

Development of a Scoring Application for Indonesian Language Essay Questions

Qurratul Aini
Master of Hospital Management,
Postgraduate Studies Universitas
Muhammadiyah Yogyakarta,
Indonesia
Jl. Brawijaya, Tamantirto Bantul
Yogyakarta Indonesia 55183
+62 274 387656 Ext. 320
q.ainiumy@gmail.com

Achmad Eko Julianto
Information Technology Department,
Universitas Muhammadiyah
Yogyakarta, Indonesia
Jl. Brawijaya, Tamantirto Bantul
Yogyakarta Indonesia 55183
+628 2221383722
achmadeko003@gmail.com

Dwijoko Purbohadi
Information Technology Department,
Universitas Muhammadiyah
Yogyakarta, Indonesia
Jl. Brawijaya, Tamantirto Bantul
Yogyakarta Indonesia 55183
+62 274 387656 Ext. 274
dwijoko.purbohadi@umy.ac.id

ABSTRACT

This article discusses the development of Automated Essay Scoring (AES) in the Learning Management System (LMS). In Indonesia, e-learning mostly implements using LMS as the central facility for teaching and learning activities. LMS is used to facilitate in providing many exam questions for a student. One of it is essay question with short answer or explanation. The use of suitable technology can help teachers in assessing student answers in large number. Technology also can increase the level of accuracy in assessing student answers. Generally, at the end of the semester, school or university will assess student answer of essay questions. If there are many participants, teachers working load will increase. Usually, LMS does not have AES. Thus, the teacher not interest using essay question because still uses the traditional method to assess the student's answer to an essay question. This method takes much time, and this does not look accurate, mainly when meet so many questions with a large student. In this case, LMS needs AES facilities. Using AES for LMS technology will help the teacher in assessing participant answers. In this experiment, AES works by comparing the student and teacher answers using Levenshtein's comparison. If LMS has AES facilities, the teacher can access many participants answer although there are many students. The research shows that AES scoring is not much different from the teacher assessment results.

CCS Concepts

Applied Computing → Education → Computer-assisted instruction

Keywords

AES; CAI, SCORM; Levenshtein; LMS;

1. INTRODUCTION

Writing is an essential activity in education. School is a basic human need. Schools aim to educate students to be better through academic activities. Measuring student abilities is an essential indicator for knowing the quality of schools and students. Measurement activities can use the test. One of the best-known tests is essay questions. The teacher gives questions and students fill out the answer sheet. After the test is complete, the teacher checks the answer sheets one by one and scores. Activities like this are traditional models; there are many limitations if a teacher uses this method [1]. First, the assessment is inconsistent because the teacher does not use standard answers. Second, teachers need more time if the number of questions increases and the number of students also increases [2]. Third, the teacher takes longer because he has to

collect students' answers one by one. Today, teachers begin to leave essay questions. They prefer to use multiple choice questions because more practical. One way for teachers to develop essay questions is that schools must apply technology. Schools can use the internet so students can collect answers online. To be faster, we can use an automatic scoring system. By using the right technology, we can make the essay answer the assessment process more practical, faster, and more accurate [3].

Automated Essay Scoring (AES) is a software that able to check and assess answers of essay questions automatically. This attracts many researchers to find the best methods and applications. Researchers began developing this AES in 1966 called the Essay Grade (PEG) Project. This program takes twenty years [4]. E-Rater is one of the most complicated methods that still used in the Graduate Management Admissions Test (GMAT), developed by the Educational Test Service (ETS) [5]. Intelligent Essay Assessor (IEA) is an AES that uses the Latent Semantic Analysis (LSA) method. This method works by comparing the contents contained in the essay. LSA compares words in a dataset with words to answers. Another system is IntelliMetric, which is an AES that has used Artificial Intelligence (AI). IntelliMetric is the same as E-Rater that uses NLP [6].

In Indonesia, universities generally start using the concept of Student-Centered Learning (SCL) and leave the concept of Teacher Center Learning (TCL), because SCL can improve the quality of learning. SCL proved to be more effective than TCL [7]. In the SCL system students are directed to independent learning while the teacher becomes supervisor. One way to develop SCL is to use e-learning [8]. With the SCL concept, students become more independent in learning or developing skills. Although students' understanding is slow at first, in the end, students can develop better [9]. Learning Management System or LMS is the most developed of e-learning today.

LMS has not been equipped with essay question module, even though the time to collect answers is shorter but the teacher still needs a lot of time to assess the answers [10]. Actually, LMS can be equipped with a rating system that uses Intelligent Essay Assessor (IEA) technology. This technology is a computational model as a representation of human knowledge by extracting words and text equations. This system can help the teacher's work in correcting essay answers [11]. One weakness of this method is that this ignores the wording and structure of the text so that this is still rarely used [12].

Moodle is one of LMS software that very popular in Indonesia, but Moodle is also not equipped with an automatic scoring system. If

the school applies Moodle, the teacher still needs time to check the answers to essay questions. What's more, if the essay question requires a long explanation. So in Indonesia, this requires an automatic scoring system that supports Moodle and is specifically designed to recognise the structure of Indonesian Language.

This study aims to develop AES that can apply in Shareable Content Object Reference Models (SCORM) or Computer Aided Instruction. The AES facility developed in this study aims to facilitate teachers in assessing essay questions as well as a tool to monitor the development of students' understanding. Aside from being a tool to monitor learning outcomes, AES can be useful to assist teachers in giving guidance [13]. From previous studies, known that AES e-learning facilities have relatively stable accuracy compared to manual methods [14].

In Indonesia, development of LMS application snowballs, e.g., the school has been implementing LMS because it is a significant role in the use of this e-learning. It is an application to assist teachers in learning activities [15]. This application makes the teacher easier to see which students receive information provided by the teacher [16]. The advantage of this LMS is to make students and instructors meet in a virtual class [17]. LMS also helps students and instructors in sharing, sending and returning assignments [18]. With this application, a teacher can assist with the existence of automatic assessment system, and a summary of student results [19]. LMS allows students to be able to see instructors form videos, images and other multimedia [20].

In this study, we designed AES that can automatically check and assess answers to essay questions with the structure of the Indonesian language. We develop AES as a CAI Module that can communicate with Moodle LMS to be more useful. The results of this study are to reduce the time teachers evaluate and assess students' answers and reduce the element of subjectivity. If the time needed to assess down and be more qualified, ultimately motivating Indonesian teachers to be willing to use and develop essay questions again.

2. DESIGN APPLICATION

Figure 1 shows the block diagram of the AES design. Two parts become the core of AES, namely the Indonesian dictionary and the Levenshtein algorithm. Levenshtein distance is a method to measure of the similarity between two strings. AES works by comparing student answers to teacher answers. There are three possibilities, students' answers are shorter, equal, or longer than the teacher's answer. How to compare using the Levenshtein algorithm to find the closeness of sentences and words. If the number of sentences is equal, then arrange the sentence with the principle of closeness. If this is different, then there is a process of pasting sentences or throwing sentences (selection process) into the students' answers so that the number of sentences is the same.

There are four steps in AES. First step is to change all words in all sentences into basic words. Both the sentence on the teacher's answer and the student's answer. The basic word is determined based on the proximity of the word in the answer (which is likely to have changed form due to the addition of a prefix or suffix) to the words in the dictionary.

The second step is to separate the sentence into an array of sentences for both answers. It is possible that students' answers have fewer, equal or number of arrangement than the number of teacher answers. Before comparing and printing the number of arrays must be the same. If the number of arrays of answers to an array of students is less than one or more, the system will enter one

or more sentences so that the number of arrays in both answers is the same. Insertion run by selecting the sentence from the answer that has the Levenshtein distance farthest from the sentence in the student's answer. If the student answers arrays is higher than teacher arrays answer, then the selection and disposal will be done so that the number of arrays is the same. The answer sentence for discarded.

The third step is to compare the sentences of each array. Before comparing, the system will sort the array of student answers based on the closeness of the sentence with the array in the teacher's answer. If found the Levenshtein distance value same then the first one is chosen as the equivalent array.

Figure 2 shows that the architecture features automatic scoring. Architecture explains how the assessment process automatically passes several steps from the question to the results that have been calculated by an comparing and scoring algorithm. This architecture is the basis for creating assessment features that will be integrated with LMS using SCORM.

The first step is to read the students 'answers, then compare them with the lecturers' answers. Before comparing, the system changes the lecturer and student's answers to a sentence form that is composed of root words. Changing to this core sentence uses the Indonesian dictionary (Kamus besar Bahasa Indonesia or KBBI). This step repeated for each sentence. Students answer questions through the CAI (SCORM Module) interface. KBBI stored on CAI as a collection of words stored in text format.

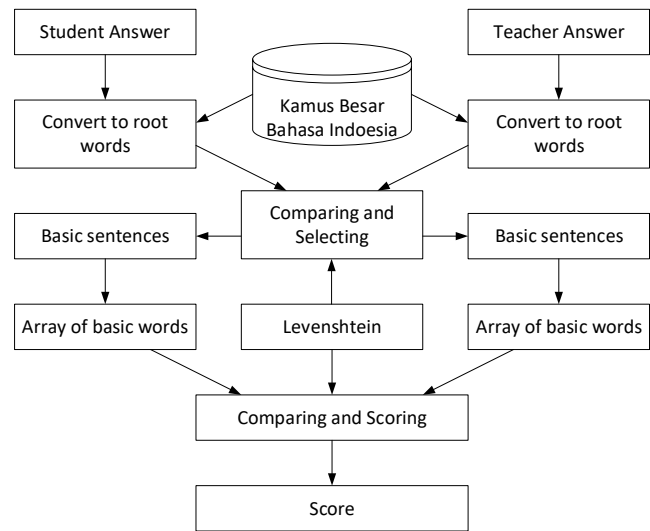


Figure 1. Design of AES

The fourth step is to score. Each array has a sentence, and each sentence is composed of necessary words. The total number of words from the entire array of teacher answers is the highest number. The system will compare each word in the array of student answers with each word in the teacher's answer array. The score obtained from the reduction in the total number of words with the total number of Levenshtein distances divided by the total number of words multiplied by 100. The score calculation is held by using the equation below:

$$score = \frac{(tw - rw)}{tw} \times 100$$

- tw: answer total word
- rw: comparing student and key answer

The Equation 1 explains the formula in calculating AES score. “rw” is the result of calculating all the number of word contained in a sentence and the result is the result of a comparison between student answers and answer keys. To find the score result using “rw” sub “tw” and divided to “tw”. Afterwards, the result of the share multiplied by 100. After completing the calculation, the system shows the results of the calculation as the student score. Figure 2 shows the AES process how to score an answer.

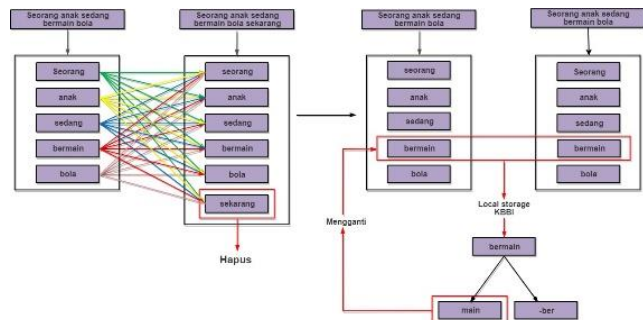


Figure 2. AES process

3. RESULT

3.1 CAI Module

CAI software is created using CourseLab 2.4 and Executable HTML 4.9. This combination produces an executable file (*.exe) called the CAI module. Figure 3 shows the main page interface. The CAI module can distribute through Moodle.

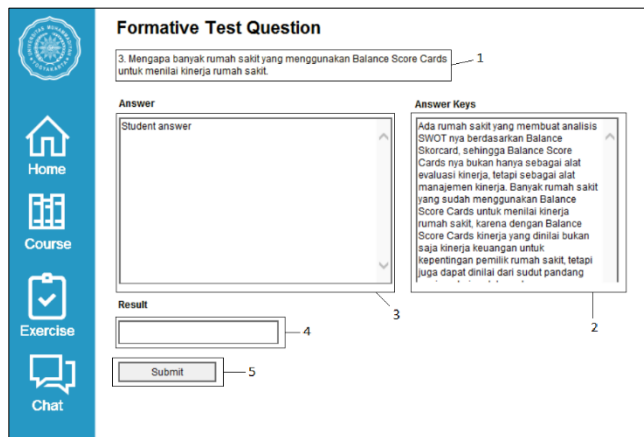


Figure 3 CAI module interface

Figure 3. Illustrates the part of the interface which consists of:

1. Questions
2. Answer key
3. Input area to answer
4. Display score
5. The button to process the answer at once to drive data to Moodle

In practice, students cannot see the answer key, but if the teacher wants to appear after the student answers this can be arranged using a button.

CourseLab 2.4 can also generate modules using the SCORM 1.2 standard. If using a module like this, students can answer questions online.

3.2 Experiment

The experiment was conducted using an essay questions for Hospital Management subject. There are 27 students as users. The purpose of the experiment is to gather information about the level of accuracy and effectiveness of this AES. Before the student answers the question using AES, the teacher evaluating and scoring the student answer manually. From the results of calculations using the t-test between the teacher's and the AES score, we found that there is no significant difference between the teacher's score and the AES score with significance levels 0.01. However, if we compare the student answer and teacher answer one by one, there are many possibilities that the score will be very different, this shows that the accuracy of the score still needs improvement.

4. DISCUSSION

4.1 Algorithm limitation

From the experimental results, we concluded that with AES presented in the form of CAI it has the potential to increase speed, reduce the teacher subjectivity, and increase the accuracy of assessment. In general, teachers teach many classes, so it is easy to understand that AES is handy for teachers because it can reduce the teacher's burden in examining the answers to essay questions. The experiment results consistent with the teacher's comments after trying, the teacher comments when compared to the old method, AES is more promising. After seeing the results of the trial, several teachers felt motivated. They will prepare essay question immediately. They will create questions with more varied answers.

This automatic scoring facility is still in the development stage; it still requires much testing for various courses with various questions and answers. One disadvantage in testing is that the teacher's assessment is always round with a difference of 5, for example, 65, 70, 75, 80, and so on. Meanwhile, the computer answer is always the difference 1. That is, there is a high probability that there is a difference between the students' answers and the teacher's answers, this is what affects accuracy.

If we analyse, the Levenshtein distance method only measures the distance of the difference between the two words is. That is, how to analyse student answers does not consider the meaning but the closeness of the character. This AES has not been able to see different words but has the same meaning. If the student's answer is "Ayah dapat keputusan" and the teacher's answer key is "Bapak mampu membuat keputusan". In Indonesian language, both answers have same meaning but have difference Levenshtein distance. This AES design only for Indonesian language essay question.

This CAI module equipped with AES needs re-evaluating because it still has some disadvantages. One of them is still relying on the Levenshtein algorithm. Thus comparable answers only use the proximity of the word used, not the meaning of the sentence. Accuracy still needs to be improved using an intelligent system.

Sometimes students answer using long sentences with the writing structure not following the rules of Indonesian writing. The meaning of the sentence easy understood by the teacher and is considered correct. However, if the sentence compared with the teacher's answer, high possibility that the Levenshtein distance is high, and the score is low. To reduce this inaccuracy, the teacher

can advise so that in answering questions use simple sentences and follow the structure of formal writing.

4.2 SCORM and CAI

All CourseLab program code write in Javascript and HTML. The publication of CourseLab can use SCORM standards and a collection of files that can be run using internet explorer. This collection of files can convert into an executable file that can work well on a Laptop or Computer (CAI module). The main difference between SCORM and CAI modules is flexibility. SCORM is placed in Moodle so students have to run it online, while if using CAI students can run it even if there is no internet connection. However, both can work as desired. From a functional trial, these two ways produce the same assessment results.

From the discussion, information obtained that CAI containing AES needed to be developed for the smartphone so that teachers more freely monitor student answers.

For building a CAI module so that can work on mobile devices, the authoring tool must use CourseLab 3 which can recognise HTML standards 5. CourseLab 2.4 only recognizes HTML 4 standards. However, this two software still rely on Javascript so that to change to the mobile version is very easy. In the next research, mobile-based AES modules will be developed using CourseLab 3 and Adobe PhoneGap framework.

5. CONCLUSION

The AES facility for CAI for supporting the Moodle LMS has many advantages. The current education system is in dire need of AES because it has the potential to reduce the workload of teachers. From discussions with teachers who have used AES, they are motivated because the exam using essay questions is no longer an obstacle. Thanks to AES, teachers have the opportunity to make essay questions more varied. Although accuracy is still not satisfactory, teachers can receive the results of AES because they assume the results are not much different, this is consistent with the t-test calculation which shows that the results of the AES assessment with the teacher are not significantly different. AES is very potential because the way to collect answers is much more practical because it can be done online. AES eliminates teacher subjectivity. Although not related to this research, from the side of students, the use of CAI has the advantage of high flexibility because students can answer questions anytime and anywhere without relying on the internet.

6. ACKNOWLEDGEMENTS

This research is part of main research with title Self Directed Learning (SDL) using Gamma Feedback Learning Models (GFLM). We would like thank to The Ministry of Research Technology and Higher Education of the Republic of Indonesia with Project Number: SP DIPA-042.06-0/2017. We also appreciate to the Research and Community Service Universitas Muhammadiyah Yogyakarta for partially support to this research. We would also like to show our gratitude to the lecturers of Hospital Management Graduate Program Universitas Muhammadiyah Yogyakarta who has been participating in this research.

7. REFERENCES

[1] Miri Shonfeld., and Hagit Meishar-Tal. 2016. Writing and Reading Preferences for Student Learning in a Paperless Classroom. Conference: SITE 2016 (Savannah, GA, United States, March 21-26, 2016), 787-792. DOI=<https://doi.org/10.1080/10494820.2018.1504306>.

- [2] Elsa F. H., Nelli P. T., Ruth Roux. 2017. Assessing EFL University Students' Writing: A Study of Score Reliability. *Revista Electrónica de Investigación Educativa*. 19 (2017), 91-103. DOI= <http://dx.doi.org/10.24320/redie.2017.19.2.928>.
- [3] Nora Mogeey., Jessie Paterson., John Burk and Michael Purcell. 2010. Typing compared with handwriting for essay examinations at university: letting the students choose. *ALT-J, Research in Learning Technology* (March 2010), 29-47. DOI= <https://doi.org/10.1080/09687761003657580>.
- [4] M.A. Hearst. 2000. The debate on automated essay grading. *IEEE Intelligent Systems and their Applications* (Sept.-Oct. 2000), 22-37. DOI= <https://ieeexplore.ieee.org/document/889104>.
- [5] Siddhartha Ghosh., Sameen S Fatima. 2009. Design of an Automated Essay Grading (AEG) System in Indian Context. *International Journal of Computer Applications* (February 2010), 60-65. DOI= <https://ieeexplore.ieee.org/document/4766677>.
- [6] Sha Liu., and Antony John Kunnan. 2016. Investigating the Application of Automated Writing Evaluation to Chinese Undergraduate English Majors: A Case Study of WriteToLearn. 2016 *calico journal*. 33 (January 2016), 71-91. DOI= <https://journals.equinoxpub.com/index.php/CALICO/article/view/26380>.
- [7] Mohammad H. Asoodeh., Mohammad B. Asoodeh., Maryam Zarepour., 2012. The impact of student-centered learning on academic achievement and social skills. *Procedia - Social and Behavioral Sciences* (December 2012), 560-564. DOI= <https://doi.org/10.1016/j.sbspro.2012.05.160>.
- [8] Martin Thomas. 2013. Teachers' Beliefs about Classroom Teaching — Teachers' Knowledge and Teaching Approaches. *Procedia - Social and Behavioral Sciences*. 89 (October 2013), 31-39. DOI= <https://doi.org/10.1016/j.sbspro.2013.08.805>.
- [9] Md. Solaiman Jony. 2016. Student-Centered Instruction for Interactive and Effective Teaching Learning: Perceptions of Teachers in Bangladesh. *International Journal of Advanced Research in Education & Technology (IJARET)*. 3 (July - Sept. 2016), 172-178.
- [10] M Ali Fauzi., Djoko Cahyo Utomo., Budi Darma Setiawan., Eko Sakti Pramukantoro. 2017. Automatic Essay Scoring System Using N-Gram and Cosine Similarity for Gamification Based E-Learning. *ICAIP 2017 Proceedings of the International Conference on Advances in Image Processing* (August 2017), 151-155. DOI= [10.1145/3133264.3133303](https://doi.org/10.1145/3133264.3133303).
- [11] Mingqing Zhang., Shudong Hao., Yanyan Xu., Dengfeng Ke., Hengli Peng. 2014. Automated Essay Scoring Using Incremental Latent Semantic Analysis. *Journal of Software*. 9 (February 2014), 429-436. DOI= [10.4304/jsw.9.2.429-436](https://doi.org/10.4304/jsw.9.2.429-436).
- [12] Kaja Zupanc., and Zoran Bosnic. 2015. Advances in the Field of Automated Essay Evaluation. *Informatica an International Journal of Computing and Informatics* (June 2015), 383-395.
- [13] Hairulliza Mohamad Judia., and Noraidah Saharia. 2013. Student-Centered Learning in Statistics: Analysis of Systematic Review. *Procedia - Social and Behavioral Sciences* (November 2013), 844-851. DOI= <https://doi.org/10.1016/j.sbspro.2013.10.406>

- [14] Paul Deane., Frank Williams., Vincent Weng., Catherine S. Trapani. 2013. Automated Essay Scoring in Innovative Assessments of Writing from Sources. *The Journal of Writing Assessment*. 6 (August 2013), 1-44.
- [15] William R. Watson., Sunnie Lee Watson., 2007. An argument for clarity: what are learning management systems, what are they not, and what should they become?. *TechTrends* (January 2007), 28-34.
- [16] Gali Naveh., and Amit Shelef. 2017. Does Platform Matter? A Case Study of Learning Management System. *International Journal of Information and Education Technology*. 7 (Oct. 2017), 749-752. DOI= 10.18178/ijiet.2017.7.10.966.
- [17] Nadire Cavus., Huseyin Uzunboylu., Dogan Ibrahim 2007. Assessing the Success Rate of Students Using a Learning Management System Together with a Collaborative Tool in Web-Based Teaching of Programming Languages. *J. Educational Computing Research*. 36 (April 2007), 301-321. DOI= <https://doi.org/10.2190%2FT728-G676-4N18-6871>.
- [18] Steven Lonna., and Stephanie D. Teasley. 2009. Saving time or innovating practice: Investigating perceptions and uses of Learning Management Systems. *Computers & Education*. 53 (November 2009), 686–694. DOI= <https://doi.org/10.1016/j.compedu.2009.04.008>.
- [19] Stephen P. Balfour. 2013. Assessing Writing in MOOCs: Automated Essay Scoring and Calibrated Peer Review™. *Research & Practice in Assessment*. 8 (2013), 40-48.
- [20] Gorgi Kakasevski., Martin Mihajlov., Sime Arsenovski., Slavcho Chungurski. 2008. Evaluating Usability in Learning Management System Moodle. *ITI 2008 - 30th International Conference on Information Technology Interfaces* (June 2008), 613-618. DOI= [10.1109/ITI.2008.4588480](https://doi.org/10.1109/ITI.2008.4588480).

Authors' background

| Your Name | Title* | Research Field | Personal website |
|----------------------------|------------------------|-----------------------------------|--|
| Qurratul Aini | Senior Lecturer | Hospital Management System | http://mmr.umy.ac.id/en/dosen/dr-qurratul-aini-skg-m-kes/ |
| Achmad Eko Julianto | Student | E-Learning | http://blog.umy.ac.id/makoto/ |
| Dwijoko Purbohadi | Senior Lecturer | E-Learning | http://ti.umy.ac.id/dosen/dr-ir-dwijokopurbohadi-s-t-m-t/ |