The Purchase Intention of Green Personal-Care Products: An Artificial Neural Networks Approach

Shamali Perera[1]*
Arfian Zudana[2]

1 School of Applied Business, Unitec, Tāmaki Makaurau Auckland, Aotearoa New Zealand

2 Oxford Brookes Business School, Oxford Brookes University, UK

* Corresponding author: Shamali Perera, ksl772008@gmail.com

https://doi.org/10.34074/jabr.19103.13

Declaration of competing interest:
The authors declare that they have no conflict of interest.

This publication may be cited as:
Perera, S., & Zudana, A. (2025). The Purchase Intention of Green Personal-Care Products: An Artificial Neural Networks Approach. *New Zealand Journal of Applied Business Research Research, 19*(1): 46–71.

Manuscript received: 03/03/2025
Manuscript revised: 22/06/2025
Manuscript accepted: 10/07/2025

Handling editor: Dr Khaled Ibrahim

Contact:
www.unitec.ac.nz/epress
ePress
Unitec
Private Bag 92025, Victoria Street West
Tāmaki Makaurau Auckland 1142
Aotearoa New Zealand

ISSN 2816-1602

The Purchase Intention of Green Personal-Care Products: An Artificial Neural Networks Approach

Shamali Perera
Arfian Zudana

Abstract

This study explores the use of deep learning (DL) to predict consumer purchase intentions for green personal-care products. Using primary data collected from 110 consumers via online questionnaires, the study applies artificial neural networks (ANNs) to analyse purchasing behaviour, diverging from traditional statistical methods. Through a rigorous optimisation process, the ANN model achieves outstanding performance, with 90% Accuracy, 100% Recall, and F1 and Precision scores of 95%. These results outperform other machine learning models tested, including random forest, support vector machine (SVM), gradient boosting, XGBoost and decision tree, highlighting the ANN's superior ability to handle complex, non-linear relationships within small datasets. This study demonstrates that deep learning models are highly adaptable to business requirements in the green personal-care sector, whether in traditional retail or online settings, and can be applied effectively to forecast consumer trends regardless of whether the datasets are complex or relatively simple.

Keywords

Purchase intention
Green product
Deep learning
Artificial neural network

1. Introduction

The demand for green personal-care products is rising in Aotearoa New Zealand, with one in ten consumers shifting their spending for environmental reasons (Taunton, 2024). Products featuring non-toxic ingredients and biodegradable or recyclable packaging are also gaining traction due to increasing awareness of allergies and skin sensitivities linked to harsh chemicals (Biswas & Roy, 2015). Aotearoa New Zealand's personal-care market reached USD 1.53 billion in 2024 and is projected to grow at a CAGR of 5.5% through 2028, outpacing Australia (Statista, 2024). This growth is driven by an increasing preference for sustainability, prompting companies to integrate biodegradable components, eco-conscious packaging and green marketing campaigns into their business strategies. As product safety and environmental friendliness become core brand values, businesses are leveraging these trends to gain a competitive advantage (Choudhury et al., 2024; Coriolis Research, 2022).

However, as personal-care companies adapt to increasing demand for green products in Tāmaki Makaurau Auckland, understanding and predicting green purchase intentions is crucial. Consumer behaviour is shaped by a mix of attitudinal, behavioural, and demographic factors. Traditional predictive methods struggle to capture these non-linear, multifaceted relationships. Therefore, advanced tools such as deep learning (DL) are needed to develop robust, data-driven models for more precise forecasting. This study adopts artificial neural networks (ANNs), a DL method, to fill that gap. Predictive models, including machine learning (ML) and DL approaches, are increasingly being used to assess consumer intentions, segment markets and support strategic business decisions (Choudhury et al., 2024; Witek & Kuźniar, 2021). Such models not only forecast future sales, but also help identify key consumer segments, thereby enhancing marketing efficiency (Atienza et al., 2022).

A variety of analytical techniques have been used in similar contexts. ML methods such as decision trees (DTs), random forest (RF), XGBoost and support vector machines (SVMs) have gained popularity for their ability to handle non-linear variable interactions and produce classification-based predictions (Choudhury et al., 2024). More recently, DL models such as ANNs and recurrent neural networks (RNNs) have demonstrated even higher accuracy and learning capabilities in predicting purchase behaviour, particularly in the green product space (Prashar et al., 2016). Previous research has identified multiple factors that influence green purchase intention. These include environmental attitudes, behavioural motivations, and demographic characteristics such as age, gender and education level (Lasuin & Ng, 2014). This study incorporates these variables to build a comprehensive predictive model tailored to green personal-care consumption.

Using primary data from consumers, this study tested ANN algorithms to classify purchase intentions. Following model optimisation, the ANN model achieved an Accuracy of 90%, Recall of 100%, and both F1 and Precision scores of 95%. These results significantly outperformed traditional ML models – including random forest, SVM, gradient boosting, XGBoost and decision tree – demonstrating the ANN's superior capability in learning patterns and generalising effectively to new data. While ML models are prone to overfitting, ANN showed robustness by minimising sensitivity to noise and maintaining consistent performance across training and test datasets.

This research offers practical value for Aotearoa's personal-care sector. By pinpointing key drivers of purchase intention, businesses can optimise product positioning, personalise outreach and allocate resources more effectively. Additionally, this study contributes to the business analytics field by showcasing the application of neural networks in consumer behaviour prediction. The methodology offers a framework for analysing non-linear relationships in complex datasets and highlights how ANN can elevate forecasting capabilities in the consumer packaged goods (CPG) sector. Ultimately, the findings support the use of AI-powered tools in strategic decision-making and market innovation.

2. Literature Review

This section reviews prior research relevant to this study. Section 2.1 outlines the theoretical basis for purchase intentions and the importance of predicting green purchase intentions. Section 2.2 presents machine learning techniques used for purchase prediction. Section 2.3 examines attitudinal, behavioural and demographic influences on green personal-care purchase intention.

2.1. Theoretical background

The theory of reasoned action (TRA), developed by Fishbein and Ajzen (1975), has been widely adopted to explain behavioural intention. It states that people's intentions are the strongest predictors of their behaviour, and defines intention as a person's readiness to act, shaped by attitudes toward behaviour and perceived social norms. The key determinants of intention are attitudes toward behaviour and subjective norms, which also reflect perceived social pressure from family, friends and society (Fishbein & Ajzen, 1975). TRA is particularly effective in explaining deliberate, non-routine consumer decisions, and has been extensively applied in marketing and consumer behaviour research, particularly in understanding purchase intentions (Ha & Janda, 2012).

Purchase intention reflects a consumer's likelihood or plan to buy a product (Albayrak et al., 2013) and is a strong predictor of actual behaviour (Chen & Yang, 2020). Green purchase intention (GPI) refers to the willingness to choose eco-friendly products over conventional alternatives, signifying concern for environmental impact while maintaining quality expectations (Sabeen et al., 2022).

Biswas and Roy (2015) define green products as those produced with non-toxic materials and eco-friendly packaging. The rise in environmental crises and consumer advocacy for sustainability has shifted consumer behaviour. Individuals increasingly prefer products with non-toxic, safe ingredients and recyclable packaging, driving the awareness and demand for green personal-care products (Kim & Seock, 2009; Biswas & Roy, 2015). Understanding purchase intentions enables companies to tailor marketing strategies, segment audiences and gain a competitive edge. In the green personal-care market, this insight helps identify key decision drivers – such as environmental awareness, ethical sourcing and product performance and supports targeted communication and long-term customer loyalty (Atienza et al., 2022; Choudhury et al., 2024).

2.2. Predictive models for prediction of purchase intention

2.2.1. Machine learning models

Prediction models for purchase intention typically fall into statistical or ML categories. While statistical models such as linear and logistic regression rely on assumptions such as linearity and no multicollinearity, they often fail to capture the complexity of real-world behavioural data, which is seen to be complex, non-linear and influenced by latent factors (Chen et al., 2021). ML algorithms address these limitations by transforming unstructured behavioural data into predictive models without relying on assumptions of linearity (Zhou et al., 2019).

Classification-based ML techniques, such as SVM, DT, RF and gradient boosting (including XGBoost), are widely used to predict binary outcomes such as purchase or no purchase (Larose & Larose, 2014). ML has proven effective in predicting consumer purchase decisions. Choudhury et al. (2024) applied six ML algorithms, including RF and XGBoost, to a dataset of 310 respondents and 12 variables related to green purchasing. All models achieved over 75% accuracy, with

RF producing the highest AUC (0.84). Feature importance plots highlighted the most influential predictors of green purchase intention.

2.2.2. *Deep learning*

DL models, or neural networks, are computational systems that mimic human neural biology by using multiple simple processors to learn relationships between variables and simulate neural processes through multiple layers of interconnected nodes (Prashar et al., 2016). These models use backpropagation and hidden layers to detect complex patterns in non-linear, high-dimensional data (Prashar et al., 2016; Vora & Bhatia, 2023). Their robustness against noise and adaptability make them ideal for classification problems in consumer prediction (Larose & Larose, 2014). Neural networks, such as ANN, consist of input, hidden and output layers, with activation functions such as ReLU and tanh determining neuron outputs (Bigarella, 2024). Optimisation techniques such as stochastic gradient descent (SGD) and Adam adjust weights to minimise error. Key parameters such as batch size, learning rate and number of epochs significantly impact performance and require fine tuning (Kalliola et al., 2021). While some studies favour ReLU for efficiency, others find that performance varies by dataset and problem type. For example, Bigarella (2024) found both tanh and ReLU to be effective under different conditions. Hence, model performance and optimisation are achieved by configuring the above parameters.

Different types of neural networks offer different, unique strengths in predictive modelling. RNNs have been extensively used for sequential data processing, making them ideal in prediction tasks involving sequential or time-dependent data due to their ability to retain historical context (Nayebi et al., 2022). Deep neural networks (DNNs), with their inherent multiple hidden layers, are capable of learning deep abstract representations; hence, they are effective in handling complex, high-dimensional datasets where shallow models fail to capture underlying structures, such as human decision-making (Fintz et al., 2022). DNNs use multiple fully connected layers, thereby enhancing their learning capability (Chaudhuri et al., 2021).

Various DL architectures serve unique functions. RNNs process sequential data and are effective in retaining historical context, making them suitable for time-series prediction tasks (Nayebi et al., 2022). DNNs, with multiple hidden layers, capture abstract relationships in complex datasets, such as decision-making behaviour (Fintz et al., 2022; Chaudhuri et al., 2021). Convolutional neural networks (CNNs), typically used for image and spatial data, are also applied to high-dimensional datasets in scientific contexts, such as chemical analysis (Peleato, 2022). ANNs, with a feedforward architecture and no backward connections, are well suited for static classification tasks. Chaudhuri et al (2021) reported 89% accuracy using DNNs for behavioural prediction. Similarly, Kumar et al. (2022) achieved 93.5% accuracy with ANNs on internet purchase intention, and Prashar et al. (2015) reached 96% accuracy in an impulse-purchase prediction model involving 24 variables. These results underscore the reliability of ANNs when carefully optimised.

2.3. Key influencing factors of green personal-care purchase intention

2.3.1. *Perceptions and attitudinal factors*

The theory of reasoned action (Fishbein & Ajzen, 1975) identifies attitudes as a key factor influencing purchase intention. Research highlights environmental consciousness as a significant predictor of green consumption. Environmental consciousness reflects a consumer's emotional engagement with environmental issues and willingness to act (Lasuin & Ng, 2014). Studies confirm that heightened environmental consciousness positively impacts green purchasing behaviour (Qomariah & Prabawani, 2020; Vinoth, 2023). Another crucial factor is environmental knowledge, which refers to a consumer's awareness of ecological and sustainability issues (Qomariah &

Prabawani, 2020; Choudhury et al., 2024). Consumers with high environmental knowledge often exhibit strong pro-environmental attitudes, increasing their likelihood of purchasing eco-friendly products (Mostafa, 2009). Choudhury et al. (2024) incorporated this variable into machine learning models to predict purchase intention.

Product and brand-related factors also influence green purchasing. Brand image plays a crucial role in shaping consumer perceptions (Tsarenko et al., 2013) and purchase decisions (Swaminathan et al., 2020). Many companies invest in green marketing to enhance their brand's perception of sustainability (Samarasinghe, 2012). A green brand-image formed through sustainability initiatives fosters positive associations with green products and increases purchase intention (Chen & Yang, 2020). Studies confirm that a strong green brand-image significantly impacts consumer buying behaviour (Wu & Chen, 2014), making it a vital predictor in purchase intention models.

Green self-image, or self-identity, also drives green purchasing. It represents how individuals perceive themselves as being environmentally responsible (Van der Werff et al., 2013). Choudhury et al. (2024) found green self-image to be the strongest predictor of green purchase intention in India and integrated it into ML models. Consumers align purchases with personal values and seek social recognition, as seen in Malaysian university students (Lasuin & Ng, 2014). Egoistic values also positively impact green purchase intention (Atienza et al., 2022). Therefore, green self-image is a critical factor influencing eco-friendly purchasing decisions.

2.3.2. Behavioural factors

Social influences play a significant role in shaping purchase intentions. Solomon et al. (2006) describe peer influence as a key factor, where opinions of others impact purchasing decisions. Choudhury et al. (2024) emphasise that social consciousness in green purchasing creates peer pressure, leading to feelings of guilt when individuals fail to conform. Studies support this notion, with Larsson and Arif Khan (2012) finding that Swedish consumers buy organic food to engage in socially acceptable behaviours. Similarly, Öhman (2011) highlights Sherman's (1980) findings that consumers tend to over-indulge in green purchases due to social responsibility and adherence. However, not all studies confirm the impact of social pressure. A study among Jakarta university students found that social influence ranked as the second-lowest factor affecting green purchase intentions (Irawan & Darmayanti, 2012).

Price sensitivity is another crucial factor in green purchasing. Price significantly influences purchasing decisions and consumer satisfaction (Herrmann et al., 2007). As green products are often more expensive than conventional ones (Witek & Kuźniar, 2021), price acts as a deterrent to purchase. Consumers tend to prioritise financial advantages over eco-friendliness (Kim & Seock, 2009). However, some studies suggest that consumers who highly value environmental benefits are willing to pay a premium, reducing price sensitivity (Moslehpour et al., 2021). Their research confirms the significant impact of price sensitivity on green purchase intentions, highlighting the importance of understanding consumer price perceptions when marketing eco-friendly products.

2.3.3. Demographic factors

The impact of demographic factors on green personal-care product purchases varies across studies. Some research suggests that women show stronger environmental concerns and a higher affinity for green products (Chekima et al., 2016). A study in Hong Kong found that female adolescents had greater environmental awareness, positively influencing their green purchase intentions (Lee, 2008). However, research in Malaysia reported no significant gender effect (Chen & Chai, 2010). Some scholars argue that gender differences are not universal, as men with greater environmental knowledge may also engage in green purchasing (Diamantopoulos et al., 2003).

Age also plays a role in purchase intention. A Malaysian study found that younger consumers make simpler purchasing decisions, whereas older consumers take a more cautious approach (Che In & Ahmad, 2018). Fisher et al. (2012) found that individuals over 55 were among the most frequent green product users in the United States. However, other studies suggest that younger consumers are more responsive to eco-friendly products (Chan, 2001).

Education level influences green purchasing behaviour. Highly educated consumers tend to buy more green products due to increased awareness of environmental issues (Che In & Ahmad, 2018). Fisher et al. (2012) and Chan (2001) confirm that higher education correlates with eco-friendly behaviours. However, Yin et al. (2010) found only a marginal impact in China, as educated consumers prioritised personal needs over environmental concerns.

This study developed a prediction model using artificial neural networks (ANNs) to classify green personal-care product purchase intentions. It incorporated factors such as environmental knowledge, green brand image, green self-image, social influence, price sensitivity, age, gender and education level.

3. Research Methodology

3.1. Data collection and sampling

The study entailed primary data collection and obtained full ethical approvals from Unitec's Research Ethics Committee, and thus has adhered to ethical standards by ensuring informed consent, avoiding harm, protecting privacy and preventing deception. For example, researchers provided a statement to participants that declared: "*Your responses will be anonymous, and the responses of this survey will be used only for the purpose stated* ...". In addition, participants were fully informed about the research purpose and their voluntary participation, with the option to review questions beforehand and withdraw at any time. The questionnaire administered online posed no risk of emotional stress or cultural insensitivity, and no personal information was collected. Data confidentiality was strictly maintained.

The target population was current and potential users of personal-care products, which included members of all genders above the age of 18 years and residing in Tāmaki Makaurau Auckland. Purposive sampling was used and included the following segments to reduce sample bias:

- Male, female, transgender and non-binary
- Age group above 18 years
- Working and non-working
- Education qualifications: high school qualifications, certificates and diplomas, bachelor's degrees, and postgraduate qualifications

Purposive sampling was used in this study to ensure representation across key consumer segments relevant to personal-care product usage (e.g., oral care, haircare, skincare, cosmetics, etc.). The questionnaire responses were regularly reviewed to purposively select individuals who brought in diverse experiences not yet captured. A similar approach was taken to ensure diversity in education levels, as prior research (Che In & Ahmad, 2018; Fisher et al., 2012; Chan, 2001) has shown that education influences green purchasing behaviour. This ensured a well-rounded sample reflective of Aotearoa New Zealand consumers, helping to reduce selection bias and enabling validation through feature importance analysis in the predictive model.

Since the target population is unknown, this study used Wibisono's (2003) formula to calculate the sample size.

$$N = \left(\frac{\frac{Z\alpha}{2} . \sigma}{e} \right)^2$$

$$N = \left(\frac{1.96 . 0.25}{0.05} \right)^2$$

$$N = 96.04$$

Where,

N = sample size

$\frac{Z\alpha}{2}$ = confidence level (95%)

σ = standard deviation (0.25)

E = error rate

To determine an appropriate sample size for this study, Wibisono's (2003) formula was employed, given its wide application in applied business research involving model development and validation. This approach ensures statistical adequacy and data representativeness, both of which are essential for building reliable deep learning models capable of generalisation and avoiding overfitting. Wibisono's method is particularly suited to research contexts where multiple independent variables, such as demographic and psychographic factors, are analysed. It supports the practical execution of predictive modelling while maintaining academic rigour, enabling the development of models that are both robust and generalisable across diverse consumer segments.

The minimum sample size was 96, as per the above formula. The data collection instrument was a structured questionnaire incorporating question items in Table 1. Additional questions were included as control questions. Since the sample included both current and potential users, the questionnaire included questions for both segments about the use of environmentally friendly products. The questionnaire was administered online through Google Forms.

3.1.1 *Operationalisation of factors*

Factors including Environmental Knowledge, Green Brand Image, Green Self-Image, Social Influence, Price Sensitivity and Purchase Intention each incorporated several questions. While two questions were included for Green Brand Image, three questions each were included for Environmental Knowledge, Green Self-Image, Social Influence, Price Sensitivity and Purchase Intention. These questions were fine-tuned and structured separately in the questionnaire to suit current and potential users of environmentally friendly personal care products.

A 5-point Likert scale was used to obtain responses for each of these sub-questions. One question was used to collect responses for each of the demographic factors: Age, Gender and Education Level. The operationalisation of the nine factors is depicted in Table 1.

Table 1. Operationalisation of factors.

Factors	Questions	Reference	Question No.
Environmental Knowledge	• I am aware of the general environmental problems. • I understand what environmental protection is about. • I understand the primary causes of global warming.	Moslehpour et al., 2021.	E1 E2 E3
Green Self-Image	• I think of myself as a socially responsible consumer. • I think of myself as an environmentally conscious consumer. • I think of myself as someone who is concerned about social issues.	Hustvedt & Dickson, 2009.	SF1 SF2 SF3
Green Brand Image	• I think/will think of brands that care about the environment when purchasing environmentally friendly personal-care products. • I think/will think of brands that are trustworthy, environmentally friendly personal-care products.	Agmeka et al., 2019. Rahmi et al., 2017.	B1, B3 B2, B4
Social Influence	• I learn so much about environmentally friendly products from friends and relatives. • I learn so much about environmental issues from friends and relatives. • I consider/will consider the opinions of friends, relatives and peers when buying eco-friendly products.	Mei et al., 2012.	SC1 SC2 SC3, SC4
Price Sensitivity	• When I buy environmentally friendly products, I compare/will compare prices with ordinary products. • I remember/will remember the prices of environmentally friendly products that I use. • I check/will check my budget before buying environmentally friendly personal-care products.	Moslehpour et al., 2021.	P1, P4 P2, P5 P3, P6
Purchase Intention	• Over the next month, I am likely to purchase any environmentally friendly personal care products. • Over the next month, I will consider switching to other brands for ecological reasons. • Over the next month, I plan to switch to an environmentally friendly version of products I purchase now.	Chan, 2001. Che In, 2018.	PC1 PC2 PC3
Age	What is your age?		3
Gender	• Male • Female • Other		4

Education Level	• Highest level of education • High school • Certificate/Diploma • Bachelor's degree • Postgraduate certificate/diploma • Master's degree • PhD	5

Note: The table above presents the factors, corresponding survey questions, and their references. Each factor represents a key construct related to green consumer behaviour, with associated questions designed to measure respondents' attitudes and intentions.
Source: Authors' own creation.

3.2. Data analysis

3.2.1. Reliability and validity of data

The reliability and validity of variables used in this study were measured using Cronbach's alpha and composite reliability for internal consistency and average variance extracted (AVE) for convergent validity. The discriminant validity was used to measure how strongly the constructs of purchase intention are related. These metrics were computed using SmartPLS structural equation modelling and tested against recommended thresholds (0.7 for reliability, 0.5 for AVE) (Hair et al., 2019). If the computed values of some constructs fell below these thresholds, this study removed those constructs in accordance with literature guidelines (Sarstedt et al., 2022). The final validated results are presented in Tables 2 and 3 after adjustments.

Table 2. The reliability and convergent validity assessment

Variables	Cronbach's Alpha	Composite Reliability	AVE
Green Brand Image	0.983	0.992	0.983
Environmental Knowledge	0.764	0.863	0.678
Price Sensitivity	0.935	0.969	0.939
Purchase Intention	0.879	0.925	0.804
Green Self-Image	0.829	0.898	0.748
Social Influence	0.878	0.942	0.891

Source: Authors' own creation.

As depicted in Table 2, all the values of Cronbach's alpha and composite reliability for the variables were greater than the minimum value of 0.7 recommended (Hair et al., 2019). Similarly, the AVE values for all variables were above the recommended minimum value of 0.5 (Hair et al., 2019). Hence, the variables' internal consistency, reliability and convergent validity were confirmed.

To test the discriminant validity of the variables, the indicator loadings were compared with the cross-loadings. Table 3 represents the results of this comparison.

Table 3. Discriminant validity assessment.

	Green Brand Image	Environmental Knowledge	Price Sensitivity	Purchase Intention	Social Influence	Green Self-Image
B1	0.992	0.342	0.941	0.521	0.344	0.389
B2	0.991	0.342	0.925	0.479	0.334	0.404
E1	0.252	0.822	0.195	0.222	0.181	0.368
E2	0.199	0.811	0.208	0.271	0.241	0.429
E3	0.387	0.836	0.390	0.293	0.272	0.532
P1	0.933	0.342	0.967	0.465	0.327	0.364
P2	0.892	0.300	0.971	0.502	0.406	0.341
PC1	0.543	0.298	0.547	0.880	0.376	0.433
PC2	0.361	0.299	0.375	0.912	0.366	0.411
PC3	0.431	0.268	0.399	0.898	0.401	0.386
SC1	0.391	0.218	0.421	0.385	0.939	0.400
SC2	0.260	0.319	0.300	0.417	0.948	0.426
SF1	0.351	0.515	0.316	0.392	0.370	0.890
SF2	0.402	0.422	0.372	0.436	0.400	0.915
SF3	0.274	0.493	0.243	0.358	0.366	0.783

Source: Authors' own creation.

3.2.2. *Data processing*

Following data collection via Google Forms, responses were exported to Excel, converted to CSV format and analysed. Since the Google Form design ensures complete responses, no missing values existed.

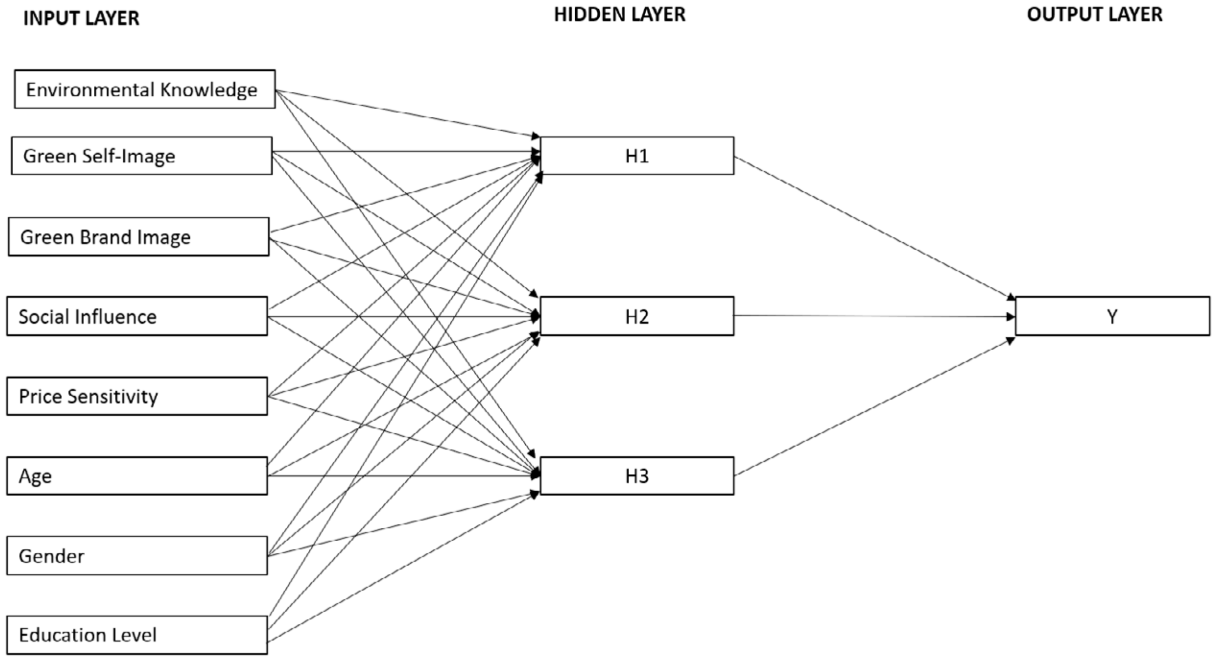
Data transformation involved aggregating mean scores of sub-questions for each of the eight independent variables. Categorical variables such as Gender and Education Level were converted into dummy variables using one-hot encoding. Following Choudhury et al. (2024), Purchase Intention was binarised (scores >2.5 = 1). Python software was used for data preprocessing.

3.2.3. *The prediction model*

This study employed ANN for classification-based prediction, as ANN effectively handles non-linear relationships and non-normal distributions (Leong et al., 2019). It is particularly suited for predicting purchase intention, which involves complex variables such as brand imagery, and behavioural and attitudinal factors (Witek & Kuźniar, 2021).

Studies have found ANN models effective in analysing cause-and-effect relationships in complex systems. In fields such as health sciences, where data is limited and complexity is high, feedforward ANNs with backpropagation have effectively modelled nonlinear relationships while avoiding overfitting through careful design (Pasini, 2015). Similarly, in small-sample purchase prediction studies (Prashar et al., 2016), ANN has proven effective. Thus, ANN was chosen for this study due to its suitability for small datasets with nonlinear variables. The proposed model entailed the model architecture depicted in Figure 1.

Figure 1: Model architecture.



Source: Author's creation.

An 80:20 split in the data was used for training and testing datasets, respectively. The proportions of the binary class distribution in the overall dataset were maintained in both the training and testing datasets. The proposed ANN model consisted of an input layer with nine neurons representing input and target variables. The hidden layers were optimised using hyperparameter tuning, adjusting layer depth, neuron count, activation functions and optimisers to identify the best configuration. The output layer had two neurons for binary classification, using the sigmoid activation function for probability-based interpretation (Bigarella, 2024).

The network depth was found to be crucial for effective predictive power (Kalliola et al., 2021). A study used to predict purchase behaviour tested three and five hidden layers with 64 and 128 neurons for optimal learning. Previous studies have suggested that increasing hidden layers and neurons up to five layers with 128 neurons enhances prediction (Chaudhuri et al., 2021).

Additionally, to improve learning and generalisation, various optimisers and activation functions were tested. ReLU and tanh, effective in deep learning and preventing vanishing gradients, were explored (Bonaccorso, 2018; Kalliola et al., 2021). Hyperparameters such as batch sizes, epochs and optimisation functions were also fine-tuned to maximise predictive accuracy. A summary of these hyperparameters is presented in Table 4.

Table 4. Hyperparameters for ANN model.

Hyperparameters	Selected Values
Neurons in the input layer	8
Number of hidden layers	3, 5
Number of neurons in hidden layer	32, 64, 128
Batch size	32, 64
No. of epochs	100, 200
Activation function – hidden layer	ReLU, tanh
Activation function – output layer	Sigmoid
Optimiser	Adam, SGD

Note: This table presents the hyperparameters tested in the tuning phase.
Source: Authors' own creation.

3.2.4. *Model performance and evaluation*

The ANN model's performance was evaluated by assessing its accuracy in distinguishing between 1s (purchase) and 0s (non-purchase). Performance was measured using the test dataset, ensuring the model generalised well. Overfitting occurs if the model performs better on training data than on testing data (Vora & Bhatia, 2023). Key performance metrics included Accuracy, F1 score, Recall and Precision, calculated for both training and testing datasets to detect overfitting or underfitting issues. Furthermore, the ANN model was compared with RF, SVM, XGBoost, gradient boosting and decision tree models, following methodologies from similar studies (Chaudhuri et al., 2021; Choudhury et al., 2024). The model was implemented in Python, utilising NumPy and Pandas for data structuring, Matplotlib and Excel for visualisation, and TensorFlow and Keras for ANN construction.

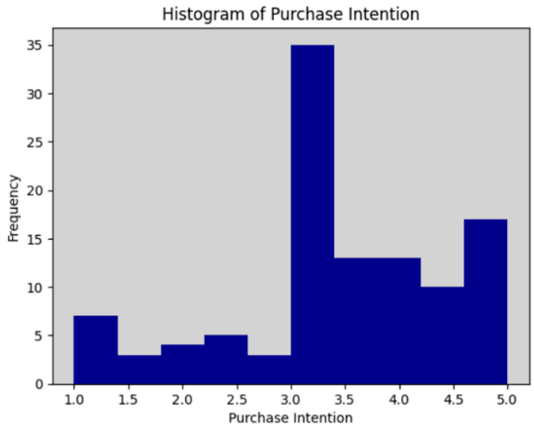
4. Results and Discussion

4.1. Results

4.1.1. *Preprocessed data*

After the preprocessing tasks outlined in 3.2.2 were carried out, the data distribution was depicted (Figure 2 and Figure 3).

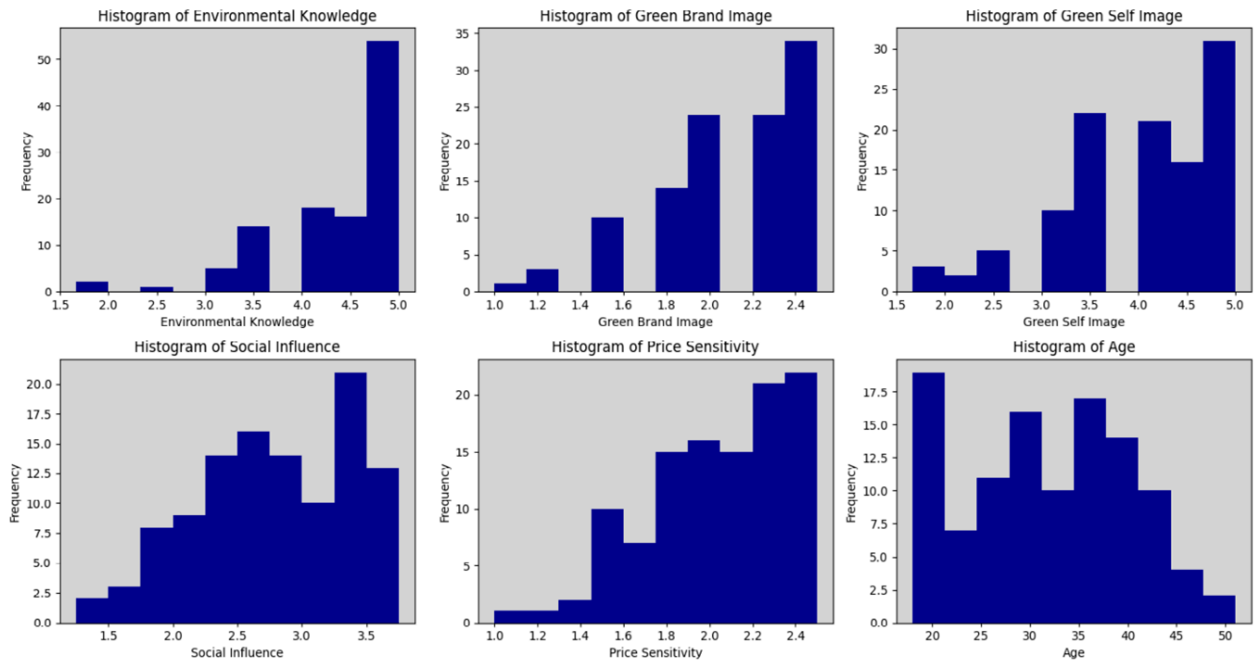
Figure 2: Histogram of Purchase Intention.



Source: Authors' own creation.

The Purchase Intention histogram depicts the highest frequency of responses around the mid-point of 3, indicating that most respondents were indecisive about purchasing green personal-care products. A smaller number of respondents favoured purchasing green products, with fewer observations between the ranges of 1 and 2.5, indicating a lower inclination towards negative purchase behaviour.

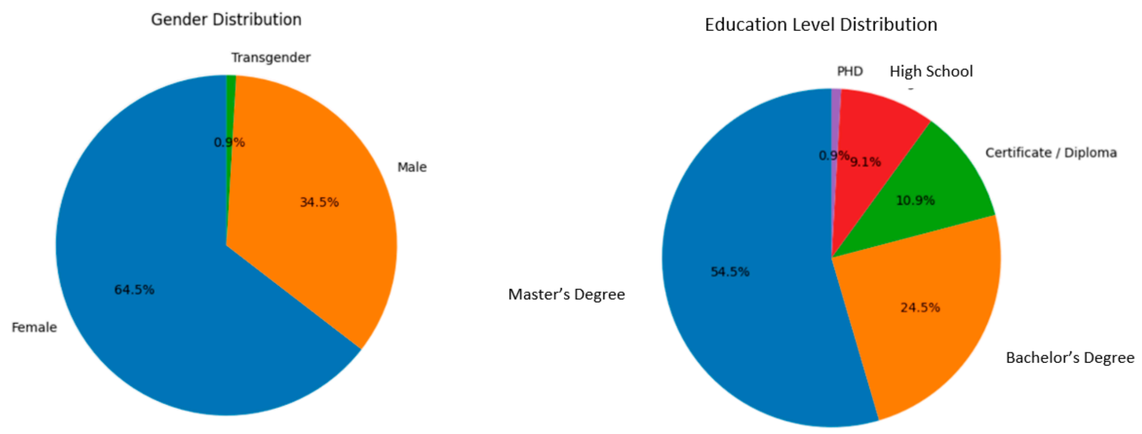
Figure 3: Histograms for independent variables.



Source: Authors' own creation.

The histograms for six variables (excluding Gender and Education Level) show their aggregated mean scores after data transformation. None follows a normal distribution, with most variables exhibiting positive skewness. Environmental Knowledge, Green Brand Image, and Green Self-Image have responses clustered around 4.5–5, while Price Sensitivity is more evenly distributed but slightly concentrated between 4.5–5. The Age variable displays a bimodal distribution, peaking at 20, 30 and 35–40 years.

Figure 4: Proportions of Gender and Education Level variables.



Source: Authors' own creation.

The Gender composition was skewed, with 64.5% being female, followed by 34.5% male and less than 1% in other genders. In terms of Education Level, over 75% of respondents had bachelor's and master's degrees.

As the next steps, categorical variables such as Gender and Education Level were converted to dummy variables using one-hot encoding, enabling the preparation of their use in the prediction model. Table 5 depicts the resulting types of data after the data preprocessing.

Table 5. Data types.

Variable	**Data Type**
Environmental Knowledge	Float
Green Brand Image	Float
Price Sensitivity	Float
Purchase Intention	Float
Green Self-Image	Float
Social Influence	Float
Age	Integer
Gender_Male	Boolean
Gender_Transgender	Boolean
Education Level_Certificate/Diploma	Boolean
Education Level_High School	Boolean
Education Level_Master's Degree	Boolean
Education Level_PhD	Boolean

Source: Authors' own creation.

4.1.2. *Descriptive statistics: Post-processing of data*

The dataset consisted of data for nine variables, of which Environmental Knowledge, Green Brand Image, Price Sensitivity, Purchase Intention, Green Self-Image and Social Influence are categorical in nature. Gender_Male, Gender_Female, Gender_Non-Binary, Gender_Transgender, Education Level_High School, Education Level_Certificate/Diploma, Education Level_Bachelor's Degree, Education Level_Master's Degree and Education Level_PhD are Boolean-type data which result from the transformation of the Gender and Education Level variables. Age is presented as integers.

The statistics in Table 6 depict significant variability in the respondents' attitudes and behaviours concerning the purchase intention of green personal-care products. Environmental Knowledge and Green Brand Image depict high mean scores over 4 with moderate variability, indicating respondents feel well informed about environmental issues. Green Self-Image (3.94) and Price Sensitivity (3.99) also score relatively high, though Green Self-Image shows greater individual differences (SD = 0.84). Social Influence reflects a moderate mean (3.36) but high variability (SD = 1.1). Purchase Intention scores of 3.42 with moderate variability (SD = 1.06) indicate mixed intentions toward eco-friendly personal-care products.

Table 6. Descriptive statistics of the numeric variables.

	Environmental Knowledge	Green Brand Image	Price Sensitivity	Purchase Intention	Green Self-Image	Social Influence	Age
Count	110.00	110.00	110.00	110.00	110.00	110.00	110.00
Mean	4.29	4.20	3.99	3.42	3.94	3.36	31.70
Std	0.71	0.73	0.81	1.06	0.84	1.10	8.60
Min	1.67	2.00	2.00	1.00	1.67	1.00	18.00
25%	4.00	3.62	3.50	3.00	3.33	2.50	26.00
50%	4.33	4.50	4.00	3.33	4.00	3.50	32.00
75%	5.00	5.00	4.50	4.00	4.67	4.00	38.75
Max	5.00	5.00	5.00	5.00	5.00	5.00	51.00

Source: Authors' own creation.

Table 7 depicts the distribution in respondents' demographics: respondents' ages ranged from 18 to 51 years (mean = 31.7, SD = 8.6), representing a young sample. Gender distribution was somewhat unbalanced (65% female, 35% male, 1% transgender). Most respondents were highly educated, with 55% holding a master's degree and 25% a bachelor's degree. The dataset includes 110 responses.

Table 7. Descriptive statistics of the categorical variables.

	Gender_ Female	Gender_ Male	Gender_ Transgender	Education Level_ Bachelor's Degree	Education Level_ Certificate/ Diploma	Education Level_ High School	Education Level_ Master's Degree	Education Level_ PhD
Count	110.00	110.00	110.00	110.00	110.00	110.00	110.00	110.00
Mean	0.65	0.35	0.01	0.25	0.11	0.09	0.55	0.01
Std	0.48	0.48	0.10	0.43	0.31	0.29	0.50	0.10
Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50%	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
75%	1.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00
Max	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Source: Authors' own creation.

4.1.3. *Deep learning model*

This study developed a prediction model using ANN to forecast purchase intention. Taking the learnings of similar studies (Prashar et al., 2016), a base model was constructed consisting of an input layer with 16 neurons, three hidden layers with 64 neurons each, and an output layer with one neuron. A batch size of 64 and 100 epochs per run was adopted. The ReLU activation function was applied to hidden layers, while the Adam optimiser was used for optimisation. The architecture was based on similar studies utilising ANN for purchase behaviour prediction (Chaudhuri et al., 2021). The performance of this base model is presented in Table 8.

Table 8. Performance of the base model.

Epochs: 100 Batch No: 64 Activation (Hidden Layers): ReLU Optimiser: Adam

Configuration			Evaluation: Training				Evaluation: Testing				Difference			
Run #	Hidden Layers	Neurons	Acc	F1	Recall	Precision	Acc	F1	Recall	Precision	Acc	F1	Recall	Precision
Baseline	3	64	1.0	1.0	1.0	1.0	0.77	0.86	0.84	0.88	0.23	0.14	0.16	0.12

Note: This table presents the evaluation results of the base model for both training and testing data. The configuration includes three hidden layers with 64 neurons, ReLU activation function, and Adam optimiser. The metrics shown are Accuracy (Acc), F1 score, Recall and Precision for both the training and testing sets, along with the differences between training and testing scores.
Source: Authors' own creation.

To optimise the model for improved accuracy, hyperparameter tuning was carried out in three stages, where hyperparameters such as the number of hidden layers, neurons per layer, epochs, batch size, activation functions and optimisers were systematically optimised. The first stage of tuning explored configurations with three and five hidden layers consisting of 64 and 128 neurons per layer. The batch number, number of epochs, optimiser and activation were maintained the same as the base model. The performance scores, as indicated in Table 9, reflect that the training accuracy reached 1.0 for all models, but testing accuracy ranged between 0.77 and 0.88. However, runs with higher neurons (128) and additional layers (five) demonstrated an increased performance, F1, Recall and Precision scores of 0.89, suggesting better classification ability compared to the baseline model in Table 8 (0.86, 0.84, 0.88), with the exception of five layers with 128 neurons. However, differences between training and testing scores indicate overfitting, requiring further refinements.

Table 9. Performance of the model with different network configurations.

Epochs: 100 Batch No: 64 Activation (Hidden Layers): ReLU Optimiser: Adam

Configuration			**Evaluation: Training**				**Evaluation: Testing**				**Difference**			
Run #	**Hidden Layers**	**Neurons**	**Acc**	**F1**	**Recall**	**Precision**	**Acc**	**F1**	**Recall**	**Precision**	**Acc**	**F1**	**Recall**	**Precision**
1	3	128	1.0	1.0	1.0	1.0	0.81	0.89	0.89	0.89	0.19	0.11	0.11	0.11
2	5	64	1.0	1.0	1.0	1.0	0.81	0.89	0.89	0.89	0.19	0.11	0.11	0.11
3	5	128	1.0	1.0	1.0	1.0	0.77	0.86	0.88	0.84	0.23	0.14	0.12	0.16

Note: This table presents the evaluation results of the model with different configurations for hidden layers and number of neurons, ReLU activation function, and Adam optimiser. The metrics shown are Accuracy (Acc), F1 score, Recall and Precision for both the training and testing sets, along with the differences between training and testing scores. Thus, the configurations in Runs 1, 2 and 3 are used for the next stage in hyperparameter tuning.
Source: Authors' own creation.

The impact of batch number and epoch configurations was evaluated in the second hyperparameter tuning stage. To mitigate overfitting, the number of epochs was increased to 200, and the batch size was reduced to 32. This stage evaluated models with the best-performing configurations from Stage 1. As illustrated in Table 10, despite changes in batch size and epoch configurations, the testing performance remained unchanged compared to Stage 1. The model with three hidden layers and 128 neurons (Run 4) performed slightly worse than those with five hidden layers (Runs 5 and 6), indicating that increasing epochs and reducing batch size does not improve generalisation but may exacerbate overfitting. The results are compared with Runs 1, 2 and 3, along with the baseline model, presented in Table 10.

Table 10. Performance of the model configurations with 200 epochs and a batch size of 32.

Epochs: 200 Batch No: 32 Activation (Hidden Layers): ReLU Optimiser: Adam

Configuration			**Evaluation: Training**				**Evaluation: Testing**				**Difference**			
Run #	**Hidden Layers**	**Neurons**	**Acc**	**F1**	**Recall**	**Precision**	**Acc**	**F1**	**Recall**	**Precision**	**Acc**	**F1**	**Recall**	**Precision**
4	3	128	1.0	1.0	1.0	1.0	0.77	0.86	0.84	0.88	0.19	0.11	0.11	0.11
5	5	64	1.0	1.0	1.0	1.0	0.81	0.89	0.89	0.89	0.19	0.11	0.11	0.11
6	5	128	1.0	1.0	1.0	1.0	0.81	0.89	0.89	0.89	0.19	0.11	0.11	0.11

Note: This table presents the evaluation results of the model, with the same configurations for number of hidden layers and number of neurons as in Table 9, but different configurations for the number of epochs and batch sizes, ReLU activation function, and Adam optimiser. The metrics shown are Accuracy (Acc), F1 score, Recall and Precision for both the training and testing sets, along with the differences between training and testing scores.
Source: Authors' own creation.

The impact of the activation and optimiser functions was evaluated in Stage 3. The tanh activation function and SGD optimiser were introduced to further improve generalisation while maintaining the best configurations from Stage 2. Table 11 depicts the results showing increased testing Accuracy (0.90) and F1, Recall and Precision scores (0.95–1.0), indicating better generalisation. Notably, Run 7 (three hidden layers, 128 neurons, 100 epochs, batch size 64) demonstrated the best performance, with testing accuracy slightly surpassing training accuracy. The reduced gap between training and testing scores suggested a decline in overfitting compared to ReLU-Adam models. However, models with five hidden layers (Runs 8 and 9) showed slightly lower performance, emphasising that additional layers do not necessarily enhance generalisation.

Table 11. Model performance with tanh activation function and SGD optimiser.

Epochs: 100 Batch No: 64 Activation (Hidden Layers): Tanh Optimiser: SGD

Configuration			**Evaluation: Training**				**Evaluation: Testing**				**Difference**			
Run #	**Hidden Layers**	**Neurons**	**Acc**	**F1**	**Recall**	**Precision**	**Acc**	**F1**	**Recall**	**Precision**	**Acc**	**F1**	**Recall**	**Precision**
7	3	128	0.86	0.92	0.97	0.92	0.90	0.95	1.0	0.95	-0.04	-0.03	-0.03	-0.03
8	5	64	0.86	0.92	0.95	0.88	0.86	0.92	0.94	0.90	0.00	0.00	0.01	-0.02
9	5	128	0.85	0.91	0.94	0.88	0.90	0.95	1.0	0.95	-0.05	-0.04	-0.06	-0.07

Note: This table presents the evaluation results of the model, with different configurations, the activation function, and the optimiser. The number of hidden layers and neurons is the same as in Table 9. The metrics shown are Accuracy (Acc), F1 score, Recall and Precision for both the training and testing sets, along with the differences between training and testing scores.
Source: Authors' own creation.

4.1.4 *Comparison test*

The model performance was further assessed by comparing its performance with other machine learning prediction models. The comparison analysis is represented in Table 12. At first glance, the ANN model appears to have lower training scores (Accuracy = 0.86, F1 = 0.92, Recall = 0.97, Precision = 0.92), compared to random forest, XGBoost, gradient boosting, and decision trees, which achieve perfect training scores (1.0). However, ANN excelled in testing performance, achieving higher testing scores than training scores, and demonstrating superior generalisation compared to other models that struggle with unseen data. ANN's training–testing differences (-0.04, -0.03, -0.03, -0.03) across Accuracy, F1, Recall, and Precision indicated minimal overfitting.

Table 12. Comparison of ANN model performance with other ML models.

Configuration	Evaluation: Training				Evaluation: Testing				Difference			
	Acc	F1	Recall	Precision	Acc	F1	Recall	Precision	Acc	F1	Recall	Precision
ANN	0.86	0.92	0.97	0.92	0.90	0.95	1.00	0.95	-0.04	-0.03	-0.03	-0.03
Random Forest	1.0	1.0	1.0	1.0	0.90	0.95	1.00	0.90	0.10	0.05	0.00	0.10
SVM	0.87	0.92	0.95	0.89	0.86	0.92	0.94	0.90	0.01	0.00	0.01	-0.01
XGBoost	1.0	1.0	1.0	1.0	0.81	0.89	0.89	0.89	0.19	0.11	0.11	0.11
Gradient Boosting	1.0	1.0	1.0	1.0	0.81	0.90	0.94	0.85	0.11	0.10	0.06	0.15
Decision Tree	1.0	1.0	1.0	1.0	0.86	0.92	0.94	0.90	0.14	0.08	0.06	0.10

Note: This table presents a comparison of the performance of the ANN model with other ML models. The metrics shown are Accuracy (Acc), F1 score, Recall and Precision for both the training and testing sets, along with the differences between training and testing scores. Source: Authors' own creation.

ANN achieved the highest testing Accuracy (0.90), while F1 and Recall scores match random forest. ANN and random forest achieved perfect Recall scores (1.0) in testing, meaning they consistently detected positive instances. ANN's Precision improved from training (0.86) to testing (0.95), with only a -0.03 difference, whereas gradient boosting showed the largest discrepancy (0.15). SVM achieved a lower Precision difference (-0.01), but its testing score (0.90) was lower than ANN (0.95), indicating overfitting. Random forest, gradient boosting and decision trees suffered from overfitting, showing a significant performance drop from training to testing (0.81–0.95 in testing). XGBoost and gradient boosting performed the worst.

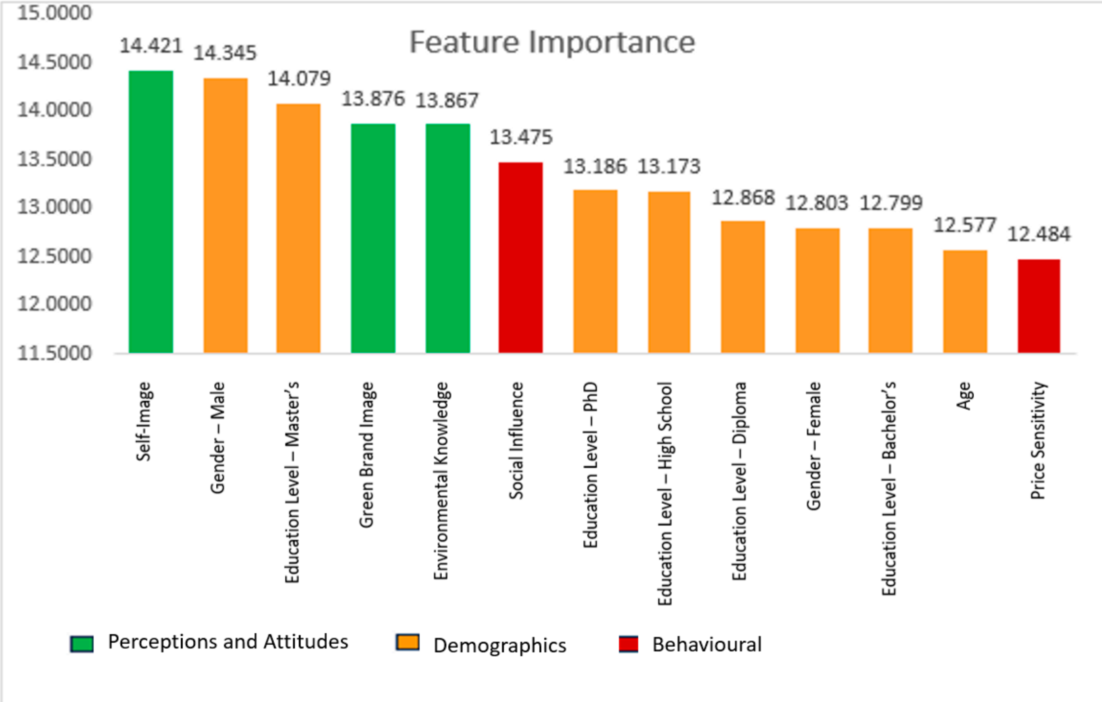
In summary, the ANN prediction model demonstrated consistent performance compared to other machine learning models. While the random forest algorithm achieved comparable F1 and Recall scores during testing, ANN exhibited a closer alignment between training and testing results, reinforcing its robustness and predictive effectiveness. Similar studies on purchase-intention prediction have also demonstrated ANN's superiority over traditional machine learning models, further validating its applicability in consumer behaviour forecasting.

Our finding that ANN can provide a good prediction is similar to Prasad and Ghosal's (2021) findings. In their study, they used ANN to predict the purchase intention for direct-to-consumer (DTC) brands with non-linear variables such as 'safety of transactions', 'availability of innovative products' and 'product quality', where a small dataset of 100 valid responses was used. They found that the ANN's accuracy was 96%. Moreover, a study to predict restaurant check-ins using ANN with a dataset of 120,825 check-ins in a geographical area in New York found that the prediction accuracy was 93%. Despite using different model architectures with a range of hidden layers and number of neurons, the highest accuracy was found to be with one hidden layer (Zheng et al., 2013).

4.1.5. *Feature importance*

The individual importance of each independent variable to the overall prediction of purchase behaviour of green personal-care products was assessed through the 'feature importance' analysis. This analysis was constructed based on the learned weights of the first layer of the neural network. This approach assumed that the larger weight magnitudes in the first layer contribute more significantly to the model's prediction, making these features more important. The feature importance was constructed based on the optimised model used in Run 7. Figure 5 depicts the independent variables' importance in descending order in this model.

Figure 5: Feature importance of the independent variables of the ANN model.



Source: Authors' own creation.

The feature importance scores indicate that all variables contributed meaningfully to predicting purchase intention, though some stood out more than others. The analysis highlighted that perception and attitudinal factors, particularly Green Self-Image, ranked the highest, with a score of 14.421, suggesting that self-perception of environmental consciousness strongly influences purchase decisions. Demographic factors such as Gender_Male also played a significant role, indicating that gender differences may influence purchasing behaviour, though Gender_Female was slightly less important. Similarly, Education Level_Master's followed closely, with scores of 14.345 and 14.078, demonstrating that education, particularly at higher levels, is a critical predictor.

Brand perception also played a moderate role, with Green Brand Image scoring 13.875, reflecting the impact of sustainability perceptions on purchase decisions. Environmental Knowledge followed closely, with a score of 13.866, emphasising the importance of awareness regarding environmental issues. Social Influence was another moderate factor, scoring 13.475, indicating that peer opinions contribute to green purchasing decisions. Various educational qualifications contributed differently, with Education Level_PhD ranking in the top half at 13.185, while Diploma, Master's, and High School qualifications showed varied contributions in the lower half of the ranking.

While age played a role in predicting purchase intention, a lower impact on Purchase Intention than other variables was observed. Similarly, Price Sensitivity was seen to have the least influence, suggesting that price was not a primary driver of environmentally conscious purchasing behaviour. Overall, attitudinal factors such as Green Self-Image and Environmental Knowledge, along with demographic factors such as Gender and Education Levels, were the strongest predictors of purchase intention. Age, some Gender considerations, and Price Sensitivity contributed to a lesser extent. These findings confirm that the ANN prediction model effectively learns and generalises on unseen data, making it a reliable tool for real-world applications.

5. Conclusion, Implications, Limitations and Future Research

This study developed a deep-learning model to predict purchase intentions for green personal-care products. A multi-layer ANN classification model was rigorously optimised and evaluated against key performance metrics and other machine learning models. The findings confirmed the ANN's superiority in predictive accuracy, highlighting its potential as a robust tool for forecasting consumer behaviour in sustainability-driven markets. The study also underscored the importance of network structure optimisation, revealing that deeper architectures can enhance training performance and may lead to overfitting in smaller datasets. The optimal model configuration consisted of three hidden layers with 128 neurons each, using the tanh activation function and the stochastic gradient descent (SGD) optimiser. These refinements ensured a stable and generalised model capable of delivering consistent performance on unseen data.

Feature importance analysis further validated the impact of attitudinal, behavioural and demographic factors, aligning with prior research. The study has significant academic and managerial implications. From an academic perspective, it advocates for the adoption of machine learning models – particularly ANNs – over traditional regression-based methods, which are often constrained by assumptions of linearity and normal distribution. For managers, the predictive insights generated by ANNs can enhance green marketing strategies, improve customer segmentation and optimise CRM systems. Businesses can also tailor marketing efforts to high-intent consumers through targeted promotions while increasing awareness among lower-intent buyers. Additionally, ANN-based predictions can improve demand forecasting, optimise inventory and supply-chain management, and assess marketing campaign effectiveness by tracking changes in purchase intent. In summary, by leveraging ANNs, businesses can better identify key purchase drivers and forecast demand across various personal-care product categories.

Despite its contributions, the study acknowledges limitations, including potential sample and respondent biases due to its focus on Tāmaki Makaurau Auckland and the use of purposive sampling via online questionnaires. Responses collected from 110 participants may constrain the generalisability of the findings. Small datasets also increase the risk of overfitting, where the model captures patterns specific to the sample rather than generalisable trends. Therefore, future research should consider validating the model on larger and more diverse datasets to enhance the robustness and predictive reliability of the results. Expanding the sample across different demographic groups, industries or geographic regions would strengthen the model's external validity and practical applicability.

Future research could explore more diverse data-collection methods, such as face-to-face interviews or social media surveys, to enhance representativeness. Further refinement of model hyperparameters, including learning rates and loss functions, could also be explored to enhance predictive accuracy and adaptability. In addition, future research may re-examine the issue using larger or more diverse samples from various regions and segments to provide additional evidence regarding purchase intentions for green personal-care products.

References

Agmeka, F., Wathoni, R. N., & Santoso, A. S. (2019). The influence of discount framing towards brand reputation and brand image on purchase intention and actual behaviour in ecommerce. *Prodedia Computer Science*, e 161, 851–858. https://www.sciencedirect.com/science/article/pii/S1877050919319039?ref=pdf_download&fr=RR-2&rr=9520cd244dffd9b6

Albayrak, T., Aksoy, Ş., & Caber, M. (2013). The effect of environmental concern and scepticism on green purchase behaviour. *Marketing Intelligence & Planning*, *31*(1), 27–39. https://doi.org/10.1108/02634501311292902

Atienza, J. P., Arturo, M. R., Bautista, A., & Bonguit, P. G. (2022). Influence of environmental values on purchase intention of selected personal care products among consumers. *International Journal of Social and Management Studies*, *3*(5), Article 5. https://doi.org/10.5555/ijosmas.v3i5.198

Bigarella, C. A. P. (2024). *Influence of hyperparameters on the performance of neural network and random forest algorithms* [Master's thesis, University of Applied Sciences, Vienna]. https://epub.technikum-wien.at/obvftwhsm/content/titleinfo/10097682/full.pdf

Biswas, A., & Roy, M. (2015). Green products: An exploratory study on the consumer behaviour in emerging economies of the East. *Journal of Cleaner Production*, *87*, 463–468. https://doi.org/10.1016/j.jclepro.2014.09.075

Bonaccorso, G. (2018). *Mastering machine learning algorithms: Expert techniques to implement popular machine learning algorithms and fine-tune your models*. Packt Publishing Limited.

Chan, R. M. (2001). Determinants of Chinese consumers' green purchase behavior. *Psychology & Marketing*, *18*(4), 389–413. https://doi.org/10.1002/mar.1013

Chaudhuri, N., Gupta, G., Vamsi, V., & Bose, I. (2021). On the platform but will they buy? Predicting customers' purchase behavior using deep learning. *Decision Support Systems*, *149*, 113622. https://doi.org/10.1016/j.dss.2021.113622

Che In, F., & Ahmad, A. (2018). The effect of demographic factors on consumer intention to purchase green personal care products. In *Proceedings of INSIGHT 2018 1st International Conference on Religion, Social Sciences and Technological Education*. https://www.researchgate.net/publication/328494976_THE_EFFECT_OF_DEMOGRAPHIC_FACTORS_ON_CONSUMER_INTENTION_TO_PURCHASE_GREEN_PERSONAL_CARE_PRODUCTS

Chekima, B., Syed Khalid Wafa, S. A. W., Igau, O. A., Chekima, S., & Sondoh, S. L. (2016). Examining green consumerism motivational drivers: Does premium price and demographics matter to green purchasing? *Journal of Cleaner Production*, *112*, 3436–3450. https://doi.org/10.1016/j.jclepro.2015.09.102

Chen, N., & Yang, Y. (2020). The impact of customer experience on consumer purchase intention in cross-border E-commerce: Taking network structural embeddedness as mediator variable. *Journal of Retailing and Consumer Services*, *59*, 102344. https://doi.org/10.1016/j.jretconser.2020.102344

Chen, S., Wang, X., Zhang, H., & Wang, J. (2021). Customer purchase prediction from the perspective of imbalanced data: A machine learning framework based on factorization machine. *Expert Systems with Applications*, *173*, 114756. https://doi.org/10.1016/j.eswa.2021.114756

Chen, T. B., & Chai, L. T. (2010). *Attitude towards the environment and green products: Consumers' perspective.* ResearchGate. https://www.researchgate.net/publication/228500580_Attitude_towards_the_Environment_and_Green_Products_Consumers'_Perspective

Choudhury, N., Mukherjee, R., Yadav, R., Liu, Y., & Wang, W. (2024). Can machine learning approaches predict green purchase intention? A study from Indian consumer perspective. *Journal of Cleaner Production*, *456*, 142218. https://doi.org/10.1016/j.jclepro.2024.142218

Coriolis Research. (2022). *Creating beauty from New Zealand: Supporting investment in the emerging natural or bio-based cosmetics platform*. https://www.coriolisresearch.com/reports/coriolis-bio-siii-biocosmetics-100

Diamantopoulos, A., Schlegelmilch, B. B., Sinkovics, R. R., & Bohlen, G. M. (2003). Can socio-demographics still play a role in profiling green consumers? A review of the evidence and an empirical investigation. *Journal of Business Research*, *56*(6), 465–480. https://doi.org/10.1016/S0148-2963(01)00241-7

Fintz, M., Osadchy, M., & Hertz, U. (2022). Using deep learning to predict human decisions and using cognitive models to explain deep learning models. *Scientific Reports*, *12*(1), 4736. https://doi.org/10.1038/s41598-022-08863-0

Fishbein, M. A., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research.* Addison-Wesley.

Fisher, C., Bashyal, S., & Bachman, B. (2012). Demographic impacts on environmentally friendly purchase behaviors. *Journal of Targeting, Measurement and Analysis for Marketing*, *20*(3), 172–184. https://doi.org/10.1057/jt.2012.13

Ha, H., & Janda, S. (2012). Predicting consumer intentions to purchase energy-efficient products. *Journal of Consumer Marketing*, *29*(7), 461–469. https://doi.org/10.1108/07363761211274974

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Herrmann, A., Xia, L., Monroe, K. B., & Huber, F. (2007). The influence of price fairness on customer satisfaction: An empirical test in the context of automobile purchases. *Journal of Product & Brand Management*, *16*(1), 49–58. https://doi.org/10.1108/10610420710731151

Hustvedt, G., & Dickson, M. A. (2009). Consumer likelihood of purchasing organic cotton apparel: Influence of attitudes and self-identity. *Journal of Fashion Marketing and Management: An International Journal*, *13*(1), 49–65. https://doi.org/10.1108/13612020910939879

Irawan, R., & Darmayanti, D. (2011). *The influence factors of green purchasing behavior: A study of university students in Jakarta*. StudyModeResearch. https://www.studymode.com/essays/The-Influence-Factors-Of-Green-Purchasing-39444346.html#google_vignette

Kalliola, J., Kapočiūtė-Dzikienė, J., & Damaševičius, R. (2021). Neural network hyperparameter optimization for prediction of real estate prices in Helsinki. *Peer Journal of Computer Science*, *7*, e444. https://doi.org/10.7717/peerj-cs.444

Kim, S., & Seock, Y. (2009). Impacts of health and environmental consciousness on young female consumers' attitude towards and purchase of natural beauty products. *International Journal of Consumer Studies*, *33*(6), 627–638. https://doi.org/10.1111/j.1470-6431.2009.00817.x

Kumar, M. R., Pradeep, E., & Vakayil, S. (2022). Predicting the dynamics of customer repurchase in internet shopping: An artificial neural network (ANN) approach. *Academy of Marketing Studies Journal*, *26*(2). https://www.proquest.com/docview/2718660875/abstract/250E277901854205PQ/1

Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: An introduction to data mining.* John Wiley & Sons, Inc.

Larsson, S., & Arif Khan, M. (2012). *A study of factors that influence green purchase* [Master's thesis, Umeå Universitet]. https://www.diva-portal.org/smash/get/diva2:614312/FULLTEXT02.pdf

Lasuin, C. A., & Ng, Y. C. (2014). Factors influencing green purchase intention among university students. *Malaysian Journal of Business and Economics (MJBE)*, *1*(2), Article 2. https://jurcon.ums.edu.my/ojums/index.php/mjbe/article/view/116

Lee, K. (2008). Opportunities for green marketing: Young consumers. *Marketing Intelligence & Planning*, *26*(6), 573–586. https://doi.org/10.1108/02634500810902839

Leong, L-Y., Hew, T-S., Ooi, K-B., & Tan, G. W-H. (2019). Predicting actual spending in online group buying: An artificial neural network approach. *Electronic Commerce Research and Applications*, *38*, 100898. https://doi.org/10.1016/j.elerap.2019.100898

Mei, O. J., Ling, K. C., & Piew, T. H. (2012). The antecedents of green purchase intention among Malaysian consumers. *Asian Social Science*, *8*(13), 248. https://doi.org/10.5539/ass.v8n13p248

Moslehpour, M., Chaiyapruk, P., Faez, S., & Wong, W-K. (2021). Generation Y's sustainable purchasing intention of green personal care products. *Sustainability*, *13*(23), Article 23. https://doi.org/10.3390/su132313385

Mostafa, M. M. (2009). Shades of green: A psychographic segmentation of the green consumer in Kuwait using self-organizing maps. *Expert Systems with Applications*, *36*(8), 11030–11038. https://doi.org/10.1016/j.eswa.2009.02.088

Nayebi, A., Tipirneni, S., Foreman, B., Ratcliff, J., Reddy, C. K., & Subbian, V. (2022). Recurrent neural network based time-series modeling for long-term prognosis following acute traumatic brain injury. *AMIA Annual Symposium Proceedings*, *2021* (pp. 900–909). AMIA. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8861707/

Öhman, N. (2011). Buying or lying: The role of social pressure and temporal disjunction of intention assessment and behavior on the predictive ability of good intentions. *Journal of Retailing and Consumer Services*, *18*(3), 194–199. https://doi.org/10.1016/j.jretconser.2010.09.008

Pasini, A. (2015). Artificial neural networks for small dataset analysis. *Journal of Thoracic Disease*, *7*(5). https://doi.org/10.3978/j.issn.2072-1439.2015.04.61

Peleato, N. M. (2022). Application of convolutional neural networks for prediction of disinfection by-products. *Scientific Reports*, *12*(1), 612. https://doi.org/10.1038/s41598-021-03881-w

Prasad, B., & Ghosal, I. (2021). Forecasting buying intention through artificial neural network: An algorithmic solution on direct-to-consumer brands. *FIIB Business Review*, *11*, 231971452110461. https://doi.org/10.1177/23197145211046126

Prashar, S., Parsad, C., & Sai Vijay, T. (2015). Application of neural networks technique in predicting impulse buying among shoppers in India. *Decision*, *42*(4), 403–417. https://doi.org/10.1007/s40622-015-0109-x

Prashar, S., Vijay, T. S., & Parsad, C. (2016). Predicting online buying behavior among Indian shoppers using a neural network technique. *International Journal of Business & Information*, *11*(2), 175–198.

Qomariah, A., & Prabawani, B. (2020). The effects of environmental knowledge, environmental concern, and green brand image on green purchase intention with perceived product price and quality as the moderating variable. *IOP Conference Series: Earth and Environmental Science*, *448*(1), 012115. https://doi.org/10.1088/1755-1315/448/1/012115

Rahmi, D., Rozalia, Y., Chan, D., Anira, Q., & Lita, R. (2017). Green brand image relation model, green awareness, green advertisement, and ecological knowledge as competitive advantage in improving green purchase intention and green purchase behavior on creative industry products. *Journal of Economics, Business & Accountancy Ventura*, *20*. https://doi.org/10.14414/jebav.v20i2.1126

Sabeen, Z., Arshad, F., & Ayub, N. (2022). Impact of brand image on green purchase intentions: Mediating role of CSR perceptions. *NICE Research Journal*, 61–88. https://doi.org/10.51239/nrjss.vi.351

Samarasinghe, D. S. R. (2012). Green consumerism individual's ethics and politics as predictors of pro-environmental behaviour. *Delhi Business Review*, *13*(1), 41–48. https://doi.org/10.51768/dbr.v13i1.131201204

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2022). Partial least squares structural equation modeling. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of market research* (pp. 587–632). Springer International Publishing. https://doi.org/10.1007/978-3-319-57413-4_15

Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology*, *39*(2).

Solomon, M., Bamossy, G. J., Askegaard, S., & Hogg, M. K. (2006). *Consumer behavior: A European perspective* (3rd edition). Prentice Hall.

Statista. (2024, March). *Beauty & personal care – New Zealand*. https://www.statista.com/outlook/cmo/beauty-personal-care/new-zealand

Swaminathan, V., Sorescu, A., Steenkamp, J-B. E. M., Clayton, T., O'Guinn, G., & Schmitt, B. (2020). Branding in a hyperconnected world: Refocusing theories and rethinking boundaries. *Journal of Marketing*, *84*(2). https://doi.org/10.1177/0022242919899905

Taunton, E. (2024, April 21). Consumer NZ accuses popular beauty brands of greenwashing. *Stuff*. https://www.stuff.co.nz/business/350252144/consumer-nz-accuses-popular-beauty-brands-greenwashing

Tsarenko, Y., Ferraro, C., Sands, S., & McCleod, C. (2012). Environmentally conscious consumption: The role of retailers and peers as external influences. *Journal of Retailing and Consumer Services*, *20*(3), 302–310. https://doi.org/10.1016/j.jretconser.2013.01.006

Van der Werff, E., Steg, L., & Keizer, K. (2013). The value of environmental self-identity: The relationship between biospheric values, environmental self-identity and environmental preferences, intentions and behaviour. *Journal of Environmental Psychology*, *34*, 55–63. https://doi.org/10.1016/j.jenvp.2012.12.006

Vinoth, S. (2023). Impact of environmental consciousness and knowledge on purchase intention and behavior of green consumers. *IUP Journal of Marketing Management*, *22*(4), 91–108.

Vora, D. R., & Bhatia, G. S. (2023). *Python machine learning projects: Learn how to build machine learning projects from scratch* (English edition). BPB Publications.

Wibisono, D. (2003). *Riset bisnis panduan bagi praktisi & akademisi*. Gramedia Pustaka Utama.

Witek, L., & Kuźniar, W. (2021). Green purchase behavior: The effectiveness of sociodemographic variables for explaining green purchases in emerging market. *Sustainability*, *13*(1), Article 1. https://doi.org/10.3390/su13010209

Wu, S-I., & Chen, Y-J. (2014). The impact of green marketing and perceived innovation on purchase intention for green products. *International Journal of Marketing Studies*, *6*(5), 81. https://doi.org/10.5539/ijms.v6n5p81

Yin, S., Wu, L., Du, L., & Chen, M. (2010). Consumers' purchase intention of organic food in China. *Journal of the Science of Food and Agriculture*, *90*(8), 1361–1367. https://doi.org/10.1002/jsfa.3936

Zheng, B., Thompson, K., Lam, S. S., Yoon, S. W., & Gnanasambandam, N. (2013). *Customer's behaviour prediction using artificial neural networks*. ResearchGate. https://www.researchgate.net/publication/287367042_Customers'_behavior_prediction_using_artificial_neural_network

Zhou, A., Ren, K., Li, X., & Zhang, W. (2019). MMSE: A multi-model stacking ensemble learning algorithm for purchase prediction. *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)* (pp. 96–102). IEEE. https://doi.org/10.1109/ITAIC.2019.8785711