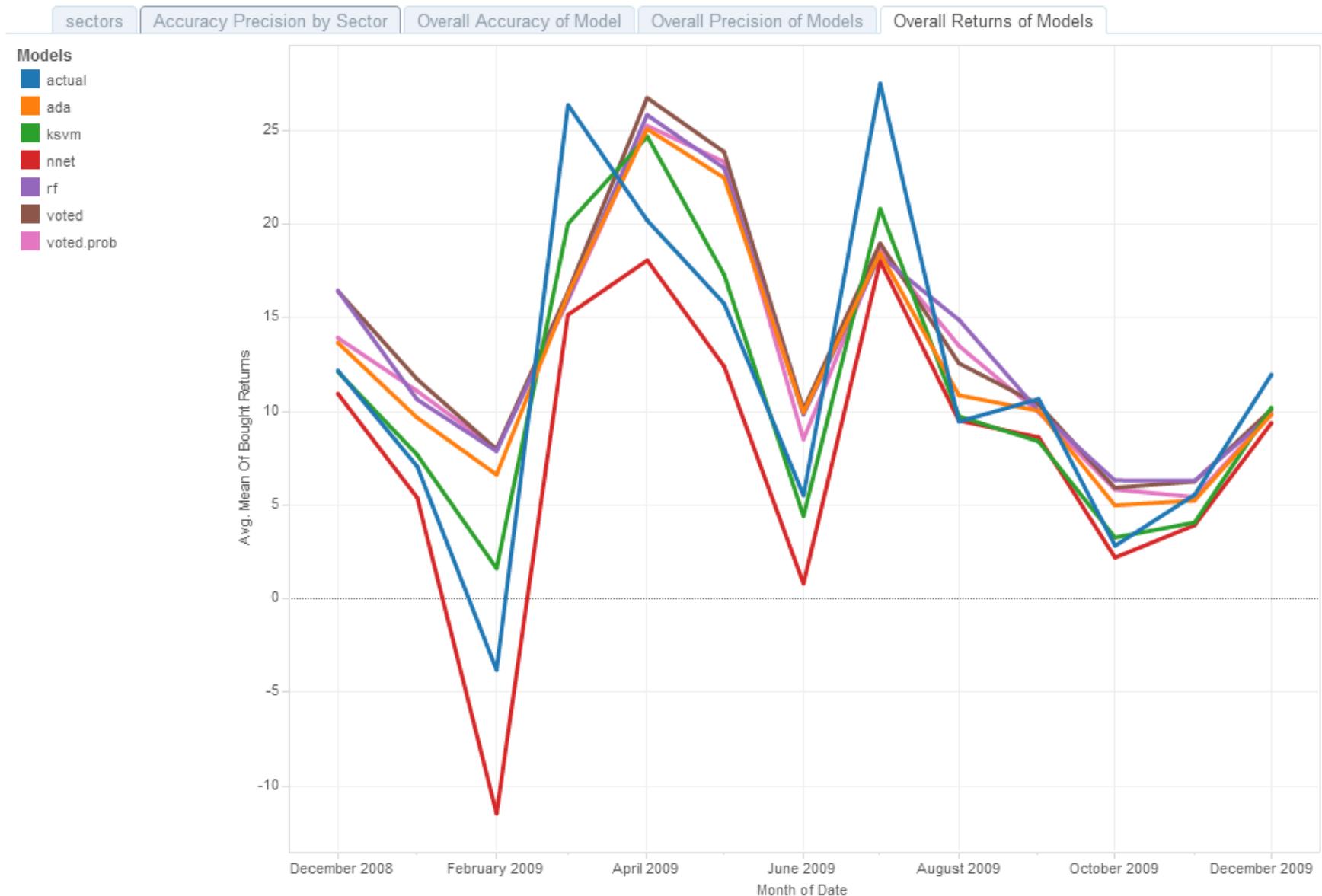
8.6 Sector - Information Technology



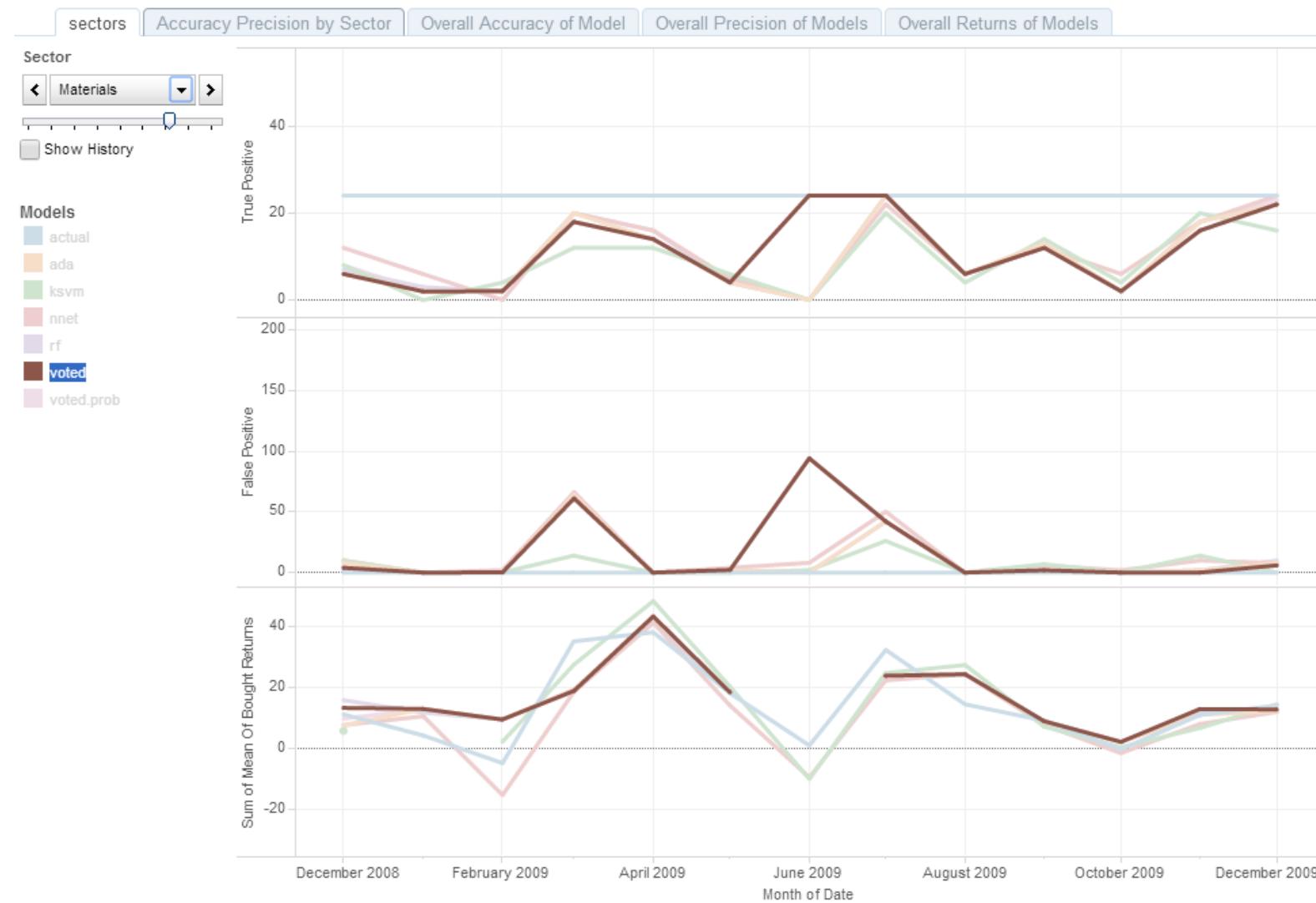




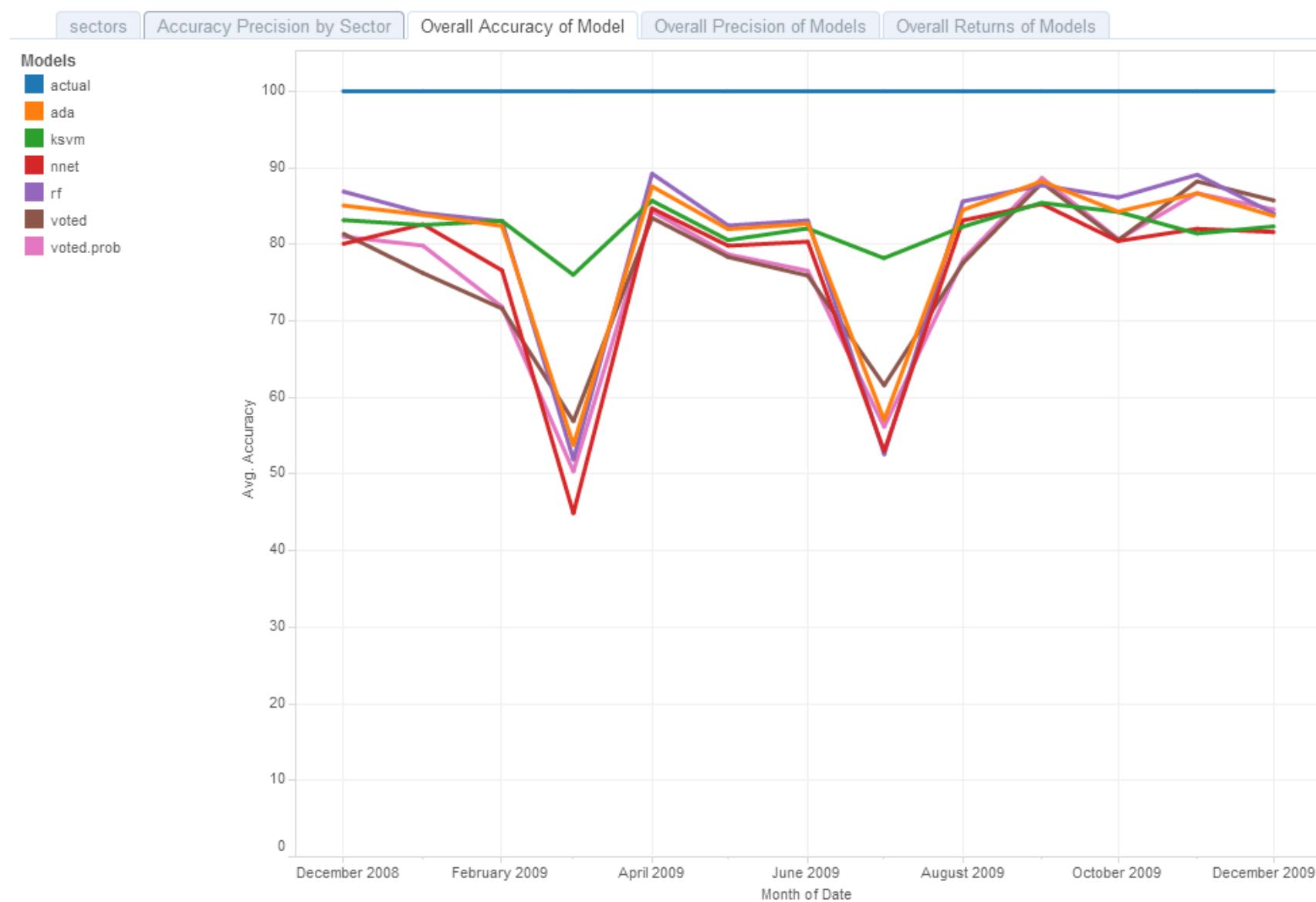




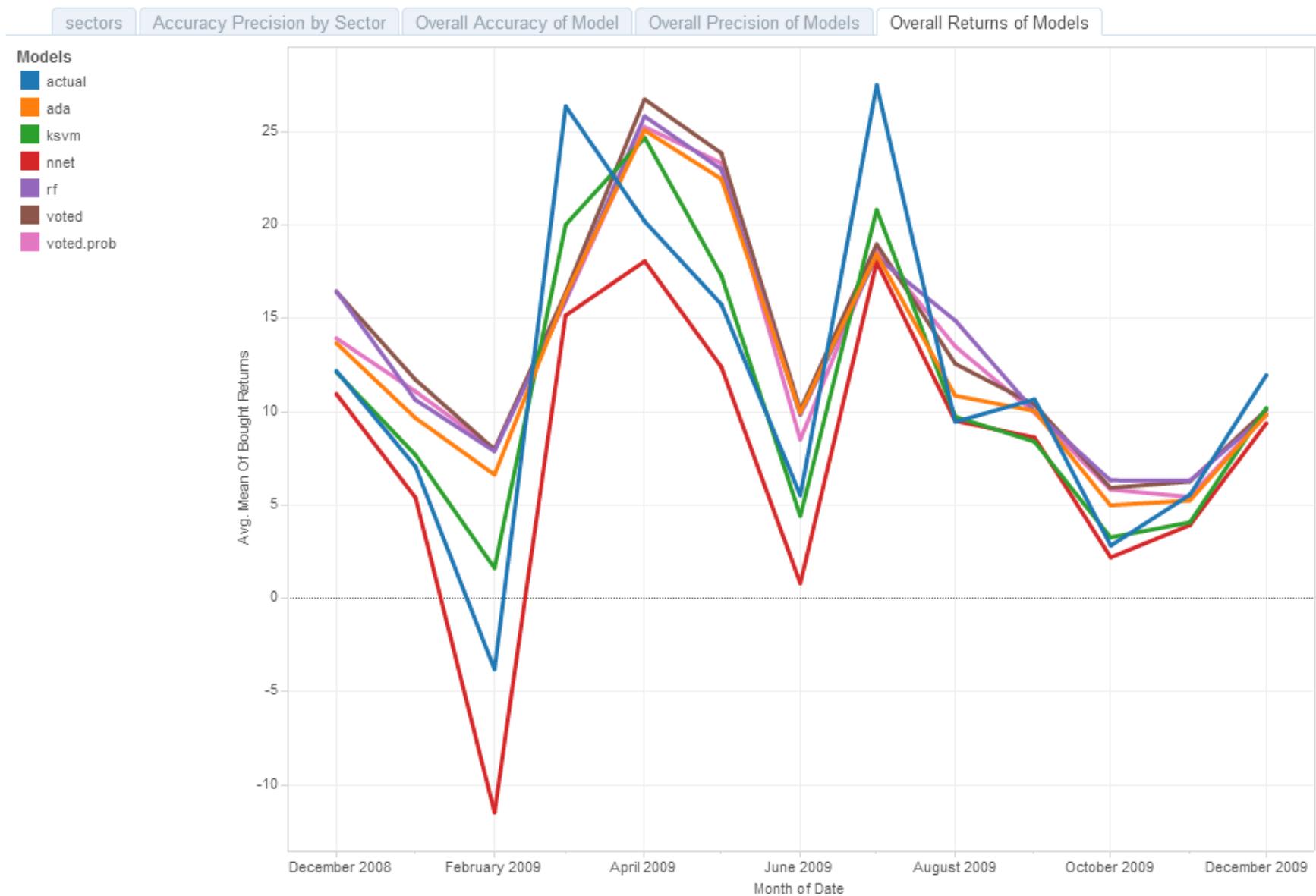
8.7 Sector - Materials



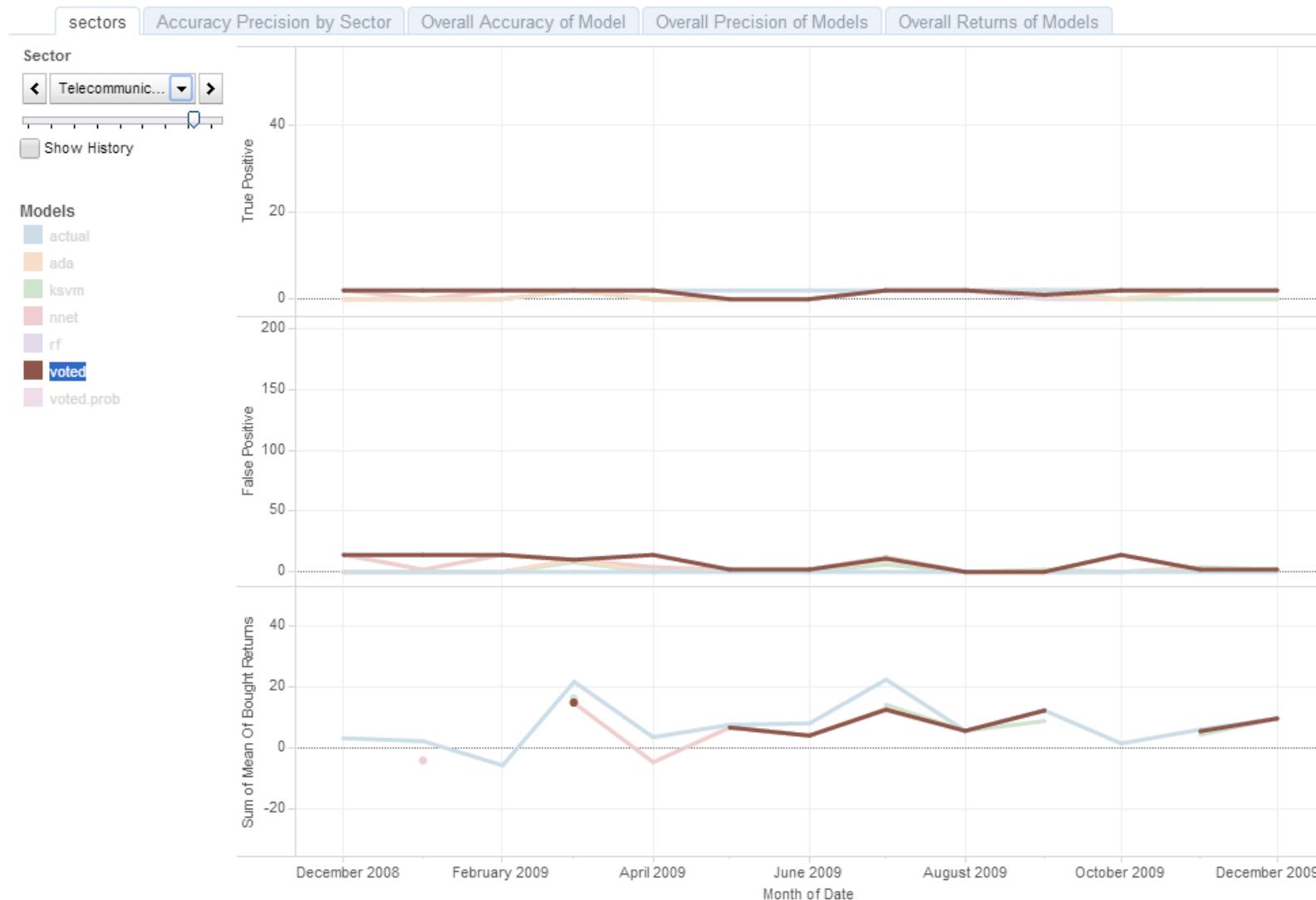




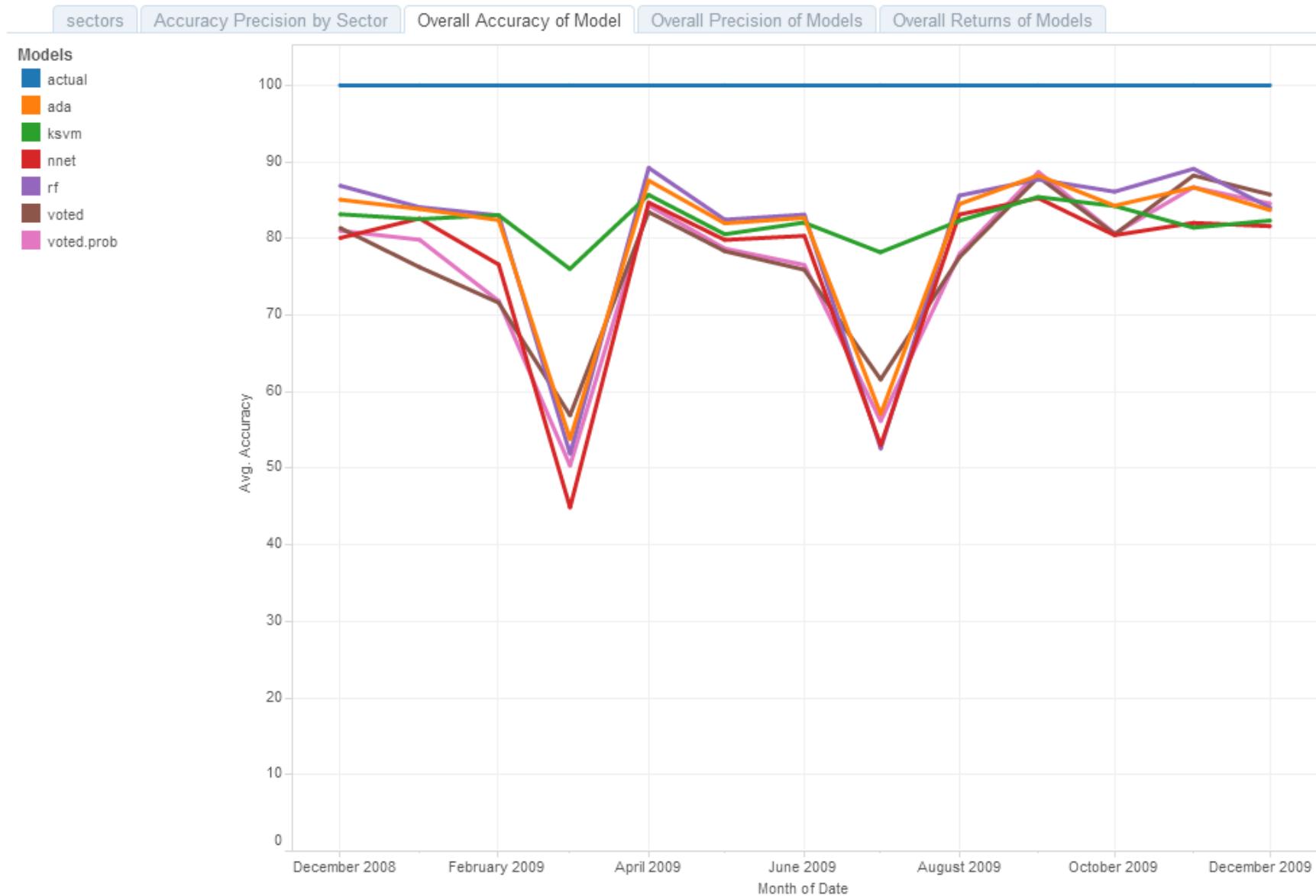


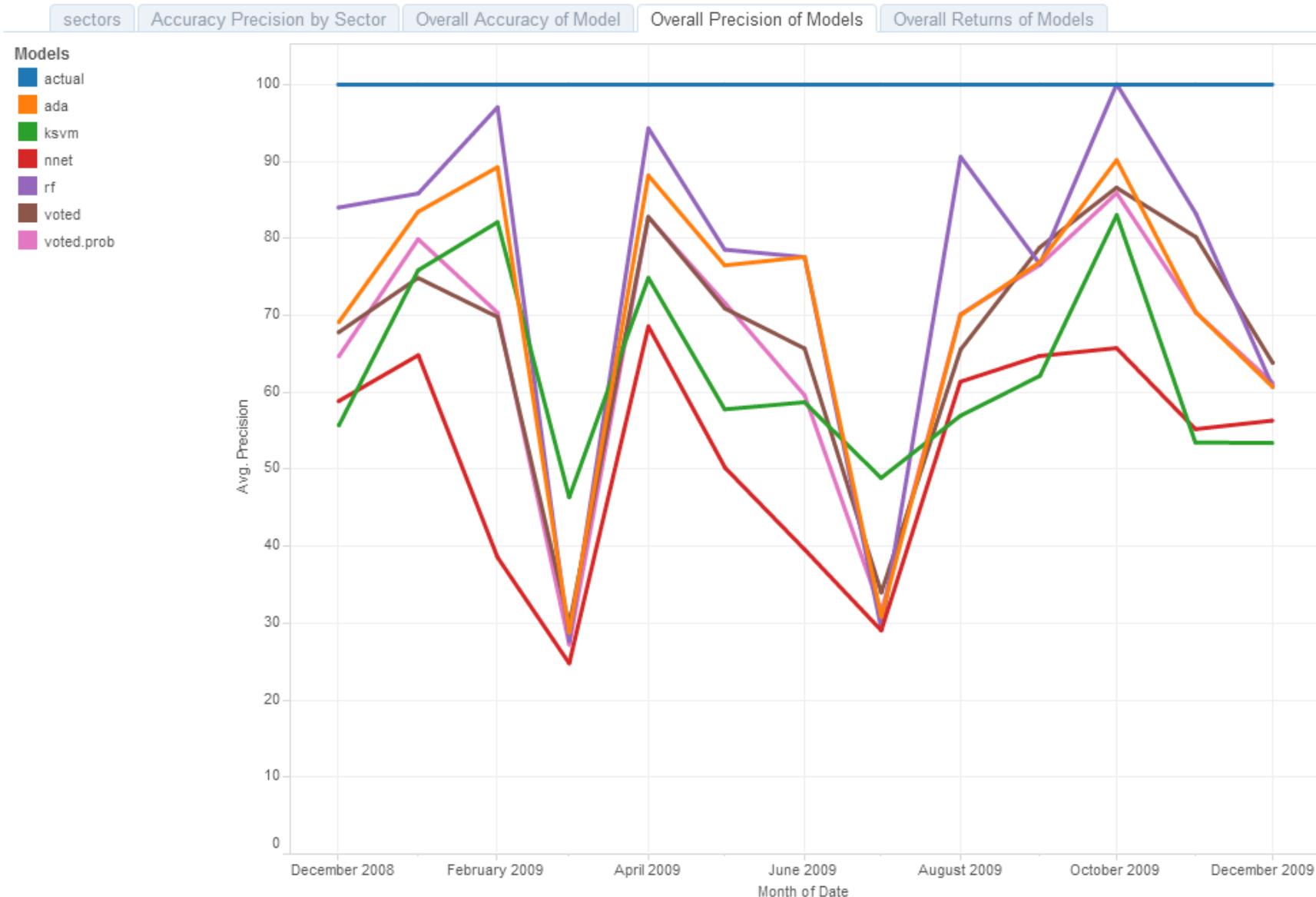


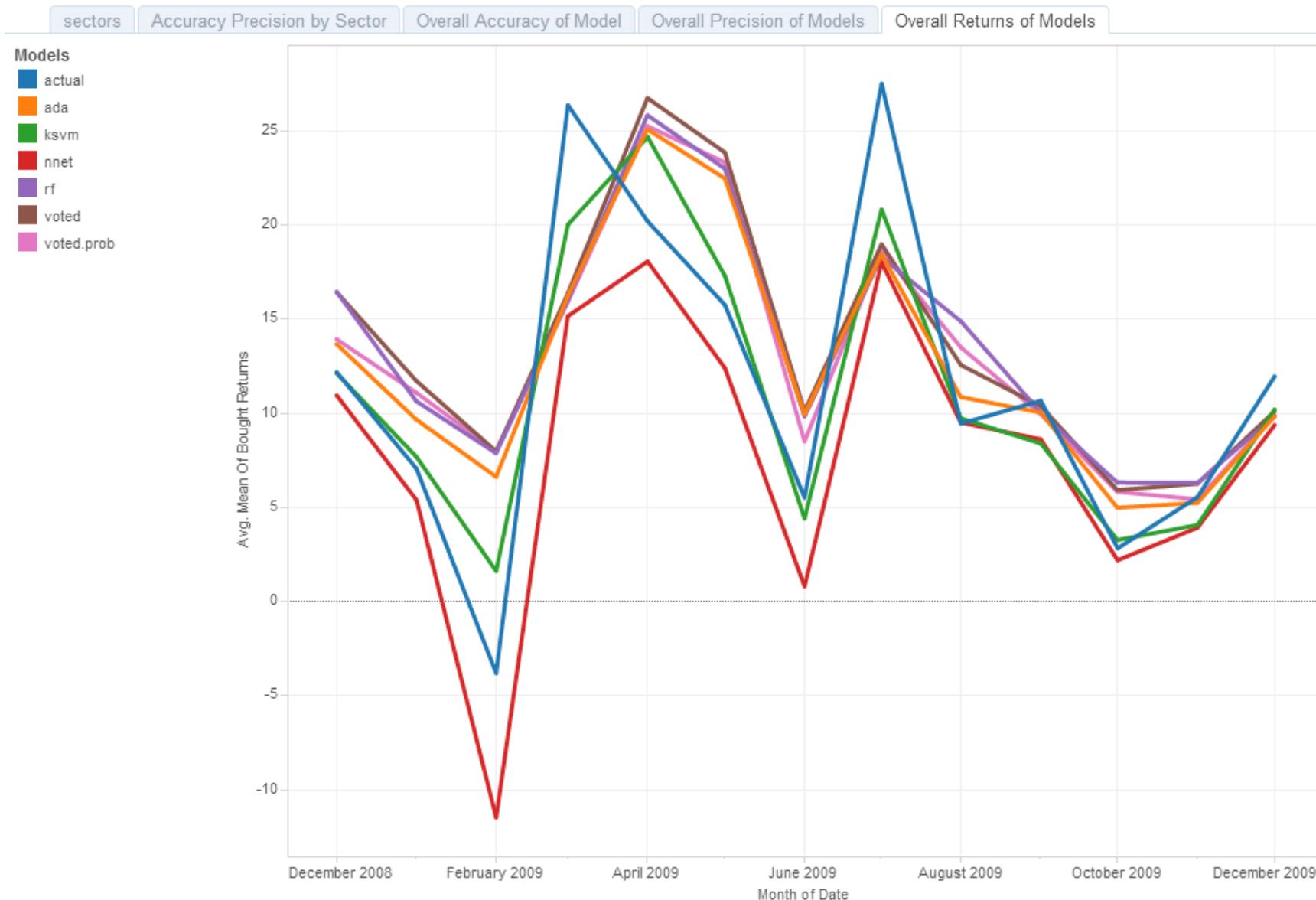
8.8 Sector - Telecommunications



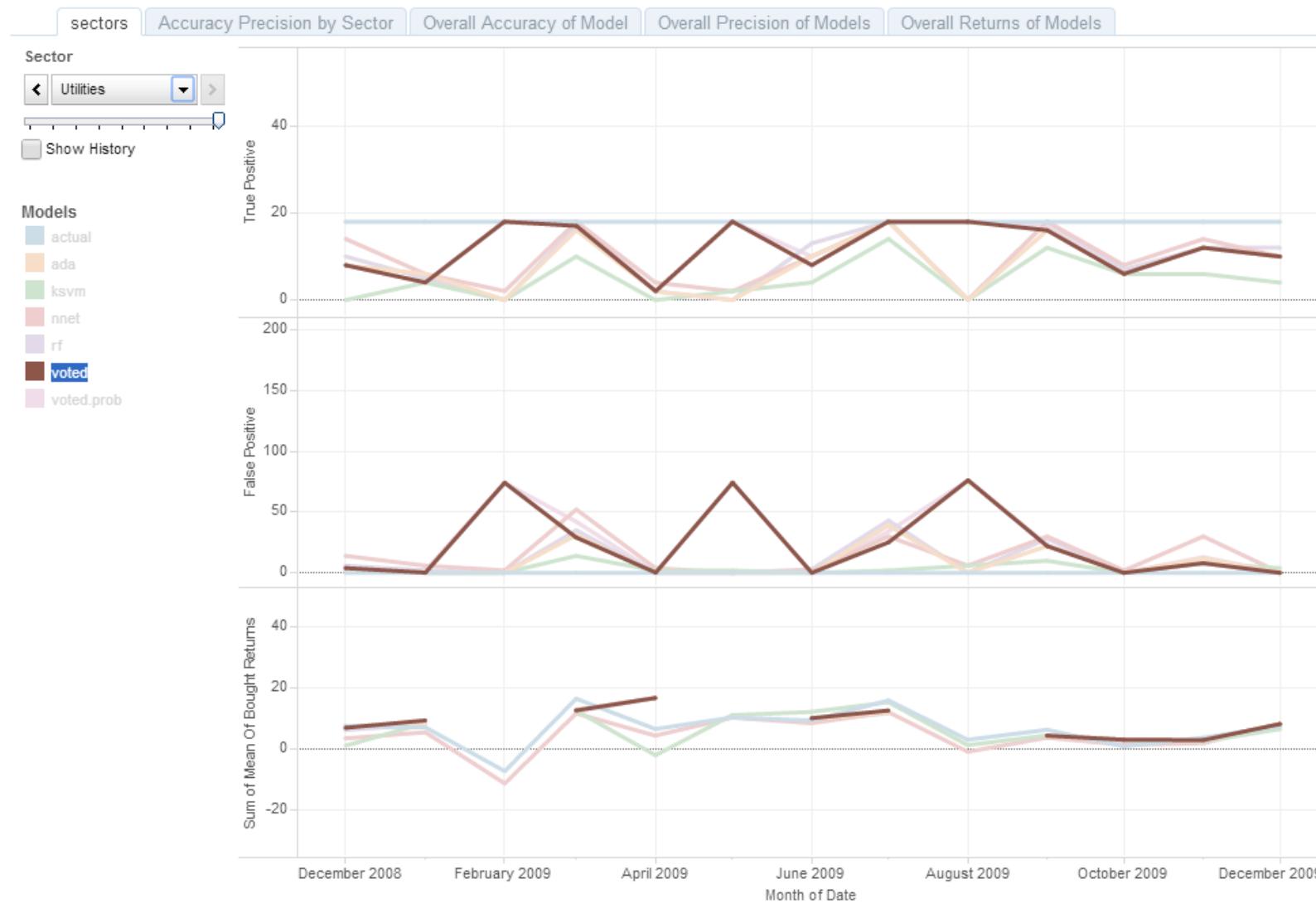


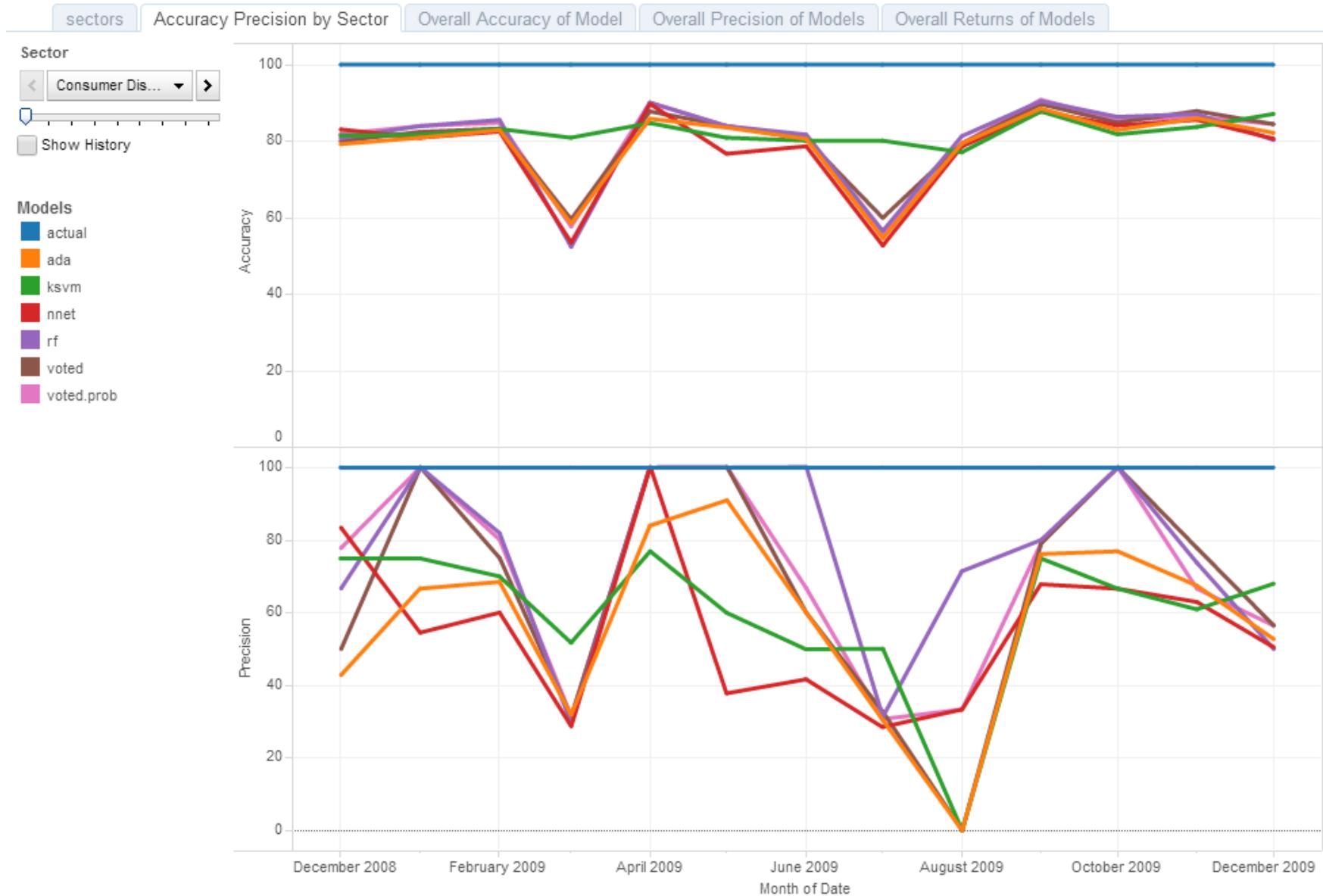


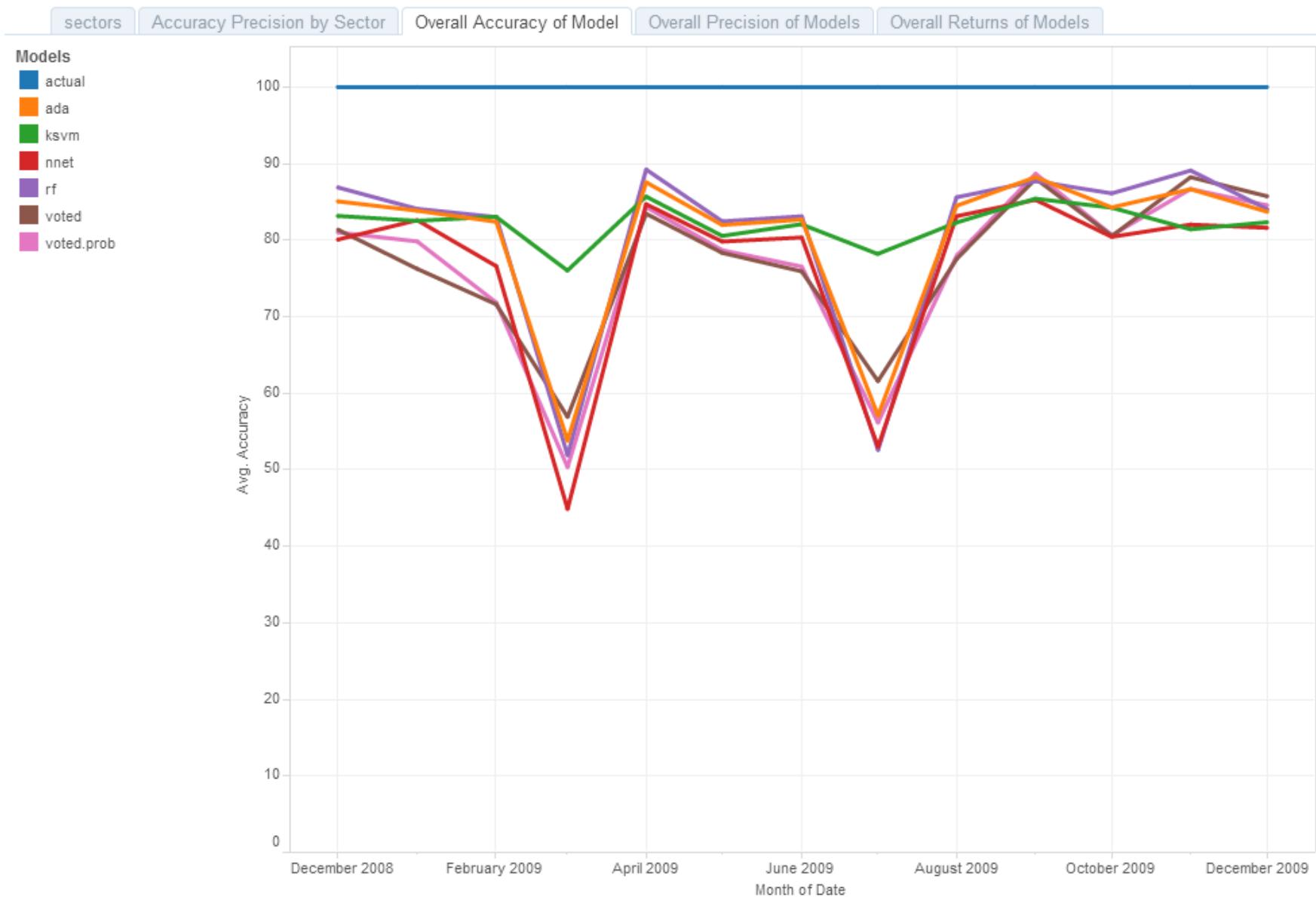




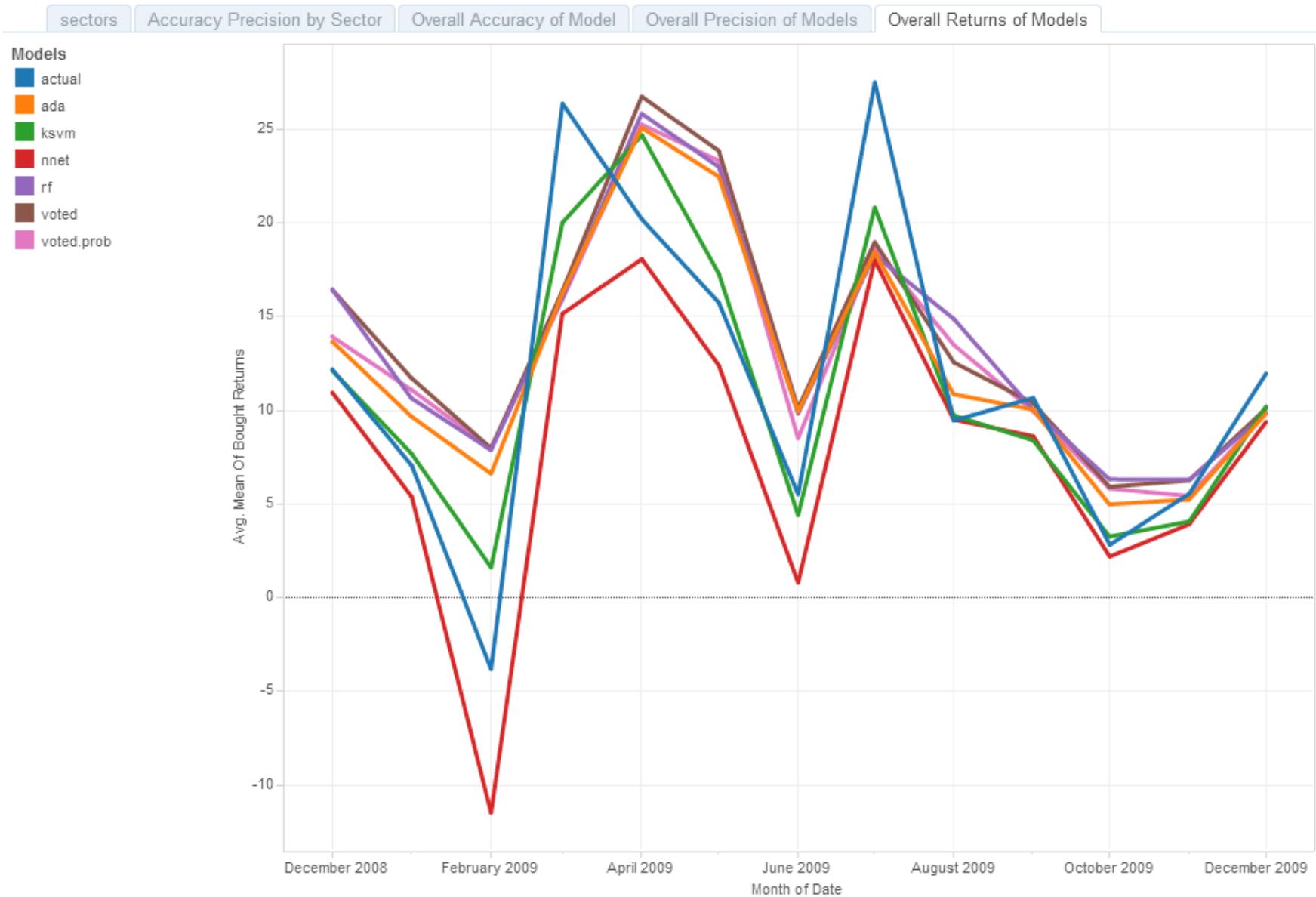
8.9 Sector - Utilities











9 Appendix B: System Design - US Stock Knowledge Model

Please find below the Domain Familiarization of the US Stock Knowledge Model

```
/*****************/
/* US Stock Knowledge Model */
/*****************/

/********************* S T A R T ********************/

/* KNOWLEDGE-MODEL us-stock;
 */
/* DOMAIN-KNOWLEDGE us-stock-domain;
 */
/* DOMAIN-SCHEMA stock-selection-schema;
 */

/* 1st of 3 main concepts stock-list */
CONCEPT stock-list;
DESCRIPTION:
    "A description of the basic characteristics of each stock";
ATTRIBUTES:
    company: STRING;
    symbol: STRING;
    market-cap (millions): FLOAT;
    volume (millions): FLOAT;
    cap-category: {large, mid, small};
    style-selection: {dividend, growth, momentum, value};
    price-chg%10d: FLOAT
    price-chg%1m: FLOAT
    price-chg%3m: FLOAT
    price-chg%6m: FLOAT
    pos-earning-growth2y: BOOLEAN
```

```

pos-earning-growth4y: BOOLEAN
pos-earning-growth6y: BOOLEAN
earning-growth3y: FLOAT
earning-growth6y: FLOAT
earning-growth9y: FLOAT
pe-fy1: FLOAT
p-sales-fy1: FLOAT
div-yield%: FLOAT
pos-div-growth5y: BOOLEAN
pos-div-growth10y: BOOLEAN
pos-div-growth15y: BOOLEAN
risk-match-cf: {-1 - 1}
stock-attractiveness-cf: {-1 - 1}
recommendation-rank: {-1 - 1}

/* RULE-TYPE 1 determines 'sector' */
sector: {consumer-staples, consumer-discretionary, energy,
financials, health-care, industrials, materials, technology,
telecommunication-services, utilities};

/* RULE-TYPE 2 determines 'risk-match-certainty-factor' */
risk-match-certainty-factor: {between -1 and 1}

/* RULE-TYPE 3 determines 'stock-attractiveness-certainty-factor' */
stock-attractiveness-certainty-factor: {between -1 and 1}

/* RULE-TYPE 4 determines 'recommendation-rank' */
recommendation-rank: {between -1 and 1};

END CONCEPT stock-list;

/* SUB SET OF 1st of 3 main concepts stock-list */
CONCEPT stock-risk-feature;
DESCRIPTION:
    "features of a stock that can affect the riskiness of a stock";
SUB-TYPE-OF: stock-list
ATTRIBUTES:
    cap-category: {large, mid, small}
    liquidity: volume-category;
    price-volatility: FLOAT;
    earning-volatility: FLOAT;

```

```

sector: {consumer-staples, consumer-discretionary, energy,
financials, health-care, industrials, materials, technology,
telecommunication-services, utilities};
stock-risk-cf: {-1 - 1}
END CONCEPT stock-risk-feature;

VALUE-TYPE volume-category;
TYPE: ORDINAL;
VALUE-LIST: {<500000,>=500000};
END VALUE-TYPE volume-category;

/* SUB SET OF 1st of 3 main concepts stock-list */
CONCEPT stock-attractiveness-feature-inherent-market-feature;
DESCRIPTION:
"features of a stock that can affect the attractiveness of a
stock";
SUB-TYPE-OF: stock-list
ATTRIBUTES:
peg: FLOAT;
ebitda-margin: FLOAT;
roic: FLOAT;
net-debt-equity-ratio: FLOAT;
stock-attractiveness-inherent-market-feature-cf: {0 - 1};
END CONCEPT stock-attractiveness-feature-inherent-market-feature;
/* SUB SET OF 1st of 3 main concepts stock-list */
CONCEPT stock-attractiveness-feature-inherent-derived;
DESCRIPTION:
"derived features of a stock that can affect the attractiveness of a
stock";
SUB-TYPE-OF: stock-list
ATTRIBUTES:
market-favourability: {high, medium, low}
//(derived from Market Favourability CF)
cap-category-attractiveness-cf: {0 - 1}
sector-attractiveness-cf: {0 - 1}
stock-attractiveness-from-market-favourability-cf: {0 - 1}
END CONCEPT stock-attractiveness-feature-inherent-derived;

/* SUB SET OF 1st of 3 main concepts stock-list */

```

```

CONCEPT stock-style-feature;

DESCRIPTION:
"Stock level style feature to be matched with Customer Style Profile ";

SUB-TYPE-OF: stock-list

ATTRIBUTES:

style-selection: {dividend, growth, momentum, value};
price-momentum-preference: {<4weeks,4weeks-12weeks,>12weeks};
no-consecutive-years-increasing-earnings: {2,4,6};
earning-growth-rate: {10%-20%,20%-30%,>30%};
yrs-used-for-earning-growth: {3,6,9}
price-earning-ratio: {<6,6-10,10-15};
price-sales-ratio: {0-0.5,0.5-1,1-1.5};
dividend-yield-required: {0-2%,2%-4%,>4%};
dividend-growth-desired: {5%-8%,8%-12%,>12%};
no-consecutive-positive-years-dividend-growth: {5-10,10-15,>15,};

END CONCEPT stock-style-feature;

```

/ 2nd of 3 main concepts macroeconomic-and-stock-market-conditions */*

```

CONCEPT macroeconomic-and-stock-market-conditions;

```

DESCRIPTION:

"overall macroeconomic and stock market conditions";

ATTRIBUTES:

```

market-favourability-cf: {between 0 and 1};
macroeconomic-conditions-cf: {between -1 and 1};
stock-market-conditions-cf: {between -1 and 1};

```

END CONCEPT macroeconomic-and-stock-market-conditions;

/ SUB SET OF 2nd of 3 main concepts macroeconomic-and-stock-market-conditions */*

```

CONCEPT macroeconomic-conditions;

```

DESCRIPTION:

"Economic factors that influence the state of the whole economy, such as changes in employment levels, gross national product (GNP), and prices (deflation or inflation)";

SUB-TYPE-OF: macroeconomic-and-stock-market-conditions

ATTRIBUTES:

```

inflation-rate: {<0% or >4%: bearish, 0% - 2%: bullish, 2% - 4%: neutral};
unemployment-rate: {<5%: bullish, 5%-6%: neutral, >6%: bearish};
unemployment-trend: {uptrend,flat,downtrend}
average-weekly-unemployment-claim: {<350,000: bullish, 350,000 - 400,000: neutral, >400,000: bearish};

```

```

unemployment-claim-trend: {uptrend,flat,downtrend};
yield-curve-slope: {upward-sloping,flat,downward-sloping};
retail-sale-trend: {uptrend,flat,downtrend};
pmi: {<50,>=50};

END CONCEPT macroeconomic-conditions;

/* SUB SET OF 2nd of 3 main concepts macroeconomic-and-stock-market-
conditions */

CONCEPT stock-market-conditions;
DESCRIPTION:
"Market conditions is a term that refers to the state of an stock";
SUB-TYPE-OF: macroeconomic-and-stock-market-conditions

ATTRIBUTES:
indicator-1: {very-bearish,bearish,neutral,bullish,very-bullish};
indicator-2: {very-bearish,bearish,neutral,bullish,very-bullish};
indicator-3: {very-bearish,bearish,neutral,bullish,very-bullish};
indicator-4: {very-bearish,bearish,neutral,bullish,very-bullish};
indicator-5: {very-bearish,bearish,neutral,bullish,very-bullish};
END CONCEPT stock-market-conditions;

/* 3rd of 3 main concepts customer-profile */

CONCEPT customer-profile;
DESCRIPTION:
"Customer level profile; each customer has one profile";
ATTRIBUTES:
customer-name: STRING;
customer-address: STRING;
customer-phone: NATURAL;
END CONCEPT customer-profile;

/* SUB SET OF 3rd of 3 main concepts customer-profile */

CONCEPT customer-risk-profile;
DESCRIPTION:
"features of a customer's ability to buy stock";
SUB-TYPE-OF: customer-profile

ATTRIBUTES:
overall-risk-tolerance-cf: {between -1 and 1};
risk-tolerance-ability-cf: {between -1 and 1};
risk-tolerance-willingness-cf: {between -1 and 1};
END CONCEPT customer-risk-profile;

```

```
/* SUB SET OF customer-risk-profile, which is SUB SET OF 3rd of 3 main
concepts customer-profile */

CONCEPT customer-risk-profile-ability;

DESCRIPTION:
"features of a customer's ability to buy stock";

SUB-TYPE-OF: customer-risk-profile

ATTRIBUTES:
age: {20-35,36-45,46-55,56-65}
net-assets: {<25000,25000-100000,100000-250000,250000-500000,>500000};
job-stability: {very-unstable,rather-unstable,a-bit-stable,stable,very-
stable};
monthly-income: {<2500,2500-5000,5000-7500,7500-10000,>10000};
liquidity-needs: {very-low,low,medium,high,very-high};

END CONCEPT customer-risk-profile-ability;
```

```
/* SUB SET OF customer-risk-profile, which is SUB SET OF 3rd of 3 main
concepts customer-profile */

CONCEPT customer-risk-profile-willingness;

DESCRIPTION:
"features of a customer's willingness to buy a stock";

SUB-TYPE-OF: customer-risk-profile

ATTRIBUTES:
thrill-factor: {very-low,low,medium,high,very-high};
job-security-preference: {very-high,high,medium,low,very-low};
investment-risk-tolerance: {very-low,low,medium,high,very-high};
job-compensation-preference: {all-salary,mainly-salary,equal-mix-salary-
commission,mainly-commission,all-commission};
aversion-to-previous-stock-loss: {very-high,high,medium,low,very-low};

END CONCEPT customer-risk-profile-willingness;
```

```
/* SUB SET OF 3rd of 3 main concepts customer-profile */

CONCEPT customer-style-profile;

DESCRIPTION:
"features of customer's personality";

SUB-TYPE-OF: customer-profile

ATTRIBUTES:
style-selection: {dividend, growth, momentum, value};
price-momentum-preference: {<4weeks,4weeks-12weeks,>12weeks};
no-consecutive-years-increasing-earnings: {2,4,6};
earning-growth-rate: {10%-20%,20%-30%,>30%};
```

```

yrs-used-for-earning-growth: {3,6,9}
price-earning-ratio: {<6,6-10,10-15};
price-sales-ratio: {0-0.5,0.5-1,1-1.5};
dividend-yield-required: {0-2%,2%-4%,>4%};
dividend-growth-desired: {5%-8%,8%-12%,>12%};
no-consecutive-positive-years-dividend-growth: {5-10,10-15,>15,};
END CONCEPT customer-style-profile;

```

```

/*****************/
/* 4 RULE TYPES */
/*****************/

/* 1st of 4 RULE TYPES */
RULE-TYPE update-stock-risk;
ANTECEDENT: customer-risk-profile; CARDINALITY: 1;
stock-risk-feature; CARDINALITY: 1+;
CONSEQUENT: stock-risk-feature; CARDINALITY: 1+;
CONNECTION-SYMBOL: update-stock-risk;
END RULE-TYPE update-stock-risk;

/* 2nd of 4 RULE TYPES */
RULE-TYPE filter-stock-style;
ANTECEDENT: customer-style-profile;
CARDINALITY: 1;
style-feature; CARDINALITY: 1+;
CONSEQUENT: stock-style-feature: 1+;
CONNECTION-SYMBOL: filter-stock-style;
END RULE-TYPE filter-stock-style;

/* 3rd of 4 RULE TYPES */
RULE-TYPE update-stock-favourability;
ANTECEDENT: macroeconomic-and-stock-market-conditions;
CARDINALITY: 1;
stock-attractiveness-feature-derived;
CARDINALITY: 1+;
CONSEQUENT: stock-attractiveness-feature-derived;
CARDINALITY: 1+;
CONNECTION-SYMBOL: update-stock-favourability;
END RULE-TYPE update-stock-favourability;

```

```

/* 4th of 4 RULE TYPES */

RULE-TYPE recommend-stock;
ANTECEDENT: stock-list;
    CARDINALITY: 1;
CONSEQUENT: stock-list;
    CARDINALITY: 1;
CONNECTION-SYMBOL: recommend-stock;
END RULE-TYPE recommend-stock;

/*****************/
END DOMAIN-SCHEMA stock-selection-schema;
/*****************/

/*****************/
/* Detailed RULEs */
/*****************/

/* 1st of 4 KNOWLEDGE-BASEs */
KNOWLEDGE-BASE filter-stock-style-model;
USES:
    filter-stock-style FROM filter-stock-style-kb;

/* Filter By dividend Style */
customer-style-profile.style-selection = 'dividend'

AND
/* First By dividend Yield */
(customer-style-profile.dividend-yield-required = '0% - 2%'
filter-stock-style stock-list.div-yield% >=0 AND stock-list.div-yield%
< 2

customer-style-profile.dividend-yield-required = '2% - 4%'
filter-stock-style stock-list.div-yield% >= 2
AND stock-list.div-yield% < 4

customer-style-profile.dividend-yield-required = '>=4%'
filter-stock-style CHARACTERISTICS stock-list.div-yield% >= 4

```

)

AND

```
/* Second By minimum no of years of consecutive dividend growth */
(
customer-style-profile.no-consecutive-positive-
years-dividend-growth = '>=5 years'
filter-stock-style CHARACTERISTICS stock-list.pos-div-growth5y = TRUE
```

OR

```
customer-style-profile.no-consecutive-positive-
years-dividend-growth = '>=10 years'
filter-stock-style CHARACTERISTICS stock-list.pos-div-growth10y = TRUE
```

OR

```
customer-style-profile.no-consecutive-positive-
years-dividend-growth = '>=15 years'
filter-stock-style stock-list.pos-div-growth15y = TRUE
```

AND

```
/* Third By dividend growth desired */
(customer-style-profile.dividend-growth-desired = '>=5%'
filter-stock-style CHARACTERISTICS stock-list.Divgrowth% >=5
```

OR

```
customer-style-profile.dividend-growth-desired = '>=8%'
filter-stock-style CHARACTERISTICS stock-list.Divgrowth% >=8
```

OR

```
customer-style-profile.dividend-growth-desired = '>=12%'
filter-stock-style stock-list.Divgrowth% >=12
```

)

```

/* Filter By growth Style */
customer-style-profile.style-selection = 'growth'

AND

/* First By no of consecutive years of increasing earnings */
(customer-style-profile.no-consecutive-years-increasing-earnings = 2
filter-stock-style stock-list.pos-earning-growth2y = TRUE

OR

customer-style-profile.no-consecutive-years-
increasing-earnings = 4
filter-stock-style stock-list.pos-earning-growth4y = TRUE

OR

customer-style-profile.no-consecutive-years-
increasing-earnings = 6
filter-stock-style stock-list.pos-earning-growth6y = TRUE

)

AND

/* Second By earnings growth rate */
(customer-style-profile.earnings-growth-rate = '10% - 20%'
AND
customer-style-profile.yrs-used-for-earnings-growth = '3 years'

filter-stock-style stock-list.earning-growth3y > 10 and stock-
list.earning-growth3y <= 20

OR

(customer-style-profile.earnings-growth-rate = '20% - 30%'
AND
customer-style-profile.yrs-used-for-earnings-growth = '3 years')

filter-stock-style stock-list.earning-growth3y > 20 and
stock-list.earning-growth3y <= 30

```

OR

```
(customer-style-profile.earnings-growth-rate = '>=30%'  
AND  
customer-style-profile.yrs-used-for-earnings-growth = '3 years')
```

```
filter-stock-style stock-list.earning-growth3y >= 30%
```

OR

```
(customer-style-profile.earnings-growth-rate = '10% - 20%'  
AND  
customer-style-profile.yrs-used-for-earnings-growth = '6 years')
```

```
filter-stock-style stock-list.earning-growth6y > 10 and  
stock-list.earning-growth6y <= 20
```

OR

```
(customer-style-profile.earnings-growth-rate = '20% - 30%'  
AND  
customer-style-profile.yrs-used-for-earnings-growth = '6 years')
```

```
filter-stock-style stock-list.earning-growth6y > 20 and  
stock-list.earning-growth6y <= 30
```

OR

```
(customer-style-profile.earnings-growth-rate = ' >=30%'  
AND  
customer-style-profile.yrs-used-for-earnings-growth = '6 years')
```

```
filter-stock-style stock-list.earning-growth6y >= 30
```

OR

```
(customer-style-profile.earnings-growth-rate = '10% - 20%'  
AND  
customer-style-profile.yrs-used-for-earnings-growth = '9 years')
```

```
filter-stock-style stock-list.earning-growth9y > 10 and stock-
list.earning-growth9y <= 20
```

OR

```
(customer-style-profile.earnings-growth-rate = '20% - 30%'  
AND  
customer-style-profile.yrs-used-for-earnings-growth = '9 years')
```

```
filter-stock-style stock-list.earning-growth9y > 20 and stock-
list.earning-growth9y <= 30
```

OR

```
(customer-style-profile.earnings-growth-rate = '>=30%'  
AND  
customer-style-profile.yrs-used-for-earnings-growth = '9 years')
```

```
filter-stock-style stock-list.earning-growth9y >= 30
```

```
/* Filter By momentum Style */  
customer-style-profile.style-selection = 'momentum'
```

AND

```
(customer-style-profile.price-momentum-preference = '< 4 weeks'  
filter-stock-style stock-list.price-chg%10d > 5 and stock-list.price-
chg%1m > 10
```

OR

```
customer-style-profile.price-momentum-preference = '4 weeks - 12 weeks'  
filter-stock-style stock-list.price-chg%1m > 10 and stock-list.price-
chg%3m > 20
```

OR

```
customer-style-profile.price-momentum-preference = '> 12 weeks'  
filter-stock-style stock-list.price-chg%3m > 20 and stock-list.price-  
chg%6m > 40  
  
/* Filter By value Style */  
customer-style-profile.style-selection = 'value'  
AND  
  
/* First by price-earnings-ratio */  
(customer-style-profile.price-earnings-ratio = '5-8'  
filter-stock-style stock-list.pe-fy1 >5 and stock-list.pe-fy1 <=8  
  
OR  
  
customer-style-profile.price-earnings-ratio = '8-11'  
filter-stock-style stock-list.pe-fy1 >8 and stock-list.pe-fy1 <=11  
  
OR  
  
customer-style-profile.price-earnings-ratio = '11-15'  
filter-stock-style stock-list.pe-fy1 >11 and stock-list.pe-fy1 <=15  
)  
  
AND  
/* Second by price-sales-ratio */  
(customer-style-profile.price-sales-ratio = '0 - 0.5'  
filter-stock-style stock-list.p-sales-fy1 >0 and stock-list.p-sales-fy1  
<=0.5  
  
OR  
  
customer-style-profile.price-sales-ratio = '0.5 - 1'  
filter-stock-style stock-list.p-sales-fy1 >0.5 and stock-list.p-sales-  
fy1 <=11  
  
OR  
  
customer-style-profile.price-sales-ratio = '1 - 1.5'  
filter-stock-style stock-list.p-sales-fy1 >1  
and stock-list.p-sales-fy1 <=1.5
```

```

    )

END KNOWLEDGE-BASE filter-stock-style-model;

/* 2nd of 4 KNOWLEDGE-BASEs */

KNOWLEDGE-BASE update-stock-favourability-model;
USES:
    update-stock-favourability FROM stock-selection-schema;
EXPRESSIONS:
/* Abstraction rules */
    inflation-rate: {bearish,bullish,neutral};
    unemployment-rate: {bullish,neutral,bearish};
    unemployment-trend: {uptrend,flat,downtrend}
    average-weekly-unemployment-claim: {bullish,neutral,bearish};
    unemployment-claim: {uptrend,flat,downtrend};
    yield-curve-slope: {upward-sloping,flat,downward-sloping};
    retail-sale: {uptrend,flat,downtrend};
    pmi: {<50,>=50};

/* Macroeconomic conditions */

/* 1. Inflation Rate for the past 6 months */
macroeconomic-conditions.inflation-rate < 0%
OR
macroeconomic-conditions.inflation-rate > 4%
update-stock-favourability inflation-rate-cf = -0.8

macroeconomic-conditions.inflation-rate = '0% - 2%'
update-stock-favourability inflation-rate-cf = +0.9

macroeconomic-conditions.inflation-rate = '2% - 4%'
update-stock-favourability inflation-rate-cf = +0.5

/* 2. Unemployment Rate over the past 6 months (Absolute) */
macroeconomic-conditions.unemployment-rate < 5%
update-stock-favourability unemployment-rate-cf = +0.7

macroeconomic-conditions.unemployment-rate = '5% - 6%'
update-stock-favourability unemployment-rate-cf = +0.5

```

```

macroeconomic-conditions.unemployment-rate > 6%
update-stock-favourability unemployment-rate-cf = +0.3

/* 3. Unemployment trend over the past 6 months */
macroeconomic-conditions.unemployment-trend = uptrend
update-stock-favourability unemployment-trend-cf = -0.7

macroeconomic-conditions.unemployment-trend = flat
update-stock-favourability unemployment-trend-cf = +0.5

macroeconomic-conditions.unemployment-trend = downtrend
update-stock-favourability unemployment-trend-cf = +0.8

/* 4. Average-weekly-unemployment-claim for the past 4 weeks */
macroeconomic-conditions.average-weekly-unemployment-claim < 350,000
update-stock-favourability unemployment-claim-cf = +0.7

macroeconomic-conditions.average-weekly-unemployment-claim =
'350,000 - 400,000'
update-stock-favourability unemployment-claim-cf = +0.5

macroeconomic-conditions.average-weekly-unemployment-claim > 400,000
update-stock-favourability unemployment-claim-cf = -0.6

/* 5. Unemployment-claims over the past 6 months - trend */
macroeconomic-conditions.unemployment-claim-trend = uptrend
update-stock-favourability unemployment-claim-trend-cf = -0.7

macroeconomic-conditions.unemployment-claim-trend = flat
update-stock-favourability unemployment-claim-trend-cf = +0.5

macroeconomic-conditions.unemployment-claim-trend = downtrend
update-stock-favourability unemployment-claim-trend-cf = +0.8

/* 6. Yield-curve-slope */
macroeconomic-conditions.yield-curve-slope = upward-sloping
update-stock-favourability yield-curve-slope-cf = +0.9

macroeconomic-conditions.yield-curve-slope = flat

```

```

update-stock-favourability yield-curve-slope-cf = +0.5

macroeconomic-conditions.yield-curve-slope = downward-sloping
update-stock-favourability yield-curve-slope-cf = -0.8

/* 7. Retail Sales Trend */
macroeconomic-conditions.retail-sale-trend = Uptrend
update-stock-favourability retail-sale-trend-cf = +0.8

macroeconomic-conditions.retail-sale-trend = Flat
update-stock-favourability retail-sale-trend-cf = +0.5

macroeconomic-conditions.retail-sale-trend = Downtrend
update-stock-favourability retail-sale-trend-cf = -0.7

/* 8. ISM Manufacturing PMI */
macroeconomic-conditions.pmi > 50
update-stock-favourability pmi-cf = 0.7

macroeconomic-conditions.pmi <= 50
update-stock-favourability pmi-cf = 0.4

/* Working Goal - macroeconomic-conditions */
inflation-rate-cf
AND
unemployment-rate-cf
AND
unemployment-trend-cf
AND
unemployment-claim-cf
AND
unemployment-claim-trend-cf
AND
yield-curve-slope-cf
AND
retail-sale-trend-cf
AND
pmi-cf

CREATE WORKING GOAL =

```

```
macroeconomic-and-stock-market-conditions.macroeconomic-conditions-cf
```

```
/* stock market conditions */

/* 1. Indicator 1 - 20D MA Ratio of no of positive returns/no of
negative
returns */
stock-market-conditions.indicator-1 = '0 - 0.3 (very-bearish)'
update-stock-favourability indicator-1-cf = -0.8

stock-market-conditions.indicator-1 = '0.3 - 0.8 (bearish)'
update-stock-favourability indicator-1-cf = -0.6

stock-market-conditions.indicator-1 = '0.8 - 1.2 (neutral)'
update-stock-favourability indicator-1-cf = +0.5

stock-market-conditions.indicator-1 = '1.2 - 2 (bullish)'
update-stock-favourability indicator-1-cf = +0.6

stock-market-conditions.indicator-1 = '>2 (very-bearish)'
update-stock-favourability indicator-1-cf = +0.8

/* 2. Indicator 2 - 20D MA Ratio of no of daily +4% moves/no of -4%
moves
*/
stock-market-conditions.indicator-2 = 0 - 0.3 (very-bearish)
update-stock-favourability indicator-2-cf = -0.8

stock-market-conditions.indicator-2 = 0.3 - 0.8 (bearish)
update-stock-favourability indicator-2-cf = -0.6

stock-market-conditions.indicator-2 = 0.8 - 1.2 (neutral)
update-stock-favourability indicator-2-cf = +0.5

stock-market-conditions.indicator-2 = 1.2 - 2 (bullish)
update-stock-favourability indicator-2-cf = +0.6

stock-market-conditions.indicator-2 = >2 (very-bearish)
update-stock-favourability indicator-2-cf = +0.8
```

```

/* 3. Indicator 3 - 20D MA Ratio of no of quarterly +25% moves/no of -25%
moves */

stock-market-conditions.indicator-3 = 0 - 0.3 (very-bearish)
update-stock-favourability indicator-3-cf = -0.8

stock-market-conditions.indicator-3 = 0.3 - 0.8 (bearish)
update-stock-favourability indicator-3-cf = -0.6

stock-market-conditions.indicator-3 = 0.8 - 1.2 (neutral)
update-stock-favourability indicator-3-cf = +0.5

stock-market-conditions.indicator-3 = 1.2 - 2 (bullish)
update-stock-favourability indicator-3-cf = +0.6

stock-market-conditions.indicator-3 = >2 (very-bearish)
update-stock-favourability indicator-3-cf = +0.8

/* 4. Indicator 4 - No of names above 200D MA/Number of names below 200D
MA */

stock-market-conditions.indicator-4 = 0 - 0.3 (very-bearish)
update-stock-favourability indicator-4-cf = -0.8

stock-market-conditions.indicator-4 = 0.3 - 0.8 (bearish)
update-stock-favourability indicator-4-cf = -0.6

stock-market-conditions.indicator-4 = 0.8 - 1.2 (neutral)
update-stock-favourability indicator-4-cf = +0.5

stock-market-conditions.indicator-4 = 1.2 - 2 (bullish)
update-stock-favourability indicator-4-cf = +0.6

stock-market-conditions.indicator-4 = >2 (very-bearish)
update-stock-favourability indicator-4-cf = +0.8

/* 5. Indicator 5 - VIX (measure of implied volatility of S&P500 out of the money options) */

stock-market-conditions.indicator-5 = <10

```

```

OR (stock-market-conditions.indicator-5 = >=25
AND stock-market-conditions.indicator-5 = <35)
update-stock-favourability indicator-5-cf = -0.6

stock-market-conditions.indicator-5 = >=10
AND stock-market-conditions.indicator-5 = <18
update-stock-favourability indicator-5-cf = +0.8

stock-market-conditions.indicator-5 = >=18
AND stock-market-conditions.indicator-5 = <25
update-stock-favourability indicator-5-cf = +0.5

stock-market-conditions.indicator-5 = >=35
update-stock-favourability indicator-5-cf = -0.9

/* Working Goal - stock-market-conditions indicators */
indicator-1-cf
AND
indicator-2-cf
AND
indicator-3-cf
AND
indicator-4-cf
AND
indicator-5-cf

CREATE WORKING GOAL =
macroeconomic-and-stock-market-conditions.stock-market-conditions-cf

/* Working Goal - market favourability-cf */
macroeconomic-and-stock-market-conditions.macroeconomic-conditions-cf
AND
macroeconomic-and-stock-market-conditions.stock-market-conditions-cf

CREATE WORKING GOAL =
macroeconomic-and-stock-market-conditions.market-favourability-cf

/* Determine market favourability */
macroeconomic-and-stock-market-conditions.market-favourability-cf >=
0.8

```

AND
macroeconomic-and-stock-market-conditions.market-favourability-cf < 1
update-stock-favourability market-favourability = high

macroeconomic-and-stock-market-conditions.market-favourability-cf >= 0.5
AND
macroeconomic-and-stock-market-conditions.market-favourability-cf < 0.8
update-stock-favourability market-favourability = medium

macroeconomic-and-stock-market-conditions.market-favourability-cf >= 0
AND
macroeconomic-and-stock-market-conditions.market-favourability-cf < 0.5
update-stock-favourability market-favourability = low

/* Determine stock features attractiveness as a result of market
favourability */

/* Determine stock cap category (Feature 1) attractiveness */

/* High market favourability */
market-favourability = high
AND
(stock-list.cap-category = large

update-stock-favourability

stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness-
cf = 0.6

OR

stock-list.cap-category = mid

update-stock-favourability

stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness-
-cf = 0.6

OR

```
stock-list.cap-category = small
```

```
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness
-cf = 0.8
```

)

```
/* Medium market favourability */
```

```
market-favourability = medium
```

AND

```
(stock-list.cap-category = large
```

```
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness
-cf = 0.7
```

OR

```
stock-list.cap-category = mid
```

```
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness
-cf = 0.6
```

OR

```
stock-list.cap-category = small
```

```
update-stock-favourability

stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness
-cf = 0.6

)

/* Low market favourability */
market-favourability = low
AND
(stock-list.cap-category= large

update-stock-favourability

stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness
-cf = 0.8

OR

stock-list.cap-category = mid

update-stock-favourability

stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness
-cf = 0.4

OR

stock-list.cap-category = small

update-stock-favourability

stock-attractiveness-feature-inherent-derived.cap-category-
attractiveness
-cf = 0.2
```

```
)  
  
/* Determine sector category (Feature 2) attractiveness */  
  
/* High market favourability */  
market-favourability = high  
AND  
(stock-list.sector = 'consumer-staples'  
OR stock-list.sector = 'health-care'  
OR stock-list.sector = 'utility')  
  
update-stock-favourability  
  
stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf  
= 0.4  
  
market-favourability = high  
  
AND  
(stock-list.sector = 'materials'  
OR stock-list.sector = 'industrials'  
OR stock-list.sector = 'Telecommunications')  
  
update-stock-favourability  
  
stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf  
= 0.6  
  
market-favourability = high  
AND  
(stock-list.sector = 'technology'  
OR stock-list.sector = 'consumer-discretionary'  
OR stock-list.sector = 'financials'  
OR stock-list.sector = 'energy')  
  
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf  
= 0.8
```

```
/* Medium market favourability */
```

```
market-favourability = medium
```

```
AND
```

```
(stock-list.sector = 'consumer-staples'  
OR stock-list.sector = 'health-care'  
OR stock-list.sector = 'utility')
```

```
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf  
= 0.5
```

```
market-favourability = medium
```

```
AND
```

```
(stock-list.sector = 'materials'  
OR stock-list.sector = 'industrials'  
OR stock-list.sector = 'Telecommunications')
```

```
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf  
= 0.5
```

```
market-favourability = medium
```

```
AND
```

```
(stock-list.sector = 'technology'  
OR stock-list.sector = 'consumer-discretionary'  
OR stock-list.sector = 'financials'  
OR stock-list.sector = 'energy')
```

```
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf  
= 0.7
```

```
/* Low market favourability */  
market-favourability = low
```

AND

```
(stock-list.sector = 'consumer-staples'  
OR stock-list.sector = 'health-care'  
OR stock-list.sector = 'utility')
```

```
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf  
= 0.9
```

```
market-favourability = low
```

AND

```
(stock-list.sector = 'materials'  
OR stock-list.sector = 'industrials'  
OR stock-list.sector = 'Telecommunications')
```

```
update-stock-favourability
```

```
stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf  
= 0.4
```

```
market-favourability = low
```

AND

```
(stock-list.sector = 'technology'  
OR stock-list.sector = 'consumer-discretionary'  
OR stock-list.sector = 'financials'  
OR stock-list.sector = 'energy')
```

```
update-stock-favourability
```

```

stock-attractiveness-feature-inherent-derived.sector-attractiveness-cf
= 0.2

/* Stock attractiveness inherent market features */

/* Valuation (PEG) */
stock-attractiveness-feature-inherent-market-feature.peg >= 0
And stock-attractiveness-feature-inherent-market-feature.peg < 1
update-stock-favourability peg-cf = 0.9

stock-attractiveness-feature-inherent-market-feature.peg >= 1
AND stock-attractiveness-feature-inherent-market-feature.peg < 1.7
update-stock-favourability peg-cf = 0.6

stock-attractiveness-feature-inherent-market-feature.peg >= 1.7
AND stock-attractiveness-feature-inherent-market-feature.peg < 2.4
update-stock-favourability peg-cf = 0.3

stock-attractiveness-feature-inherent-market-feature.peg >= 2.4
AND stock-attractiveness-feature-inherent-market-feature.peg < 3
update-stock-favourability peg-cf = -0.2

stock-attractiveness-feature-inherent-market-feature.peg < 0
OR stock-attractiveness-feature-inherent-market-feature.peg >= 3
update-stock-favourability peg-cf = -0.9

/* EBITDA-Margin (%) */
stock-attractiveness-feature-inherent-market-feature.ebitda-margin >=
25
update-stock-favourability ebitda-margin-cf = 0.9

stock-attractiveness-feature-inherent-market-feature.ebitda-margin >=
18
AND stock-attractiveness-feature-inherent-market-feature.ebitda-margin
< 25
update-stock-favourability ebitda-margin-cf = 0.7

stock-attractiveness-feature-inherent-market-feature.ebitda-margin >=
10

```

```

AND stock-attractiveness-feature-inherent-market-feature.ebitda-margin
< 18
update-stock-favourability ebitda-margin-cf = 0.4

stock-attractiveness-feature-inherent-market-feature.ebitda-margin >= 0
AND stock-attractiveness-feature-inherent-market-feature.ebitda-margin
< 10
update-stock-favourability ebitda-margin-cf = -0.2

stock-attractiveness-feature-inherent-market-feature.ebitda-margin < 0
update-stock-favourability ebitda-margin-cf = -0.7

/* ROIC (Return on invested capital %) */
stock-attractiveness-feature-inherent-market-feature.roic > 15
update-stock-favourability roic-cf = 0.9

stock-attractiveness-feature-inherent-market-feature.roic >= 10
AND
stock-attractiveness-feature-inherent-market-feature.roic <= 15
update-stock-favourability roic-cf = 0.7

stock-attractiveness-feature-inherent-market-feature.roic >= 5
AND
stock-attractiveness-feature-inherent-market-feature.roic < 10
update-stock-favourability roic-cf = 0.3

stock-attractiveness-feature-inherent-market-feature.roic >= 0
AND
stock-attractiveness-feature-inherent-market-feature.roic < 5
update-stock-favourability roic-cf = -0.3

stock-attractiveness-feature-inherent-market-feature.roic < 0
update-stock-favourability roic-cf = 0.9

/* Net debt equity ratio */
stock-attractiveness-feature-inherent-market-feature.net-debt-equity-
ratio < -20
update-stock-favourability net-debt-equity-ratio-cf = 0.9

```

```

stock-attractiveness-feature-inherent-market-feature.net-debt-equity-
ratio >= -20
AND
stock-attractiveness-feature-inherent-market-feature.net-debt-equity-
ratio < 0
update-stock-favourability net-debt-equity-ratio-cf = 0.7

stock-attractiveness-feature-inherent-market-feature.net-debt-equity-
ratio >= 0
AND
stock-attractiveness-feature-inherent-market-feature.net-debt-equity-
ratio < 20
update-stock-favourability net-debt-equity-ratio-cf = 0.4

stock-attractiveness-feature-inherent-market-feature.net-debt-equity-
ratio >= 20
AND
stock-attractiveness-feature-inherent-market-feature.net-debt-equity-
ratio < 40
update-stock-favourability net-debt-equity-ratio-cf = -0.4

stock-attractiveness-feature-inherent-market-feature.net-debt-equity-
ratio >= 40
update-stock-favourability net-debt-equity-ratio-cf = -0.8

/* Compute stock-attractiveness-inherent-market-feature-cf */
peg-cf
ebitda-margin-cf
roic-cf
net-debt-equity-ratio-cf

AVERAGES TO

stock-attractiveness-inherent-market-feature.stock-attractiveness-
inherent-market-feature-cf

/* Determine stock attractiveness inherent derived due to market
favourability */

stock-list.risk-match-cf

```

```

AND
stock-attractiveness-inherent-derived.cap-category-attractiveness-cf
AND
stock-attractiveness-inherent-derived.sector-attractiveness-cf

AVERAGES TO

stock-attractiveness-inherent-derived.stock-attractiveness-from-
market-favourability-cf

/* Determine overall stock attractiveness */

stock-attractiveness-inherent-market-feature.stock-attractiveness-
inherent-market-feature-cf
AND
stock-attractiveness-inherent-derived.stock-attractiveness-from-
market-favourability-cf

AVERAGES TO stock-list.stock-attractiveness-cf

END KNOWLEDGE-BASE update-stock-favourability-model;

/* 3rd of 4 KNOWLEDGE-BASEs */
KNOWLEDGE-BASE update-stock-risk-model
USES:
update-stock-risk FROM update-stock-risk-kb;

/* market capitalization size */
stock-risk-feature.cap-size = large
update-stock-risk cap-size-cf = 0.9

stock-risk-feature.cap-size = mid
update-stock-risk cap-size-cf = 0.5

stock-risk-feature.cpa-size = small
update-stock-risk cap-size-cf = 0.3

/* liquidity */
stock-risk-feature.liquidity >= 500000
update-stock-risk liquidity-cf = 0.7

```

```
stock-risk-feature.liquidity <500000
update-stock-risk liquidity-cf = 0.3

/* sector */
stock-risk-feature.sector = consumer-staples
OR stock-risk-feature.sector = health-care
OR stock-risk-feature.sector = utility
update-stock-risk sector-cf = 0.8

stock-risk-feature.sector = materials
OR stock-risk-feature.sector = industrials
update-stock-risk sector-cf = 0.5

stock-risk-feature.sector = technology
OR stock-risk-feature.sector = consumer-discretionary
OR stock-risk-feature.sector = financials
update-stock-risk sector-cf = 0.3

/* price volatility */
stock-risk-feature.price-volatility <=18%
update-stock-risk price-volatility-cf = 0.9

stock-risk-feature.price-volatility >18%
AND stock-risk-feature.price-volatility <=24%
update-stock-risk price-volatility-cf = 0.7

stock-risk-feature.price-volatility >24%
AND stock-risk-feature.price-volatility <=40%
update-stock-risk price-volatility-cf = 0.5

stock-risk-feature.price-volatility >40%
AND stock-risk-feature.price-volatility <= 60%
update-stock-risk price-volatility-cf = -0.1

stock-risk-feature.price-volatility >60%
update-stock-risk price-volatility-cf = -0.6

/* earning volatility */
stock-risk-feature.earning-volatility <=14%
```

```
update-stock-risk earning-volatility-cf = 0.9

stock-risk-feature.earning-volatility >14%
AND
stock-risk-feature.earning-volatility <=24%
update-stock-risk earning-volatility-cf = 0.7

stock-risk-feature.earning-volatility >24%
AND
stock-risk-feature.earning-volatility <=40%
update-stock-risk earning-volatility-cf = 0.4

stock-risk-feature.earning-volatility >40%
AND
stock-risk-feature.earning-volatility <=60%
update-stock-risk earning-volatility-cf = -0.2

stock-risk-feature.earning-volatility >60%
update-stock-risk earning-volatility-cf = -0.7

/* Compute stock-risk-cf */
market-cap-cf
liquidity-cf
price-volatility-cf
earning-volatility-cf

AVERAGES TO stock-risk-feature.stock-risk-cf

/*Determine customer's overall risk tolerance* /

/* Risk ability */
/* Age*/
customer-risk-profile-ability.age = '20 - 35'
update-stock-risk age-cf = 0.8

customer-risk-profile-ability.age = '36 - 45'
update-stock-risk age-cf = 0.6

customer-risk-profile-ability.age = '46 - 55'
update-stock-risk age-cf = 0.4
```

```
customer-risk-profile-ability.age = '56 - 65'
update-stock-risk age-cf = -0.2

customer-risk-profile-ability.age = '>65'
update-stock-risk age-cf = -0.4

/* Net assets */
customer-risk-profile-ability.net-assets = '<$25,000'
update-stock-risk net-assets-cf = -0.2

customer-risk-profile-ability.net-assets = '$25,000 - $100,000'
update-stock-risk net-assets-cf = 0.2

customer-risk-profile-ability.net-assets = '$100,000 - $250,000'
update-stock-risk net-assets-cf = 0.4

customer-risk-profile-ability.net-assets = '$250,000 - $500,000'
update-stock-risk net-assets-cf = 0.6

customer-risk-profile-ability.net-assets = '>$500,000'
update-stock-risk net-assets-cf = 0.8

/* Job stability */
customer-risk-profile-ability.job-stability = 'very-unstable'
update-stock-risk job-stability-cf = -0.2

customer-risk-profile-ability.job-stability = 'rather-unstable'
update-stock-risk job-stability-cf = 0.2

customer-risk-profile-ability.job-stability = 'a-bit-unstable'
update-stock-risk job-stability-cf = 0.4

customer-risk-profile-ability.job-stability = 'stable'
update-stock-risk job-stability-cf = 0.6

customer-risk-profile-ability.job-stability = 'very-stable'
update-stock-risk job-stability-cf = 0.8
```

```

/* Monthly income */

customer-risk-profile-ability.monthly-income = '<$2,500'
update-stock-risk monthly-income-cf = -0.2

customer-risk-profile-ability.monthly-income = '$2,500 - $5,000'
update-stock-risk monthly-income-cf = 0.2

customer-risk-profile-ability.monthly-income = '$5,000 - $7,500'
update-stock-risk monthly-income-cf = 0.4

customer-risk-profile-ability.monthly-income = '$7,500 - $10,000'
update-stock-risk monthly-income-cf = 0.6

customer-risk-profile-ability.monthly-income = '>$10,000'
update-stock-risk monthly-income-cf = 0.8

/* Liquidity needs */

customer-risk-profile-ability.liquidity-needs = 'very-low'
update-stock-risk liquidity-needs-cf = -0.2

customer-risk-profile-ability.liquidity-needs = 'low'
update-stock-risk liquidity-needs-cf = 0.2

customer-risk-profile-ability.liquidity-needs = 'medium'
update-stock-risk liquidity-needs-cf= 0.4

customer-risk-profile-ability.liquidity-needs = 'high'
update-stock-risk liquidity-needs-cf = 0.6

customer-risk-profile-ability.liquidity-needs = 'very-high'
update-stock-risk liquidity-needs-cf = 0.8

/* Compute customer risk ability cf */

age-cf
net-assets-cf
job-stability-cf
monthly-income-cf
liquidity-needs-cf

```

AVERAGES TO customer-risk-profile.risk-tolerance-ability-cf

```
/* Risk willingness */

/* Thrill factor */
customer-risk-profile-willingness.thrill-factor = very-low
update-stock-risk thrill-factor-cf = -0.5

customer-risk-profile-willingness.thrill-factor = low
update-stock-risk thrill-factor-cf = 0.2

customer-risk-profile-willingness.thrill-factor = medium
update-stock-risk thrill-factor-cf = 0.4

customer-risk-profile-willingness.thrill-factor = high
update-stock-risk thrill-factor-cf = 0.6

customer-risk-profile-willingness.thrill-factor = very-high
update-stock-risk thrill-factor-cf = 0.8

/* Job security preference */
customer-risk-profile-willingness.job-security-preference = very-high
update-stock-risk job-security-preference.cf = -0.5

customer-risk-profile-willingness.job-security-preference = high
update-stock-risk job-security-preference.cf = 0.2

customer-risk-profile-willingness.job-security-preference = medium
update-stock-risk job-security-preference.cf = 0.4

customer-risk-profile-willingness.job-security-preference = low
update-stock-risk job-security-preference.cf = 0.6

customer-risk-profile-willingness.job-security-preference = very-low
update-stock-risk job-security-preference.cf = 0.8

/* Investment risk tolerance */
customer-risk-profile-willingness.investment-risk-tolerance = very-low
update-stock-risk investment-risk-tolerance-cf = -0.2
```

```

customer-risk-profile-willingness.investment-risk-tolerance = low
update-stock-risk investment-risk-tolerance-cf = 0.1

customer-risk-profile-willingness.investment-risk-tolerance = medium
update-stock-risk investment-risk-tolerance-cf = 0.2

customer-risk-profile-willingness.investment-risk-tolerance = high
update-stock-risk investment-risk-tolerance-cf = 0.4

customer-risk-profile-willingness.investment-risk-tolerance = very-high
update-stock-risk investment-risk-tolerance-cf = 0.6

customer-risk-profile-willingness.investment-risk-tolerance = Very,
very-high
update-stock-risk investment-risk-tolerance-cf = 0.8

/* Job compensation preference */
customer-risk-profile-willingness.job-compensation-preference = 'all-
salary'
update-stock-risk job-compensation-preference-cf = -0.2

customer-risk-profile-willingness.job-compensation-preference =
'mainly-salary'
update-stock-risk job-compensation-preference-cf = 0.1

customer-risk-profile-willingness.job-compensation-preference = 'equal-
mix-of-salary-and-commission'
update-stock-risk job-compensation-preference-cf = 0.2

customer-risk-profile-willingness.job-compensation-preference =
'mainly-commission'
update-stock-risk job-compensation-preference-cf = 0.5

customer-risk-profile-willingness.job-compensation-preference = 'all-
commission'
update-stock-risk job-compensation-preference-cf = 0.7

/* Aversion to previous stock loss */

```

```

customer-risk-profile-willingness.aversion-to-previous-stock-loss =
very-low
update-stock-risk aversion-to-previous-stock-loss-cf = -0.2

customer-risk-profile-willingness.aversion-to-previous-stock-loss = low
update-stock-risk aversion-to-previous-stock-loss-cf = 0.1

customer-risk-profile-willingness.aversion-to-previous-stock-loss =
medium
update-stock-risk aversion-to-previous-stock-loss-cf = 0.2

customer-risk-profile-willingness.aversion-to-previous-stock-loss =
high
update-stock-risk aversion-to-previous-stock-loss-cf = 0.5

customer-risk-profile-willingness.aversion-to-previous-stock-loss =
very-high
update-stock-risk aversion-to-previous-stock-loss-cf = 0.7

/* Compute customer risk willingness cf */
thrill-factor-cf
AND
job-security-preference-cf
AND
investment-risk-tolerance-cf
AND
job compensation-preference-cf
AND
aversion-to-previous-stock-loss-cf

AVERAGES TO customer-risk-profile.risk-tolerance-willingness-cf

/* Compute customer overall risk tolerance cf */
customer-risk-profile.risk-tolerance-ability-cf
AND
customer-risk-profile.risk-tolerance-willingness-cf

CREATE WORKING GOAL = customer-risk-profile.overall-risk-tolerance-cf
/* Determine risk match */
customer-risk-profile.overall-risk-tolerance-cf

```

```

AND
stock-risk-feature.stock-risk-cf

update-stock-risk stock-list.risk-match-cf

END KNOWLEDGE-BASE update-stock-risk-model

/* 4th of 4 KNOWLEDGE-BASEs */

KNOWLEDGE-BASE update-recommendation-rank-model
USES:
    update-recommendation-rank FROM update-recommendation-rank-kb;

/* Determine recommendation rank based on risk match and stock
attractiveness */

stock-list.risk-match-cf
AND
stock-list.stock-attractiveness-cf

recommend-stock stock-list.recommendation-rank

END KNOWLEDGE-BASE update-recommendation-rank-model

/*****************/
END DOMAIN-KNOWLEDGE us-stock-domain
/*****************/

/*****************/
INFERENCE-KNOWLEDGE us-stock-recommendation;
/*****************/

/* 1st of 4 INFERENCEs: filter-stock-style */
KNOWLEDGE-ROLE customer-style-profile;
TYPE: DYNAMIC;
DOMAIN-MAPPING: customer-style-profile;
END KNOWLEDGE-ROLE customer-style-profile;

KNOWLEDGE-ROLE stock-style-feature;
TYPE: DYNAMIC;

```

```

DOMAIN-MAPPING: stock-style-feature SET OF stock-list;
END KNOWLEDGE-ROLE stock-style-feature;

KNOWLEDGE-ROLE filter-stock-style-model;
  TYPE: STATIC;
  DOMAIN-MAPPING: filter-stock-style FROM filter-stock-style-kb;
END KNOWLEDGE-ROLE filter-stock-style-model;

INFERENCE filter-stock-style;
  ROLES:
    INPUT:      customer-style-profile;
                stock-style-feature;
    OUTPUT:     stock-style-feature;
    STATIC:     filter-stock-style;
  SPECIFICATION: "Input1 is a set of customer style profile (customer level), Input2 is a set of stock style features (stock level). Output is the same set of stock list rated with style features (stock level).";
END INFERENCE filter-stock-style;

/* 2nd of 4 INFERENCES: update-stock-risk */
KNOWLEDGE-ROLE customer-risk-profile;
  TYPE: DYNAMIC;
  DOMAIN-MAPPING: customer-risk-profile;
END KNOWLEDGE-ROLE customer-risk-profile;

KNOWLEDGE-ROLE stock-risk-feature;
  TYPE: DYNAMIC;
  DOMAIN-MAPPING: stock-risk-feature SET OF stock-list;
END KNOWLEDGE-ROLE stock-risk-feature;

KNOWLEDGE-ROLE update-stock-risk-model;
  TYPE: STATIC;
  DOMAIN-MAPPING: update-stock-risk FROM update-stock-risk-kb;
END KNOWLEDGE-ROLE update-stock-risk-model;

INFERENCE update-stock-risk;
  ROLES:
    INPUT:      customer-risk-profile;
                stock-risk-feature;

```

```

        OUTPUT:      stock-risk-feature;
        STATIC:      update-stock-risk;
SPECIFICATION: "Input1 is a set of customer risk profile (customer
level), Input2 is a set of stock risk features (stock level). Output
is the same set of stock list rated with risk features (stock
level).";
END INference update-stock-risk;

/* 3rd of 4 INFERENCES: update-stock-favourability */
KNOWLEDGE-ROLE macroeconomic-and-stock-market-conditions;
    TYPE: DYNAMIC;
    DOMAIN-MAPPING: macroeconomic-and-stock-market-conditions;
END KNOWLEDGE-ROLE macroeconomic-and-stock-market-conditions;

KNOWLEDGE-ROLE stock-attractiveness-feature-inherent-Market-Feature;
    TYPE: DYNAMIC;
    DOMAIN-MAPPING: stock-attractiveness-feature-inherent-Market-Feature
SET OF stock-list;
END KNOWLEDGE-ROLE stock-attractiveness-feature-inherent-Market-Feature;

KNOWLEDGE-ROLE stock-attractiveness-feature-inherent-derived;
    TYPE: DYNAMIC;
    DOMAIN-MAPPING: stock-attractiveness-feature-inherent-derived SET OF
stock-list;
END KNOWLEDGE-ROLE stock-attractiveness-feature-inherent-derived;

KNOWLEDGE-ROLE update-stock-favourability-model;
    TYPE: STATIC;
    DOMAIN-MAPPING: update-stock-favourability FROM
update-stock-favourability-kb;
END KNOWLEDGE-ROLE update-stock-favourability-model;

INference update-stock-favourability;
    ROLES:
        INPUT:      macroeconomic-and-stock-market-conditions;
                    stock-attractiveness-feature-inherent-Market-Feature;
        OUTPUT:     stock-attractiveness-feature-inherent-derived;
        STATIC:     update-stock-favourability;
SPECIFICATION: "Input1 is a set of macroeconomic-and-stock-market-
conditions (world level & stock market/sector level), Input2 is a set

```

```

of stock Favourability (attractiveness) features (stock level).
Output is the same set of stock list rated with Favourability
features (stock level).";
END INference update-stock-favourability;

/* 4th of 4 INFERENCES: update-recommendation-rank */
KNOWLEDGE-ROLE stock-risk-feature;
  TYPE: DYNAMIC;
  DOMAIN-MAPPING: stock-risk-feature SET
  OF stock-list;
END KNOWLEDGE-ROLE stock-risk-feature;

KNOWLEDGE-ROLE stock-attractiveness-feature-inherent-derived;
  TYPE: DYNAMIC;
  DOMAIN-MAPPING: stock-attractiveness-feature-inherent-derived
  SET OF stock-list;
END KNOWLEDGE-ROLE stock-attractiveness-feature-inherent-derived;

KNOWLEDGE-ROLE update-recommendation-rank-model;
  TYPE: STATIC;
  DOMAIN-MAPPING: update-recommendation-rank
  FROM update-recommendation-rank-kb;
END KNOWLEDGE-ROLE update-recommendation-rank-model;

INference update-recommendation-rank;
  ROLES:
    INPUT: stock-risk-feature;
    stock-attractiveness-feature-inherent-derived;
    OUTPUT: stock-risk;
    STATIC: update-recommendation-rank;
  SPECIFICATION:
    "Input1 is a set of stock risk feature (stock level), Input2 is a
     set of stock favourability
     features (stock level). Output is the same set of stock list rated
     with final/overall
     Recommendation Rank (stock level).";
END INference update-recommendation-rank;

*****/*
END INference-KNOWLEDGE us-stock-recommendation;

```

```

/***** */

/* **** */

TASK-KNOWLEDGE us-stock-recommendation-task;
/* **** */

/* Main Task & Task Method: us-stock-assessment-process-task */
/* Main Task & Task Method have 4 respective SUB-Task & Task Method*/
TASK us-stock-assessment-process-task;
  DOMAIN-NAME: us-stock-domain;
  GOAL: "Update/Rank all the stock in stock list base on Customer Risk
         profile, Customer Style profile, and parameters of
         Macroeconomic-and-Stock-market-conditions, ";
  ROLES:
    INPUT: stock-list;
           customer-risk-profile SET OF customer-profile;
           customer-style-profile SET OF customer-profile;
           macroeconomic-and-stock-market-conditions;
    OUTPUT: stock-list.recommendation-rank: "with Recommendation Rank";
  END TASK us-stock-assessment-process;

TASK-METHOD us-stock-assessment-process-task-method;
  REALIZES: us-stock-assessment-process-rask;
  DECOMPOSITION:
    TASKS: filter-stock-style, update-stock-risk, update-stock-
           favourability, update-recommendation-rank;
  ROLES:
    INTERMEDIATE:
      stock-list.recommendation-rank: "with Recommendation Rank";
  CONTROL-STRUCTURE:
    filter-stock-style(stock-list.stock-style-feature + customer-
                      profile.customer-style-profile -> stock-list.style-selection);
    update-stock-risk(stock-list.stock-risk-feature + customer-
                      profile.customer-risk-profile -> stock-list.risk-match-cf);
    update-stock-favourability(stock-list.stock-attractiveness-feature-
                               inherent-market-feature + macroeconomic and stock-market-conditions -
                               > stock-list.stock-attractiveness-cf);
    update-recommendation-rank(stock-list.risk-match-cf + stock-
                               list.stock-attractiveness-cf -> stock-list.recommendation-rank);

```

```

END TASK-METHOD us-stock-assessment-process-task-method;

/* 1st of 4 SUB Task/Method of Main Task & Task Method: filter-stock-style
*/
TASK filter-stock-style;
DOMAIN-NAME: us-stock-domain;
GOAL: "Compare between the stock-style-feature and customer-style-features,
       identify and retain stock in stock-list if the stock-style-feature
       matches customer-style-features.";
ROLES:
INPUT: stock-style-feature SET OF stock-list;
       customer-style-profile SET OF customer-profile;
OUTPUT: stock-list.style-selection: "dividend, growth, momentum, value";
END TASK filter-stock-style;

TASK-METHOD filter-stock-style-task-method;
REALIZES: filter-stock-style;
DECOMPOSITION:
INFERENCES: filter-stock-style;
ROLES:
INTERMEDIATE:
stock-list.style-selection: "dividend , growth, momentum,
value";
CONTROL-STRUCTURE:
WHILE NOT END OF stock-list DO
    IF      stock-list.style-selection =
           customer-style-profile.style-selection;
           RETAIN the stock in stock-list
    ELSE
           REMOVE the stock from stock-list
END WHILE
END TASK-METHOD filter-stock-style-task-method;

/* 2nd of 4 SUB Task/Method of Main Task & Task Method: update-stock-risk
*/
TASK update-stock-risk;
DOMAIN-NAME: us-stock-domain;

```

GOAL: "Compare between the stock-risk-feature and customer-risk-features, identify and update stock-risk-feature when it's matched with customer-risk-features.";

ROLES:

INPUT: stock-risk-feature SET OF stock-list;
customer-risk-profile SET OF customer-profile;
OUTPUT: stock-list.risk-match-cf: {-1 - 1};
END TASK update-stock-risk;

TASK-METHOD update-stock-risk-task-method;

REALIZES: [update-stock-risk](#);

DECOMPOSITION:

INFERENCES: update-stock-risk;

ROLES:

INTERMEDIATE:

stock-list.risk-match-cf: {-1 - 1};

CONTROL-STRUCTURE:

```
WHILE NOT END OF stock-list DO
    stock-list.risk-match-cf := normalized-gap between customer-risk-
    profile and stock-list.stock-risk-feature;
END WHILE
```

END TASK-METHOD update-stock-risk-task-method;

/ 3rd of 4 SUB Task/Method of Main Task & Task Method: update-stock-favourability */*

TASK update-stock-favourability;

DOMAIN-NAME: [us-stock-domain](#);

GOAL: "Compare between stock-list.stock-attractiveness-feature-inherent-Market-Feature and stock attractiveness due to macroeconomic-and-stock-market-conditions, to update stock-attractiveness certainty factor.";

ROLES:

INPUT:

stock-attractiveness-feature-inherent-market-feature SET OF stock-list
macroeconomic-and-stock-market-conditions

OUTPUT: stock-list.stock-attractiveness-cf: {-1 - 1};

END TASK Update-Stock-Favourability;

TASK-METHOD update-stock-favourability-task-method;

REALIZES: [update-stock-favourability](#);

DECOMPOSITION:

```

INFERENCES: update-stock-favourability;

ROLES:
    INTERMEDIATE:
        stock-list.stock-attractiveness-cf: {-1 - 1};

CONTROL-STRUCTURE:
    WHILE NOT END OF stock-list DO
        stock-list.stock-attractiveness-certainty-factor:= normalized-gap
        between macroeconomic-and-stock-market-conditions and stock-
        list.stock-attractiveness-feature-inherent-market-feature;
    END WHILE

END TASK-METHOD update-stock-favourability-task-method;

/* 4th of 4 SUB Task/Method of Main Task & Task Method: update-stock-
recommendation-rank */

TASK update-stock-recommendation-rank;
DOMAIN-NAME: us-stock-domain;
GOAL: " compute the final stock-list.recommendation-rank based on CF
merging between: stock-list.risk-match-certainty-factor and stock-
list.stock-attractiveness-certainty-factor. ";

ROLES:
INPUT:      stock-list.risk-match-cf;
            stock-list.stock-attractiveness-cf;
OUTPUT:     stock-list.recommendation-rank: {-1 - 1};
END TASK update-stock-recommendation-rank;

TASK-METHOD update-stock-recommendation-rank-task-method;
REALIZES: update-stock-recommendation-rank;
DECOMPOSITION:
    INFERENCES: update-stock-recommendation-rank;
ROLES:
    INTERMEDIATE:
        stock-list.recommendation-rank: {-1 - 1};

CONTROL-STRUCTURE:
    WHILE NOT END OF stock-list DO
        stock-list.recommendation-rank-match := CF-merger between stock-
        list.risk-match-cf and stock-list.stock-attractiveness-cf;
    END WHILE

END TASK-METHOD update-stock-recommendation-rank-task-method;

```

```
*****  
END TASK-KNOWLEDGE us-stock-recommendation-task;  
*****  
*****  
END KNOWLEDGE-MODEL us-stock;  
*****  
  
***** END *****
```

10 Appendix B: Knowledge Acquisition Transcript

Following is the summary of the transcript of the knowledge elicitation and acquisition process.

1. We will like to find out more about investing in stocks. What are the advantages and disadvantages of investing the stock market?

The stock market is a great way of putting your capital to work for you. In the long run, the right stocks will create returns that should more than sufficiently beat inflation and help to build a portfolio of wealth that should put one in a strong financial situation and possibly help with retirement needs. It should preferably at least be considered in a part of anyone's portfolio given its strong return potential compared to other instruments such as savings or even bonds. That said, the stock market is a treacherous place and is not for the fainted hearted with its constant swing and gyrations and even occasional market crash. If one chooses the wrong stocks or invest at the wrong time, he can even cause his wealth to erode rapidly. As such, it is important that one takes the trouble to educate himself about how best to go about investing in stocks. The effort in doing so can potentially pay dividends in the long term.

2. We have heard that the U.S market is a great place to start an investing journey. Is that true?

What advantages does the U.S market offer?

Indeed, the U.S market is one of the most attractive, if not the most attractive market in the world to invest in. There are several advantages that I can mention here. They include the fact that the U.S market is the most highly capitalized market in the world, highly liquid and efficient and is home to many of the top companies in the world such as McDonalds, Google, Starbucks, Microsoft and Procter and Gamble to name a few. Any investor should be able to find his niche or favoured method of investing in this market given such a wide range of choices.

3. Is the stock market for everyone? Are there any personal circumstances that one should evaluate when considering an investment in the stock market?

Well, I guess it is not for everyone. As the stock market can involve a high level of risk, it is definitely not for the uninformed and risk adverse person. What I mean is if you do not understand what to look up for when selecting stocks and rely frequently on hearsay or advice from friends who are

equally as clueless as you are, then you might be better off not touching it. When considering any investment in the stock market, it is important to first identify traits that will have a bearing on the type of risk that you can undertake. I will say there are two main components to evaluate one's overall risk tolerance. There are his ability to take risk as well as his willingness to take risk. Ability to take risk will include factors such as his age, his personal wealth, future liquidity needs and their size, stability of his job and size of his paycheck and so on. This aspect of risk tolerance has more to do with a person's existing circumstances. On the other hand, willingness to take risk is a person's inherent characteristic and can be determined based on how risk adverse the person is when it comes to losing money, the amount of drawdown in his stock trading account he can accept before feeling very uncomfortable, how often he invest simply for the thrill of seeing his stock trading account goes up in value or how adverse he is towards making a new investment in a stock which he previously lost money in but which appears to have huge potential again.

4. Can stocks be classified as having different level of risks? What can affect the risks affecting a stock?

Definitely. Various attributes of a stock can affect its riskiness. Not all stocks are the same in terms of return and risk profile. Factors that can affect stock risk include its market capitalization (Generally, large cap stocks are safer than small cap stocks not least because they are more liquid, such to less manipulation because of their large size as well as being more well established and less prone to failures.), the sector it is in (Consumer staples, health care, utility stocks are generally safer compared to technology and financials or even consumer discretionary stocks whose performance can vary closely to the state of the economy). Earnings volatility over the years is an important factor as well as a high value on this measure means that the stock can be highly cyclical and very dependent on the economy, causes it to possibly have more wild swings in price. Likewise, price volatility can also affect stock risk as it may potentially mean a large fall in price.

5. What are the criteria that can be used to select a stock investment or how attractive it is?

This is where fundamental analysis comes in. It is the favoured method used by many investors, including investment greats like warren buffet and peter lynch. It considers various components such as the financial health of the company, how profitability the company is as well as how reasonably valued it is. In considering financial health, the amount of debt relative to equity can determine if a company is likely to experience financial hardship when times are hard. In considering profitability, profit margins and return on invested capital can indicate how good the company is in turning each dollar of sales into actual earnings and how much earnings the company is generating on the capital used. Finally, even with strong profitability and financial health, it is important to

understand if the stock of the company is reasonably valued and whether there is continued good growth going forward. After all, a good company does not necessarily mean a good stock and a good company might be unnecessarily bid up in price and becomes too expensive. A good valuation metric that combines both valuation and growth is the price-earnings growth ratio which measures the price-earnings per percentage of growth. Typically, a PEG under 1 is considered to be a favourable value.

6. Are there various styles of investing in the stock market? If so, what are they and what should an investor look out for?

Yes. There are different ways that an investor can play the stock market. There are quite a few dominant themes in the stock market. There are themes such as value investing, growth investing, dividend investing as well as momentum trading. Each has its merits and suits different kinds of investors. In value investing, the key is to identify stocks which are cheap so that there is a chance of buying them low and selling them high. That said, prices can continue going lower without necessarily turning higher and other fundamental criteria need to be evaluated as well. Metrics to determine valuation can be price- earnings ratio, price-sales ratio, price-cashflow ratio etc. For price earnings ratio, a value of 15 is considered reasonable and a value below that can be considered cheap. For price-sales ratio, 1.5 is about fair and below 1 is considered cheap. For growth investing, the key is to identify stocks that experience earnings growth year over year. These stocks can experience a good surge in price especially in earnings continue to stay strong in the coming years. However, growth stocks tend to have higher expectations embedded in them and as such can turn lower in price quickly if they do not perform as well as expected in earnings growth in future. In dividend investing, an investor identifies stocks that distributes a good amount of dividends every year and preferably increase its dividend year over year over the past years. This means the investor can enjoy much profit in terms of dividends over the years with a low cost basis even as the price does not appreciate much. Finally, in momentum trading, the investor's tries to ride on strong upwards price momentum for a relatively short period of time without necessarily considering the company's underlying fundamental characteristics and instead rely on strong price movement. The investor has to be nimble enough to get out of the market when the tide turns against him as momentum can fade as quickly as it goes up.

7. Are there external factors that can affect the price of a stock besides the inherent characteristics of each stock?

Yes. While the stock market averages tend to move up over a long period of time, there exist secular bull and bear markets along the way. One should vary his allocation to stocks according to how favourable the market is. To have a sense of whether the market is favourable, two broad factors

can be taken into consideration. First, macroeconomic factors should be considered. This will affect the corporate earnings of companies, how much companies are willing to invest and hence the price of their stocks. There are various important economic factors to consider, mainly the leading indicators. They include the inflation rate, the interest rate which can be determined by the yield curve and fed reserve actions, the unemployment rate and its recent month-over-month trend, weekly unemployment claims, trend of retail sales and manufacturing PMI. All these factors can foretell potential turning points in the stock market and alert the investor to the possibility of reducing exposure to equities. The second broad factor more directly measures stock market conditions and the underlying bullishness of the market. This has to do with aggregates of metrics relating to stock prices. They include measures such as 20 Day moving average of no of positive returns versus negative returns, 20 Day moving average of no of +4% moves versus -4% moves, 20 Day moving average of quarterly +25% versus -25% moves and ratio of no of names above 200 Day moving average versus no of names below 200 Day moving average. VIX which is the implied volatility of various S&P500 out of the money options otherwise known as the fear factor is also a good gauge of a potential turning point in the stock market.

8. Does market favourability impact different types of stock differently? Are they stocks which generally perform better during bull markets and versus during bear markets?

Yes. When the market is favourable, it is better to invest in cyclical stocks. Such stocks should be in sectors such as consumer discretionary (including housing stocks), technology and financials. On the other hand, when the market is not favourable, it is better to invest in defensive stocks in sectors such as health care, utilities and consumer staples. The stocks in these sectors tend to do better because they deal with products that are still consumed whether or not times are good. Also, small cap stocks will do better in a strong market environment compared to large cap stocks because investors are more optimistic and willing to take on more risk when they invest in stocks.

11 Appendix C: Project Schedule

	Name	Duration	Start	Finish	Predecessors	Nov 3 -
			S	S	M	
□	Phase 0 - Proposal	9d	11/01/2013	11/15/2013		
	Determine project scope	4d	11/01/2013	11/07/2013		
	Secure project sponsorship	1d	11/07/2013	11/09/2013		
	Project Proposal (frame up, review and sign off)	5d	11/07/2013	11/15/2013		
□	Phase 1 - Knowledge Modelling	40.5d	02/07/2014	04/11/2014		
	Requirements Gathering	22.5d	02/07/2014	03/14/2014		
	Phase 1 Report (frame up, review)	13d	03/15/2014	04/04/2014		
	Phase 1 Presentation (Preparation also)	4d	04/05/2014	04/11/2014		
□	Phase 2 - Software Development (Iterative)	188.5d	11/01/2013	08/21/2014		
	Data Collection and Assessment	50d	11/01/2013	01/18/2014		
	Business Understanding and Hypothesis Derivation	10d	11/01/2013	11/16/2013		
	Modeling	50d	04/05/2014	06/21/2014		
	Model Tuning	10d	07/30/2014	08/15/2014		
	Model Evaluation	25d	07/06/2014	08/15/2014	12	
	System Design Documentation	17d	04/11/2014	05/07/2014	8	
	Phase 2 Report (frame up, review)	13d	07/26/2014	08/15/2014		
	Phase 2 Presentation (Preparation also)	4d	08/15/2014	08/21/2014		
□	Phase 3 - Project report and Research paper preparation	95d	08/21/2014	01/16/2015		
	Model Refinement	10d	08/21/2014	09/06/2014		
	UI and Configuration	20d	09/07/2014	10/08/2014		
	User Acceptance Test / Training	20d	11/01/2014	11/30/2014		
	Reporting of Results	10d	10/24/2014	11/08/2014	12	
	Final Project Report (frame up, review)	30d	11/15/2014	12/31/2014		
	Technical Paper (frame up, review)	30d	11/15/2014	12/31/2014		
	Final Presentation	10d	12/31/2014	01/16/2015	23,24	

12 References

- <http://piquantcapital.com/home.html>
- <http://www.kdnuggets.com/datasets/index.html>
- http://en.wikipedia.org/wiki/Data_mining
- <http://scholarbank.nus.edu.sg/handle/10635/35562>
- <http://www.nasdaq.com/investing/glossary/>
- <http://teamgantt.com/>
- <http://www.r-project.org/>
- <http://rdotnet.codeplex.com/>
- <http://www.microsoft.com/en-us/server-cloud/products/sql-server/>
- <http://finance.yahoo.com/stock-center/>
- <http://www.stocknod.com/stock-market-for-beginners/>
- And NUS Library Resources