Measurement of Learning Evaluation Against Assisted by Laboratory Assistants

Sulis Sandiwarno, Ariyani Wardhana

Faculty of Computer Science
Universitas Mercu Buana, Indonesia

Abstract - *The development of good information now triggers the use of very broad and very easy to use technology. In its development this technology is known as Information Technology (IT), where IT is a good tool in connecting the giver and recipient of information. An example of the application of IT that can measure the level of user satisfaction with the system is by using sentiment analysis. Measuring the level of satisfaction by using this sentiment analysis can also be applied in the field of learning. Previous research has been conducted to measure the level of student satisfaction with the learning process assisted by laboratory assistants based on questionnaires. However, the assessment of student satisfaction in previous studies is said to fail if only limited to questionnaires. We propose using sentiment analysis to evaluate the learning process for students assisted by laboratory assistants. In conducting research using the concept of sentiment analysis we use logistic regression (LR) and naïve Bayes (NB) methods. As for several stages such as: first, collecting data about opinions or reviews from students whose learning process is assisted by laboratory assistants. Second, we will conduct training data with both methods. Third, we will make conclusions, what methods are best used in measuring the evaluation of learning carried out by laboratory assistants. The results of this study will provide results that NB is a good algorithm in evaluating student opinion levels with an accuracy value of 80.32%.*

Keywords: *learning, evaluation, machine learning classifiers*

I. Introduction

Information plays an important role in an organization for the survival of an organization that specifically underlies decision making at the tactical level and strategic decisions. At present organizations are faced with a large amount of information that plays an important role in decision making. Information technology has a great opportunity value in the world of education, the dissemination of information is very rapid and very useful for its users. Decision-making systems are used in an institution to get the best steps in choosing information that will be used in development. The development of Information Technology is supported by the lives of everyone who wants to use it in helping solve existing problems. Some layers that use information technology as in the education sector, with better utilization, the results will be very good.

In developing a very good learning process, this is inseparable from the activities of lecturers and laboratory assistants who always uphold the concept of learning. A university usually requires an evaluation of the learning process assisted by the lecturer, but in this studies we tried to focus on laboratory assistants (aslab). In order to see how far aslab understands and is able to convey the material given to students. As well as the provision of learning materials delivered with the help of laboratory assistants (ASLAB) will result in better or less good grades. According to Yeh [1] in building the learning process carried out by Aslab, it must be based on collaboration with participants, so that in the end the learning process will succeed. Aslab must be able to provide an overview of the knowledge that will be given to students, especially if the learning process is carried out in the laboratory. There are two things that can be seen, the ease and difficulty in the process, if students feel they do not understand the material presented, aslab will provide good direction or generally aslab will guide students slowly (step by step) to be able to understand the material provided. Added by the opinion of Šumak [2] to build a good learning process, aslab to create a study group. The purpose of forming this learning group is to make students easier to communicate with their members in groups and fellow members in the group can also provide motivation to other members, so that other members can feel the meaning of a learning process [3]. Combining from the three previous studies, Islam [4] modeled the learning process using the Technology Acceptance Model (TAM) method. Based on the TAM model, researchers include students who understand about the material taught in the group whose content participants from the group do not understand the material. The researcher then measures the evaluation of learning that has been done by giving questionnaires to the students. The evaluation results obtained can provide information to the instructor to what extent the evaluation of the learning process has been carried out.

Previous research has evaluated the learning process, but the evaluation process carried out in previous studies only focused on the use of questionnaires. Questionnaires are classic or old models in evaluating the learning process. Another failure in the questionnaire is that it cannot capture information about the responses of students during the learning process. So according to us, research based on questionnaires cannot provide good conclusions about the results of evaluation of the learning process. The results of evaluations assisted by laboratory assistants (aslab) cannot be used as an indicator that the learning process is better. We capture information that in previous studies did not provide an opportunity for students to convey responses in the form of opinions to instructors on the learning process. In this study we propose to utilize sentiment analysis to evaluate the opinions of students on the learning process assisted by ASLAB. In utilizing sentiment analysis, we need methods that can help measure good (positive) responses, negative responses (negative) methods are logistic regression (LR) and naïve Bayes (NB).

The writing structure of the research proposal is as follows: in the second section, we provide information about the research that is reflected in the research strategic plan. In the third section we provide information about the literature review (reference) used in helping research writing. In the fourth section, we provide information about what methods will be used in the study. In the fifth section, we provide information about the schedule of research to be conducted.

II. Related Word

A system can consist of several subsystems or part systems. Components or subsystems in a system cannot stand alone independently[1], [2], [15]. Components or subsystems interact and interact with each other to form a unity so that the goals or objectives can be achieved [3], [4]. Information is the meaning of relationships and interpretation of data that allows someone to make a decision. Information is said to be valuable if the information affects the decision-making process better. Information system is a concept to adopt people, technology, and information to develop the decision support system in an internal organization[5], [6], [16]. In the process of information systems designing the problem of optimal information resources distribution in computer systems is one of the most paramount tasks. Such problems arise upon designing information systems on the base of computer systems [7], [8]. Analysis is the decomposition of a complete information system into its component parts with the intention of identifying and evaluating problems, opportunities, obstacles, and expected needs so that improvements can be concluded. SDLC models consist of analysis, design, code, and testing [9], [10]. Discussion forums are a good tool in shaping communication between students and lecturers. Online forum can be used to complement learning and teaching, particularly in blended or hybrid learning courses[11], [12]. The asynchronous discussion forum may contribute to understanding the learning content, knowledge construction and student achievement. In asynchronous discussion forums, students are less involved or not willing to ask their peers [13], [14].

III. Methodology

A. Evaluation Model

In this section we will explain how machine learning algorithms classify data. To classify data, machine learning algorithm has several stages, opinions pre-processing, feature extraction, and opinion classifiers. Meanwhile, the stages in classifying the data will is shown in Figure 4.1.

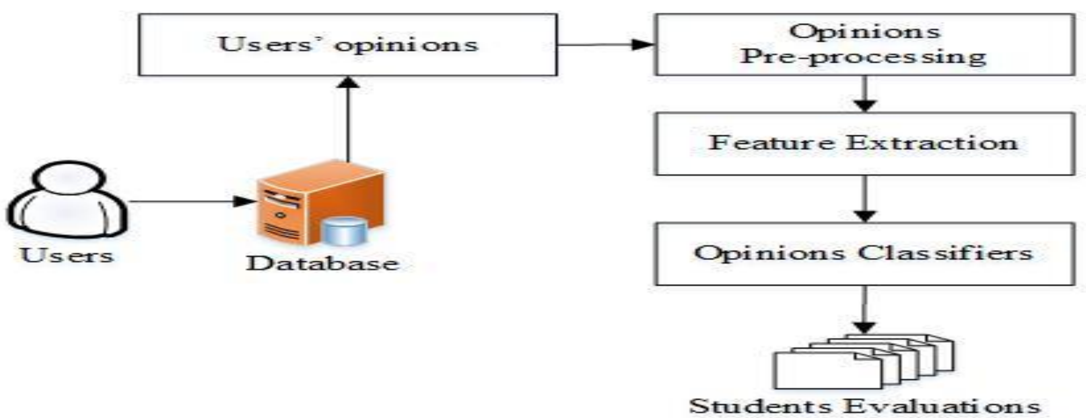

Figure 1. Classification Model

a) Opinion Pre-processing

This section removes some identities from a text, where the identity is HTML decoding, remove stop words, and remove bad characters in opinion.

b) Feature extraction

To select features, we will extract data from the features of student opinion results that will be more effective. Then we will analyse and evaluate the results of student opinions to identify "feature words". To extract this data, we use the term frequency inverse document frequency (TF-IDF). TF-IDF is a method used to calculate the weight of each word that is most commonly used in data classification. This method is known to be effective, efficient, and easy to use. This method will calculate the value of TF and IDF on each data (text) in each body. Where is the metric of the TF-IDF especially as $Wdt = tfdt * I$, where, d describes the document to - d, t describes the word in the text - t from the keyword, W describes it as weighting the document to - t to the word in the text to - t, tf describes as the number of words in the text searched in the document (W), while the IDF describes the results of the process from $IDF = log2\frac{d}{df}$

c) Opinions classifiers

In this study we will select opinion data for training (training) as much as 70% and as much as 30% taken as testing data (testing). In classifying data, we will do testing for 10 times by utilizing the fold cross validation. Fold cross validation is an algorithm used in iterating data classifications. Each iteration will produce different results, and the results obtained from the two algorithms will also be different. Following this we will explain the algorithm that we use in classifying data.

1) Logistic Regression (LR)

LR is an easy technique commonly used in binary and multi class classification. For binary classification, Malanga [54] illustrated the techniques as:

$$P(Y(T) = i|X) = \frac{1}{(1 + e^{-\theta^T X})} \qquad \text{(Eq. 1)}$$

where, $P(Y(T) = i|X)$ is the posterior probability, Y is presenting the class, X describes the values of the feature of Y, T represents the test samples, and θ is the vector of parameters to be predicted. Moreover, $x = [x_1, x_2, \dots, x_n]^T$ is defined as the vector of feature values for identifying the documents. We are encoding the fact that a document included in a class $K \in \{1, .., K\}$ by a K-dimensional vector $y = [y_1, y_2, \dots, y_n]^T$ valued 0/1. Where, $y_k = 1$ and the rest of the coordinates are 0. Ndenga et al. [24] describes the multinomial logistic regression as a model of the conditional probability of the form parameterized matrix $B = [\beta_1, \dots, \beta_k]$. To generalize binary class to the multi class of logistic regression, in every column of B is a parameter vector corresponding of the class $\beta_k = (\beta_{k1}, \dots \beta_{kd})^T$. Therefore, the multi-label classifier can be formalized as follows:

$$P(Y_k = 1|x, B) = \frac{\exp(\beta_k^T x)}{\sum_{k'} \exp(\beta_k^T x)} \qquad \text{(Eq. 2)}$$

2) Naïve Bayes (NB)

Given the test description of the document d of an opinion represented by the vector $< w_1, w_2, \dots, w_m >$, to classify the document d, MNB is defined as:

$$C_{MNB^{(d)}} = P(c) \prod_{i=1}^{n} P(w_i|c)^{fi} \qquad \text{(Eq. 3)}$$

where, $P(c)$ is a prior probability that a document d belongs to class c, n is a number of the features, $P(w_i|c)$ is the conditional probability that a word w_i occurring in the class c, w_i is the word feature occurred in d, fi is the number of frequency count of a word w_i in reporting d, and $C_{MNB^{(d)}}$ is the class label of d predicted by the classifier [24].

IV. Evaluation

This section will explain the results of the research that has been done in evaluating student opinions regarding the learning process which is assisted by a lab assistant. The results of data classification carried out by naïve Bayes and logistic regression are then calculated using several techniques, such as: precision, recall, F1 and accuracy. Madani, explained about precision, recall, F1 and accuracy as follows. precision describes the level of accuracy between the information requested by the user and the answers given by the system.

Table 1 data classification results from logistic regression (LR)

Fold (#)	Accuracy (%)	Recall (%)	Precision (%)	F1 (%)
1	80.05	66.41	71.90	69.05
2	77.49	66.00	72.79	69.23
3	78.97	71.43	72.41	71.92
4	82.82	72.03	79.23	75.46
5	78.72	66.67	77.04	71.48
6	78.21	67.36	71.85	69.53
7	80.77	79.26	69.48	74.05
8	78.72	66.91	71.54	69.14
9	77.95	79.03	62.03	69.50
10	81.28	75.52	73.97	74.74
Average	79.50	71.06	72.22	71.41

From the results obtained from the two methods, it can be seen that NB is the best method of classifying the data in this study. The accuracy value obtained from NB is 80.32% while LR is 79.5% with a difference of 8.2%. From the fold (#) value that has been done that the smallest result of NB is in the second iteration with a value of 77.49%, and the highest value is 82.82 in the 4th iteration. Meanwhile, the smallest value in LR is in the 2nd iteration with a value of 78.77% and the highest value is 82.05% in the 4th iteration. The Recall value at NB is 76.66% and the value at LR is 71.06%. with a difference in value of 5.6 Meanwhile, the value of fold (#) in NB displays the lowest result in the 7th iteration with a value of 75.49%, while in LR is 66% in the second iteration.

Fold (#)	Accuracy (%)	Recall (%)	Precision (%)	F1 (%)
1	81.33	76.00	81.82	73.06
2	78.77	76.53	83.09	73.14
3	81.79	77.65	84.14	77.46
4	82.05	76.90	83.85	75.69
5	79.74	76.57	86.67	74.76
6	77.69	76.48	77.78	70.71
7	82.31	75.49	83.12	78.77
8	80.00	76.71	78.46	72.34
9	79.23	77.28	77.85	75.23
10	80.26	76.99	82.88	75.86
Average	80.32	76.66	81.96	74.70

If seen in the 2nd iteration NB displays the results of 76.53%, then there is a difference of 10.53%, while in the 7th iteration LR displays the results of 79.26% with a value of 3.77% difference from NB. In the precision section, NB displays the results of 81.96% and LR 72.22% with a difference in value of 9.74%. Whereas in section F1, NB displays the data classification results of 74.7% and LR displays the results of 71.41% with a difference in value of 3.29%. based on the results of data classification that has been done by both methods, the biggest difference in value occurs in precision. In Figure 5.2, there are steps in fold (#) in classifying data.

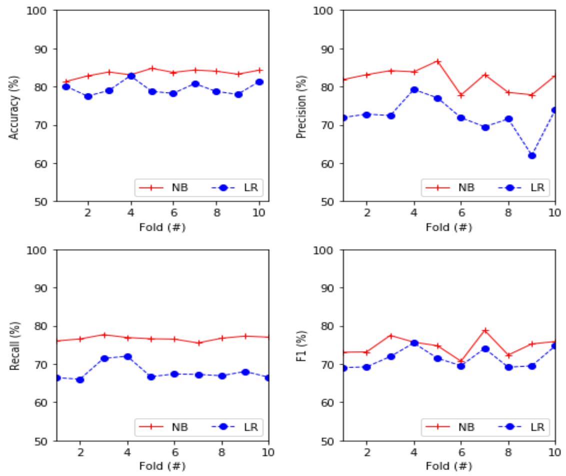

Figure 2 Results of fold (#) of two different methods

After classifying student opinion data, then we will calculate how much the percentage of positive and negative values of student opinion in the learning process is assisted by a lab assistant. The results obtained were 2512 students gave positive opinions (64%) and 1390 students gave negative opinions (36%), this can be seen in Figure 5.2. To prove this result we will present a portion of the opinions of students regarding the learning process which is assisted by a lab assistant.

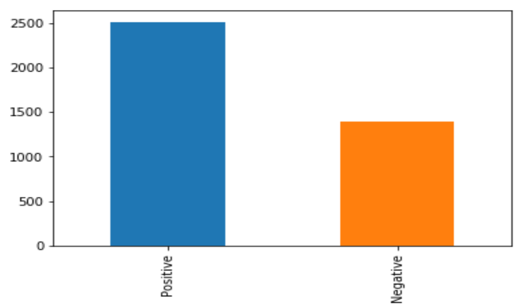

Figure 3. Results of classification of student opinion data

V. Conclusion

In this study, we utilized 2 methods of machine learning, namely naïve Bayes (NB) and logistic regression (LR) to classify student opinion data during the learning process carried out by lab assistants. The results obtained showed results, that more than 50% of students expressed positive opinions during the learning process, and less than 40% of students gave negative opinions. The results obtained conclude that during the lab assistant learning process contributes very well and to correct some negative opinions from students, in the future we will choose lab assistants who do have the knowledge to convey information to students. Then, we will continue to evaluate this learning process by adding a method that is currently famous, namely deep learning. We will do a comparative process between machine learning and deep learning; which algorithm is the best in classifying student opinion data.

REFERENCES

[1] Yeh, Y. chu: Integrating collaborative PBL with blended learning to explore preservice teachers' development of online learning communities. Teach. Teach. Educ. 26, 1630–1640 (2010). doi:10.1016/j.tate.2010.06.014

[2] Šumak, B., Heričko, M., Pušnik, M.: A meta-analysis of e-learning technology acceptance: The role of user types and e-learning technology types, (2011)

[3] Dascalu, M.I., Bodea, C.N., Lytras, M., De Pablos, P.O., Burlacu, A.: Improving e-learning communities through optimal composition of multidisciplinary learning groups. Comput. Human Behav. 30, 362–371 (2014). doi:10.1016/j.chb.2013.01.022

[4] Islam, A.K.M.N.: E-learning system use and its outcomes: Moderating role of perceived compatibility. Telemat. Informatics. 33, 48–55 (2016). doi:10.1016/j.tele.2015.06.010

[5] Ortigosa, A., Martín, J.M., Carro, R.M.: Sentiment analysis in Facebook and its application to e-learning. Comput. Human Behav. 31, 527–541 (2014). doi:10.1016/j.chb.2013.05.024

[6] Muñoz, M.Á.C., Hernández, M.L.G.: Teacher and School Counsellor Training to Promote a Collaborative Family-school Relationship: An Empirical Study. Procedia - Soc. Behav. Sci. 23, 667–671 (2017). doi:10.1016/j.sbspro.2017.02.039

[7] López-Alcarria, A., Gutiérrez-Pérez, J., Poza-Vilches, F.: Sustainable Management of Pre-school Education Centers: A Case Study in the Province of Granada. Procedia - Soc. Behav. Sci. 237, 541–547 (2017). doi:10.1016/j.sbspro.2017.02.104

[8] Colace, F., Casaburi, L., De Santo, M., Greco, L.: Sentiment detection in social networks and in collaborative learning environments. Comput. Human Behav. 51, 1061–1067 (2015). doi:10.1016/j.chb.2014.11.090

[9] Zheng, L., Huang, R.: The effects of sentiments and co-regulation on group performance in computer supported collaborative learning. Internet High. Educ. 28, 59–67 (2016). doi:10.1016/j.iheduc.2015.10.001

[10] Binali, H.H., Wu, C., Potdar, V.: A new significant area: Emotion detection in E-learning using opinion mining techniques. In: 2009 3rd IEEE International Conference on Digital Ecosystems and Technologies, DEST '09. pp. 259–264 (2009)

[11] Song, D., Lin, H., Yang, Z.: Opinion mining in e-learning system. In: Proceedings - 2007 IFIP International Conference on Network and Parallel Computing Workshops, NPC 2007. pp. 788–792 (2007)

[12] Yoo, S.Y., Song, J.I., Jeong, O.R.: Social media contents based sentiment analysis and prediction system. Expert Syst. Appl. 105, 102–111 (2018). doi:10.1016/j.eswa.2018.03.055

[13] Cross-domain aspect extraction for sentiment analysis: A transductive learning approach. Decis. Support Syst. 114, 70–80 (2018). doi:https://doi.org/10.1016/j.dss.2018.08.009

[14] Ravi, K., Ravi, V.: A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowledge-Based Syst. 89, 14–46 (2015). doi:10.1016/j.knosys.2015.06.015

[15] Mining: Students Comments about Teacher Performance Assessment using Machine Learning Algorithms. Int. J. Comb. Optim. Probl. Informatics. 9, 26–40 (2018)

[16] Wen, M., Yang, D., Rosé, C.P.: Sentiment analysis in MOOC discussion forums: What does it tell us? In: Proceedings of Educational Data Mining. pp. 1–8 (2014)

[17] Schiller, S.Z.: CHAT for chat: Mediated learning in online chat virtual reference service. Comput. Human Behav. (2016). doi:10.1016/j.chb.2016.06.053

[18] Serrano-Guerrero, J., Olivas, J.A., Romero, F.P., Herrera-Viedma, E.: Sentiment analysis: A review and comparative analysis of web services. Inf. Sci. (Ny). 311, 18–38 (2015). doi:10.1016/j.ins.2015.03.040

[19] Bing Liu: Sentiment Analysis and Subjectivity. In: Handbook of Natural Language Processing. pp. 1–38 (2010)

[20] Pham, D.H., Le, A.C.: Learning multiple layers of knowledge representation for aspect based sentiment analysis. Data Knowl. Eng. 57, 26–39 (2018). doi:10.1016/j.datak.2017.06.001

[21] Aldoğan, D., Yaslan, Y.: A comparison study on active learning integrated ensemble approaches in sentiment analysis. Comput. Electr. Eng. 57, 311–323 (2017). doi:10.1016/j.compeleceng.2016.11.015

[22] Yadollahi, A., Shahraki, A.G., Zaiane, O.R.: Current State of Text Sentiment Analysis from Opinion to Emotion Mining. ACM Comput. Surv. 50, 1–33 (2017). doi:10.1145/3057270

[23] Madani, Y., Erritali, M., Bengourram, J.: Sentiment analysis using semantic similarity and Hadoop MapReduce, (2018)

[24] Ndenga, M.K., Ganchev, I., Mehat, J., Wabwoba, F., Akdag, H.: Performance and cost-effectiveness of change burst metrics in predicting software faults, (2018)

[25] Andi Nugroho. (2016). Aplikasi Web Informasi Dan Registrasi Peserta Seminar, Workshop, Talkshow Pada Acara Seminar Nasional Pengamplikasian Telematika (Sinaptika) Tahun 2016. Seminar Nasional Sistem Informasi Indonesia, 1-8.

[26] Raka Yusuf[1], Yossi Susanto[2]. (2010). Pemanfaatan SMS Gateway untuk Absensi Sekolah Siswa. Seminar Nasional Pengaplikasian Telematika SINAPTIKA, 1-4.

[27] Fajar Masya[1], Elvina[2], Fitri Maria Simanjuntak[3]. (2012). Sistem Pelayanan Pengaduan Masyarakat pada Divisi HUMAS POLRI Berbasis Web. Seminar Nasional Aplikasi Teknologi Informasi, 1-6.

[28] Raka Yusuf[1], Gilang Widi Darmawan[2]. (2016). Aplikasi Berbasis Web Dengan Menggunakan Pustaka Javascript Fabricjs Untuk Pembuatan Komik Strip Punakawan. Seminar Nasional Teknologi Informasi dan Multimedia, STMIK AMIKOM Yogyakarta, 1-6.