Author:

P. Carmel Marie Frédérique ZAGRÉ

1 Introduction

The study of the human decision-making process has long been a focal point in research, offering invaluable insights across various domains. Accurately modeling human decision-making holds significance in fields ranging from economic analysis to forecasting human behavior in scenarios such as autonomous vehicle operation (Plonsky et al. (2018)).

Despite its importance, human decision-making frequently diverges from rational, optimal choices. Moreover, existing decision-making models often apply only to specific contexts and face challenges related to dimensionality. Capturing all the factors influencing individual choices proves to be exceptionally challenging.

Insights from behavioral economics, cognitive science, and psychology have provided valuable tools for understanding specific behavioral phenomena associated with decision-making. However, significant room for improvement remains.

In efforts to advance the field, two prediction competitions, CPC15 and CPC18, were launched. These competitions tasked researchers with developing a unified model capable of addressing 14 classical choice anomalies documented in behavioral decision research (Plonsky et al., 2018). The outcomes of these competitions suggested the feasibility of creating models with high predictive accuracy, capable of encompassing various human behavioral phenomena.

Our project focuses on utilizing the data from these competitions to predict choice behavior. Specifically, our objective involves employing neural network modeling along with risk preference and attention-related covariates to forecast choice behavior. We have engineered our own metrics for assessing risk preference and attention-related factors and present an analysis of their correlations with choice behaviors.

2 Dataset

In this project, we used the individual data provided by CPC18 in which a total of 60 problems are presented to a decision maker. The data was collected by the experiment paradigm described in (Plonsky et al., 2018). For each individual, we observe diverse demographics.

In each problem, the decision maker gets descriptions of two monetary prospects and choose between them. Such process is repeated 25 times for each problem. These 25 trials are divided into 5 blocks and 5 trials each, and in the first block the decision maker receives no feedback about the payoffs while receiving full feedback in all later blocks. The rate of choosing option B in each block is the mean of the five trials within it. A problem is defined by 12 parameters as shown in Table 1, including the descriptions of two lotteries, and game settings :

3 Methods

We aim to forecast the likelihood of selecting option B within each block of the five trials. Our approach involves predicting individual choices across these trials and subsequently

Tableau 1: Problem Parameters

Parameter	Meaning
Ha	Expected value of payoff of Lottery A
pHa	Probability of getting payoff of Lottery A
La	Alternative outcome of Lottery A
LotNumA	Number of possible outcomes in Lottery A
LotShapeA	Shape of distribution of Lottery A
Hb	Expected value of payoff of Lottery B
pHb	Probability of getting payoff of Lottery B
Lb	Alternative outcome of Lottery B
LotNumB	Number of possible outcomes in Lottery B
LotShapeB	Shape of distribution of Lottery B
Amb	Whether prob. of outcomes in B are revealed
Corr	Correlation between two Lotteries

computing the rate of choosing option B within each block. We then compare these predicted rates against the actual outcomes.

Our neural network incorporates critical variables, including the payouts and probabilities associated with options A and B, as well as participant responses and demographics. Additionally, we integrate two primary types of input: raw data and engineered features designed to capture latent variables such as risk preference and attention.

Initially, our baseline model includes the risk feature. Subsequently, we enhance the model by incorporating attention features (augmented model) and assess their respective contributions to predicting choice behavior. This comparative analysis allows us to evaluate the relative importance of these features in our predictive framework.

3.1 Data cleaning

To be able to feed some of the raw data, we needed to work on them. To be more specific, some the raw variable were string variable which we needed to encode into e numeric variable. In table 2, we present these variables, the meaning of their new numeric values.

3.2 Engineering data

In addition to raw variables, we construct features related to risk preference and attention. To generate these variables, we initially reconstruct the distributions of options A and B. Since the dataset lacks complete information on the lottery distributions, we employ functions established by prior researchers to recover these distributions. Subsequently, we define the risk preference and attention variable.

- Risk preference variable :

Tableau 2: Encoding variable

Variables	String value	Numeric value
gender	F	1
gender	M	2
location	Rehovot	1
location	Technion	2
condition	ByProb	1
button	L	1
button	R	2
lotshapea (lotshapeb)	-	1
lotshapea (lotshapeb)	L-skew	2
lotshapea (lotshapeb)	R-skew	3
lotshapea (lotshapeb)	Symm	4

– **Variance (varA, varB)**: Variance measures the dispersion of possible outcomes in a lottery. Individuals with higher risk tolerance may prefer lotteries with high variance, as they offer the possibility of large gains despite higher likelihood of losses.

– **Dominance**: Equal to 1 if neither lottery stochastically dominates the other, 2 if A stochastically dominates B, and 3 otherwise. Studying choices between such lotteries reveals risk attitudes and decision-making strategies. Risk-averse individuals favor lotteries they dominate, while risk-seekers may prefer dominated ones, showing their willingness to take on higher risk for greater rewards.

– **Skewness (skewA, skewB)**: Skewness measures the asymmetry of outcome distribution. Positive skewness in lotteries may attract risk-seeking individuals who are willing to accept higher likelihood of losses for the chance of substantial gains.

– **Expected Value (EV_A, EV_B)**: Expected value represents the average outcome of a lottery. Risk-seeking individuals may prefer lotteries with higher expected values, prioritizing potential gains over likelihood of losses.

– **Difference in Expected Value (diffEV)**: Significant differences in expected value between lotteries may indicate risk-seeking behavior, as individuals prioritize options with higher potential gains even if they come with higher variability.

– **Probability Amplification (HAPWA, LAPWA, HAPWB, LAPWB)**: Amplifying probabilities associated with high outcomes may suggest a preference for high-risk, high-reward scenarios among risk-seeking individuals.

- Attention variable

– **Change in Reaction Time (change_reactiontime)**: Variations in reaction time across blocks may indicate shifts in attentional allocation. Longer reaction times suggest more deliberation, while shorter times may indicate quicker, automatic responses.

– **Behavior Change (behavior_change)**: Measures how often a participant changes behavior within a block. Salient events or stimuli may trigger these changes, reflecting shifts in attention and decision-making.

– **Behavior Change after Feedback (behavior_change_af)**: Binary outcome indicating whether behavior changes after feedback. Reflects participants' sensitivity to feedback information and its role in guiding decision-making.

– **Relative Order (RelativeOrder)**: Reflects changes in attentional focus over time. Fluctuations may indicate shifts in attentional engagement or arousal levels throughout the task.

3.3 Model

The model is the following :

```
 model1 = Chain(
Dense(size(X_train1, 2), 64, tanh),
Dense(64, 64, relu),
Dense(64, 64, tanh),
Dense(64, num_classes1),
Dense(num_classes1, 2, ),
softmax
)
```

The use of tanh and ReLU activation functions in the hidden layers introduces non-linearity into the model. Tanh is effective for capturing non-linear relationships, while ReLU accelerates convergence during training. The output layer configuration is tailored to the specific requirements of the prediction task. By using a sigmoid activation function followed by softmax, the model produces probabilistic predictions for each class, allowing for multi-class classification with interpretable outputs.

4 Results and Discussion

Both baseline and augmented models use the same architecture, consisting of three hidden layers with tanh and ReLU activation functions, followed by an output layer with softmax activation.

- Accuracy Rate: For the Baseline Variables, the accuracy rate is 99.76%. This indicates that the model using only the baseline variables achieved a very high accuracy in

Tableau 3: Variables by model

Raw Variables	age set ha pha la lotnuma hb phb lb lotnumb amb corr block Gender Location Condition ShapeA ShapeB order trial payoff forgone
Baseline Variables	Raw variable + hapwa lapwa hapwb lapwb eva evb diffev skewa skewb dominance
Augmented Variables	Baseline variable + relativeorder change_reactiontime behavior_change_af feedback rt

Tableau 4: Results

	Accuracy rate	MSE
Baseline Variables	0.9976	21.07%
Augmented Variables	0.8728	37.33%

predicting outcomes. However, when using the attention Variables, the accuracy rate drops to 87.28%. This suggests that the model's performance decreased when attention variables were included.

- For the Baseline Variables, the MSE is 21.07%. This indicates a relatively low error rate in prediction. In contrast, the "Augmented Variables" have a higher MSE of 37.33%. This suggests that the model's predictions were less accurate when using the augmented variables.

The results suggest that adding the attention variables to the model did not improve its performance and may have even led to a decrease in accuracy and an increase in prediction errors.

5 Comparison with Structural Models

In this section, let's compare the BLP model and the neural network model.

Firstly, let's discuss the trade-off between flexibility and interpretability. Neural network models offer more flexibility in capturing complex relationships, but they may lack the interpretability of BLP models. BLP models provide clear insights based on microeconomic theory, making it easier to understand the economic implications of the model's coefficients. While the variables used in neural networks can be derived from economic theory, discerning their relative importance in our prediction task can be challenging.

Next, let's consider data-driven learning versus theory-driven Modeling. Neural networks learn directly from data, allowing them to capture patterns and relationships without relying on specific economic theory assumptions. They aim to recover the true Data Generating

Process (DGP) without imposing any preconceived notions. In contrast, BLP models are built on microeconomic theory, offering interpretable insights grounded in economic principles. While this theory-driven approach ensures consistency with economic theory, it may overlook patterns in the data that are not explicitly accounted for in the theory.

Finally, the complexity of decision Making. Neural networks excel at capturing complex decision-making processes, where choices depend on numerous factors interacting in non-linear ways. On the other hand, while BLP models are theoretically consistent, they may struggle to capture the same level of complexity with numerous factors. BLP models typically rely on specific functional forms and assumptions about consumer behavior, which may oversimplify the decision-making process in certain scenarios.

Bibliography

Plonsky, O., Apel, R., Erev, I., Ert, E., & Tennenholtz, M. (2018). When and how can social scientists add value to data scientists? a choice prediction competition for human decision making [Unpublished Manuscript].