Using Artificial Intelligence to Develop a Robo Advisor

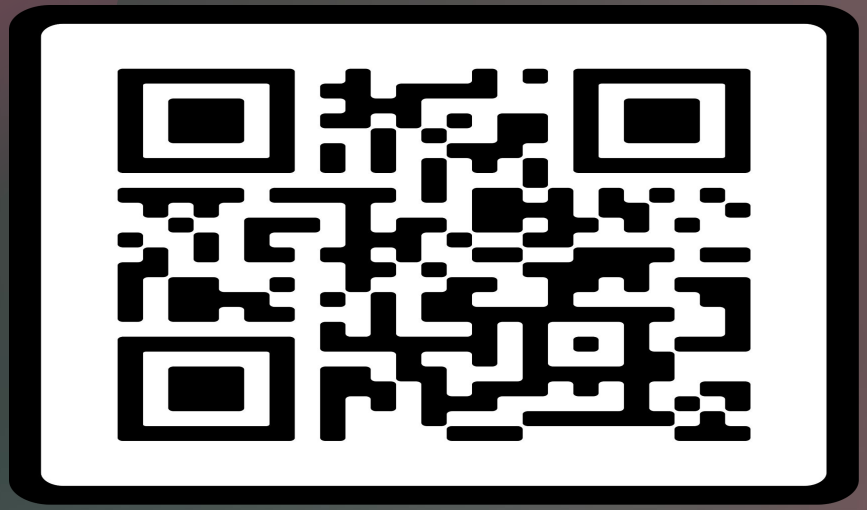
By Wes Lee

Contents

Problem Statement

- ***Predict risk tolerance*** of an individual from various demographic, financial and behavioural attributes ***using supervised ML***
- ***Use reinforcement learning*** (value-based deep Q-network (DQN)) to ***dynamically change the portfolio allocation weights*** and compare cumulative return against the ***mean variance optimized (MVO) portfolio***
- ***Build a robo-advisor dashboard*** using Plotly Dash and implement the ***risk tolerance prediction model in the dashboard***
- Using the ***predicted value of risk tolerance*** and choosing from any selection of equities from the S&P 500, ***calculate the optimal portfolio weights*** using ***mean variance optimization (MVO)*** with the aid of the CVXOPT library (uses convex optimization)

Demonstration



SCAN ME

Datasets

- ***2019 Survey of Consumer Finances (SCF)***. Conducted triennially by the United States Federal Reserve
  - Out of 351 variables, ***28 were selected***
  - 15 represent various measures of ***asset values***
  - 13 represent various ***demographic, financial and behavioural*** attributes
- ***Adjusted closing prices for all S&P 500 stocks*** from 1st week of 2000 to 1st Week of September 2023 downloaded using yfinance library
  - The ***full list of S&P 500 stocks*** was scraped from Wikipedia

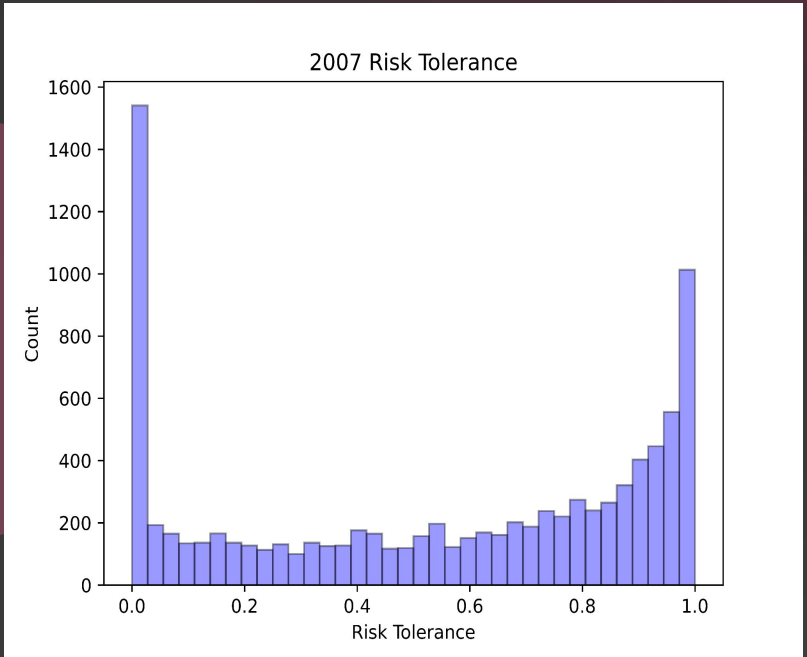
Methodology

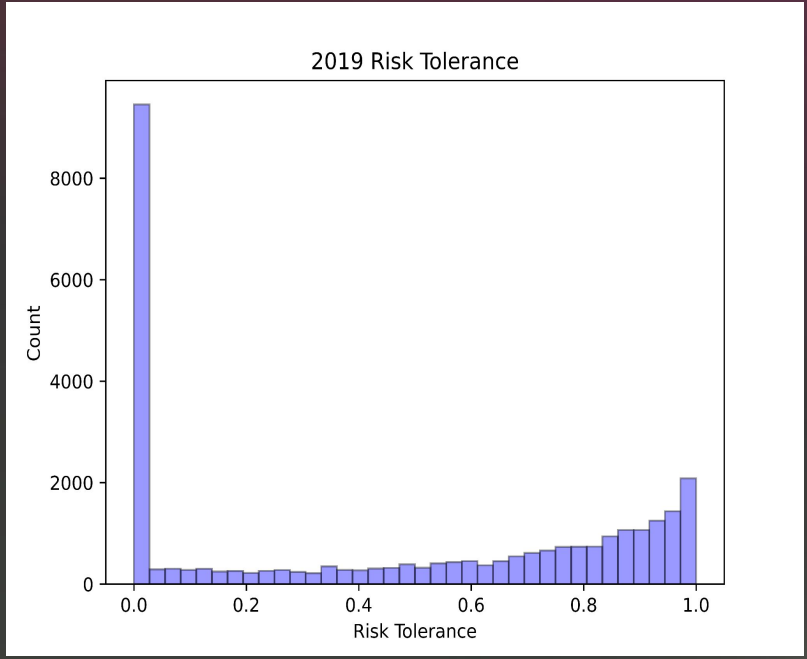
1. The 15 variables measuring asset values were used to calculate an ***a posteriori*** measure of ***risk tolerance***
    a. Risk tolerance calculated as the ***ratio of risky to (risky + risk-free) assets***
    b. Risk tolerance is the ***target value to predict***
2. The 13 various demographic, financial and behavioural attributes used as ***predictor variables to predict risk tolerance***
3. For DQN, a ***random sample of adjusting closing prices of 10 equities*** from start of 2020 to present day used to ***train an RL Agent class*** (contains variable and member functions that perform Q-learning) ***in a StockEnvironment class*** (simulation environment for training the agent)

Methodology (DQN)

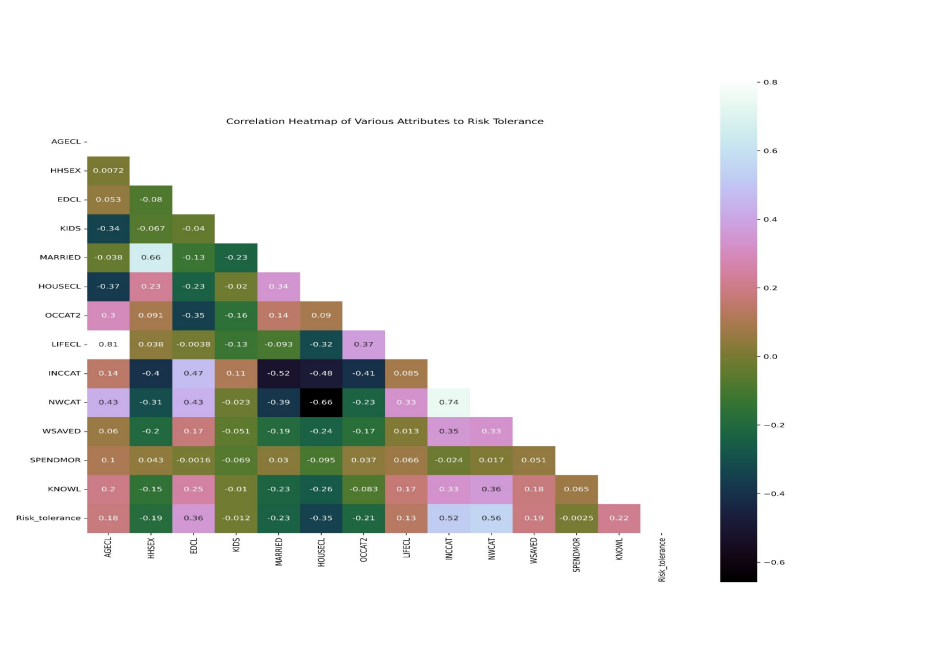
1. Deep Q-network is a ***value-based method that combines deep learning*** (using an ANN) with Q-learning, which sets the learning objective to ***optimize the estimates of Q-value***
2. The deep Q-learning algorithm ***approximates the Q-values by learning a set of weights, θ***, of a multilayered deep Q-network that maps states to actions and ***learns to find the right weights*** by iteratively adjusting them to ***maximize rewards (sharpe ratio)***
3. Maximizing this reward will lead to ***portfolio allocation weights*** that ***maximize the return of the portfolio for a given level of risk***
4. The ANN ***applies gradient descent to minimize a loss function*** (essentially MSE) which is the squared difference between the ***NN's estimate of the target and its estimate of the Q-value of the current state-action pair, $Q(s,a:\theta)$***
5. Since the Q-value is the expected reward for the state-action pair following a policy π, when the algorithm ***iteratively converges to the optimal Q-value***, it learns an ***optimal policy*** which is how to act to ***maximize the return/reward in every state***, which is the ***optimal*** sharpe ratio of the portfolio and also the ***optimal*** portfolio allocation weights

EDA - Changes in Risk Tolerance from 2007 to 2019





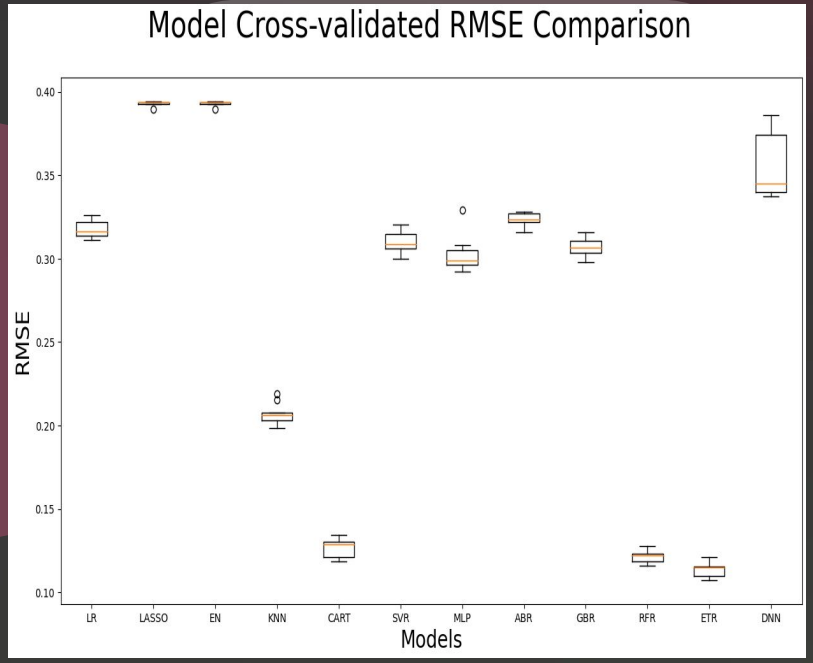
EDA – Correlation of Attributes to Risk Tolerance



Modeling - Models Used To Predict Risk Tolerance

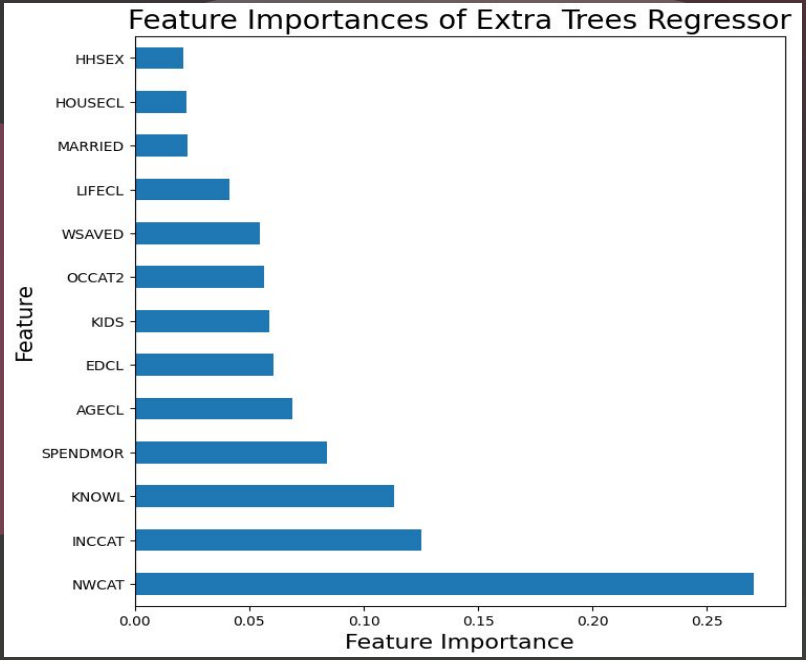
1. Linear Regression (regularized and non-regularized)
2. k-Nearest Neighbours Regression
3. Support Vector Regression
4. Decision Tree Regression
5. Random Forest Regression
6. Extra Trees Regression
7. Gradient Boosted Regression (including AdaBoost)
8. Multi-layer Perceptron Regression
9. Sequential Deep Neural Network

Modeling - Model Cross-validated RMSE Comparison



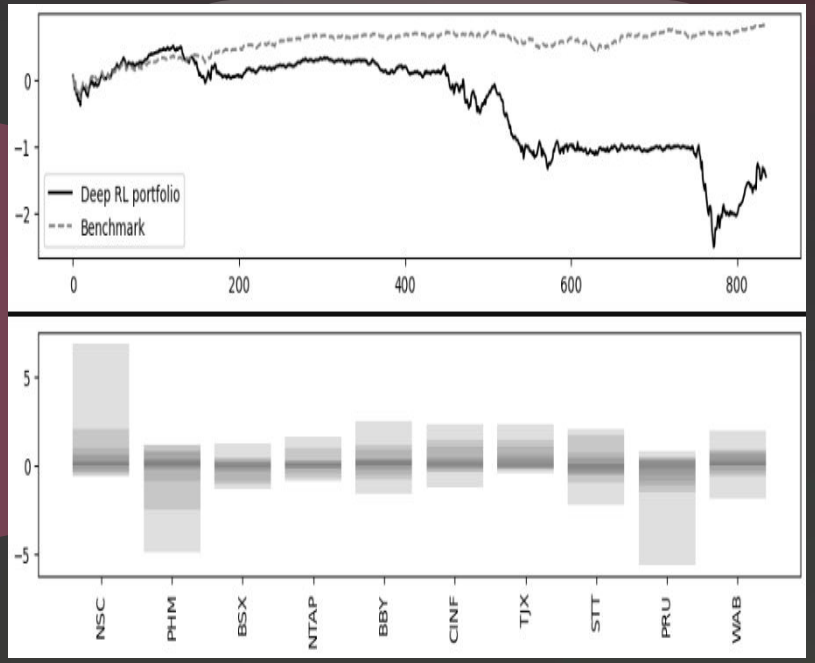
Model	Cross-validated RMSE
Decision Tree Regression	0.12771
Random Forest Regression	0.12156
Extra Trees Regression	0.11418

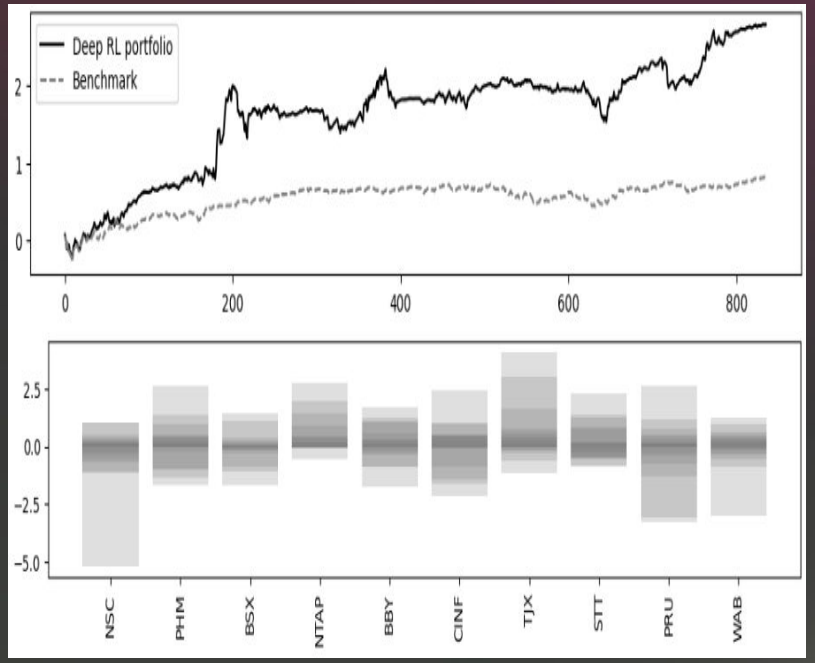
Modeling - Feature Importances + RMSE & R2 Scores



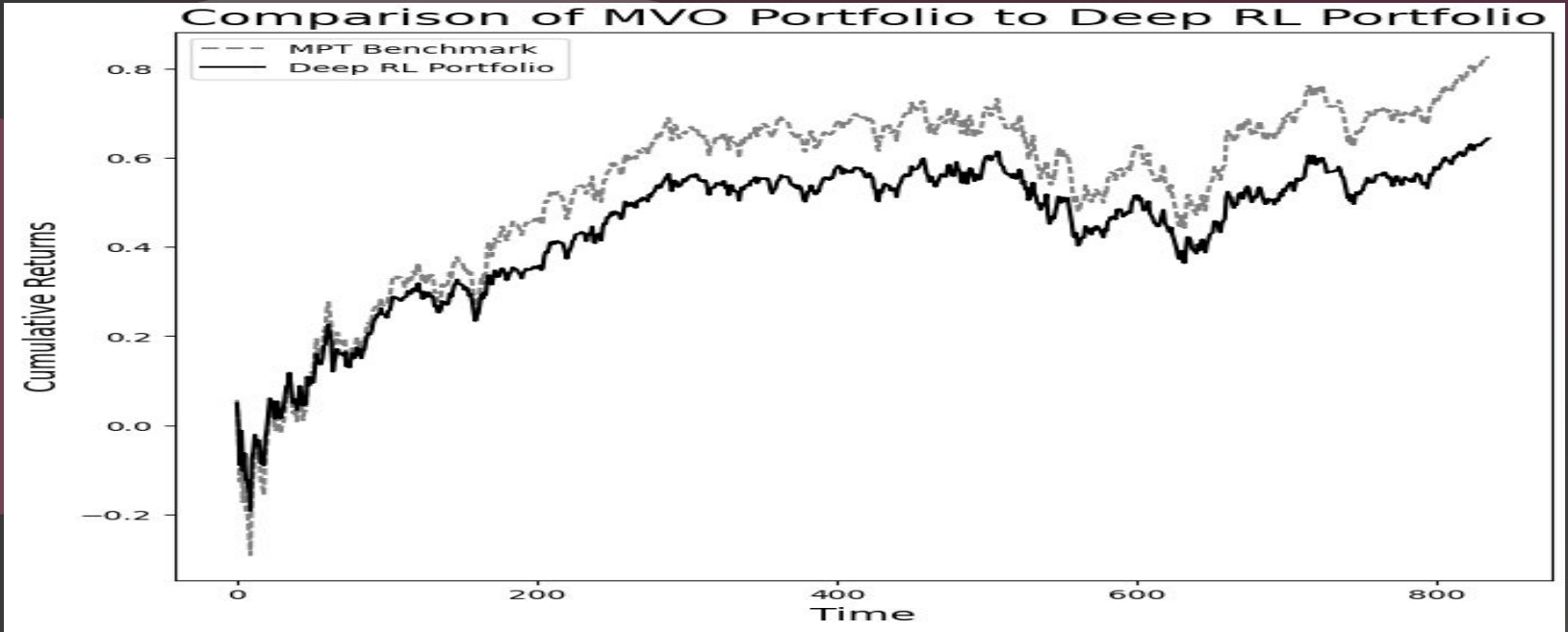
Train RMSE	0.08337
Cross-validated RMSE	0.11407
Test RMSE	0.11519
Train R2	0.95501
Test R2	0.91427

Modeling (DQN) - Comparison of First Vs Last Training Episode

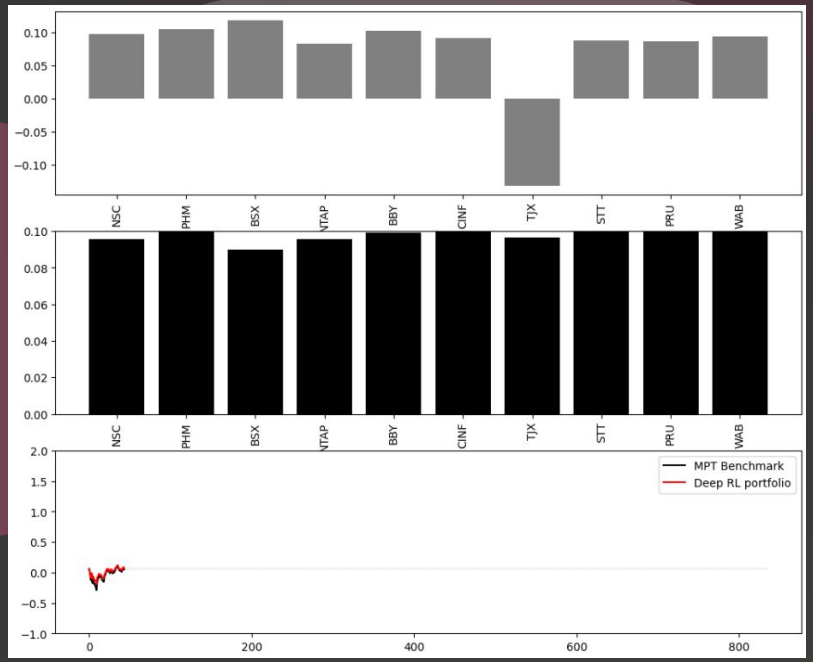


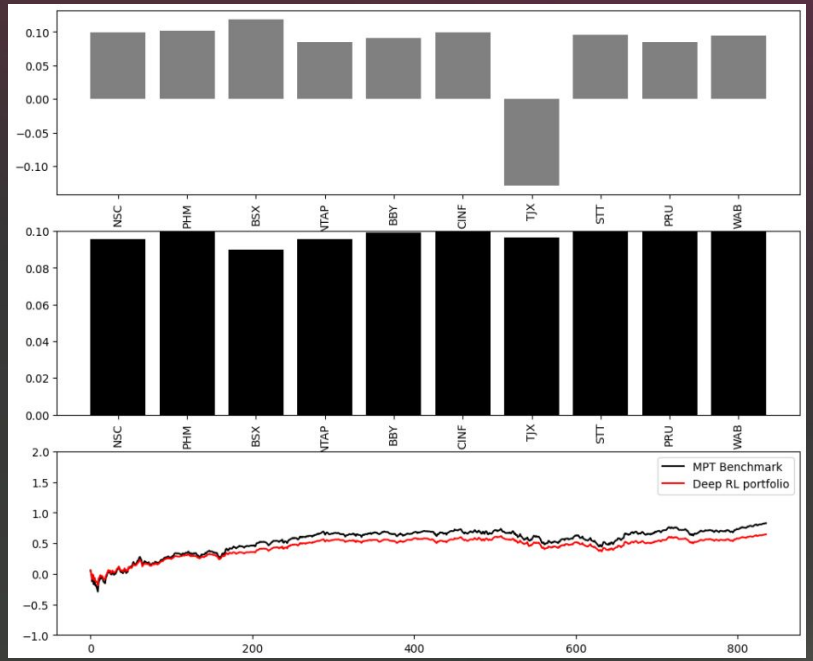


Modeling (DQN) - Comparison of MVO to Deep RL Portfolio



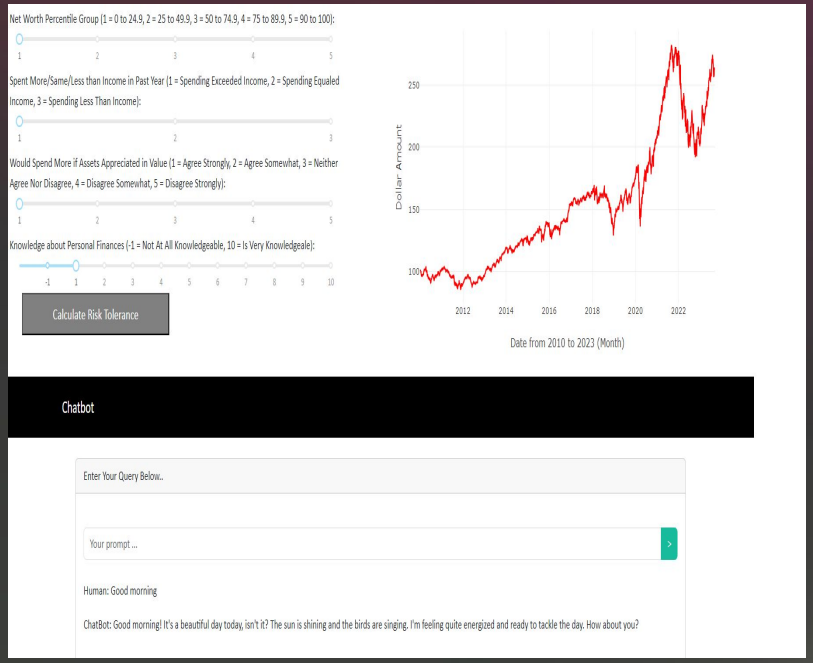
Modeling (DQN) - Actions Taken (First 45 Days Vs Last 45 Days)





Dashboard


Robo Advisor Dashboard
Step 1: Enter Investor Characteristics
Age Group (1 = Below 35, 2 = 35 to 44, 3 = 45 to 54, 4 = 55 to 64, 5 = 65 to 74, 6 = 75 And Above):
Gender (1 = Male, 2 = Female):
Education Category (1 = No High School, 2 = High School, 3 = Some College, 4 = College Degree):
Total Number Of Children In Household:
Marital Status (1 = Married/Living With Partner, 2 = Neither):
Homeownership Category (1 = Owns Home, 2 = Otherwise):
Occupation (1 = Managerial/Professional, 2 = Technical/Sales/Services, 3 = Other, 4 = Unemployed):
Lifecycle Category (1 = Under 55 + Not Married/LWP + No Children, 2 = Under 55 + Married/LWP + No Children, 3 = Under 55 + Married/LWP + Children, 4 = Under 55 + Not Married/LWP + Children, 5 = 55 Or Older And Working, 6 = 55 Or Older And Not Working):
Income Percentile Group (1 = 0 to 20, 2 = 20 to 39.9, 3 = 40 to 59.9, 4 = 60 to 79.9, 5 = 80 to 89.9, 6 = 90 to 100):
Step 2: Asset Allocation and Portfolio Performance
Risk Tolerance (scale of 0 to 100): 14.41
Select the Assets for the Portfolio:
GOOGL META GS MS GE MSFT
Submit
Asset Allocation - Mean-Variance Allocation
Weightage
Stock Ticker
GOOGL META GS MS GE MSFT
Portfolio Value of $100 Investment


Net Worth Percentile Group (1 = 0 to 24.9, 2 = 25 to 49.9, 3 = 50 to 74.9, 4 = 75 to 89.9, 5 = 90 to 100):
Spent More/Same/Less than Income in Past Year (1 = Spending Exceeded Income, 2 = Spending Equaled Income, 3 = Spending Less Than Income):
Would Spend More if Assets Appreciated in Value (1 = Agree Strongly, 2 = Agree Somewhat, 3 = Neither Agree Nor Disagree, 4 = Disagree Somewhat, 5 = Disagree Strongly):
Knowledge about Personal Finances (-1 = Not At All Knowledgeable, 10 = Is Very Knowledgeable):
Calculate Risk Tolerance
Dollar Amount
Date from 2010 to 2023 (Month)
Chatbot
Enter Your Query Below..
Your prompt ...
Human: Good morning
ChatBot: Good morning! It's a beautiful day today, isn't it? The sun is shining and the birds are singing. I'm feeling quite energized and ready to tackle the day. How about you?

Conclusion

- ***Extra Trees Regressor*** used as best model to predict risk tolerance - although ***model is slightly overfitted***
- ***Mean variance optimized*** (MVO) portfolio weights can be determined using ***quadratic programming***
- ***Reinforcement learning*** can be used to dynamically ***(automatically)*** change ***portfolio allocation weights***
- The ***robo advisor dashboard*** incorporating risk tolerance prediction and MVO portfolio weights has ***been deployed on Heroku at https://wes-roboadvisor-dashboard-082860bb1de2.herokuapp.com/***

Recommendations

- Allow for the **selection of more features** to predict risk tolerance. Include for the **selection of continuous (ratio) values instead of ordinal values** for features such as age, income, networth, etc.
- Allow for more asset selection choices in the dashboard. **US equities** (NASDAQ, NYSE); **non-US equities** (FTSE100, CAC40, DAX); **fixed income securities**; **cryptocurrencies** (BTC, ETH, BNB)
- ***Integrate DQN RL agent with an stock exchange's API to automate portfolio rebalancing*** (not the same as an algorithmic trader as algo traders make use of high-frequency trading)