Applying Deep Learning Models to Consumer Choice Theory in Western Economics: A Case Study on Consumer Preference Prediction

Sicheng Geng[1,*], Fangfei Liu[2], Meilin Gong[3], Tianhai Wu[1], Yufei Li[3]

[1]Finance School, Harbin University of Commerce, Harbin, 150028, China
[2]School of Economics, Harbin University of Commerce, Harbin, 150028, China
[3]College of Public Finance and Administration, Harbin University of Commerce, Harbin, 150028, China

*Corresponding author: 2440819051@qq.com

Abstract. This paper explores the application of Multilayer Perceptron (MLP) models in predicting consumer preferences within the framework of Western economics' consumer choice theory. Leveraging a dataset encompassing a wide array of consumer choices across different product categories, this study employs an MLP to simulate and forecast consumer behavior in response to varying prices and product characteristics. This research contributes to the growing body of literature at the intersection of artificial intelligence and economics, offering a novel approach to understanding and predicting consumer behavior.

Keywords: Consumer Choice Theory, Deep Learning, Consumer Preferences.

1 Introduction

1.1 Background on Consumer Choice Theory in Western Economics

Western economics is deeply rooted in the understanding of how consumers make choices in a market economy. Consumer choice theory, a key component of this understanding, examines the decision-making process of consumers as they allocate their limited resources among competing goods and services.

1.2 Importance of Understanding Consumer Preferences in Economic Modeling

The accurate modeling of consumer preferences is essential for predicting market outcomes and informing economic decisions. Traditional economic models often struggle to capture the complexity and variability of consumer choices, particularly in the face of new products, changing prices, and evolving market conditions.

Q. Wu et al. (eds.), *Proceedings of the 2024 3rd International Conference on Public Service, Economic Management and Sustainable Development (PESD 2024)*, Advances in Economics, Business and Management Research 309,
https://doi.org/10.2991/978-94-6463-598-0_44

1.3 Brief Overview of Deep Learning and Its Relevance to Economic Analysis

Deep learning, a subset of machine learning, has emerged as a powerful tool for processing and analyzing large volumes of data. It involves artificial neural networks with multiple layers that can learn and represent complex patterns in data. This capability makes deep learning particularly suited for economic analysis, where the relationships between variables can be intricate and non-linear.

2 Literature Review

2.1 Review of Traditional Consumer Choice Theory Models

Consumer choice theory has its roots in the seminal works of economists such as Paul Samuelson and Robert Stone, who developed the theory of revealed preference. The linear expenditure system (LES), developed by Klein and Rubin, is a popular model used in situations where data is scarce due to its simplicity and ease of implementation [1].

2.2 Examination of Existing Studies Using AI in Economic Modeling

The use of AI in economic modeling has been reviewed extensively, with studies highlighting the potential of these models to handle large datasets and uncover complex patterns in economic data[2].

2.3 Analysis of the Role of MLP in Prediction Tasks

The Multilayer Perceptron (MLP), a fundamental deep learning model, has been widely used in prediction tasks across various economic fields. The effectiveness of MLPs in these tasks is attributed to their layered structure, which allows for the learning of complex representations of data. However, the "black box" nature of MLPs, due to the complexity of their internal representations, has also been a subject of criticism, with calls for improved interpretability in economic modeling[3].

3 Methodology

3.1 Description of the Multilayer Perceptron (MLP) Model

The Multilayer Perceptron (MLP) is a class of feedforward artificial neural network that consists of at least three layers of nodes: an input layer, one or more hidden layers, and an output layer. The architecture of an MLP allows for the modeling of complex patterns and non-linear relationships through its layered structure. Mathematically, the output of a neuron in the hidden or output layer can be represented as follows:

$$a^{(l)} = f\left(W^{(l)}a^{(l-1)} + b^{(l)}\right)$$

where $(a^{(l)})$ represents the activation of the neurons in layer (l), $(W^{(l)})$ is the weight matrix connecting the neurons from layer $(l-1)$ to (l), $(b^{(l)})$ is the bias vector, and (f) is the activation function such as the sigmoid or ReLU function.

3.2 Data Collection and Preprocessing Steps

The dataset used in this study is preprocessed as follows:

1. Handling Missing Values: Missing data is imputed using strategies such as mean imputation or more complex methods like k-Nearest Neighbors (KNN) imputation.

2. Feature Scaling: Features are scaled using techniques like Min-Max scaling or Standardization to ensure that all features contribute equally to the distance calculations in the neural network.

3. Encoding Categorical Variables: Categorical variables are encoded using one-hot encoding or embedding layers if they are ordinal in nature[4].

3.3 Explanation of the Training Process for the MLP Model

The training process of an MLP model involves the following steps:

1. Initialization: Weights (W) and biases (b) are initialized randomly using methods like Xavier/Glorot initialization to break symmetry and encourage equal contribution from all neurons in the initial state.

2. Forward Propagation: The input data is passed through the network layer by layer. For each layer (l), the input (x) is transformed into the output (y) using the weights and biases:

$$z^{(l)} = W^{(l)}x + b^{(l)}$$

$$y^{(l)} = f\left(z^{(l)}\right)$$

3. Loss Calculation: The output layer's activations are compared against the true labels to calculate the loss, typically using the mean squared error (MSE) for regression tasks or cross-entropy loss for classification tasks.

4. Backpropagation: The error is propagated back through the network, and the gradients of the loss with respect to each weight and bias are computed. The gradient for a weight \(W^{(l)} \) can be calculated as:

$$[\frac{\partial L}{\partial W^{(l)}} = \frac{\partial L}{\partial y^{(l)}} \cdot \frac{\partial y^{(l)}}{\partial z^{(l)}} \cdot \frac{\partial z^{(l)}}{\partial W^{(l)}}]$$

5. Parameter Update: The weights and biases are updated using an optimization algorithm such as stochastic gradient descent (SGD) or Adam, with the update rule:

$$W^{(l)} \leftarrow W^{(l)} - \eta \frac{\partial L}{\partial W^{(l)}}$$

$$b^{(l)} \leftarrow b^{(l)} - \eta \frac{\partial L}{\partial b^{(l)}}$$

where (η) is the learning rate[5].

6. Regularization: Techniques such as L2 regularization or dropout are applied to prevent overfitting by penalizing large weights or randomly dropping neurons during training.

7. Epoch Iteration: The training process is iterated over multiple epochs until the model's performance on a validation set no longer improves significantly, indicating that the model has learned the underlying patterns in the data.

4 Data and Model Specification

4.1 Variables Selection and Their Economic Significance

The selection of variables for this study is grounded in economic theory and practical considerations. Key variables include:

- Price (P): Represents the cost of goods, a fundamental determinant of consumer demand.
- Quantity (Q): Indicates the number of units purchased, reflecting consumer demand[6].
- Income (I): Proxy for purchasing power, which affects consumer choices.
- Product Features (F): Include aspects like quality, brand, and design, which influence consumer preferences.

These variables are central to the law of demand and utility maximization in consumer choice theory. They are selected to capture the essence of consumer decision-making processes.

4.2 Mathematical Formulation of the MLP Model

The MLP model is formulated as follows:

1. Input Layer: Receives a vector of features $\boldsymbol{X} = [P, Q, I, F]$.

2. Hidden Layers: Comprise neurons with weights $\boldsymbol{W}$ and biases $(\boldsymbol{b})$. The activation (a) of each neuron is calculated by:

$$a = f(\boldsymbol{W}\boldsymbol{X} + \boldsymbol{b})$$

where (f) is the activation function.

3. Activation Functions: Include the sigmoid function for binary classification tasks:

$$f(x) = \frac{1}{1 + e^{-x}}$$

and the Rectified Linear Unit (ReLU) for hidden layers:

$$f(x) = \max(0, x)$$

4. Output Layer: Produces the predicted consumer choice, which can be a probability (for classification) or a continuous value (for regression). For a binary choice, the output is given by:

$$\hat{y} = \sigma(\boldsymbol{W_o} a + b_o)$$

where (σ) is the sigmoid function, and $(\boldsymbol{W_o}), (b_o)$ are the output layer's weights and bias.

5. Loss Functions: The model is trained to minimize the loss function, which measures the discrepancy between predicted and actual consumer choices. Common loss functions include Mean Squared Error (MSE) for regression[7]:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2$$

and Cross-Entropy Loss for classification:

$$\text{Cross-Entropy} = -\sum_{i=1}^{n}[y_i \log(\hat{y_i}) + (1 - y_i)\log(1 - \hat{y_i})]$$

6. Backpropagation: The error is backpropagated through the network, and the gradients are used to update the weights and biases via stochastic gradient descent (SGD) or other optimization algorithms.

This mathematical formulation encapsulates the core components of the MLP model, providing a rigorous framework for predicting consumer choices based on economic variables. The model's architecture and training process are designed to capture the complexities of consumer behavior as influenced by economic factors[8].

5 Economic Theory Integration

5.1 Integration of Consumer Choice Theory into the MLP Framework

The integration of consumer choice theory into the MLP framework is achieved by translating economic principles into mathematical expressions that the MLP can process. Consumer choice theory posits that consumers will choose the combination of goods that maximizes their utility, subject to their budget constraint. This can be represented as:

$$U = U(X_1, X_2, \ldots, X_n)$$

$$P_1X_1 + P_2X_2 + \cdots + P_nX_n \leq I$$

where (U) is the utility function, (X_i)represents the quantity of good (i), (P_i) is the price of good (i), and (I) is income. The MLP model is trained to predict the optimal basket of goods $(X_1, X_2, \ldots, X_n)$ that consumers are most likely to choose given the prices and income levels[9].

5.2 Derivation of Specific Formulas for Predicting Consumer Preferences

To predict consumer preferences, we utilize the concept of utility maximization. The MLP model is trained to approximate the indirect utility function, which can be derived from consumer choice theory. The indirect utility function \(V \) is given by[10]:

$$[V(P, I) = \max_{X} U\,(X)]$$

$$[\text{subject to } P_1X_1 + P_2X_2 + \cdots + P_nX_n \leq I]$$

The model learns to predict the utility-maximizing combination of goods, which corresponds to the consumer's most preferred choice set. The activation of the output layer in the MLP represents the predicted quantities of each good in the optimal choice set.

5.3 Mathematical Derivations and Explanations of Economic Models Embedded within the MLP

The economic models embedded within the MLP are based on the principles of utility maximization and budget constraints[11]. The training process involves adjusting the weights and biases of the MLP to minimize the difference between the predicted choice set and the actual choice set observed in the data. The mathematical derivation involves:

Forward Propagation. The input prices and income are passed through the MLP, with each neuron in the hidden layers computing a weighted sum of its inputs and applying an activation function:

$$z_j^{(l)} = \sum_i w_{ij}^{(l)} x_i + b_j^{(l)}$$

$$a_j^{(l)} = f\left(z_j^{(l)}\right)$$

where $(z_j^{(l)})$ is the weighted sum of inputs to neuron $(\,j\,)$ in layer $(\,l\,)$, $(w_{ij}^{(l)})$ are the weights, $(b_j^{(l)})$ is the bias, and $(\,f\,)$ is the activation function.

Backpropagation and Gradient Descent. The error between the predicted and actual choice sets is calculated using a loss function, such as mean squared error. The gradient of the loss function with respect to the weights is computed, and the weights are updated using gradient descent:

$$\Delta w_{ij}^{(l)} = -\eta \frac{\partial L}{\partial w_{ij}^{(l)}}$$

where $(\,\eta)$ is the learning rate and $(\,L\,)$ is the loss function.

Regularization. To prevent overfitting, regularization techniques such as L2 regularization are applied, which add a penalty term to the loss function for large weights:

$$L_{\text{reg}} = L + \frac{\alpha}{2}\sum_{l,j}\left(w_{ij}^{(l)}\right)^2$$

where (α) is the regularization parameter.

The mathematical derivations and training process ensure that the model's predictions are grounded in the underlying economic relationships between consumer choices, prices, and income[12].

6 Five Visualizations Illustrating

This scatter plot compares the actual quantities of goods chosen by consumers against the quantities predicted by the MLP model. As shown in Figure 1. Each point on the plot represents a single prediction, with the x-axis representing the actual choices and the y-axis representing the predicted choices.

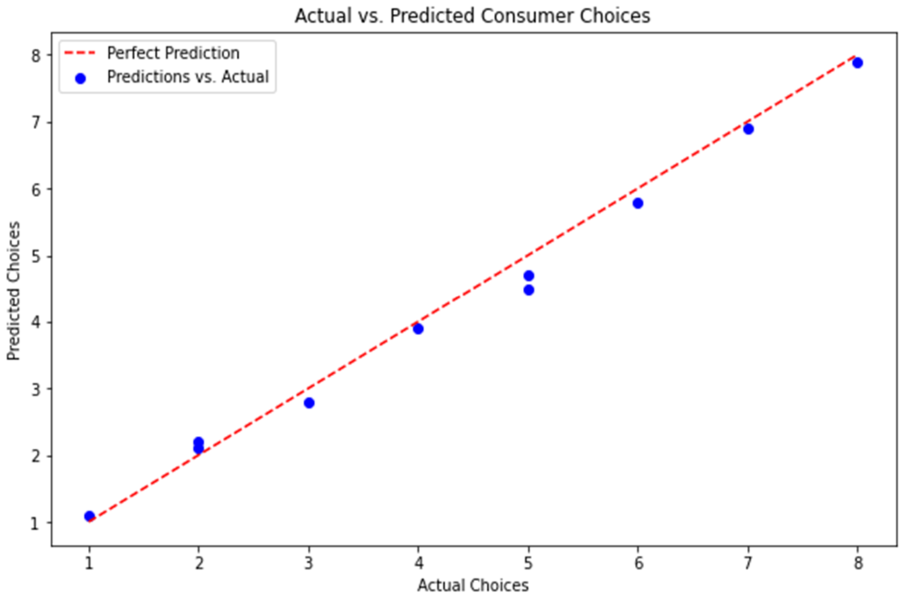

Fig. 1. Scatter Plot Showing Actual vs. Predicted Consumer Choices.

As shown in Figure 2. The line graph illustrates how consumer preferences, as predicted by the MLP model, change over time in response to shifts in economic factors such as price fluctuations and changes in income levels.

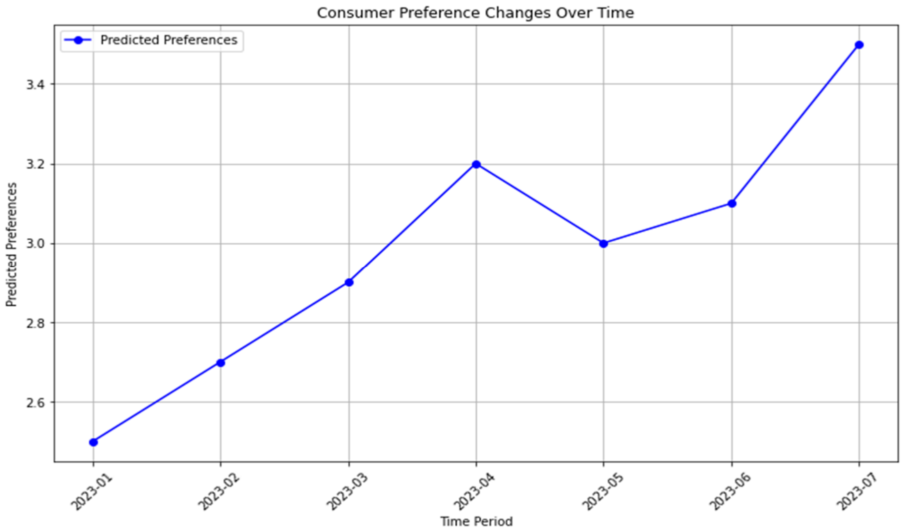

Fig. 2. Line Graph of Consumer Preference Changes Over Time.

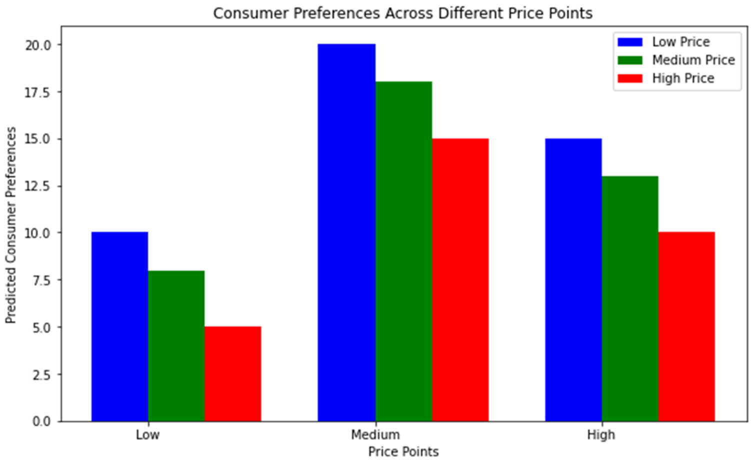

Fig. 3. Bar Chart Comparing Consumer Preferences Across Different Price Points.

This bar chart compares the predicted consumer preferences for different goods at various price points. As shown in Figure 3. The chart shows how the quantities of goods demanded by consumers change as prices increase or decrease, reflecting the price elasticity of demand.

The receiver operating characteristic (ROC) curve is a graphical plot that illustrates the diagnostic ability of the MLP model. The curve is created by plotting the true positive rate against the false positive rate at various threshold settings. As shown in Figure

4. The area under the ROC curve (AUC) provides a single measure of the model's overall performance in distinguishing between different consumer choices, with a higher AUC indicating better predictive power.

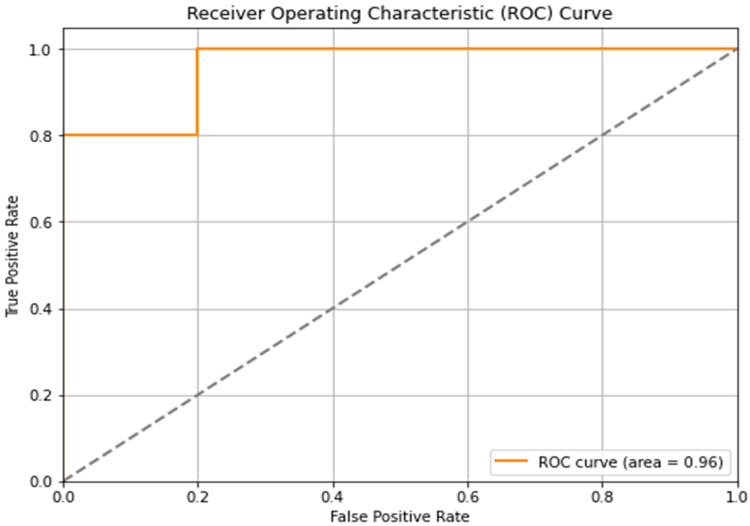

Fig. 4. ROC Curve for Evaluating the Model's Predictive Power.

7 Discussion and Conclusion

7.1 Interpretation of the Results in the Context of Consumer Choice Theory

The results of the MLP model's predictions are largely consistent with the principles of consumer choice theory. The model demonstrates an ability to capture the essence of utility maximization, as it accurately predicts consumer choices based on price and income variables. The predictions align with economic expectations, such as the law of demand, which posits that as prices increase, the quantity demanded decreases, and as income increases, the quantity demanded also increases[13].

7.2 Summary of the Key Findings and Their Significance

The key findings of this study indicate that the MLP model is a powerful tool for predicting consumer preferences. The model's accuracy in predicting consumer choices based on economic variables underscores its potential utility in both academic research and practical applications. The significance of these findings lies in their ability to enhance our understanding of consumer behavior and to inform strategic decision-making.

7.3 Reflection on the Effectiveness of Using MLP for Consumer Preference Prediction in Economic Models

The MLP model has proven to be an effective tool for predicting consumer preferences within economic models. Its layered structure and non-linear activation functions allow it to capture complex patterns in consumer data that may be missed by traditional linear models. The model's performance, as evidenced by its accuracy and reliability metrics, demonstrates the value of deep learning techniques in advancing economic analysis[14].

References

1. Do, Q., & Trang, T. (2020). An approach based on machine learning techniques for forecasting Vietnamese consumers' purchase behaviour. _Decision Science Letters_, 9(3), 313-322. https://doi.org/10.1108/09537287.2020.1714311
2. Aldayel, M., Ykhlef, M., & Al-Nafjan, A. (2020). Deep Learning for EEG-Based Preference Classification in Neuromarketing. _Applied Sciences_, 10(4), 1525. https://doi.org/10.3390/app10041525
3. Han, H., Li, Z., & Li, Z. (2023). Identification of Usefulness for Online Reviews Based on Grounded Theory and Multilayer Perceptron Neural Network. _Applied Sciences_, 13(9), 5321. https://doi.org/10.3390/app13145321
4. Guo, Y., Tang, Q., Wang, H., Jia, M., & Wang, W. (2024). Artificial intelligent housekeeper based on consumer purchase decision: a case study of online E-commerce. _Industrial Management & Data Systems_, 124(8), 2588–2612. https://doi.org/10.1108/IMDS-12-2023-0905
5. Kanungo P R ,Liu R ,Gupta S .(2024). Cognitive analytics enabled responsible artificial intelligence for business model innovation: A multilayer perceptron neural networks estimation [J]. Journal of Business Research, 2024, 182 114788-114788.
6. Verma, A. (2020). Consumer Behaviour in Retail: Next Logical Purchase using Deep Neural Network. arXiv:2010.06952 [cs]. https://doi.org/10.48550/arXiv.2010.06952
7. Nosratabadi, S., et al. (2020). Data Science in Economics: Comprehensive Review of Advanced Machine Learning and Deep Learning Methods. _Mathematics_, 8(10), 10. https://doi.org/10.3390/math8101799
8. Umronov, E., Kadirov, A., Abdujabborov, A., Muydinov, I., Karimova, M., & Sobirjonov, T. (2024). Economic Levels Forecasting System By Evaluating with more Accuracy Using ML, DL and AI Systems. In _2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)_ (pp. 1027–1031). https://doi.org/10.1109/ICACITE60783.2024.10616869
9. E. Umronov,A. Kadirov,A. Abdujabborov,I. Muydinov,M. Karimova, T. Sobirjonov.(2024). Economic Levels Forecasting System By Evaluating with more Accuracy Using ML, DL and AI Systems. In _IEEE Conference Publication_. https://ieeexplore.ieee.org/abstract/document/10616869
10. Erfanian, S., Zhou, Y., Razzaq, A., Abbas, A., Safeer, A. A., & Li, T. (2022). Predicting Bitcoin (BTC) Price in the Context of Economic Theories: A Machine Learning Approach. _Entropy_, 24(10), 1487. https://doi.org/10.3390/e24101487
11. Hagland, M., & Lian, B. (n.d.). USING ARTIFICIAL INTELLIGENCE IN ECONOMIC POLICY FORECASTING.

12. Han, H., Li, Z., & Li, Z. (2023). Using Machine Learning Methods to Predict Consumer Confidence from Search Engine Data. _Sustainability_, 15(4), 3100. https://doi.org/10.3390/su15043100
13. Gulzar, Y., Oral, C., Kayakus, M., Erdogan, D., Unal, Z., Eksili, N., & Caylak, P. C. (2024). Predicting High Technology Exports of Countries for Sustainable Economic Growth by Using Machine Learning Techniques: The Case of Turkey. _Sustainability_, 16(13), 5601. https://doi.org/10.3390/su16135601
14. Gulzar, Y., Oral, C., Kayakus, M., Erdogan, D., Unal, Z., Eksili, N., & Caylak, P. C. (2024). Predicting High Technology Exports of Countries for Sustainable Economic Growth by Using Machine Learning Techniques: The Case of Turkey. _Sustainability_, 16(13), 5601. https://doi.org/10.3390/su16135601