Emotion Detection based on Column Comments in Material of Online Learning using Artificial Intelligence

Irawan Dwi Wahyono[1](✉), Djoko Saryono[1], Hari Putranto[1], Khoirudin Asfani[1],
Harits Ar Rosyid[1], Sunarti[1], Mohd Murtadha Mohamad[2],
Mohd Nihra Haruzuan Bin Mohamad Said[2], Gwo Jiun Horng[3], Jia-Shing Shih[3]
[1]Universitas Negeri Malang, Malang, Indonesia
[2]Universiti Teknologi Malaysia, Johor Bahru, Malaysia
[3]Southern Taiwan University of Science and Technology, Tainan, Taiwan
`irawan.dwi.ft@um.ac.id`

Abstract—Many universities use online learning as media learning that each material of media which includes videos, textual content, or audio may be given remarks from college students. The lecture desires to recognize approximately the feelings of college students which include happy, disappointed, or unhappy when they accessed the media and instructors get an assessment of pleasant from their media. This study constructed a utility cellular for the detection of emotion from column remarks in the media online. The mobile application makes use of synthetic intelligence to type textual content from remarks and to decide the emotion of college students. The mobile application on a cellular device. The set of rules with inside the utility is k-Nearest Neighbour for the textual content mining feature in this study. The information of trying out these studies is commenting on YouTube channels and online studying which include SIPEJAR. The result of trying it out is that the common accuracy is 0,697, the value of recall is 0.5595, and the common precision is 0, 4421 and the accuracy for the utility of this mobile app is 70% for detection emotion-primarily based totally on a column of remark in the media online.

Keywords—paper publishing, online journals, styles, how-to

1 Introduction

In times of a pandemic, all schools use online learning for home learning. The teacher has created a multimedia tutorial such as video, audio, and documents for the online class. The teachers have just created a medium with their version. They do not care whether the media is appropriate or not to have an impact on students [1–3]. The effects are sometimes worse, which students cannot understand because the media is not interesting, boring, or moody. This issue can be resolved if the teacher receives feedback from the student after the students have viewed or read the media. The comments are obtained in the comments of columns online like Youtube or in the online learning like SIPEJAR. Another problem with column comments is the language, which does not use an official language and sometimes uses symbols or characters in the comment that

makes it very difficult to understand the meaning of the comment. The teacher needs a quick response to get feedback so that the quick response can take into account the emotions of the viewer or student based on the student's comment. The purpose of feedback for the teacher is that the teacher can make good means [2], [4]. It is clear that the purpose of creating good media requires quick feedback from viewers, and media-creating media can assess online media learning.

On the other hand, it is necessary to make good means of online learning. One of the ways to create a good learning medium is to get feedback from both viewers and students. For example, many learning media use videos like YouTube channels, all videos receive comments from viewers. These comments are used as input for video owners to create other videos with good content [5–7]. Other comments are visible in the online learning materials, as there is a comment function. The form of the comments is a text document, and many users add text or icons to the comments, making it difficult to say what they want to say or what their emotions are when they watch videos that are online learning materials. For special purposes, teachers or media creators should know the reactions of the viewers, such as the emotions of the audience, or rate the learning media [6–7]. Processing the transformation of text documents into emotions requires artificial intelligence algorithms to determine emotions based on the comment column in online learning media. Artificial intelligence in text documents, such as text mining, is necessary and is integrated into online or web-based learning, especially in mobile applications [7–8]. The integration of AI on the web or mobile devices is necessary to help teachers or video producers improve the quality of the videos that are converted based on comments in the form of viewers' emotions.

However, there is a lot of research on the use of Artificial Intelligence (AI) in a text document to detect emotions. This is because AI can classify the text document by weighting each text in the document for all categories of emotions [9–12]. For example, text mining has many algorithms to classify the document, but you need another algorithm to determine the precision of the emotion, such as naive Bayes, clustering, and nearest neighbor (kNN). Naive Bayes requires more resources to determine emotion, so it is not suitable for mobile use. The KNN algorithm does not need a lot of resources to determine the recognition of emotion [12–13]. The problem with text mining is training and testing data. If there are no resources, it is very difficult to create tokenization, filtering, and bypassing. Another problem is that the weighting in text mining has the same value in the different texts in the Indonesian language, so an algorithm is required that provides the determination with the close value of the weight and can calculate it quickly in mobile applications [12–15]. It is clear that to calculate the text document, this research requires a text mining algorithm that is already a data training and can provide a clear weighting and grouping with a close value.

This study creates an emotion recognition mobile application based on the comment column on online learning materials using artificial intelligence. The comment field exists and is created on all video media like YouTube etc. The comment type is a text document, so this research uses a text-mining algorithm and the k-NN algorithm. Text mining is used for document classification to identify emotions, and k-NN is used to determine the accuracy of emotions based on comments in online learning media. The test verifies the accuracy of the algorithm and the validation of the emotional results. Obviously, this app will help teachers to create online learning materials that may or

may not be satisfactory in emotion recognition through the online learning media feed-back column.

2 Method

The problem with recognizing the emotion in this mode is the text mode when the user enters the emotion in the comments of the online column. This research uses 5 steps to solve the problem of column comment-based emotion recognition in online learning.

Step is

1. Create data record
2. Create data training
3. Tokenization, filtering, and derivation
4. Weighting algorithm
5. KNN algorithm and
6. Validation

Three processing at the beginning is the text mining algorithm to classify using a text document.

2.1 Text mining

Text mining is the extraction of information from text by the user using analysis tools. In general, text mining uses form data mining for both processing and technology. Word pre-processing is an early stage in text mining. Pre-processing is done to eliminate parts of the text that are not necessary to obtain quality data for execution [15–17]. First, in the pre-processing phase, namely, tokenization, which aims to break up sentences into individual words known as terms or tokens. In addition, filtering is done by removing punctuation marks, changing uppercase letters to lowercase letters, and removing stop words that aim to eliminate words that are not useful or that do not affect the process [15–17]. The last is to derive to get basic words from words that have been given affixes or other information. The derivation used is derivation and since this algorithm is more precise than Porter's algorithm [16–17].

To recognize emotions, training data is required as a machine learning process. This study uses the ISEAR dataset as training data. One of the training data in the ISEAR dataset is shown in Table 1.

Table 1. Example of data

No.	Label	Indonesia Text
1	Happy	Saya merasa beruntung dan telah membantu
2	Afraid	Video ini tidak membantu
3	Angry	Tidak useless

This study used test data of up to 200 comments tagged based on the emotions written in the emotion hashtag at the end [18–19]. The test data for this study was collected from April to June 2020 and was only in Indonesian. One of the test data samples is shown in Table 2.

Table 2. The test of data samples

Name	Comment	Label
Xin	berbahaya#takut	Afraid
FF Sam	Porno sara najis ...#jijik	Disgusted
Heryi	Sudah tidak anggap #malu	Shy

To preprocess the text in this study, the tokenization phase was performed with data. When converting to token, omit symbols and numbers [18]. The results of the tokenization of the training data are shown in Table 3.

Table 3. The sample of Tokenizing in data training

No.	Label	Text	Tokenizing
1	Happy	Saya merasa dekat dengan teman-teman yang lain. Setelah menonton video ini	Saya merasa damai dengan diri saya sendiri dan juga berhubungan dekat dengan orang-orang Saya menghargai
2	Afraid	sakit parah sampai mati	Saya merasa sakit parah sampai mati
3	Angry	kemungkinan ketidakadilan	Saya merasa ketidakadilan untuk video ini

2.2 Weighting

Term frequency and inverse document frequency (TFIDF) are weights that are often used in information retrieval and text mining [18]. The equation to calculate TF is [18–19]

$$TF\,(d,t) = f\,(d,t) \tag{1}$$

Equation 1 describes that f(d,t) is the prevalence of the phrase t in record d. The equation for calculating the IDF price may be visible in equation 2 [19]:

$$IDF(t) = \left(\log \frac{N}{df\,(t)} \right) \tag{2}$$

The equation for calculating TF-IDF is proven in equation 3 [19]:

$$TFIDF = TF\,(d,t).IDF\,(t) \tag{3}$$

TF IDF weighting may be finished if there may be check facts this is used as a question within side the TF-IDF approach. The manner of weighting the usage of the TF IDFDF approach is shown in Table 4.

Table 4. The sample of data weighting

No.	Text								Label
221	Video ini bermanfaat dan menyenangkan pikiran								Happy
	ups	sampe	yahh	Ngelike	Tmn	Yg	status	senang	
	0	0	0	0	0	0	0	9.56	

2.3 K-NN

The K-Nearest Neighbor (KNN) set of rules is a technique for classifying gadgets primarily based totally on education information this is closest to the object [18–19]. Near or a ways friends are normally calculated primarily based totally at the Euclidean distance via way of means of the overall formulation this is proven in equation 4[18–19].

$$d = \sqrt{\sum_{i=1}^{n}(a_i - b_1)^2} \tag{4}$$

In equation 4, the factor is d = distance, a = check statistics/testing, b = pattern statistics, i = variable statistics and n = Dimension of statistics

2.4 Validation

K-Fold Cross Validation is one method to determine the average success of a system [17]. Calculating the accuracy value can be done with equation 5 [17]

$$accuracy = \frac{sum\,of\,the\,right}{sum\,of\,data} \times 100\% \tag{5}$$

3 Result and discussion

This phase discusses the result and checking out of the utility cellular. The utility makes use of a web-primarily based totally and cellular base to make it user-friendly.

3.1 Implementation

This study makes use of YouTube and online studying as a useful resource for recording textual content emotion with in the column of remark. The useful resource from the YouTube channel containing studying cloth is proven in Figure 1. The records creating a detection emotion are from the column of remarks. The remark is stored through a txt record after which is calculated through Text mining and k-NN. The processing of load record txt primarily based totally at the column of remarks is proven within side the utility in Figure 2.
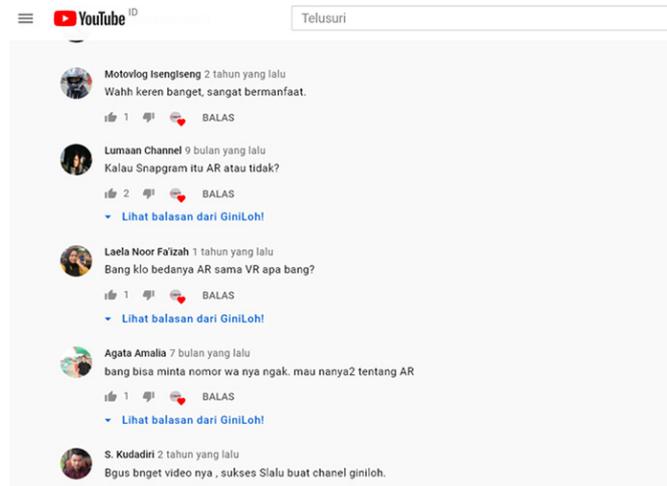
Fig. 1. Example of comment in YouTube
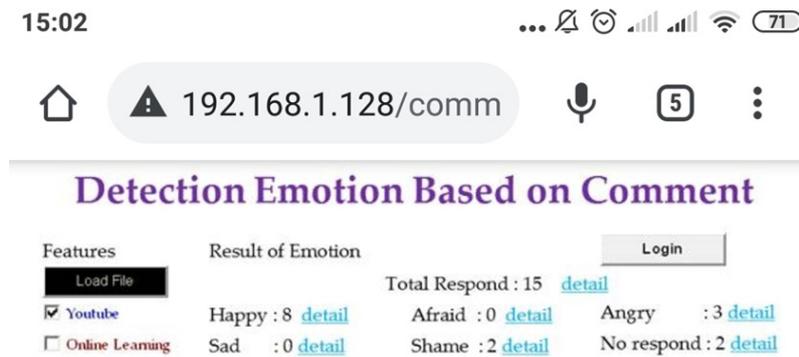
Fig. 2. Application for the recognition of emotions in YouTube media and online learning

The processing with in the utility proven in Figure 2 is

1. Load a record textual content-primarily based totally on a touch upon YouTube Channel.
2. Users should pick the aid is from YouTube or online leaning
3. After the consumer's choosing, the utility confirmed the result along with the whole of outcomes and emotion category.

Users can recognize the element of emotion and who's the consumer with the aid of using clicking the element within side the utility.

Most of the styles are intuitive. However, we invite you to read carefully the brief description below.

3.2 Testing and discussion

Testing the algorithm. The consequences of the category of feelings that have been completed are looked at by examining the checking out the use of the confusion matrix approach to decide Recall, Precision, and Accuracy with inside the category completed. The category completed on this take a look at the makes use of 35 check records and 4566 education records from the ISEAR dataset. Experimental records turned into received from YouTube social media from April 2020 to June 2020. The check records had been categorized primarily based totally on hashtags primarily based totally on emotional categories. There are nine check records within side the glad category, 22 check records within side the worry category, 35 check records within side the irritating category, 187 check records in the unhappy category, 17 check records in the disgusting category, 24 check records within side the shy category.

While the education dataset incorporates 1091 education records with inside the glad category, 1093 education records within side the worry category, 1094 education records with inside the irritating category, 1094 education records with inside the unhappy category, 1094 education records within side the disgusting category, and 1094 education records. Within side the shy category. These records are categorized to stumble on feelings with the proposed approach to make comparisons. From the submission assessment stage, the consequences of the emotional category check had been received. The consequences of checking out the use of the k-nearest neighbour approach are proven in Table 5.

Table 5. The result of emotion

Emotion	Precision	Recall	Accuracy
Disgusted	0.27327	0.52394	0.89312
Shy	0.23333	0.29317	0.86339
Angry	0.25349	0.37314	0.79359
Sad	0.90328	0.34376	0.56312
Happy	0.07341	0.44344	0.81329
Afraid	0.31348	0.77327	0.85371
Average	0.44321	0.55395	0.63397

The result of trying out is that the common accuracy is 0,697, the recollect is 0,5595 and the common precision is 0,4421.

Testing the validation of the application. In this research, checking out in utility cellular for a person in actual case used 10 college students to present feedback in online learning. The remark is processed to decide the emotion of the scholar. After the utility is given a cautioned emotion, the scholar offers remarks approximately the notion this is legitimate or not. The result of validation for 10 college students is shown in Table 6.

Table 6. The result of the validation

No.	Student	Suggest Emotions	Feedback
1	1705313628629	happy	Not Valid
2	1705313628506	sad	Valid
3	1705313628565	happy	Not Valid
4	1705313628559	happy	Valid
5	1705313628516	sad	Valid
6	1705313628552	happy	Valid
7	1705313628562	happy	No respond
8	1705313628566	sad	Valid
9	1705313628601	happy	Valid
10	1705313628600	sad	Valid

Analysis data-primarily based totally in Table 6,

1. There are 10 college students, simply 9 college students provide feedback.
2. 7 college students provide legitimate for ensuing their emotion from their remark
3. Just 2 college students provide no legitimate fee in utility.
4. Just one scholar did now no longer provide a reaction in his sentences.

Needless to say, the accuracy for this utility is 70% for detection emotion-primarily based column of remark in media online gaining knowledge of SIPEJAR.

4 Conclusion

This look builds a cellular software to discover YouTube-primarily based totally on students' feelings and online gaining knowledge of within side the remarks column while customers view online gaining knowledge of substances. Comments in web-primarily based are transformed to textual content documents for calculations in Artificial Intelligence algorithms. This look makes use of a textual content-mining set of rules for textual content-primarily based emotion class type and k-NN to decide every students' feelings primarily based totally on consumer remarks on YouTube and online gaining knowledge of substances. This software is beneficial for instructors to discover college students' feelings after college students see motion pictures of studying substances they make on YouTube and online learning.

5 Acknowledgment

6 References

[1] Dubovi, I., & Tabak, I. (2020). An empirical analysis of knowledge co-construction in YouTube comments. *Computers & Education*, *156*, 103939. https://doi.org/10.1016/j.compedu.2020.103939

[2] Nawaila, M. B., Kanbul, S., & Uzunboylu, H. (2018). Toward making social media effective in our classrooms: A 9-point method. *International Journal of Interactive Mobile Technologies*, *12*(4). https://doi.org/10.3991/ijim.v12i4.9201

[3] Wahyono, I., Saryono, D., Asfani, K., Ashar, M., & Sunarti, S. (2020). Smart online courses using computational intelligence. https://doi.org/10.1109/ICVEE50212.2020.9243193

[4] Purvis, A. J., Rodger, H. M., & Beckingham, S. (2020). Experiences and perspectives of social media in learning and teaching in higher education. *International Journal of Educational Research Open*, *1*, 100018. https://doi.org/10.1016/j.ijedro.2020.100018

[5] Anksorus, H., & Bradley, C. L. (2020). Using social media and focused learning activities to impact self-efficacy of empathy. *Currents in Pharmacy Teaching and Learning*, *12*(6), 741–750. https://doi.org/10.1016/j.cptl.2020.01.019

[6] Sholihin, M., Sari, R. C., Yuniarti, N., & Ilyana, S. (2020). A new way of teaching business ethics: The evaluation of virtual reality-based learning media. *The International Journal of Management Education*, *18*(3), 100428. https://doi.org/10.1016/j.ijme.2020.100428

[7] Wahyono, I. D., Fadlika, I., Asfani, K., Putranto, H., & Hammad, J. (2019, October). New adaptive intelligence method for personalized adaptive laboratories. In 2019 International Conference on Electrical, Electronics and Information Engineering (ICEEIE) (Vol. 6, pp. 196–200). IEEE. https://doi.org/10.1109/ICEEIE47180.2019.8981477

[8] Divayana, D. G. H., Sappaile, B. I., Pujawan, I., Dibia, I. K., Artaningsih, L., Sundayana, I., & Sugiharni, G. A. D. (2017). An evaluation of instructional process of expert system course program by using mobile technology-based CSE-UCLA model. *International Journal of Interactive Mobile Technologies*, *11*(6). https://doi.org/10.3991/ijim.v11i6.6697

[9] Gong, Y. (2021). Application of virtual reality teaching method and artificial intelligence technology in digital media art creation. *Ecological Informatics*, *63*, 101304. https://doi.org/10.1016/j.ecoinf.2021.101304

[10] Wahyono, I. D., Putranto, H., Asfani, K., & Afandi, A. N. (2019, September). VLC-UM: A novel virtual laboratory using machine learning and artificial intelligence. In 2019 International Seminar on Application for Technology of Information and Communication (iSemantic) (pp. 360–365). IEEE. https://doi.org/10.1109/ISEMANTIC.2019.8884288

[11] Manikanthan, S. V., Padmapriya, T., & Hussain, A. (2020). Artificial intelligence techniques for enhancing smartphone application development on mobile computing. *International Journal of Interactive Mobile Technologies*, *14*(17). https://doi.org/10.3991/ijim.v14i17.16569

[12] Karabatzaki, Z., Stathopoulou, A., Kokkalia, G., Dimitriou, E., Loukeri, P. I., Economou, A., & Drigas, A. (2018). Mobile application tools for students in secondary education. An evaluation study. *International Journal of Interactive Mobile Technologies*, *12*(2). https://doi.org/10.3991/ijim.v12i2.8158

[13] Wahyono, I. D., Saryono, D., Ashar, M., & Asfani, K. (2019, September). Face emotional detection using computational intelligence based ubiquitous computing. In 2019 International Seminar on Application for Technology of Information and Communication (iSemantic) (pp. 389–393). IEEE. https://doi.org/10.1109/ISEMANTIC.2019.8884320

[14] Jung, H., & Lee, B. G. (2020). Research trends in text mining: Semantic network and main path analysis of selected journals. *Expert Systems with Applications*, *162*, 113851. https://doi.org/10.1016/j.eswa.2020.113851

[15] Ribeiro, J., Duarte, J., Portela, F., & Santos, M. F. (2019). Automatically detect diagnostic patterns based on clinical notes through Text Mining. *Procedia Computer Science*, *160*, 684–689. https://doi.org/10.1016/j.procs.2019.11.027

[16] Estrada, M. L. B., Cabada, R. Z., Bustillos, R. O., & Graff, M. (2020). Opinion mining and emotion recognition applied to learning environments. *Expert Systems with Applications*, *150*, 113265. https://doi.org/10.1016/j.eswa.2020.113265

[17] Wahyono, I. D., Ashar, M., Fadlika, I., Asfani, K., & Saryono, D. (2019, October). A new computational intelligence for face emotional detection in ubiquitous. In 2019 International Conference on Electrical, Electronics and Information Engineering (ICEEIE) (Vol. 6, pp. 148–153). IEEE. https://doi.org/10.1109/ICEEIE47180.2019.8981420

[18] Ma, C. J., & Ding, Z. S. (2020, December). Improvement of k-nearest neighbor algorithm based on double filtering. In *2020 5th International Conference on Mechanical, Control, and Computer Engineering (ICMCCE)* (pp. 1567–1570). IEEE. https://doi.org/10.1109/ICMCCE51767.2020.00343

[19] Pal, K., & Patel, B. V. (2020, March). Data classification with k-fold cross-validation and holdout accuracy estimation methods with 5 different machine learning techniques. In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 83–87). IEEE. https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00016

7 Authors

Irawan Dwi Wahyono is a lecture on Department of Engineering in Universitas Negeri Malang, Indonesia (email: irawan.dwi.ft@um.ac.id).

Djoko Saryono is a Professor on Department of Literature in Universitas Negeri Malang, Indonesia (email: djoko.saryono.fs@um.ac.id).

Hari Putranto is a lecture on Department of Engineering in Universitas Negeri Malang, Indonesia (email: Hari.putranto.ft@um.ac.id).

Khoirudin Asfani is a lecture on Department of Engineering in Universitas Negeri Malang, Indonesia (email: khoirudin.asfani.ft@um.ac.id).

Harits Ar Rosyid is a lecture on Department of Engineering in Universitas Negeri Malang, Indonesia (email: harits.ar.ft@um.ac.id).

Sunarti is a lecture on Department of Literature in Universitas Negeri Malang, Indonesia (email: sunarti.fs@um.ac.id).

Mohd Murtadha Mohamad is a lecture on School of Computing in Universiti Teknologi Malaysia, Malaysia (email: murtadha@utm.my).

Mohd Nihra Haruzuan Bin Mohamad Said is a lecture on Department of Educational Sciences, Mathematics and Creative Multimedia Universiti Teknologi Malaysia, Malaysia (email: nihra@utm.my).

Gwo Jiun Horng is a lecture on Department of Computer Science and Information Engineering in Southern Taiwan University of Science and Technology, Taiwan (email: grojium@stust.edu.tw).

Jia-Shing Shih is a lecture on Department of Electrical Engineering in Southern Taiwan University of Science and Technology, Taiwan (email: jasonshih@stust.edu.tw).