Offering New Insights for Behavioral Economics Via Machine Learning

Tamerlan Ibrahimzade
Department of Management, Faculty of Economics, Azerbaijan State University of Oil and Industry
Baku, Azerbaijan
0009-0000-6209-3474

***Abstract*—Machine learning methods have been widely utilized across fields, delivering enhanced computational performance and uncovering hidden data patterns. In economics, particularly behavioral economics, their application depends on factors such as dataset nature, relationship complexity, interpretability, computational resources, and research questions. Deep learning, specifically deep neural networks, has revolutionized artificial intelligence. Based on the Turing-Church understanding of computation, these interactive methods surpass traditional algorithmic approaches like "divide and conquer" or dynamic programming, mimicking cognitive processes and enabling a deeper exploration of bounded rationality. Additionally, we examine recent applications of deep reinforcement learning (DRL) in economics. As a subset of AI, DRL effectively tackles complex problems through interactive learning, offering powerful tools for economic analysis. Balancing model complexity and interpretability is critical. This study focuses on personal consumption expenditures from a public dataset, employing Linear Regression and Decision Trees in Python for comparison.**

Keywords— machine learning, deep learning, behavioral economics, decision trees, linear regression, personal consumption expenditures, deep neural networks, deep reinforcement learning

I. Introduction

Over the past few years, the surge in data generation and sharing has sparked a keen interest in the realms of big data and machine learning analysis. While these tools have been extensively employed in domains like biology, genetics, engineering, and astronomy, their application in economics, notably in behavioral economics, is relatively nascent. The characteristics inherent in big data, including its extensive sample sizes and complex dimensions, drive the necessity and fascination with methodologies and algorithms offered by disciplines like machine learning for research within the realm of economics. These datasets, enabling a more comprehensive and intricate exploration of economic activities and interactions, are poised to shape the nature of inquiries by economists and offer a wealth of information in the future. [1]. While “rationality” based on Cambridge Dictionary definition, is the behavioral pattern where all decision- making process is done by only taking into account all the information with disregard to preference or emotions and choosing the most optimal alternative with the highest given utility. Thus, all agents make their decisions solely based on reason but history as well as human psychology has proven repeatedly that human agents usually act upon given them situations and are included by variety of factors coming from their complex cognitive process. Knowing this, American political scientist as well as economist Herbert Simon introduced the term "bounded rationality" into academia which has become one of the defining parts of field of behavioral economics. "Bounded rationality" a human decision-making process where a person attempts to satisfice, rather than optimize [2]. Alternatively speaking, person seeks a decision that will be good enough, rather than the best possible decision in the given setting. Also, the concept of "rationality" in Herbert Simon's works as a principle where humans make decisions by step-by-step or algorithmic reasoning utilizing systematic rules of logic to maximize utility. Within the context of bounded rationality, human decision maker is treated like a machine with limited resources, data as well as computational ability.

According to Stanford Encyclopedia of Philosophy, under the principle of embodied cognition, a cognitive mind is an interactive machine. So, Turing [3] and Church [4] computations are not interactive, and interactive machines can do operations Turing-Church computation can’t accomplish. That being said, history as well as human psychology has proven repeatedly that human agents usually act upon given them situations and are included by variety of factors coming from their complex cognitive process. Deep neural networks, which are one of the fundamental architectures of deep learning, have led to a revolution in artificial intelligence. These methods are interactive and not fundamentally algorithmic; their concept is usually based on the Turing understanding of computations proposed by Turing and Church. Thus, their ability to mimic certain cognitive abilities far exceeds previous algorithmic methods such as "divide and conquer," dynamic programming, greedy heuristics, etc. This will allow a deeper understanding of bounded rationality. In addition, we explore recent applications of deep reinforcement learning in economics. Deep reinforcement learning (DRL) as a subset of artificial intelligence (AI) is one of the most effective approaches to solving complex problems through interactive learning.

II. Statement of Purpose

A. Modern understanding of bounded rationality

What is special about deep learning methods such as deep neural networks (DNNs) are not their high level of computational prowess and data absorption capacity but also their ability to mimic or simulate neural process that occurs inside human brain. Hence, deep learning methods are closest process that we have than can pattern human behavior. Interesting to note that many optimization solutions in engineering are being commonly surpassed by deep neural networks (DNNs). It is common for DNNs to be described as algorithms but in reality, they are not capable of rational decision-making level. Despite the fact that DNNs are usually calculated algorithmically on computers, when DNNs deliver a result, for instance, classifying an image as a stop sign, there is no step of rational or logical steps that would be possible to point to that could be rationally classified. That being said, the classification is closer to intuition rather than rationality. The human mind is not a procedure in a brain interacting the world via sensors. On the contrary, the human mind is an interaction between processes in a brain and the world around it [5]. Also, the decision-making process implemented by DNNs are quite inexplicable, resisting any classification as rational decisions. Considering the fact that DNNs could be achieved through the computers, these achievements are brute-force simulations of procedures that are not inherently algorithmic. While the field of reservoir computing [6], for instance, gives very different architectures that have little similarity to Turing-Church calculations and would be challenging to classify as rational decision makers in the traditional sense.

The field of feedback control, which is rests fundamentally upon the interaction, requires neither computers nor algorithms. Truly, earliest applications of feedback control were in the 1920s through the 1950s which predate even digital computers. According to philosophy of science, observation overtakes interaction. It is commonly accepted that the best science is objective, not subjective. In other words, allow the data to speak for itself meaning that designing your instruments to minimally disturb what you are observing. Additionally, science also shows us that observation without interaction is unattainable. Thus, it possible to achieve much more if we embrace feedback and interaction.

Concerning brief history of Reinforcement Learning (RL), in the late 1980s, RL reached the machine learning scientific circles, where it was implemented to problems like simple mobile robots and chess. Watkin (1989) merged the temporal-difference learning and the optimal control literature in his development of Q-learning algorithm, which was the first modern RL algorithmic model. Furthermore, in the late 1990s, we observed that RL merged itself with artificial neural networks that enabled RL to be utilized to cases like learning how to play video games. While in the early 2000s, RL merged itself with evolutionary computation which enabled RL to implement to cases where controlling a robot arm was the priority. Then finally in in the late 2000s, RL was merged with deep learning that enabled RL to be utilized to cases like learning how to play video games or navigate 3D environments, thus creating Deep Reinforcement Learning (DRL) method as we know it today.

It could be said that DRL approach is a very modern approach that incorporates reinforcement learning techniques from machine learning and combines it with deep learning algorithms. The result is the fusion of these two different although highly related methods to realize end-to-end learning from perception to action. Despite different researchers having similar thoughts on initial DRL concepts it was the publication of “Playing Atari with Deep Reinforcement Learning” paper that gave DRL academic attention [7]. Afterwards, Google's DeepMind improved upon the deep Q-network's (DQN) algorithm while the studies done by LeCun [8] made DRL as one of important development directions of deep learning in the future. The premise of DRL lies in describing as well as solving problems in which agents interact and adapt learning strategy to maximize the return or realize some particular targets in the process of interacting with environment. Moreover, when comparing DRL with conventional methods, the superior performance of the proposed DRL becomes evident. This is attributed to the capability of DRL not only to statistically predict the direction of change trends but also to capture the discrete nature of environmental states, thereby significantly aiding humans in making prompt and efficient decisions.

B. Applications of deep reinforcement learning in macroeconomics and microecnomics

The utilizations of DRL methods in the field of economics and in macroeconomics more precisely speaking are developing very fast and extend from primitive single-agent optimization to comprehensive multi-agent settings. RL and DRL methods are quite novice to the field of economics, but economists have been utilizing similar models for longer period of time. The various utilizations of DRL algorithms in the field of economics and more precisely in macroeconomics are fast developing and extend from primitive single-agent optimization to sophisticated multi-agent settings. RL as well as DRL algorithms are very novice to the field of economics, but economics practitioners have been utilizing similar models for longer period of time. Implementation of DRL methods have gained recognition in recent time. More than 30 years ago, two profound academic researches on RL methods in computational economics have published by [9] as well as [10]. Also important to mention that, application of RL to particular economic topics has been done by [11]. Concerning financial economics and mathematics, [12] and [13] made important research on literature review of DRL application in risk analysis, portfolio management, insurance, pricing and bidding optimization. Although on relative scale the research focus was primarily on macroeconomics, there has been also research of RL application in microeconomics field. The work done by [14] examines RL players who learn to make resources in a spatially complex environment, trade them with each other, and consume those that they have a preference. Paper shows that the supply and demand shifts explain emergent production, consumption and pricing patterns.

It’s been observed that in case of emergence of price differences, some RL players learn to take arbitrage opportunities for profit. Overall, this academic work is part of a research initiative for the purpose of building human-like artificial general intelligence via multi-agent interactions in simulated societies. Also, this model includes heterogeneous tastes and physical abilities, and RL players negotiate with each other as a grounded type of communication. DRL has also application in monetary policy as well where it is used to determine optimal interest rate reaction function that accomplishes the inflation and output gap targets of central banks [15]. Initially, ANN is utilized to make estimation of transitions equations of an economy which is determined by the three equation New Keynesian model in [16]. Particular emphasis is given to the aggregate demand equation (i.e., the dynamic investment-saving curve), and the aggregate supply equation estimation. Finally, this estimated transition equations are used as a representation of the RL environment and train the central bank RL players to determine the optimal interest rate reaction function.

C. Utilzing machine learning models for understanding personal consumption expenditures

Here, we utilize two widely used machine learning methods namely Linear Regression and Decision Trees and compare their results by selecting a relevant variable called "Personal Consumption Expenditures" column from a publicly available "US Macroeconomic Factors and Growth" dataset. Afterwards, we extract necessary results from this dataset by using statistical evaluation metrics like mean absolute error (MAE), mean squared error (MSE), root mean squared Error (RMSE), r-squared (R2 score) for evaluation purposes in order to determine for the given data types and study purposes which method offers a superior purpose.

III. Methodology

For this study, we used a publicly available "US Macroeconomic Factors and Growth" dataset from Kaggle for comparison and evaluation purposes. Data consists of 88 rows and 83 columns in total, from which we will select "Personal Consumption Expenditures" column as it better captures the nature of research with regards to behavioral economics. As regards to software, we utilized Anaconda package which is an open-source Python/R data science distribution with over 250 pre-installed libraries as well as *Jupyter Notebook IDE* where we wrote and executed our relevant Python scripts as well as Linear Regression & Decision Trees machine learning algorithms. Initially, we imported relevant libraries for each of the selected machine learning method and common economics dataset. Linear regression is a supervised machine learning technique that establishes a linear relationship between a dependent variable and one or more independent features by fitting a linear equation to observed data. In instances where there is only a single independent feature, it is termed as Simple Linear Regression, whereas when multiple features are present, it is referred to as Multiple Linear Regression. Similarly, if there is just one dependent variable, it is labeled as Univariate Linear Regression, while the presence of multiple dependent variables characterizes it as Multivariate Regression. The strength of linear regression lies in its interpretability. The model's equation reveals clear coefficients that delineate the influence of each independent variable on the dependent variable, facilitating a deeper comprehension of the underlying dynamics. Its simplicity is regarded as advantageous, as linear regression is transparent, straightforward to implement, and forms the basis for more intricate algorithms.

For Linear Regression method, we import *numpy, pandas* for data analysis, *matplotlib, pyplot* and *seaborn* for data visualization, filter warnings sub-library from warnings library and all relevant sub-libraries (pipeline, impute, compose, preprocessing, model_selection, linear_model, metrics) from *sklearn* Python machine learning library. The Decision Tree algorithm falls within the realm of supervised learning methods. Differing from other techniques in this category, Decision Trees can address both regression and classification tasks. The objective behind employing a Decision Tree is to construct a model based on training data, which can then be utilized to forecast the class or value of the target variable by deducing straightforward decision rules from historical data. While for the Decision Tree method, we import all libraries and sub-libraries that we did for Linear Regression method plus additional tree and ensemble sub-libraries from *sklearn* Python machine learning library. Afterwards, for both methods we come to data preparation stage which involves splitting data into features (X) and target variable (y) as well as splitting the dataset into training and testing sets using train_test_split(). Then, we proceed to model initialization by assigning each relevant method for Linear Regression and Decision Trees respectively and building their pipelines. After that, comes the model training part where we fit the training variables into the machine learning pipe that built. Finally, we use statistical evaluation benchmarks like mean absolute error (MAE), mean squared error (MSE), root mean squared Error (RMSE), r-squared (R2 score) for evaluation purposes and compare each model in order to see which is a better fit.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \ldots \ldots \beta_n X_n \qquad [17]$$

Where is Y is the dependent variable, $X_1, X_2, \ldots, X_p$ are the independent variables, β_0 is the intercept and $\beta_1, \beta_2, \ldots, \beta_n$ are the slopes.

MAE function can be calculated as:

$$\text{Cost function}(J) = \frac{1}{n}\sum_{n}^{i} |\widehat{y_i} - y_i| \qquad [18]$$

Where the predicted value is (y) with hat and the true y value is (y) while n is the total number of observations.

MSE function can be calculated as:

$$\text{Cost function}(J) = \frac{1}{n}\sum_{n}^{i} (\widehat{y_i} - y_i)^2 \qquad [19]$$

Where the predicted value is (y) with hat and the true y value is (y) while n is the total number of observations.

While RMSE can be calculated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{n}^{i} (\widehat{y_i} - y_i)^2} \qquad [20]$$

Where the predicted value is (y) with hat and the true y value is (y) while n is the total number of observations.

$$Entropy(p) = -\sum_{i=1}^{N} p_i log_2 p_i \qquad [21]$$

Where p is the whole dataset, N is the number of classes, and pi is the frequency of class i in the same dataset

$$R^2 = 1 - RSS/TSS) \qquad [22]$$

Where R^2 represents the required R Squared value, *RSS* represents the residual sum of squares, and *TSS* represents the total sum of squares.

$$Information\ Gain = Entropy(before) - \sum_{j=1}^{K} Entropy(j, after) \quad [23]$$

Where "before" is the dataset before the split, K is the number of subsets generated by the split, and (j, after) is subset j after the split.

IV. Models' Comparison Results

Two Machine Learning Models' Performance Summary for Personal Consumption Expenditures as follows:

Linear Regression

```
Evaluation Metrics for Train Data:
Mean Absolute Error (MAE): 169.5276151352347
Mean Squared Error (MSE): 92090.73684006045
Root Mean Squared Error (RMSE): 303.46455615122574
R-squared (R2 Score): 0.9860129384569957
```

Decision Trees

```
Evaluation Metrics for Train Data:
Mean Absolute Error (MAE): 3.0143252973045624e-12
Mean Squared Error (MSE): 2.5997104965952297e-23
Root Mean Squared Error (RMSE): 5.098735624245711e-12
R-squared (R2 Score): 1.0
```

V. Conclusion

This study introduces bounded rationality, a concept by Herbert Simon, emphasizing that individuals prioritize satisfactory decisions over optimal ones. Deep neural networks (DNNs) are highlighted for their potential to simulate neural processes, offering insights into bounded rationality. The progression from reinforcement learning (RL) to deep reinforcement learning (DRL) is explored, showcasing the integration of artificial neural networks with deep learning. DRL combines RL and deep learning, enabling end-to-end learning from perception to action. By interacting with the environment to maximize rewards, DRL models effectively handle multidimensional, continuous state-action spaces and dynamic environments, enhancing their capability to learn and recognize patterns. The adoption of machine learning techniques in applied and behavioral economics provides methodological and practical advantages over traditional statistical methods. In analyzing the "US Macroeconomic Factors and Growth" dataset, linear regression outperformed decision trees in regression tasks, yielding lower mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE), alongside higher R-squared metrics. These results suggest that linear regression is a better fit for this dataset across statistical benchmarks. However, ensemble methods like Random Forests or Gradient Boosting, often superior to decision trees, were not considered here. While linear regression delivers stronger performance in standard cases, it requires careful fine-tuning and more complex steps compared to decision trees. Ultimately, selecting between these methods depends on the dataset's characteristics, problem nature, and the trade-offs between model complexity and interpretability. Nonetheless, there are also alternative approaches like Fuzzy Recurrent Neural Network (FRNN) where experimental result on FRNN show that, FRNN is a good forecasting tool for forecasting electric energy consumption and sunspot problem [24]. FRNN could also be utilized for personal consumption expenditure forecasting in future. All being said, together with DRL advancements, these tools demonstrate how machine learning methods can address both theoretical and practical challenges in economics.

References

[1] Einav, L. and Levin, J. (2014). Economics in the Age of Big Data. Science, 346, 715-721.

[2] Simon, H. A. (1972). "Theories of bounded rationality," in Decision and Organization, eds C. B. McGuire and R. Radner (Amsterdam: North-Holland Publishing Company), 161–176.

[3] Turing, A. M. (1936). On computable numbers with an application to the entscheidungsproblem. Proc. Lond. Math. Soc. 42, 230–265.

[4] Church, A. (1932). A set of postulates for the foundation of logic. Ann. Math. 32, 346–366.

[5] Thelen, E. (2000). Motor development as foundation and future of developmental psychology. *International Journal of Behavioral Development,24*(4), 385-397.

[6] Tanaka, G., Yamane, T., Héroux, J. B., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D., & Hirose, A. (2019). Recent advances in physical reservoir computing: A review. *Neural Networks, 115*, 100–123.

[7] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). *Playing Atari with Deep Reinforcement Learning*. arXiv preprint arXiv:1312.5602.

[8] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.

[9] Barto, A. G., & Singh, S. P. (1991). On the computational economics of reinforcement learning. In *Connectionist Models: Proceedings of the 1990 Summer School* (pp. 35–44). Morgan Kaufmann.

[10] Arthur, W. B. (1991). Designing economic agents that act like human agents: A behavioral approach to bounded rationality. *American Economic Review, 81*(2), 353–359.

[11] Hughes, N. (2014). *Applying reinforcement learning to economic problems*. ANU Crawford PhD Conference, Canberra.

[12] Mosavi, A., Faghani, Y., Ghamisi, P., Duan, P., Faizollahzadeh Ardabili, S., Salwana, E., & Band, S. S. (2020). Comprehensive review of deep reinforcement learning methods and applications in economics. *Mathematics, 8*(10), 1640.

[13] Charpentier, A., Elie, R., & Remlinger, C. (2020). Reinforcement learning in economics and finance. *Computational Economics*

[14] Johanson, M. B., Hughes, E., Timbers, F., & Leibo, J. Z. (2022). Emergent bartering behaviour in multi-agent reinforcement learning. *arXiv preprint arXiv:2205.06760*.

[15] Hinterlang, N., & Tänzer, A. (2022). Optimal Monetary Policy Using Reinforcement Learning. *Deutsche Bundesbank Discussion Paper No. 51/2021*.Jkjl

[16] Rotemberg, J. J., & Woodford, M. (1997). An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy. In *NBER*

Macroeconomics Annual 1997, Volume 12 (pp. 297–361). MIT Press.
[17] https://www.geeksforgeeks.org/ml-linear-regression/;
[18] https://www.geeksforgeeks.org/how-to-calculate-mean-absolute-error-in-python/;
[19] https://www.geeksforgeeks.org/mean-squared-error/;
[20] https://www.geeksforgeeks.org/step-by-step-guide-to-calculating-rmse-using-scikit-learn/;
[21] https://towardsdatascience.com/from-a-single-decision-tree-to-a-random-forest-b9523be65147;
[22] https://www.geeksforgeeks.org/r-squared/;
[23] https://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html;
[24] Gardashova L., Hasanli N., Akhmadov S. (2019). Benchmark Fuzzy Time-Series Forecasting Using Recurrent Neural Network. *Black Sea. Scientific journal of academic research, volume 51, issue 08,* 48-56.