Information Management and Business Review (ISSN 2220-3796)
Vol. 16, No. 3(S), pp. 844-854, Sep 2024

Gender-Based Analysis of Online Shopping Patterns on Shopee in Malaysia: A J48 Decision Tree Approach

Nurul Ain Mustakim, Zatul Himmah Abdul Karim*, Muna Kameelah Sauid, Noorzalyla Mokhtar, Zuhairah Hassan, Nur Hazwani Mohamad Roseli
Faculty of Business & Management, Universiti Teknologi MARA, Cawangan Melaka Kampus Bandaraya Melaka, Melaka, Malaysia
ainmustakim@uitm.edu.my, *zatul446@uitm.edu.my, munasauid@uitm.edu.my, noorza2904@uitm.edu.my , zuhairah612@uitm.edu.my , nurha5338@uitm.edu.my
Corresponding Author: Zatul Himmah Abdul Karim

Abstract: The purpose of this study is to investigate the gender differences of the Shopee platform for online shopping behavior by using the J48 decision tree algorithm to classify and predict shopping frequency among male and female consumers in the Malaysian context. WEKA software was used in this study to analyze the datasets. From the experiments, the majority of Shopee users were female consumers. The findings show that female consumer behavior is more complicated and more varied regarding purchasing behavior. The study's findings demonstrate the potential of gender-specific insights to enhance e-commerce strategies, particularly in product recommendations and targeted marketing. Although the J48 model performed well in predicting male shopping patterns, it was less effective for females, indicating the need for more advanced modeling techniques to better capture the complexities of female consumer behavior. This research also emphasizes the significance of using machine learning tools like the J48 decision tree to analyze consumer data, providing valuable insights for improving customer satisfaction and business performance. However, limitations such as sample size and the focus on a single platform suggest that further research is needed, including the exploration of alternative algorithms and broader demographic factors.

Keywords: *Gender, Shopping Patterns, E-Commerce, Decision Tree, J48, Shopee*

1. Introduction

Starting in 2004, Malaysia's government took the first step towards e-commerce with local websites and attempts to introduce people with the help of promotional activity. Major retailers such as Lazada then helped rapid increases in the e-commerce industry. Among the many other available e-commerce websites, Shopee has become the most widely used in Southeast Asia, including Malaysia. Wong et al. (2023) Express that in their study, though Lazada was created in 2012, Shopee has grown to become the go-to platform for Malaysian consumers who intend to shop online. This is because this platform wants products to be shopped online easily, and effortlessly for buyers and sellers.

Nowadays, understanding consumer behavior in e-commerce is important for business. This is because most e-commerce platforms provide a unique environment where consumers can access numerous products and services simply and immediately without leaving the comfort of their homes. Align with the statement by Arif et al (2021) The trend of online shopping is growing with time as digital technology and the internet take over our lives. Along with the growth of the digital market, the amount of data created through online purchases, customer interactions, and other e-commerce activities has also increased dramatically. In this case, businesses must understand consumer preferences, emerging purchasing trends, and factors influencing purchasing decisions. By knowing these purchasing trends, businesses can improve their marketing strategy, consumer experiences, and overall performance. However, the amount of data generated by the e-commerce platform, especially Shopee, is enormous and complex. It consists of a variety of information such as purchase history, demographics, consumer preferences, and others. Moreover, it provides opportunities for companies to collect, analyze, and use data to gain valuable insight. This is attributed to the fact that growth in e-commerce has been rapid, hence making it important for businesses to keep themselves updated for them to maintain some level of competitiveness. Hence, in analyzing the data, especially the complex purchasing patterns, innovative technologies such as machine learning or any other predictive algorithm can be employed by businesses.

With innovative technologies today, machine learning approaches can effectively predict the characteristic customer purchasing behavior. Machine learning is aimed at developing algorithms that could be learned and improved over time and applied to a prediction context. Liu et al. (2020) say that machine learning is a method that is quick, accurate, and very advanced. Furthermore, Choudhury & Nur (2019) also applied machine learning in their study for the classification of probable customers based on their past buying habits. In addition, understanding consumer purchasing trends is important so that one can align offerings to meet the needs of the consumers at a better level. By finding trends and preferences within users' consumption, businesses can work toward better inventory management, personalization of marketing campaigns, and enhancement of user experiences. Also, the insights into purchasing behavior on gender help predict future trends to proactively strategize business. This analysis not only creates value for businesses through increasing sales and consumers but also increases consumer satisfaction by offering products and services that cater to the preferences of female and male consumers.

This paper identifies the frequency of online shopping on the Shopee platform and the gender factors that determine the frequency of shopping among consumers in Malaysia. Research has continually shown that, regarding the frequency of buying goods online, there is a difference between genders, with different factors influencing the behavior, such as loyalty, discount, and general shopping patterns. Filipas et al. (2023). Females and males present different preferences when shopping online. It also shows that, while purchasing something, males consider fewer factors than females. Rianto et al. (2017). Furthermore, Malaysian findings report that, in terms of frequency of usage of mobile platforms for online shopping, education, and traveling, females use them more than males, while males are interested in downloading, gaming, and other online leisure activities. (Kovacevic & Kascelan, 2020). This also agrees with the local consumer behavior whereby Shopee has gained wider support because its mobile-compatible platform is user-friendly and easy to use by both genders. On the other hand, it is equally important to find out about the gender-based differences in online shopping experiences within Malaysia's emergent e-commerce market and diverse consumer base. In this regard, the local players will profit from how to design their marketing strategy according to the gender preferred to increase the satisfaction of consumers and thereby increase the sales record in the direction required. Whereas various research works have been done to investigate the differences between genders in online shopping behavior, few of them have utilized the J48 decision tree algorithm in analyzing or predicting the said behavior. By applying this machine learning technique, this present study fills a very important gap by enabling a localized understanding of gender-based online shopping behavior on the Shopee platform in Malaysia.

2. Literature Review

Consumer Behavior Models in E-Commerce

Employing various theoretical frameworks to analyze and predict purchasing patterns to understand consumer behavior in online purchasing, has been a significant area of research nowadays. (Arora & Sahney, 2018) Deploy an integrated technology acceptance model (TAM) and theory of planned behavior (TPB) framework to examine showrooming behavior. In this study, they found two factors significantly influence consumer decisions which are the perceived ease of online purchasing and the usefulness of showrooming. Similarly, the study by Ha et al. (2019) Used an integrated TAM and TPB models to examine online purchase intention. This study reveals that perceived usefulness, attitude, ease of use, subjective norm, and trust positively influence shopping intention.

To understand consumer behavior in specific situations, an extended TAM model has been applied. Study by (Sun et al. (2009) Two elements have been added to the extended TAM model which are trust and resources in mobile commerce adoption. From this study, the finding shows that perceived credibility and self-confidence are the main factors that influence. It can be concluded that using an extended TAM model is very relevant in understanding the adoption of mobile commerce. Different study by Mou & Benyoucef (2021), they employed different theoretical frameworks such as TAM, TPB, and the Theory of Acceptance and Use of Technology (UTAUT). In this study, they identified that UTAUT and TAM models have a strong correlation with each other. It shows that mutually of this theory is very beneficial in recognizing how people perform in social commerce.

Gender Differences in Consumer Behavior

Study by Hasan (2010) Identified there is a difference between females and males especially in shopping perspective such as in terms of thinking, feeling, and shopping attitude. In this study, the data were collected from 80 students from Midwestern University who enrolled in an e-commerce course. The experiments reveal that thinking attitude is the most important factors that give differences between females and males, and it shows that females think online shopping is less useful compared to males. A different study done by Akhlaq & Ahmed (2016) Which focuses on Pakistan to find the relationship between gender and online shopping factors. This study was distributed through an online survey via email to students, alumni, and local university in Karachi, Pakistan. The study found that a big gap had arisen between females and males through perceived enjoyment (PE) and legal framework (LF) online shopping factors. The results show that Eastern females are more enjoyable during online shopping because they can have fun shopping in a secure virtual environment.

In addition, a study from Indonesian by Pradhana & Sastiono (2019) Focusing on finding out how different genders affect consumer choice during they shop online. From their experiments, they found out that females are more towards online shopping compared to males, but in terms of total spending, males are more than females, this study also reveals that trust is the main factors that influence males when they decide how much money they spend and how frequently they shop online. It can be concluded that different genders can be influenced by online shopping factors.

Moreover, a study by An et al. (2022) Emphasize predicting gender for the system that can make predictions based on online shopping behavior specifically for Vietnam FPT groups. The system testing, reveals that the system can predict gender with almost a 78% success rate, it demonstrates that based on online shopping behavior data can be used to figure out gender. However, this study has a hole because it is only based on prediction accuracy rather than gender-specific shopping patterns or preferences. Therefore, this study provides valuable insight into the differences between females and males in online shopping behavior, and it can help to tailor their marketing strategies more effectively to each gender, especially in the Shopee platform.

Data Mining in E-Commerce

In e-commerce, it involves a massive volume of data which this data is collected from various sources. Utilization of data mining or big data is the best way to handle the data. Data mining is the procedure of determining patterns and knowledge from large amounts of data. To enhance an e-commerce strategy on consumer behavior and help optimize business operations utilizing data mining is very important. A study by Arif et al., (2021) Investigates the challenges and advantages of utilization mining on e-commerce data. The study highlights the need to design e-commerce sites for better data collection and analysis. While data mining can reveal useful customer trends and preferences, it also comes with challenges that need to be addressed to fully benefit from it.

Each of the businesses can use data mining to increase their market to consumers and improve their prediction capabilities. Both of the studies by Asniar and Surendro (2019) and Alghanam et al. (2022), agreed that using data mining techniques in business can help make the prediction more reliable and accurate. Similarly study by Li et al. (2019), shows that using decision trees and cluster analysis, both of these categories under machine learning methods, would help businesses better understand consumer behavior and preferences. Additionally, Moon et al. (2021) Apply four types of machine learning methods such as Apriori, Decision Tree, Random Forest Tree, and Naïve Bayes to find an effective algorithm to analyze consumer satisfaction in e-commerce. In this study, it used a large dataset that included 40,000 records. After the experiments have been done, the results show that the Apriori algorithm is the best among the other machine learning methods, with an accuracy of 88%. From the several experiments and findings from this study, it can be concluded that selecting the right data mining techniques will help companies easily understand consumer trends to make a proper marketing plan and product recommendation. Similarly study by Chaubey et al. (2022) Conducted a comparison of numerous machine learning algorithms to predict consumer purchase behavior. The study examined methods such as decision trees, KNN, SGD, AdaBoost, and hybrid algorithms. The findings show that using advanced and hybrid data mining techniques can improve predicting consumer behavior, especially in an e-commerce context. This will guide the companies leading to more effective marketing strategies.

J48 Decision Tree Algorithm

The J48 decision tree algorithm is commonly used in data mining, especially in predicting consumer behavior. It uses accuracy, and efficiency are examined in the literature below from different contexts. Study by Alghanam et al. (2022) In the context of e-commerce recommended J48 algorithm is a data mining model to predict consumer purchase behavior. They used the Northwind data set, and the finding reveals that 95.2% accuracy was achieved in this study. It shows that using the J48 algorithm also can predict consumer behavior.

Additionally, different contexts of study by Afolabi et al. (2019), for predicting consumer behavior, in this study, they apply integration between structured and unstructured data. They found that prediction based on structured data is more accurate than unstructured data. It shows that the J48 algorithm can work with different types of data and can be used as a tool for predicting consumer behavior. At the same time Kolahkaj & Madjid Khalilian (2015) The Study uses hybrid methods which are J48 and frequent pattern mining. The goal is to improve the recommendation system. From the experiment, the result shows that the J48 algorithm has the highest accuracy and recall. After frequent pattern mining is added into J48, it not only makes the system recommendation better but also contributes to the easy-to-handle big dataset. Meanwhile, the study by Brunello et al. (2019) Employed the new version of the J48 algorithm, known as the J48 security system. Both the sequential and temporal series of data are things that this new version is intended to handle. It is not only advantageous to use J48SS to categorize the data, but it also makes the process of compiling the data much easier. The result is thus an indication that this new version of the J48 method would be acceptable and appropriate for use in industry.

3. Methodology

This study used questionnaires to understand the purchase behavior of the client. Customers who have purchased Shopee are randomly selected for the survey. A total of 156 respondents were recorded, and each was allowed to answer questions that shadowed their buying habits through an online poll. The survey was administered using an online format backed by a Google form. For your convenience, a description of the data is provided in detail below.

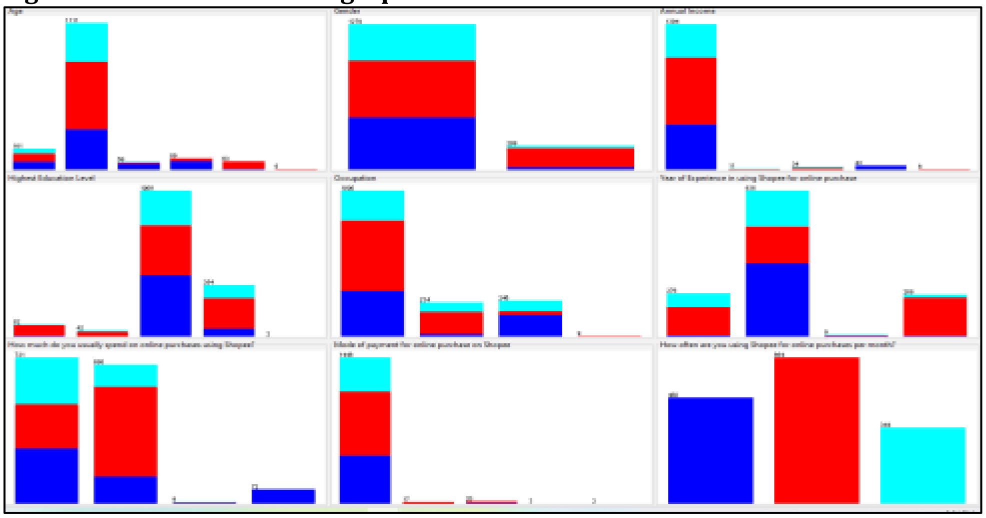
Dataset Description

The dataset used in this study was built with 9 attributes with 156 instances. These attributes cover the demographic and behavioral characteristics of the consumers. Table 1 below shows the dataset description of each attribute. Meanwhile, Figure 1 displays an overview of the demographic and behavior attributes. It shows deeper insights into the consumer demographic and behavior patterns of Malaysian consumers on the Shopee platform.

Table 1: Dataset Description

No	Attribute	Description
1	Age	Consumer's age
2	Gender	Consumer's gender
3	Annual Income	Consumer's annual income
4	Highest Education Level	Consumer's education
5	Occupation	Consumer's occupation
6	Year of experience in using Shopee for online purchase	Consumer's year experiences
7	How much do you usually spend on online purchases using Shopee	Consumer's spending
8	Mode of payment for online purchases on Shopee	Consumer's mode of payment
9	How often are you using Shopee for online purchases per month	Consumer's shopping frequency

Figure 1: Overview of Demographics and Behaviour Attributes



Pre-Processing

Preprocessing is the most crucial phase in data mining. During this phase, pre-processing transforms the data into an appropriate format, with which data mining and machine learning become much easier and smoother. It consists of a few steps required for cleaning the data by removing the missing value if any, followed by a selection of the attribute as per requirement of the study. In this work, the Synthetic Minority Over-sampling Technique is applied because this study possesses class imbalances in frequency attributes. After applying the SMOTE, the number of instances in the dataset has grown from 156 to 1488.

Data Modelling

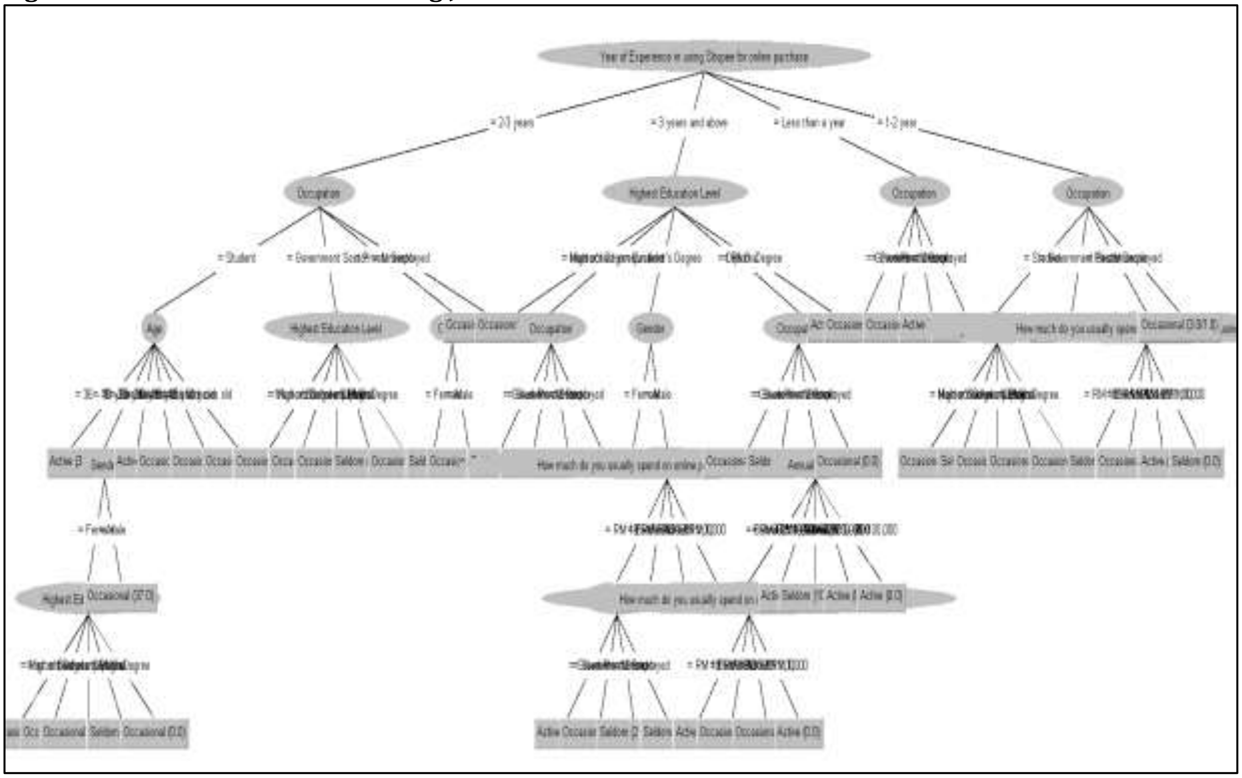
The data is then divided into training and test sets after cleaning. In this paper, for the evaluation of the performance of the model, a 10-fold cross-validation has been implemented. This process is a common approach in machine learning for the model to work well on unseen data. In the modelling stage, the J48 Decision Tree algorithm was applied to the dataset. The reason the J48 algorithm was selected in this study is because of its effectiveness in producing predictions, especially in the prediction of consumer behavior focusing on e-commerce.

Decision Tree

This study employs a classifier based on the decision tree. The reason why this study picks the decision tree is that it is the most common and simplest classification algorithm to be learned and implemented. The tree is made up of three components: root node, branch (edge or link), and leaf node. The root represents the test condition for different attributes, the branch represents all possible outcomes that can be there in the test, and the leaf nodes contain the label of the class to which it belongs.

The root node is at the start of the tree which is also called the top of the tree. It has been used in many studies, as an example by Abana (2019) Use the decision tree to foresee the behavior of consumer purchases since it makes it easy to see how logical rules work. As stated by Safarkhani & Moro (2021), the decision tree is a model that can be read and explained by humans. This research uses J48 as the model-building process. J48 is the open-source Java implementation in WEKA, and J48 is the C4.5 algorithm. The generated decision tree for the J48 algorithm is represented in Figure 2. This decision tree has 59 leaves and a tree size of 78.

Figure 2: Decision Tree Result Using J48



4. Experiments and Results

The J48 decision tree algorithm was also used for the analysis and performance evaluation of each machine learning classifier based on its results. Performance was evaluated in terms of accuracy, precision, recall, and F1-measure. In further analysis, the dataset was split into two subsets based on gender.

Exploratory Data Analysis

EDA is an approach that summarizes and analyzes the main characteristics of a dataset. The result from EDA will be plotted into some form of visualization, like a histogram or graph, which will be easy to understand. The goal of this analysis is to understand gender-specific patterns in online shopping behavior on the Shopee platform. The considerations are the frequency of shopping, modes of payment, and other behaviors that have differed between male and female shoppers. This can consequently help in improving the predictive models and provide insight into how e-commerce strategies can best be optimized on the Shopee platform to meet the needs of male and female consumers. The data for the above analysis is extracted from the survey among Malaysian consumers. This dataset includes demographic variables and behavior of consumer behavior on the Shopee platform. Descriptive analyses were conducted to realize the differences between male and female consumers of the Shopee platform.

As indicated from Figure 3 to Figure 6, the categories of frequency using the Shopee platform for purchases include active, occasional, and seldom. The results indicate that most of the consumers, 66.4%, use the Shopee platform occasionally to make purchases. This is followed by active consumers, who represent 48%, and those who shop seldom amount to 34.4%. Meanwhile, figure 4 shows gender distribution, reflecting the main objective of finding the gender difference between females and males. It highlights that most consumers who make their purchases through the Shopee platform are females. They contribute approximately 85.95% of the total consumers while males contribute about 14.05%.

According to Figure 5, for the female buyers in Shopee, the number of consumers who shop on an occasional basis is greater with 480 out of 1279. The number of active consumers is 384, while the ones who shop seldomly are 415 each month. Thus, it can be identified that most of the female buyers are occasional shoppers of the Shopee online application. Figure 6 also depicts the analysis of male consumers on the Shopee platform. Out of 209 consumers, the highest 92 shop occasionally each month, while 56 shop seldom, and only 61 are active users. This therefore suggests that most male consumers use Shopee on an occasional basis, like the pattern observed in female consumers.

Figure 3: Shopee Categories on Consumer Usage Frequency

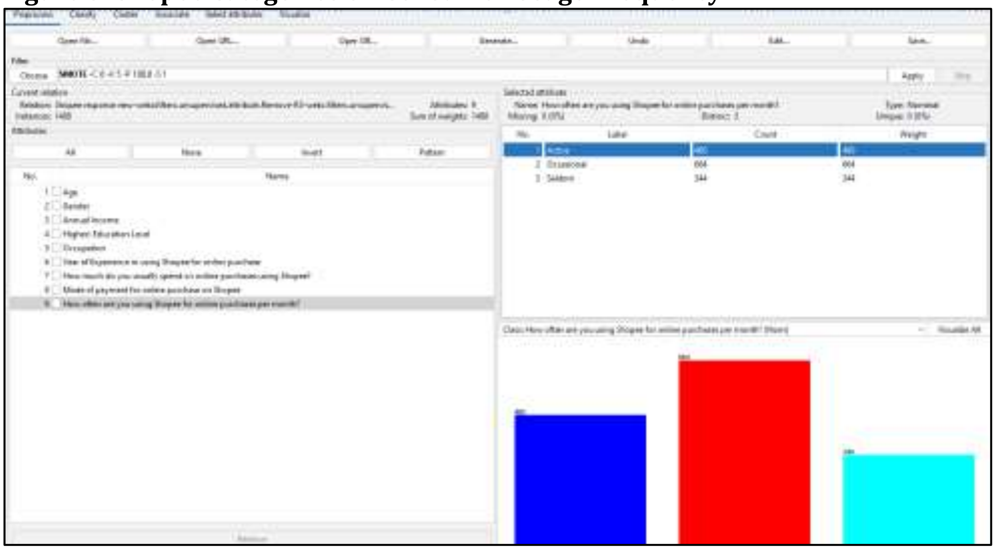


Figure 4: Shopee Usage Frequency by Gender

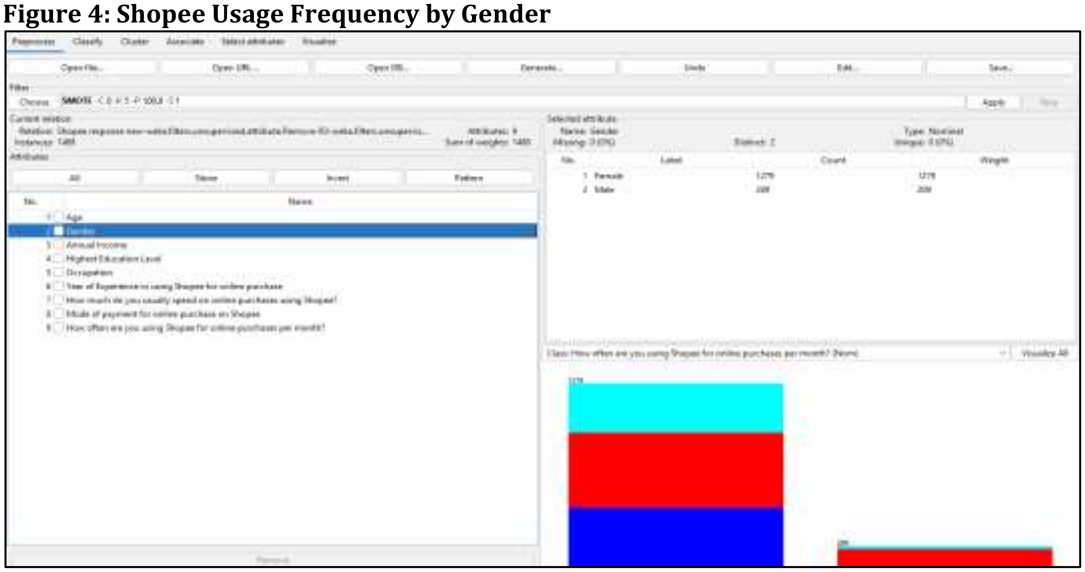


Figure 5: Shopee Usage Frequency for Female

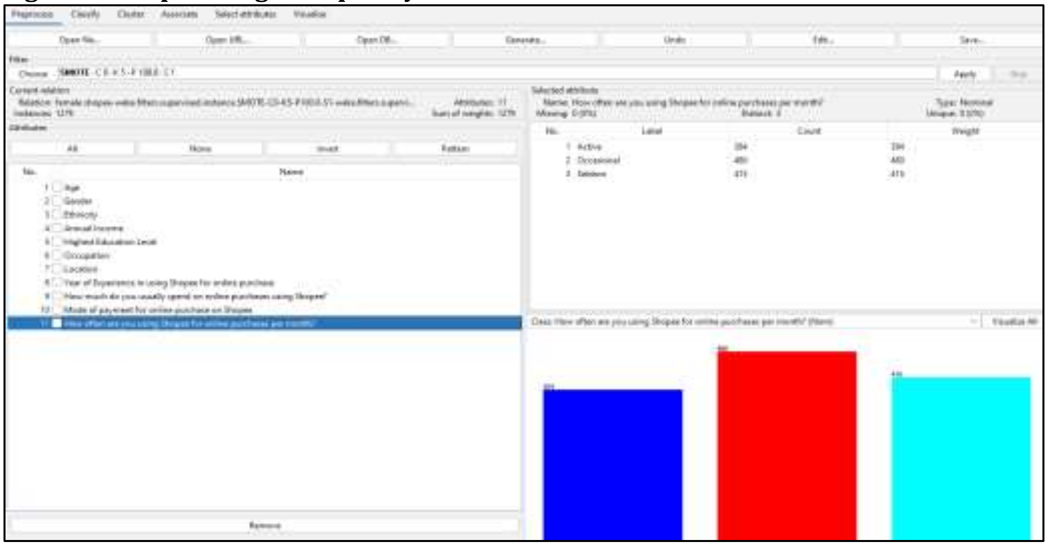


Figure 6: Shopee Usage Frequency for Male

Performance Measurement

This section explains the performance measurement of the J48 decision tree model in terms of accuracy, precision, recall, and F1 measure. The model's evaluation uses cross-validation, which was utilized for both training and testing to ensure the strength and reliability of the performance metrics. Accuracy in the J48 decision tree model is an evaluation metric that measures the number of correct predictions made by a model about the total number of predictions being made. The calculation is performed by dividing the number of correct predictions by the total number of predictions made. Meanwhile precision is a metric that measures the proportion of positive class predictions that correctly belong to the positive class. A higher precision score indicates that the model exhibits greater accuracy and reliability in correctly identifying positive samples. Recall is a performance metric in classification. It is a performance metric that scores the ability of the model presented to it to retrieve its positive samples from all the actual positive samples in that data set. Recall is the ratio of the number of true positive instances the model has correctly classified as positive to the number of instances that are considered positive. Finally, combining precision and recall into a single important performance metric f-measure gives a balanced evaluation of how well a classification model performs. This is extremely important when the data set is highly imbalanced and the classes are very badly skewed, as it is then not a great choice to rely on any kind of accuracy measuring technique.

Table 2 summarizes the classification results of the J48 decision tree, divided into three datasets: the general dataset, the female dataset, and the male dataset. The accuracy for men is 91.38%, and for women, it is 70.68%. It can be seen from the results that each dataset performed differently in the J48 decision tree algorithm. The highest performance was recorded in the male dataset to depict that the algorithm had been more effective in the prediction of male shopping behavior on the Shopee platform compared to the prediction of female shopping behavior. Lower accuracy in the female dataset may refer to greater complexity or diversity in women's shopping habits for which another approach is needed to obtain better results. However, the general dataset that included both genders was moderately accurate, reflecting the mix of male and female shopping behaviors. This implies that separate models for females and males will yield more accuracy in the results while predicting shopping behavior.

Table 2: J48 Decision Tree Performance by Dataset

Algorithm	Dataset	Accuracy	Precision	Recall	F-Measure
	General	76.41%	0.784	0.716	0.748
J48	Female	70.68 %	0.751	0.712	0.731
	Male	91.38 %	0.913	0.912	0.912

Discussion

This paper, therefore, classifies the frequency of shopping online through the Shopee platform and determines the significant gap in online shopping behavior between female and male consumers. Based on the findings, it can be understood that, among the users of this study, the majority are female consumers, and most of the users are those who shop occasionally. It follows that, while the number of male consumers is fewer, their shopping behavior remains more consistent. The findings also testify to the fact that the J48 decision tree model is better at predicting male shopping patterns. The fact allows one to conclude that male purchasing patterns on Shopee platforms are more straightforward and easily predictable, whereas, for females, it's complex and varied. There is a large variance in the J48 decision tree algorithm performance among the three datasets. Specifically, it built the model with the highest performance in terms of accuracy when applied to predict male consumer behavior. This may be because male consumers have relatively homogeneous shopping preferences, thereby making it easier for algorithms to make more accurate predictions. On the other hand, the low accuracy of the female dataset may indicate greater diversity in shopping behaviors, where women consumers using the Shopee platform consider a broader range of factors when deciding which item to buy. Thus, because female consumers are so complex, enhancing the prediction accuracy requires coming up with a more advanced modeling technique or separate models.

These findings consequently bear an applied value for e-commerce platforms, such as Shopee, which may apply the gained gender-specific insights to improve their marketing activities and recommend products more efficiently. The knowledge gained concerning the differences in shopping behavior by male and female customers allows businesses to adapt their offerings better, possibly leading to higher customer satisfaction and improved sales performance. For example, targeted marketing campaigns might be more effective in entertaining the predictable patterns of male consumers, while in the case of females, more personalized strategies may be necessary to handle their diverse preferences. Despite these promising findings, some limitations of the present study need to be acknowledged. Although the sample size in this research is adequate, it cannot be representative of the population of Shopee users. The research covers only the Shopee platform; therefore, the generalization of findings may not be necessitated on every e-commerce platform. It is suggested that future studies investigate other machine learning algorithms that might be sensitive to capturing the complexity of the shopping behavior of females. Extension can also be made to other demographic variables, including age and income, for a holistic understanding of the online shopping pattern.

5. Conclusion and Limitations

These results are useful for e-commerce websites such as Shopee, where the knowledge will help suggest better products and set up better marketing strategies, both of which can tap into gender-specific information. These findings have contributed significantly to the comprehension of online shopping behavior by underlining clear gender differences in buying trends, especially within the Malaysian context. Value is added by the research study, as it makes use of a machine learning model J48 decision tree to analyze consumer behavior, hence providing pragmatic insights useful in the making of effective marketing strategies. Knowledge acquired will be able to help businesses such as Shopee develop focused marketing strategies to achieve customer satisfaction and business growth. From a policy standpoint, this research study supports the notion of tailored marketing strategies to be based on gender-specific behaviors. These can, therefore, be useful to policymakers in businesses in advocating for gender-sensitive e-commerce policies that ensure inclusiveness in meeting diverse consumers' needs. This could imply that government agencies work out policies with e-commerce platforms that ensure marketing campaigns are targeted more toward serving varied groups of consumers. Besides, policies can be instituted that will incentivize local businesses to adopt advanced data analytics and machine learning models since this is the only way they could remain competitive in the rapid sea changes of the digital economy.

There are a few limitations regarding this research study. First, this is related to the sample size. Though the sample size is adequate in this research design, it may not represent the consumer base of Shopee, particularly in the Malaysian context. Because of this, it focuses only on the Shopee platform. Its results may therefore not apply to other e-commerce platforms like Lazada, eBay, and others. Future studies will research other machine learning algorithms to understand such complicated female shopping behaviors. Furthermore, extending other demographic aspects, for example, age and income, can also complement a study in the aspects of online shopping pattern understanding. The study brings out amply the differences in consumer behavior between the genders on Shopee and identifies areas that most definitely will help businesses in the improvement of marketing strategies through predictive analytics. E-commerce platforms will be able to handle the complexity of female shopping patterns by leveraging the more direct male behaviors to better position their offerings for improved customer experiences and better business outcomes. These findings set up a concrete foundation for further research and policy efforts that would help implement more inclusive and data-driven marketing strategies in e-commerce.

Acknowledgment: The authors extend their sincere appreciation to Universiti Teknologi MARA, Cawangan Melaka, for the invaluable support provided in the completion of this article.

References

Abana, E. C. (2019). A decision tree approach for predicting student grades in a Research Project using Weka. *International Journal of Advanced Computer Science and Applications*, 10(7), 285–289. https://doi.org/10.14569/ijacsa.2019.0100739

Afolabi, I. T., Worlu, R. E., Adebayo, O. P., & Jonathan, O. (2019). Predicting Customer Behavior with Combination of Structured and Unstructured Data. *Journal of Physics: Conference Series*, *1299*(1), 1–15. https://doi.org/10.1088/1742-6596/1299/1/012041

Akhlaq, A., & Ahmed, E. (2016). Gender differences among online shopping factors in Pakistan. *Organizations and Markets in Emerging Economies*, *7*(1), 74–89. https://doi.org/10.15388/omee.2016.7.1.14216

Alghanam, O. A., Al-Khatib, S. N., & Hiari, M. O. (2022). Data Mining Model for Predicting Customer Purchase Behavior in e-Commerce Context. *International Journal of Advanced Computer Science and Applications*, *13*(2), 421–428. https://doi.org/10.14569/IJACSA.2022.0130249

An, Y., Meng, S., & Wu, H. (2022). Discover Customers' Gender From Online Shopping Behavior. *IEEE Access*, *10*(1), 13954–13965. https://doi.org/10.1109/ACCESS.2022.3147447

Arif, M., Malik, M. H., & Ghous, H. (2021). E-Commerce Customer Purchase Behavior Using Data Mining Prediction. *International Journal of Scientific & Engineering Research*, *12*(2), 2021. http://www.ijser.org

Arora, S., & Sahney, S. (2018). Antecedents to consumers' showrooming behavior: an integrated TAM-TPB framework. *Journal of Consumer Marketing*, *35*(4), 438–450. https://doi.org/10.1108/JCM-07-2016-1885

Asniar, & Surendro, K. (2019). Predictive analytics for predicting customer behavior. *Proceeding - 2019 International Conference of Artificial Intelligence and Information Technology, ICAIIT 2019*, 230–233. https://doi.org/10.1109/ICAIIT.2019.8834571

Brunello, A., Marzano, E., Montanari, A., & Sciavicco, G. (2019). J48SS: A novel decision tree approach for the handling of sequential and time series data. *Computers*, *8*(1), 1–28. https://doi.org/10.3390/computers8010021

Chaubey, G., Gavhane, P. R., Bisen, D., & Arjaria, S. K. (2022). Customer purchasing behavior prediction using machine learning classification techniques. *Journal of Ambient Intelligence and Humanized Computing*, *April*. https://doi.org/10.1007/s12652-022-03837-6

Choudhury, A. M., & Nur, K. (2019). A machine learning approach to identify potential customers based on purchase behavior. *1st International Conference on Robotics, Electrical and Signal Processing Techniques, ICREST 2019*, *February 2021*, 242–247. https://doi.org/10.1109/ICREST.2019.8644458

Filipas, A. M., Vretenar, N., & Prudky, I. (2023). Decision trees do not lie: Curiosities in preferences of Croatian online consumers. *Zbornik Radova Ekonomskog Fakulteta u Rijeci / Proceedings of Rijeka Faculty of Economics*, *41*(1), 157–181. https://doi.org/10.18045/zbefri.2023.1.157

Ha, N. T., Nguyen, T. L. H., Nguyen, T. P. L., & Nguyen, T. Do. (2019). The effect of trust on consumers' online purchase intention: An integration of tam and tpb. *Management Science Letters*, *9*(9), 1451–1460. https://doi.org/10.5267/j.msl.2019.5.006

Hasan, B. (2010). Exploring gender differences in online shopping attitude. *Computers in Human Behavior*, *26*(4), 597–601. https://doi.org/10.1016/j.chb.2009.12.012

Kolahkaj, M., & Madjid Khalilian. (2015). A recommender system using a classification based on frequent pattern mining and the J48 algorithm. *International Conference on Knowledge-Based Engineering and Innovation*, 1–7.

Kovacevic, D., & Kascelan, L. (2020). Internet usage patterns and gender differences: A deep learning approach. *IEEE Consumer Electronics Magazine*, *9*(6), 105–114. https://doi.org/10.1109/MCE.2020.2986817

Li, J., Pan, S., Huang, L., & Zhu, X. (2019). A machine learning-based method for customer behavior prediction. *Tehnicki Vjesnik*, *26*(6), 1670–1676. https://doi.org/10.17559/TV-20190603165825

Liu, C. J., Huang, T. S., Ho, P. T., Huang, J. C., & Hsieh, C. T. (2020). Machine learning-based e-commerce platform repurchase customer prediction model. *PLoS ONE*, *15*(12 December), 1–15. https://doi.org/10.1371/journal.pone.0243105

Moon, N. N., Talha, I. M., & Salehin, I. (2021). An advanced intelligence system in customer online shopping behavior and satisfaction analysis. *Current Research in Behavioral Sciences*, *2*(June). https://doi.org/10.1016/j.crbeha.2021.100051

Mou, J., & Benyoucef, M. (2021). Consumer behavior in social commerce: Results from a meta-analysis. *Technological Forecasting and Social Change*, *167*(January). https://doi.org/10.1016/j.techfore.2021.120734

Pradhana, F., & Sastiono, P. (2019). *Gender Differences in Online Shopping: Are Men More Shopaholics Online? 72*(Icbmr 2018), 123–128. https://doi.org/10.2991/icbmr-18.2019.21

Rianto, Nugroho, L. E., & Santosa, P. I. (2017). Pattern discovery of Indonesian customers in an online shop: A case of fashion online shop. *Proceedings - 2016 3rd International Conference on Information Technology, Computer, and Electrical Engineering, ICITACEE 2016*, 313–316. https://doi.org/10.1109/ICITACEE.2016.7892462

Safarkhani, F., & Moro, S. (2021). Improving the accuracy of predicting bank depositor's behavior using a decision tree. *Applied Sciences (Switzerland)*, *11*(19). https://doi.org/10.3390/app11199016

Sun, Q., Wang, C., & Cao, H. (2009). An Extended TAM for Analyzing Adoption Behavior of Mobile Commerce. *2009 8th International Conference on Mobile Business*, 52–56. https://doi.org/10.1109/ICMB.2009.16

Wong, K. X., Wang, Y., Wang, R., Wang, M., Oh, Z. J., Lok, Y. H., Khan, N., & Khan, F. (2023). Shopee: How Does E-commerce Platforms Affect Consumer Behavior during the COVID-19 Pandemic in Malaysia? *International Journal of Accounting & Finance in Asia Pacific*, *6*(1), 38–52. https://doi.org/10.32535/ijafap.v6i1.1934