ASSESSMENT OF ESSAY QUIZ SYSTEM IN THE LMS VIDYANUSA PLATFORM USING THE NAZIEF AND ADRIANI STEMMING ALGORITHMS "COMPARISON OF TEACHER, EXPERT RUBRIC, AND SYSTEM ASSESSMENTS"

Ratih Kusuma Dewi[1], Mervin Tangguar Hutabarat[2]
[1 2]*Bandung Institute of Technology*
(*[1]ratihkusumadewi09@gmail.com, [2]mhutabar@gmail.com)*

***Abstract*— This study focuses on the assessment system of the essay quiz answers on the LMS VidyaNusa platform using Nazief and Adriani algorithms. The stemming algorithm is a word cutting process that makes use of basic words using certain rules. This method uses LSA (Latent Semantic Analysis) which determines the similarity values of two documents, to find interaction between sentences using SVD (Singular Value Decomposition) algebraic calculation of each statement then to get the similarity value from each statement using cosine similarity algorithm. This system is followed to assist the teacher in evaluating answers to student essay quizzes because this system is equipped with value weighting so that it reduces subjective assessors and increases accuracy in assessment. The results of this study from 15 students by answering 10 questions and 1 teacher obtained results, that the assessment given by the teacher without the vulnerable rubric was 8-10, the value given by the vulnerable system was 0-10, and the assessment of the system with the vulnerable rubric was equal 0 -10. So when viewed from the function of the rubric on the objectivity of the assessment, the value given by the teacher without rubrics is less objective. Correlation or relationship between expert assessment system using rubric with the results of r = 0.82166624 So that both variables have a close relationship and the form of the relationship is linearly positive.**

***Keywords:* LMS; Essay Quiz; Nazief dan Adriani Algorithms; LSA; VidyaNusa; Rubric**

I. INTRODUCTION

E-Learning is a system or concept of education that utilizes information technology in the teaching and learning process. E-Learning is a learning system that is used as a means for teaching and learning that is carried out without having to face to face directly between the teacher and students (Ardiansyah, 2013) [1]. In the concept of e-learning, all teaching and learning processes can be done online, starting from giving material, evaluating learning to the results of evaluation. Evaluation is a process of collecting data to determine how far, in what way, and how educational objectives have been achieved. If not, what is not and why (Ralp Tyler (in Arikunto, 2011: 3)) [2]. Learning evaluation is very important for students to know the understanding of a material. Evaluation of manual evaluation in the form of a test requires a very long time so that a solution is needed to minimize the assessment time of the exam results. Sometimes manual assessment there are several obstacles, namely an inconsistent assessment between students with one another even though the answer is almost the same, then because the writing of each student is different sometimes there is something that is difficult to read it also makes it difficult for the teacher to make an assessment, this can reduce the quality from the teacher's assessment of students' answers sometimes the study is no longer objective. Objectivity is an honest attitude, not influenced by personal or group opinions and considerations in taking decisions or actions. Then at the time of paper quiz collection sometimes the paper was scattered which resulted in invaluable student exam papers and it was very detrimental to students. To overcome these problems, an essay examination system is needed online. This system will correct students' answers quickly and precisely and automatically as well as assessments to be more effective and efficient because usually used paper that accumulates will become a pile of garbage.

A. Essay Question

Essay questions are questions used to measure (the goal) achievement of learning outcomes in complex aspects. And it is recommended that the test designer measure the ability of test participants in the form of analysis, organizing and expressing ideas about something [3]. Also called essay examination is a tool to evaluate learning outcomes, in the form of written questions that demand answers: describe, explain, compare, give reasons, using words and languages themselves or express ideas through written language [4]

B. Automatic Essay Assessment

Manual essay assessments have several problems, such as lack of objectivity in assessment, lack of accuracy, tend to take time [5], to reduce the problem, several automatic essay assessment systems are proposed, including: essay grader (PEG) project developed in 1960- [6] But PEG is criticized because it ignores the semantic aspects of the essay and focuses more on the structure of the word length, the number of semicolons or commas, the number of spaces [6] [7]. Then Intelligent Essay Assessor (IEA) developed by Landauer, Laham, and Foltz. This IEA system uses a semantic text analysis method called Latent Semantic Analysis (LSA). IEA

was developed in 2001 while LSA itself was developed in 1996 [6]. Furthermore Electronic Essay Rater (E-rater) developed by Educational Testing Service (ETS) in 1999. E-rate uses natural Language Processing (NLP). Currently E-rater is used to assess GMAT essay exams (Postgraduate Admissions Manager Tests) [6] [7] [8] [9].

C. Stemming

Stemming is the word cutting process to determine the basic word. The process of stemming this by eliminating all types of affix (affix).

D. Morphology Word Indonesian

Words in Indonesian can be divided into four types, namely basic words, repetition words, derivative words, compound words. Indonesian can be developed into new words or different meanings by adding affixes according to the language grammar. There are several types of affixes or affixes in Indonesian including: prefix, suffix, suffix (confix), insert (infix).

E. Rubric

The rubric is a guide used to assess consistently by the teacher and can be accounted for the quality of student work. Rubrics can also be said to be feedback on the quality of student work.

F. Assessment Formative and Summative

Formative assessment is an assessment process, which is used to obtain information and learning evidence from students to plan the next instructional activities. While summative assessment is an assessment used at the end of the student learning process.

II. METHOD

Nazief and Adriani algorithms have a high degree of accuracy in the stemming process compared to other algorithms. The method used in this study is Latent Semantic Analysis (LSA This LSA concept will look for similarity values between two segments of text without regard to word arrangement. The main function of this LSA is to calculate the similarity of documents by comparing vector representations of each document. In the formation of a vector-based term representation, LSA will form a matrix that represents the relationship between terms and documents called semantic space, that is, words and documents that are closely associated will be placed close to each other represented by vectors [6]. between the sentence and the word algebraic method used is a singular value decomposition.

III. RESULT AND DISCUSSION

There are many development projects using the Nazief and Adriani algorithms. Dheru Alam Perkasa, Eki Saputra, S.Kom, M.Kom, Mona Fronita, S.Kom used Nazief and Adriani algorithms to develop the Essay Online Examination System with Assessment using the Latent Semantic Analysis (LSA) Method. The results of their research between the expected values and the observed values generated by the difference system are only 0.04 [10]. Atqia Aulia, Dewi Khairani, and Nashrul Hakiem. Their research produced a perfect recall score of 1 and an accuracy of 0.961 [11]. Reina Setiawan, Aditya Kurniawan, Widodo Budiharto, Iman Herwidiana Kartowisastro, Harjanto Prabowo. They conducted a study using 1,704 text documents in the forum by comparing their approach better than the confix-stripping approach stemming from Nazief and Adriani [12]. Rahardyan Bisma Setya Putra, Ema Utami. They modified the Nazief and Adriani algorithm with Flexible Affix Classification to be able to stem 40 of the 60 non-formal words [13]

In making this application, several supporting devices are needed including hardware, software and users.

a. Hardware Requirements

This hardware requirement is used at the time of making the system, the following hardware specifications are used:

1. Processor intel i5-7200U CPU @ 2.50GHz 2.71 GHz
2. RAM 12.0GB
3. 1T hard drive
4. Operating system Windows 10

b. Software requirements

This software requirement is used at the time of making the system, the following software specifications are used:

1. WebStorm
2. Robo3T
3. Postman
4. Filezilla
5. Command Prompt
6. MEAN (Mongo, Express, Angular, NodeJS)
7. ArgoUML
8. Microsoft Visio 2010

c. User/ Brainware

To support the system that has been created, this system requires users who are familiar and understand using a computer.

LMS VidyaNusa is made with several menus for teachers and students.

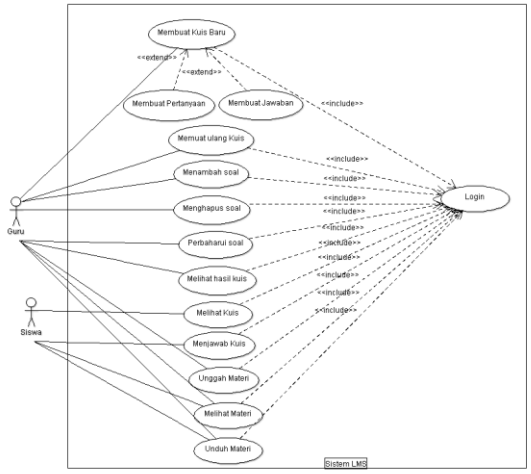

Figure 1 LMS System

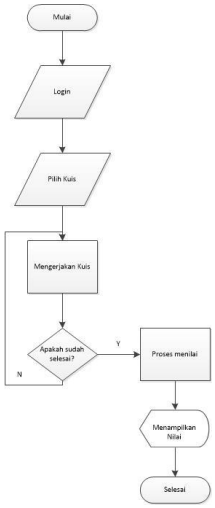

Figure 2 Student Flowchart Quiz

Actor description

1. Guru

Login is the initial step that is done by the teacher or students to enter the LMS module. In the LMS module, there are two main menus in the teacher section, namely making a quiz and uploading material. On the quiz menu the teacher can create a new quiz or reload the existing quiz. There are two quiz choices, namely multiple choices and essays. When you have entered the quiz form the teacher is required to fill out the quiz information first such as the name of the quiz, the name of the lesson, the duration of work, time to work and choose the type of quiz to proceed to the question menu. After entering the menu the teacher must fill in all available forms, namely the question form, answer form and value form. After that when students have taken the quiz, there will be a report containing the weight of the values received from each student. In the upload material menu the teacher must also fill in all available forms including subjects, discussion material, descriptions of the material to be uploaded and the material documents. After uploading the material the teacher can also see the material that has been uploaded on the dashboard menu then the teacher can also download the material.

2. Siswa

The LMS module on the student side has two main menus, namely taking a quiz and downloading material. Before students take the quiz, students are required to log in first, then students choose the quiz menu to take part in the quiz. Students will find out the amount of grades obtained after completing the available quiz. Then besides participating in the quiz, students can also view and download material that has been uploaded by the teacher.

In the flowchart, the student's process flow answers the quiz starting from the login, then the student chooses the quiz that will be followed. After choosing the quiz, students will work on the quiz that has been selected if they have finished working on the student pressing the submit or send answer then the system will process the student's answer and produce the final grade of the student.

Quiz essay application is a web-based application. Users of this application are teachers and students. The teacher will enter the question, answer key, along with the weight of each question into the system. Students see and answer questions, the answers are typed in the form provided. The system will process the student's answers and match the answer keys that have been provided, then will calculate the value of the similarity between the student's answers and the teacher's answers. This process will go through two stages, namely preprocessing and correction of answers using the LSA algorithm. After going through the final score is obtained.

I. Assessment Design Using Nazief and Adriani Stemming

a. Designing questions and Answers

Question: Apa yang dimaksud dengan kuman
Answer Key: Organisme mikroskopis yang bisa menyebabkan penyakit dan infeksi bila mereka masuk ke dalam tubuh
Answer:

- Siswa 1: Organisme yang bisa menyebabkan penyakit dan infeksi pada manusia dan hewan
- Siswa 2: Organisme yang membuat manusia sakit jika masuk kedalam tumbuh
- Siswa 3: Organisme yang dapat menimbulkan rasa sakit dan infrksi nagi masnua jika masuk kedalam tubuh
- Siswa 4: Organisme yang dapat menimbulkan penyakit bagi manusia jika masuk kedalam tubuh.

b. Processing

After the key answers and answers are implemented, the next process is the processing process. In the process there are several stages, namely:
1. Stage of the removal of punctuation marks (stoplist)

2. The phase of solving the sentence becomes a fragment of words (tokenization)
3. Removal of prepositions, links that are not related to the essence of the sentence (stopword)
4. The word cutting stage becomes the basic word (stemming)

c. Answer Correction

1. Correction of answers or the application of the Latent Semantic Analysis method is to determine each query, including the answer key query (Q), query the student answer (D), and query the combined student answer and answer key (QS).
2. Weighting the word or TR (term frequency) from the answer key query (Q), querying the student answer (D), and querying the combined key of the student's answer and answer (QS).
3. Form a matrix with Amxn size
4. Calculate the value of the matrix U, S, VT using the Singular Value Decomposition (SVD) calculation. Manual calculations can use the matrix calculator found in the toolmat website http://www.bluebit.gr/matrix-calculator/calculate.aspx
5. Reduction of k = 2
6. Calculate the key vector answers and answers

The key vector answers with the following formula:

Matrix Q = tf x matrix Uk x matrix Sk

Q = | 1010011010010 | x matrix Uk x matrix Sk

Then the results

Q = 6,296 4,256

The answer vector is obtained from the corresponding column in the Vt matrix of the SVD results with the following results:

D1 = 0.317 0.916

D2 = 0544 -0.030

D3 = 0.526 -0.376

D4 = 0.572 0.133

7. Calculate CoSim values from each matrix Q and D, with calculations using the following formula:

$$\text{CoSim}\ (D_n Q) = \frac{\sum \text{Q} D_n}{|\text{Q}||D_n|}$$

The CoSim calculation uses the Cosine Similarity Calculator with a Web address http://scistatcalc.blogspot.com/2015/11/cosine-similarity-calculator.html#.

The results of the calculation are as follows:

CoSim (D_1 Q) = 08002
CoSim (D_2 Q) = 0.8284
CoSim (D_3 Q) = 0.3483
CoSim (D_4 Q) = 0.9338

8. Multiply the cosim value by the value weight

In this case the weight of the problem is 100. The results are as follows:

Value D1 = 80.02 (Student final score 1)

Value of D2 = 82.84 (Student's final grade 2)

Value D3 = 34.83 (Student final score 3)

Value of D4 = 93.38 (Student's final grade 4)

II. Testing Value Validation

The test is done by giving a quiz in the form of essay questions to students then the teacher assesses manually and compares the value given by the teacher to the assessment carried out by the system. The following is an assessment conducted by the teacher and system:

Question : Apa yang dimaksud dengan merangkum
Answer Key : Kegiatan menyusun gagasan pokok dari suatu bacaan/ buku menjadi bentuk yang ringkas atau pendek

Table 1 Testing Value Validation

No	Student Id	Answer	Grade		
			Teacher	System	Rubric
1	599q	menyimpulkan kata yang pentingnya saja	10	7.41	6
2	59d2	persingkat uraian dengan perbandingan secara proporsional	8	1.3	5
3	59fw	meringkas suatu paragraf dengan cara mengambil poin-poin pentingnya saja	10	7.32	7
4	5a0z	suatu kegiatan yang mengambil pokok-pokok suatu tulisan atau pembicaraan atau menjadi suatu uraian yang lebih singkat dengan perbandingan secara proporsional antara bagian yang dirangkum dengan rangkuman	10	3.73	4
5	5a0x	mencatat pokok-pokok teks dan menyusunnya menjadi singkat dan jelas	10	8.23	8
6	5a0c	memendekan suatu paragraf yang panjang menjadi lebih pendek	8	2.52	4
7	5a0v	mempersingkat uraian dengan	10	2.52	4

No	Student Id	Answer	Grade		
			Teacher	System	Rubric
		perbandingan secara proporsional			
8	5a0b	merangkum atau kegiatan meringkas yang isinya penting-penting saja	10	2.54	6
9	5a0n	merangkum adalah mempersingkat uraian dengan perbandingan secara proporsional	10	1.3	4
10	5a1s	merangkum adalah suatu tek yang hanya di ambil intinya saja, jadi mengambil yang pentingnya saja	10	6.91	6
11	5a0w	merangkum adalah mengambil inti-inti dari sebuah teks yang membuatnya menjadi lebih ringkas.	10	9	6
12	5a0u	merangkum adalah pengambilan pokok-pokok tulisan atau pembicaraan yang penting saja menjadi uraian yang lebih singkat	10	8.33	9
13	5a0	merangkum adalah menyimpulkan kata yang pentingnya saja	10	7.18	7
14	5a0m	merangkum adalah menyimpulkan kata yang pentingnya saja	10	7.18	7
15	5a0s	merangkum adalah kegiatan menyusun gagasan pokok dari suatu bacaan / buku menjadi bentuk yang ringkas atau pendek	10	9.47	9

Vulnerable values that are given by the teacher manual system and assessment are far enough, while assessments that use rubrics and systems are not far enough. For more details, here is a comparison chart of assessments given by the teacher, system, and using rubrics.

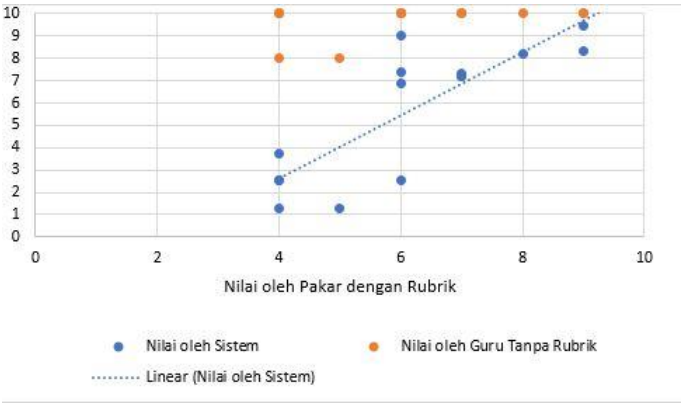

Figure 3 Analisys Beween System Assessment, Experts With Rubrics, and Teacthers without Rubrics

From the diagram above shows that the value given by the teacher without the vulnerable rubric is 8-10, the value given by the vulnerable system is 0-10, and the evaluation of the system with the vulnerable rubric is equal to 0-10. So when viewed from the function of the rubric on the objectivity of the assessment, the value given by the teacher without rubrics is less objective. In the essay quiz assessment, a rubric is needed to help guide the objective assessment, information on the value weight, determine student learning strategies. So with the assessment system, the essay quiz can help the teacher's assignment for assessment.

III. Discussion

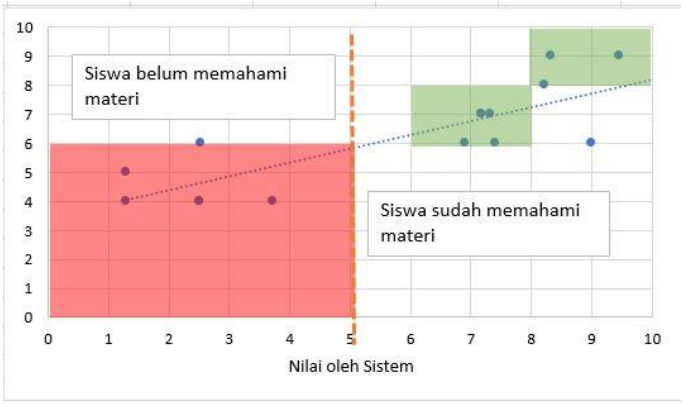

Figure 4 Corelation of System Assessment and using Rubrics

From the picture, shows that there is a correlation or relationship between the appraisal system and expert judgment using a rubric with the results of r = 0.82166624 So that both variables have a close relationship and the form of the relationship is linearly positive. When checking the variance value of the two data, it is found that the value for variance using the rubric is s2 = 2.98095238 and the variance using the system is s2 = 8.75746381. From the results of the variance can be seen the distribution of its value. If the value of variance uses the rubric s2 = 2.98095238, it means that it is close to the average value of the student as a whole. Because it is getting closer to 0, the data is getting more homogeneous and vice versa. The value of the variance of the system value indicates

that the results of system data vary, between students one with other students. Assessment using this system can be used as a formative assessment..

IV. Conclusion

This system needs further research, because there is still a considerable difference between the assessment of the system and the teacher's assessment, for example students get a score of 10 using a manual assessment by the teacher and get a score of 1.39 using the assessment by the system, although this assessment has been assisted by the rubric assessment.

This appraisal system can help teachers, but it still needs system development that is adding a method that can recognize the similarity of words with different representations but has the same meaning.

V. Acknowledgment

This is my thesis project, , I would like to thank Mr. Kokon Komarudin for his assistance in making assessment rubric and assessing the student's quiz using the rubric and Mrs. Wati who helped in getting a comparison in the assessment of this system.References

VI. REFERENCES

[1] M. Riyadi, "Pengertian, Karakteristik dan Manfaat E-Learning," 17 June 2014. [Online]. Available: https://www.kajianpustaka.com/2014/06/pengertian-karaktiristik-dan-manfaat-elearning.html. [Diakses 03 July 2018].

[2] Nurwahyuningsih, "ANALISIS KUALITAS SOAL TRY OUT ... Bab 2 - 04201244078," 1 Februari 2013. [Online]. Available: http://eprints.uny.ac.id/9672/3/bab%202%20-%20%2004201244078.pdf - http://eprints.uny.ac.id/9672/. [Diakses 03 July 2018].

[3] S. Anwar, Penilian Berbasis Kompetensi, Padang: UNP Press, 2009.

[4] N. Sudjana, Penilaian Hasil Proses Belajar Mengajar, Bandung: PT. Remaja Rosdakarya, 1991.

[5] A. A. P. Ratna, B. Budiardjo dan D. Hartanto, "SIMPLE: SISTIM PENILAI ESEI OTOMATIS UNTUK MENILAI UJIAN," *MAKARA, TEKNOLOGI*, vol. 11, pp. 5- 11, 2007.

[6] Turkish Online Journal of Distance Education-TOJDE, "Automated Essay Scoring," *Semire DIKLI. Florida State University Tallahassee, FL, USA*, vol. 7, no. January 2006 ISSN 1302-6488, p. Number: 1 Article: 5, 2006.

[7] A. A. P. Ratna, A. Henry dan A. B. Anantasatya, "GLSA Based Online Essay Grading System," dalam *IEEE - TALE*, Bali, 2013.

[8] S. D. S. L. T. F. G. a. H. R. Deerwester, "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science & Landauer*, 1990.

[9] J. e. a. Burstein, "Enriching automated essay scoring using discourse marking. Proceedings of the Workshop on Discourse Relations and Discourse Marking," dalam *Annual Meeting of the Association of Computational Linguistics*, Montreal, Canada, 1998.

[10] D. A. Perkasa, Sistem Ujian Online Essay Dengan Penilaian Menggunakan Latent Semantic Analysis (LSA) Pada Jurusan Sistem Informasi Universitas Islam Negeri Sultan Syarik Kasim Riau, Riau: Fakultas Sains dan Teknologi. UIN Suska, 2014.

[11] A. A. , D. K. dan N. H. , "Development of a Retrieval System for Al Hadith in Bahasa (Case Study: Hadith Bukhari)," dalam *5th International Conference on Cyber and IT Service Management (CITSM)*, Denpasar, pp. 1-5, 2017.

[12] K. A. B. W. K. I. H. P. H. S. Reina, "Flexible Affix Classification for Stemming Indonesian Language," dalam *Internasional Conference IEEE*, 2016.

[13] P. R. B. S. U. Ema, "Non-formal Affixed Word Stemming in Indonesian Language," dalam *International Conference on Information and Communications Technology (ICOIACT)*, Yogyakarta, 2018.

[14] R. B. Aji, "Automatic Essay Grading System Menggunakan Metode Latent Semantic Analysis," dalam *Seminar Nasional Aplikasi Teknologi Informasi 2011 (SNATI 2011)*, Yogyakarta 17-19 Juni 2011, 2011.