Eco-Kansei design for retailing packaging: a current research progress

A Aprilia\textsuperscript{1}, T Dhatna\textsuperscript{2}, N S Indrasti\textsuperscript{2} and Sugarto\textsuperscript{2}

\textsuperscript{1}Doctoral Student, of Agroindustrial Engineering School of Graduate Studies, IPB University
\textsuperscript{2}Department of Agroindustrial Technology Faculty of Agricultural Technology and Engineering, IPB University (Bogor Agricultural University)

E-mail: arrahmah_aprilia@apps.ipb.ac.id

Abstract. The ultimate goal of this work is to support product packaging design by providing several guidelines in accordance with environmental issues related to users and perspectives of their knowledge. Currently, one needs to consider important moments from consumer experience with sustainable products model that includes complex dimensions such as Kansei Engineering (KE) by integrating the customers themselves in the design process. Eco-Kansei brings specific approaches to design products and services by taking into account their environmental impacts during the complete lifecycle. This must be considered in an integrated holistic perspective, where the adequacy of products with industrial processes must consider many sustainability criteria, including the scope of profit, people and of course the planet. The sustainable criteria must be able to be adjusted to market demands, and be able to be translated by packaging designers in packaging design elements. This work has completed modeling the classification criteria of sustainability with logistic regression, and evaluating it with the confusion matrix. However, the results of the confusion matrix concluded that the sustainable criteria for tuna packaging are still far from sufficient and balanced in the three subjective classification groups.

1. Introduction
Kansei Engineering aims to produce new products based on the feelings and demands of consumers. There are four general stages in Kansei Technology \cite{1}; 1) understanding user's feelings (Kansei) about the product; 2) identifying product design characteristics; 3) building Kansei Engineering; 4) adjust product design with changes in current community or community preference trends. Community trends in the fishing industry are related to ecolabeling and sustainable industry schemes. This encourages the fishing industry to improve marketing strategies in approaching the desires of consumers who also support ecolabeling. On the other hand, consumer knowledge about green packaging is relatively low \cite{2}. Although, most consumers have certain environmental awareness and willingness to pay for environmentally friendly products, the purchase of green products still depends on many factors that influence it \cite{3}.

In brief, eco-Kansei is a "way to produce" new products based on the user's environmental awareness that is closely related to subjective ideas that are usually covered by Kansei Engineering. The Eco-Kansei
design is then evaluated with values that carry a specific approach to designing products and services taking into account environmental impacts during the complete packaging life cycle [4].

The objectives of this work are to (1) identify sustainability criteria from Eco-Design Literature (2) compile sustainability word vectors from words or sentences of sustainability criteria for Kansei-word packaging-labeling which is aesthetics, environmental, and functional as the Kansei-goal label (3) predict Kansei label class of each word, phrase and sentence obtained by logistic regression. (4) evaluate proposed models.

2. Literature study

2.1. Sustainability criteria from eco-design literature
Eco-design brings a specific approach to designing products and services taking into account their environmental impacts during the complete life cycle [5]. In a business context, environmentally friendly design significantly reduces environmental impact while providing a competitive advantage in accordance with customer expectations [6]. Generally, eco-design articles raise labels of sustainable criteria either as an introduction to knowledge about sustainability or as a target of research. The sustainable labels are mentioned in words and sentences. The author has compiled words and sentences that can be used as a label of sustainability from eco-design literature. The label is sourced from 5 articles including Wendy Jedlicka, 2009 [7], R. Holdway, D. Walker, and M. Hilton, 2010 [8], C. Chen, C. Yeh, and Y. Lin, 2009 [9], Chen, CF, Yeh, CH, & Lin, YC, 2010 [10], VL Suárez, G. Barrera, and RM Naveiro, 2019 [11]. These words and sentences are hereinafter referred to as sustainability criteria. Sustainability criteria are very abstract criteria if they are related to functional packaging that must be referred to as the basis for the design concept.

In the sentence label about sustainability there are words that may be repeated on other labels, but they will have a different meaning. An example is the “Form Simplicity” phrase, which is labeled as a functional target. But it will be different if the word "simplicity" is made into the phrase "Simplicity and Minimalism Mechanism (reducing the number of components and materials)" then this word must be labeled as environmental.

2.2. Eco-Kansei
Packaging design can be defined as the relationship of form, structure, material, color, image, typography, and regulatory information with additional design elements to make a product suitable for marketing. The main objective is to create a kind of vehicle that functions to contain, protect, transport, issue, store, identify and differentiate a product on the market [12]. Packaging designs with variable elements must be able to carry environmentally friendly messages and even be able to influence changes in consumer attitudes to support the concept of environmental care.

Eco-Kansei has been explored by several researchers in various fields of application and method development. As did Chen et al [10][13] and Hartono [14][15], who explored and developed Kansei’s packaging design relationships based on environmental awareness, with methods such as multidimensional scaling analysis, cluster analysis, Kano, Quality function Deployment, and TRIZ (a Russian acronym for “Teoriya Resheniya Izobretatelskikh Zadatch” which means “Theory of Inventive Problem Solving”). The Kansei stimulus used in the study is still around verbal and image. Image stimulation development as done by Kim et al [16], which develops image stimuli with various angles of product factor images (eg: product appearance and darkness variations), and Reality Sets (Uninominal-Binominal), such as front-facing cars, side-cars, multi-aspect cars (as a Uninominal Reality Set) and a combination of front and side cars (as a Binominal Reality Set).

Chen et al [10] and Ushada et al [17], studied machine learning to explore verbal perceptions and their relationship with design elements using the same stimulus, that is the picture. Wagner et al [18] also explored
ecodrawing stimuli and the development of decision concepts for packaging design elements. This study analyzes consumers' perceptions about environmental values on web pages as a means of communicating visual information. Mapping the relationship of packaging design elements with these studies is modeled by the semantics of Verbal and Visual Reports on eco-images and questionnaires. The questionnaire is very sensitive to the problem of ambiguity of words. This points to the problem of lack of understanding and misunderstanding in the emotional list of semantic words and emphasizes the absence of meaningful contexts that can help reduce ambiguity. Many studies [19] [20] [21] explore the meaning of various emotional words in context by extracting texts. The use of complete sentences can provide better clarity because of the inclusion of context, both in the linguistic and situational sense. The use of complete sentences can be clearer than isolated words and can reduce ambiguity [22]. Therefore, we need to quantify the sentence and the word, so we can make evaluation model of every word, phrase and sentence that can encourage the manifestation of semantic imperfections from various sources. However, to facilitate understanding in accordance with the general designer's target, author classifies words, phrases and sentences according to what Chen, C., F., Yeh, C., H., & Lin, YC, 2010 have described as a design target. There are aesthetic, environmental and functional goal labels.

2.3. Word vectorization
Kansei uses the self-report method to document the word Kansei in the form of interviews and questionnaires. As information technology advances, interviews and questionnaires are distributed and stored in texts that can be analyzed by computers, using the concept of Natural Language Processing. The main challenge in extracting text is to transform unstructured texts into structured models. This must be done before carrying out further analytics. Natural Language Processing for text preprocessing steps may be the same for all text mining tasks, even though the processing steps selected depend on the task. From a Machine Learning perspective, there are five main steps that must be taken to make unstructured text data into data that is ready to be analyzed. The five main steps involved in NLP are: Reading the corpus, Tokenization, Cleaning/Stop word removal, Stemming, Converting into Numerical Form [23].

2.4. Classification Model with Logistic Regression
Chen et al [10], divides 3 target class concepts from the development of design elements, namely: aesthetic, functional, and environmentally friendly attributes. The attributes Chen listed in the sentence are still very limited, however this is needed in order to simplify the complexity of packaging design targets. Moreover, attributes about the environment, which like many studies on sustainable development mentioned that these attributes must balance the interests of people, profits, and the planet. Sustainability criteria from the study of literature will enrich the verbal semantics of sustainable development by adjusting the subjective classification grouped by user of packaging design system, which is in line with the grooves of increasing packaging design cognition.

Logistic regression analysis (LRA) extends the techniques of multiple regression analysis to research situations in which the outcome variable is categorical [24]. Logistic regression calculates the probability of an event occurring over the probability of an event not occurring, the impact of independent variables is usually explained in terms of odds. If the probability of an event occurring is p, the probability of the event not occurring is (1-p). Then the corresponding odds is a value given by:

\[ \text{odds} = \frac{p}{1-p} \]  

The model is not a good model because the extreme value of x will give a value of \( \alpha + \beta x \) that does not fall between 0 and 1. Then the logistic regression solution for the categorical prediction problem is to change...
the opportunity to use natural logarithms so that the relationship between the class variable and the feature variable is as follows [25], on equation (2) and (3).

\[
\text{Logit}(y) = \ln(\text{odds}) = \ln\left(\frac{p}{1-p}\right)
\]

\[
P = \frac{e^{a+bx}}{1+e^{a+bx}}
\]

p is Probability where y is interested outcome e of X, and a as a specific value

3. Method

3.1. Research framework

This work can be described in the following framework (figure 1). As stated in the objectives, there are 3 working groups namely count word vectorization, compile learning models with logistic regression, and finally, draw conclusions by analyzing model evaluation parameters for each classification class.

3.2. Word vectorization classification model for prediction with logistic regression

There are 4 columns in the logistic regression calculation table. The first column is Raw Prediction, which is the raw output of the logistic regression classifier. Probability is the result of applying the logistics function to raw prediction (array of length Probabilities are the same as raw prediction). All 44 words and sentences are then processed by extracting the basic words by eliminating parts of the sentence that have no meaning in building the model. This step consists of tokenization and stemming. Of the 44 words and sentences, we obtain basic words.
3.3. Evaluation parameter
The evaluation technique uses confusion matrix as a comparison between actual data and predicted data. We can see at the Table 1, confusion matrix for this work consists of 9 variables. TP means true positive, which means the results of the actual data are similar to the predicted data. While E means the actual data do not match the predicted data.

Table 1. Confusion matrix with 3 class target classification.

| Actual | Aesthetics (A) | Functional (B) | Environmental (C) |
|--------|----------------|-----------------|-------------------|
|        | Prediction     |                 |                   |
|        | A              | B               | C                 |
| Actual | TP<sub>AA</sub> | E<sub>AB</sub> | E<sub>AC</sub>    |
|        | E<sub>BA</sub> | TP<sub>BB</sub> | E<sub>BC</sub>    |
|        | E<sub>CA</sub> | E<sub>CB</sub> | TP<sub>CC</sub>   |

Parameters formed to evaluate are Precision (<em>P</em>), Recall, Overall accuracy and Specificity. Precision parameter is the comparison of the true positive value of class A (Aesthetics) with all the values that contain the predicted value of A (Aesthetics), both the true positive value and the actual error data B (Functional) or C (Environmental) but the predicted words / sentences fall into group A (Aesthetics). Recall or sensitivity parameter is a comparison of true positive A with all values that contain the actual data value A, both the true positive value and the error data which is the actual data A but the predicted words / sentences fall into groups B or C. Overall accuracy is a comparison of all predicted precise data with all data. Meanwhile specificity is the possibility of prediction being negative when the variable does not exist. It is possible that the prediction results are other than environmentally friendly classes when actual environmentally friendly data choices do not exist.

\[
\text{Precision}_A = \frac{|TP_{AA}|}{|TP_{AA}| + |E_{BA}| + |E_{CA}|} \tag{4}
\]

\[
\text{Recall}_A = \frac{|TP_{AA}|}{|TP_{AA}| + |E_{AB}| + |E_{AC}|} \tag{5}
\]

\[
\text{Accuracy} = \frac{|TP_{AA}| + |TP_{BB}| + |TP_{CC}|}{|TP_{AA}| + |TP_{BB}| + |TP_{CC}| + |E_{AB}| + |E_{AC}| + |E_{BA}| + |E_{BC}| + |E_{CA}| + |E_{CB}|} \tag{6}
\]

\[
\text{Specificity}_C = \frac{|TP_{AA}| + |E_{AB}| + |E_{BA}| + |TP_{BB}|}{|TP_{AA}| + |TP_{BB}| + |E_{AB}| + |E_{AC}| + |E_{BA}| + |E_{BC}|} \tag{7}
\]

4. Result and discussion
Author has collected 44 words and phrases are mentioned as criteria for sustainability. Words and phrases such as: (1) System design that extends product life,(2) Minimize and value by product,(3) Aim for biocompatibility,(4) Design to optimize transportation,(5) Reducing toxicity,(6) Improve equality and integrate actors in the system,(7) Promotes responsible consumption,(8) Fortify and promote local resources,(9) Maintain economic viability in product development,(10) Useful, safe & healthy for individuals and communities throughout their life cycle,(11) Is beneficial, safe and healthy for individuals and communities throughout its life cycle,(12) Meet market criteria for performance and cost,(13) Using renewable energy,(14) Maximizing the use of renewable or recycled source materials,(15) Produced using clean production technology and best practices,(16) Made from healthy...
ingredients in all possible end-of-life scenarios'. (17) "Physically designed to optimize materials and energy'; (18) 'Restored and used effectively in the Cradle to Cradle cycle'. (19) 'Considering Environmental health - toxic effects of all components'. (20) 'Considering VOC content, the mass of volatile organic compounds'. (21) 'Considering Use of materials, products and packaging / shipping systems throughout the entire life cycle'. (22) 'Considering The use-reduction of pallets, the theoretical volume fraction is not used', (23) 'Harmonious Forms Proportion', (24) 'Consistent Form Elements', (25) 'Fashionable', (26) 'Timeless Style', (27) 'High Cultural Reference or Identity', (28) 'Elegant', (29) 'Culture', (30) 'Elegant', (31) 'Form Simplicity', (32) 'Comfortable', (33) 'Stable', (34) 'Ease of Mobility', (35) 'Update and Modularity (ease of maintenance/ease of replacement, and extend product life', (36) 'Space Adaptation (suitable for a variable of office spaces, eg large or small spaces, living rooms or other public areas)', (37) 'Simplicity and Minimalism Mechanism (reducing the number of components and materials)', (38) 'Ease of Disassembly (shorter process and using simple tools)', (39) 'Harmony with Environment (Harmonious relationship between the consumer, product, and environment)', (40) 'Cyclic, the product is made from compostable organic materials or from minerals that are continuously recycled in a closed loop', (41) 'Solar, the product in manufacture and use consumes only renewable energy that is cyclic and life', (42) 'Safe, all releases to air, water, land, or space are food for other systems', (43) 'Efficient, the product in manufacture and use requires 90% less energy, materials and water than products providing equivalent utility did in 1990', (44) 'Social, product manufacture and use supports basic human rights and natural justice'.

Each base word is then labeled according to its position in the word group and is made into a vector so that it can be modeled in the continuous criteria classification model.

| Word or sentences | features_vec | Label |
|-------------------|--------------|-------|
| 1 System design that Extends product life | (185,[0,3,18,24,136,184],[1.0,1.0,1.0,1.0,0.5,0]) | 2 |
| 2 Minimize and value by product | (185,[0,102,139,184],[1.0,1.0,0.1,0.0,3.0]) | 3 |

The next step is to quantify the unstructured data of the words and sentences that are collected. After Tokenization, Cleaning / Stop word removal, Stemming, Converting into Numerical Forms the vector form of each word and sentence will be classified as a packaging label. Examples of first ten sentences can be seen in the following Table 2.

There are 4 values in the raw prediction and Probability columns. The first value is for the raw prediction and the probability of each word or sentence input that is not included in class 1, class 2, or class 3. While the second array value is the value of raw predictions and word probability if entered into class 1. While the value of the array to three is the raw prediction value and word probability if it enters class 2. Finally, the fourth array value is the raw prediction value and word probability if it enters class 3. The array probability takes its maximum value, and the argument label is given according to array position. Prediction is a class label argument whether the words or sentences included include 1 (aesthetic class), 2 (functional class) or class 3 (environmentally friendly class). Here are the first four data from the twenty-two data testing models as an example.

| Table 2. Word vector of sustainability criteria. |
Sustainability criteria that have been obtained and extracted from 44 words and sentences have become 185 data in the form of words. However, these 185 words still have words that are included as stop words. Therefore, the author still reduces 4 symbols that are not sustainable criteria words. These symbols include '&, ‘eg’, ”, ‘-’ and ’/’. Labeling sustainable criteria with numeric 1, 2 and 3 instead of aesthetic, functional and environmentally friendly classes, is done subjectively. This causes the amount of data per class is not the same. The data consists of 9 data in class 1, 10 data in class 2 and 25 in class 3.

Table 3. Prediction result of sustainability criteria with logistic regression model.
Sustainability compilation of the extracted words is then trained with a logistic regression model to obtain alpha and beta values that can predict 4 array values for each class called Raw prediction. This can be seen in the raw prediction column in Table 3. This compilation generates the prediction value of the word compilation label argument for the ongoing criteria in the prediction column. Of the 22 training data used, there are 9 data that do not fit the logistic regression prediction model. All data that is wrong prediction, determine the prediction argument in the value of class 3. This is caused by the amount of data training class 3 (environmentally friendly) almost three times as many as classes 2 and 1.

Logistic prediction result model evaluation based on 2 parameters, namely precision and recall. We can conclude that the amount of data is still not balanced when compared to each label respectively. Therefore, the overall accuracy value is still below 50%. The Environmentally friendly class specificity is only 0.125, whereas none of the precision and recall values are close to 90%.

5. Conclusion
Confusion matrix presents the degree to which predicted probabilities agree with actual outcomes in a classification table. The overall correct prediction, 0.461% shows an improvement under the chance level which is 50%. Recall measures the proportion of correctly classified events. Specificity for environmentally friendly measures the proportion of correctly classified environmentally friendly, which is just 33%. Evaluation results show that the data used is still not enough. The model used cannot support environmentally friendly variables.
There are still a lot of words and sentences that support these 3 classes. Sources of data can be obtained from interviews and extracting data in the form of extract comments on the website. It is recommended that the amount of data classified per class is balanced in number. Criteria data should be complete sentences, which have a good language structure to obtain the concept of objects and emotions that will be extracted in accordance with the concept of eco-Kansei.

6. References

[1] Nagamachi M 1995 Kansei engineering: A new ergonomic consumer-oriented technology for product development Int. J. Ind. Ergon. 15 3–11
[2] Hao Y, Liu H, Chen H, Sha Y, Ji H, and Fan J 2018 What affect consumers’ willingness to pay for green packaging? Evidence from China Resour. Conserv. Recycl. 141 21–29
[3] Joshi Y and Rahman Z 2015 Factors Affecting Green Purchase Behaviour and Future Research Directions vol. 3, no. 1–2 (Holy Spirit University of Kaslik)
[4] Rasamoelina F, Bouchard C and Aoussat A 2013 Towards a kansei-based user modeling methodology for eco-design International Journal of Affective Engineering 12 337-348 <hal-00962434>
[5] Bouchard C 2010 User kansei modeling and eco-design User kansei modelling and eco-design King Mongkut’s Univ. Technol. North Bangkok Int. J. Appl. Sci. Technol. 3 1–6
[6] Goepp V, Caillaud E and Rose B A 2013 framework for the design of knowledge management systems in eco-design Int. J. Prod. Res. 51 5803–5823
[7] Jedlicka W 2009 Packaging sustainability: Tools, systems, and strategies for innovative package design
[8] Holdway R, Walker D and Hilton M 2010 Eco-design and successful packaging Des. Manag. J. Former Ser. 13 45–53
[9] Chen C, Yeh C and Lin Y 2009 A Kansei Engineering Approach to Eco-product Form Design Proc. of IASDR09
[10] Chen C F, Yeh C H, and Lin Y C 2010 A neural network approach to eco-product form design Proc. 2010 5th IEEE Conf. Ind. Electron. Appl. ICIEA 2010 pp. 1445–1450
[11] Suárez V L, Barrera G and Naveiro R M 2019 Materials Design and Applications II, vol. 98. (Springer International Publishing)
[12] Tang S. 2014 Sustainable Packaging Design - A Business Case Study
[13] Chen H C 2018 Green Innovation Design Model of Kansei Engineering Adv. Intell. Syst. Comput. 739 267–276
[14] Hartono M 2016 A conceptual integrative model of Kansei Engineering, Kano and TRIZ towards sustainability in service 8th Widyatama Int. Semin. Sustain. Widyatama Univ., pp. 303–307
[15] Hartono M, Santoso A, and Prayogo D N 2017 How Kansei Engineering, Kano And QFD can Improve Logistics Services Int. J. Technol. 8 1070
[16] Kim S K, Cho Y, Niki K and Yamanaka T 2016 Integrating affective values to sustainable behaviour focused on Kansei engineering Int. J. Sustain. Eng. 9 378–389
[17] Ushada M and Murase H 2009 Design of customisable greening material using swarm modelling Biosyst. Eng. 104 2 169–183
[18] Wagner M, Curteza A, Chen Y, Thomassey S and Nationale E 2018 Environmentally-Friendly Perception of Fashion Products: A Kansei Study Int. Conf. Kansei Eng. Emot. Res 369–378
[19] Che W, Zhao Y, Guo H, Su Z and Liu T 2015 Sentence Compression for Aspect-Based Sentiment Analysis IEEE/ACM Trans. Audio Speech Lang. Process. 23 1211–1224
[20] Chaturvedi I, Cambria E, Welsch R E and Herrera F 2018 Distinguishing between facts and opinions for sentiment analysis: Survey and challenges Inf. Fusion 44 65–77.
[21] Alm C O, Roth D and Sproat R 2005 Emotions from text Proc. Conf. Hum. Lang. Technol. Empir. Methods Nat. Lang. Process. - HLT ’05 October 579–586.

[22] Burgess P 2016 Integrating the Packaging and Product Experience in Food and Beverages: A Road-Map to Consumer Satisfaction.

[23] Bandi R, Amudhavel J and Karthik R 2018 Machine learning with PySpark – Review vol. 12 no. 1.

[24] Dayton C M 1992 Logistic Regression Analysis Stat, 474-574

[25] Park H 2013 An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain J. Korean Acad. Nurs. 43 154-164

Acknowledgment
A gratitude is expressed to Ministry of Research and Higher Education and Ministry of Finance for support given to this work through LPDP-BUDI DN scholarship, as well as to Politeknik Negeri Media Kreatif, Indonesia.