COSINE SIMILARITY FOR ESSAY ANSWER DETECTION

Laurensius Rendi Setiawan[a], Dewi Nasien[b*]

[a,b]Department of Computer Science, Institut Bisnis dan Teknologi Pelita Indonesia, 28156, Pekanbaru, Indonesia

Abstract

Saint Mary Senior High School is the one of the famous private schools in Pekanbaru. examination in Saint Mary Senior High School has used a smartphone for multiple choice questions, while in essay questions the pen-and-paper is still used (manual). The number of examination scripts received from students is a problem for a teacher to see the similarity of students' answers with the answer key. At present, there is no system that can assist in checking answers. This study began with the selection stage, which is done by the interview process and filling out the questionnaire. Then proceed with an analysis of this case is used to look for possible development needs of the system. Then the programming language used is PHP and uses MySql database. Web-Based Essay Answer Detection Application Using Cosine Similarity Method in Saint Mary Senior High School. With the application of the essay answer detection application in Saint Mary Senior High School, it is hoped that it will be easier for the teacher to conduct the exam. Besides, it is expected that the application of a computerized system and the use of a database can accelerate the processing of student ranking grades effectively and efficiently.

Keywords: Detection, Essay Answer, Web, Cosine Similarity.

1.0 INTRODUCTION

Along with the development of Information and Computer Technology (ICT) today, the world of education is experiencing a very significant system change. Some educational institutions have begun implementing e-learning based learning or often also referred to as electronic learning. According to (Hartanto, 2016), e-learning is information and communication technology to enable students to learn whenever and wherever. Electronic learning or e-learning began in the 1970s. Various terms are used to express opinions or ideas about electronic learning, such as: on-line learning, internet-enabled learning, virtual learning, or web-based learning. E-learning aims to bring a new atmosphere in the effort to develop learning media, by shortening learning time, making students more practical, and can facilitate interaction between students with the material to be learned. Students can share information with each other and as often as possible to learn, so it can be deeper to strengthen in the mastery of the material provided.

Santa Maria High School is under the auspices of the Prayoga Riau Foundation and the educational reference is based on the 2013 curriculum. To be able to enter the school, students must take many tests. One of the tests given is a written test. With the written test, students can find out the level of understanding of the material presented. Santa Maria High School has 2 majors, namely Natural Sciences and Social Sciences. At present the school has as many as 24 classrooms with a total of 42 teachers.

In making questions, the teacher himself makes the questions, except for semester questions that are made based on the Subject Teachers' Consultation (MGMP). Because the number of exams received by Santa Maria High School students, is a problem faced by a teacher in the similarity of answers between students with one another with the answer key. The current test, the teacher only uses the manual method in the examination. The absence of a system that helps in the examination of these answers. Based on this background the researchers tried to solve the problems faced by teachers in Santa Maria High School by using the Cosine Similarity Method.

   The selection of Cosine Similarity method is based on previous research. According to (Firdaus, 2014) the Cosine Similarity method can be applied in determining the similarity values in two text document files. Whereas according to (Fitri, 2015), the cosine similarity method aims to equalize the two sentences in a document by comparing and seeing how big the results of the similarity are. The cosine similarity method is a method that has been around since 2014. Although this method is an old method, this method is still used today.

2.0 METHODOLOGY

System Development Life Cycle (SDLC) Method SDLC is a system development cycle consisting of Problem Identification, Analysis, Design, Programming, Testing, Maintenance.

1. Identification of Problems In the identification and selection stage, the interview and questionnaire filling processes were carried out. The interview was conducted on July 1, 2019 with one of the teachers named Ibu Mega Veronika who was the source of this research. The interview process can be seen on the website: (http://bit.ly/wawancara-penelitian). While the results of filling out the research questionnaire can be seen on the website: (http://bit.ly/dataangket2019).

2. Analysis At this stage the analysis process is carried out such as software and hardware requirements in the implementation process. Required hardware as listed below:

   a) Hardware device 1. Server computer specifications: Intel (R) Xeon (R) CPU E3-1225 v5 @ 3.30GHz (4 CPUUs), 3.3GHz, 4GB RAM, and 1TB HDD (HDD) 2. Media client specifications: - Android: Android 5.0 (Lollipo) - Android 9.0 (Pie), 4 GB RAM, and 32 GB Memory -Computer: Intel (R) Core (TM) i3-2100, 8GB RAM, and 500GB HDD b) Software Devices

   1.OS Windows 7

   2. Adobe Dreamweaver (As a Web Editor)

   3. XPAMP Control Panel

   4. Apache

   5. MySQL

   6. Web Hosting

3. Design This third sub step is carried out to produce a number of essay examination system design that will be compared and it is ensured that only one of the best new designs is implemented according to the needs using mobile android or computer.

4. Programming After the design process is complete, the next step is making the coding script. The programming language used is PHP while for XAMPP database creation.

5. Testing At this stage the test includes: a.Do repair and test the system from indications of defects, bugs, errors and so on in the future that will be used by the user. b. Making a database in accordance with the concept and design scheme. c. Take responsibility and guarantee that the new system can run optimally, effectively and efficiently.

6. Maintenance At the maintenance stage, maintenance must be carried out after the program is declared running and is appropriate to use, which is to maintain and maintain the program periodically so that it can be used optimally, effectively and efficiently, by updating every time there is the latest information.
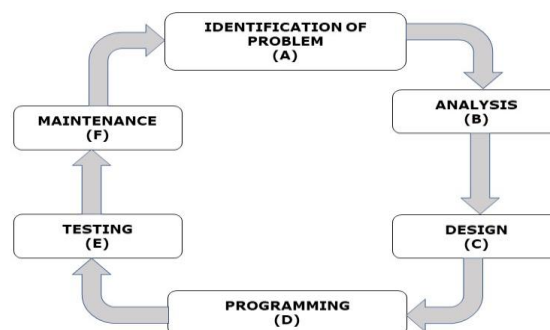
Picture 1. System Development Life Cycle

3.0 RESULTS AND DISCUSSION

Application Design
Detailed Design
   a. Login Menu Design
   The input design is the display on the android screen or on the computer, before it is operated by the user. Display input functions as a liaison between the application and the user to input the data needed to produce output in accordance with what is needed.

Picture 2. Login Menu Design

   b. Teacher's Menu Design
Figure 2 explains where the teacher will check the students 'answers by looking for the students nis first, then, selecting the answer key code that was created by the teacher beforehand. After completion the data will appear in the available column. Then the teacher just presses the process button so automatically the similarity results appear between the student's answer and the answer key. The teacher can save the results obtained, before saving the teacher is required to fill in grades manually, this aims if the calculation of the method is not right, or one of the students is not complete answer then the teacher can enter grades directly.

Picture 3. Teacher's Menu Design

   c.   Student Menu Design
   Figure 4 explains the form of student worksheets when working on essay questions, students are required to choose the question code provided by the teacher. After selecting the appropriate question code, it automatically appears in the question column and then students can answer the question. After completion, students must press the send button, so the answers sent are stored in the database.

REPORT OF STUDENT EXAM RESULTS

SMA SANTA MARIA PEKANBARU

Part : Kepala Sekolah

Name : Norita Rismawati, S.Pd

No	Nis	Name	Gender	Subjects	Score	Ranking	Class	Majors
1	8781	CLARISSA KIMBERLY TANUWIJAYA	Perempuan	MATEMATIKA	100	1	XII	IPS 1

Back

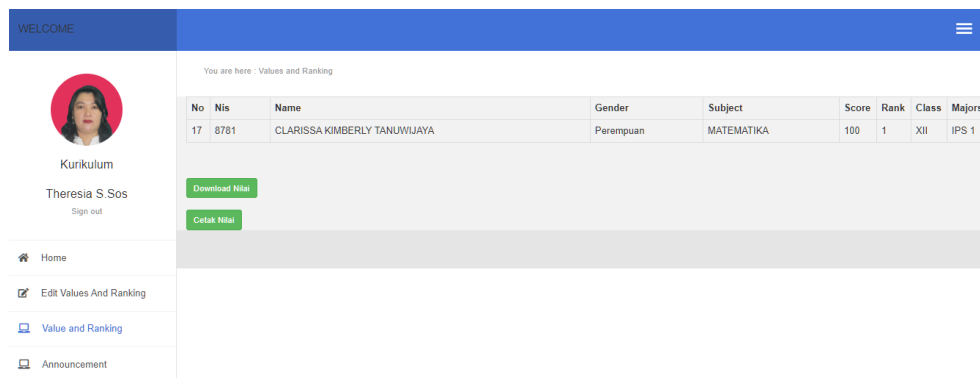**Picture 4.** Student Menu Design

d.   Curriculum Menu Design

Figure 5 explains where the curriculum can see the grades and ranks of students after the teacher checks the answers and the curriculum provides the ranking, in the form there are two features namely download grades and print grades, this aims so that the curriculum has an archive of grades and scores in each subject.

Picture 5. Curriculum Menu Design

e.   Headmaster Menu Design

Figure 6 explains the printout of the essay exam report. It is intended that the principal also has a printed archive of grades and ranks in each subject.

Picture 6. Menu Design Report To School

4.0 CONCLUSION

1. With the cosine similarity method in Santa Maria High School, it can increase the effectiveness in checking each student's answer accurately.
2. With the application of the essay answer detection application in Santa Maria High School, it can make it easier for teachers to conduct monthly exams so that they can quickly find out every ability of their students.

Acknowledgement

References

Arie, Azhar, Ernawati, et. al. (2014). Aplikasi Pendeteksi Kemiripan pada Dokumen Teks Menggunakan Algoritma Nazief & Andriani Dan Metode Cosine Similirity. *Jurnal Teknologi Informasi*, *10*(April), 96–109.

Azhar Firdaus, Ernawati, A. V. (2014). Aplikasi Pendeteksi Kemiripan pada Dokumen Teks Menggunakan Algoritma Nazief & Andriani Dan Metode Cosine Similirity. *Jurnal Teknologi Informasi*, *10*(April), 96–109.

Azmi, A. N., Nasien, D., & Shamsuddin, S. M. (2013). *A review on handwritten character and numeral recognition for Roman, Arabic, Chinese and Indian scripts*. Retrieved from http://arxiv.org/abs/1308.4902

Firstiara Maudi, Arief Laila Nugraha, B. S. (2014). DESAIN APLIKASI SISTEM INFORMASI PELANGGAN PDAM BERBASIS WebGIS (STUDI KASUS :   KOTA   DEMAK).   *Jurnal   Geodesi   Undip*,   *3*(3),   98–110.   Retrieved   from https://ejournal3.undip.ac.id/index.php/geodesi/article/view/5840

Hartanto, W. (2016). Penggunaan E-Learning Sebagai Media Pembelajaran Wiwin Hartanto adalah staf mengajar Prog. Studi Ekonomi FKIP UNEJ. *Jurnal Pendidikan Ekonomi*, *10*(1), 1–18.

Ogie Nurdiana, Jumadi, Di. N. (2016). Perbandingan Metode Cosine Similarity Dengan Metode Jaccard Similarity Pada Aplikasi Pencarian Terjemah Al-Qur'an Dalam Bahasa Indonesia. *Jurnal Online Informatika*, *1*(1), 59. https://doi.org/10.15575/join.v1i1.12

Rahimi dan Arifin. (2015). Aplikasi Penilaian Ujian Essay Otomatis Menggunakan Metode Cosine Similarity. *Poros Teknik*, *7*(2), 88–94. https://doi.org/10.31961/porosteknik.v7i2.218

Rito Putriwana Pratama, Muhammad Faisal, A. H. (2019). *Deteksi Plagiarisme pada Artikel Jurnal Menggunakan Metode Cosine Similarity*. *5*(1), 22–26.

Rizki Tri Wahyuni, Dhidik Prastiyanto, E. S. (2017). Jurnal Teknik Elektro. *Jurnal Teknik Elektro*, *9*(1), 18–23. Retrieved from https://journal.unnes.ac.id/nju/index.php/jte/article/view/10955/6659

Silitonga, J. (2013). *Mahasiswa Pendaftaran Baru Berbasis Web - jurnal.umrah.ac.id*. 3–5.