ICIC Express Letters
Volume **16**, Number **5**, May **2022**

ICIC International ©2022 ISSN 1881-803X
pp. 505–512

AI-SUPPORTED EVALUATION OF KANO MODEL FEATURES FOR ONLINE COURSES

Daniel Moritz Marutschke[1] and Yugo Hayashi[2]

[1]College of Global Liberal Arts
[2]College of Comprehensive Psychology
Ritsumeikan University
2-150 Iwakura-cho, Ibaraki, Osaka 567-8570, Japan
{ moritz; yhayashi }@fc.ritsumei.ac.jp

Received September 2021; accepted November 2021

Abstract. *Online courses produce multivariate, often unstructured, large dataset. Challenges remain to use quantitative methodologies to extract information about users' learning abilities, satisfaction rates, expectation, and consumption experience. This research proposes a method combining machine learning and customer satisfaction analysis. Undergraduate students at a Japanese university were asked to fill out questionnaires ex-ante (before) and ex-post (after) taking an online course to contrast expectation and consumption experience. The questionnaire was designed to perform the Kano model analysis on 12 factors related to online course satisfaction and to use the artificial intelligence tool Word2Vec on freeform text answers to create a language model. Findings from the language model were used to support the feature classifications provided by the Kano model. As a proof of concept, the findings support the usefulness of Word2Vec supplementing the Kano model classifications, even with small text corpora.*
Keywords: E-learning, Online course, Kano model, Word2Vec

1. **Introduction and Background.** Online courses in higher education have gained attention in the past decades mainly with high-brand universities offering free Massive Open Online Courses (MOOCs). With growing popularity, shortcomings were identified and research fields in student retention, automated grading, test generation, and others emerged. The machine learning community has addressed these issues and progress has been made in the respective fields.

E-learning has still many unsatisfactory answers in research priorities and challenges, such as learner adaptation, pedagogical advantage and learner success, learner motivation, and learner satisfaction.

The authors of this paper address the question of how to measure and improve student satisfaction. Identifying factors that contribute to satisfactory online courses will be pursued as a basis for improved learner outcome.

When MOOCs started out, their strong drivers were brand recognition, accessibility, and being free of charge. Anyone with an Internet access could enroll in such an online course at no cost. Soon, however, one of the still persistent challenges was identified as student retention [1]. With growing popularity, the investigation into the reasons behind dropout rates also gained traction [2,3] with some offering strategies for solutions [4].

Other previous research includes learning motivation and exploring multiple hypothesis that tie into student satisfaction [5]. Others investigate the satisfaction to more mechanical implementations, such as medical software usage [6] as a virtual, web-based 3D electroencephalogram to propose design references for future learning platforms. The Kano method was used to poll students on several e-learning factors in a previous study

DOI: 10.24507/icicel.16.05.505

by Dominici and Palumbo [7] to construct a theoretical framework of online course design (ex-ante questionnaire without implementation and testing the ex-post consumption experience). Another research was published by Wang et al. to investigate a hybrid online/face-to-face course using blended learning [8].

Language models have gained popularity in data mining and big data analysis, but also garnered discussions about implicit bias in large models [9-11]. An excerpt of previous publications and research is presented below.

Altszyler et al. investigated small corpora for word associations and found that Latent Semantic Analysis (LSA) can provide better results for low-frequency words than Word2Vec [12]. They apply their algorithms to dream journals and are focused on psychological research.

A large corpus of abstracts and papers related to biomedical research was analyzed by Zhu et al. [13]. They conclude that increasingly large datasets can result in more relations in biomedical terms, but do not guarantee better precision. This ties into the strength of word associations, but poor semantic properties of Word2Vec.

Further limits and strengths of Word2Vec were investigated by Di Gennaro et al. [14]. While overfitting was mostly ruled out for the training process and a strength in analogy models was found, syntactic relationships were not able to be achieved by Word2Vec.

The research in using models for customer satisfaction in online courses is unable to explain the distribution in Kano model results.

To allow to compare future studies and allow context shifting vector arithmetic, the Word2Vec tool was chosen to combine with the Kano model in this research. This combination allows novel insight into customer satisfaction variance and distribution, which cannot be understood by the Kano model alone.

The rest of the paper is structure in the following way. Section 2 describes the research method, the Kano model in Section 2.1 and the Word2Vec machine learning in Section 2.2. Section 3 covers findings and discussions. The paper is concluded in Section 4.

2. **Research Method.** In this section, the authors describe how the Kano model from customer satisfaction research and the machine learning tool Word2Vec were combined to strengthen the insight into students' perception of online courses. Figure 1 shows the overview of the experiment. Printed questionnaires were distributed before (ex-ante) and after (ex-post) students take the online course.

To allow the Kano model analysis, functional and dysfunctional questions were asked in both the ex-ante and ex-post questionnaire (explained in Section 2.1):

- How do you feel, if the following function **is** included in an online course? (functional question)
- How do you feel, if the following function **is not** included in an online course? Please answer independently from previous question. (dysfunctional question)

The following list of 12 features was considered for online courses, based on previous research and online courses system functionality [7]: • User-friendly platform • Certificate of completion • Download of course material • Profile and account page • Quizzes and exercises • Interactive quizzes and exercises • Comment function • Personal tutor • User manual for the platform • Videos • Photos/Graphics • Text.

In addition, questions about age, gender, university grade, English language proficiency, and photography knowledge were asked.

The online course was created with a wide audience in mind and without special background knowledge. An introductory course to photography (lenses and depth of field) was chosen. The online course was self-hosted with WordPress, one of the most popular open source Content Management Systems (CMS) and a premium plugin LearnDash to allow scalable e-learning and online course functionality.

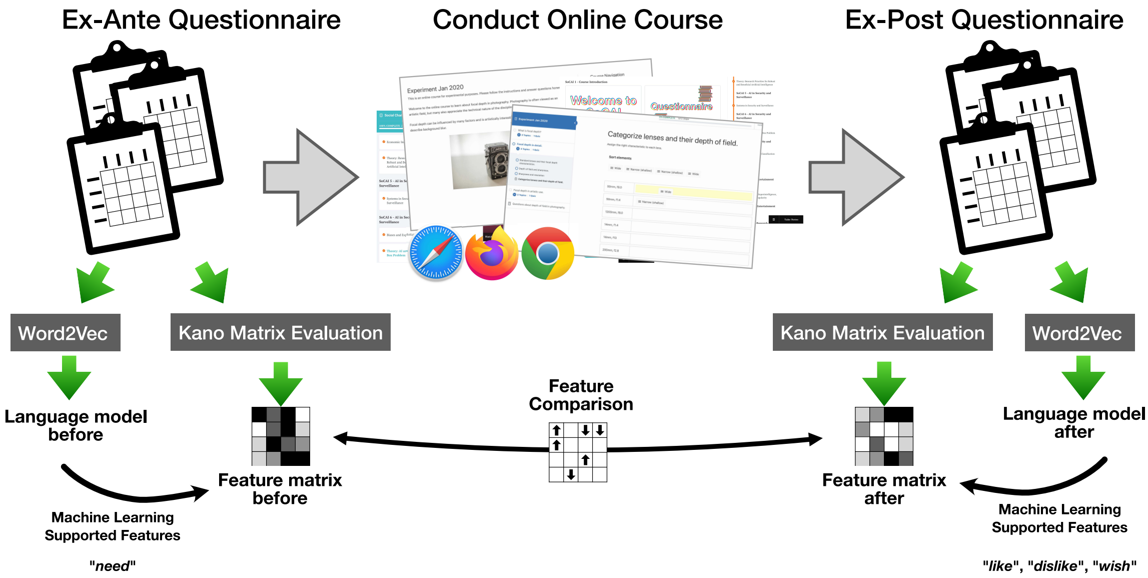

FIGURE 1. Experiment setup showing the ex-ante and ex-post questionnaires and their analysis using the Kano model and the machine learning tool Word2Vec

Other freely available plugins were used to protect content and create preset user accounts in bulk.

2.1. **Kano model.** The Kano model is a well established tool to conduct customer satisfaction research [15]. At the core is a questionnaire that asks (prospective) customers to answer functional and dysfunctional questions regarding a product's feature. The *functional* question asks customers how they feel if a feature is present (implemented) and the *dysfunctional* question asks customers how they feel if a feature is missing. Their response is typically rated from highly satisfied to highly dissatisfied for both of the aforementioned functional and dysfunctional questions. Combining these answers results in one of six categories of the Kano model.

The pair of functional and dysfunctional questions is presented in Figure 2. Participants have to rate a feature in five levels for both these questions – highly satisfied, as expected, neutral, can live with it, and highly dissatisfied.

Kano Model Evaluation Matrix

			Dysfunctional				
			How do you feel if the following feature **is not included**				
			Highly satisfied	As expected	Neutral	Can live with it	Highly dissatisfied
Functional	How do you feel if the following feature **is included**	Highly satisfied	Q	A	A	A	O
		As expected	R	I	I	I	B
		Neutral	R	I	I	I	B
		Can live with it	R	I	I	I	B
		Highly dissatisfied	R	R	R	R	Q

B = basic requirement	Q = questionable requirement
O = one-dimensional requirement	A = attractive requirement
I = indifferent requirement	R = reverse requirement

FIGURE 2. Kano model evaluation matrix, with resulting classification based on functional and dysfunctional questions

The list below details the categories that the Kano model produces. As the original publication describing the categories is written in Japanese, the terminology has been adjusted. Namely the term *quality* has been replaced by *requirement* to reflect terminology from software engineering and the term *must-be* was replaced by *basic*. The categories are pre-defined by Kano and described individually in the list below.

Most researches that incorporate the Kano model have a business-centric approach. For this research, the term *customer* is replaced by *user*, *student*, or *participant*.

- `B` (basic requirement)
  This is also called a *dissatisfier* or *must-be* factor, as these could be viewed as minimal (basic) requirements. The user expects these to be implemented well, and they are taken for granted. An increasingly well implemented variation does not increase the level of satisfaction, but with poor execution, the satisfaction will diminish. When a feature is judged as basic requirement by a high number of users, creators and developers should implement these with priority.
- `O` (one-dimensional requirement)
  The degree of satisfaction has a positive linear correlation with the degree of implementation. The proportional increase and decrease means that developers should be competitive here. That is why this requirement is also referred to as *performance attribute*. If an increase in customer satisfaction is a priority of a system or service, improving one-dimensional requirements is the most straight forward approach.
- `I` (indifferent requirement)
  As the name suggests, users do not care about these requirements. Identifying features that users frequently judge as an indifferent requirement can help to reduce unnecessary development efforts.
- `Q` (questionable requirement)
  If users judge a feature highly satisfactory for both functional and dysfunctional questions or highly dissatisfactory for both functional and dysfunctional questions, the result is contradictory. To still allow working with the rest of the results, the Kano model allows a classification as *questionable requirement*. Low numbers of `Q` point to reliable questionnaire design.
- `A` (attractive requirement)
  This is also called *satisfier* or *excitement factor*. These are attributes that are usually not anticipated by users. A good implementation and performance will greatly increase the customer satisfaction. Missing this requirement typically does not impact the satisfaction negatively. In terms of product or service development, these requirements can differentiate from the competition.
- `R` (reverse requirement)
  This factor can be viewed as the inverse of the one-dimensional requirement. The degree of satisfaction has a negative linear correlation with the degree of implementation. Developers should look out for features that users frequently judge as reverse requirements.

To illustrate the meaning of the Kano model categories covered in the case of an e-learning platform or online course, the following examples are explained for each requirement.

A working website, consistent URLs, and mobile-friendly content can be considered basic requirements (`B`). They do not increase the satisfaction if properly included, but would reduce the satisfaction if missing or implemented poorly.

An attractive feature (`A`) could be a live note-taking function or engaging implementation of gamification. Over time, attractive features might become one-dimensional requirements (`O`) or basic requirements (`B`).

One-dimensional requirements (`O`) increase the user's satisfaction proportional to the degree of implementation. Factors that result in a good User Experience (UX) increase the satisfaction proportionally to the degree of implementation. The opposite is also the case as poor implementation of factors that decrease UX also decreases the user's satisfaction.

The Kano model includes states that cannot be captured with other methods in this field. Reverse requirements (`R`) have the opposite effect of one-dimensional requirements (`O`) or in some cases of attractive requirements (`A`), i.e., they decrease (proportionally) the satisfaction when implemented. An example for web-based applications would be pop-up dialogue boxes. This is one of the reasons for choosing the Kano over others commonly used ones, such as SERVQUAL [16], E-Learning Satisfaction (ELS) [17], or e-SERVQUAL [18]. Dominici et al. also detailed two more reasons for this methodology: the first is crucial to this research and allows ex-ante and ex-post analysis, and the second is that this model does not assume a linear relationship between the product or service performance and the user satisfaction [7,19].

Questionable requirements (`Q`) are inconsistencies in the questionnaire answers. If a user answers both highly satisfied if a feature is included and is not included, the answer is unusable, in other words questionable (`Q`).

2.2. **Language model based on Word2Vec.** Artificial intelligence and language models have gained popularity and were successfully used in predictive modeling, translation, context-based analogies, and others. With growing size in datasets, concerns were also raised about potential biases introduced into large language models [9-11].

The original paper was published by Mikolov et al., proposing a simple neural network to train on large text corpora, representing words in multidimensional vector space [20]. Numerical representation of words as vectors allowed to calculate distances and similarities of word vectors [21,22]. This representation could also be used to perform arithmetic on word vectors to transform and compare words in different contexts [23].

With several test-runs a vector size of 60 was found to be optimal, also considering the vocabulary size being 99, to avoid overfitting. As indicated by Mikolov et al. and Aggarwal, accuracy for rare words is higher in the skip-gram architecture, rather than the Continuous Bag-of-Words (CBOW) architecture. Each entry in the free form part of the questionnaire was converted into an array by applying Part of Speech (POS) analysis using Stanza, a Natural Language Processing (NLP) toolkit developed by the Stanford NLP group [24]. The tokenized arrays were then used to train the Word2Vec model. This step might be unfeasible for larger datasets and there are mechanisms to counter frequent stop-words with a probability of selecting words based on how often they occur. In the case of the rather small corpus size, each value should be taken into account by the training process. The context window size of the Word2Vec training was set to 5, essentially taking all words in one array into account (five words to the left and right of each word respectively). The minimum count of words was set to 1, ensuring that unique words also get processed. This results in a Bernoulli model, predicting the context, providing a particular word, given by $P(w_1 \dots w_m|w)$, where w is a single target word as input and m context words $w_1 \dots w_m$ as output. For higher accuracy, negative sampling is used on contexts [21]. The initial learning rate was set to $\alpha = 0.001$. As the training corpus was small for this experiment, the results could slightly be improved compared to the default $\alpha = 0.025$.

To prime the categories, each was prefaced by a keyword – *need*, *like*, *dislike*, and *wish*. The keyword *need* indicates user expectation and was surveyed before taking the online course. After completing the online course, factors that users *liked* or *disliked* were surveyed, as well as a *wish* list.

2.3. **Combining the Kano model and Word2Vec.** The Kano model is an established tool in customer satisfaction research and provides a tool to allow quantitative analysis in a qualitative research field. A still difficult task is to untangle impactful factors, which are also influenced by cultural aspects. All participants in the experiment summarized in this paper are young Japanese students.

Open ended questions can prompt students to more freely express themselves and add valuable information. The qualitative nature of freeform text requires a method to extract quantitative information that can be combined with the Kano model. Here, the Word2Vec tool was implemented to combine the feature classification provided by the Kano model with similarity values of words associated with *need*, *like*, *dislike*, and *wish*.

3. **Results and Discussion.** This section will detail the findings from the Kano model evaluation and then tie the findings from the machine learning algorithm to these factors.

A total of 16 Japanese students participated, 5 males, 11 females, and 0 non-binary. The average age was 19.82 ($\sigma = 1.70$), ranging from 18 to 24.

As can be seen from Tables 1 and 2 and in accordance with Hofstede et al. of a collectivist culture, a majority of the factors are classified as *Indifferent* [25]. The evaluations show a redistribution from before to after in most cases. Only the factor *Comment function* shows almost no changes.

Table 1. Kano model evaluation ex-ante (expectation)

Factor\Kano category	B	O	I	Q	A	R
User-friendly platform	2	3	7	0	4	0
Certificate of completion	0	0	13	0	3	0
Download of course material	3	1	6	0	4	2
Own profile and account page	0	0	13	0	3	0
Quizzes and exercises	2	0	10	0	2	2
Interactive quizzes and exercises	0	1	14	0	1	0
Comment function	0	1	14	0	1	0
Personal tutor	0	1	9	0	4	1
User manual for the platform	4	1	10	0	1	0
Videos	0	0	9	0	7	0
Photos/Graphics	0	2	9	0	5	0
Text	2	2	11	0	1	0

Table 2. Kano model evaluation ex-post (consumption experience)

Factor\Kano category	B	O	I	Q	A	R
User-friendly platform	0	5	9	0	2	0
Certificate of completion	0	2	14	0	0	0
Download of course material	1	2	11	0	2	0
Own profile and account page	0	1	12	0	1	2
Quizzes and exercises	1	1	9	0	3	2
Interactive quizzes and exercises	0	1	12	0	2	1
Comment function	0	0	15	0	1	0
Personal tutor	0	0	12	0	4	0
User manual for the platform	1	4	8	0	3	0
Videos	0	1	9	0	6	0
Photos/Graphics	0	2	7	0	7	0
Text	2	2	7	1	4	0

The machine learning approach resulted indicative answers for *need*, *like*, *dislike*, and *wish* as listed below with corresponding similarity score.

- *need*: video (0.35), understand (0.29), test (0.26)
- *like*: professional (0.34), photography (0.24), {restart/test (0.21/0.21)}
- *dislike*: manual (0.39), complex (0.28), frustration (0.25)
- *wish*: explanation (0.36), hint (0.24), video (0.20)

Using this additional information, the two heuristic approaches can be combined to provide a stronger quantitative result.

As the Word2Vec tool also allows vector arithmetic on multidimensional word vectors, two more aspects were taken into account as listed below.

- *need* – *wish* (expectation removed consumption experience):
  system (0.22), motivation (0.16), support (0.16)
- *wish* – *need* (consumption experience removed expectation):
  explanation (0.43), quiz (0.32), manual (0.18)

Given the small text corpus and takeing the results form the Word2Vec alone, it would be difficult to interpret the findings. However, taking the classifications from the Kano model into account, additional inferences can be made.

A clear emphasis can be seen on multimedia, especially on videos. This is supported by Tables 1 and 2 as well as Word2Vec's *need* and *wish* vector similarities.

A user manual for the online course site was as well supported by the machine learning algorithm. A support in the form of a manual or personal tutor could also be found.

The attractiveness of retaking quizzes was a new insight that will be considered for an extended Kano model questionnaire in the future. This has to be balanced, however with motivational factors and frustration avoidance.

The findings are still strongly dependent on human interpretation. The authors will expand the experiments to include more participants, but also the amount of textual data for each participant for future research.

This novel approach allows large-scale surveys with a scalable AI implementation. Researchers and practitioners can support their questionnaire findings with word-embeddings to identify the gap between product expectation and consumption experience.

4. **Conclusions.** Both the Kano model and Word2Vec have been successfully implemented in their respective fields. While the former is used in customer satisfaction research for more than four decades, the latter is a recent addition to text corpora analysis.

The combination of the Kano model and Word2Vec provides a novel way to analyze student perceptions of online courses. The machine learning algorithm can support findings from the Kano categorization matrix. The numerical word relationships can also help with a deeper insight into student expectation and consumption experience by priming the ex-ante and ex-post texts.

Future research will include an extended list of features to be analyzed by the Kano method, as well as larger text corpora for Word2Vec training.

REFERENCES

[1] Y. Chen and M. Zhang, MOOC student dropout: Pattern and prevention, *Proc. of the ACM Turing 50th Celebration Conference – China ACM (TUR-C'17)*, New York, NY, USA, 2017.

[2] T. R. Liyanagunawardena, P. Parslow and S. A. Williams, Dropout: MOOC participants' perspective, *The 2nd MOOC European Stakeholders Summit (EMOOCs2014)*, pp.95-100, 2014.

[3] C. Chen, G. Sonnert, P. M. Sadler, D. D. Sasselov, C. Fredericks and D. J. Malan, Going over the cliff: MOOC dropout behavior at chapter transition, *Distance Education*, vol.41, no.1, pp.6-25, 2020.

[4] J. Goopio and C. Cheung, The MOOC dropout phenomenon and retention strategies, *Journal of Teaching in Travel & Tourism*, pp.1-21, 2020.

[5] P.-C. Sun, R. J. Tsai, G. Finger, Y.-Y. Chen and D. Yeh, What drives a successful e-learning? An empirical investigation of the critical factors influencing learner satisfaction, *Computers & Education*, vol.50, no.4, pp.1183-1202, 2008.

[6] M. G. Violante and E. Vezzetti, Virtual interactive e-learning application: An evaluation of the student satisfaction, *Computer Applications in Engineering Education*, vol.23, no.1, pp.72-91, 2015.

[7] G. Dominici and F. Palumbo, How to build an e-learning product: Factors for student/customer satisfaction, *Business Horizons*, vol.56, no.1, pp.87-96, 2013.

[8] Y.-S. Wang, S. Bauk, S. Šćepanović and M. Kopp, Estimating students' satisfaction with web based learning system in blended learning environment, *Education Research International*, 2014.

[9] M. Nissim, R. van Noord and R. van der Goot, Fair is better than sensational: Man is to doctor as woman is to doctor, *Computational Linguistics*, vol.46, no.2, pp.487-497, 2020.

[10] A. R. Rodgers and M. Puterbaugh, Digital badges and library instructional programs: Academic library case study, *Journal of Electronic Resources Librarianship*, vol.29, no.4, pp.236-244, 2017.

[11] J. Zou and L. Schiebinger, AI can be sexist and racist – It's time to make it fair, *Nature*, vol.559, no.7714, pp.324-326, 2018.

[12] E. Altszyler, S. Ribeiro, M. Sigman and D. F. Slezak, The interpretation of dream meaning: Resolving ambiguity using latent semantic analysis in a small corpus of text, *Consciousness and Cognition*, vol.56, pp.178-187, 2017.

[13] Y. Zhu, E. Yan and F. Wang, Semantic relatedness and similarity of biomedical terms: Examining the effects of recency, size, and section of biomedical publications on the performance of Word2Vec, *BMC Medical Informatics and Decision Making*, vol.17, no.1, 2017.

[14] G. Di Gennaro, A. Buonanno and F. A. N. Palmieri, Considerations about learning Word2Vec, *The Journal of Supercomputing*, 2021.

[15] N. Kano, N. Seraku, F. Takahashi and S. I. Tsuji, Attractive quality and must-be quality, *Journal of the Japanese Society for Quality Control*, vol.14, no.2, pp.147-156, 1984.

[16] A. Parasuraman, V. A. Zeithaml and L. Berry, SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality, *Journal of Retailing*, vol.64, pp.12-40, 1988.

[17] Y.-S. Wang, H.-Y. Wang and D. Y. Shee, Measuring e-learning systems success in an organizational context: Scale development and validation, *Computers in Human Behavior*, vol.23, no.4, pp.1792-1808, 2007.

[18] A. Parasuraman, V. A. Zeithaml and A. Malhotra, E-S-QUAL: A multiple-item scale for assessing electronic service quality, *Journal of Service Research*, vol.7, no.3, pp.213-233, 2005.

[19] A. Chaudha, R. Jain, A. R. Singh and P. K. Mishra, Integration of Kano's model into quality function deployment (QFD), *The International Journal of Advanced Manufacturing Technology*, vol.53, no.5, pp.689-698, 2011.

[20] T. Mikolov, K. Chen, G. Corrado and J. Dean, Efficient estimation of word representations in vector space, *Proc. of the 1st International Conference on Learning Representations*, 2013.

[21] C. C. Aggarwal, *Neural Networks and Deep Learning*, Springer International Publishing AG, 2018.

[22] S. Skansi, *Introduction to Deep Learning – From Logical Calculus to Artificial Intelligence*, Springer International Publishing AG, Part of Springer Nature 2018, 2018.

[23] G. Finley, S. Farmer and S. Pakhomov, What analogies reveal about word vectors and their compositionality, *Proc. of the 6th Joint Conference on Lexical and Computational Semantics (SEM2017)*, Vancouver, Canada, 2017.

[24] P. Qi, Y. Zhang, Y. Zhang, J. Bolton and C. D. Manning, Stanza: A python natural language processing toolkit for many human languages, *Proc. of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 2020.

[25] G. Hofstede, G. J. Hofstede and M. Minkov, *Cultures and Organizations: Software of the Mind*, 3rd Edition, McGraw-Hill, 2010.