

IMPLEMENTATION OF GENERALIZED VECTOR SPACE MODEL METHOD AT AUTOMATIC ASSESSMENT OF ONLINE ESSAY EXAM

Muhammad Arafah¹, Pasmur², Fahmi Idrus³

¹STMIK AKBA, Indonesia
arafah@akba.ac.id

²STMIK AKBA, Indonesia
pasmur@akba.ac.id

³STMIK AKBA, Indonesia
fahmi@akba.ac.id

Abstract-- The aim of the study was to design and implement automatic testing of online essay examinations using the Generalized Vector Space Model (GVSM) method. This data is obtained through (1) Literature Study (2) Observation (3) Documentation. The results of this study indicate that the automatic scoring system with the GVSM weighting method and the cosine similarity similarity calculation method have the accuracy of the assessment with an average of 66%.

Keywords: Automatic Assessment System; Cosine Similarity; Essay Online Exam, Generalized Vector Space Model.

I. INTRODUCTION

Aside from being a teacher, lecturers also act as evaluators. One way lecturers evaluate their students is by giving written tests in the form of essay questions. The evaluation technique that is still applied by lecturers at this time is manual evaluation techniques. This evaluation technique is done by giving a number of questions to the student then the student answers the questions on the answer sheet that has been prepared beforehand. In addition, the lecturer must also check the answers of the student manually.

Assessment of the results of the exams conducted manually has a weakness, namely there are different assessments of the participants with almost the same answers. In addition, the teacher takes a long time to examine the essay's answers. The more number of examinations and the greater number of students taking the exam, number of examinations have to be examined by the teacher then more. This certainly causes the quality of the assessment to decline and the quality of the assessment is not objective anymore.

This study proposes a solution to the lack of manual evaluation techniques by implementing an online test-based system. Besides paperless, this application can also analyze the answers of each student quickly.

To support system performance in assessing the results of the exam, a method that able to measure the similarity of the text in online essay exam evaluations is needed, so the instructor work will more effective and make easier to do a

more accurate assessment.

Methods related to measuring text similarity have been proposed by some researchers. Some of these methods include a research by Amin that entitled Information Retrieval System with Vector Space Model Method [1]. Wibowo, et al. also doing Implementation of Generalized Vector Space Models Using WordNet [2].

These methods can be used to measure the similarity of text in evaluating online essay exams with certain advantages and disadvantages. The main weakness of VSM is that this method assumes that each term in the query and document is independent, in this case it does not consider the value of correlation in each query and document. Therefore, the GVSM method was chosen which could be a solution to find the accuracy of text similarity in online essay exam assessment.

This study aims to measure the similarity of student essay answers and answer keys by implementing GVSM so that an automatic assessment of the online essay exam will be easier to do.

II. METHOD

A. Assessment of Essay Exams

An essay-based exam is one of evaluation form that didn't provided an answer choices, so students will be train to convey something verbally. Beside that, essay exams also require a better understanding of a science and can be used to measure the level of understanding of human knowledge deeply [3].

B. Information Retrieval

Information retrieval is a system used to retrieve information that is elevated to the needs of users of a collection of information automatically. In addition, the information retrieval system does not help (does not change the knowledge) of the user regarding the problem being asked. The system only notifies the existence (or non-existence) and the description of the documents related to the request [4].

C. Preprocessing Text

Preprocessing Text is a stage before the calculation of word weights needed to clean, eliminate, change data sources, both

in the form of non-alphabet characters and unnecessary words [6]. The following is a preprocessing text process consisting of folding cases, tokenizing, filtering, and stemming. The flow chart can be seen in Figure 1.

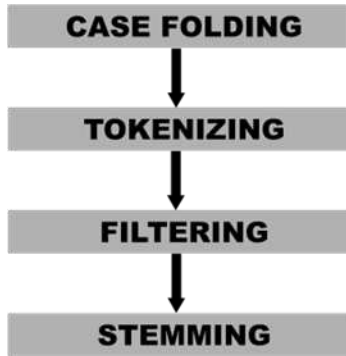


Fig.1. Preprocessing Text

Case folding is the process of converting all text into lowercase characters and removing all characters other than a-z. If there are punctuation marks, numeric numbers and symbols, it will be omitted. Tokenizing is a process of changing sentences into single words.

In the filtering process, the stopword is removed. Stopword is words that do not have meaning or words that are less meaningful and often appear in a collection of words. The next process is stemming, which is to return a root word.

D. Cosine Similarity

Cosine similarity is a measure of the similarity between two vectors in a dimensional space obtained from the angular cosine value of the multiplication value of two vectors compared because the cosine of 0 is 1 and less than 1 for other angular values, then the similarity value of two vectors said to be similar when the value of cosine similarity is 1 [6].

$$\text{sim}(d_j, q) = \frac{d_j \cdot q}{\|d_j\| \|q\|}$$

$$\text{sim}(d_j, q) = \frac{\sum_{i=1}^N w_{i,j} w_{i,q}}{\sqrt{\sum_{i=1}^N w_{i,j}^2} \sqrt{\sum_{i=1}^N w_{i,q}^2}}$$

Here is the formula for cosine similarity:

When:

d_j = Document j

q = query Document

$$\sqrt{\sum_{i=1}^N w_{i,j}^2} = \text{weight of } i \text{ words in document } j$$

$$\sqrt{\sum_{i=1}^N w_{i,q}^2} = \text{weight of } i \text{ words in query document}$$

E. Generalized Vector Space Model Method

The Generalized Vector Space Model (GVSM) is a development of the Vector Space Model that considers the proximity of the sense between terms more accurately, in representing documents. Wong et al made the first GVSM,

which introduced a correlation between terms, which assumed that each term was expressed as a linear combination of 2-dimensional vectors [2].

GVSM is one model of the Information Retrieval (IR) system which is included in the Algebraic Model which is an extension of VSM, the process that occurs in GVSM is divided into two namely preprocessing stages which consist of reading text, tokenizing, filtration, stemming and parse queries. The second process is to calculate the relevance between a set of preprocessing user answers and system answers [7].

There are several steps or processes to get results from the entered query:

1. Discard prepositions and conjunctions.
2. Using stemmer in documents and queries, namely applications that are used to eliminate additions (prefix, suffix).

Example: majesty = great; eternity = eternal.

3. Determines the minterm to determine the possibility of the word frequency pattern. This minterm length is based on many words entered in the query. Then it is converted into an orthogonal vector according to the minterm pattern that appears. A possible pattern will appear:

$$M_1 = (0,0,0)$$

$$M_2 = (1,0,0)$$

$$\dots\dots\dots$$

$$M_n = (0,0,0)$$

4. Calculate the number of frequencies or occurrences of words in a document that matches the query.
5. Calculate the index terms that can be stated with:

$$Kl = \frac{\sum \forall r, g_i(Mr) = 1^{Ci,r} \overline{Mr}}{\sqrt{\sum \forall r, g_i(Mr) = 1^{Ci,r^2}}}$$

When

\overrightarrow{Kl} : Index term i

\overrightarrow{Mr} : orthogonal Vector according to the used minterm pattern

Ci, r : Correlation factor between Index term i with mint r

Calculate the correlation factor as follows:

$$\overrightarrow{d_j} = \sum_{i=1}^n w_{i,j} \times \overrightarrow{K_i} \dots\dots\dots (3)$$

$$\overrightarrow{q} = \sum_{i=1}^n q_i \times \overrightarrow{K_i} \dots\dots\dots (4)$$

When:

$w_{i,j}$: weight index term i in document j

$g_i(Mr)$: weight index term K_i in minterm M_r

6. Turn documents and queries into vectors:

$$\text{sim}(\overrightarrow{d_j}, \overrightarrow{q}) = \frac{\overrightarrow{d_j} \cdot \overrightarrow{q}}{\|\overrightarrow{d_j}\| \|\overrightarrow{q}\|} \dots\dots\dots (5)$$

when :

$\overrightarrow{d_j}$: Vector document ke- j

\vec{q} : Vector query

$W_{i,j}$: weight index term i in document j

q_i : weight index term in query i

n : amount of index term

7. Sort documents based on similarity, by calculating vector multiplication:

$$C_{i,r} = \sum_{d_j | g_i(\vec{d}_j) = g_i(Mr)} W_{i,j}$$

When:

\vec{d}_j : Vector document j

\vec{q} : Vector query

F. Software Used

1. PHP

Hypertext Preprocessor or abbreviated as php is a scripting language specifically used for web development. Because of its server side scripting, running PHP must use a web server to run PHP [8].

2. HTML

Hypertext Markup Language is a protocol used to format web documents that can be read in browsers from various compute platforms [8].

3. MySQL

MySQL (My Structure Query Language) or the usual "maise-kuel" is a database creation program that is open-source, meaning anyone can use it and not block it [9].

G. System Design

1. Use Case

The use case diagram is a series of groups that are interrelated and form a system regularly that is carried out or supervised by an actor. Use case diagram serves to describe what features will be run on the software that will be created. Use Case system diagram can be seen in Figure 2.

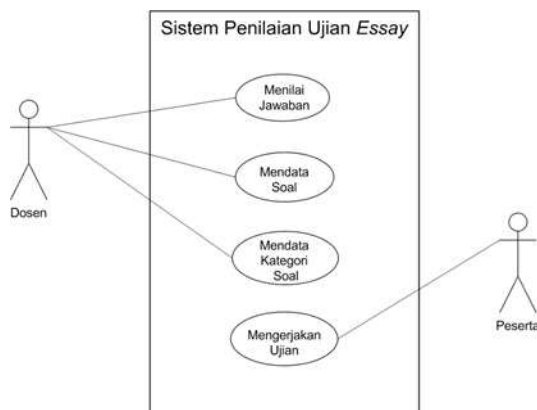


Fig. 2. Use Case Diagram

1.1 Lecturers

Lecturers are actors who can see, add, change and delete question data and essay exam evaluations.

1.2 Participant

The Participant is an actor who can see and answer the essay exam.

2. Flow Chart of System Overview

The design of the system model that used in this study is as shown in Fig. 3 and Fig. 4.

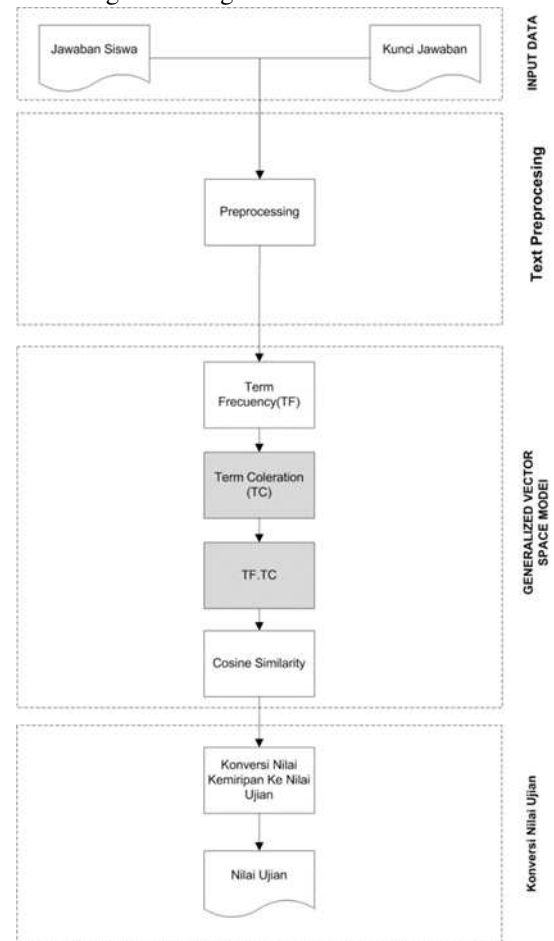


Fig 3. Research Design

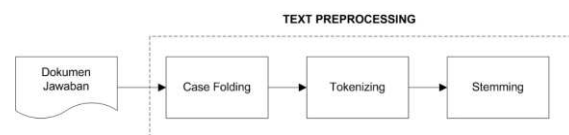


Fig 4. Preprocessing Text

Based on the design of the system model above, there are several things that must be done in the system, namely:

2.1 Data Input

The data entered into the system is the data entered by teaching staff and students, namely, question questions, answer keys after the answers of the Participants.

2.2 Preprocessing Text

In this stage all the answer documents entered into the system will be made a Preprocessing Text. The stages in this processing are as shown in Figure 1.

2.3 Term Frequency (TF)

In this stage, weights the word in each document by calculating the frequency of occurrence of words in each document.

2.4 Term Colleration (TC)

At this stage the weighting of the correlation values between the terms in all the answer documents that reference the frequency term.

2.5 Weighting TF.TC

At this stage, the GVSM is weighted by multiplying the TF value with TC to get vector length in each document.

2.6 Calculating Similarities With Cosine Similarity

This stage calculates similarities from the students' essay answers with the key answers that the teacher has, so at this stage only compares two documents. This stage produces output in the form of similarity values between the two documents which will then be converted into the test participant's score.

2.7 Conversion of Similarity Values to Essay Exam Values

The similarity values obtained previously will be converted into the exam participants' essay scores based on the range of values of the human assessment

3. Tabel Relation

Database design (database) aims to store data that are interconnected with one another. In designing this essay assessment system consists of one database and several tables which can be seen in Fig. 5.

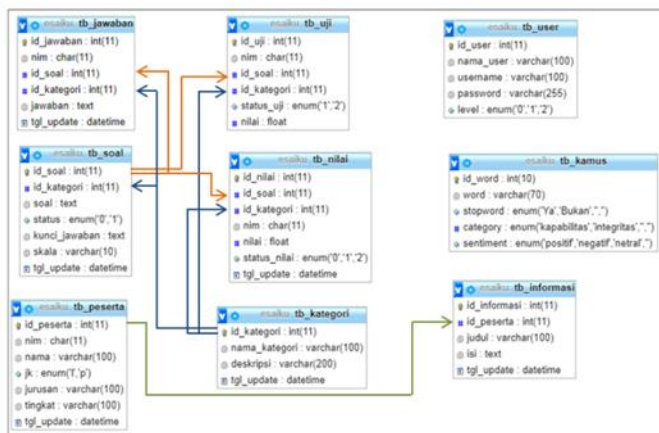


Fig. 5. Tabel Relation

III. RESULT AND DISCUSSION

A. Test Result

System accuracy testing is done by measuring the level of closeness between the automatic values provided by the system and the correction value (manual) given by the lecturer. The following formula calculates accuracy on the system:

$$Y > X \rightarrow Akurasi = \frac{(Y-(Y-X))}{Y} \quad || \quad Y < X \rightarrow Akurasi = \frac{(Y-(X-Y))}{Y}$$

when:

X = automatic assessment results

Y = manually assessment results

B. Access level

In the implementation of the Generalized Vector Space Model method for automatic assessment of online essay examinations there are two levels of access, namely, the level of access for lecturers and the level of access for Participants.

1. Access Level for Lectures

At the access level for lecturers there are several features including:

1.1 Form Login

The login form at the access level for lecturers is a feature provided to limit users who can enter the system by verifying based on their username and password.

1.2 Main Page

The main page will be displayed for the lecturer access level if the username and password entered on the login form have been validated, as for the display of the main page for the lecturer access level as shown in Fig. 6.



Fig. 6. Main Page Lectures

1.3 Question Category

The question category feature is used to add data to the question categories which will be tested to the test participants. The appearance of this feature shown in Fig 7.



Fig.7. Question Category

1.4 Question Data

In the question data feature, lecturers can display question data and add data and edit and delete question data that will be tested. The appearance of this feature show in Fig.8.

| No | Soal | Kategori | Materi | Jawaban | Skor | Pilihan |
|----|----------------------------------------------------------------------------------------------------------|----------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|------|---------|
| 1 | Sebutkan bunyi pasal 21 ayat 1 | PKN | Setiap warga negara berhak mendapatkan pendidikan. | Esok | 20 | 100% |
| 2 | Jelaskan keterkaitan Pancasila sebagai dasar negara dengan UUD 1945 sebagai konstitusi negara Indonesia? | PKN | Sebagai dasar negara Pancasila berfungsi dalam pembentukan UUD 1945 karena ke-4 sila itu Pancasila mempunyai kekuatan mengikat secara hukum sehingga semua peraturan hukum yang bertentangan dengan Pancasila harus dihapus. | Esok | 35 | 100% |
| 3 | Apakah yang mendorong lahirnya suatu bangsa? | PKN | Yang mendorong lahirnya suatu bangsa adalah adanya keinginan dan cita-cita serta dorongan oleh berbagai latar belakang baik etnis, geografis, kultural maupun politik. | Esok | 20 | 100% |
| 4 | Apakah manfaat norma dalam kehidupan sehari-hari? | PKN | Manfaat norma dalam kehidupan sehari-hari yaitu menciptakan masyarakat yang tertib, damai dan sejahtera. | Esok | 15 | 100% |
| 5 | Jelaskan apa yang dimaksud dengan norma? | PKN | norma adalah aturan-aturan atau kebiasaan yang berlaku dalam masyarakat. | Esok | 10 | 100% |
| 6 | Jelaskan, apa yang dimaksud dengan Tatawika? | PAI | Tatawika artinya berakhlak diri kepada Allah SWT yaitu menyerahkan segala urusan kepada Allah SWT sebagai Pencipta (beribadah) serta mematuhi aturan-aturan yang ditetapkan Allah yang dijunjung. | Esok | 20 | 100% |
| 7 | Jelaskan apa yang dimaksud dengan Tawakal? | PAI | Tawakal adalah menyerahkan segala urusan kepada Allah SWT dan tidak bergantung kepada manusia lain. | Esok | 25 | 100% |
| 8 | Jelaskan pengertian shalat sebagai ibadah kepada Allah? | PAI | Shalat yang dilakukan secara beraturan yang dilakukan oleh dua orang atau lebih, satu orang memimpin dan lainnya mengikuti, sebagai pengingat untuk Allah dan diri sendiri. | Esok | 10 | 100% |
| 9 | Jelaskan mengapa Allah SWT disebut Penyayang (Ar-Rahman)? | PAI | Karena Allah SWT itu baik hati dan penuh kasih sayang kepada makhluk-Nya yang beriman dan beramal saleh. Dia menurunkan hujan sebagai rahmat-Nya yang melimpah, Dia menurunkan angin sebagai rahmat-Nya yang menyejukan, Dia menurunkan air sebagai rahmat-Nya yang menyiratkan kehidupan. | Esok | 30 | 100% |

Fig. 8. Question Data

1.5 Participant Data

In the participant data feature, the lecturer can display and add exam participant data, as for the display to add participant data shown in Fig. 9.

Fig. 9. Add Participant

1.6 Evaluation

The evaluation feature is a feature to display the value of student exam results, where this value will be made automatically by the system by referring to the assessment method. The appearance of this feature is shown in Fig. 10.

| No | NIM | Nama | Jurusan | Tingkat | Nilai | Pilihan |
|----|------------|----------------------|-----------------------|------------|-------|---------|
| 1 | 0007077707 | HAFIDH NURROHMAN | SI Teknik Informatika | SEMESTER 8 | 55.8 | 100% |
| 2 | 2017200081 | Dell Widya Anggraini | SI Teknik Informatika | SEMESTER 2 | 7.2 | 100% |
| 3 | 2017200073 | Rahmatullah | SI Teknik Informatika | SEMESTER 3 | 44.1 | 100% |
| 4 | 2017200087 | Alvin Adhoni | SI Teknik Informatika | SEMESTER 2 | 51.2 | 100% |
| 5 | 2015200044 | Edward Widyawati | SI Teknik Informatika | SEMESTER 6 | 34.5 | 100% |
| 6 | 2014200108 | Alvin Firdaus | SI Teknik Informatika | SEMESTER 8 | 52.3 | 100% |
| 7 | 2014200089 | A. Aiman Wira | SI Teknik Informatika | SEMESTER 8 | 13.5 | 100% |
| 8 | 2014200080 | SA Sufatrina | SI Teknik Informatika | SEMESTER 8 | 100 | 100% |
| 9 | 2014200078 | Kiki Risa Lestari | SI Teknik Informatika | SEMESTER 8 | 50.8 | 100% |
| 10 | 2014200071 | Fahmi Idrus | SI Teknik Informatika | SEMESTER 8 | 100 | 100% |

Fig. 10. Evaluation System

2. Access Level for Participant

At the access level for Participants there are only three features provided, namely:

2.1 Form Login

In the login form feature for participant, it is used to limit registered participant who can log in to the system with the

conditions that the username and password entered must be correct, while the display for this login form is shown in Figure 11.

Fig. 11. Form Login

2.2 Main Page

The main page on the exam participant's access level will be displayed if the authentication of username and password is valid, while the main display of this access level is shown in Figure 12.

Fig. 12. Participant Main Page

2.3 Question List

In the question list feature there is a process of selecting the question categories to be done by the examinees, currently there are three choices of problem categories that have been provided, namely Islamic Education (PAI), Citizenship Education (PKN), and Indonesian (Bahasa). The appearance of this feature is shown in Fig. 13 - 14.

Fig. 13. Choice of Problem Categories

Fig. 14. Work Problem Example

IV. CONCLUSION

Based on the implementation of the Generalized Vector Space Model method in the automatic assessment of online essay exams, it can be concluded that the implementation of the Generalized Vector Space Model method on automatic assessment of online essay exams can be done easily and the system is fairly fast.

Based on the results of the percentage calculation between manual and automatic values of one of the examinees, it was found that the participants' answers were different from the keywords having an average percentage value of 79% while the examinees who had the exact answer to the keyword had an average percentage of 100%. In addition, the automatic scoring system with the GVSM weighting method and the cosine similarity calculation method have the results of the accuracy of the assessment with an average of 66% of the 70 answers processed.

Based on the explanation of the Generalized Vector Space Model method for automatic assessment of online essay exams, suggestions for the next development are to add automatic correction features to word spelling errors, in order to reduce the user's error. As well as the addition of algorithms that can determine the similarity of the meaning of the word, so that the accuracy of the assessment does not see the similarity of the sentence.

V. ACKNOWLEDGMENT

The author would like to thank all those who helped complete this paper.

VI. REFERENCES

- [1] Amin, F. (2012). Information Retrieval System with Vector Space Model Method. *Jurnal Sistem Informasi Bisnis*. Fakultas Teknologi Informasi, Universitas Stikubank, Semarang.
- [2] Wibowo, et. al. (2012). Generalized Vector Space Model Implementation Using WordNet.. Universitas Kristen Petra. Jurusan Teknik Informatika.
- [3] Fitri, R., & Asyikin, A. N. (2015). Automatic Essay Exam Assessment Application Using the Cosine Similarity Method. *Jurnal Poros Teknik*.
- [4] Bunyamin, H., & Negara, C. P. (2008). CATA Information Retrieval (IR) Application with the Generalized Vector Space Model Method.
- [5] Mujilawati, S. (2016). Pre-Processing Text Mining Pada Data Twiter. Seminar Nasional Teknologi Informasi dan Komunikasi, Program Studi Teknik Informatika, Fakultas Teknik, Universitas Islam Lamongan
- [6] Ariantini, d. (2016). Measurement of the Similarity of Indonesian Text Documents Using the Cosine Similarity Method. *E-Journal Teknik Informatika*.
- [7] Pardede, dkk. (2013). Implementasi Metode Generalized Verctor Space Model Pada Aplikasi Infromation Retrieval. Institut Teknologi Nasional Bandung. Jurusan Teknik Informatika.
- [8] Hidayatullah, & Kawistara. (2014). Pemograman Web. Bandung: Informatika.
- [9] Nugroho. (2004). Php dan Mysql dengan Editor Dreamweaver Mx. Cosine Similarity. *Jurnal Poros Teknik*.