Computers in Industry 144 (2023) 103790

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.sciencedirect.com/journal/computers-in-industry

Spec guidance for engineering design based on data mining and neural networks

Seyoung Park [a], Junegak Joung [b], Harrison Kim [a,*]

[a] Department of Industrial and Enterprise Systems Engineering, University of Illinois Urbana-Champaign, Urbana, IL, 61801, USA
[b] Division of Interdisciplinary Industrial Studies, Hanyang University, Seoul, 04763, Republic of Korea

ARTICLE INFO

Keywords:
Data mining
Explainable neural networks
Online reviews

ABSTRACT

Recently, many studies on product design have been utilizing online data. They analyze user-generated online data and draw design implications. However, most of them provide customers' tendency for feature categories rather than spec ranges for sub-features, which are crucial in industrial applications. This paper proposes an approach based on data mining and neural networks to extract spec guidance for engineering design from online data. First, product sub-features are extracted from online data, and customer choice sets are constructed. Next, a neural network choice model is trained based on these choice sets. Finally, the model is interpreted by SHAP (SHapley Additive exPlanations). In the final stage, this study proposes a method for analyzing the obtained SHAP values to draw new design implications. The suggested approach was tested on smartphone review data, and the result provides a set of recommended spec values for each sub-feature. The resultant spec guidance can help companies design a product with spec configuration preferred by customers.

1. Introduction

With the increasing amount of online user-generated data and the development of data analysis techniques, many researchers have been utilizing online data in the product design area. They propose various methods for analyzing online customers and discovering design implications for new products. Some of them focus on numeric parts, such as review numbers and ratings (Chevalier and Mayzlin, 2006; Sun, 2012; Chong et al., 2017). Others analyze textual data using NLP (Natural Language Processing) techniques. These studies extract customer needs (Ali et al., 2020; Jiao and Qu, 2019) and preferences (Wang et al., 2020) for product attributes, discover new product features (Goldberg and Abrahams, 2022), and identify changes in usage patterns (Zhou et al., 2019). Although these studies provide some design implications, they have limitations in practical design applications.

In industry, a company releases a new product through various tasks, including product design, development, component sourcing, assembly, and production (Pahl et al., 2007). As can be inferred from 'component sourcing and assembly', the company manufactures the product by combining multiple components. Therefore, a general product feature consists of several sub-features. For example, in smartphones, the camera feature includes two parts—rear and front camera modules. Also, a part is described by its features. Specifically, screen components have different sizes, resolutions, and types. In this paper, the term 'sub-feature' means both the part and part features. In the product design stage, the range of specs for sub-features must be determined so that the company can start sourcing necessary parts. However, most studies focus on product feature categories instead of specific sub-features.

A choice model (McFadden, 1986), which analyzes the customer's behavior within given alternatives, can be a solution to this problem. A logit model based on random utility is the widely used choice model in market-based engineering design (Chen et al., 2012). It analyzes the relationship between relevant attributes and customers' choices. Since the model considers product sub-features, it can draw design implications on them. However, the logit model has limitations in that it cannot capture non-linearity between input terms and customers' choices. The resultant design implications only provide customers' tendency for sub-features, not their preferences for spec ranges. For example, a negative coefficient for the price means that customers prefer lower prices. However, the reality can be that customers do not prefer too low prices due to out-of-date specs, and they may want middle-range priced products. The previous model could not capture this nonlinear property. Recently, many studies on choice models adopted neural networks (NN) to capture non-linearity (Lee et al., 2018; Nam and Cho, 2020; Sifringer et al., 2018). The NN model identifies linear and nonlinear relationships between input and output (customer choice), so it provides higher prediction accuracy. However, the NN model also

* Corresponding author.
E-mail addresses: seyoung7@illinois.edu (S. Park), june30@hanyang.ac.kr (J. Joung), hmkim@illinois.edu (H. Kim).

https://doi.org/10.1016/j.compind.2022.103790
Received 30 November 2021; Received in revised form 24 August 2022; Accepted 27 September 2022
Available online 3 November 2022
0166-3615/

Table 1
Literature review (Data-Driven Design).

Literature	Method	Automated data collection	Feature importance	Spec guidance
Wang and Chen (2015), Suryadi and Kim (2019)	Choice Model	✓	✓	
Archak et al. (2007)	Linear Regression	✓	✓	
Joung and Kim (2021)	NN + SHAP	✓	✓	
Tuarob and Tucker (2015)	Sentiment Analysis	✓	✓	
Zhou et al. (2019), Bi et al. (2019)	Kano Model	✓	✓	
Our study	NN + SHAP	✓	✓	✓

has limitations in that it is difficult to interpret. As a result, there is a trade-off between accuracy and interpretability.

This research addresses the above limitations by interpreting the NN choice model with SHAP (SHapley Additive exPlanations). A method for analyzing the obtained SHAP values is proposed to draw new design implications. The result shows that the suggested approach provides spec guidance, i.e., recommended spec ranges for product sub-features.

The rest of the paper is organized as follows. In Section 2, previous works relevant to this research will be presented. In Section 3, the detailed process of the proposed methodology will be explained. Section 4 will show the simulation of the proposed methodology on actual data collected from online sources. The simulation result will be presented in Section 5. Section 6 will evaluate the performance of the choice model and discuss managerial implications of the result. Finally, Section 7 will summarize the whole idea of the research and discuss future works.

2. Related works

In this section, three main topics relevant to this research will be presented. The first topic is data-driven design. Previous studies in this field and their limitations will be discussed. In the second topic, research about customer choice models will be introduced. The final topic is neural networks. In this topic, the state-of-the-art method for explaining neural networks will be discussed.

2.1. Data-driven design

Data-Driven Design can be defined as the design based on the use of data science algorithms supporting a specific phase of the product development process (Bertoni, 2020). Product designers can make smart decisions by analyzing the data that brings new opportunities to enhance the production efficiency and product competitiveness (Feng et al., 2020). Chiarello et al. (2021a) conducted a literature review at the intersection between data science and engineering design and identified the challenges to be tackled to maximize the synergies between the two fields. They pointed out that while data-driven methods are common practices in medicine, engineering, defense, and other safety–critical systems, using these methods in consumer product design is a relatively recent phenomenon. In this section, recent studies about consumer product design based on data mining are presented.

Regarding data sources, various types of sources are used in engineering design research (Chiarello et al., 2021b). The dominant data source is human interactions such as group discussion and interviews, but it requires much time and cost. As an alternative, recent studies utilize online data from web sources. The most powerful benefit of online data is that we can collect a large amount of data in a short period with little cost. This study also used online data, and relevant studies are presented below.

The basic step of online data analysis for consumer product design is product feature extraction. Various NLP (Natural Language Processing) techniques such as association mining (Hu and Liu, 2004; Spreafico and Spreafico, 2021), Word2Vec (Mikolov et al., 2013; Giabelli et al., 2022), and LDA (Latent Dirichlet Allocation) (Blei et al., 2003) were applied to online data to identify feature-related words mentioned by customers. Based on these feature terms, researchers proposed diverse approaches to draw design implications.

One of the approaches is to analyze the relationship between product features and desirable outputs such as high sales ranking and high star ratings. Wang and Chen (2015) and Suryadi and Kim (2019) constructed choice sets using online user data and analyzed the influence of product features on the customer's purchase decision. Archak et al. (2007) analyzed online review data and modeled product demands as a linear function of product attributes. The coefficients in the regression result show the effect of each product feature on sales ranking. Joung and Kim (2021) suggested a methodology to identify the importance of product features based on review ratings. They built a neural network model, where the input data is the sentiment scores, and the output data is the customer's rating for the product. By analyzing the trained model, the authors obtained influence scores of product features on the ratings.

Another approach is to observe the changes in customer sentiments for features. Tuarob and Tucker (2015) detects feature-related terms in Twitter mentions by utilizing a bootstrapping algorithm. Based on the extracted keywords, the authors compared the customers' sentiments for the same feature of two consecutive products. For example, they observed how the ratio of positive sentiments for the 'screen' feature was changed between iPhone4 and iPhone5. The result gives clues about which features to improve or maintain for the next generation of products. Zhou et al. (2019) and Bi et al. (2019) presented a Kano model based on online data analysis. After identifying feature words in online reviews using LDA, the authors detected corresponding sentiments. Then, by analyzing the polarity and score of sentiments, the authors divided all features into four categories—must-be, performance, excitement, and reverse. The characteristics of categories are defined in the Kano model, so the result provides product design strategies for features in each category.

Table 1 summarizes the studies mentioned above. They proposed various data-driven design methods using online data. However, the design implications derived in the results are about feature importance, which is insufficient for industrial applications. This study is distinct from other data-driven design research by providing spec guidance for sub-features based on customer preference from online data. The following subsections explain the background of the proposed approach.

2.2. Customer choice model

Discrete choice analysis (DCA) uses the principle of utility maximization. In DCA, a decision-maker selects the alternative with the highest utility among available options (Ben-Akiva and Lermna, 1985). Regarding utility, the concept of random utility is adopted assuming that the individual's true utility consists of a deterministic part V and a random disturbance ϵ as shown in Eq. (1). U_{ni} represents the utility of customer n obtained by purchasing product i. The deterministic part V_{ni} can be represented as a function of observable independent variables such as product features (Chen et al., 2012). In Eq. (1), V_{ni} is the weighted sum of product features where x_{ik} represents the spec of feature k of product i. β_{nk} is the importance that customer n has for feature k. The probabilistic distribution assumed for ϵ determines the analytical relation between the choice probability and the observed component (V_{ni}) of the utility functions, and hence the type of discrete choice model (Ramanujam and Balakrishnan, 2011). A multinomial logit (MNL) model yields the choice probability shown in Eq. (2). Pr_{ni} indicates the probability of customer n choosing product i. J represents a set of products, and V_{nj} comes from Eq. (1).

$$U_{ni} = V_{ni} + \epsilon_{ni} = \sum_{k} \beta_{nk} x_{ik} + \epsilon_{ni} \tag{1}$$

$$Pr_{ni} = \frac{e^{V_{ni}}}{\sum_{j \in J} e^{v_{nj}}} \tag{2}$$

There have been a few studies about online data-based choice models (Wang and Chen, 2015; Suryadi and Kim, 2019). Suryadi and Kim (2019) suggested a probabilistic sampling method called normalized sampling. Their methodology consists of four stages. First, a dataset of laptop products is collected from Amazon.com. The data contains product specifications and customer reviews. The products are clustered based on the collected spec data. The second stage is customer clustering. In this stage, product reviews are used for defining customer attributes. The authors identified product feature words in the review data using word embedding and extracted sentiments associated with the feature words. These sentiments become attributes that characterize customers. Then, reviewers (customers) are clustered by X-means clustering based on these customer attributes. The next stage is generating sampling distributions. Based on the clustering results from the first two stages, the method creates the sampling distribution for each customer cluster. Unlike the random distribution where all products are uniformly weighted, products have different weights for different customer clusters. The product weight is determined based on the sales record within the customer cluster. In the final stage, choice sets are constructed based on this normalized distribution. The constructed choice sets become the input data for the MNL model.

In this study, random sampling and normalized sampling methods are used for choice set construction. The detailed process will be explained in Section 3.

2.3. Explainable neural networks

An artificial neural network (NN) has become a popular model in many disciplines (Abiodun et al., 2018). In the manufacturing industry, NN provides methods and tools to address issues due to the massive data scale (Dekhtiar et al., 2018). In DCA, there have been studies adopting NN (Lee et al., 2018; Sifringer et al., 2018; Nam et al., 2017) to enhance the performance of choice models. They compare the logit and NN model and show that the NN model gives higher accuracy in predicting choices. However, the NN model has a downside, the difficulty of interpretation. None of the above researches provides explanations for the resultant NN model. Adopting an NN model was a trade-off between accuracy and explainability.

Lundberg and Lee (2017) proposed SHAP to address this problem. SHAP is an approach based on Shapley values in a game theory introduced by Roth (1988). As 'Additive' in the name implies, SHAP uses an additive feature attribution method shown in Eq. (3). $g(x')$ approximates actual output $f(x)$. x'_i is a binary variable that maps to the original input x_i, and M is the number of input features. ϕ_i represents the effect that attribute x'_i has on the output $g(x')$. Therefore, the method approximates the output $f(x)$ by attributing an effect ϕ_i to each feature x'_i and summing all influences.

$$f(x) = g(x') = \phi_0 + \sum_{i=1}^{M} \phi_i x'_i \tag{3}$$

In Lundberg and Lee (2017), Lundberg & Lee mentioned three desirable properties for a solution in the class of additive feature attribution method: local accuracy, missingness, and consistency. The only possible model for Eq. (3) that satisfies all three properties is shown in Eq. (4). z' is a subset of input variables x', and $|z'|$ is the number of non-zero entries in z'. It compares the output value for subset z' and the value when attribute i is excluded from subset z', i.e., $f_x(z') - f_x(z' \setminus i)$. The model evaluates the difference in the output made by including input i for all combinations of features.

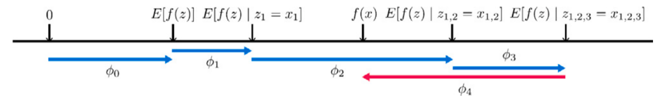

Fig. 1. SHAP diagram (Lundberg and Lee, 2017).

$$\phi_i = \sum_{z' \subseteq x'} \frac{|z'|!(|M| - |z'| - 1)!}{|M|!} [f_x(z') - f_x(z' \setminus i)] \tag{4}$$

As a solution to Eq. (4), Lundberg & Lee proposed SHAP values. They are the Shapley values of a conditional expectation function of the original model. Specifically, in Eq. (4), $f_x(z') = f(h_x(z')) = E[f(z) \mid z_S]$ where S is the set of non-zero indexes in z'. Fig. 1 shows how SHAP values explain an NN model. The figure illustrates the change in the expected model prediction when conditioning on each feature. Specifically, ϕ_i represents the influence of input x_i on the model output. The direction of arrows implies that x_1, x_2, and x_3 have positive influence ($\phi_i > 0$) on the output, and x_4 has negative impact ($\phi_i < 0$) on the output value.

Recent studies have adopted SHAP to interpret the resultant NN models (Joung and Kim, 2021; García and Aznarte, 2020; Sujith et al., 2020). They explain the NN model by graphically representing SHAP values or analyzing statistical properties of the SHAP values. This paper applies SHAP to the NN choice model to interpret customers' decisions. A method for analyzing SHAP values is proposed to draw novel design implications from the NN choice model. Specifically, the result provides a range of spec values for product sub-features that customers may prefer or accept. The obtained design implications will be discussed in Section 5.

3. Methodology

The proposed methodology consists of three stages, as shown in Fig. 2. In the first stage, choice sets are constructed based on online data. In the second stage, a neural network choice model is trained based on the choice sets. Finally, the choice model is interpreted based on SHAP, and the SHAP values are further analyzed to draw design implications.

3.1. Choice sets

In this stage, the method collects two types of data: (i) customer reviews for products; (ii) product attributes. Both are obtained online. Then customer choice sets are constructed based on the collected data.

3.1.1. Collecting data

Customer reviews can be collected from e-commerce websites (similarweb, 2022) such as Amazon[1] and eBay.[2] Regarding product attributes, it is recommended to collect data on all attributes that influence consumers' purchasing decisions (Gowharji and Whitefoot, 2021). The proposed methodology identifies features of customer interest by analyzing customer reviews. Among various approaches for feature extraction (Archak et al., 2007; Joung and Kim, 2021; Tuarob and Tucker, 2015), the method based on Word2Vec (Park and Kim, 2022) was adopted because the sub-feature level is required in engineering design. For example, the 'screen' feature consists of multiple sub-features such as screen size, screen resolution, screen type, etc. In engineering design, decisions are made on these sub-features, not the feature category.

[1] Available: https://www.amazon.com.

[2] Available: https://www.ebay.com.

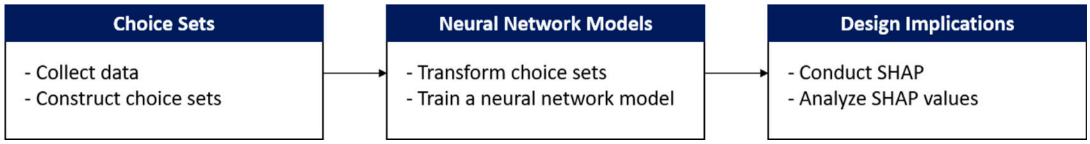

Fig. 2. Overview of the proposed methodology.

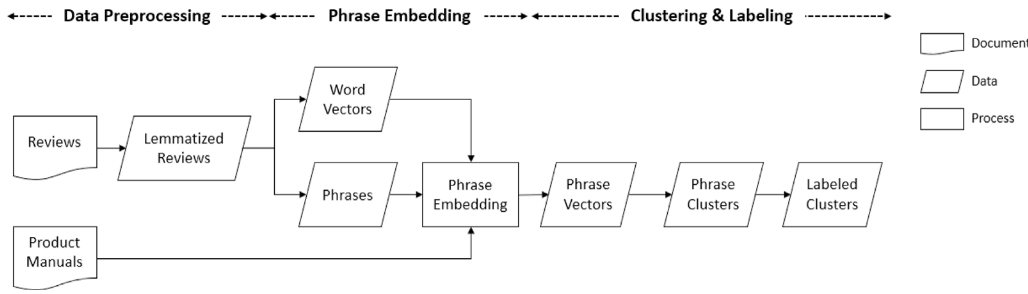

Fig. 3. Sub-feature extraction.

The method of Park and Kim (2022) extracts sub-features while most methods extract feature categories.

It is considered necessary to summarize the method for sub-feature extraction while the details are available in Park and Kim (2022). Fig. 3 shows the overall process of sub-feature extraction. In data preprocessing, online customer reviews and product manual documents are collected and preprocessed. In specific, special characters are removed, and all punctuations are replaced with a period. Upper case letters are transformed into lower cases, and all words are lemmatized. Stopwords are not removed as it affects the phrase extraction to be performed in this step.

For phrase embedding, the lemmatized words from the review data are embedded into vectors. Next, the method extracts phrases in the review data. Among the extracted phrases, the ones that contain words never mentioned in the product manuals are removed. The remaining phrases are embedded into a vector space by Eq. (5) where W_i denotes word i and $\vec{W}_i$ represents a vector for word i. $Freq(W_i)$ means the frequency of word i in the manual documents.

$$\text{Phrase} = \alpha_i \times \vec{W}_i + \alpha_j \times \vec{W}_j$$
$$\alpha_i = \frac{\text{Freq}(W_i)}{\text{Freq}(W_i) + \text{Freq}(W_j)} \tag{5}$$

In clustering & labeling, phrase vectors are grouped by two clustering methods—HDBSCAN (Mclnnes et al., 2017) and spectral clustering (Cortesy et al., 2012). The resultant clustered are labeled based on the TF (Term Frequency) analysis of cluster members. The clusters with feature-relevant labels are selected, and the phrases in them represent sub-features mentioned by customers.

This study collects product specifications for all the identified sub-features. The spec data can be collected online. The online commerce websites provide product spec information. Also, there exist websites specialized in certain product categories. For example, GSMArena.com and phoneArena.com are focused on smartphone products and provide detailed spec data for each product.

3.1.2. Constructing choice sets

The traditional methods for collecting choice sets are surveys and interviews, but they require much time and cost. As an alternative, this study constructs customer choice sets by utilizing online data with little time and cost. Fig. 4 shows the process of choice set construction for each reviewer. Let us assume that the first reviewer purchased product $P12$. Then, it is assigned to the choice set, and the choice value of $P12$ becomes 1. Next, the remaining alternatives are selected from product candidates, which include all products in the review data except $P12$. This study adopts two sampling methods for alternative selection—random sampling and normalized sampling. The random sampling method assumes a uniform distribution, i.e., all products have an equal probability of being selected as an alternative. The normalized sampling method (Suryadi and Kim, 2019) assumes a non-uniform distribution. As explained in Section 2.2, the method assigns a different probability to each product based on the sales record. In Fig. 4, $P42$ and $P35$ are selected completing the choice set. Since the choice always goes to the first alternative, the completed choice set should be shuffled.

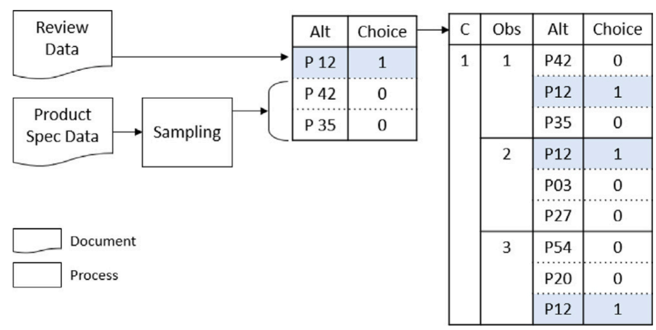

Fig. 4. Constructing choice sets.

While the previous studies (Wang and Chen, 2015; Suryadi and Kim, 2019) made one choice set for each customer, this study constructs multiple-choice sets for each reviewer. In surveys, the questionnaire presents more than one choice set for each respondent (WHO, 2012). So, in the survey data, each choice set is managed by an observation ID, which is different from a customer ID. Referring to the survey questionnaire, this study constructs three observations for each reviewer, and mange choice sets by customer Id (C) and observation ID (Obs), as shown in the rightmost table in Fig. 4. The process is repeated for all reviewers. Then, the resultant choice sets are divided into the train and test sets. The former is for training a choice model, and the latter is for evaluating the performance of the trained model.

3.2. Neural network models

In this stage, the method transforms the choice sets from the previous section to make them fit for neural network input. Then a neural network model is trained based on these choice sets.

Table 2
Example of SHAP results.

Input data				SHAP values				Output	C
x_1	x_2	..	x_k	ϕ_1	ϕ_2	..	ϕ_k	$\sum \phi_i$	
0.3	0.5	..	1.0	0.7	−1.6	..	−0.4	−0.3	0
0.3	1.0	..	0.8	1.4	−0.8	..	0.5	0.6	1
1.0	0.5	..	0.5	0.6	−1.2	..	0.3	−0.3	0

3.2.1. Transforming choice sets

The MNL and NN model have different structures, so the choice sets need to be transformed to train an NN model. In MNL, the input data is each observation in Fig. 4, and the model calculates the product utility for each alternative. The dimension of the input data is (set size) × (# product attributes). In this study, the choice set contains three alternatives, and the product has eleven attributes. Therefore, the input for MNL has a size of 3 × 11. On the other hand, the NN model analyzes the features of all products in a choice set. For this, product attributes in each choice set are merged into a one-dimensional array. Specifically, 3 × 11 data is transformed to a 1 × 33 array. The output of the NN model is a customer's decision in a choice set, and the decision result is also merged into a one-dimensional array. For example, the choice result of Obs 1 in Fig. 4 is transformed to [0, 1, 0].

3.2.2. Training a neural network model

In this study, an NN with one hidden layer is used for training a choice model. The previous studies about NN choice models used different network structures. One of the commonly used structures is the three layers network with one hidden layer (Lee et al., 2018; Sifringer et al., 2018; Alwosheel et al., 2018). Spec values are normalized for each attribute. For example, the largest screen size among the entire products becomes 1, and the smallest screen size becomes 0. Regarding sample size, Alwosheel et al. (2018) tested the required sample size when using NN for discrete choice analysis. They suggested using a minimum sample size of fifty times the number of weights in the NN model. After training, a neural network choice model is obtained. The performance of the NN model will be evaluated by comparing it with the multinomial logit model. The result will be discussed in Section 6.

3.3. Design implications

In this stage, SHAP (SHapley Additive Explanations) is conducted on the trained NN model. And then, the method analyzes the resultant SHAP values to draw design implications.

3.3.1. Conducting SHAP

As explained in Section 2.3, SHAP analyzes the influence of input values on the output. Since it calculates the effect of each input value, the number of SHAP values is the same as the dimension of the input layer. Table 2 shows the general form of the SHAP result. When the input data is k-dimensional, SHAP produces k SHAP values for each input data. These values are summed up to the output, and this output value determines the expected choice.

The SHAP library in PYTHON provides diverse functions to visualize SHAP results (Lundberg, 2018), and one of them is shown in Fig. 5. In this graph, the color bar on the right indicates feature values. The colors with high saturation represent comparably high spec values. If the color is close to white (the lowest saturation), it means a comparatively low spec value. The x-axis represents the SHAP value, the impact on model output. The y-axis lists product features in order of greatest influence. In the case of the choice model, this graph shows the relationship between product specs and customer choices. For example, the lower the battery capacity of product 1 (our product), the greater the positive influence on the model output (customer choice). And the lower the battery capacity of product 2 or 3 (competitors), the greater the negative effect on model output. Although this graph shows the tendency for feature specs, it is insufficient for practical application. As mentioned in Section 1, product designers need to determine the spec values of sub-features because a product consists of multiple parts. They need guidance for spec ranges rather than a tendency for feature specs.

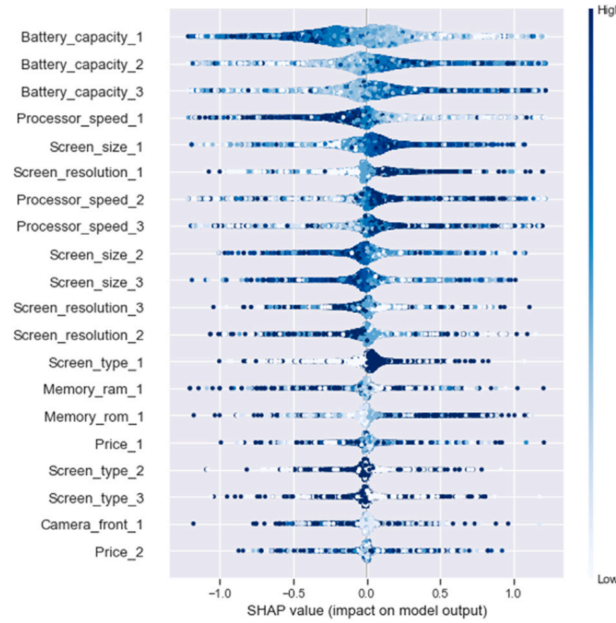

Fig. 5. SHAP values.

3.3.2. Analyzing SHAP values

This study proposes a method for analyzing SHAP values to derive design implications about spec ranges and customer decisions. First, the structure of the SHAP result needs to be modified. When an NN model predicts the output values such as rating and price, SHAP produces one set of SHAP values. If the NN model is for classification with N classes, SHAP generates N datasets. Since the choice model is a type of classification, this study obtains N datasets of SHAP values. The tables on the left of Fig. 6 are the resultant datasets for the choice model with three alternatives. In the table, x_{ij} is the spec value for feature j of product i, and ϕ_{ij} represents the corresponding SHAP value for x_{ij}. In each dataset, SHAP values show the influence of input data on the decision for the Nth alternative. In specific, the first dataset shows the influence of spec values on whether the customer selects the first alternative. The second dataset shows the impact of product specifications on whether the customer chooses the second option and the same for the third set. The result shows the misalignment in the datasets where the target alternative highlighted in blue is located on a different column in each table. Therefore, the resultant datasets need to be properly modified before any further analysis. This study rearranges the columns in each dataset to place the target alternative in the first column. It also organizes SHAP values according to the order of products. And then, the modified SHAP datasets are concatenated, as shown on the right of Fig. 6. The first column is the test design in that the model determines the customer's purchase of this product, and the following two columns are competitors.

Next, SHAP values in the modified dataset are analyzed. For each feature of the test design, spec values are sorted in ascending order.

Table 3
Choice sets.

Obs	Alt	S_{size}	S_{resol}	S_{type}	P_{speed}	P_n	M_{ram}	M_{rom}	C_{front}	C_{rear}	B_{cap}	$Price$	Choice
1	1	6.2	2	2	1.8	8	4	64	8	12	3000	299.99	0
	2	5.8	3	3	2.8	8	4	64	8	12	3000	499.99	0
	3	6.1	3	3	2.8	8	8	128	10	12	3400	649.99	1
2	1	6.3	1	2	2.0	8	3	32	13	13	3500	149.99	0
	2	6.1	3	3	2.8	8	8	128	10	12	3400	649.99	1
	3	5.8	1	2	1.6	8	2	32	5	8	3000	124.96	0

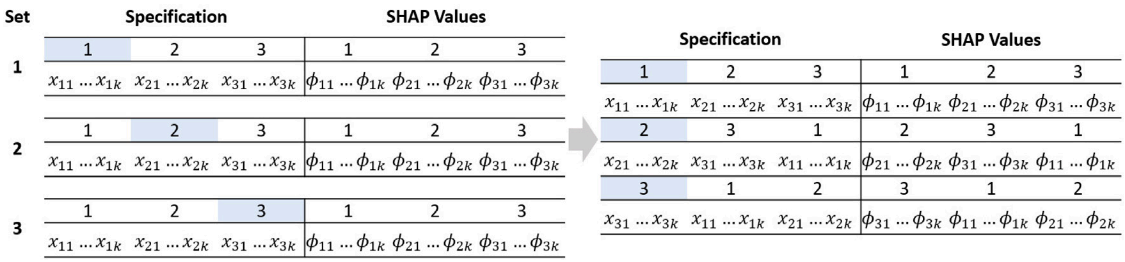

Fig. 6. SHAP result modification.

Then, for each spec value, the mean SHAP values are calculated. By analyzing the spec values and corresponding SHAP mean values, design implications for the spec range are obtained.

4. Case study

The proposed methodology was tested on the smartphone for two reasons: (i) It is a highly integrated product with multiple sub-features; (ii) With 85% penetration rate in US (O'Dea, 2021), most people are familiar with product features.

4.1. Choice sets

This study collected review data from Amazon.com. First, the target products were selected from the top 100 products in the cell phone category in Amazon. After removing non-smartphone products and those with no reviews, 58 products remained. Then, review data was collected by using a PYTHON package. This study used BeautifulSoup, but other packages or methods can be used. The raw data includes both verified and unverified purchases. For the authenticity of the data, only the reviews marked as 'verified purchase' were selected and used in this study. The final data contains 25,340 reviews about 58 products, and the reviews were written between May 2017 and July 2020. Products of the same category have common features, so the study collected product manual documents for six smartphones: Samsung Galaxy fold, Galaxy S10, Apple iPhone, OnePlus 7T, Xiaomi Mi, ZTE Blade Z Max. The collected datasets were cleaned and lemmatized using the Spacy library in PYTHON. Based on these two datasets, this study extracted 11 sub-features mentioned by customers: {Screen: [size, resolution, type], Processor: [speed, count], Memory: [ram, rom], Camera: [front, rear], Battery: [capacity], Price}. Details are presented in Table A.9 in Appendix A. For product specifications, this study collected data for 58 products from phoneArena.com and GSMArena.com. Although Amazon provides spec data in the product description, some sub-features are missing. The two websites mentioned above are focused on mobile devices and provide information for all the sub-features.

Next, customer choice sets were constructed based on the collected reviews and product specs. This study used a set size of 3 referring to the previous research (Wang and Chen, 2015; Suryadi and Kim, 2019). Three choice sets were constructed for each customer by the random sampling and utility sampling methods. Then, the entire choice sets were divided into training sets and test sets. The choice model is trained and evaluated by 5-fold validation, so the ratio of the training and testing sets is 8:2. Table 3 shows the choice sets from the random sampling method. The 'Obs' column shows the choice set ID, and each choice set has three different alternatives. In the 'Choice' column, the product purchased by a customer is marked as 1.

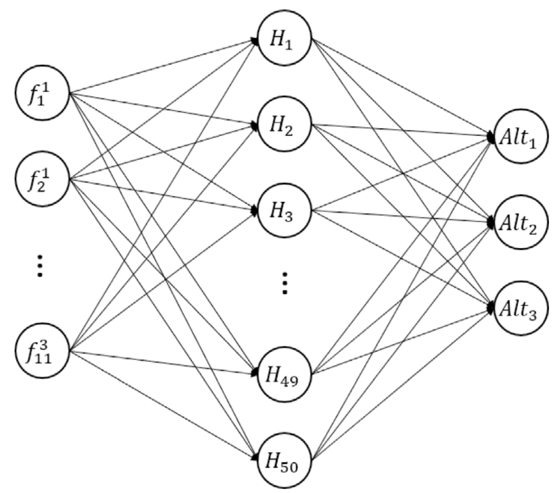

Fig. 7. The neural network structure.

4.2. Neural network choice models

In this section, an NN choice model was trained based on the previously constructed choice sets. Regarding the model structure, this study adopts an NN with one hidden layer, as shown in Fig. 7. With 11 product attributes and three alternatives, the size of the input layer in the NN model is 33. With three available choices in each choice set, the size of the output layer is 3. In training, this study conducted a grid search that evaluates all combinations of hyperparameters. The list of parameters tested in this research is presented in Table B.10 in Appendix B. Those with the highest prediction accuracy are as follows: The model was trained for 500 epochs with a batch size of 100. The optimizer is Adam, with a learning rate of 0.001. The initializer is GlorotNormal, and the number of neurons in the hidden layer is 50. The model was implemented using TensorFlow Keras in PYTHON.

This study also tested a deep neural network (DNN) consisting of 10 hidden layers with a 0.1% dropout rate and the same hyperparameters, but the DNN did not show better performance (Table C.11). Similar

Table 4
SHAP dataset after modification.

Product features (Input values)						SHAP values					
S^T_{size}	S^T_{resol}	S^T_{type}	…	$B^{C_2}_{cap}$	$Price^{C_2}$	S^T_{size}	S^T_{resol}	S^T_{type}	…	$B^{C_2}_{cap}$	$Price^{C_2}$
6.1	1	2	…	1821	279.99	−0.0004	−0.0469	−0.0166	…	−0.3543	−0.0310
5.5	2	3	…	2659	477.07	−0.0396	0.0029	0.0036	…	−0.2051	0.0086
6.6	2	1	…	3340	229.5	0.1152	0.0089	−0.1509	…	0.0338	−0.0098
6.7	2	3	…	3179	624.95	0.1038	−0.0044	−0.0049	…	−0.0075	0.0300
4.7	1	2	…	3500	149.99	−0.1895	−0.1166	0.0136	…	0.1910	−0.0358

results were presented in relevant studies (Lee et al., 2018; Joung and Kim, 2021). This study also trained an MNL model using the Pylogit package in PYTHON to compare the performance of choice models. The Pylogit requires more RAM resources as the data size increases. This study trained the models on the computer with 32 GB RAM, and the Pylogit could take up to 60,000 choice sets. The performance of the models is measured by 5-fold validation and will be evaluated in Section 6.

4.3. Design implications

The resultant NN model was analyzed by the DeepExplainer method of the SHAP package in PYTHON. For 60,000 choice sets with three alternatives, 180,000 sets of SHAP values were obtained. After the modification process, the values were aligned, as shown in Table 4. On the top row, the superscript T denotes the target design, and C means the competitor. The subscript indicates the sub-features. For example, S^T_{size} means the screen size of the target design, and $B^{C_2}_{cap}$ means the battery capacity of the second competitor.

Next, the SHAP values were analyzed for each product attribute of the target design. Among 11 product attributes shown in Table A.9, the screen size has 15 spec values ranging from 4.7“ to 6.8”. The screen resolution and screen type are ordinal variables with three classes. The CPU speed has 12 specs between 1.4 GHz and 2.8 GHz. The CPU count means the number of cores, and the specs include Quad-core, Hexa-core, and Octa-core. Regarding the memory feature, RAM has 7 specs between 1.5 GB and 12.0 GB, and ROM has 6 specs ranging from 16 GB to 1 TB. The front camera has 11 specs between 2 MP and 32 MP. The rear camera also has 11 specs between 5 MP and 108 MP. The battery capacity ranges from 1821 mAh to 5260 mAh with 29 specs. Finally, the price ranges from $109.99 to $949.99 with 56 specs. For each of these spec values, the mean SHAP was calculated. The result will be discussed in the following section.

5. Result & validation

5.1. Result from the NN choice model

Table 5 shows the result from the NN choice model. For each sub-feature, spec values are presented in ascending order. The corresponding SHAP values are also presented. The specs with positive SHAP values are highlighted in blue. As mentioned in Section 3.3, the SHAP value represents the influence of the spec value on customers' choices. Therefore, the highlighted values show the range of specifications recommended for product design. These ranges have two types of design implications, linear and nonlinear.

First, the NN model provides linear design implications for the screen size, CPU speed, battery capacity, screen resolution, CPU count, and ROM. As shown in Table 5, spec values are divided into two groups in each of these features. For the screen size, the group with higher spec values has a positive influence on customers' choice. For the CPU speed, the group with lower specs has a positive effect on the customer's decision. In the same manner, the other four features show the customers' preferences for higher/lower specs.

The NN choice model provides nonlinear design implications as well. In Table 5, the front and rear camera, price, screen type, and RAM have a nonlinear relationship with the customer's choice. There is no distinct trend, e.g., the higher, the better, or the lower, the better. Instead, a set of preferred spec values is identified. Regarding the camera, the recommended specs are {5, 8, 25} MP for the front camera and {12, 108} MP for the rear camera. The low specs (5, 8 MP front, 12 MP rear) would be proper for the low-mid-tier smartphones, and the high specs (25 MP front, 108 MP rear) would be appropriate for the flagship models. Regarding RAM, the recommended specs are {1.5, 2.0, 3.0, 12.0} GB. In the same context, 12 GB would be for the high-tier smartphones, and the others would be for the low-mid-tier products. In the price column, the middle range has a positive influence on customers' choices. In specific, the $208.99 - $749.00 range is recommended. All prices with a positive SHAP value fall within this range. And all prices outside of this range have a negative SHAP value.

5.2. Validation

The resultant spec guidance is validated by two approaches: (i) compares the design implications from the NN choice model and the previous MNL model; (ii) compares the result with design insights from customer reviews.

Table 6 shows the MNL result. All sub-features extracted from customer review data are significant factors with P-values less than 0.05. Design implications are drawn from coefficients. The coefficient's sign implies customers' tendency for the feature, and the magnitude of the coefficient indicates how much influence the feature has on the product's utility. Based on Table 6, it is recommended to decrease the spec values of the CPU speed, rear camera, battery capacity, and price. For other features, higher spec values are recommended. As explained in Section 5.1, the NN choice model provides both linear and nonlinear design implications. And the linear implications are consistent with that of MNL. Specifically, the NN model in Table 5 implies that customers prefer higher specs for the screen size, screen resolution, CPU count, and ROM. The MNL result has positive coefficients for all these features. For the CPU speed and battery capacity, the NN model implies that customers prefer lower specs. And the coefficients for these features have negative values in the MNL model. Although two models tell the same tendency, the NN model has an additional benefit. It provides a lower or upper bound for spec decisions. For example, when a company has a tight budget, it may want to select the minimum spec without compromising customer satisfaction. Then, the company can choose a 6.1" screen component and 128 GB ROM because they are the lowest spec values with the positive SHAP value, as shown in Table 5.

The NN model also provides non-linear preferences for spec values, and this gives novel design implications. For example, in MNL, the negative coefficient for the price implies that the lower the price, the better. However, the NN result shows that too low prices have a negative effect on customer choice. In addition, it provides recommended price ranges with specific numbers. This spec guidance cannot be obtained from the previous choice models. The practical application of these novel design implications will be discussed in Section 6.

Another approach for validation is to compare the result with design insights from customer opinions. Table 7 shows the reviews mentioning three sub-features (screen size, battery capacity, and price). For each sub-feature, 100 reviews were randomly selected, and then

Table 5
NN choice model result.

Screen size		CPU speed		Camera front		Camera rear		Battery cap		Price	
Spec	SHAP	Spec	SHAP	Spec	SHAP	Spec	SHAP	Spec	SHAP	Spec	SHAP
4.7	-0.233	1.4	0.437	2	-0.008	5	-0.015	1821	0.633	109.99	-0.053
5.0	-0.190	1.6	0.276	5	0.018	8	-0.014	1960	0.631	119.99	-0.058
5.5	-0.116	1.7	0.235	7	0.000	12	0.007	2600	0.150	124.96	-0.135
...	...	1.8	0.224	8	0.004	13	-0.001	...	...	...	...
5.8	-0.046	2.0	0.110	10	-0.005	16	-0.005	3179	0.112	361.99	0.004
6.0	-0.010	2.1	0.022	12	-0.003	24	-0.022	3300	0.093	399.95	0.019
6.1	0.027	2.2	0.013	13	-0.014	25	-0.011	3340	-0.013	409.00	0.025
6.2	0.042	2.3	-0.013	16	-0.005	32	-0.042	3400	-0.011	421.78	0.003
...	...	2.4	-0.041	20	-0.018	48	-0.034	...	...	...	...
6.6	0.082	2.5	-0.084	25	0.006	64	-0.022	5000	-0.369	835.99	-0.129
6.7	0.118	2.7	-0.110	32	-0.008	108	0.014	5020	-0.334	849.99	-0.030
6.8	0.078	2.8	-0.141					5260	-0.394	949.99	-0.098

Screen resol		Screen type		CPU count		Memory RAM		Memory ROM	
Spec	SHAP	Spec	SHAP	Spec	SHAP	Spec	SHAP	Spec	SHAP
1	-0.136	1	0.037	4	-0.123	1.5	0.051	16	-0.093
2	0.000	2	-0.045	6	-0.012	2.0	0.023	32	-0.046
3	0.115	3	0.023	8	0.007	3.0	0.004	64	-0.019
						4.0	-0.010	128	0.025
						6.0	-0.012	256	0.108
						8.0	-0.007	1024	0.310
						12.0	0.025		

Table 6
MNL result.

Product attribute	Coefficient
Screen size	1.008
Screen resolution	0.809
Screen Type	0.687
CPU speed	−1.316
CPU count	0.029
Memory RAM	0.025
Memory ROM	0.028
Camera rear	−0.027
Camera front	0.017
Battery capacity	−0.247
Price	−0.018

* All coefficients of product attributes are significant for the MNL model.

Table 7
Reviews for sub-features.

S_size	*"This phone is way better than my Samsung 7 - much better performance, **bigger screen** and [...]"* *"I also love that this phone is relatively lightweight (even with a ballistics case), and has a **large screen**."*
B_cap	*"Battery life is amazing due to the big **4000 mah**"* *"Battery capacity is not enough as described. About **5 h** for screen on time."* *"It's not the **biggest battery** but it still gives me **18 to 20 h** standby time."*
Price	*"Nice **price** for the combo that came with the wireless charger and Samsung earbuds."* *"The screen quality is really good for the **price**."* *"Graphics is out dated. At a certain point it stops letting you download apps. Don't waste your **money** spring for something that technically advanced."*

the context of each review was manually analyzed by the authors. Regarding the screen size, 70% of the customers show their preferences for the larger screen. 24% talk about the screen size, but they do not specify larger/smaller ones. The remaining 6% prefer the smaller screen. This result shows that most customers want to purchase a smartphone with a larger screen, which matches the spec guidance shown in Table 5. Regarding the battery, the spec guidance recommends a smaller capacity. It may seem unmatched with the common perception that customers prefer a larger battery. Customer opinions in the review data can explain this. Among all the reviews about the battery capacity, 100 items were randomly selected and analyzed. Half of them are concerned about battery health since they purchased refurbished products. The other half express their opinion on the battery size. As shown in Table 7, most of them care about usage time rather than the capacity itself. It implies that customers are satisfied with the small battery capacity if it provides decent usage time. This study does not include the spec values for battery life because the data is missing for some products. Unlike component specs, battery life is measured by simulations such as talk time and video play. So, it is hard to collect for all products. Further analysis of the battery life will be conducted in future research. The review analysis also explains the non-linearity in design implications. Among the reviews about price, 12% express satisfaction with renewed products due to the lower price than new ones. 7% complain about issues such as refunds. In the remaining 81%, most reviews evaluate the price based on the product specs, as shown in Table 7. They are satisfied with the product not because it has the lowest price but because it has reasonable specs for the given price. In the same context, reviewers give negative feedback for low-priced products when their specs do not meet the expectations. Therefore, it can be expected that customers do not prefer very cheap products due to their out-of-date specs. This preference is captured by the proposed method in Table 5, while MNL in Table 6 cannot capture it.

Table 8
Prediction accuracy (5-Fold Validation).

	MNL	NN
Random Sampling	66.61%	72.45%
Normalized Sampling	59.08%	64.36%

6. Discussion

6.1. Model evaluation

In Section 5, the nonlinear relationships between product attributes and customer choice are explained. Not only do they provide spec guidance for product design but also they contribute to higher prediction accuracy. Table 8 shows the prediction accuracy of the MNL

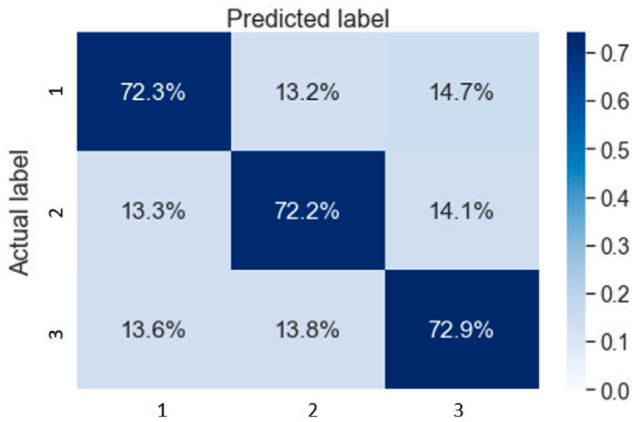

Fig. 8. Confusion matrix (NN, Random sampling).

and NN model based on 5-fold validation. For both sampling methods, the NN model provides better performance than MNL. Although the resultant accuracy seems lower than in other cases, it can be accepted considering the output type. The models with binary outputs (Onchis and Gillich, 2021; Lee et al., 2022) show 80–90+% of accuracy. On the other hand, those predicting multiple classes (Giabelli et al., 2022; Pan and Stark, 2022) provide 60–70+% accuracy. The proposed model predicts customers' choices among three alternatives. In this case, the guess is 33%, and the current result shows accuracy much higher than the guess. We will continue discussing the accuracy of the model in Section 7.

The performance of the NN model is further analyzed by a confusion matrix, as shown in Fig. 8. It contains all results from 5-fold validation. The percentage on the matrix represents True Positive (TP) and False Negative (FN). TP means that the case is predicted as positive and is actually positive. FN is the incident that is predicted as negative but is actually positive. By the definition of TP and FN, the values are calculated row-wise. In Fig. 8, the diagonal values represent TP since the predicted labels are the same as the actual labels. Off-diagonal values denote FN. Specifically, [row 1, column 2] shows the case where customers are predicted to choose option 2 when they actually purchase option 1. Similarly, [row 2, column 1] shows the case where customers are expected to select option 1, but they actually buy option 2. The confusion matrix shows that the NN choice model provides the same level of accuracy for all classes.

6.2. Managerial application

In product design, there exist correlations among design factors such as components, product dimensions, and cost. For example, in smartphone design, the battery capacity is dependent on the display size and product volume. And battery cost is determined by the battery capacity. There are also physical restrictions to be considered. Specifically, smartphone thickness cannot keep decreasing because it affects the surface temperature and thus the performance of core processors (Kang et al., 2019). Ergonomic evaluations (Lee et al., 2019) restrict product dimensions. They cannot be too small or too large to ensure comfort in the grip. In practical applications, product design is formulated as an engineering design optimization (EDO) problem (Gowharji and Whitefoot, 2021; Shiau et al., 2007). The above factors – correlations and physical restrictions – are included as constraints. Solving the problem gives an optimal design or design alternatives. However, there is no guarantee that customers will prefer or accept the resultant design solution. For instance, the EDO problem may result in different dimension options with a 6.0" screen and 3500 mAh battery. According to Table 5, customers prefer a screen size ≥ 6.1" and battery capacity ≤ 3300 mAh. Then, the solution is not a preferred design, although it is a feasible design satisfying all the restrictions. This issue can be addressed by including the spec guidance

The spec guidance provides upper/lower bounds of spec ranges and discrete spec values preferred by customers. This guidance will change the EDO problem by replacing previous constraints or adding new ones. Consequently, the feasible design space changes, and active constraints change too. Those related to spec ranges will be active to preserve the recommended spec values, and previously active ones may be lifted. It will result in new solutions with different product dimensions and spec values. In this way, the spec guidance obtained in this study can help companies design a product with spec configuration preferred by customers.

7. Conclusion & future works

Recently, online data has become a popular resource for consumer product design research. Many studies have been utilizing online data to understand customers and extract design implications. However, most of them focus on customers' tendency for product features, which is insufficient for industrial applications. Since a product consists of multiple components, a company needs information about product sub-features and spec ranges. The previous studies give implications for product feature categories, not sub-features. Moreover, the implications are about the customer's tendency for the features, not the range of sub-feature specs. This research addresses the above problem by extracting product spec guidance from online data.

The suggested approach consists of three stages. The first stage is constructing customer choice sets based on data mining. The method identifies sub-features mentioned in customer reviews and collects specifications for them. Then, customer choice sets are constructed based on the random and normalized sampling method (Suryadi and Kim, 2019). In the second stage, the method transforms these choice sets into one-dimensional arrays for NN input. Then, an NN choice model is trained based on these choice set arrays. In the final stage, the resultant NN model is interpreted by SHAP. The obtained SHAP values are further analyzed by the method proposed in this study. The presented approach was tested on smartphone products, and the result provides novel design implications for engineering design. Specifically, the recommended spec range for each sub-feature is obtained.

In future works, the limitations of the proposed approach will be addressed. First, this study provides spec guidance for only existing features. When designing a new product, companies make decisions on both new features and existing features. For example, in iPhone X, Apple introduced a new technology for security. The company removed the fingerprint sensor and installed the Face ID function. At the same time, the company fixed the spec configuration of the basic features such as screen, memory, and battery. This study helps companies determine these existing features. But guidance for new features is also necessary for product design. In future works, we will address this limitation by including new features in product attributes. Also, we plan to improve the feature extraction part by adopting state-of-the-art NLP techniques such as BERT (Devlin et al., 2018). Including domain knowledge in the data analysis process (Fantoni et al., 2021) is a possible approach for the dataset with technical terms. It can help further analysis on review data to extract reasons behind customer preference for certain spec values. In Section 5, text analysis on review data discovered that customers use 'battery capacity' for different contexts. This word sense disambiguation can be solved by BERT, enabling more precise text analysis. Based on the analysis result, additional data such as battery life will be included in spec guidance. Another limitation is that the current result provides somewhat lower accuracy than in other cases, as discussed in Section 6.1. To improve the accuracy, we tested different NN structures, including feedforward neural networks (FFNN) with multiple hidden layers, recurrent neural networks (RNN), and convolutional neural networks (CNN). The result is presented in Table C.11 in Appendix C. FFNN with multiple layers

Table A.9
Product attributes.

Main	Sub	Unit
Screen	Size	Inch
	Resolution	HD = 1, FHD = 2, QHD = 3
	Type	TFT = 1, IPS = 2, OLED = 3
CPU	Speed	GHz
	Count	
Memory	RAM	GB
	ROM	GB
Camera	Rear	MP
	Front	MP
Battery	Capacity	mAh
Price		USD

Table B.10
Hyperparameters for NN.

Parameter	Option
Epochs	200, 300, 400, 500
Batch size	20, 50, 100
Optimizer	SGD, RMSprop, Adagrad, Adadelta, Adam, Adamax, Nadam
Learning rate	0.0005, 0.001, 0.005, 0.01, 0.05
Initializer	uniform, lecun_uniform, normal, zero, glorot_normal, glorot_uniform, he_normal, he_uniform
Hidden layer size	30, 50, 80, 100

resulted in slightly lower accuracy than SNN adopted in this study, RNN provides the accuracy around 57%, and CNN gives the average accuracy of 72%. Although no improvement was observed in the current test result, different NN architectures other than the listed may enhance the prediction accuracy. We will work on new structures for NN to improve the model performance. The enhanced accuracy will increase the reliability of this research. Lastly, this study has a limitation in that it mines historical data for new product design. Due to the nature of the data, the method cannot identify radical innovations in the market that are very different from the past. It is a common drawback of research in data-driven design based on user-generated data, and new techniques for radical innovation may be developed in future works.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This material is based on the work supported by the National Research Foundation of Korea under Grant No. 2021R1I1A1A01044552.

Appendix A. Appendix 1

Table A.9 shows the description for product attributes.

Appendix B. Appendix 2

Table B.10 shows the hyperparameters tested in the grid search.

Appendix C. Appendix 3

Table C.11 shows the prediction accuracy of different NN structures. The NN was implemented by the keras library in PYTHON (keras.layers - LSTM/ GRU, keras.applications - DenseNet121/ ResNet50/ InceptionV3/ VGG16) The result is based on the 5-fold validation.

The results provide consistent accuracy according to their architectures, i.e., FFNN, RNN, and CNN. One exception is VGG16. Although

Table C.11
Prediction accuracy.

Model	Accuracy
FFNN (H = 1)	72.45%
FFNN (H = 5)	72.39%
FFNN (H = 10)	71.64%
GRU	56.55%
LSTM	56.54%
DenseNet121	72.22%
ResNet50	72.10%
InceptionV3	72.05%
VGG16	32.83%

Table C.12
Prediction distribution of VGG16.

Class	1	2	3	Total
Number of predictions	3947	3932	4121	12000
	33%	33%	33%	

the prediction is evenly distributed among three classes, as shown in Table C.12, the resultant accuracy is very low (32.83%). It is probably because the number of training sets is insufficient for VGG16. While the other CNN models have parameters $\leq$ 25.6M, VGG16 has much more parameters of 138.4M (Keras, 2022). So, it requires a much larger sample size for training.

References

Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A., Arshad, H., 2018. State-of-the-art in artificial neural network applications: A survey. Heliyon 4.
Ali, M.M., Doumbouya, M.B., Louge, T., Rai, R., Karray, M.H., 2020. Ontology-based approach to extract product's design features from online customers' reviews. Comput. Ind. 116, 103175.
Alwosheel, A., Cranenburgh, S., Chorus, C.G., 2018. Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. J. Choice Model. 28, 167–182. http://dx.doi.org/10.1016/j.jocm.2018.07.002.
Archak, N., Ghose, A., Ipeirotis, P.G., 2007. Show me the money! Deriving the pricing power of product features by mining consumer reviews. In: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 56–65.
Ben-Akiva, M., Lermna, S.R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. The MIT Press.
Bertoni, A., 2020. Data-driven design in concept development: Systematic review and missed opportunities. In: Proceedings of the Design Society: DESIGN Conference, Vol. 1. Cambridge University Press, pp. 101–110.
Bi, J.-W., Liu, Y., Fan, Z.-P., Cambria, E., 2019. Modelling customer satisfaction from online reviews using ensemble neural network and effect-based kano model. Int. J. Prod. Res. 57 (22), 7068–7088.
Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent Dirichlet allocation. J. Mach. Learn. Res..
Chen, W., Hoyle, C., Wassenaar, H.J., 2012. Decision-Based Design. Springer.
Chevalier, J.A., Mayzlin, D., 2006. The effect of word of mouth on sales: Online book reviews. J. Mar. Res. 43 (3), 345–354. http://dx.doi.org/10.1509/jmkr.43.3.345.
Chiarello, F., Belingheri, P., Fantoni, G., 2021a. Data science for engineering design: State of the art and future directions. Comput. Ind. 129, 103447.
Chiarello, F., Coli, E., Giordano, V., Fantoni, G., Bonaccorsi, A., 2021b. Data for engineering design: Maps and gaps. Proc. Des. Soc. 1, 821–830.
Chong, A.Y.L., Ch'ng, E., Liu, M.J., Li, B., 2017. Predicting consumer product demands via Big Data: The roles of online promotional marketing and online reviews. Int. J. Prod. Res. 55 (17), 5142–5156. http://dx.doi.org/10.1080/00207543.2015.1066519.
Cortesy, A., Chaki, N., Saeed, K., Wierzchon, S., 2012. Computer Information Systems and Industrial Management. Springer.
Dekhtiar, J., Durupt, A., Bricogne, M., Eynard, B., Rowson, H., Kiritsis, D., 2018. Deep learning for Big Data applications in CAD and PLM–Research review, opportunities and case study. Comput. Ind. 100, 227–243.
Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
Fantoni, G., Coli, E., Chiarello, F., Apreda, R., Dell'Orletta, F., Pratelli, G., 2021. Text mining tool for translating terms of contract into technical specifications: Development and application in the railway sector. Comput. Ind. 124, 103357.

Feng, Y., Zhao, Y., Zheng, H., Li, Z., Tan, J., 2020. Data-driven product design toward intelligent manufacturing: A review. Int. J. Adv. Rob. Syst. 17 (2), 1729881420911257.
García, M.V., Aznarte, J.L., 2020. Shapley additive explanations for NO2 forecasting. Ecol. Inf. 25.
Giabelli, A., Malandri, L., Mercorio, F., Mezzanzanica, M., 2022. WETA: Automatic taxonomy alignment via word embeddings. Comput. Ind. 138, 103626.
Goldberg, D.M., Abrahams, A.S., 2022. Sourcing product innovation intelligence from online reviews. Decis. Support Syst. 157, 113751.
Gowharji, W., Whitefoot, K.S., 2021. Influence of omitted variables in consumer choice models on engineering design optimization solutions. J. Mech. Des. 143 (12), 121–703. http://dx.doi.org/10.1115/1.4051192.
Hu, M., Liu, B., 2004. Mining opinion features in customer reviews. In: 19th National Conference on Artificial Intelligence, San Jose, CA. pp. 755–760.
Jiao, Y., Qu, Q.-X., 2019. A proposal for Kansei knowledge extraction method based on natural language processing technology and online product reviews. Comput. Ind. 108, 1–11.
Joung, J., Kim, H.M., 2021. Explainable neural network-based approach to kano categorisation of product features from online reviews. Int. J. Prod. Res. 1–21.
Kang, S., Choi, H., Park, S., Park, C., Lee, J., Lee, U., Lee, S.-J., 2019. Fire in your hands: Understanding thermal behavior of smartphones. In: The 25th Annual International Conference on Mobile Computing and Networking. pp. 1–16.
Keras, 2022. Keras applications. https://keras.io/api/applications/.
Lee, D., Derrible, S., Pereira, F.C., 2018. Comparison of four types of artificial neural network and a multinomial logit model for travel mode choice modeling. Transp. Res. Rec. 2672, 101–112.
Lee, S., Kyung, G., Yi, J., Choi, D., Park, S., Choi, B., Lee, S., 2019. Determining ergonomic smartphone forms with high grip comfort and attractive design. Hum. Factors 61 (1), 90–104.
Lee, J., Yeo, C., Kim, H., Mun, D., 2022. Deep learning-based digitalization of a part catalog book to generate part specification by a neutral reference data dictionary. Comput. Ind. 139, 103665.
Lundberg, S., 2018. Welcome to the SHAP documentation. https://shap.readthedocs.io/en/latest/index.html.
Lundberg, S.M., Lee, S., 2017. A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems, Vol. 30. pp. 4765–4774.
McFadden, D., 1986. The choice theory approach to market research. Mark. Sci. 5 (4), 275–297. http://dx.doi.org/10.1287/mksc.5.4.275.
Mclnnes, L., Healy, J., Astels, S., 2017. hdbscan: Hierarchical density based clustering. J. Open Source Softw. 1 (11), http://dx.doi.org/10.21105/joss.00205.
Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.
Nam, D., Cho, J., 2020. Deep neural network design for modeling individual-level travel mode choice behavior. Sustainability 12.
Nam, D., Kim, H., Jayakrishnan, R., 2017. A model based on deep learning for predicting travel mode choice. In: Transportation Research Board 96th Annual Meeting Transportation Research Board.
O'Dea, S., 2021. Smartphones in the U.S. - Statistics & facts. https://www.statista.com/topics/2711/us-smartphone-market/.
Onchis, D.M., Gillich, G.-R., 2021. Stable and explainable deep learning damage prediction for prismatic cantilever steel beam. Comput. Ind. 125, 103359.
Pahl, G., Beitz, W., Feldhusen, J., Grote, K.H., 2007. Engineering Design. Springer, http://dx.doi.org/10.1007/978-1-84628-319-2.
Pan, Y., Stark, R., 2022. An interpretable machine learning approach for engineering change management decision support in automotive industry. Comput. Ind. 138, 103633.
Park, S., Kim, H.M., 2022. Phrase embedding and clustering for sub-feature extraction from online data. J. Mech. Des. 144 (5).
Ramanujam, V., Balakrishnan, H., 2011. Estimation of maximum-likelihood discrete-choice models of the runway configuration selection process. In: Proceedings of the 2011 American Control Conference. pp. 2160–2167.
Roth, A.E., 1988. The Shapley Value. Cambridge University Press, pp. 31–40.
Shiau, C.-S., Tseng, I.H., Heutchy, A.W., Michalek, J., 2007. Design optimization of a laptop computer using aggregate and mixed logit demand models with consumer survey data. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol. 48078. pp. 175–185.
Sifringer, B., Lurkin, V., Alahi, A., 2018. Enhancing discrete choice models with neural networks. In: 18th Swiss Transport Research Conference.
similarweb, 2022. Top websites ranking for e commerce and shopping in the world. https://www.similarweb.com/top-websites/category/e-commerce-and-shopping/.
Spreafico, C., Spreafico, M., 2021. Using text mining to retrieve information about circular economy. Comput. Ind. 132, 103525.
Sujith, M., Hwang, S., Jeon, J., 2020. Failure mode and effects analysis of RC members based on machine learning-based shapley additive explanations (SHAP) approach. Eng. Struct. 219.
Sun, M., 2012. How does the variance of product ratings matter? Manage. Sci. 58 (4), 696–707. http://dx.doi.org/10.1287/mnsc.1110.1458.
Suryadi, D., Kim, H.M., 2019. A data-driven methodology to construct customer choice sets using online data and customer reviews. J. Mech. Des. 141.
Tuarob, S., Tucker, C., 2015. Quantifying product favorability and extracting notable product features using large scale social media data. J. Comput. Inf. Sci. Eng. 15 (3).
Wang, M., Chen, W., 2015. A data-driven network analysis approach to predicting customer choice sets for choice modeling in engineering design. J. Mech. Des. 137.
Wang, A., Zhang, Q., Zhao, S., Lu, X., Peng, Z., 2020. A review-driven customer preference measurement model for product improvement: Sentiment-based importance–performance analysis. Inform. Syst. E-Bus. Manag. 18 (1), 61–88.
WHO, 2012. How to Conduct a Discrete Choice Experiment for Health Workforce Recruitment and Retention in Remote and Rural Areas: A User Guide with Case Studies. World Health Organization.
Zhou, F., Ayoub, J., Xu, Q., Yang, X.J., 2019. A machine learning approach to customer needs analysis for product ecosystems. J. Mech. Des. 142.

Seyoung Park is currently a Ph.D. candidate in the department of industrial and enterprise systems engineering at the University of Illinois Urbana-Champaign (UIUC). She received her BS in electrical engineering at POSTECH and MS in industrial engineering at UIUC. Her main research interests include data mining and data-driven product design.

Junegak Joung is currently an assistant professor in the division of interdisciplinary industrial studies at Hanyang University. He received BS and Ph.D. in industrial and management engineering at POSTECH. His main research interests include product/service design, text mining, and interpretable machine learning.

Harrison Kim is currently a professor in the department of industrial and enterprise systems engineering at the University of Illinois Urbana-Champaign. He received Ph.D. in mechanical engineering at the University of Michigan. His research interests include engineering systems design and optimization, green, sustainable product design, manufacturing and remanufacturing, product design analytics, and trend-mining design.