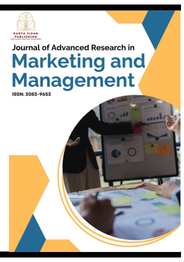
KARYA ILHAM PUBLISHING
CONNECTING MINDS, BRIDGING IDEAS, INSPIRING READERS

Journal of Advanced Research in Marketing and Management

Journal homepage:
https://karyailham.com.my/index.php/jarmm/index
ISSN: 3083-9653



Predicting Customer Behaviour on Buying Life Insurance using Machine Learning

Abu Ayub A Ghafar[1], Norhana Abd Rahim[1,*] , Yumn Suhaylah Yusoff[1]

[1] Faculty of Science and Technology, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

ARTICLE INFO

Article history:
Received 14 August 2025
Received in revised form 28 August 2025
Accepted 5 September 2025
Available online 15 September 2025

Keywords:
Decision trees; logistics regression; Naïve Bayes; machine learning algorithms; life insurance; prediction

ABSTRACT

Despite modest growth in insurance penetration, Malaysia remains below the benchmark set by Bank Negara Malaysia, with life insurance still underutilized as a personal risk management tool. There are many factors that contribute to these issues. To address this, the present study investigates the potential of machine learning (ML) algorithms to predict consumer behavior in purchasing life insurance in Malaysia. The research identifies demographic factors, household income classifications, and risk perception as key determinants influencing insurance purchase decisions. Data were collected from 400 respondents through online surveys, and ML models, including logistic regression, Naïve Bayes, and decision trees, were employed for analysis. The findings reveal that gender, income group (T20), and risk perception significantly impact the likelihood of purchasing life insurance. Logistic regression emerged as the most effective algorithm, achieving an accuracy of 96.25%, followed by Naïve Bayes (93.75%) and decision trees (92.50%). These results underscore the utility of ML in analysing complex datasets and providing precise predictions, outperforming traditional methods. Future studies are recommended to utilize larger datasets to further refine predictive capabilities and explore additional influencing factors. The integration of ML models offers a promising pathway to improve insurance penetration and strengthen financial security in Malaysia.

1. Introduction

Insurance is crucial in protecting a person from any unfortunate events in their life. These dangers may relate to their own belongings such as their house, car, or health [1]. People are exposed to a variety of dangers on a daily basis since accidents and damages might occur when least expected. One way to lower risk exposure and protect oneself against suffering excessive losses and damages in the event of an unforeseen circumstance is to purchase an insurance [2]. There are various types of insurances that a person could have. One could purchase life insurance, health insurance, property insurance, auto insurance, and many other specialty insurances. These insurances provide different types of coverage for a variety of situations.

* *Corresponding author.*
E-mail address: norhana@usim.edu.my

https://doi.org/10.37934/jarmm.2.1.3950

There are many factors which influence a customer's decision in buying an insurance. Customer buying behaviour significantly depends on the cultural, social, personal, and psychological aspects of said customer [3]. In terms of buying insurance, the customer may be persuaded by factors such as household income levels, demographic factors [4], and risk perception [5]. One notable characteristic of Malaysians is their unfavourable attitudes on life insurance even when they are aware of its value [5]. Understanding the factors that influence customer decisions in buying insurance is pivotal for insurers as it is important for them to realize who their target market should be.

A modern way of predicting customer behaviour is by using machine learning (ML). ML is defined as the branch of research that enables computers to learn without explicit programming as cited in [7]. ML is, at its core, an area of artificial intelligence that gives computers the ability to reason and learn for themselves [8]. Compared to traditional methods, ML is a tool to make life easier as it is used to handle data more efficiently [6]. The simulation of ML models holds a strong connection to computational statistics, which is based on computer-assisted prediction [8].

Several methods are used in ML to address data-related issues. It is very unlikely to have just one kind of algorithm that works for every situation [6]. In order to apply the best ML algorithm to a problem, it must first be appropriately classified before the problem solving can begin. Problems can be classified as anomaly detection, classification, regression, clustering, and reinforcement learning problems [8]. Solutions to these problems can be solved with various algorithms such as Bayesian, Decision Tree, Regression, Clustering, Instance Based, Artificial Neural Network, Support Vector Machine, and Ensemble Learning algorithms [9]. In this study, the type of problem that will be faced is a problem regarding a binary output of either "yes" or "no". An algorithm that addresses this is the Logistic Regression algorithm since the response variable in a logistic regression analysis has a binomial distribution [10]. Another algorithm that may address this problem is the Naïve Bayes algorithm as it is useful for categorization [6]. A decision tree approach may also be appropriate, as it classifies instances by sorting them from root to certain leaf nodes based on feature values [8].

According to Statista [11], Malaysia's insurance industry only contributes to 1.4% of the country's GDP in 2023. Additionally, the insurance industry in Malaysia has seen some growth in penetration rates but is still far from the initial target made by Bank Negara Malaysia in 2017. PricewaterhouseCoopers [12] states that Malaysia's Life Insurance and Family Takaful penetration rates are 19% in 2021 and 54% in 2022. But it is still below the 75% target established by Bank Negara Malaysia. This shows a significant protection gap among Malaysian citizens. Mahdzan and Victorian [4] claimed that financial advisors and life insurance company personnel would confirm that many people are frequently reluctant to invest in life insurance and do not recognise its inherent advantages as a tool for personal risk management. This indicates that Malaysians have poor levels of risk management knowledge and practice, which might leave them unable to protect themselves against certain types of dangers [2]. Because of this, individuals should be aware of the advantages of life insurance as well as the various plans offered by different insurers [13]. Therefore, it is crucial to study customer buying behaviour for insurance products in order to raise insurance penetration rates and, subsequently, improve Malaysians' awareness of risk management and protection [2].

Understanding customer behaviour is a way of making sure that insurance penetration rates improve over time. Nowadays, with the utilization of technology, it allows us to transition from traditional methods for predicting customer behaviour to more modern methods which involve the use of ML models. Do and Trang [14] stated that computational intelligence models such as ML can provide predictions that are more accurate than traditional techniques. In addition, Mahesh [6] acknowledges that the need for ML is growing due to the number of datasets that are available, and that ML is used by many sectors to retrieve relevant data. This allows for the optimization of ML over traditional methods as ML is capable of executing tasks that are beyond human capabilities in an effective manner as it is able to analyse large and complex datasets [8]. Thus, ML models will be much more efficient compared to traditional methods in analysing large datasets.

Luciano *et al.*, [15] acknowledges that there is a strong correlation between insurance demand and employment status, family structure, and income. A study by Mahdzan and Victorian [4] also supports that the demand for life insurance is significantly affected by demographics such as marital status, education level, and income level. Meanwhile, a study by Masud *et al.,* [5] suggests that risk perception is also a significant factor in determining customer purchase behaviour of life insurance. Therefore, the purpose of this research is to investigate the effects of demographic factors, household income classification, and risk perception on customer behaviour in purchasing life insurance. By using ML models, this research also aims to predict whether a customer would buy life insurance based on their demographic factors, household income classification, and risk perception. Following that, this research will also go deeper into assessing the performance of the chosen ML algorithms using criterion such as accuracy, precision, recall, and F1-score.

2. Methodology

2.1 Data Collection and Research Instruments

The data used in this study was collected through online questionnaires. The questionnaires were distributed through social media platforms such as WhatsApp and Instagram. The questions and responses were prepared and collected in Google Forms. A pilot test of 40 responses was first carried out to test the reliability and validity of the instruments before distributing the questionnaires further to the public. A total of 400 responses including the pilot test were collected containing information on the dependent and independent variables of the study from each respondent. The questionnaire is divided into three sections. The first section is to collect data on the respondent's demographics. The second section is to gauge the respondent's level of risk perception. The final section is to identify the respondent's income group and life insurance ownership.

2.2 Conceptual Framework

Based on the conceptual framework illustrated in Figure 1, the study proposes the following hypotheses to examine the determinants influencing the decision to own or purchase life insurance in Malaysia:

- H_1: Gender significantly affects the decision to own or purchase life insurance.
- H_2: Marital status significantly affects the decision to own or purchase life insurance.
- H_3: Educational attainment significantly affects the decision to own or purchase life insurance.
- H_4: Employment status significantly affects the decision to own or purchase life insurance.
- H_5: Household income classification significantly affects the decision to own or purchase life insurance.
- H_6: Risk perception significantly affects the decision to own or purchase life insurance.

To test these hypotheses, a logistic regression model is employed, regressing each independent variable against the dependent variable which is purchase decision. The statistical significance of each factor is assessed using *p*-values, with a threshold set at $p < 0.05$. Variables meeting this criterion are considered to have a statistically significant influence on life insurance purchasing behaviour.

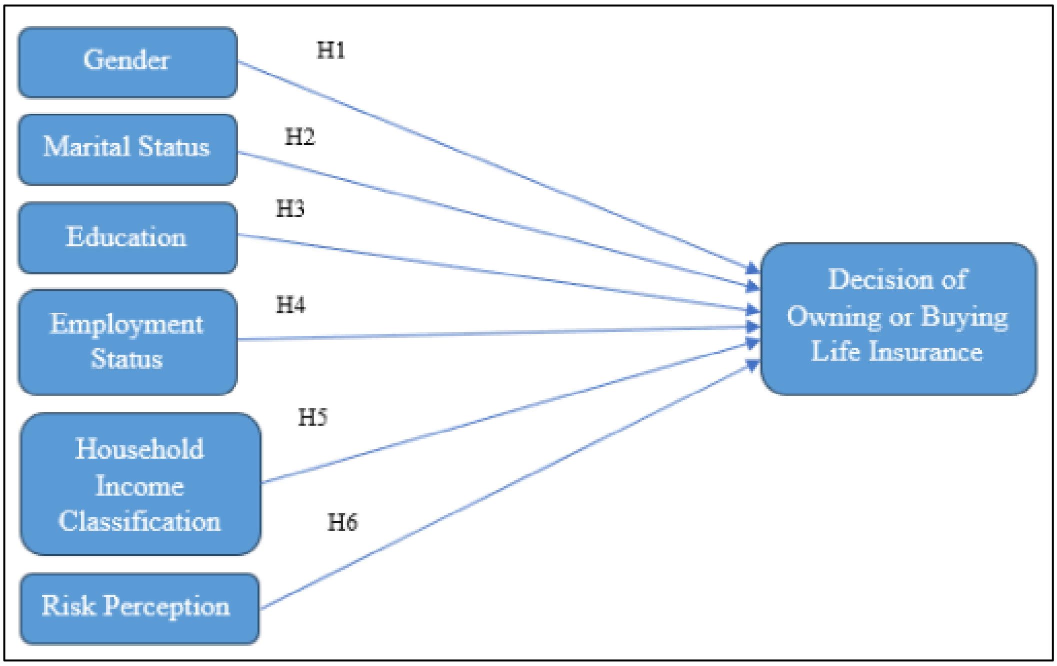


Fig. 1. Conceptual framework

2.3 Reliability and Validity Test

To test the reliability and validity of the research instruments, two tests were carried out during the pilot test phase to ensure suitability of the questionnaire before further distribution to the public. The tests consist of measuring the Cronbach's Alpha and Kaiser-Meyer-Olkin (KMO) values of the collected data. The values must be above 0.7 for it to be considered reliable and suitable for factor analysis.

2.4 Multicollinearity Test

In any predictive model, multicollinearity must be tested between the independent variables of the model. Variance inflation factor (VIF) is a method to determine the level of multicollinearity between the independent variables. When one or more independent variables or inputs have a linear connection, or correlation, multicollinearity is present. Because all of the inputs are impacting one another, multicollinearity causes issues for the regression model because it signifies that they are not truly independent of each other. VIF is useful in determining the extent of multicollinearity problems so that the model may be modified. Thus, the formula for VIF is presented in Eq. (1):

$$VIF_i = \frac{1}{1-R_i^2} \tag{1}$$

where R_i^2 is the unadjusted coefficient of determination for regressing the i^{th} independent variable on the remaining ones. VIF values between 0 and 5 are considered acceptable, while values above 5 indicate high multicollinearity.

2.4 Machine Learning Models
2.4.1 Naïve Bayes Algorithm

The Naïve Bayes algorithm is a classification method which assumes the independence of predictors and is based on the Bayes Theorem as shown in below:

$$P(A|B) = P(B|A) * P(A)P(B) \tag{2}$$

where A and B are observable events. The probabilities of detecting A and B independently are denoted by $P(A)$ and $P(B)$. The conditional probability, or the likelihood of observing A if B is true, is denoted by $P(A|B)$. In contrast, $P(A|B)$ represents the likelihood of observing B, supposing that A is true. The Naive Bayes algorithm proves useful for classification and grouping tasks that rely on the conditional probability of occurrences. Despite its straightforward design and presumptions, the Naive Bayes algorithm offers the advantages of being simple to implement, effective for big datasets, and performing well in challenging contexts like real-world scenarios.

2.4.2 Logistic regression algorithm

The logistic regression algorithm is a predictive modelling approach that looks into the connection between one or more independent variables and a dependent variable. Within the field of regression analysis, logistic regression determines a probability between 0 and 1 to the data, indicating its likelihood of falling into a certain category and classifying the data accordingly. The classification of the data depends on what the classification threshold is defined as. As opposed to a linear regression, a logistic regression estimates the coefficients in terms of the logit function which is the log-odds value as seen in Eq. (3).

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{3}$$

By taking the exponential values of the coefficients as shows in Eq. (4), the effects of beta values can be interpreted as the odd ratios of the variable.

$$Odds\ Ratio = e^{\beta_n} \tag{4}$$

Thus, to calculate the probability of the logistic regression, it is shown in Eq. (5).

$$p = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}} \tag{5}$$

2.4.3 Decision trees algorithm

A decision tree algorithm is an ML algorithm that is used for classification and regression tasks. It models decisions and their possible consequences as a tree-like structure. The structure resembles a flowchart and is used to make predictions or decisions. It is composed of nodes that indicate choices or attribute tests, branches that show how these choices turned out, and leaf nodes that provide the results or outcomes. Every internal node represents an attribute test, every branch represents the test's outcome, and every leaf node represents a class label or a continuous value. A visualisation of a decision tree is shown in Figure 3.

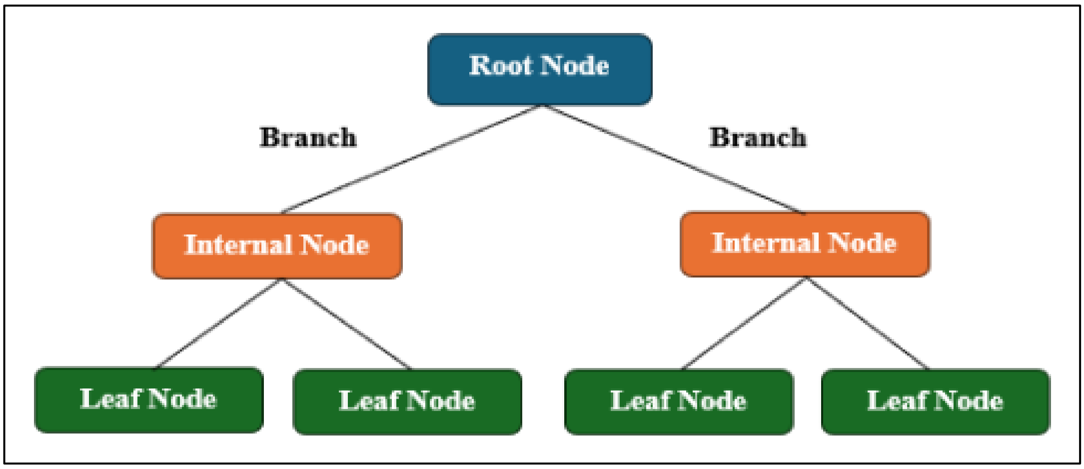


Fig. 2. Elements of decision tree

2.5 Model Accuracy Testing
2.5.1 Confusion matrix

A confusion matrix is a tool that evaluates a classification performance for a machine learning (ML) model. It labels the predicted data as a True Positive (TP), True Negative (TN), False Positive (FP), or False Negative (FN). Not only does the confusion matrix help with assessing the performance of a classification ML model, but it also helps with understanding the type of error that the ML model is making. The confusion matrix can be visualised as shown in Figure 3.

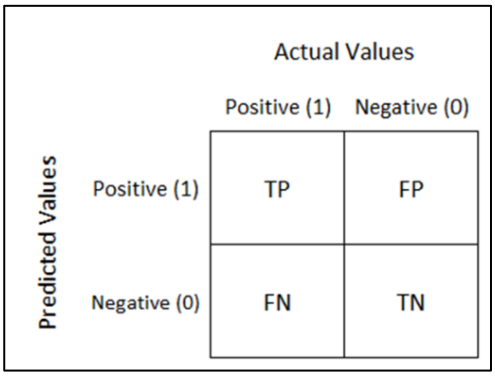


Fig. 3. Confusion matrix

2.5.2 Accuracy

Accuracy is a simple way to measure the performance of a ML model. It is the proportion of all the accurately predicted samples to the total number of samples. The model is better when the value is higher. It can be calculated using Eq. (6).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Although accuracy is simple and easy to understand, it cannot accurately represent the performance of an ML model if the classes of the data are imbalanced. Hence, other metrics such as the precision, recall, and F1 score will also be calculated.

2.5.3 Precision

An ML model's precision is a metric that indicates how frequently it accurately predicts the positive class. Precision may be measured by dividing the total number of correct positive predictions that the model correctly anticipated by the total number of positive predictions. When the cost of a false positive is significant, precision is a useful metric. The formula for precision is shown in Eq. (7).

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

2.5.4 Recall

On the other hand, recall is a metric that calculates how often an ML model can identify the positive outcomes from all the actual positive samples in the dataset. This is done by dividing the number of accurate positive predictions with the number of actual positive predictions. When the cost of a false negative is significant, recall is a useful metric. The calculation for recall can be seen in Eq. (8).

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

2.5.5 F1 score

The F1 score is the harmonic mean of precision and recall. It combines both metrics into a single number, providing a balance between them. When it's necessary to provide equal weight to false positives and false negatives, the F1 score is practical. The measurement of the F1 score is presented in Eq. (9).

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision+Recall} \tag{9}$$

3. Results

3.1 Descriptive Analysis

As shown in Table 1, the respondents' profiles revealed that 54% were male, while 46% were female. This shows relatively balanced distribution allows for meaningful analysis of gender-based differences in insurance ownership. A significant majority of the respondents were employed (85.5%), with smaller proportions being self-employed (8%) or unemployed (6.5%). This skew distribution suggests that stable employment is a key factor in life insurance ownership, likely due to regular income and access to employer-sponsored insurance plans.

Additionally, 65.25% of the respondents had attained tertiary-level education, 27.75% had completed secondary-level education, and only 7% had only received primary-level education. The data also indicated that most respondents belonged to the T20 income group (58.25%). This distribution aligns with the finding that income classification significantly affects insurance ownership, as higher-income individuals are more likely to afford and prioritize life insurance. Finally, 65.75% of respondents own life insurance, while 34.25% do not.

Table 1
Profile of respondents

Variables	Frequency	Percentile (%)
Gender		
Male	216	54.00
Female	184	46.00
Marital status		
Married	205	51.25
Other	195	48.75
Employment status		
Employed	342	85.50
Self-employed	32	8.00
Unemployed	26	6.50
Education level		
Primary education	28	7.00
Secondary education	111	27.75
Tertiary education	261	65.25
Household income group		
B40	74	18.50
M40	93	23.25
T20	233	58.25
Life insurance ownership		
Own	263	65.75
Do not own	137	34.25

3.2 Reliability and Validity Testing

The risk perception construct was measured using three items designed to assess respondents' views on the relative safety and financial risk associated with life insurance or takaful products. The analysis includes measures of central tendency, variability, and reliability, offering a robust understanding of how respondents perceive life insurance as a risk management tool. An analysis of the pilot test yielded a Cronbach's Alpha of 0.883 and a KMO value of 0.728, as presented in Table 2. Both values exceeded the acceptable threshold of 0.7, indicating strong internal consistency and suitability of the data for factor analysis.

Table 2
Reliability and Validity Analysis

Constructs	Mean	Standard deviation	Cronbach's Alpha	KMO
Risk perception			0.883	0.728
Investing in life insurance/takaful has less risk compared to other financial investments	3.83	0.903		
Investing in life insurance/takaful is more secure than investing in the stock market	3.85	1.001		
Purchasing life insurance/takaful does not expose you to financial risks	3.53	1.198		

3.3 Multicollinearity Testing

VIF values were calculated to investigate any multicollinearity between independent variables. Table 3 shows that the VIF values attained are all below 5 with the highest value being 4.348 for Tertiary Education. This proves that there is no multicollinearity between the independent variables.

Table 3
Multicollinearity analysis

Variables	VIF
Gender	1.058
Marital status	1.246
Employed	2.234
Self-employed	2.260
Tertiary education	4.348
Secondary education	3.731
T20	2.224
M40	2.084
Risk perception	1.448

3.4 Logistic Regression Analysis

A logistic regression analysis was conducted to examine the effects and significance of the independent variables on the dependent variable. The results in Table 4 indicate that three variables which are gender, T20 income group, and risk perception, are statistically significant in influencing a customer's decision to purchase life insurance.

Gender emerged as a significant factor, with the analysis indicating that men are more likely to purchase life insurance than women. This aligns with the findings of Luciano *et al.,* [15], who reported that women are generally less inclined to buy insurance contracts, even after controlling for other variables. One plausible explanation is that men, having shorter life expectancies, perceive a greater need for life insurance to provide financial security for their dependents [16]. Additionally, societal roles and income dynamics may contribute to this disparity, as men are more often positioned as primary earners, thereby increasing their perceived responsibility to secure financial protection.

The T20 income group also showed a significant positive association with life insurance ownership. This aligns with the findings of [15] and [4]. Individuals in this top income group are more likely to afford and prioritize insurance products, reflecting the role of financial capacity and economic stability in shaping insurance behaviour. Higher income levels often correlate with greater financial literacy and asset accumulation, both of which enhance the perceived value of life insurance as a risk management tool.

Additionally, the *p*-value for risk perception is less than 0.01, indicating that customers are much more likely to buy life insurance when they view it as a low-risk asset. This finding is consistent with Masud *et al.,* [5] assertion that people's perception of risk is the main motivator for them to purchase life insurance, with those who perceive a higher level of risk being more likely to do so whether they do so online or offline.

The coefficient value of 1.988 for gender suggests that males are over seven times more likely to purchase life insurance compared to females. Similarly, the T20 income group has a coefficient of 2.917, indicating that individuals in the T20 household income group are more than 18 times more likely to purchase life insurance compared to those in the B40 household income group.

Table 4
Logistics regression analysis

Variables	β	e^{β}	*p*-value
Gender	1.988	7.30	0.001
Marital status	-0.283	0.75	0.626
Employed	1.003	2.73	0.168
Self-employed	1.787	5.97	0.161
Tertiary education	-3.844	0.02	0.125
Secondary education	-2.778	0.06	0.264
T20	2.917	18.49	0.001
M40	1.194	3.30	0.070
Risk perception	2.360	10.59	<0.001

3.5 Situational Prediction

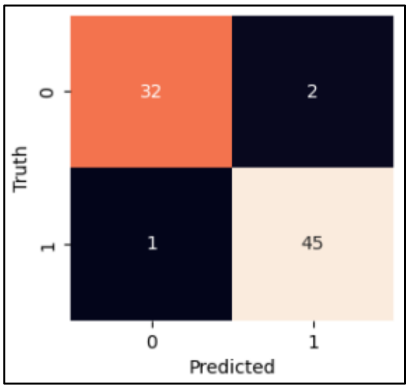
The purpose of using ML algorithms is to predict whether a customer would purchase life insurance based on prior information about the customer. To evaluate the performance of each algorithm, a test is conducted assuming the customer is a married female, employed, has attained tertiary-level education, belongs to the T20 income group, and has a risk perception score of 3.6. As shown in Table 5, both the logistic regression and Naïve Bayes algorithms predicted that the customer would purchase life insurance, with probabilities of 0.6583 and 0.9999, respectively. In contrast, the decision tree algorithm predicted that the same customer would not purchase life insurance.

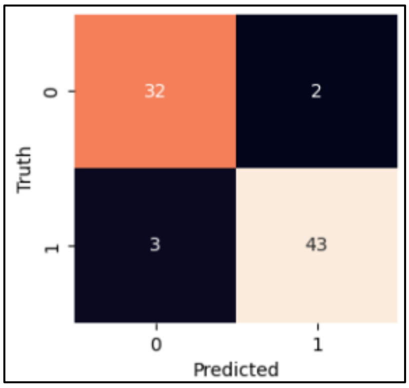
Table 5
Prediction results

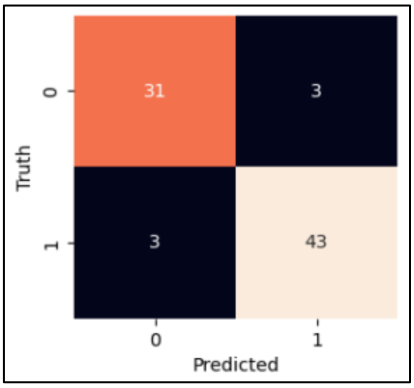
ML Algorithm	Probability (Buy)	Prediction
Logistic Regression	0.6583	Buy
Naïve Bayes	0.9999	Buy
Decision Tree	NA	Do not buy

3.6 Model Accuracy

To assess the performance of each ML algorithm, a confusion matrix was used to calculate accuracy, precision, recall, and F1-score. A test set of 80 respondents was used to evaluate each model, consisting of 46 customers who purchased life insurance and 34 customers who did not. As shown in Figure 4, the logistic regression algorithm predicted 47 customers would purchase life insurance and 33 would not. The Naïve Bayes algorithm, on the other hand, predicted 45 customers would purchase and 35 would not. Meanwhile, the decision tree algorithm predicted 46 customers would purchase life insurance and 34 would not. The values from each confusion matrix were then used to compute the accuracy, precision, recall, and F1-score for each ML algorithm, as summarized in Table 6. The results indicate that the logistic regression algorithm outperformed the others, providing the highest accuracy, precision, recall, and F1-score, followed by the Naïve Bayes algorithm, with the decision tree algorithm coming last. Each metric for all three algorithms comfortably exceeded 0.9, indicating that all three models are suitable for commercial use.







(a) (b) (c)

Fig. 4. Confusion matrix (a) Logistic Regression (b) Naïve Bayes (c) Decision Tree

Table 6
Performance assessment

ML Algorithm	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.9625	0.9627	0.9625	0.9624
Naïve Bayes	0.9375	0.9380	0.9375	0.9376
Decision Tree	0.9250	0.9250	0.9250	0.9250

4. Conclusions

This study investigates the influence of demographic characteristics, household income classification, and risk perception on consumer behaviour in purchasing life insurance in Malaysia. The analysis reveals that gender, membership in the T20 income group, and especially risk perception are significant predictors of life insurance ownership. These findings underscore the importance of understanding psychological and socioeconomic factors in shaping buying decision of life insurance. This study employed machine learning models namely logistic regression, Naïve Bayes, and decision trees and demonstrates the practical utility of ML techniques in predicting consumer behaviour with high accuracy. Among these machine learning models, logistic regression emerged as the most effective, achieving a predictive accuracy of 96.25%. These findings underscore the utility of ML in analysing complex datasets and providing precise predictions, outperforming traditional statistical methods.

Acknowledgement
This research was not funded by any grant.

References
[1] Zakaria, Zainuddin, Nurul Marina Azmi, Nik Fakrul Hazri Nik Hassan, Wan Anisabanum Salleh, Mohd Tajul Hasnan Mohd Tajuddin, Nur Raihana Mohd Sallem, and Jannah Munirah Mohd Noor. "The intention to purchase life insurance: A case study of staff in public universities." *Procedia Economics and Finance* 37 (2016): 358-365. https://doi.org/10.1016/S2212-5671(16)30137-X

[2] Guan, Lim Phui, Dayang Hasliza Muhd Yusuf, and Mohd Rosli Abdul Ghani. "Factors influencing customer purchase intention towards insurance products." *International Journal of Business and Management* 4, no. 5 (2020): 70-79. https://doi.org/10.26666/rmp.ijbm.2020.5.9

[3] Khaniwale, M. (2015). "Consumer Buying Behavior." *International Journal of Innovation and Scientific Research, 14*(2), 278–286. http://www.ijisr.issr-journals.org/abstract.php?article=IJISR-14-129-01

[4] Mahdzan, Nurul Shahnaz, and Sarah Margaret Peter Victorian. "The determinants of life insurance demand: A focus on saving motives and financial literacy." *Asian social science* 9, no. 5 (2013): 274. https://doi.org/10.5539/ass.v9n5p274
[5] Masud, Muhammad Mehedi, Md Rajibul Ahsan, Noor Azina Ismail, and Md Sohel Rana. "The underlying drivers of household purchase behaviour of life insurance." *Society and Business Review* 16, no. 3 (2021): 442-458. https://doi.org/10.1108/SBR-08-2020-0103
[6] Mahesh, B. (2018). "Machine Learning Algorithms - A Review." *International Journal of Science and Research (IJSR).* https://doi.org/10.21275/ART20203995
[7] Samuel, A. L. (1959). "Some Studies in Machine Learning Using the Game of Checkers." *IBM Journal of Research and Development, 3*(3), 210-229. https://doi.org/10.1147/rd.33.0210
[8] Alzubi, Jafar, Anand Nayyar, and Akshi Kumar. "Machine learning from theory to algorithms: an overview." In *Journal of physics: conference series*, vol. 1142, p. 012012. IOP Publishing, 2018.. https://doi.org/10.1088/1742-6596/1142/1/012012
[9] Liakos, Konstantinos G., Patrizia Busato, Dimitrios Moshou, Simon Pearson, and Dionysis Bochtis. "Machine learning in agriculture: A review." *Sensors* 18, no. 8 (2018): 2674. https://doi.org/10.3390/s18082674
[10] Rymarczyk, Tomasz, Edward Kozłowski, Grzegorz Kłosowski, and Konrad Niderla. "Logistic regression for machine learning in process tomography." *Sensors* 19, no. 15 (2019): 3400. https://doi.org/10.3390/s19153400
[11] Statista. (2024). *Share of insurance to the GDP in Malaysia 2015-2023*. https://www.statista.com/statistics/1423604/malaysia-share-of-insurance-to-the-gdp/
[12] PricewaterhouseCoopers. (2023). *Revolutionising financial inclusion through Digital Insurance and Takaful.* https://www.pwc.com/my/en/publications/2023/revolutionising-financial-inclusion-through-digital-insurance-and-takaful.html
[13] Saad, Norma Md, Nur Edzalina Haji Idris, and N. Edzalina. "Efficiency of life insurance companies in Malaysia and Brunei: a comparative analysis." *International Journal of Humanities and Social Science* 1, no. 3 (2011): 111-122.http://irep.iium.edu.my/id/eprint/3263
[14] Do, Q., and T. Trang. "An approach based on machine learning techniques for forecasting vietnamese consumers' purchase behaviour." *Decision Science Letters* 9, no. 3 (2020): 313-322. https://doi.org/10.5267/j.dsl.2020.5.004
[15] Luciano, Elisa, J. François Outreville, and Mariacristina Rossi. "Life insurance demand: evidence from Italian households; a micro-economic view and gender issue." (2015): 010*.* https://doi.org/10.2139/ssrn.2608482
[16] Jahan, Tasmin, and Md Mahiuddin Sabbir. "Analysis of consumer purchase intention of life insurance: Bangladesh perspective." *Business Review–A Journal of Business Administration Discipline* 13, no. 2 (2018): 13-28. https://doi.org/10.35649/KUBR.2018.13.12.2