ISSN 2355-0082

Cross-Platform Mobile Based Crowdsourcing Application for Sentiment Labeling Using Gamification Method

Elaine[1], Farica Perdana Putri[2], Alethea Suryadibrata[3]
Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia
[1]elaine@student.umn.ac.id, [2]farica@umn.ac.id, [3]alethea.suryadibrata@lecturer.umn.ac.id

Accepted 17 December 2024
Approved 17 January 2025

***Abstract*— Sentiment analysis is the application of natural language processing which aims to identify the sentiment of texts. To carry out sentiment analysis, data which has been labeled sentiment is needed to be included in the training model. Crowdsourcing is considered as the most optimal method to label data because it has a high level of accuracy at a relatively low cost. However, the use of crowdsourcing platforms has its own challenge, which is to increase user interest and motivation. A solution which can be applied is to design and build a crowdsourcing platform or application using the gamification method. The definition of gamification is an effort to increase one's intrinsic motivation for an activity by applying game elements to it. Therefore, a cross-platform mobile based crowdsourcing application for sentiment labeling using gamification method was carried out. The gamification design process was done based on the 6D framework and the application was developed using the Ionic-React framework. Application was examined through black box testing and the result showed that the application was functioning properly and according to the design requirements. There was also an evaluation carried out by distributing Intrinsic Motivation Inventory questionnaires to users who had used the application for 2 weeks. From a total of 40 respondents, the result showed that the level of user motivation and interest in using the application was high with a percentage of 83.10%.**

***Index Terms*— Crowdsourcing; Gamification; Ionic-React Platform; Sentiment Labeling**

I. INTRODUCTION

As human civilization and information technology develop, expectations arise for computer interfaces so that they can communicate with humans using language which is easily understood by humans [1]. Therefore, natural language processing (NLP) is needed to achieve these expectations. NLP is a branch of computer science in the field of artificial intelligence which studies the interaction between computers and human language to assist computers in understanding, interpreting, and manipulating natural language or the language used by humans so that it can be used practically [2]. NLP combines computational linguistics or rule-based modeling of human language with statistical, machine learning, and deep learning models which enable computers to process human language and understand the meaning, intent, and sentiments it embodies [3].

Sentiment analysis is the practice of applying NLP in identifying and obtaining subjective information from texts to determine human opinions, attitudes, and emotions towards an individual, event, or topic [4, 5]. The main objective of sentiment analysis is to identify the sentiments of a text and determine its polarity, whether the text has positive, negative, or neutral sentiments. Sentiment analysis plays an important role in many fields, including analyzing reviews, political views, economic market developments, historical movements, and also emotional level of social media users [6]. In fact, based on data obtained in April 2013, it was known that 90% of customers decided to buy based on online reviews [4]. Not only for customers, sentiment analysis is also needed by companies to monitor discussions and evaluations regarding the company's reputation and products issued by the company. Therefore, sentiment analysis is one of the areas which has been extensively researched regarding decision-making process.

In carrying out sentiment analysis, a data set which has been labeled with sentiment is needed to be included in the training model. The process of sentiment labeling can be done automatically or manually [7,8]. Automatic labeling can be done using dictionaries or machine learning, while manual labeling can be done with the help of expert or with the help of groups of people by crowdsourcing. Based on the research [9], it was found that the sentiment analysis where the labeling was done automatically had the lowest accuracy with the highest alpha value of 0.50 and the sentiment analysis where the labeling was done by experts had the highest accuracy with the highest alpha value of 0.90. Sentiment analysis where the labeling was done by crowdsourcing had a relatively good level of accuracy with the highest alpha value of 0.81. Beyond performance, there were cost considerations where labeling by an expert was significantly more expensive than the other two

methods and could increase as more data became available [9]. Therefore, sentiment labeling by crowdsourcing is considered to be the most optimal method.

Based on the availability of wages, crowdsourcing platforms can be classified into paid crowdsourcing and unpaid crowdsourcing [10]. Paid crowdsourcing may tend to be more desirable but requires more capital from the employer. Meanwhile, unpaid crowdsourcing depends on the workload and the identity of the employer, so there is a risk of a lack of interest from the participants because the task is considered difficult and only benefits the employer. In overcoming this problem, a solution which can be applied is to use gamification method or application of game mechanisms to increase the motivation of the participants through a fun game experience [8, 11]. In addition, the gamification method is considered to be able to improve the process and the end result of the activity.

There was a previous research, namely the development of a web system for crowdsourcing opinion labeling using the gamification method, which was carried out by Cirqueira et al in 2017. The evaluation in this research focused on aspects related to the functionality and ease of use of the system as well as the clarity of the game mechanism. The results obtained were 2 aspects considered good and 7 aspects considered very good. Based on the results of this study, Cirqueira et al made several suggestions for further research, including providing a feature for downloading labeled data sets, implementing a global leaderboard to motivate players, using more appropriate colored buttons, and developing a better badge system [8].

This research was appointed after considering the things which had been described. The application was built as a mobile cross-platform based application so that it could be accessed as a web via a browser or downloaded to a device, according to user needs. In addition, the application did not use much device features, so it did not need to be built as a native application. The application was built specifically for text sentiment labeling in Indonesian and English, was designed using the 6D gamification framework, and was built using Ionic-React framework. The objectives of this research were to design and develop a cross-platform mobile based crowdsourcing application for sentiment labeling using gamification method and to determine the level of motivation and interest of users in using the application.

II. Literature Review

Sentiment Analysis

Sentiment analysis is a type of text classification which focuses on sentiment orientation analysis which shows individual polarity towards a topic [12]. The polarity classification in sentiment analysis can be divided into a number of different polarity classes, such as two classes (positive and negative), three classes (positive, neutral, and negative), as well as a more specific classification into five classes (very positive, positive, neutral, negative, very negative) [4]. In addition, in sentiment analysis, there are three main classification levels, namely the document level, sentence level, and aspect level [5]. Document level analysis analyzes a document as a unit of basic information. Sentence level analysis analyzes based on opinion on one sentence. Aspect level analysis considers that one person may give different opinions on different aspects of an entity.

Sentiment analysis techniques can be classified into three types of approaches which are machine learning approaches, lexicon-based approaches, and hybrid approaches [5]. The machine learning approach uses linguistic features in machine learning algorithms. The lexicon-based approach relies on a known and pre-arranged set of sentiment terms. The hybrid approach combines the other two approaches in which the sentiment lexicon generally plays the most important role.

Based on the presence or absence of supervision, sentiment analysis using a machine learning approach can be classified into supervised learning methods and unsupervised learning methods. The supervised method makes use of a large number of training documents which have been labeled with sentiments, whereas the unsupervised method is used when the labeled document or data set is hard to find. Based on data sources, the lexicon-based approach can be classified into two methods, namely the dictionary-based approach and the corpus-based approach. Both lexicon-based approaches rely on finding a lexicon of opinion which is used to analyze the text. Therefore, regardless of which approach is used, the data set used is an important factor in sentiment analysis [5].

Crowdsourcing

Crowdsourcing is a type of online activity in which a party invites the participation of a group of individuals through an open call to do a task [10]. The four pillars which underlie crowdsourcing include the employer, individual groups, tasks, and platforms [13]. Two main aspects which need to be considered in implementing crowdsourcing are that crowdsourcing must be able to motivate groups of individuals in carrying out tasks and that crowdsourcing implementation must be able to achieve the required solution [14].

To mediate between employers and community groups, a platform is needed in the form of an application which provides functions for carrying out tasks and has a group management system . Based on the types of functions provided, crowdsourcing platforms can be classified into two types, namely

applications with specific functions, such as InnoCentive and ClickWorker, and applications with general functions, such as Amazon Mechanical Turk (MTurk), Microworker, and CrowdFlower [10]. Based on task specifications, crowdsourcing platforms are classified into applications which perform microtasks or tasks which are simple, repetitive, independent, and short, such as labeling tasks, and applications which perform macrotasks or tasks which depend on context, require a lot of effort, and take a long time, such as code writing assignments. In addition, crowdsourcing platforms can be differentiated based on the presence or absence of wages. Procurement of wage rewards is usually used to encourage people to participate and contribute to crowdsourcing. However, not only motivated by money, research by Kaufmann et al. in 2011 found that crowdsourcing participants at MTurk were also motivated by the sense of satisfaction, togetherness, and pride which they got from doing tasks [15].

In the implementation of crowdsourcing, the nature and complexity of the task has a significant impact on the level of community participation and motivation so it is important for a crowdsourcing project to have an effective task design. Several things need to be considered when the task is to be designed, namely the description of the task, the scope of the task, the skills needed to complete the task, the division of tasks if necessary, the need for solutions, the targets of participants, and the duration of work. In addition, it is also important to identify and analyze the relationship between elements of task design and the ultimate goal of crowdsourcing [14].

Gamification

Gamification is an effort to increase intrinsic motivation for various activities by implementing game elements into the design. Examples of applicable game elements are point system, leaderboards, awards, badges, bonuses, targets, and narrative [16, 17]. Gamification plays an important role in triggering game-like psychological experiences by leveraging elements found in games. The main goal of gamification is to increase the user's positive motivation for a given activity so that the quantity and quality of the results of these activities can also increase [18].

In general, gamification approaches follow a process which can be divided into seven phases. First, the project preparation phase, namely all activities which need to be carried out before the project starts. Second, the analysis phase, namely the process of identifying problems related to the users, the processes, and the project. Third, the idea generation phase, namely the idea search stage related to the gamification design. Fourth, the design phase, namely the activity of designing a gamification approach and prototyping. Fifth, the implementation phase, namely the application of the gamification approach which was previously designed. Sixth, the evaluation phase, which is the activity of testing and assessing the results of implementing the gamification approach. Finally, the observation phase, which is to monitor the result of the implementation of the gamification approach on users.

The application of gamification needs to be well designed to achieve the desired positive impact, so it is important to choose the right framework [19]. A literature study conducted in 2017 stated that there were 40 recorded gamification design frameworks. One of the most popular and most widely referenced gamification design frameworks in research is the 6D framework initiated by Werbach and Hunter [20].

The 6D framework consists of the following six stages [21, 22]. First, define business objectives or the stage where the final goal and the specific positive impact expected from the application are determined. Second, define the target behavior or the stage where the expected player behavior and how the system should support and provide feedback on this behavior are determined. Third, describe the players or the stage where the characteristics of the players are determined. Fourth, develop the activity loop or the stage where it is planned how to motivate players by using engagement and progress cycles, how to attract new players, and how to encourage further action from players. Fifth, don't forget the fun or the stage where the application is reviewed to identify aspects which motivate players intrinsically. Finally, deploy appropriate tools or the stage where mechanics, metrics, and tools, including relevant game elements, are defined.

Intrinsic Motivation Inventory

Intrinsic Motivation Inventory (IMI) is a multidimensional measurement tool used to measure participants' subjective experiences related to specific activities in research. Based on the research by McAuley et al. in 1987 and by Goudas et al. in 2000, it was concluded that the IMI model was a valid measurement tool, in accordance with common factors, and was three-dimensional, supported by strong evidence [23, 24]. Monteiro et al conducted research on the level of reliability of each aspect in the IMI model in 2015 and obtained Cronbach's alpha values ranging from 0.82 to 0.91, stated that the IMI model showed adequate validity and reliability. IMI has been used in many studies focusing on intrinsic motivation in various fields where the question models used apply different aspects and questions of IMI depending on the characteristics of task and participants.

The IMI model measures seven aspects, namely interest/enjoyment, perceived competence, perceived choice, pressure/tension, effort/importance, value/usefulness, and relatedness [23]. Interest is a

form of personal assessment of intrinsic motivation and it is considered as the main aspect of IMI and often has more questions. Perceived competence and perceived choice are used as positive predictors in measuring intrinsic motivation, while pressure is a negative predictor of intrinsic motivation. Effort is a separate variable which is considered relevant and not necessarily but can be included in the IMI model to measure the level of motivation related to a specific issue or context [23, 24]. The value aspect aims to study the participants' self-internalization where participants are more likely to live up to the tasks being done if the participants judge that the task has benefits or is of value to the participants. There is also an aspect of relatedness which is used to measure interpersonal interactions between participants.

Each aspect measured in IMI has a different number of question choices with a total of 45 questions [23]. For each of these questions, the user's response was measured using a 7-point Likert scale where 1 is “not at all true” and 7 is “very true”. Researchers may choose the aspects to be measured according to the relevance of the aspects to the research being conducted [23]. In designing a list of questions using the IMI model, measured aspects are first determined based on variables in the problem formulation then which questions to use are determined based on these aspects in random order [23]. In measuring the result of an IMI question model, it is done by calculating the average value of each aspect, then the average results can be used according to the purposes of analysis in research.

Likert Scale

The Likert scale was first mentioned by Rensis Likert in 1932 on his research article "A Technique for the Measurement of Attitudes" in the journal "Archive of Psychology" [25]. The Likert scale was designed to measure attitudes in a scientifically acceptable and validated way [26]. Likert in his research discussed the possibility of classifying an infinite number of attitudes of a person into groups of responses [26]. In applying the Likert scale, participants were asked to indicate their level of conformity with the statements given on the metric scale [26]. Likert stated that two polar choices on the scale needed to be assigned an exact value, while the choices between them were left without an exact value [27].

Likert himself did not consider the number of choices to be important and implied that each researcher was free to determine it according to their individual needs [27]. However, several studies stated that the 7-point Likert scale showed better performance compared to the 5-point scale. The 7-point scale has more choices of responses which can trigger the respondent's reasoning abilities and increase the possibility of achieving objectivity [26]. As a result, the 7-point scale produces data with lower measurement errors and a higher level of accuracy [27].

To apply the result of Likert scale measurements to a case, class interval calculations can be performed [28]. The interval class formula can be seen in Equation 1 [29].

$$i = \frac{x_n - x_1}{k} \qquad (1)$$

Where i = class interval, x_n = highest value, x_1 = lowest value, and k = number of classes.

III. Research Method

Requirements Analysis

In the requirements analysis stage, a simple literature study was carried out regarding sentiment analysis, especially about sentiment labeling, regarding crowdsourcing, regarding gamification methods, and regarding Likert scale to find out what the application needs in order to achieve the stated research objectives.

Application Design

At the design stage, it was determined thoroughly, in detail, and specifically about the application to be built. It included functions, features, and how game elements are implemented. The design process was done using the 6D gamification framework. The expected design results were in the form of design report, flowcharts, relational database scheme, mockup, and questionnaire design.

Application Development

At the application development stage, based on the design that had been made, the application was built using the Ionic-React framework through a source-code editor application.

Application Testing

This stage was carried out using the black box testing method. The purpose of the black box testing method was to test the external functions of the application in accordance with the design requirements which had been determined. Black box testing was done by trying each feature in various scenarios as input to observe whether the application was functioning and giving the expected results or output.

Evaluation

The evaluation stage was carried out with a quantitative approach. Sample participants were taken randomly from people who were willing. The evaluation began by giving application access to participants so that participants could try using the application directly. After the usage time period of 2 weeks had passed, each participant was given a motivational and interest level questionnaire related to

crowdsourcing activities for sentiment labeling in the application. The questionnaire was prepared based on the Intrinsic Motivation Inventory (IMI) question model and was measured using a Likert scale. The IMI question model was chosen after considering the validity and reliability of the IMI as an adequate measuring tool. Responses to the questionnaire on levels of motivation and interest were then calculated and analyzed to determine the effectiveness of the application which has been developed.

IV. Results and Discussion

Gamification Design

The application design based on the 6D gamification framework by Werbach and Hunter can be seen as follows:

- Define business objectives

The purpose of gamification is to increase user motivation and interest in sentiment labeling activities. The expected benefits are the increase in quantity of training data or data which has been labeled with sentiment.

- Delineate target behavior

The expected target behavior is that users are motivated and interested in engaging in sentiment labeling activities. The achievement of the target behavior is measured by giving a questionnaire compiled based on the Intrinsic Motivation Inventory (IMI) model to users to measure the level of intrinsic motivation of users.

- Describe the players

Target users are individuals from various demographic groups who understand Indonesian or English well and understand the concept of sentiment so that they can provide appropriate sentiment labels.

- Devise activity loop

The engagement cycle applied, namely: (1) Users are motivated to increase player levels, enter the leaderboards, or get other rewards; (2) Users participate in sentiment labeling activities; (3) The system provides feedback in the form of points to users; (4) Points impact the final game score which is then calculated into the player level and leaderboard. The stages of progress applied are: (1) Users have a short-term goal to get the best final score and enter the leaderboard; (2) Users have long-term goals of increasing levels, achieving various achievements, obtaining various badges, obtaining mysterious gifts, and completing profiles.

- Don't forget the fun

Aspects which motivate intrinsically, namely: (1) The hard fun aspect is implemented through sentiment labeling activities in the form of games where users need to consider carefully which sentiment label is the most appropriate and most likely to be chosen by other users to get points; (2) The easy fun aspect is implemented through achievement system, badge system, and mysterious reward system which can be obtained by completing certain relatively simple tasks, which in addition, badges and other rewards, such as profile icons, profile frames, and namecards, can be used to decorate and complete user profiles; (3) The aspect of altered states is implemented through sentiment labeling activities in the form of games where users can encounter various sets of statements which are always different; (4) The aspect of the people factor is implemented through leaderboard system which compares and displays the final scores and user profiles.

- Deploy appropriate tools

Based on the previous stages, the following were determined. The application has a game language filter feature which consists of three choices, namely Indonesian, English, and Indonesian-English. Game elements applied to application gamification, namely point system, player level, leaderboards, achievements, badges, mysterious reward, and player profile customization.

Questionnaire Design

To measure the level of user motivation and interest in the application, a questionnaire was designed based on IMI model which can be seen as follows. Questions with the code "(R)" are reverse coded questions in which the score data from the questionnaire responses for these questions needs to be calculated first with the formula, which is 8 minus the original score, then the calculation results can be processed as usual.

- Interest/enjoyment

I enjoyed doing this activity very much;
I would describe this activity as very interesting;
I thought this activity was quite enjoyable.

- Perceived competence

After working at this activity for awhile, I felt pretty competent;

I am satisfied with my performance at this task;
I was pretty skilled at this activity.

- Perceived choice

I believe I had some choice about doing this activity;
I did this activity because I wanted to;
I did this activity because I had to (R).

- Value/usefulness

I think that doing this activity is useful;
I would be willing to do this again because it has some value to me;
I believe doing this activity could be beneficial to me.

Application Results

The application can be seen in Figure 1. Users who have logged in to their accounts will be directed to the home page which includes options to play and options to access other menus. When the user chooses to play, the user is shown a list of sets along with general information about the sets. After selecting a set, the user enters the game. When the user has finished specifying the sentiment label for each given data, the user is redirected to the results page which displays the final score and game leaderboard.

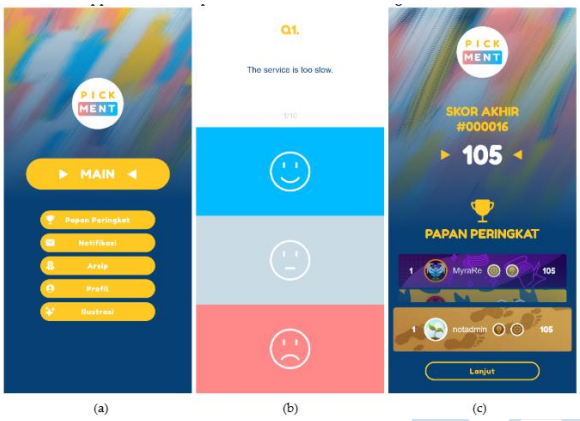

Fig 1. Screenshots of Application: (a) Home; (b) Game; (c) Leaderboard

Application Testing

The results of application functionality tests carried out using the black box testing method can be seen in Table 1.

TABLE I. BLACK BOX TESTING RESULTS

Function Name	Result
Welcome Page	Works Accordingly
Register	Works Accordingly
Login	Works Accordingly
Home page	Works Accordingly
Dataset List Page	Works Accordingly
Sentiment Labeling Game	Works Accordingly
Set Leaderboard	Works Accordingly
Weekly Leaderboard	Works Accordingly
Notification	Works Accordingly
Game History	Works Accordingly
Achievement	Works Accordingly
Badge	Works Accordingly
Profile	Works Accordingly
Edit Username	Works Accordingly
Edit Profile Badge	Works Accordingly
Edit Profile	Works Accordingly
Edit User Information	Works Accordingly
Settings	Works Accordingly
Upload Dataset	Works Accordingly
Uploaded Dataset List Page	Works Accordingly
Dataset Labeling Result Page	Works Accordingly
Illustration	Works Accordingly

Evaluation

To determine the level of user motivation and interest obtained from measurements using Likert scale, the average index of each aspect of the Intrinsic Motivation Inventory (IMI) and the overall average index are categorized into 7 class intervals. Therefore, it is necessary to find out in advance the value of the interval length through Equation 1 (interval class formula) which results in an interval length value of 0.86. Based on the obtained value, interval classes to determine the level of user motivation can be seen in Table 2.

TABLE II. USER MOTIVATION LEVEL INTERVAL CLASSES

Class Interval	Description
1-1.85	Very Low
1.86-2.71	Low
2.72-3.57	Somewhat low
3.58-4.45	Neutral
4.46-5.29	Somewhat High
5.30-6.15	High
6.16-7	Very High

By comparing to the predetermined interval classes, the results of measuring the motivation level of application users can be seen in Table 3. The measurement results were obtained through questionnaire responses from a sample of 40 users who had tried the application during a 2-week usage period.

TABLE III. USER MOTIVATION LEVEL INTERVAL CLASSES

Aspect	Average	Standard Deviation	Percentage	Desctiption
Interest/enjoyment	6.06	0.82	86.55 %	High
Perceived competence	5.57	1.07	79.52 %	High
Perceived choice	6.17	0.78	88.10 %	Very High
Value/usefulness	5.48	1.12	78.21 %	High
All aspects	5.82	0.75	83.10 %	High

From Table 3, it can be seen that the level of motivation of application users based on all aspects of IMI had an average value of 5.82 with a standard deviation of 0.82 and a percentage of 83.10%. The value of 83.10% itself was obtained by calculating the average of the percentage of each aspect. Several things which need to be considered, namely the perceived choice aspect was the aspect with the highest average value of 6.17 and the value/usefulness aspect was the aspect with the lowest average value of 5.48. There was also the interest/enjoyment aspect as the main determining aspect in the IMI model which had an average value of 6.06 and was the aspect with the second highest average value. In addition, it can be seen that 3 of the 4 total aspects of the level of motivation considered had an average value which was included in the high category and 1 aspect of another level of motivation, namely the aspect of perceived choice, was included in the very high category. Based on these data, it can be stated that the level of user motivation and interest in using the application was relatively high.

V. CONCLUSION

Based on the results and discussion, the following conclusions can be drawn. The design and development of a cross-platform mobile based crowdsourcing application for sentiment labeling using gamification method has been well completed. The application was designed by following the 6D gamification framework and built using Ionic-React framework in a source-code editor application. The gamification elements applied to the application included point system, player level, leaderboards, achievements, badges, mysterious reward, and player profile customization.

Application testing was carried out through black box testing and the result showed that all the features tested have worked according to the design requirements. Results also showed that the level of user motivation and interest in using the application was high. Determining the level of user motivation and interest was done by analyzing response data from a questionnaire created using the Intrinsic Motivation Inventory (IMI) model with a 7-point Likert measurement scale. Questionnaires were distributed to users after 2 weeks of application usage period. From a total of 40 respondents from various backgrounds, the result was that the average level of user motivation based on all aspects reached a percentage of 83.10% which could be classified as high.

ACKNOWLEDGMENT

Authors thank Universitas Multimedia Nusantara for sponsoring the research.

REFERENCES

[1] Y. Goldberg, Neural Network Methods in Natural Language Processing, G. Hirst, Ed., Morgan & Claypool Publishers, 2017.

[2] D. Oliva and E. H. Houssein, Metaheuristics in Machine Learning: Theory and Applications, S. Hinojosa, Ed., Springer International Publishing, 2021.

[3] A. Saxena, A. K. Bhatt and R. Kumar, Artificial Intelligence for Societal Development and Global Well-Being, IGI Global, 2022.

[4] D. M. E.-D. M. Hussein, "A survey on sentiment analysis challenges," Journal of King Saud University - Engineering Sciences, vol. 30, no. 4, pp. 330-338, 7 2018.

[5] W. Medhat, A. Hassan and H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, vol. 5, no. 4, pp. 1093-1113, 7 2014.

[6] B.L. Supriyatna and F. P. Putri "Optimized Support Vector Machine for Sentiment Analysis of Game Reviews," International Journal of Informatics and Communication Technology (IJ-ICT), Vol. 13, No. 3, pp 344-353, 2024.

[7] G. Haralabopoulos and E. Simperl, "Crowdsourcing for Beyond Polarity Sentiment Analysis A Pure Emotion Lexicon," 7 2017.

[8] D. Cirqueira, L. Vinícius, M. Pinheiro, A. J. Junior, F. Lobato and Á. Santana, "Opinion Label: A Gamified Crowdsourcing System for Sentiment," Companion Proceedings of the 23rd Brazilian Symposium on Multimedia and Web, pp. 209-213, 2017.

[9] W. van Atteveldt, M. A. C. G. van der Velden and M. Boukes, "The Validity of Sentiment Analysis: Comparing Manual Annotation, Crowd-Coding, Dictionary Approaches, and Machine Learning Algorithms," Communication Methods and Measures, vol. 15, no. 2, pp. 121-140, 7 2021.

[10] R. M. Borromeo and M. Toyama, "An investigation of unpaid crowdsourcing," Human-centric Computing and Information Sciences, vol. 6, no. 1, p. 11, 7 2016.

[11] C. Busch, ECGBL2014-8th European Conference on Games Based Learning: ECGBL2014, Academic Conferences and Publishing International, 2014.

[12] H. Agatha, F. P. Putri, "Sentiment Analysis on Song Lyrics for Song Popularity Prediction Using BERT Algorithm," Ultimatics : Jurnal Teknik Informatika, vol. 15, no. 2, pp 99-105, 2023

[13] M. Hosseini, K. Phalp, J. Taylor and R. Ali, "The four pillars of crowdsourcing: A reference model," 2014 IEEE Eighth International Conference on Research Challenges in Information Science (RCIS), pp. 1-12, 7 2014.

[14] R. Karachiwalla and F. Pinkow, "Understanding crowdsourcing projects: A review on the key design elements of a crowdsourcing initiative," Creativity and Innovation Management, vol. 30, no. 3, pp. 563-584, 7 2021.

[15] H. Xia, Y. Huang and J. Stromer-Galley, "A Collective Action Inspired Motivation Framework for Crowdsourcing: A Collective Action Inspired Motivation Framework for Crowdsourcing," vol. 13, no. 1, pp. 41-48, 2016.

[16] S. S. Setiawan and A. Suryadibrata, "Fitrust: Promoting Healthy Lifestyle Through Gamified Mobile Health Application," 5th International Conference on New Media Studies (CONMEDIA), 2019.

[17] E. Sanchez, H. van Oostendorp, J. D. Fijnheer and E. Lavoué, "Gamification," Encyclopedia of Education and Information Technologies, pp. 816-827, 2020.

[18] B. Morschheuser, K. Werder, J. Hamari and J. Abe, "How to gamify? A method for designing gamification," Proceedings of the 50th Hawaii International Conference on System Sciences, pp. 1298-1307, 2017.

[19] W. Oliveira, A. M. Toda, P. T. Palomino, L. Rodrigues and S. Isotani, "Which one is the best? A quasi-experimental study comparing frameworks for unplugged gamification," RENOTE, vol. 18, no. 1, 7 2020.

[20] A. Mora, D. Riera, C. González and J. Arnedo-Moreno, "Gamification: a systematic review of design frameworks," Journal of Computing in Higher Education, vol. 29, no. 3, pp. 516-548, 7 2017.

[21] F. Silva, A. Toda and S. Isotani, "Towards a link between Instructional Approaches and Gamification - A Case Study in a Programming Course," Anais do XXIV Workshop de Informática na Escola (WIE 2018), pp. 157-165, 7 2018.

[22] A. Odincova, "Using Gamification to Motivate Self-Regulated Learning in Adults. Research through Design: Developing a Quiz-Based Learning Tool," 7 2017.

[23] "Intrinsic Motivation Inventory (IMI)," Center of Self-Determination Theory.

[24] V. Monteiro, L. Mata and F. Peixoto, "Intrinsic Motivation Inventory: Psychometric Properties in the Context of First Language and Mathematics Learning," Psicologia: Reflexão e Crítica, vol. 28, no. 3, pp. 434-443, 7 2015.

[25] Z. Awang, A. Afthanorhan and M. Mamat, "The Likert scale analysis using parametric based Structural Equation Modeling (SEM)," Computational Methods in Social Sciences, vol. 4, pp. 13-21, 7 2016.

[26] A. Joshi, S. Kale, S. Chandel and D. K. Pal, "Likert scale: Explored and explained," British journal of applied science & technology, vol. 7, no. 4, p. 396, 2015.

[27] J. Munshi, "A Method for Constructing Likert Scales," SSRN Electronic Journal, 2014.

[28] A. M. Soamole, A. K. Hadi and Watono, "Analisis Penerapan Sistem Manajemen Mutu ISO 9001 pada PT. Semen Tonasa," Journal Flyover, vol. 2, no. 1, 2022.

[29] A. D. Miqdad, "Analisis pengembangan model pembelajaran berbasis Project Based Learning dengan berbantu media audio visual VideoScribe dalam Pembelajaran Komputer dan Jaringan dasar kelas X SMK Unitomo Surabaya," IT-Edu: Jurnal Information Technology and Education, vol. 2, no. 02, 2017.