

Choice Prediction Competition 2018
Final exam submission for ECON 832: Computational Economics
Derek Situ
04/20/2024

Overview

[Choice Prediction Competition 2018](#) (CPC18) was a competition focused on predicting human choices over lotteries. The dataset for this competition comes from an experiment in which decision-makers are given descriptions of two monetary gambles. Each respondent faces 25 trials of 30 games. For the first 5 trials, respondents do not receive feedback regarding the outcomes generated by each lottery within the trial. From trial 6 onwards, respondents receive this feedback. To incentivize honest answers, respondents were compensated based on a lottery they chose, selected at random.

Our goal is to train a model to predict respondents' choices using the data provided which includes information on the decision-makers' demographics, the options within trials, and the games.

Of particular interest is the role attention may play in affecting choice. In this analysis, I consider two main mechanisms through which attention may affect choice. The first is fatigue. As a decision-maker progresses through the experiment, it is possible that they become fatigued, reducing their ability to pay attention. Secondly, it is possible that the number of sources of information a decision-maker is faced with can affect their attention. That is, it is possible that having too many sources of information demanding attention at once can affect decision-making.

Another important driver of choice-making in this context is risk preference. As the lotteries present varying levels of payoffs and their probabilities, an individual's preference for risk can systematically influence their choices in this experiment. I compare the performance of two models, one with features related to attention and risk, and another without these features, in order to evaluate their importance in predicting choice behaviour.

Methodology and Feature Analysis

The outcome to be predicted, called B in the raw data, is whether the respondent selected lottery B rather than A. Each observation in the raw data represents a trial played by a particular individual.

I use 36 total features for the model with attention and risk features, and 16 features for the model with no attention and risk features. From the raw data I use **Gender**, **Age**, **Set** (identifier for the set of games the respondent faced), **GameID**, **Ha** (expected value of high lottery in option A), **pHa** (probability to get payoff drawn from lottery in option A), **La** (low payoff in option A), **LotShapeA** (shape of distribution of outcomes in option A), **LotNumA** (number of outcomes in option A), **Hb** (expected value of high lottery in option B), **pHb** (probability to get payoff drawn from lottery in option B), **Lb** (low payoff in option B), **LotShapeB** (shape of distribution of outcomes in option B), **LotNumB** (number of outcomes in option B), **Amb** (whether option B is ambiguous i.e. probabilities not described), **Corr** (whether payoffs generated by the two options are correlated and the sign of correlation (-1/0/1), **Order** (serial

position of the current game, 1-30), ***Trial*** (trial number within a game, 1-25), **Button** (on-screen side of the chosen button (L/R)), **Payoff** (payoff obtained by the subject in the trial), **Forgone** (payoff the subject would have obtained if they had chosen the other option), **RT** (reaction time in milliseconds), **Apay** (payoff provided by option A), **Bpay** (payoff provided by option B), **Feedback** (whether feedback was provided for the trial), and **block** (serial position of the current 5-trial block within the game). Features I deem related to attention or risk, and remove for the second model, are in italics. From each of *LotShapeA* and *LotShapeB* I generated 4 dummies, as these variables were categorical with 4 possible responses (blank, symmetric, left-skewed, or right-skewed). From *RT* I generated 2 dummies: *RT_NA* if *RT* is missing, and *RT_high* if *RT* is above the median.

I generate three additional features related to attention or risk, for use in the first model. The first is ***fatigue***, a fatigue index calculated as $0.66 \times Order + 0.33 \times Trial$. As the experiment progresses, it is possible that participants become fatigued, and less able to focus. The second is ***info_overload***, an index intended to capture signs of potential information overload which might overwhelm a participant's focus. It is calculated as $LotNumA + LotNumB - Amb + RT_high + Feedback$. The third engineered feature is ***diff***, calculated as $Hb - Ha$, which captures the disparity in expected values of the high outcomes. This is intended to measure the riskiness of the choice.

Through trial and error, I found that the location of the experiment (either Rehovot or Technion) and *Condition*, a legacy variable from CPC15, were poor predictors of the outcome. In order to achieve improved model performance and parsimony, I did not include these variables from the raw dataset in my analysis.

I train a feedforward neural network with 5 hidden layers using the leakyReLU activation function, and softmax as the output layer. For the model with attention and risk features, the first layer has 36 inputs and outputs, the second layer also has 36 outputs, the third layer has 20, the fourth layer has 10, and the fifth layer has 2 outputs. For the model with no attention and risk features, the first layer has 16 inputs and outputs, the second layer also has 16 outputs, the third layer has 12, the fourth layer has 10, and the fifth layer has 2 outputs. The loss function used is mean squared error (MSE), and the optimiser used is gradient descent with a learning rate of 0.5.

Results and Discussion

Model outcomes are shown in Table 1. In-sample and out-of-sample loss and accuracy are reported for the model with attention and risk features as well as the model without attention and risk features.

Table 1: Results	Model	
	With Attention/Risk	No Attention/Risk
Loss (MSE, in-sample)	0.0376	0.0311
Loss (MSE, out-of-sample)	0.0433	0.0358
Accuracy (in-sample)	0.9601	0.9660
Accuracy (out-of-sample)	0.9509	0.9723

The model without features deemed related to attention or risk performed slightly better in all aspects, which could suggest that these were not significant factors affecting choice, or that they were not modelled in a way that provided more predictive power.

Comparison with Structural Models

The standard comparative advantages between neural networks and structural models applies. Structural models explicitly model the data-generating process, quantifying the relationships between the predictors and the outcome. For example, using the BLP method of demand estimation, one can obtain estimates of elasticities and then quantify the effect of prices on demand (while taking into account simultaneous causality!). Structural models are interpretable by design, as they are made to describe relationships. However, interpretability typically comes at the cost of not modeling complexity.

On the other hand, neural networks tend to model complexity well. The nested layers and activation functions are able to capture complex relationships that are difficult to describe with closed-form expressions, but are harder to interpret. However, neural networks often provide highly accurate out-of-sample prediction power. Applied to the context of consumer choice behaviour, neural networks can allow for accurate prediction of consumer choice, but are not suited for identifying the channels through which the variables affect the outcome.

This primary difference between structural models and neural networks can also be seen in their dependence on assumptions. Structural models often make explicit assumptions about a data-generating process, while neural networks are typically free of such assumptions.