A RESEARCH REGARDING DESIGNING IMPORTANT CROSS THE IMPACT IN PURCHASE INCIDENCE: AN ANALYSIS OF ARTIFICIAL NEURAL NETWORK METHODS AND MULTIVARIATE PROBIT DESIGNING

[1]CHENG NANNAN, [2]Divya Midhunchakkaravarthy

Abstract

The goal of this research is to provide light on customer habits across similar product types by simulating massive cross-effects in purchase occurrence. Complex cross-category interactions are tested to see how effectively various modeling tools, such as Multivariate Probit (MVP) and Artificial Neural Networks (ANNs), capture them. In contrast to ANNs' renowned flexibility in handling big datasets and nonlinear interactions, MVP models provide a transparent statistical foundation based on economic theory. Research compares the precision of these methodologies' predictions, the lucidity of their presentations, and their capacity to grasp the subtleties of consumer behavior across different product categories. Even while ANNs do very well in terms of prediction accuracy, the results show that MVP models provide much better interpretability and theoretical consistency. The results provide useful information for marketing and decision-making while also contributing to the expansion of consumer choice modeling approaches. This research analyzes and contrasts the efficacy of ANN techniques with MVP models to evaluate large-scale cross effects on sales incidence. By comparing and analyzing two state-of-the-art approaches, this study aims to provide insight on their respective capabilities in identifying complex linkages and predicting consumer behavior in big datasets. The purpose of this research is to evaluate the efficacy of applying ANN and MVP algorithms to purchase incidence data in order to get more relevant and useful insights about consumer behavior. Focusing on model correctness, computational efficiency, and interpretability, this comparison study seeks to aid researchers doing marketing or consumer research in choosing the most suitable modeling methodologies.

Keywords: *Prevalence Of Sales, Data Interpretation, Multimodal Probit Simulation, Artificial Intelligence and Neural Networks.*

Introduction

Companies need to know what factors influence purchase incidence in today's consumer behavior analysis landscape if they want to adapt their strategy and succeed. The amount of data and the intricacy of client interactions heighten the requirement for high-level modeling tools to undertake meaningful analysis and projections. The most cutting-edge methods for investigating large-scale cross effects in purchase incidence include multivariate probit modeling and ANNs. In this research, researchers will compare and contrast these two methods. ANNs are a well-known alternative to traditional methods for understanding consumer behavior because of their ability to depict complicated data patterns and non-linear interactions. However, when dealing with several simultaneous purchase possibilities, Multivariate Probit Modelling provides a statistical framework for examining many connected binary outcomes. This study analyzes and contrasts two approaches so that readers may get a better feel for the benefits and drawbacks of each. Estimating the incidence of purchases and capturing the cross effects that influence customers' choices are the objectives. Improving the predictive power of economic and marketing research and painting a more accurate picture of customer behavior were two of the expected advantages of the study. One main goal of marketing research is to find out when and how individuals buy different kinds of things. Customers are often swayed by interactions between similar items, which are called cross-category effects, rather than making purchases individually. Precise simulations of these interactions are crucial to marketing techniques including product packaging, pricing, and promotional activities. Customer behavior analysis has traditionally made heavy use of traditional methodologies like MVP models. When it comes to accounting for linkages between purchase occurrences across categories, MVP models provide a systematic and theoretically based framework using probabilistic approaches. Their dependence on established assumptions about the connections among variables may limit their capacity to represent the complete complexity of consumer decision-making, especially when non-linear interactions or substantial cross-effects are involved. Flexible and data-driven ANNs may represent complicated interactions and very nonlinear connections rather than relying on rigid parametric assumptions. Their versatility makes them a viable option for large-scale consumer behavior research. Despite their adaptability, ANNs are often criticized for not being interpretable, which hinders their ability to provide useful insights. Because it compares and contrasts conventional econometric models with cutting-edge machine learning methods, the results of this study can help academics and marketers choose the best consumer choice research tools (Arslan et al., 2022).

Background of study

In order to understand market dynamics and create successful marketing strategies, accurate estimation of the elements influencing purchase frequency is essential. This may be achieved via consumer behavior research. Understanding the consequences of consumer decision-making is necessary for sophisticated analytical tools, since there are many interdependent factors and non-linear interactions. What researchers call a consumer's "purchase incidence" is just how often they buy something during a certain time frame (Désir et al., 2022). Although capturing these cross-effects was crucial for proper modeling, it becomes more complex when working with large-scale datasets that include several commodities and consumer characteristics. Studies with complicated datasets and many binary outcomes have traditionally made use of statistical methods like Multivariate Probit Modeling. When applied to situations involving linked purchase decisions, this method provides valuable insights into the myriads of variables that influence choices made simultaneously (Cao & Sun, 2019).

Purpose of the study

In order to explore large-scale cross effects on purchase occurrence, this research sought to investigate and evaluate two state-of-the-art modeling approaches: Multivariate Probit modeling and ANNs. In order to better understand the intricate network of factors that impact consumers' purchase choices, the researchers set out to identify the best successful ways for mining massive datasets. To better understand the strengths and weaknesses of these two approaches to consumer behavior analysis and purchase prediction, this research compares artificial neural networks with multivariate probit models. The findings have implications for both academic and industry-level studies of economics and marketing. Better models for assessing and predicting customer purchases will also emerge as a result of this.

Literature review

To get there, the researcher has tried a few different strategies. Researchers have been focusing on building sophisticated modeling ways to handle large volumes of intricate modern consumer data. A powerful method for analyzing several related binary outcomes, such as buying many things at once, is Multivariate Probit Modeling. Analyzing the interdependencies between multiple purchases is made possible by this strategy, which helps to illuminate the factors that influence buying decisions. Because of its ability to anticipate connections between binary outcomes, it became very useful for processing massive datasets with several dimensions of data and understanding intricate consumer patterns (Echenique & Saito, 2019). When artificial neural networks first emerged, they posed a serious challenge to traditional statistical models. ANNs were created to identify complicated patterns and non-linear correlations in data; they were modeled after the way the human brain functions (Chen et al., 2019). They are great for modeling intricate consumer interactions because they can learn from large datasets and adapt to novel circumstances. Recent research suggests that ANNs may outperform more traditional methods in predicting various consumer outcomes, including purchase frequency, due to their ability to represent complex and non-linear interactions. Although both ANNs and Multivariate Probit Modeling have their uses, research comparing them in the context of large-scale cross effects on purchase incidence is lacking. The research shown that when presented with highly dimensional data, ANNs outperform traditional models in terms of accuracy and adaptability. Multivariate probit modeling remains significant due to its interpretability and the insights it offers on the linkages between linked binary outcomes (Chen & Chen, 2021).

Research question

- How does price sensitivity influence on comparing artificial neural networking cadence

Research methodology

Quantitative research refers to studies that examine numerical readings of variables using one or more statistical models. The social environment may be better understood via quantitative research. Quantitative approaches are often used by academics to study problems that impact individuals. Objective data presented in a graphical format is a byproduct of quantitative research. Numbers are crucial to quantitative research and must be collected and analyzed in a systematic way. Averages, predictions, correlations, and extrapolating findings to larger groups are all possible with their help.

Research design: In order to analyse quantitative data, SPSS version 25 was used. When analysing the statistical association, the odds ratio and 95% confidence interval were used to determine its direction and size. A statistically significant threshold was suggested by the researchers at $p < 0.05$. The primary features of the data were identified by a descriptive analysis. Mathematical, numerical, or statistical evaluations using quantitative methodologies are often used for data gathered from surveys, polls, and questionnaires, or by modifying existing statistical data using computing tools.

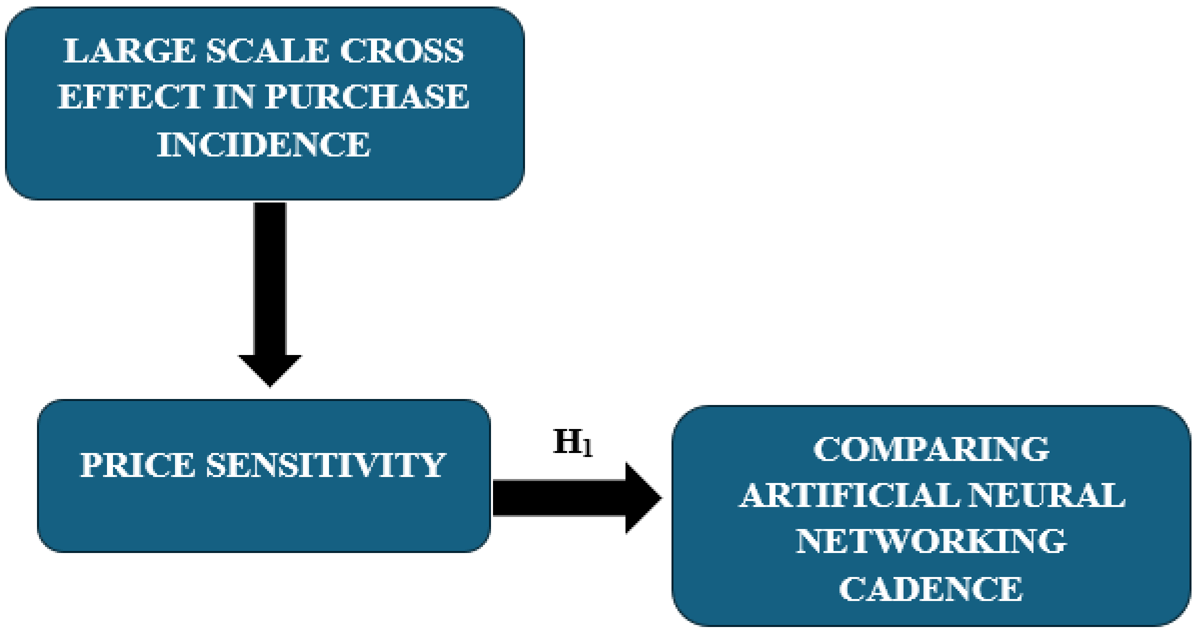
Sampling: Research participants filled out questionnaires to provide information for the research. Using the Rao-soft programme, researchers determined that there were 735 people in the research population, so researchers sent out 850 questionnaires. The researchers got 810 back, and they excluded 32 due to incompleteness, so the researchers ended up with a sample size of 778.

Data and Measurement: A questionnaire survey functioned as the primary data collection instrument for the investigation. The survey had two sections: (A) General demographic information and (B) Responses on online and non-online channel factors on a 5-point Likert scale. Secondary data was obtained from many sources, mostly on internet databases.

Statistical software: The statistical analysis was conducted using SPSS 25 and MS-Excel.

Statistical Tools: To grasp the fundamental character of the data, descriptive analysis was used. The researcher is required to analyze the data using ANOVA.

Conceptual framework



Result

- **Factor analysis**

One typical use of Factor Analysis (FA) is to verify the existence of latent components in observable data. When there are not easily observable visual or diagnostic markers, it is common practice to utilize regression coefficients to produce ratings. In FA, models are essential for success. Finding mistakes, intrusions, and obvious connections are the aims of modelling. One way to assess datasets produced by multiple regression studies is with the use of the Kaiser-Meyer-Olkin (KMO) Test. They verify that the model and sample variables are representative. According to the numbers, there is data duplication. When the proportions are less, the data is easier to understand. For KMO, the output is a number between zero and one. If the KMO value is between 0.8 and 1, then the sample size should be enough. These are the permissible boundaries, according to Kaiser: The following are the acceptance criteria set by Kaiser:

A dismal 0.050 to 0.059, subpar 0.60 to 0.69
Middle grades often range from 0.70 to 0.79.
Exhibiting a quality point score between 0.80 and 0.89.
They are astonished by the range of 0.90 to 1.00.
Table 1: KMO and Bartlett's Test for Sampling Adequacy Kaiser-Meyer-Olkin measurement: .899
The outcomes of Bartlett's test of sphericity are as follows: Approximately chi-square degrees of freedom = 190 significance = 0.000
This confirms the legitimacy of claims made just for sampling purposes. Researchers used Bartlett's Test of Sphericity to ascertain the significance of the correlation matrices. A Kaiser-Meyer-Olkin value of 0.899 indicates that the sample is sufficient. The p-value is 0.00 according

to Bartlett's sphericity test. A positive outcome from Bartlett's sphericity test indicates that the correlation matrix is not an identity matrix.

Table 10: KMO and Bartlett's

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.899
Bartlett's Test of Sphericity	**Approx. Chi-Square**	3252.968
	df	190
	Sig.	.000

The overall significance of the correlation matrices was further confirmed by using Bartlett's Test of Sphericity. A value of 0.899 was the Kaiser-Meyer-Olkin sampling adequacy. By using Bartlett's sphericity test, researchers found a p-value of 0.00. A significant test result from Bartlett's sphericity test demonstrated that the correlation matrix was not a correlation matrix.

Test For Hypothesis:

- ❖ **INDEPENDENT VARIABLE**
  - ➢ **Large-Scale Cross Effect on Purchase Incidence**

You have a lot of leeway to affect related or comparable products that a buyer buys. This type of conversation might develop in a variety of ways, depending on the details. There is a plethora of add-ons that can be purchased for most phones, including cases and headphones. The researcher may want to consider getting a fancy coffee maker if they want to elevate their coffee-making game. One of the several times this is required is when advertising two separate products simultaneously. It is possible that more broad-based tendencies and patterns were involved. For instance, if a recently established health movement were to get significant popularity, the demand for organic food and exercise gear would see a spectacular surge. Customers' actions in one area may have repercussions on their buying patterns in other areas. This was one of its subsequent outcomes (Chen et al., 2022).

- ❖ **FACTOR**
- ➢ **Price Sensitivity**

The term "price sensitivity" describes how much customers' buying habits are affected by small changes in product or service prices. It shows that consumers are sensitive to changes in product prices and are prepared to pay that amount. Small changes in price may have a big impact on demand when price sensitivity is strong, since people would look for cheaper alternatives or cut down on their purchases altogether. When customers have low price sensitivity, they are less impacted when prices change. This might be because they are very loyal to a certain brand, there aren't any alternatives, or they think the product is very valuable. Several variables impact price sensitivity, such as disposable income, product requirement, availability of alternatives, and competitive pricing. In order to maximize profits, create successful promotions, and satisfy customers, businesses analyze sales data, conduct market tests, and conduct surveys to determine how price sensitive their target audience is. When trying to maximize profits without alienating customers or falling behind the competition, knowing how price sensitivity works is essential (Cheung et al., 2019).

- **DEPENDENT VARIABLE**
- **Comparing Artificial Neural Networking Cadence**

A variety of artificial neural network topologies were used to tackle various sorts of issues. How they handle learning rates, training durations, generalizability, and task suitability may give researchers a sense of their efficiency, or cadence. When the data links aren't sequential, algorithms like FNNs come into play since they are easily trainable. However, its lack of complexity could be an issue when working with elaborate designs. Traditional CNNs failed miserably when forced to deal with spatial data-intensive applications like image recognition. The hierarchical features and patterns in images are recognized by means of convolutional layers. Convolutional neural networks provide great results on visual tasks, but they take a long time to train and infer since they are computationally and resource intensive. Long Short-Term Memory Networks (LSTMs) and Recurrent RNNs both originated with the intention of processing sequential input. Language modeling and time-series prediction are two examples of jobs where they shine: keeping an internal state while analyzing sequences. LSTM models aim to compensate for RNNs' slower training speed and potential difficulties with long-term dependencies caused by their complex design (Cho et al., 2024).

- **Relationship between Price sensitivity and Comparing Artificial Neural Networking Cadence**

Understanding how fast and efficiently firms can react to what consumers are ready to pay is the relationship between price sensitivity and comparing the cadence of ANNs. The degree to which consumers are affected by changes in pricing is known as price sensitivity. Some consumers may seek less expensive alternatives in the event of a price increase, while others may be willing to pay a premium for a product or service they really love. The frequency with which an intelligent system (such as an AI model) updates itself to improve its prediction or decision-making abilities is referred to as ANN cadence. Companies may react instantly to market changes by using ANNs with a high cadence, which means the system updates often or in real-time. If a product is selling well, the system may decide to slightly increase the price; conversely, if sales are poor, it may decide to provide a discount in the hopes of attracting more consumers. Businesses may remain competitive and satisfy price-sensitive clients with this form of dynamic pricing. The converse is also true: ANNs with slow update cadences risk missing these rapid changes, which may cause pricing to seem out of line with what consumers anticipate. Businesses may discover the optimal ANN cadence by evaluating several options; this is the frequency with which the system should update pricing so that they remain fair, competitive, and in line with what consumers are prepared to pay. Maintaining a successful company while making sure consumers feel appreciated is like finding the perfect rhythm to dance to the ever-changing beat of the market. Businesses may maintain their agility, responsiveness, and customer-focus with the correct cadence, especially in this price-sensitive era (Derakhshan et al., 2022).

Analysis of the above discussion, the researcher formulated the following hypothesis, which was to analyse the relationship between Price sensitivity and Comparing Artificial Neural Networking Cadence.

***H_{01}*: There is no significant relationship between Price Sensitivity and Comparing Artificial Neural Networking Cadence.**

***H_1*: There is a significant relationship between Price Sensitivity and Comparing Artificial Neural Networking Cadence.**

Table 2: H_1 ANOVA Test

ANOVA					
Sum					
	Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Between Groups	39588.620	166	6635.417	1516.320	.000
Within Groups	492.770	611	4.376		
Total	40081.390	777			

In this study, the result was significant. The value of F was 1516.320, which reaches significance with a p-value of .000 (which is less than the .05 alpha level). This means the ***"H_1: There is a significant relationship between Price Sensitivity and Comparing Artificial Neural Networking Cadence."*** was accepted and the null hypothesis was rejected.

Discussion

The researcher may examine the similarities and differences between ANNs and Multivariate Probit Modeling while investigating the issue of simulating large-scale cross interactions in purchase incidence. Their ability to capture and comprehend consumer behavior is dependent on the relative strengths of the two methods. ANNs stand out from other models due to their ability to simulate complex, non-linear interactions in datasets with a high degree of dimensionality. Thanks to their flexibility, they may still uncover intricate patterns and connections, even when traditional statistical methods have failed. Because of their flexibility, ANNs shine in situations with complicated interdependencies, as those involving several connected purchase decisions. Since ANN performance is very sensitive to hyperparameter tweaking, data amount, and data quality, it may be more difficult to implement and understand ANN. When it came to understanding the relationships between several binary outcomes, however, Multivariate Probit Modeling offered a more systematic approach. By using this approach, researcher was able to better understand the statistical correlations at work and the connection between quick purchasing decisions. One advantage is that it can handle cases where the dependent variables were organically related; another is that it produces conclusions that are simple to comprehend and absorb. Many intricate and non-linear correlations may go unnoticed by interdisciplinary probit models when dealing with large-scale data sets.

Conclusion

The research discovered that both Artificial Neural Networks and Multivariate Probit Modeling perform equally well for analyzing large-scale cross effects on purchase incidence. Before settling on a strategy, several factors had to be considered, including the significance of interpretability and the need for precise prediction. In order to better understand and predict consumer behavior, it was proposed that these tactics be studied in more depth in future research, with an emphasis on hybrid models and powerful computational tools.

References

1. Arslan, A., R. Easley, R. Wang, and O. Yilmaz. (2022). "DataDriven Sports Ticket Pricing for Multiple Sales Channels with Heterogeneous Customers". Manufacturing & Service Operations Management. 24(2): 1241–1260.
2. Cao, J. and W. Sun. (2019). "Dynamic learning of sequential choice bandit problem under marketing fatigue". In: Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. 3264–3271.
3. Chen, N. and Y. J. Chen. (2021). "Duopoly competition with network effects in discrete choice models". Operations Research. 69(2): 545–559.
4. Chen, N., G. Gallego, and Z. Tang. (2019). "The Use of Binary Choice Forests to Model and Estimate Discrete Choices". arXiv preprint arXiv:1908.01109.
5. Chen, X., J. Li, M. Li, T. Zhao, and Y. Zhou. (2022). "Assortment optimization under the multivariate MNL model". arXiv preprint arXiv:2209.15220.
6. Cheung, W. C., V. Tan, and Z. Zhong. (2019). "A thompson sampling algorithm for cascading bandits". In: The 22nd International Conference on Artificial Intelligence and Statistics. 438–447.
7. Cho, S., M. Ferguson, P. Pekgun, and J. Im. (2024). "Robust demand estimation with customer choice-based models for sales transaction data". Production and Operations Management
8. Derakhshan, M., N. Golrezaei, V. Manshadi, and V. Mirrokni. (2022). "Product ranking on online platforms". Management Science. 68(6): 4024–4041.
9. Désir, A., V. Goyal, and J. Zhang. (2022). "Technical Note—Capacitated Assortment Optimization: Hardness and Approximation". Operations Research. 70(2): 893–904.
10. Echenique, F. and K. Saito. (2019). "General Luce Model". Economic Theory. 68(4): 811–826.