IMPLEMENTATION OF THE COSINE SIMILARITY METHOD USING THESAURUS IN THE SEARCH APPLICATION OF ENGLISH AL-QURAN TRANSLATION

UNDERGRADUATE THESIS

By:
AHMAD SIROJUN NUHA
NIM. 16650093

JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2020

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

IMPLEMENTATION OF THE COSINE SIMILARITY METHOD AND THESAURUS IN THE SEARCH APPLICATION OF ENGLISH AL-QURAN TRANSLATION

UNDERGRADUATE THESIS

An Undergraduare Thesis is submitted to
Fakultas Sains dan Teknologi,
Universitas Islam Negeri Maulana Malik Ibrahim Malang
for the Requirements for degree of Bachelor of Computer (S.Kom)

By:
AHMAD SIROJUN NUHA
NIM. 16650093

JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG
2020

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

APPROVAL SHEET

IMPLEMENTATION OF THE COSINE SIMILARITY METHOD USING THESAURUS IN THE SEARCH APPLICATION OF ENGLISH AL-QURAN TRANSLATION

UNDERGRADUATE THESIS

By:
AHMAD SIROJUN NUHA
NIM. 16650093

Have been reviewed and approved for examination
Date: December 6, 2020

Supervisor I	Supervisor 2
Prof. Dr. Suhartono, M.Kom	A'la Syauqi, M.Kom
NIP. 19680519 200312 1001	NIP. 19771201 200801 007

Acknowledged,
Head of Jurusan Teknik Informatika

Dr. Cahyo Crysdian
NIP. 19740424 200901 008

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

RATIFICATION SHEET

IMPLEMENTATION OF THE COSINE SIMILARITY METHOD USING THESAURUS IN THE SEARCH APPLICATION OF ENGLISH AL-QURAN TRANSLATION

UNDERGRADUATE THESIS

By:
AHMAD SIROJUN NUHA
NIM. 16650093

Has been conducted in the present of the board of undergraduate thesis examiner and stated accepted as the one of requirements to obtain the title Bachelor of Computer (S.Kom)

Board of Examiner:			**Signature**
Chairperson	:	Dr. Cahyo Crysdian NIP. 19740424200901008	( )
Head	:	Khadijah F. H. Holle, M.Kom NIDT.19900626201608012007	( )
Secretary	:	Prof. Dr. Suhartono, M.Kom NIP. 196805192003121001	( )
Member	:	A'la Syauqi, M.Kom NIP. 19771201200801007	( )

Assesed and Approved by:
Head of Jurusan Teknik Informatika

Dr. Cahyo Crysdian
NIP. 19740424 200901 008

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

LETTER OF STATEMENT

The undersigned below:

Name : Ahmad Sirojun Nuha

NIM : 16650093

Department : Teknik Informatika

Faculty : Sains dan Teknologi

Undergraduate Thesis : Implementation of the Cosine Similarity Method using Thesaurus in the Search Application of English Quran Translation.

State that an thesis above is my original work, not taking over data, except for the quotation and statement whose resources are acknowledged on the references. Any shortcoming on this present work, therefore, are entirely my own responsibility.

Moreover, this work is not plagiarism result and if it is found that this statement is false, my academic record will attest to invalidation and I will be responsible for that as well.

Malang, December 5th 2020

Signed by,

Ahmad Sirojun Nuha

NIM. 16650093

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

MOTTO PAGE

"Don't underestimate the slightest kindness"

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

PRESENT PAGE

This thesis is dedicated to Father Samsul Hadi, Mother. Nurul Hidayah, Brother Aan, Brother Aim, Grandma Siti Badriyah who have supported, encouraged, and prayed for and helped complete this thesis.

To the thesis supervisor, Prof. Dr. Suhartono, M.Kom and Mr. A'la Syauqi M.Kom who have provided valuable experience and knowledge. Dear my friend, friends of organizations, CSSMoRA, HTQ, the fellow in arms who have accompanied me all this time. Dear friends of Informatics Engineering of 2016, friends who have supported this thesis is complete.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

FOREWORD

All praise belongs only to Allah SWT for his mercy and blessings to the author in order to complete this thesis at the Department of Informatics Engineering, State Islamic University of Maulana Malik Ibrahim Malang. Prayers and greetings to the Prophet Muhammad who became Uswatun Hasanah for his people. Alhamdulillah. The author has completed the thesis entitled "Implementation of the Cosine Similarity Method and the Thesaurus in the Search Application of English Quran Translation".

The thesis is intended to fulfill one of the requirements in completing the study of the Department of Informatics Engineering as a form and participation of the author in developing and actualizing the knowledge that the author has acquired while in college.

There are no words that the author can say other than just an expression of deep gratitude to all those who have helped writing this thesis, either directly or indirectly. Therefore, the authors would like to thank:

1. Both of My Beloved Parents Samsul Hadi, Nurul Hidayah, who always provide support both materially and spiritually.
2. Dr. Cahyo Crysdian, as chairman of Informatics Engineering Department, UIN Maulana Malik Ibrahim Malang.
3. Prof. Dr. Suhartono, M.Kom as the first supervisor who has been patiently guiding the author, providing advice, suggestions and directions so that the author is not only able to complete the thesis but also can take a lot of lessons.
4. A'la Syauqi, M.Kom as the second supervisor who has been careful to help correct the deficiencies of this thesis.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

5. Mr and Mrs Lecturers of Informatics Engineering at the State Islamic University of Maulana Malik Ibrahim Malang who have provided their knowledge to the author for 4 years.
6. Informatics Engineering Students of 2016 State Islamic University Maulana Malik Ibrahim Malang.

The author is fully aware that in this world nothing is perfect. Likewise the writing of this thesis, which is not free from flaws and mistakes. Therefore, with all sincerity and humbly, the author really expects suggestions and criticism that are constructive for the sake of completing this thesis.

The last with all forms of shortcomings and mistakes, the author hopes that with his mercy and permission, I hope this thesis is useful for author in particular and for the parties concerned.

Malang, December 6th 2020

Author

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

TABLE OF CONTENTS

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

LIST OF TABLES

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

LIST OF FIGURES

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

ABSTRAK

Nuha, Ahmad Sirojun (2020). **Penerapan Metode *Cosine Similarity* menggunakan Tesaurus dalam Aplikasi Pencarian Terjemah Alquran Bahasa Inggris**. Skripsi. Jurusan Teknik Informatika. Fakultas Sains dan Teknologi. Universitas Islam Negeri Maulana Malik Ibrahim Malang.
Pembimbing : (I) Prof. Dr. Suhartono. M,Kom
(II) A'la Syauqi, M.Kom

Kata Kunci : Pencarian Terjemah Alquran, *Cosine Similarity, Query Expansion.*

Mengetahui Terjemah Alquran adalah kunci untuk mengungkapkan semua mutiara kebijaksanaan dan ajaran yang terkandung dalam kitab suci Alquran. Tanpa kunci ini, kita tidak dapat mencapai kesempurnaan mutiara makna yang terkandung dalam Alquran. Dengan memahami terjemah Alquran kita akan bisa belajar untuk memahami Tafsir Alquran. Sebagai Muslim, kita tidak hanya harus pandai membacanya tetapi kita juga perlu mempelajari Alquran sehingga kita bisa mendapatkan berkah dari Alquran dengan sempurna. Text Similarity menentukan seberapa dekat dua teks atau dokumen dalam leksikal atau semantik. Salah satu metode Text Similarity adalah Cosine Similarity yang mengukur sudut cosinus antara vektor, yaitu vektor term terjemahan dan vektor term query. Hasil berupa angka antara 0-1, nilai yang lebih tinggi menghasilkan kecocokan dokumen terbaik. Text similarity melihat kesamaan antara 2 dokumen yang mana semakin banyak kata yang sama maka nilai similarity juga tinggi, padahal ada kata yang berbeda akan tetapi memiliki kesamaan dalam makna, maka Text Similarity saja tidak cukup. Oleh karena itu perlu ditambahkan metode untuk mendapatkan persamaan kata (sinonim) apalagi penelitian ini memiliki objek Terjemah Alquran yang begitu kental dengan kemiripan arti kata. Salah satu metode untuk mendapatkannya yaitu dengan *Query Expansion* (Perluasan Query). Tujuan penelitian ini adalah untuk mengetahui nilai presisi, *recall*, *f-measure* dan akurasi penerapan metode *Cosine Similarity* tanpa menggunakan *Query Expansion* dan dengan menggunakan *Query Expansion*. Uji coba yang telah dilakukan memberikan kesimpulan bahwa metode *Cosine Similarity* memiliki nilai uji evaluasi yang tinggi dalam aplikasi pencarian terjemah Al-Quran Bahasa Inggris yakni dengan nilai presisi sebesar 84,52%, *recall* sebesar 68,25%, bobot *f-measure* 72,94% dan akurasi sebesar 98,62%, akan tetapi dengan menggunakan *Query Expansion* bisa memberikan nilai lebih tinggi, yakni bisa menghasilkan nilai presisi sebesar 91,64%, *recall* sebesar 88,40%, bobot *f-measure* 88,89% dan akurasi sebesar 99,37%.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

ABSTRACT

Nuha, Ahmad Sirojun (2020). **Implementation of the Cosine Similarity Method uses a Thesaurus in the English Quran Translation Search Application.** Undergraduate Thesis. Informatics Engineering Departement. Faculty of Science and Technology. State Islamic University of Maulana Malik Ibrahim Malang.
Advisor : (I) Prof. Dr. Suhartono. M,Kom
(II) A'la Syauqi, M.Kom

Keywords: Al-Quran Translation Search, Cosine Similarity, Query Expansion.

Knowing the translation of the Al-Quran is the key to revealing all the pearls of wisdom and teachings contained in the holy Al-Quran. Without this key, we cannot reach the perfection of meaning contained in the Qur'an. By understanding the translation of the Al-Quran we will be able to learn to understand the interpretation of the Al-Quran. As Muslims, we not only have to be good at reading it but we also need to study the Al-Quran so that we can get the blessings of the Quran perfectly. Text Similarity determines how close two texts or documents are in lexical or semantic. One of the Text Similarity methods is Cosine Similarity, which measures the cosine angle between vectors, namely the translation term vector and the query term vector. The result is a number between 0-1, the higher value is the best document match. Text similarity sees the similarity between 2 documents where the more words are the same, the similarity value is also high, even though there are different words but have similarities in meaning, then Text Similarity is not enough. Therefore, it is necessary to add a method to obtain synonyms, especially since this research has the object of translating the Al-Quran which is so near with the similarity of the word meaning. One of the methods to get it is by Query Expansion. The purpose of this study was to determine the value of precision, recall, f-measure and accuracy of the application of the Cosine Similarity method without using Query Expansion and by using Query Expansion. The experiments that have been carried out have concluded that the Cosine Similarity method has a high evaluation test value in the English translation search application, namely with precision value of 84.52%, recall of 68.25%, f-measure weight of 72.94%. and an accuracy of 98.62%, but using Query Expansion can provide a higher value, which can produce precision value of 91.64%, recall of 88.40%, f-measure weight of 88.89% and accuracy of 99.37%.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

ملخص

النهى ، أحمد سيراج (٢٠٢٠). **تنفيذ طريقة كوسين سملاريتي باستخدام قاموس المرادفات في تطبيق بحث ترجمة القران الانجليزية** بحث جامعية. قسم هندسة المعلوماتية. كلية العلوم والتكنولوجيا. الجامعة الإسلامية الحكومية مولانا مالك إبراهيم مالانج.

المسشرف : (I) الأستاذ الدكتور سوهارتونو الماجستر

(II) اعلى شوقي الماجستر

الكلمات الدالة: بحث ترجمة القرآن ، كوسين سملاريتي ، توسع الاستعلام (Query Expansion).

إن معرفة ترجمة القرآن هي مفتاح الكشف عن كل لآلئ الحكمة والتعاليم الواردة في القرآن الكريم. بدون هذا المفتاح ، لا يمكننا الوصول إلى كمال المعنى الموجود في القرآن. من خلال فهم ترجمة القرآن ، سنتمكن من تعلم فهم تفسير القرآن. كمسلمين ، لا يجب أن نكون جيدين في قراءته فحسب ، بل نحتاج أيضًا إلى دراسة القرآن حتى نتمكن من الحصول على بركات القرآن الكاملة . يحدد تشابه النص مدى قرب نصين أو وثيقتين في المعجم أو الدلالات. إحدى طرق تشابه النص هي كوسين سملارتي ، والتي تقيس زاوية كوسين بين المتجهات ، أي متجه مصطلح الترجمة ومتجه مصطلح الاستعلام. والنتيجة هي رقم بين ٠ – ١ ، وكلما زادت القيمة أفضل تطابق وثيقة. يرى تشابه النص التشابه بين وثيقتين حيث كلما كانت الكلمات متشابهة ، تكون قيمة التشابه عالية أيضًا ، بالرغم من وجود كلمات مختلفة ولكن لها أوجه تشابه في المعنى ، فإن تشابه النص ليس كافيًا. لذلك لا بد من إضافة طريقة للحصول على المرادفات ، خاصة وأن الهدف من هذا البحث هو ترجمة مصحف كثيف مع تشابه معاني الكلمة. إحدى طرق الحصول عليها هي من توسيع الاستعلام (Query Expansion). كان الغرض من هذه الدراسة هو تحديد قيمة الدقة والاسترجاع والقياس والدقة لتطبيق طريقة كوسين سملارتي بدون استخدام توسيع الاستعلام وباستخدام توسيع الاستعلام. نتجت التجربة التي تم إجراؤها أن طريقة كوسين سملارتي لها قيمة اختبار تقييم عالية في تطبيق بحث الترجمة الإنجليزية ، وهي precision تبلغ ٨٤,٥٢٪ ، recall ٦٨,٢٥٪ ، f-measure ٧٢و٤٩ ٪. accuracy تبلغ ٩٨,٦٢٪ ، ولكن استخدام توسيع الاستعلام يمكن أن يوفر قيمة أعلى الذي يمكن أن ينتج قيمة precision تبلغ ٩١,٦٤٪ ، recall ٨٨,٤٠٪ ، f-measure ٨٨,٨٩٪ accuracy ٩٩,٣٧٪.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

CHAPTER I

INTRODUCTION

1.1 Research Background

Understanding the Al-Quran is the main key to reveal all the pearls of wisdom and teaching contained in the Holy Qur'an. Without this key, we cannot achieve the perfection of pearls of meaning contained in the Al-Quran.

It is unfortunate if Muslims in their daily lives only read the Al-Quran and then read it with a melodious voice but do not imprint on their hearts because of the lack of understanding of the meanings contained in the verses of the Al-Quran. They do not get a part of the Al-Quran's glory except only because of the joy of listening to Al-Quran and also just to take blessings from their reading without knowing the purpose of the verse they are listening to.

It is undeniable or not, what has been said by the Prophet Muhammad PBUH perhaps is what is happening today, as this hadith.

وَنَشْنًا يَنْشَنُونَ يَتَّخِذُونَ الْقُرْآنَ مَزَامِيرَ وَسَفْكَ الدَّمِ

Meaning :

"the generation that grows up by making the Quran as a flute and shed blood"
(Musnad Ahmad : 22845)

So this is where the importance of understanding the Al-Quran that we need to learn so that we can get perfect glory. We sometimes forget that the most important *barakah* of the Al-Quran is to analyze and understand it, then make it a guide in daily life, do its

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

teachings, do its instructions, and stay away from its prohibitions. Allah said in Surah Saad verse 29.

كِتَابٌ أَنزَلْنَاهُ إِلَيْكَ مُبَارَكٌ لِيَدَّبَّرُوا آيَاتِهِ وَلِيَتَذَكَّرَ أُولُو الأَلْبَابِ

Meaning :

"[This is] a blessed Book which We have revealed to you, [O Muhammad], that they might reflect upon its verses and those who understanding would be reminded.". (QS. Saad :29).

As Muslims, we not only have to be good at reading Al-Quran but we also need to study it so we can get blessings from the Al-Quran perfectly. So this is where the importance of understanding the Al-Quran we need to learn so that we can get perfect glory. The Al-Quran is a universal way of life that is always appropriate with conditions wherever and whenever, therefore it is necessary for us to continue to understand the Al-Quran so that we can get the wisdom contained in it and make it easy to practice it in life.

Lack of misunderstanding in understanding the Al-Quran will certainly cause polemics. The polemic that arose because of the misunderstanding of the meaning of the Al-Quran was so great. A few years ago Indonesia was shocked by the polemics due to misunderstanding of the meaning of the Al-Quran, starting with Al-Maidah verse 51 polemic which dragged so many parties who argued irrespectively of personal interests or truth. Then continued with the case of the legality of LGBT argued by a number of political elites who want to legalize the LGBT law. And no less polemical that from Indonesia began to be independent until now, namely the argument about the caliphate. This polemic makes Islam split up and difficult to unite.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

The international world is also rocked by the polemic of the misunderstanding of Al-Quran. The most fatal polemic till now is about jihad, as revealed by Warraq (1995: 15) the problem is rooted in Islamic teachings, namely about jihad. So with this reason makes him why he decided to leave Islam, then turned into a harsh critic of Islam itself. As it proves the statement above, actions for actions of violence in the name of jihad are rife in various parts of the world: from the Khawarij and Assassin sects which allowed the killing of political opponents in the period of al-Mamun and al-Mu'tasim, to the violence by the splinter group around the 1980s which led to the killing of President Anwar Sadat. Likewise the Bali Bombing on October 12, 2002, till a series of terrorist acts of bombing in the homeland that continued rife to this day (Nuzul: 2019). With the various polemics that occur above, it shall be for us Muslims to begin to learn the meaning of the content of Al-Quran so that it can be used as a guide and not easy to washed his understanding.

To learn the Al-Quran, of course we need references. Al-Qur'an Tafseer references have been widely published starting from Arabic such as Tafseer Al-Jalalain's, Tafseer Ibn Kathir, Tafseer Al-Qurthubi of the classical genre and Tafseer by Zaghlul An-Najar which is a modern Tafseer with a scientific interpretation approach. Indonesian Tafseer such as Tafseer Al-Mishbah by Quraish Shihab. There is also a Linguistic Miracle Tafseer by Nouman Ali Khan that reviews the Al-Quran in terms of linguistics. Reported from kompas.com there are estimated to be one-third of the world's population or around two billion people use English. That status makes English in the position of the global lingua franca. Therefore the use of the English language as

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

objects is addressed for the international world because of the increasing globalization so that peoples are more able to exchange information, especially in English. The first step before understanding the Tafseer is by knowing the translation of the Al-Quran first

The development of science and technology on various sides from the economy, government, transportation, and also education. Even many religious teachings have been adopted in the form of applications, making it easier for followers to study religion anywhere at any time. Some religious support applications include digital Al-Quran, prayer schedule, digital prayer beads, translation of the Al-Quran, murottal audio of the Al-Quran and also Tafseer of the Al-Quran. The translation applications available are still limited in features as quran.com which only has the main features of searching the Surah's name and translation from choosing verse numbers. Of course, with features like this that will be difficult to find a translation if we do not know the verse number. There is also helloquran.com which has more features with a translation search but still cannot detect the synonym of the word. Therefore needs to be a translation of the Al-Quran application that has a search feature also with the synonym for the word used, making it easier for those who already know the translation but do not know what verse or even what surah that verse.

From the explanation described above, IT-based translation of the Al-Quran application is needed which makes it easier for Muslims to learn the Al-Quran. Translating the Al-Quran is a language transition from Arabic to another language, which of course in one Arabic word can have several meanings or synonyms so that

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

there is a similarity between the words used in the translation of the Al-Quran. By knowing the essence of a translation we can look for a complete translation of a verse because of the similarities in the words used. Text Similarity determines how close two texts or documents are in lexical or semantic. One of the text similarity methods is Cosine Similarity, which measures the cosine angle between vectors, namely the query term vector and the translation term vector. A result is a number between 0-1, the higher value is the best document match. Because the translation is a semantic-based document (meaning content), it is not enough just to use Cosine Similarity to find similarities in meaning, in this study the authors applied a thesaurus so that the results obtained could match both in words and meanings. A comparison between the search process using the query expansion involving a thesaurus, and without query expansion shows that the thesaurus can improve the quality of search results (Rasyidi et al., 2013). The explanation above is the basis for the author to propose the title *Implementation of the Cosine Similarity Method and the Thesaurus in the Search Application for the English Al-Quran Translation*.

1.2 Research Question

According to the research background, the following research questions are raised:

1. How much is the accuracy, precision, recall, and f-measure of the Cosine Similarity method with thesaurus and without it in searching an English Al-Quran translation?
2. How is the comparison of Cosine Similarity with thesaurus and without it in searching an English Al-Quran translation?

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

1.3 Research Objective

The objectives of this study are as follows:

1. To measure the accuracy, precision, recall, and f-measure of Cosine Similarity method with thesaurus and without it in searching an English Al-Quran translation.
2. To compare the use of Cosine Similarity method with thesaurus and without it in searching an English Al-Quran translation.

1.4 Hypothesis

1. The accuracy of the translation search application using Cosine Similarity with the Thesaurus approach will increase if the query terms and documents in one meaning are more and more.
2. The use of the thesaurus is potentially more accurate because it does not only look at documents lexically but also semantically.

1.5 Research Benefit

The benefits of this research are seen from several sides as follows:

1. As a reference for developing a search program for English Al-Quran Translation.
2. As a supporter of the Al-Quran study institution to facilitate the search for Al-Quran verses and translations.

1.6 Research Scope

In order for this research does not deviate according to the author's wishes and so that the problem under study does not spread, it is necessary to define the research scope. The research scope of this study are:

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

1. Using the English Al-Quran Translation Juz 30 downloaded from http://www.qurandatabase.org.
2. Selection of the thesaurus from the dataset is carried out manually by the author.
3. Generation of a thesaurus is carried out by the system automatically.
4. The stemming uses the Porter method and is only able to detect regular verb similarities.
5. Thesaurus data is obtained from the WordNet application.
6. Stopword data is obtained from https://countwordsfree.com.

1.7 Writing Systematic

This research is organized into a report which consists of several chapters as follows:

1. **Chapter I Introduction:** This first chapter discusses research background, research questions, research scope, research objectives, hypotheses, research benefits and writing systematics.
2. **Chapter II Literature Review:** This chapter contains the literature review and the previous research about Cosine Similarity and the use of the thesaurus.
3. **Chapter III Research Methodology:** This chapter contains methods, system design that applies the use of Cosine Similarity and thesaurus which is applied to the search application for the English Al-Quran translation.
4. **Chapter IV Experiment and Discussion:** This chapter contains the results of data processing carried out in this study in translation search with text similarity measurement using Cosine Similarity.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

5. **Chapter V Conclusion and Suggestion:** This chapter contains the conclusions of this research, as well as useful suggestions for next this research.
6. **References:** This chapter contains references related to this research.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

CHAPTER II

LITERATURE REVIEW

2.1 Related Research

Research carried out by Rahman et al. (2010) with Sahih Bukhari's Hadith as an object with query expansion using a thesaurus and using Cosine Similarity as similarity detection, more text documents were retrieved from the test collection, resulting in decreased precision. Although the precision yield decreased, the detailed results showed that the effectiveness level increased slightly. The results of experiments carried out on the Hadith text collection prove that the thesaurus increases the effectiveness of retrieval in this domain even though the percentage increase is very small. They revealed that the thesaurus would also help users to retrieve documents relevant to the entered query by searching for synonymous terms.

Further research was carried out by Rasyidi et al. (2013) developed the use of an automatic thesaurus and manual validation. A comparison between the search process with a query expansion involving a thesaurus, and without a query expansion shows that the thesaurus can improve the quality of search results. The precision in the 10 tests increased by 34%, meaning that the thesaurus was able to help produce good search results. This means that query expansion has succeeded in finding more relevant documents. Since query expansion depends on the thesaurus being used, the process for developing the thesaurus is a critical point. A thesaurus built manually will have excellent accuracy.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Alodadi and Vandana (2015) Developed a program to detect similar health posts so that repetition of questions does not occur. This study uses a model that finds similarities in forum posts using the Cosine Similarity metric over Term Frequency-Inverse Document Frequency (TF-IDF). The results of this model using Cosine Similarity and TF-IDF are more improved than the existing models.

Not only useful with the Cosine Similarity, Khatri et al. (2017) proposed an application that filters and optimizes search queries using a thesaurus and Bayesian filters. Applications developed have achieved an average of 94.9% in terms of relevance and 90.6% of responses in terms of usage.

The survey was carried out on applications applying the thesaurus by Shchitov et al. (2017) revealed that the thesaurus showed extraordinary results in completing the task of text analysis, text classification and sentiment, text summarization.

2.2 Theoretical Basis

2.2.1 Information Retrieval System

The definition of information retrieval according to the Oxford dictionary is "A system for tracking and recovering specific information from stored data". The term information retrieval was introduced by Mooers. Mooers (1951) defined information retrieval as a process of discovery or recovery with connection to stored information that includes intellectual aspects of the information description and the specification itself for search and includes any systems, techniques or machines used to perform the operation. Over time, the definition of information retrieval has grown widely, as defined in ISO 2382/1 which defines information retrieval as actions, methods and

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

procedures to recover stored data, then provide feedback information about the subject as needed. These actions include text indexing, inquiry analysis. Information includes text, tables, images, speech and video.

There are two kinds of processes carried out by the information retrieval system (Bunyamin, 2008), namely carrying out the preprocessing step of the data stored in the database, then applying the method to calculate the closeness between documents in the database that has been processed with a query from the user. In preprocessing, steps are carried out starting from changing lowercase letters, deleting punctuation marks, deleting unimportant words, then making the verb into a root word according to the dictionary form, and finally weighting each word from the terms in the database. Likewise, the same process is carried out in the query, and the value is returned, then an approach is carried out to calculate the similarity value to the document, then a list of documents that is similar to the query is generated. Feedback given to the user is a document that is relevant to the query by the system.

2.2.2 Term Frequency – Inverse Document Frequency

Information retrieval system is inseparable from user queries which have similarities with the system collection. The collection consists of documents of various lengths with different term contents. To start the processing document, term weighting is required. The term can be a phrase, word or another indexed unit in a document that can be used to find out the meaning of the document. The terms in each document have different levels of importance, therefore an indicator is given for each term to recognize them, namely the term weight (Zafikri, 2010).

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

The TF-IDF algorithm is one of the algorithms used to give weight to a term in a document. There are two concepts in this algorithm, namely the term frequency or frequency of appearance of terms in a document and inverse document frequency that contains certain terms (Aziz, 2015). TF or term frequency assigns the term weight to a document based on the number of times that term appears in a document. This meaning that the greater number of appearance/frequency of a term in a document, the greater its weight in the document. Then, Inverse Document Frequency (IDF), which is intended to reduce the dominance of terms that often appear in various documents. Terms that often appear in documents are considered common terms so that these terms are not considered very important in value and therefore must be weakened. This weighting is raised because terms that rarely appear in documents have very valuable importance. IDF is the inverse of documents frequency containing certain terms.

2.2.3 Cosine Similarity

Cosine similarity is an algorithm used to measure document similarity. Cosine Similarity calculates the angular cosine value between two vectors, the document vectors and query vector. Cosine similarity calculates angular cosine values from queries and other documents. The returned value will indicate the degree of similarity of the query to the document. Because the Cosine Similarity value is based on the angle between two vectors, the resulting value ranges from 0 to 1. A value of 0 indicates that the two documents are not similar at all and a value of 1 indicates that the query and document are completely identical. Whereas the value between the range 0-1, if it is greater, the similarity is also greater and if the value is smaller, the degree of similarity

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

between the query and the document is small (Fikri, 2018).

2.2.4 Query Expansion

Query expansion is a technique used to modify queries that aim to fulfill and complete information needs. Generally, modifications made are in the form of adding terms to keywords as a complement, although actually they also include adjustments for weighting and deletion of keyword terms (Selberg, 1997).

In the Query Expansion, there are three types based on the method that is carried out, namely (Rahayuni, 2011):

1. Manual Query Expansion (MQE)

   In this method, users modify the query manually. That is, the user enters the additional keyword himself for the expansion of the query.

2. Automatic Query Expansion (AQE)

   In this method, the user has no role in generating the Query Expansion. The system will automatically perform the Query Expansion according to the developed program. According to previous research, there are several techniques used in Automatic Query Expansion, namely: Global Analysis consists of basic principles, namely utilizing the context of one term to determine its similarity to other terms, Local Analysis using documents that the user has retrieved to get new queries, and Local Context Analysis which is a combination of Global Analysis and Local Analysis techniques.

3. Interactive Query Expansion (IQE)

   In this method, the user and the system work together in interaction to carry out

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

the Query Expansion process.

2.2.5 System Evaluation Test

System evaluation is carried out to measure how much the ability of the system to perform its performance, in this case the application of Cosine Similarity using a thesaurus. System evaluation is also carried out to determine the performance comparison in measuring the similarity of the query with the English Al-Quran translation document between the Cosine Similarity method that uses a thesaurus and doesn't.

Information Retrieval System retrieves a set of documents as feedback from user queries. There are two categories of documents produced by the information retrieval system related to query processing, namely relevant documents (documents that are relevant to the query) and retrieved documents (documents provided by the system). To measure the relevance of documents, the Lancaster accuracy formula is used. Meanwhile, the general measure used to measure the quality of retrieval data is a combination of precision and recall. Precision is the proportion of a dataset that is obtained and is relevant to the input. Recall is the proportion of all relevant results in the collection system including the results obtained whether relevant or not. F-measure is the harmonic weight of precision and recall which is the reciprocal measure between both.

2.3 Theoretical Framework

According to the explanation in the previous sub-chapter, it can be concluded that the various uses of the thesaurus have improved the search results in Text Similarity.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

The process of searching for documents by querying to get information is called the Information Retrieval System. In general, the theoretical framework for the Information Retrieval System is described in Figure 2.1 (Zakaria and Soyusiawati, 2018):

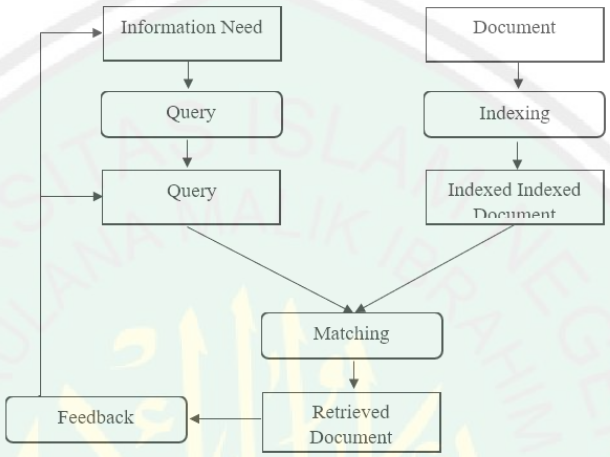

Figure 2.1 Information Retrieval System Theoretical Framework

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

CHAPTER III

RESEARCH METHODOLOGY

3.1 Research Data Source

The source of the research documents in this research is the Translation of the English Al-Quran by the International Ulama Yusuf Ali, which consists of the verses juz 30. Each verse in juz 30 is considered 1 document, so the document in this research consists of 564 documents. The document files were downloaded from https://www.qurandatabase.org. The file is in the form of a MySQL dataset which is then uploaded to this translation search application database. The downloaded file is a .sql file that can be uploaded to the database. This interpretation data of Yusuf Ali's work will be the object of this research.

3.1 System Design

The system design is the flow of the application, at this step the author describes the system design in the form of a block diagram. Figure 3.1 is a system design of the Quran Translation search application. The application form is website based.

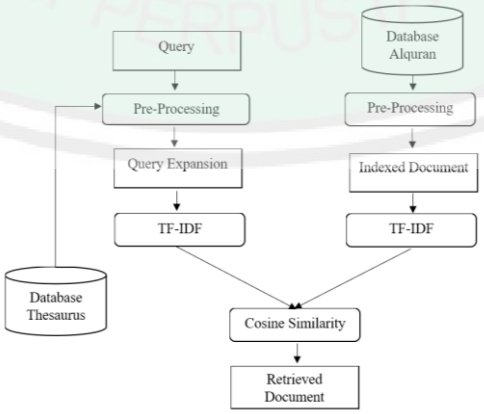

Figure 3.1 Proposed Information Retrieval System Design

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.2.1 Database Repository

The Database Repository is a place to store data used in this study in the form of Yusuf Ali's Al-Quran translation, obtained from the source http://www.qurandatabase.org. Making a database using one of the well-known DBMS, MySQL. The Al-Quran translation database consists of several fields that include entities such as letter id, verse id, translation id, letter name, verse number and the contents of the translation. The thesaurus database is obtained from the WordNet application site by taking 2-3 words in the thesaurus dictionary that have the most similarity in meaning to word clusters. The thesaurus database consists of entities such as word id, thesaurus id, word group and thesaurus group.

3.2.2 Pre-processing

At the preprocessing step, 5 processes are carried out, namely as follows:

3.2.2.1 Whitespace Intensivity

The first step is to remove irrelevant characters (whitespace intensivity) from the text document. Deletion is carried out on all letters, symbols, punctuation marks, operators and others that are not A-Z, a-z, 0-9. The flow is shown in Figure 3.22.

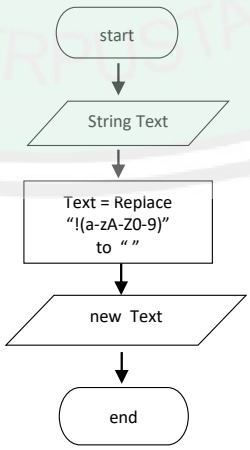

Figure 3.2 Whitespace Intensivity Flowchart

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

And here are examples of document data and queries that will be processed (Table 3.1) and the results after Whitespace Intensivity is carried out (Table 3.2).

Table 3.1 Examples of Documents and Query

ID	Isi Dokumen / Query
D1	I am submitting myself and entering into the refuge.
D2	We are humbling ourselves to Allah and going into his protection.
D3	We are looking for refuge from Allah in daybreak.
Q	Say, I seek refuge in the Lord of daybreak.

Table 3.2 Results of Document and Query Whitespace Intensivity

ID	Isi Dokumen / Query
D1	I am submitting myself and entering into the refuge
D2	We are humbling ourselves to Allah and going into his protection
D3	We are looking for refuge from Allah in daybreak
Q	Say I seek refuge in the Lord of daybreak

3.2.2.2 Case Folding

In the second step, the text is changed to lowercase letters so that only characters in the form of letters or numbers will be processed further. The flow is as illustrated in Figure 3.3 with results can be seen in Table 3.4.

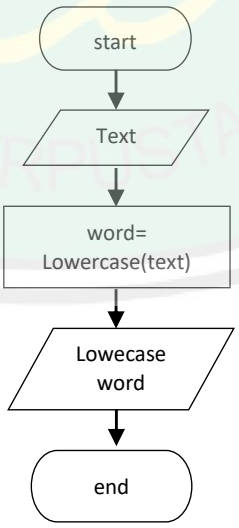

Figure 3.3 Case Folding Flowchart

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 3.3 Results of Documents and Query Case Folding

ID	Isi Dokumen / Query
D1	i am submitting myself and entering into the refuge
D2	we are humbling ourselves to allah and going into his protection
D3	we are looking for refuge from Allah in daybreak
Q	refuge in the lord of daybreak

3.2.2.3 Tokenizing

Text that has been changed to lowercase letters will be cut or separated by each word separated by space using a process as shown in Figure 3.4 and will become a separate word unit as can be seen in table 3.4.

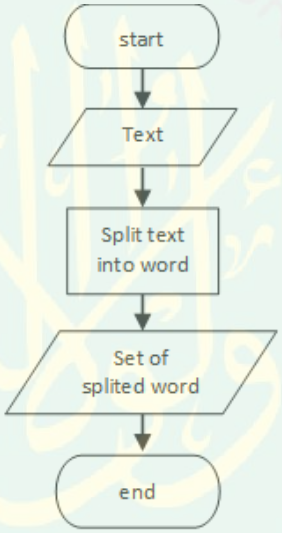

Figure 3.4 Tokenizing Flowchart

Table 3.4 Results of Document and Query Tokenizing

No.	Term	No.	Term
1	i	18	into
2	am	19	his
3	submitting	20	protection
4	myself	21	we
5	and	22	are
6	entering	23	seeking
7	into	24	refuge
8	the	25	from
9	refuge	26	allah
10	we	27	in

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 3.5 Results of Document and Query Tokenizing (rmore)

11	are	28	daybreak
12	humbling	29	refuge
13	ourselves	30	in
14	to	31	the
15	allah	32	lord
16	and	33	of
17	going	34	daybreak

3.2.2.4 Filtering

In a document, of course, there are some words that have no meaning or do not play a role in understanding the sentence. These words are in the form of conjunctions, prepositions of pronouns and others. These words can be entered into a stopword or stoplist, namely the words that will be filtered and deleted later to reduce the volume of words to be processed. So that what remains is only words that play a role and are important to be processed further. In this research, stopword data is in the form of excel dataset which is downloaded from http: //www.countwordsfree which is then entered into the program database. The filtering flow can be seen in Figure 3.5 and in this example it produces the remaining words as shown in table 3.5.

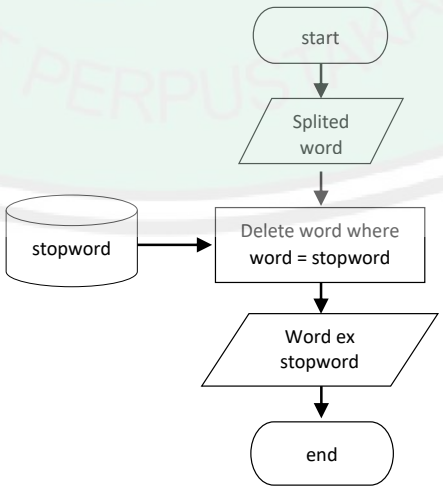

Figure 3.5 Filtering Flowchart

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 3.6 Results of Documents and Query Filtering

No.	Term
1	submitting
2	myself
3	entering
4	refuge
5	humbling
6	ourselves
7	allah
8	going

No.	Term
9	protection
10	seeking
11	refuge
12	allah
13	daybreak
14	refuge
15	lord
16	daybreak

3.2.2.5 Stemming

Basically, every word comes from a root word or dictionary form. Basic words can receive additions to have different meanings. Additional differences have different meanings. Additional words can be at the beginning of the word (affix) or at the end of the word (suffix). Two or more sentences with different words can have the same or similar understanding. Therefore, in this study the addition of prepositions in each word needs to be removed so that it becomes a dictionary form where the basic words that compose a sentence can be found so that it is easy to find similarities with the input query. However, in this study we have not applied stemming to irregular verbs. More specifically, the flow of stemming in this study is described in Figure 3.6.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

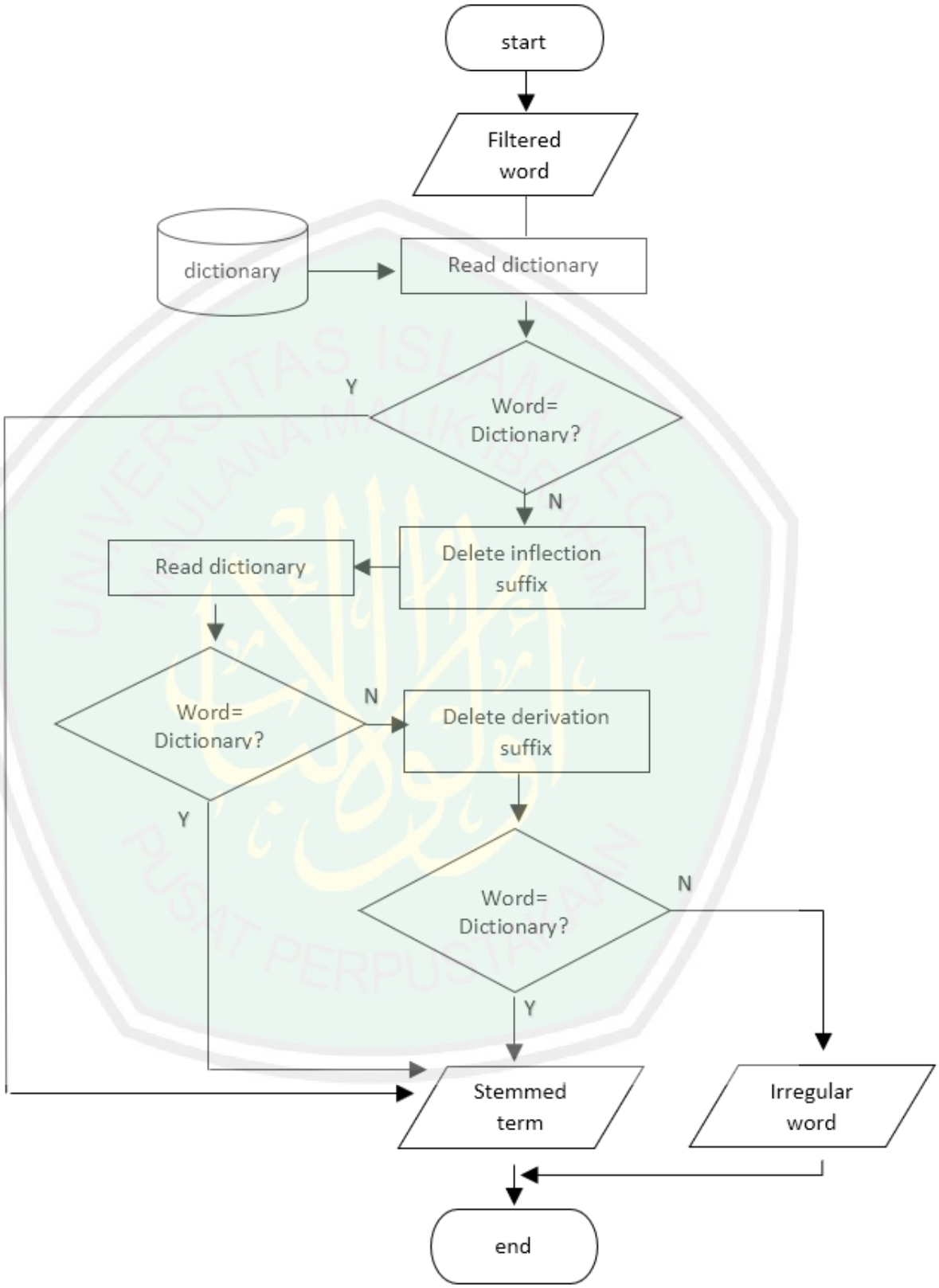

Figure 3.6 Stemming Flowchart

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

The list of documents that have been processed with the stemming process will return to the basic word or dictionary form as shown in table 3.6 so that the query and subsequent documents that will be processed are changed as in table 3.7.

Table 3.7 Results of Documents and Query Stemming

No.	Term	No.	Term
1	submit	9	protect
2	myself	10	seek
3	enter	11	refuge
4	refuge	12	allah
5	humble	13	daybreak
6	ourselves	14	refuge
7	allah	15	lord
8	go	16	daybreak

Table 3.8 Results of Documents and Query Preprocessing

ID	Isi Dokumen / Query
D1	submit myself enter refuge
D2	humble ourselves allah go protect
D3	look refuge allah daybreak
Q	refuge lord daybreak

3.2.3 Query Expansion

The query entered by the user has not been able to detect the similarity of words that have the same meaning or synonym, therefore it is necessary to add a method to get it, especially since this research has so many objects of Al-Quran translation with similar meanings of words. One method to get it is by Query Expansion. Query expansion is carried out by looking for a thesaurus or synonyms in word clusters in the thesaurus database. The thesaurus data is obtained manually from the WordNet application by taking 1-3 words which have a similarity level according to the WordNet sequence. After obtaining the thesaurus then entered into the database. In this study,

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

the thesaurus is searched based on the root word or dictionary word-stemming results. The system will look for the stemming result word in the thesaurus database, if the word is found and there is a thesaurus for that word then the thesaurus is taken directly (whatever is available in the database) in addition to the query with poses as shown in Figure 3.7, so the query used will change with the addition of thesaurus as in table 3.8.

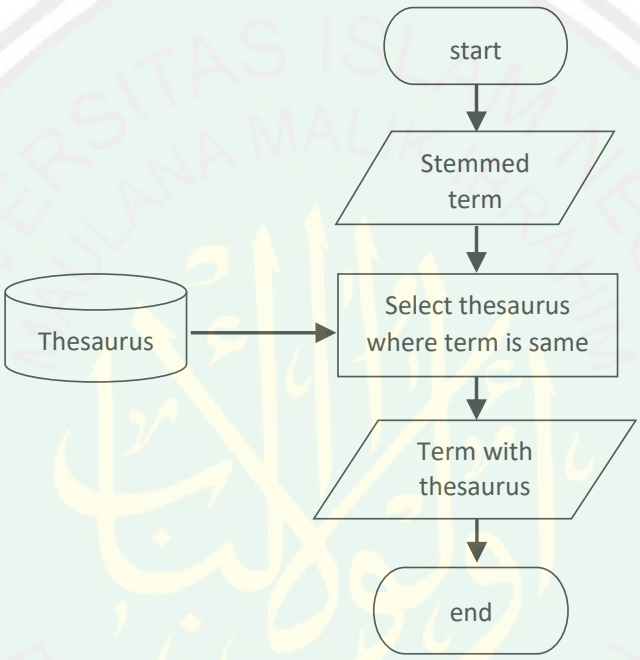

Figure 3.7 Query Expansion Flowchart

Table 3.9 Results of Query Expansion on Query

ID	Query
Q	refuge lord daybreak protect Allah morning

3.2.4 Term Frequency-Inverse Document Frequency (TF-IDF)

The next step is weighting the words on the query and the Al-Quran Translation dataset so that they can be processed numerically. The weighting used is the Term Frequency-Inverse Document Frequency (TF-IDF) method. TF is the frequency of

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

appearance of terms in a document. The TF value can be seen from the following equation (Manning, 2008).

$$TF(t,d) = f_{t,d} \tag{3.1}$$

where TF_{td} is *t* terms appearance frequency on *d* document.

Term Frequency as above suffers from a critical problem: all terms are considered as important as their relevance rating on a query. Even if certain terms have little or no distinguishing power in determining relevance, a mechanism is needed to weaken the effect of terms that appear too frequently in collections to be meaningful for determining relevance. The immediate idea is to reduce the high-frequency term weight scale, which is defined as the total number of appearance of terms in the collection (TF). The idea is to reduce the weight of the term by a factor that grows with the frequency of collection. The IDF value is obtained from the following equation (Manning, 2008).

$$IDF_t = \log(N/Df_t) + 1 \tag{3.2}$$

where *N* is the total number of documents in the collection and Df_t is a document containing the *t* term.

The weighting process uses the TF-IDF method, namely by multiplying TF to IDF so that the equation looks like this.

$$TF - IDF = TF_{t,d} \times IDF_t \tag{3.3}$$

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

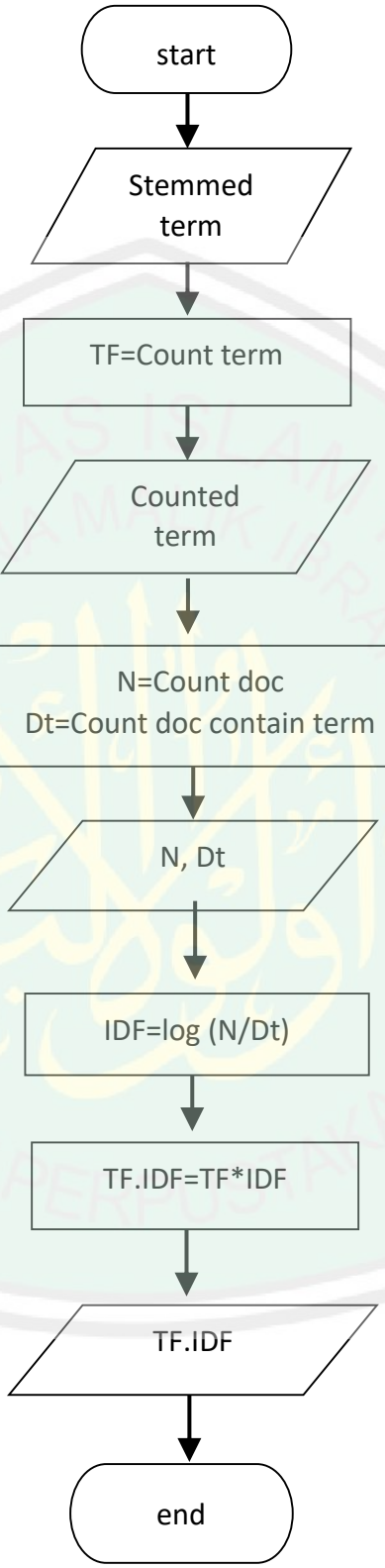

Figure 3.8 TF-IDF Weighting Flowchart

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 3.10 Results of TF-IDF Calculation

No.	Term	Q	D1(TF)	D2(TF)	D3(TF)	Df	N/Df	Log (N/Df)+1
1	Refuge	1	1		1	3	1,000	1
2	Lord	1				1	3,000	1,477
3	Daybreak	1			1	2	1,500	1,176
4	Protect	1		1		2	1,500	1,176
5	Allah	1			1	2	1,500	1,176
6	Morning	1				1	3,000	1,477
7	Submit		1			1	3,000	1,477
8	Myself		1			1	3,000	1,477
9	Enter		1			1	3,000	1,477
10	Humble			1		1	3,000	1,477
11	Ourselves			1		1	3,000	1,477
12	Allah			1		1	3,000	1,477
13	Go			1		1	3	1,477

Table 3.11 Query and Documents Weight Values using TF-IDF

No.	WQ	WD1	WD2	WD3
1.	1	1	0	1
2.	1,477	0	0	0
3.	1,176	0	0	1,176
4.	1,176	0	1,176	0
5.	1,176	0	0	1,176
6.	1,477	0	0	0
7.	0	1,477	0	0
8.	0	1,477	0	0
9.	0	1,477	0	0
10.	0	0	1,477	0
11.	0	0	1,477	0
12.	0	0	1,477	0
13.	0	0	1,477	0

3.2.5 Cosine Similarity

Weighted documents and queries will be calculated the equation with Cosine Similarity. The principle is to calculate the cosine angle between vectors, namely the

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

query vector and the Al-Quran Translation dataset so that the similarity value is obtained as represented in Figure 3.9.

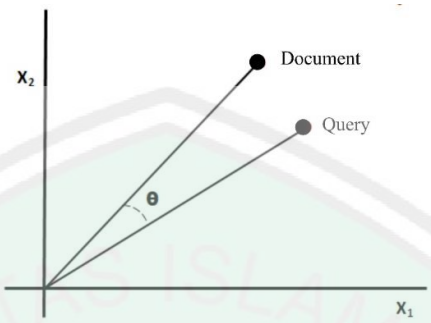

Figure 3.9 Cosine Similarity Representation

The closer result to the number 1, it shows the document is getting similar and vice versa, the closer to the number 0, the document is less similar. The following is an overview of the Cosine Similarity process (Figure 3.10) and an example of the calculation results from calculations on a document that has been given a table weight (3.12). The Cosine Similarity mathematical equation, namely

$$sim(x, y) = \frac{x \times y}{\|x\| \|y\|} \tag{3.4}$$

Where ||x|| is the eucledian norm of the element x $(x = x_1, x_2 \dots x_p)$ which is defined as $x = \sqrt{x_1^2 + x_2^2 \dots x_p^2}$ (Jiawei dkk, 2012) which in this study is applied to equation

$$sim(q, d) = \frac{w_q \times w_d}{\sqrt{\sum w_q^2 \times \sum w_d^2}} \tag{3.5}$$

Note :

$sim(q, d)$ = the similarity of query and document

w_q = query weight

w_d = document weight

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

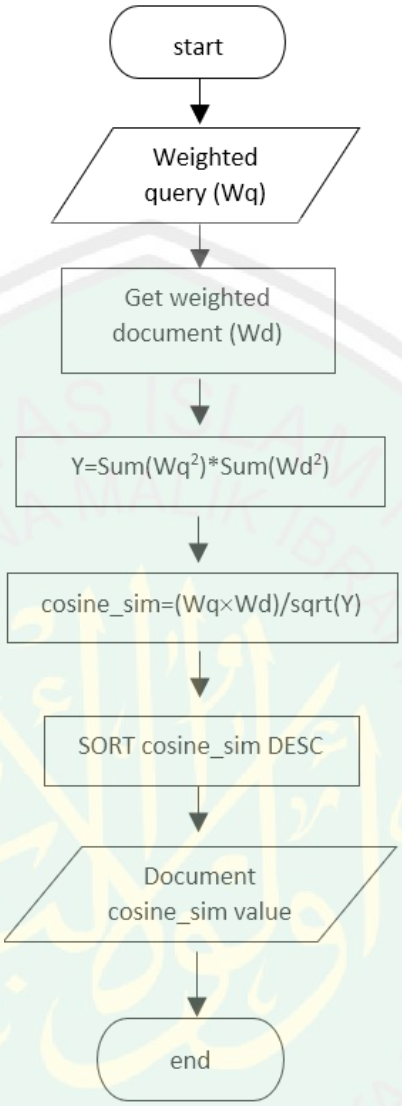

Figure 3.10 Cosine Similarity Flowchart

Table 3.12 The final Cosine Similarity calculation

	W	D1	D2	D3
$w_q \times w_d$		1	1,38	3,76
$\sqrt{\sum w_q^2 \times \sum w_d^2}$	3,084	2,74	3,17	1,94
Cosine Similarity		0,11	0,14	0,62

So it can be concluded that the third document has the highest similarity to the query, with a similarity value of 0.62.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.3 System Testing

To determine the success rate of application built in this study, testing is needed. The test is to calculate the accuracy of application output to find out how accurate the application that has been built is in creating the output desired by the user. In addition to accuracy, it is also necessary to know the precision and recall values. Precision is defined as the percentage of information accuracy that the user wants with the answer output given by the system. Recall is defined as the percentage of the system's success in retrieving relevant information.

Testing is carried out by performing a search using an input query using the Cosine Similarity method, testing the query is carried out starting from the 1st to the n-th test which will then be evaluated for its success rate. From each test result, accuracy, precision and recall values can be obtained to evaluate the system's capabilities. The formulas for accuracy, precision and recall are as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \times 100\% \quad (3.6)$$

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\% \quad (3.7)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (3.8)$$

Where value of TP, FP, FN and TN as shown in the following table :

Table 3.13 Value of TP, FP, FN and TN

Output Cosine Similarity	**Relevant**	**Irrelevant**	**Total**
Found	TP (hits)	FP (noise)	TP+FP
Not Found	FN (misses)	TN (rejected)	FN+TN
Total	TP+FN	FP+TN	TP+FP+FN+TN

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Value of TP is the number of documents that are relevant and have been successfully found by the system. The value of FP is the number of documents that are not relevant but are found by the system to be a relevant value. The value of FN is the number of documents that are relevant but not found by the system. The value of TN is the number of documents that are not relevant and are considered irrelevant by the system so that the data is not found (in the relevant data collection). To improve the evaluation system, F-measure is used. F-measure is a weight that combines the precision and recall values which are a reciprocal measure between the two. F-measure formula is as follows:

$$F\text{-}measure = 2\times\frac{precision\times recall}{precision+recall}\times 100\% \quad (3.9)$$

After evaluating the system's ability for each test from the 1st to the n-th test by obtaining the accuracy, precision, recall and f-measure values, then evaluate the Cosine Similarity method whether using a thesaurus can produce more relevant output or can't. To find out the comparison of the Cosine Similarity method between those using a thesaurus and don't, the results of the test evaluation were calculated the total and the average value was sought for each method. After obtaining the average value, the value is visualized and compared. By comparing the two methods, it can be determined which method can produce more relevant outputs.

The test used 30 keywords in the English translation of Juz Amma in the Cosine Similarity method with and without using a thesaurus. Then will be evaluated the ability of the system for each test. After obtaining the evaluation value for each test, then the application of Cosine Similarity uses a thesaurus and don't is compared. The

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

test output reference in this study is based on one of the Al-Quran translation search software, helloquran.com. This software can provide translated information feedback on user queries.

3.4 System Implementation

The system that has been designed by the researcher is built with the PHP programming language. The following is an implementation of each design that has been designed as the system design sub-chapter.

3.4.1 Pre-processing

The first process is pre-processing. Pre-processing consists of 5 steps, namely Whitespace Intensivity, Case Folding, Tokenizing, Filtering and Stemming. The input to this process is the research data document and user queries. Here is the source code for pre-processing.

```
//whitespace intensivity

foreach ($punctuation as $i => $value) {
  $query = str_replace($punctuation[$i], " ", $query);
}

//case folding
$query = strtolower(trim($query));

//tokenizing
$each_term = explode(" ", $query);

//iltering
if(in_array($each_term[$i], $stopWord)){
  unset($each_term[$i]);
}

//stemming
function cekKamus2($query, $term_id);
```

Figure 3.11 Pre-processing Souce Code

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.4.2 Query Expansion

The next process is Query Expansion. The input of this process is a user query that has gone through the pre-processing process. At this step, it produces user query output that is added with the appropriate thesaurus and is available in the database. The following is the source code for the Query Expansion process.

```
for($i=0; $i<$lengthq; $i++){
   $sql = mysqli_query($con, "SELECT * FROM thesaurus where term =
'$arrquery[$i]'");
          while ($result = mysqli_fetch_array($sql)) {
                    $sinonim = $result['thesaurus'];
                    $asinonim = explode(",", $sinonim);
                    $length = count($asinonim);
                    for($j=0; $j<$length; $j++){
                              array_push($arrquery, $asinonim[$j]);}
                    }
          }
}
$arrquery = implode(" ", $arrquery);
```

Figure 3.12 Query Expansion Source code

3.4.3 Term Frequency-Inverse Document Frequency (TF-IDF)

The next process is to give weight to each term using the TF-IDF algorithm. Input is in the form of a term that is already in a dictionary form and produces the output in the form of weight for each word. The following is the source code for weighting using the TF-IDF algorithm.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

```
$resBobot = mysqli_query($con,"SELECT * FROM indexing");
                $num_rows = mysqli_num_rows($resBobot);
while($rowbobot = mysqli_fetch_array($resBobot)) {
                $term = $rowbobot['term'];
                $tf = $rowbobot['tf'];
                $term_id = $rowbobot['term_id'];
                $idf = $rowbobot['idf'];
                $resNTerm = mysqli_query($con,"SELECT Count(*) as
N FROM indexing WHERE term = '$term'");
                $rowNTerm = mysqli_fetch_array($resNTerm);
                $NTerm = $rowNTerm['N'];
                $idf=log10($n/$NTerm)+1;
                $w = $tf * $idf;  } }
```

Figure 3.13 TF-IDF Source Code

3.4.4 Cosine Similarity

Cosine similarity calculates the similarity between two vectors, namely the query vector and the document. Queries and documents for which the similarity is calculated have a weighted value for each term, so that the length of the vector can be obtained by adding all the terms weigh of a document and then looking for the similarities. The output is the similarity value between the query and the document with a 0-1 value range. The following is the source code for Cosine similarity as in equation 3.5 in system design.

```
$resBobot = mysqli_query($con,"SELECT * FROM indexing");
        $num_rows = mysqli_num_rows($resBobot);
while($rowbobot = $doc_id = $rowdoc_id['doc_id'];
        $longdoc_id = $rowdoc_id['long_vector'];
        $resTerm = mysqli_query($con,"SELECT * FROM indexing WHERE doc_id =
$doc_id AND status=0");
        while ($rowTerm = mysqli_fetch_array($resTerm)) {
                for ($i=0; $i<count($aquery); $i++) {
                if ($rowTerm['term'] == $aquery[$i]) {
                $dotproduct = $dotproduct + $rowTerm['wegh'] * $query_weigh[$i];
                }  }  } if ($dotproduct > 0) {
$sim = $dotproduct / ($longQ* longdoc_id)}
```

Figure 3.14 Cosine SImilarity Source Code

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

3.5 Database Implementation

The database is needed to store data to be processed in this study. The database used is MySQL with the phpMyAdmin tool. The data is in the form of a translation of the Koran, stopword, and thesaurus which are then implemented into each table as shown in the following Entity Relationship Diagram (ERD).

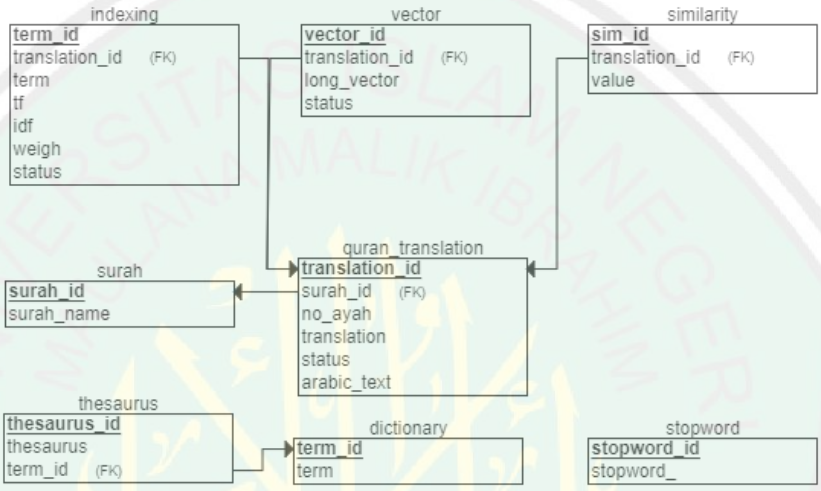

Figure 3.15 Built Application ERD

3.6 Interface Implementation

Interface design is a form of graphic display that will be seen by the user. In this study, an English translation of the Al-Quran application was built on a web-based. The following is an application interface design that is built.

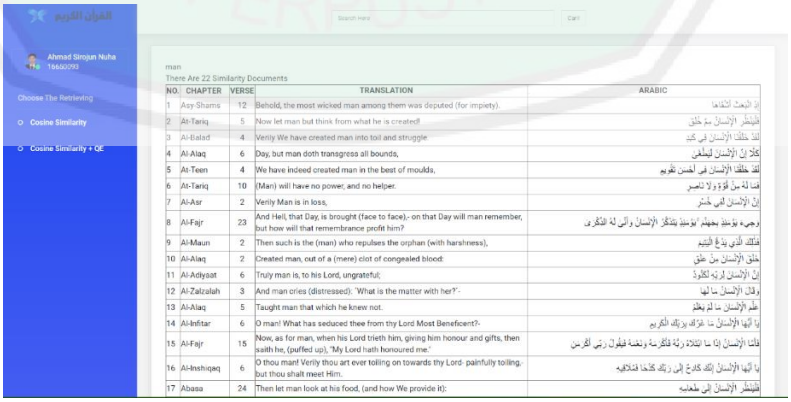

Figure 3.16 Built Application Interface

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

The interface consists of several parts. At the top, there is a search box that will be filled in by the user with the desired query and next to it is a search button to perform the search process. On the side there is a choice of application of the method used, the user will choose one as the method used. And the main part, namely in the middle there is a table that displays the search results performed by the user. Above the table, there is information on the query entered by the user and the number of documents retrieved by the system.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

CHAPTER IV

EXPERIMENT AND DISCUSSION

In this chapter, the experiment was carried out as based on the steps described in Chapter 3, which aims to compare the Cosine Similarity method which uses query expansion and those that do not. By carrying out this step, it will produce experiment results that will be processed to answer the Research Questions in Chapter 1. The results of this experiment will be calculated the value of precision, recall, F-Measure and calculated the accuracy value of each method. After obtaining the results, it will be known which method is more effective in finding an English translation of the Al-Quran.

4.1 Experiment Setup

The Experimental Setup carried out are as follows:

1. Prepare data of the Al-Quran translation that has been obtained from the site http://www.qurandatabase.org. After the data is obtained, it is stored in the phpMyAdmin database.
2. Pre-processing documents that are already in the database will be carried out so that they are easy to process. Non-relevant characters such as punctuation, numbers will be removed (Whitespace Intensivity). Changing the letter into a lower case (Case Folding). Splitting documents composed of sentences into one-word pieces or terms (tokenizing). After breaking down into each term, if there is a word which is a stopword it will be deleted (Stopword Removal). Because a word can be a root word and an affixed word, the word / term that is affixed will be

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

returned to its basic form (stemming). The results obtained from preprocessing will be processed into the next process.

3. The use of Query Expansion is carried out by looking for terms in the database which have the same meaning and the closest similarity based on the WordNet thesaurus data.
4. Give weight to each term using the TF-IDF method which consists of TF (the number of a term (x) in a document) and IDF (inverse of the number of documents containing a term (x)). Furthermore, the TF and IDF values are multiplied to get the weighed term value.
5. Calculate the vector length value of each document by adding up all the weight terms of a document.
6. Testing the system in searching for documents for the translation of the Al-Quran by entering 30 queries. In each query, the pre-processing, weighting and vector length steps are also carried out. The query testing is performed twice, the first is not applying query expansion, the second is by implementing query expansion.
7. Calculating the similarity between the query documents and each Al-Quran translation document using the Cosine Similarity method.
8. Calculating the value of precision, recall, F-Measure and the calculating accuracy value of each experiment whether it uses query expansion or doesn't. The calculation is carried out based on the formula as stated in the sub-chapter of System Testing. As for the relevance of the system results compared to the search results of one of the websites for the translation of the English Al-Quran, helloquran.com and quranindex.info.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

9. Calculating the average value of each experiment and finally comparing the four values of each method so that it is known which method is more effective.

4.2 Experiment Results

The experiment began by processing the research object, namely the data for the English Al-Quran translation Juz 30 consisting of 564 documents. Furthermore, Pre-Processing is carried out on the translated document. At the tokenizing step, all documents counted 5,596 terms which were still mixed with stopwords. After filtering the number of terms becomes 2,118 which will then be given the weight and the vector which will be processed with a query later.

The next process is to prepare a query that will be used as an input system. There are 30 queries that have been prepared, starting from one-word to four-word queries. Also, each query is pre-processed as in the document. Because this study compares the use of a thesaurus to Cosine Similarity, the process of adding a thesaurus is carried out in the query. The following table 4.1 About Query Testing and Expansion:

Table 4.1 Queries and Recommendation for Expansion

No.	Query	Recommendation of Term	Query Expansion Result	Note
1	Day	-	Day	No added thesaurus
2	Disbeliever	unbeliever	disbelief unbelieve	1 thesaurus term added
3	Sin	wicked	sin wicked	1 thesaurus term added
4	Rest	Sleep	rest sleep	1 thesaurus term added
5	Glory	honor honour	glory honor honour	2 thesaurus terms added
6	Punishment	torment torture	punish torment torture	2 thesaurus terms added
7	Say	Speak	say speak	1 thesaurus term added
8	Bad deeds	evil effort	bad evil deed effort	2 thesaurus terms added
9	Allah swears	oath affirm	allah swear oath affirm	2 thesaurus terms added
10	Messenger from god	lord	messenger god lord	1 thesaurus term added
11	Night of decree	dark rescript	night dark decree rescript	2 thesaurus terms added

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

12	Allah taught man	instruct human mankind	allah taught instruct man human mankind	3 thesaurus terms added
13	See creation	look	see look creat	1 thesaurus term added
14	When quran recited	read	quran recite read	1 thesaurus term added
15	Wealth is not eternal	abundant unend	wealth abundant eternal unend	2 thesaurus terms added
16	Moon follows sun	pursue	moon follow pursue sun	1 thesaurus term added
17	Advice patience and kindness	recommend	advice recommend patience kind	1 thesaurus term added
18	Enter my heaven	heaven eden paradise	enter heaven eden paradise	3 thesaurus terms added
19	The greatest punishment	torment torture	great punish torment torture	2 thesaurus terms added
20	Give the book	throw record	give throw book record	2 thesaurus terms added
21	Prefer the present live	choose exist world	prefer choose present exist live world	3 thesaurus terms added
22	There is no power	strength	power strength	1 thesaurus term added
23	Hell fuel	blaze fire	hell blaze fire fuel	2 thesaurus terms added
24	Back to family	return	back return family	1 thesaurus term added
25	Do you know	Recognize	know recognize	2 thesaurus terms added
26	Refuse to believe	deny trust	refuse deny believe trust	2 thesaurus terms added
27	Righteous are in pleasure	pleasant joy happy	righteous pleasure pleasant joy happy	3 thesaurus terms added
28	Moses and pharaoh	-	moses pharaoh	No added thesaurus
29	Destruction	Woe	destruct woe	1 thesaurus term added
30	Warning for whoever fear	notice fright	warn notice fear fright	2 thesaurus terms added

The determination of the thesaurus is based on the wordNet application. Some words have no expansion because the WordNet application does not mention it or only mentions the meaning in the form of a sentence. In addition, some verbs have similarities with verbs with prepositions, such as come after, come in, etc. are not included, because the preposition itself is considered to be a stopword, so a verb with a preposition will have a very different meaning without the preposition. Before proceeding to the trial results, source data was prepared which would later be compared with the system results. The source data can be seen in table 4.2

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.2 Reference Data based on Queries

No.	*Query*	Data Acuan		Total
1.	Day	[78] An-Naba 11	[83] Al-Mutaffifeen 5	39
		[78] An-Naba 17	[83] Al-Mutaffifeen 10	
		[78] An-Naba 18	[83] Al-Mutaffifeen 11	
		[78] An-Naba 39	[83] Al-Mutaffifeen 34	
		[78] An-Naba 40	[85] Al-Burooj 2	
		[78] An-Naba 8	[86] At-Tariq 9	
		[79] An-Nazi'at 46	[88] Al-Gashiyah 2	
		[79] An-Nazi'at 8	[88] Al-Gashiyah 8	
		[79] An-Nazi'at 35	[89] Al-Fajr 23	
		[79] An-Nazi'at 6	[89] Al-Fajr 25	
		[80] Abasa 40	[90] Al-Balad 14	
		[80] Abasa 38	[91] Asy-Shams3	
		[80] Abasa 37	[92] Al-Layl 2	
		[80] Abasa 34	[99] Al-Zalzalah 4	
		[81] At-Takweer 15	[99] Al-Zalzalah 6	
		[82] Al-Infitar 19	[100] Al-Adiyaat 11	
		[82] Al-Infitar 17	[101] Al-Qoriah 4	
		[82] Al-Infitar 18	[102] At-Takathur 8	
		[82] Al-Infitar 15	[108] Al-Kawthar 1	
		[83] Al-Mutaffifeen 15		
2.	Disbeliver	[78] An-Naba 40	[86] At-Tariq 17	7
		[79] An-Nazi'at 17	[88] Al-Gashiyah 23	
		[80] Abasa 42	[109] Al-Kafiroon 1	
		[83] Al-Mutaffifeen36		
3.	Sin	[81] At-Takweer 9	[91] Asy-Shams14	6
		[83] Al-Mutaffifeen7	[96] Al-Alaq 16	
		[83] Al-Mutaffifeen12	[82] Al-Infitar 14	
4.	Rest	[78] An-Naba 9	[89] Al-Fajr 27	2
5.	Punishment	[78] An-Naba 30	[85] Al-Burooj 10	8
		[78] An-Naba 40	[88] Al-Gashiyah 24	
		[79] An-Nazi'at 25	[89] Al-Fajr 25	
		[84] Al-Inshiqaq 24	[89] Al-Fajr 13	
6.	Glory	[80] Abasa 13	[89] Al-Fajr 15	5
		[85] Al-Burooj 21	[92] Al-Layl 2	
		[85] Al-Burooj 15		
7.	Say	[78] An-Naba 37	[89] Al-Fajr 27	17
		[78] An-Naba 38	[89] Al-Fajr 24	
		[78] An-Naba 40	[89] Al-Fajr 16	
		[79] An-Nazi'at 18	[89] Al-Fajr 15	
		[79] An-Nazi'at 24	[90] Al-Balad 6	
		[79] An-Nazi'at 43	[91] Asy-Shams 13	
		[79] An-Nazi'at 12	[99] Al-Zalzalah 3	
		[79] An-Nazi'at 10	[109] Al-Kafiroon 1	
		[83] Al-Mutaffifeen 17	[112] Al-Ikhlas 1	
		[83] Al-Mutaffifeen 32	[113] Al-Falaq 1	
		[83] Al-Mutaffifeen 13	[114] An-Naas 1	

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.3 Reference Data based on Queries (more)

8.	bad deeds	[78] An-Naba 40 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 9 [95] At-Teen 6 [99] Al-Zalzalah 8 [99] Al-Zalzalah 6	[101] Al-Qoriah 8 [101] Al-Qoriah 6 [103] Al-Asr 3 [107] Al-Maun 6 [113] Al-Falaq 2 [113] Al-Falaq 5 [113] Al-Falaq 3 [113] Al-Falaq 4 [114] An-Naas 4	14
9.	allah swears	[78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 25 [80] Abasa 9 [81] At-Takweer 15 [81] At-Takweer 29 [82] Al-Infitar 19 [83] Al-Mutaffifeen 28 [84] Al-Inshiqaq 16 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [85] Al-Burooj 3 [85] Al-Burooj 20 [85] Al-Burooj 1 [85] Al-Burooj 8 [85] Al-Burooj 9 [85] Al-Burooj 15 [87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24 [89] Al-Fajr 5 [89] Al-Fajr 2 [89] Al-Fajr 3 [89] Al-Fajr 1 [89] Al-Fajr 4 [89] Al-Fajr 11	[90] Al-Balad 1 [91] Asy-Shams 5 [91] Asy-Shams 6 [91] Asy-Shams 2 [91] Asy-Shams 15 [91] Asy-Shams 4 [91] Asy-Shams 13 [92] Al-Layl 1 [92] Al-Layl 5 [95] At-Teen 1 [95] At-Teen 8 [95] At-Teen 2 [96] Al-Alaq 14 [96] Al-Alaq 19 [96] Al-Alaq 12 [103] Al-Asr 1 [104] Al-Humazah [106] Al-Quraish 3 [106] Al-Quraish 2 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 2 [112] Al-Ikhlas 1 [113] Al-Falaq 1	46
10.	messenger from god	[78] An-Naba 36 [78] An-Naba 37 [79] An-Nazi'at 24 [79] An-Nazi'at 19 [79] An-Nazi'at 5 [79] An-Nazi'at 44 [79] An-Nazi'at 40 [80] Abasa 15 [81] At-Takweer 20 [81] At-Takweer 19 [81] At-Takweer 29 [82] Al-Infitar 6	[89] Al-Fajr 16 [89] Al-Fajr 22 [89] Al-Fajr 6 [89] Al-Fajr 28 [91] Asy-Shams13 [91] Asy-Shams14 [92] Al-Layl 20 [93] Ad-Duha 5 [93] Ad-Duha 3 [94] Al-Inshirah8 [96] Al-Alaq 1 [97] Al-Qadr 4	47

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.4 Reference Data based on Queries (more)

		[83] Al-Mutaffifeen 15 [83] Al-Mutaffifeen 6 [84] Al-Inshiqaq 15 [84] Al-Inshiqaq 2 [84] Al-Inshiqaq 5 [84] Al-Inshiqaq 6 [85] Al-Burooj 12 [87] Al-Alaa 1 [87] Al-Alaa 15 [89] Al-Fajr 15 [89] Al-Fajr 14 [89] Al-Fajr 13	[98] Al-Bayyinah 2 [98] Al-Bayyinah 8 [100] Al-Adiyaat 11 [105] Al-Feel 1 [106] Al-Quraish 3 [108] Al-Kawthar 2 [113] Al-Falaq 1 [114] An-Naas 3 [114] An-Naas 1 [79] An-Nazi'at 16 [78] An-Naba 39	
11.	night of decree	[78] An-Naba 10 [79] An-Nazi'at 29 [80] Abasa 41 [81] At-Takweer 17 [84] Al-Inshiqaq 17 [89] Al-Fajr 2 [89] Al-Fajr 4	[91] Asy-Shams4 [92] Al-Layl 1 [93] Ad-Duha 2 [97] Al-Qadr 1 [97] Al-Qadr 2 [97] Al-Qadr 3 [113] Al-Falaq 3	14
12.	allah taught man	[78] An-Naba 40 [78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 35 [80] Abasa 24 [80] Abasa 23 [80] Abasa 2 [80] Abasa 17 [80] Abasa 34 [81] At-Takweer 29 [81] At-Takweer 27 [82] Al-Infitar 6 [82] Al-Infitar 19 [83] Al-Mutaffifeen 6 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 6 [85] Al-Burooj 20 [85] Al-Burooj 8 [85] Al-Burooj 9 [85] Al-Burooj 15 [86] At-Tariq 5 [86] At-Tariq 10 [87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24 [89] Al-Fajr 15	[90] Al-Balad 4 [91] Asy-Shams15 [91] Asy-Shams13 [92] Al-Layl 5 [95] At-Teen 4 [95] At-Teen 6 [96] Al-Alaq 5 [96] Al-Alaq 4 [96] Al-Alaq 6 [96] Al-Alaq 2 [96] Al-Alaq 14 [96] Al-Alaq 19 [96] Al-Alaq 12 [99] Al-Zalzalah 6 [99] Al-Zalzalah 3 [100] Al-Adiyaat 6 [103] Al-Asr 2 [104] Al-Humazah6 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 2 [112] Al-Ikhlas 1 [113] Al-Falaq 1 [114] An-Naas 3 [114] An-Naas 2 [114] An-Naas 6 [114] An-Naas 1	47

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.5 Reference Data based on Queries (more)

13.	see creation	[78] An-Naba 8 [78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [80] Abasa 18 [80] Abasa 19 [82] Al-Infitar 7 [83] Al-Mutaffifeen 35 [80] Abasa 24 [84] Al-Inshiqaq 15 [86] At-Tariq 5 [86] At-Tariq 6 [86] At-Tariq 7 [88] Al-Gashiyah 17	[89] Al-Fajr 6 [90] Al-Balad 7 [92] Al-Layl 3 [95] At-Teen 4 [96] Al-Alaq 1 [96] Al-Alaq 2 [96] Al-Alaq 7 [102] At-Takathur 7 [102] At-Takathur 6 [105] Al-Feel 1 [110] An-Nasr 2 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8 [99] Al-Zalzalah 6	29
14.	when quran was recited	[81] At-Takweer 25 [83] Al-Mutaffifeen 13 [84] Al-Inshiqaq 21 [85] Al-Burooj 21	[87] Al-Alaa 6 [96] Al-Alaq 1 [96] Al-Alaq 3 [98] Al-Bayyinah 2	8
15.	wealth is not eternal	[78] An-Naba 23 [78] An-Naba 14 [89] Al-Fajr 20 [90] Al-Balad 6 [92] Al-Layl 18 [92] Al-Layl 11	[95] At-Teen 6 [100] Al-Adiyaat 8 [104] Al-Humazah 2 [108] Al-Kawthar 1 [111] Al-Lahab 2 [104] Al-Humazah3	12
16.	moon follows sun	[79] An-Nazi'at 7 [81] At-Takweer 1 [84] Al-Inshiqaq 18 [91] Asy-Shams 12	[91] Asy-Shams 1 [91] Asy-Shams 12 [91] Asy-Shams 4	7
17.	advice patience and kindness	[90] Al-Balad 17 [103] Al-Asr 3		2
18.	enter my heaven	[78] An-Naba 31 [79] An-Nazi'at 41 [81] At-Takweer 13 [81] At-Takweer 21 [82] Al-Infitar 15 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12 [87] Al-Alaa 12	[88] Al-Gashiyah 4 [89] Al-Fajr 30 [89] Al-Fajr 29 [98] Al-Bayyinah 8 [101] Al-Qoriah 7 [108] Al-Kawthar 1 [110] An-Nasr 2 [111] Al-Lahab 3	16
19.	the greatest punishment	[78] An-Naba 30 [78] An-Naba 2 [78] An-Naba 26 [78] An-Naba 40 [79] An-Nazi'at 25 [79] An-Nazi'at 20 [84] Al-Inshiqaq 24	[85] Al-Burooj 10 [85] Al-Burooj 11 [87] Al-Alaa 12 [88] Al-Gashiyah24 [89] Al-Fajr 25 [89] Al-Fajr 13	13

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.6 Reference Data based on Queries (more)

20.	give the book	[78] An-Naba 29 [80] Abasa 6 [80] Abasa 13 [82] Al-Infitar 11 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen 9 [83] Al-Mutaffifeen 20 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen 3 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 10	[84] Al-Inshiqaq 7 [84] Al-Inshiqaq 24 [86] At-Tariq 17 [89] Al-Fajr 15 [92] Al-Layl 12 [92] Al-Layl 5 [92] Al-Layl 18 [93] Ad-Duha 5 [98] Al-Bayyinah 5 [108] Al-Kawthar 1 [112] Al-Ikhlas 3	23
21.	prefer the present live	[78] An-Naba 39 [79] An-Nazi'at 38 [79] An-Nazi'at 10 [81] At-Takweer 27 [83] Al-Mutaffifeen 6	[87] Al-Alaa 16 [89] Al-Fajr 24 [101] Al-Qoriah7 [102] At-Takathur 1	9
22.	there is no power	[81] At-Takweer 20 [82] Al-Infitar 19 [86] At-Tariq 10	[90] Al-Fajr 5 [110] An-Nasr 1	5
23.	hell fuel	[78] An-Naba 21 [79] An-Nazi'at 36 [81] At-Takweer 6 [81] At-Takweer 12 [82] Al-Infitar 14 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12 [85] Al-Burooj 5 [85] Al-Burooj 10 [87] Al-Alaa 12	[88] Al-Gashiyah 4 [89] Al-Fajr 23 [90] Al-Balad 20 [92] Al-Layl 14 [98] Al-Bayyinah 6 [101] Al-Qoriah11 [101] Al-Qoriah9 [102] At-Takathur 6 [104] Al-Humazah6 [111] Al-Lahab 3	20
24.	back to family	[79] An-Nazi'at 12 [79] An-Nazi'at 22 [79] An-Nazi'at 44 [79] An-Nazi'at 10 [82] Al-Infitar 5 [84] Al-Inshiqaq 9 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 14 [83] Al-Mutaffifeen 31	[86] At-Tariq 11 [86] At-Tariq 8 [88] Al-Gashiyah 25 [89] Al-Fajr 28 [92] Al-Layl 19 [94] Al-Inshirah 3 [95] At-Teen 5 [96] Al-Alaq 8	16
25.	do you know	[78] An-Naba 4 [78] An-Naba 5 [80] Abasa 3 [82] Al-Infitar 5 [81] At-Takweer 14 [82] Al-Infitar 12 [82] Al-Infitar 17	[84] Al-Inshiqaq 23 [87] Al-Alaa 7 [90] Al-Balad 12 [96] Al-Alaq 5 [96] Al-Alaq 14 [97] Al-Qadr 2 [100] Al-Adiyaat 9	24

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.7 Reference Data based on Queries (more)

		[82] Al-Infitar 18 [82] Al-Infitar 5 [83] Al-Mutaffifeen 24 [83] Al-Mutaffifeen 8 [83] Al-Mutaffifeen 19	[101] Al-Qoriah10 [101] Al-Qoriah 3 [102] At-Takathur 3 [102] At-Takathur 4 [102] At-Takathur 5	
26.	refuse to believe	[78] An-Naba 28 [79] An-Nazi'at 21 [80] Abasa 42 [82] Al-Infitar 9 [83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 17 [84] Al-Inshiqaq 20 [84] Al-Inshiqaq 25	[91] Asy-Shams11 [91] Asy-Shams14 [95] At-Teen 7 [95] At-Teen 6 [96] Al-Alaq 13 [98] Al-Bayyinah 7 [107] Al-Maun 1 [107] Al-Maun 7	16
27.	righteous are in pleasure	[78] An-Naba 31 [78] An-Naba 38 [82] Al-Infitar 13 [83] Al-Mutaffifeen 22 [83] Al-Mutaffifeen 18 [83] Al-Mutaffifeen 24 [83] Al-Mutaffifeen 12 [84] Al-Inshiqaq 13	[84] Al-Inshiqaq 9 [84] Al-Inshiqaq 7 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 8 [90] Al-Balad 18 [95] At-Teen 6 [103] Al-Asr 3	16
28.	moses and pharaoh	[79] An-Nazi'at 15 [79] An-Nazi'at 21 [79] An-Nazi'at 17	[87] Al-Alaa 19 [89] Al-Fajr 10	5
29.	destruction	[78] An-Naba 40 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 10 [84] Al-Inshiqaq 11	[91] Asy-Shams14 [104] Al-Humazah 1 [107] Al-Maun 4	7
30.	warning for whoever fear	[78] An-Naba 40 [79] An-Nazi'at 45 [79] An-Nazi'at 26 [79] An-Nazi'at 19 [79] An-Nazi'at 8 [79] An-Nazi'at 9 [79] An-Nazi'at 40 [80] Abasa 9 [80] Abasa 11 [83] Al-Mutaffifeen 18	[87] Al-Alaa 10 [88] Al-Gashiyah 2 [91] Asy-Shams15 [92] Al-Layl 14 [92] Al-Layl 5 [92] Al-Layl 17 [96] Al-Alaq 12 [98] Al-Bayyinah 8 [106] Al-Quraish 4	19

The query is input into the system and displays the similarity results with each document in the database. In table 4.7 we can see the system results using the Cosine Similarity method only.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.8 Experiment Results of Cosine Similarity without Query Expansion

No.	Queries	Results of Cosine Similarity		Total
1.	Day	[78] An-Naba 11	[83] Al-Mutaffifeen 10	40
		[78] An-Naba 17	[83] Al-Mutaffifeen 11	
		[78] An-Naba 18	[83] Al-Mutaffifeen 34	
		[78] An-Naba 39	[83] Al-Mutaffifeen 6	
		[78] An-Naba 40	[85] Al-Burooj 2	
		[78] An-Naba 8	[86] At-Tariq 9	
		[79] An-Nazi'at 46	[88] Al-Gashiyah 2	
		[79] An-Nazi'at 8	[88] Al-Gashiyah 8	
		[79] An-Nazi'at 35	[89] Al-Fajr 1	
		[79] An-Nazi'at 6	[89] Al-Fajr 23	
		[80] Abasa 40	[89] Al-Fajr 25	
		[80] Abasa 38	[90] Al-Balad 14	
		[80] Abasa 37	[91] Asy-Shams3	
		[80] Abasa 34	[92] Al-Layl 2	
		[81] At-Takweer 15	[99] Al-Zalzalah 4	
		[82] Al-Infitar 19	[99] Al-Zalzalah 6	
		[82] Al-Infitar 17	[100] Al-Adiyaat 11	
		[82] Al-Infitar 18	[101] Al-Qoriah4	
		[82] Al-Infitar 15	[102] At-Takathur 8	
		[83] Al-Mutaffifeen 5	[108] Al-Kawthar 1	
2.	Disbeliver	[78] An-Naba 40	[83] Al-Mutaffifeen 30	8
		[79] An-Nazi'at 17	[86] At-Tariq 17	
		[80] Abasa 42	[88] Al-Gashiyah 23	
		[83] Al-Mutaffifeen36	[109] Al-Kafiroon 1	
3.	Sin	[81] At-Takweer 9	[91] Asy-Shams14	4
		[83] Al-Mutaffifeen12	[96] Al-Alaq 16	
4.	Rest	[78] An-Naba 9	[89] Al-Fajr 27	2
5.	Glory	[85] Al-Burooj 15	[92] Al-Layl 2	3
		[89] Al-Fajr 17		
6.	Punishment	[79] An-Nazi'at 25	[88] Al-Gashiyah 24	5
		[84] Al-Inshiqaq 24	[89] Al-Fajr 25	
		[85] Al-Burooj 10		
7.	Say	[78] An-Naba 40	[89] Al-Fajr 15	14
		[79] An-Nazi'at 12	[90] Al-Balad 6	
		[79] An-Nazi'at 10	[99] Al-Zalzalah 3	
		[79] An-Nazi'at 46	[109] Al-Kafiroon 1	
		[83] Al-Mutaffifeen 13	[112] Al-Ikhlas 1	
		[89] Al-Fajr 24	[113] Al-Falaq 1	
		[89] Al-Fajr 16	[114] An-Naas 1	
8.	bad deeds	[78] An-Naba 36	[85] Al-Burooj 11	17
		[78] An-Naba 40	[88] Al-Gashiyah 9	
		[78] An-Naba 8	[95] At-Teen 6	
		[81] At-Takweer 10	[99] Al-Zalzalah 6	
		[83] Al-Mutaffifeen 18	[101] Al-Qoriah 6	
		[84] Al-Inshiqaq 23	[101] Al-Qoriah 8	
		[84] Al-Inshiqaq 25	[103] Al-Asr 3	

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.9 Experiment Results of Cosine Similarity without Query Expansion (more)

		[84] Al-Inshiqaq 6 [81] At-Takweer 14	[107] Al-Maun 6	
9.	allah swears	[78] An-Naba 38 [79] An-Nazi'at 25 [79] An-Nazi'at 26 [80] Abasa 9 [81] At-Takweer 15 [81] At-Takweer 29 [82] Al-Infitar 19 [83] Al-Mutaffifeen 18 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [84] Al-Inshiqaq 16 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [85] Al-Burooj 1 [85] Al-Burooj 15 [85] Al-Burooj 20 [85] Al-Burooj 3 [85] Al-Burooj 8 [85] Al-Burooj 9 [87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24 [88] Al-Gashiyah 3 [89] Al-Fajr 1 [89] Al-Fajr 11 [89] Al-Fajr 2 [89] Al-Fajr 3	[89] Al-Fajr 4 [90] Al-Balad 1 [91] Asy-Shams 13 [91] Asy-Shams 15 [91] Asy-Shams 2 [91] Asy-Shams 4 [91] Asy-Shams 5 [91] Asy-Shams 6 [92] Al-Layl 1 [92] Al-Layl 5 [95] At-Teen 1 [95] At-Teen 2 [95] At-Teen 8 [96] Al-Alaq 12 [96] Al-Alaq 14 [96] Al-Alaq 19 [103] Al-Asr 1 [104] Al-Humazah6 [106] Al-Quraish 1 [106] Al-Quraish 2 [106] Al-Quraish 3 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 1 [112] Al-Ikhlas 2 [113] Al-Falaq 1	53
10.	messenger from god	[80] Abasa 15 [81] At-Takweer 19 [91] Asy-Shams 13 [96] Al-Alaq 12	[98] Al-Bayyinah 2 [98] Al-Bayyinah 5 [98] Al-Bayyinah 8 [114] An-Naas 3	8
11.	night of decree	[78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [84] Al-Inshiqaq 15 [85] Al-Burooj 13 [85] Al-Burooj 15 [86] At-Tariq 5	[89] Al-Fajr 6 [96] Al-Alaq 7 [99] Al-Zalzalah 6 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8 [102] At-Takathur 6 [102] At-Takathur 7 [110] An-Nasr 2	16
12.	allah taught man	[78] An-Naba 38 [78] An-Naba 40 [79] An-Nazi'at 25 [79] An-Nazi'at 26 [79] An-Nazi'at 35 [80] Abasa 17	[88] Al-Gashiyah 3 [89] Al-Fajr 11 [89] Al-Fajr 15 [90] Al-Balad 4 [91] Asy-Shams 13 [91] Asy-Shams 15	55

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.10 Experiment Results of Cosine Similarity without Query Expansion (more)

		[80] Abasa 2 [80] Abasa 23 [80] Abasa 34 [80] Abasa 9 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 6 [85] Al-Burooj 15 [85] Al-Burooj 20 [85] Al-Burooj 8 [85] Al-Burooj 9 [86] At-Tariq 10 [81] At-Takweer 29 [82] Al-Infitar 19 [82] Al-Infitar 6 [83] Al-Mutaffifeen 18 [86] At-Tariq 5 [87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24	[92] Al-Layl 5 [95] At-Teen 4 [95] At-Teen 8 [96] Al-Alaq 12 [96] Al-Alaq 5 [96] Al-Alaq 6 [99] Al-Zalzalah 3 [103] Al-Asr 2 [104] Al-Humazah6 [106] Al-Quraish 1 [106] Al-Quraish 2 [106] Al-Quraish 3 [110] An-Nasr 1 [110] An-Nasr 2 [96] Al-Alaq 14 [96] Al-Alaq 19 [96] Al-Alaq 2 [96] Al-Alaq 4 [112] Al-Ikhlas 1 [112] Al-Ikhlas 2 [113] Al-Falaq 1	
13.	see creation	[78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [84] Al-Inshiqaq 15 [85] Al-Burooj 13 [85] Al-Burooj 15 [86] At-Tariq 5	[89] Al-Fajr 6 [96] Al-Alaq 7 [99] Al-Zalzalah 6 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8 [102] At-Takathur 6 [102] At-Takathur 7 [110] An-Nasr 2	16
14.	when quran recited	[81] At-Takweer 25 [84] Al-Inshiqaq 12 [84] Al-Inshiqaq 21 [85] Al-Burooj 21	[87] Al-Alaa 1 [96] Al-Alaq 3 [96] Al-Alaq 6 [98] Al-Bayyinah 2	8
15.	wealth is not eternal	[89] Al-Fajr 20 [90] Al-Balad 6 [92] Al-Layl 11 [92] Al-Layl 18	[100] Al-Adiyaat 8 [104] Al-Humazah 2 [104] Al-Humazah3 [111] Al-Lahab 2	8
16.	moon follows sun	[79] An-Nazi'at 22 [81] At-Takweer 1 [84] Al-Inshiqaq 18	[91] Asy-Shams 1 [91] Asy-Shams 2 [91] Asy-Shams 5	6
17.	advice patience and kindness	[89] Al-Fajr 28 [90] Al-Balad 17 [103] Al-Asr 2	[103] Al-Asr 3 [104] Al-Humazah 4	5
18.	enter my heaven	[78] An-Naba 12 [78] An-Naba 19 [78] An-Naba 37	[85] Al-Burooj 9 [86] At-Tariq 1 [87] Al-Alaa 12	19

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.11 Experiment Results of Cosine Similarity without Query Expansion (more)

		[81] At-Takweer 11 [81] At-Takweer 21 [81] At-Takweer 25 [82] Al-Infitar 1 [82] Al-Infitar 15 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12	[88] Al-Gashiyah 4 [89] Al-Fajr 29 [89] Al-Fajr 30 [92] Al-Layl 21 [110] An-Nasr 2 [111] Al-Lahab 3	
19.	the greatest punishment	[78] An-Naba 30 [78] An-Naba 2 [78] An-Naba 26 [78] An-Naba 40 [79] An-Nazi'at 25	[85] Al-Burooj 10 [85] Al-Burooj 11 [87] Al-Alaa 12 [88] Al-Gashiyah24 [89] Al-Fajr 25 [89] Al-Fajr 13	11
20.	give the book	[78] An-Naba 29 [80] Abasa 6 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 20 [83] Al-Mutaffifeen 3 [83] Al-Mutaffifeen 9 [84] Al-Inshiqaq 24 [86] At-Tariq 17	[89] Al-Fajr 15 [92] Al-Layl 12 [92] Al-Layl 18 [92] Al-Layl 5 [93] Ad-Duha 5 [98] Al-Bayyinah 5 [104] Al-Humazah5 [112] Al-Ikhlas 3	16
21.	prefer the present live	[79] An-Nazi'at 10 [79] An-Nazi'at 38	[87] Al-Alaa 16 [101] Al-Qoriah 7	4
22.	there is no power	[81] At-Takweer 20 [82] Al-Infitar 19	[86] At-Tariq 10 [89] Al-Fajr 5	4
23.	hell fuel	[78] An-Naba 21 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [81] At-Takweer 12 [85] Al-Burooj 10	[85] Al-Burooj 5 [92] Al-Layl 11 [98] Al-Bayyinah 6 [101] Al-Qoriah 9 [102] At-Takathur 6	10
24.	back to family	[82] Al-Infitar 5 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 14	[84] Al-Inshiqaq 9 [94] Al-Inshirah 2 [94] Al-Inshirah 3	7
25.	do you know	[78] An-Naba 4 [78] An-Naba 5 [81] At-Takweer 14 [82] Al-Infitar 12 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 5 [83] Al-Mutaffifeen 19 [83] Al-Mutaffifeen 8 [84] Al-Inshiqaq 23	[87] Al-Alaa 7 [90] Al-Balad 12 [96] Al-Alaq 14 [96] Al-Alaq 5 [97] Al-Qadr 2 [100] Al-Adiyaat 9 [101] Al-Qoriah 10 [101] Al-Qoriah 3 [102] At-Takathur 3 [102] At-Takathur 4	20
26.	refuse to believe	[84] Al-Inshiqaq 20 [84] Al-Inshiqaq 25 [95] At-Teen 6	[98] Al-Bayyinah 7 [107] Al-Maun 7	5

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.12 Experiment Results of Cosine Similarity without Query Expansion (more)

27.	righteous are in pleasure	[78] An-Naba 31 [78] An-Naba 38 [82] Al-Infitar 13 [83] Al-Mutaffifeen 12 [83] Al-Mutaffifeen 22 [84] Al-Inshiqaq 25	[84] Al-Inshiqaq 7 [85] Al-Burooj 11 [88] Al-Gashiyah 8 [90] Al-Balad 18 [95] At-Teen 6 [103] Al-Asr 3	12
28.	moses and pharaoh	[79] An-Nazi'at 15 [79] An-Nazi'at 21	[79] An-Nazi'at 17 [87] Al-Alaa 19	4
29.	Destruction	[91] Asy-Shams 14 [84] Al-Inshiqaq 11 [83] Al-Mutaffifeen 10	[78] An-Naba 40 [107] Al-Maun 4	5
30.	warning for whoever fear	[78] An-Naba 40 [79] An-Nazi'at 45 [79] An-Nazi'at 26 [79] An-Nazi'at 19 [79] An-Nazi'at 8 [79] An-Nazi'at 9 [79] An-Nazi'at 40 [80] Abasa 9 [83] Al-Mutaffifeen 18 [87] Al-Alaa 10	[88] Al-Gashiyah 2 [91] Asy-Shams15 [92] Al-Layl 14 [92] Al-Layl 5 [92] Al-Layl 17 [96] Al-Alaq 12 [98] Al-Bayyinah 8 [106] Al-Quraish 4 [106] Al-Quraish 2	18

And the following in table 4.11 are the results of the system with the Cosine Similarity method using a thesaurus.

Table 4.13 Experiment Results of Cosine Similarity using Query Expansion

No.	***Query***	**Results of Cosine Similarity using Query Expansion**		**Total**
1.	Day	[78] An-Naba 11 [78] An-Naba 17 [78] An-Naba 18 [78] An-Naba 39 [78] An-Naba 40 [78] An-Naba 8 [79] An-Nazi'at 46 [79] An-Nazi'at 8 [79] An-Nazi'at 35 [79] An-Nazi'at 6 [80] Abasa 40 [80] Abasa 38 [80] Abasa 37 [80] Abasa 34 [81] At-Takweer 15 [82] Al-Infitar 19 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 15 [83] Al-Mutaffifeen 5	[83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 11 [83] Al-Mutaffifeen 34 [83] Al-Mutaffifeen 6 [85] Al-Burooj 2 [86] At-Tariq 9 [88] Al-Gashiyah 2 [88] Al-Gashiyah 8 [89] Al-Fajr 1 [89] Al-Fajr 23 [89] Al-Fajr 25 [90] Al-Balad 14 [91] Asy-Shams 3 [92] Al-Layl 2 [99] Al-Zalzalah 4 [99] Al-Zalzalah 6 [100] Al-Adiyaat 11 [101] Al-Qoriah 4 [102] At-Takathur 8 [108] Al-Kawthar 1	40

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.14 Experiment Results of Cosine Similarity using Query Expansion (more)

2.	Disbeliver	[78] An-Naba 40 [79] An-Nazi'at 17 [80] Abasa 42 [83] Al-Mutaffifeen36	[83] Al-Mutaffifeen 30 [86] At-Tariq 17 [88] Al-Gashiyah 23 [109] Al-Kafiroon 1	8
3.	Sin	[79] An-Nazi'at 17 [81] At-Takweer 9 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen12	[91] Asy-Shams14 [96] Al-Alaq 16 [82] Al-Infitar 14	7
4.	Rest	[78] An-Naba 9	[89] Al-Fajr 27	2
5.	Glory	[80] Abasa 13 [85] Al-Burooj 21 [85] Al-Burooj 15	[89] Al-Fajr 17 [89] Al-Fajr 15 [92] Al-Layl 2	6
6.	Punishment	[78] An-Naba 30 [78] An-Naba 40 [79] An-Nazi'at 25 [84] Al-Inshiqaq 24	[85] Al-Burooj 10 [88] Al-Gashiyah 24 [89] Al-Fajr 25 [89] Al-Fajr 13	8
7.	Say	[78] An-Naba 38 [78] An-Naba 40 [79] An-Nazi'at 43 [79] An-Nazi'at 12 [79] An-Nazi'at 10 [79] An-Nazi'at 46 [83] Al-Mutaffifeen 17 [83] Al-Mutaffifeen 32 [83] Al-Mutaffifeen 13 [89] Al-Fajr 27	[89] Al-Fajr 24 [89] Al-Fajr 16 [89] Al-Fajr 15 [90] Al-Balad 6 [91] Asy-Shams 13 [99] Al-Zalzalah 3 [109] Al-Kafiroon 1 [112] Al-Ikhlas 1 [113] Al-Falaq 1 [114] An-Naas 1	20
8.	bad deeds	[78] An-Naba 26 [78] An-Naba 8 [78] An-Naba 36 [78] An-Naba 40 [79] An-Nazi'at 18 [81] At-Takweer 10 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 6 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 9 [81] At-Takweer 14	[95] At-Teen 6 [99] Al-Zalzalah 8 [99] Al-Zalzalah 6 [101] Al-Qoriah 8 [101] Al-Qoriah 6 [103] Al-Asr 3 [107] Al-Maun 6 [113] Al-Falaq 2 [113] Al-Falaq 5 [113] Al-Falaq 3 [113] Al-Falaq 4 [114] An-Naas 4	25
9.	allah swears	[78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 25 [80] Abasa 9 [81] At-Takweer 15 [81] At-Takweer 29 [82] Al-Infitar 19 [83] Al-Mutaffifeen 21	[89] Al-Fajr 11 [90] Al-Balad 1 [91] Asy-Shams 5 [91] Asy-Shams 6 [91] Asy-Shams 2 [91] Asy-Shams 15 [91] Asy-Shams 4 [91] Asy-Shams 13	54

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.15 Experiment Results of Cosine Similarity using Query Expansion (more)

		[83] Al-Mutaffifeen 28 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 16 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [85] Al-Burooj 3 [85] Al-Burooj 20 [85] Al-Burooj 1 [85] Al-Burooj 8 [85] Al-Burooj 9 [85] Al-Burooj 15 [87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24 [88] Al-Gashiyah 3 [89] Al-Fajr 5 [89] Al-Fajr 2 [89] Al-Fajr 3 [89] Al-Fajr 1	[89] Al-Fajr 4 [92] Al-Layl 1 [92] Al-Layl 5 [95] At-Teen 1 [95] At-Teen 8 [95] At-Teen 2 [96] Al-Alaq 14 [96] Al-Alaq 19 [96] Al-Alaq 12 [103] Al-Asr 1 [104] Al-Humazah 6 [106] Al-Quraish 1 [106] Al-Quraish 3 [106] Al-Quraish 2 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 2 [112] Al-Ikhlas 1 [113] Al-Falaq 1	
10.	messenger from god	[78] An-Naba 36 [78] An-Naba 37 [79] An-Nazi'at 24 [79] An-Nazi'at 19 [79] An-Nazi'at 5 [79] An-Nazi'at 44 [79] An-Nazi'at 40 [80] Abasa 15 [81] At-Takweer 19 [81] At-Takweer 29 [82] Al-Infitar 6 [83] Al-Mutaffifeen 15 [83] Al-Mutaffifeen 6 [84] Al-Inshiqaq 15 [84] Al-Inshiqaq 2 [84] Al-Inshiqaq 5 [84] Al-Inshiqaq 6 [87] Al-Alaa 1 [87] Al-Alaa 15 [89] Al-Fajr 15 [89] Al-Fajr 14 [89] Al-Fajr 13 [89] Al-Fajr 16	[89] Al-Fajr 22 [91] Asy-Shams 13 [91] Asy-Shams 14 [92] Al-Layl 20 [93] Ad-Duha 5 [93] Ad-Duha 3 [94] Al-Inshirah 8 [96] Al-Alaq 12 [96] Al-Alaq 1 [97] Al-Qadr 4 [98] Al-Bayyinah 2 [98] Al-Bayyinah 8 [98] Al-Bayyinah 5 [100] Al-Adiyaat 11 [105] Al-Feel 1 [106] Al-Quraish 3 [108] Al-Kawthar 2 [113] Al-Falaq 1 [114] An-Naas 3 [114] An-Naas 1 [79] An-Nazi'at 16 [78] An-Naba 39	45
11.	night of decree	[78] An-Naba 10 [79] An-Nazi'at 29 [80] Abasa 41 [81] At-Takweer 17 [81] At-Takweer 15 [84] Al-Inshiqaq 17 [89] Al-Fajr 2 [89] Al-Fajr 4	[91] Asy-Shams 4 [92] Al-Layl 1 [93] Ad-Duha 2 [97] Al-Qadr 1 [97] Al-Qadr 2 [97] Al-Qadr 3 [113] Al-Falaq 3	15

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.16 Experiment Results of Cosine Similarity using Query Expansion (more)

12.	allah taught man	[78] An-Naba 40 [78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 35 [79] An-Nazi'at 25 [80] Abasa 24 [80] Abasa 9 [80] Abasa 23 [80] Abasa 2 [80] Abasa 17 [80] Abasa 34 [81] At-Takweer 29 [82] Al-Infitar 6 [82] Al-Infitar 19 [83] Al-Mutaffifeen 6 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 6 [85] Al-Burooj 20 [85] Al-Burooj 8 [85] Al-Burooj 9 [85] Al-Burooj 15 [86] At-Tariq 5 [86] At-Tariq 10 [87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24 [88] Al-Gashiyah 3 [89] Al-Fajr 11	[89] Al-Fajr 15 [90] Al-Balad 4 [91] Asy-Shams 15 [91] Asy-Shams 13 [92] Al-Layl 5 [95] At-Teen 4 [95] At-Teen 8 [96] Al-Alaq 5 [96] Al-Alaq 4 [96] Al-Alaq 6 [96] Al-Alaq 2 [96] Al-Alaq 14 [96] Al-Alaq 19 [96] Al-Alaq 12 [99] Al-Zalzalah 6 [99] Al-Zalzalah 3 [100] Al-Adiyaat 6 [103] Al-Asr 2 [104] Al-Humazah 6 [106] Al-Quraish 1 [106] Al-Quraish 3 [106] Al-Quraish 2 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 2 [112] Al-Ikhlas 1 [113] Al-Falaq 1 [114] An-Naas 3 [114] An-Naas 2 [114] An-Naas 6 [114] An-Naas 1	63
13.	see creation	[78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [80] Abasa 24 [84] Al-Inshiqaq 15 [85] Al-Burooj 13 [85] Al-Burooj 15 [86] At-Tariq 5	[88] Al-Gashiyah 17 [89] Al-Fajr 6 [96] Al-Alaq 7 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8 [99] Al-Zalzalah 6 [102] At-Takathur 7 [102] At-Takathur 6 [110] An-Nasr 2	18
14.	when quran recited	[81] At-Takweer 25 [84] Al-Inshiqaq 12 [84] Al-Inshiqaq 21 [85] Al-Burooj 21	[87] Al-Alaa 1 [96] Al-Alaq 2 [96] Al-Alaq 6 [98] Al-Bayyinah 2	8
15.	wealth is not eternal	[78] An-Naba 23 [78] An-Naba 14	[100] Al-Adiyaat 8 [104] Al-Humazah 2	11

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.17 Experiment Results of Cosine Similarity using Query Expansion (more)

		[89] Al-Fajr 20 [90] Al-Balad 6 [92] Al-Layl 18 [92] Al-Layl 11	[104] Al-Humazah 3 [108] Al-Kawthar 1 [111] Al-Lahab 2	
16.	moon follows sun	[79] An-Nazi'at 22 [81] At-Takweer 1 [84] Al-Inshiqaq 18	[91] Asy-Shams 1 [91] Asy-Shams 2 [91] Asy-Shams 5	6
17.	advice patience and kindness	[89] Al-Fajr 27 [90] Al-Balad 17 [103] Al-Asr 2	[103] Al-Asr 3 [104] Al-Humazah 4	5
18.	enter my heaven	[78] An-Naba 31 [78] An-Naba 12 [78] An-Naba 19 [78] An-Naba 37 [79] An-Nazi'at 41 [81] At-Takweer 11 [81] At-Takweer 13 [81] At-Takweer 21 [81] At-Takweer 25 [82] Al-Infitar 15 [82] Al-Infitar 1 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12	[85] Al-Burooj 9 [86] At-Tariq 1 [87] Al-Alaa 12 [88] Al-Gashiyah 4 [89] Al-Fajr 30 [89] Al-Fajr 29 [92] Al-Layl 21 [98] Al-Bayyinah 8 [101] Al-Qoriah 7 [108] Al-Kawthar 1 [110] An-Nasr 2 [111] Al-Lahab 3	25
19.	the greatest punishment	[78] An-Naba 30 [78] An-Naba 2 [78] An-Naba 40 [79] An-Nazi'at 25 [79] An-Nazi'at 20 [79] An-Nazi'at 9 [84] Al-Inshiqaq 24	[85] Al-Burooj 10 [85] Al-Burooj 11 [87] Al-Alaa 12 [88] Al-Gashiyah24 [89] Al-Fajr 25 [89] Al-Fajr 13 [106] Al-Quraish 1	14
20.	give the book	[78] An-Naba 29 [80] Abasa 6 [80] Abasa 13 [82] Al-Infitar 11 [83] Al-Mutaffifeen 9 [83] Al-Mutaffifeen 20 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen 3 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 7	[84] Al-Inshiqaq 24 [86] At-Tariq 17 [89] Al-Fajr 15 [92] Al-Layl 15 [92] Al-Layl 12 [92] Al-Layl 5 [92] Al-Layl 18 [93] Ad-Duha 5 [98] Al-Bayyinah 5 [104] Al-Humazah5 [112] Al-Ikhlas 3	23
21.	prefer the present live	[78] An-Naba 39 [79] An-Nazi'at 38 [79] An-Nazi'at 10 [79] An-Nazi'at 46 [81] At-Takweer 29	[81] At-Takweer 25 [83] Al-Mutaffifeen 6 [86] At-Tariq 13 [87] Al-Alaa 16 [88] Al-Gashiyah 3	14

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.18 Experiment Results of Cosine Similarity using Query Expansion (more)

		[81] At-Takweer 19 [81] At-Takweer 27	[101] Al-Qoriah7 [102] At-Takathur 1	
22.	there is no power	[86] At-Tariq 10 [110] An-Nasr 1 [82] Al-Infitar 19	[89] Al-Fajr 5 [81] At-Takweer 20	5
23.	hell fuel	[78] An-Naba 21 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [81] At-Takweer 6 [81] At-Takweer 12 [84] Al-Inshiqaq 12 [85] Al-Burooj 5 [85] Al-Burooj 10 [86] At-Tariq 3 [87] Al-Alaa 12 [88] Al-Gashiyah 4	[89] Al-Fajr 23 [90] Al-Balad 20 [92] Al-Layl 14 [92] Al-Layl 11 [98] Al-Bayyinah 6 [101] Al-Qoriah11 [101] Al-Qoriah9 [102] At-Takathur 6 [104] Al-Humazah6 [111] Al-Lahab 3	21
24.	back to family	[79] An-Nazi'at 12 [79] An-Nazi'at 44 [79] An-Nazi'at 10 [82] Al-Infitar 5 [83] Al-Mutaffifeen 31 [84] Al-Inshiqaq 9 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 14 [84] Al-Inshiqaq 6	[86] At-Tariq 11 [86] At-Tariq 8 [88] Al-Gashiyah 25 [89] Al-Fajr 28 [92] Al-Layl 19 [94] Al-Inshirah3 [94] Al-Inshirah2 [95] At-Teen 5 [96] Al-Alaq 8	19
25.	do you know	[78] An-Naba 4 [78] An-Naba 5 [81] At-Takweer 14 [82] Al-Infitar 12 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 5 [83] Al-Mutaffifeen 24 [83] Al-Mutaffifeen 8 [83] Al-Mutaffifeen 19 [84] Al-Inshiqaq 23	[87] Al-Alaa 7 [90] Al-Balad 12 [96] Al-Alaq 5 [96] Al-Alaq 14 [97] Al-Qadr 2 [100] Al-Adiyaat 9 [101] Al-Qoriah10 [101] Al-Qoriah 3 [102] At-Takathur 3 [102] At-Takathur 4	21
26.	refuse to believe	[80] Abasa 42 [82] Al-Infitar 9 [83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 17 [84] Al-Inshiqaq 20 [84] Al-Inshiqaq 25	[91] Asy-Shams11 [91] Asy-Shams14 [95] At-Teen 7 [95] At-Teen 6 [98] Al-Bayyinah 7 [107] Al-Maun 7	12
27.	righteous are in pleasure	[78] An-Naba 31 [78] An-Naba 38 [82] Al-Infitar 13 [83] Al-Mutaffifeen 22	[84] Al-Inshiqaq 7 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 8	14

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.19 Experiment Results of Cosine Similarity using Query Expansion (more)

		[83] Al-Mutaffifeen 12 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 9	[90] Al-Balad 18 [95] At-Teen 6 [103] Al-Asr 3	
28.	moses and pharaoh	[79] An-Nazi'at 15 [79] An-Nazi'at 21	[79] An-Nazi'at 17 [87] Al-Alaa 19	4
29.	Destruction	[78] An-Naba 40 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 10 [84] Al-Inshiqaq 11	[91] Asy-Shams14 [104] Al-Humazah 1 [107] Al-Maun 4	7
30.	warning for whoever fear	[78] An-Naba 40 [79] An-Nazi'at 45 [79] An-Nazi'at 26 [79] An-Nazi'at 19 [79] An-Nazi'at 8 [79] An-Nazi'at 9 [79] An-Nazi'at 40 [80] Abasa 9 [83] Al-Mutaffifeen 18 [87] Al-Alaa 10	[88] Al-Gashiyah 2 [91] Asy-Shams15 [92] Al-Layl 14 [92] Al-Layl 5 [92] Al-Layl 17 [96] Al-Alaq 12 [98] Al-Bayyinah 8 [106] Al-Quraish 4 [106] Al-Quraish 2	19

The relevance of the similarity query with the documents obtained by the system is matched with the reference data, namely the search results by one of the Quran Translation websites, namely helloquran.com and quranindex.info. The following table 4.16 is a comparison between the system results with the Cosine Similarity method with the reference data and then the values of TP, FP, FN and TN can be determined, based on the details as described in table 3.13. The following are the search results by reference data and the Cosine Similarity Method and Values TP, FP, FN, TN that only apply Cosine Similarity:

Table 4.20 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
1.	Day	[78] An-Naba 11 [78] An-Naba 17 [78] An-Naba 18 [78] An-Naba 39 [78] An-Naba 40 [78] An-Naba 8	[78] An-Naba 11 [78] An-Naba 17 [78] An-Naba 18 [78] An-Naba 39 [78] An-Naba 40 [78] An-Naba 8	38	2	1	523

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.21 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[79] An-Nazi'at 46 [79] An-Nazi'at 8 [79] An-Nazi'at 35 [79] An-Nazi'at 6 [80] Abasa 40 [80] Abasa 38 [80] Abasa 37 [80] Abasa 34 [81] At-Takweer 15 [82] Al-Infitar 19 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 15 [83] Al-Mutaffifeen 15 [83] Al-Mutaffifeen 5 [83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 11 [83] Al-Mutaffifeen 34 [85] Al-Burooj 2 [86] At-Tariq 9 [88] Al-Gashiyah 2 [88] Al-Gashiyah 8 [89] Al-Fajr 23 [89] Al-Fajr 25 [90] Al-Balad 14 [91] Asy-Shams3 [92] Al-Layl 2 [99] Al-Zalzalah 4 [99] Al-Zalzalah 6 [100] Al-Adiyaat 11 [101] Al-Qoriah 4 [102] At-Takathur 8 [108] Al-Kawthar	[79] An-Nazi'at 46 [79] An-Nazi'at 8 [79] An-Nazi'at 35 [79] An-Nazi'at 6 [80] Abasa 40 [80] Abasa 38 [80] Abasa 37 [80] Abasa 34 [81] At-Takweer 15 [82] Al-Infitar 19 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 15 [83] Al-Mutaffifeen 5 [83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 11 [83] Al-Mutaffifeen 34 [83] Al-Mutaffifeen 6 [85] Al-Burooj 2 [86] At-Tariq 9 [88] Al-Gashiyah 2 [88] Al-Gashiyah 8 [89] Al-Fajr 1 [89] Al-Fajr 23 [89] Al-Fajr 25 [90] Al-Balad 14 [91] Asy-Shams3 [92] Al-Layl 2 [99] Al-Zalzalah 4 [99] Al-Zalzalah 6 [100] Al-Adiyaat 11 [101] Al-Qoriah4 [102] At-Takathur 8 [108] Al-Kawthar 1				
2.	Disbeliver	[78] An-Naba 40 [79] An-Nazi'at 17 [86] At-Tariq 17 [88] Al-Gashiyah 23	[78] An-Naba 40 [79] An-Nazi'at 17 [83] Al-Mutaffifeen 30 [86] At-Tariq 17	7	1	0	556
		[80] Abasa 42 [83] Al-Mutaffifeen36 [109] Al-Kafiroon 1	[80] Abasa 42 [83] Al-Mutaffifeen36 [88] Al-Gashiyah 23 [109] Al-Kafiroon 1				

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.22 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
3.	Sin	[81] At-Takweer 9 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen12 [91] Asy-Shams14 [96] Al-Alaq 16	[[82] Al-Infitar 14 [81] At-Takweer 9 [83] Al-Mutaffifeen12 [91] Asy-Shams14 [96] Al-Alaq 16	4	1	2	557
4.	Rest	[78] An-Naba 9 [89] Al-Fajr 27	[78] An-Naba 9 [89] Al-Fajr 27	2	0	0	562
5.	Glory	[80] Abasa 13 [85] Al-Burooj 21 [85] Al-Burooj 15 [89] Al-Fajr 15 [92] Al-Layl 2	[85] Al-Burooj 15 [89] Al-Fajr 17 [92] Al-Layl 2	2	1	4	557
6.	Punishment	[78] An-Naba 30 [78] An-Naba 40 [79] An-Nazi'at 25 [84] Al-Inshiqaq 24 [85] Al-Burooj 10 [88] Al-Gashiyah 24 [89] Al-Fajr 25 [89] Al-Fajr 13	[79] An-Nazi'at 25 [84] Al-Inshiqaq 24 [85] Al-Burooj 10 [88] Al-Gashiyah 24 [89] Al-Fajr 25	5	0	3	556
7.	Say	[78] An-Naba 37 [78] An-Naba 38 [78] An-Naba 40 [79] An-Nazi'at 18 [79] An-Nazi'at 24 [79] An-Nazi'at 43 [79] An-Nazi'at 12 [79] An-Nazi'at 10 [83] Al-Mutaffifeen 17 [83] Al-Mutaffifeen 32 [83] Al-Mutaffifeen 13 [89] Al-Fajr 27 [89] Al-Fajr 24 [89] Al-Fajr 16 [89] Al-Fajr 15 [90] Al-Balad 6 [91] Asy-Shams 13 [99] Al-Zalzalah 3 [109] Al-Kafiroon 1 [112] Al-Ikhlas 1 [113] Al-Falaq 1 [114] An-Naas 1	[78] An-Naba 40 [79] An-Nazi'at 12 [79] An-Nazi'at 10 [79] An-Nazi'at 46 [83] Al-Mutaffifeen 13 [89] Al-Fajr 24 [89] Al-Fajr 16 [89] Al-Fajr 15 [90] Al-Balad 6 [99] Al-Zalzalah 3 [109] Al-Kafiroon 1 [112] Al-Ikhlas 1 [113] Al-Falaq 1 [114] An-Naas 1	13	1	9	541

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.23 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
8.	bad deeds	[78] An-Naba 40 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 9 [95] At-Teen 6 [99] Al-Zalzalah 8 [99] Al-Zalzalah 6 [101] Al-Qoriah 8 [101] Al-Qoriah 6 [103] Al-Asr 3 [107] Al-Maun 6 [113] Al-Falaq 2 [113] Al-Falaq 5 [113] Al-Falaq 3 [113] Al-Falaq 4 [114] An-Naas 4	[78] An-Naba 36 [78] An-Naba 40 [78] An-Naba 8 [81] At-Takweer 10 [81] At-Takweer 14 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 25 [84] Al-Inshiqaq 6 [85] Al-Burooj 11 [88] Al-Gashiyah 9 [95] At-Teen 6 [99] Al-Zalzalah 6 [101] Al-Qoriah 6 [101] Al-Qoriah 8 [103] Al-Asr 3 [107] Al-Maun 6	12	7	6	539
9.	allah swears	[78] An-Naba 38 [79] An-Nazi'at 26 [79] An-N azi'at25 [80] Abasa 9 [81] At-Takweer 15 [81] At-Takweer 29 [82] Al-Infitar 19 [83] Al-Mutaffifeen 28 [84] Al-Inshiqaq 16 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [85] Al-Burooj 3 [85] Al-Burooj 20 [85] Al-Burooj 1 [85] Al-Burooj 8 [85] Al-Burooj 9 [85] Al-Burooj 15 [87] Al-Alaa 10 [87] Al-Alaa 7 [89] Al-Fajr 11 [90] Al-Balad 1 [91] Asy-Shams 5 [91] Asy-Shams 6 [91] Asy-Shams 2	[78] An-Naba 38 [79] An-Nazi'at 25 [79] An-Nazi'at 26 [80] Abasa 9 [81] At-Takweer 15 [81] At-Takweer 29 [82] Al-Infitar 19 [83] Al-Mutaffifeen 18 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [84] Al-Inshiqaq 16 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [85] Al-Burooj 1 [85] Al-Burooj 15 [85] Al-Burooj 20 [85] Al-Burooj 3 [85] Al-Burooj 8 [85] Al-Burooj 9 [89] Al-Fajr 4 [90] Al-Balad 1 [91] Asy-Shams 13 [91] Asy-Shams 15 [91] Asy-Shams 2	49	4	1	510

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.24 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[91] Asy-Shams 15 [91] Asy-Shams 4 [91] Asy-Shams 13 [92] Al-Layl 1 [92] Al-Layl 5 [95] At-Teen 1 [95] At-Teen 8 [95] At-Teen 2 [96] Al-Alaq 14 [96] Al-Alaq 19 [96] Al-Alaq 12 [103] Al-Asr 1 [104] Al-Humazah [106] Al-Quraish 3	[91] Asy-Shams 4 [91] Asy-Shams 5 [91] Asy-Shams 6 [92] Al-Layl 1 [92] Al-Layl 5 [95] At-Teen 1 [95] At-Teen 2 [95] At-Teen 8 [96] Al-Alaq 12 [96] Al-Alaq 14 [96] Al-Alaq 19 [103] Al-Asr 1 [104] Al-Humazah6 [106] Al-Quraish 1				
		[88] Al-Gashiyah 24 [89] Al-Fajr 5 [89] Al-Fajr 2 [89] Al-Fajr 3 [89] Al-Fajr 1 [89] Al-Fajr 4 [106] Al-Quraish 2 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 2 [112] Al-Ikhlas 1 [113] Al-Falaq 1	[87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24 [88] Al-Gashiyah 3 [89] Al-Fajr 1 [89] Al-Fajr 11 [89] Al-Fajr 2 [89] Al-Fajr 3 [106] Al-Quraish 2 [106] Al-Quraish 3 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 1 [112] Al-Ikhlas 2 [113] Al-Falaq 1				
10.	messenger from god	[78] An-Naba 36 [78] An-Naba 37 [79] An-Nazi'at 24 [79] An-Nazi'at 19 [79] An-Nazi'at 5 [79] An-Nazi'at 44 [79] An-Nazi'at 40 [80] Abasa 15 [81] At-Takweer 20 [81] At-Takweer 19 [81] At-Takweer 29 [82] Al-Infitar 6	[80] Abasa 15 [81] At-Takweer 19 [91] Asy-Shams 13 [96] Al-Alaq 12 [98] Al-Bayyinah 2 [98] Al-Bayyinah 5 [98] Al-Bayyinah 8 [114] An-Naas 3	5	3	41	515

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.25 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[83] Al-Mutaffifeen 15 [83] Al-Mutaffifeen 6 [84] Al-Inshiqaq 15 [84] Al-Inshiqaq 2 [84] Al-Inshiqaq 5 [84] Al-Inshiqaq 6 [85] Al-Burooj 12 [87] Al-Alaa 1 [89] Al-Fajr 16 [89] Al-Fajr 22 [89] Al-Fajr 6 [89] Al-Fajr 28 [91] Asy-Shams13 [91] Asy-Shams14 [92] Al-Layl 20 [93] Ad-Duha 5 [93] Ad-Duha 3 [94] Al-Inshirah8 [96] Al-Alaq 1 [97] Al-Qadr 4 [98] Al-Bayyinah 2 [98] Al-Bayyinah 8 [100] Al-Adiyaat 11 [105] Al-Feel 1 [106] Al-Quraish 3 [108] Al-Kawthar 2 [113] Al-Falaq 1 [114] An-Naas 3 [87] Al-Alaa 15 [89] Al-Fajr 15 [89] Al-Fajr 14 [89] Al-Fajr 13 [114] An-Naas 1 [79] An-Nazi'at 16 [78] An-Naba 39					
11.	night of decree	[78] An-Naba 10 [79] An-Nazi'at 29 [80] Abasa 41 [81] At-Takweer 17 [84] Al-Inshiqaq 17 [89] Al-Fajr 2	[78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [84] Al-Inshiqaq 15 [85] Al-Burooj 13	13	1	1	549

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.26 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[89] Al-Fajr 4 [91] Asy-Shams4 [92] Al-Layl 1 [93] Ad-Duha 2 [97] Al-Qadr 1 [97] Al-Qadr 2 [97] Al-Qadr 3 [113] Al-Falaq 3	[85] Al-Burooj 15 [86] At-Tariq 5 [89] Al-Fajr 6 [96] Al-Alaq 7 [99] Al-Zalzalah 6 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8 [102] At-Takathur 6 [102] At-Takathur 7 [110] An-Nasr 2				
12.	allah taught man	[78] An-Naba 40 [78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 35 [80] Abasa 24 [80] Abasa 23 [80] Abasa 2 [80] Abasa 17 [80] Abasa 34 [81] At-Takweer 29 [81] At-Takweer 27 [82] Al-Infitar 6 [82] Al-Infitar 19 [83] Al-Mutaffifeen 6 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [90] Al-Balad 4 [91] Asy-Shams15 [91] Asy-Shams13 [92] Al-Layl 5 [95] At-Teen 4 [95] At-Teen 6 [96] Al-Alaq 5 [96] Al-Alaq 4 [96] Al-Alaq 6 [96] Al-Alaq 2 [96] Al-Alaq 14 [96] Al-Alaq 19 [96] Al-Alaq 12 [99] Al-Zalzalah 6 [99] Al-Zalzalah 3 [100] Al-Adiyaat 6	[78] An-Naba 38 [78] An-Naba 40 [79] An-Nazi'at 25 [79] An-Nazi'at 26 [79] An-Nazi'at 35 [80] Abasa 17 [80] Abasa 2 [80] Abasa 23 [80] Abasa 34 [80] Abasa 9 [81] At-Takweer 29 [82] Al-Infitar 19 [82] Al-Infitar 6 [83] Al-Mutaffifeen 18 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [88] Al-Gashiyah 3 [89] Al-Fajr 11 [89] Al-Fajr 15 [90] Al-Balad 4 [91] Asy-Shams 13 [91] Asy-Shams 15 [92] Al-Layl 5 [95] At-Teen 4 [95] At-Teen 8 [96] Al-Alaq 12 [96] Al-Alaq 14 [96] Al-Alaq 19 [96] Al-Alaq 2 [96] Al-Alaq 4 [96] Al-Alaq 5 96] Al-Alaq 6	46	9	9	500

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.27 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[84] Al-Inshiqaq 23 [84] Al-Inshiqaq 6 [85] Al-Burooj 20 [85] Al-Burooj 8 [85] Al-Burooj 9 [85] Al-Burooj 15 [86] At-Tariq 5 [86] At-Tariq 10 [87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24 [89] Al-Fajr 15 [103] Al-Asr 2 [104] Al-Humazah6 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 2 [112] Al-Ikhlas 1 [113] Al-Falaq 1 [114] An-Naas 3 [114] An-Naas 2 [114] An-Naas 6 [114] An-Naas 1	[84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 6 [85] Al-Burooj 15 [85] Al-Burooj 20 [85] Al-Burooj 8 [85] Al-Burooj 9 [86] At-Tariq 10 [86] At-Tariq 5 [87] Al-Alaa 10 [87] Al-Alaa 7 [88] Al-Gashiyah 24 [99] Al-Zalzalah 3 [103] Al-Asr 2 [104] Al-Humazah6 [106] Al-Quraish 1 [106] Al-Quraish 2 [106] Al-Quraish 3 [110] An-Nasr 1 [110] An-Nasr 2 [112] Al-Ikhlas 1 [112] Al-Ikhlas 2 [113] Al-Falaq 1				
13.	see creation	[78] An-Naba 8 [78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [80] Abasa 18 [80] Abasa 19 [82] Al-Infitar 7 [83] Al-Mutaffifeen 35 [80] Abasa 24 [84] Al-Inshiqaq 15 [86] At-Tariq 5 [86] At-Tariq 6 [86] At-Tariq 7 [88] Al-Gashiyah 17 [89] Al-Fajr 6 [90] Al-Balad 7 [92] Al-Layl 3 [95] At-Teen 4	[78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [84] Al-Inshiqaq 15 [85] Al-Burooj 13 [85] Al-Burooj 15 [86] At-Tariq 5 [89] Al-Fajr 6 [96] Al-Alaq 7 [99] Al-Zalzalah 6 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8 [102] At-Takathur 6 [102] At-Takathur 7 [110] An-Nasr 2	14	2	15	533

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.28 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[96] Al-Alaq 1 [96] Al-Alaq 2 [96] Al-Alaq 7 [102] At-Takathur 7 [102] At-Takathur 6 [105] Al-Feel 1 [110] An-Nasr 2 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8[99] Al-Zalzalah 6					
14.	when quran recited	[81] At-Takweer 25 [83] Al-Mutaffifeen 13 [84] Al-Inshiqaq 21 [85] Al-Burooj 21 [87] Al-Alaa 6 [96] Al-Alaq 1 [96] Al-Alaq [98] Al-Bayyinah	[81] At-Takweer [84] Al-Inshiqaq [84] Al-Inshiqaq [85] Al-Burooj [87] Al-Alaa [96] Al-Alaq [96] Al-Alaq [98] Al-Bayyinah	7	1	1	555
15.	wealth is not eternal	[78] An-Naba 23 [78] An-Naba 14 [89] Al-Fajr 20 [90] Al-Balad 6 [92] Al-Layl 18 [92] Al-Layl 11 [95] At-Teen 6 [100] Al-Adiyaat 8 [104] Al-Humazah 2 [108] Al-Kawthar 1 [111] Al-Lahab 2 [104] Al-Humazah3	[89] Al-Fajr 20 [90] Al-Balad 6 [92] Al-Layl 11 [92] Al-Layl 18 [[100] Al-Adiyaat 8 [104] Al-Humazah 2 [104] Al-Humazah3 [111] Al-Lahab 2	8	0	3	553
16.	moon follows sun	[79] An-Nazi'at 7 [81] At-Takweer 1 [84] Al-Inshiqaq 18 [91] Asy-Shams 12 [91] Asy-Shams 1 [91] Asy-Shams 2 [91] Asy-Shams 4	[79] An-Nazi'at 7 [81] At-Takweer 1 [84] Al-Inshiqaq 18 [91] Asy-Shams 12 [91] Asy-Shams 1 [91] Asy-Shams 4	6	0	1	557
17.	advice patience and kindness	[90] Al-Balad 17 [103] Al-Asr 3	[89] Al-Fajr [90] Al-Balad [103] Al-Asr [103] Al-Asr [104] Al-Humazah	2	3	0	559

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.29 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
18.	enter my heaven	[78] An-Naba 31 [79] An-Nazi'at 41 [81] At-Takweer 13 [81] At-Takweer 21 [82] Al-Infitar 15 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12 [87] Al-Alaa 12 [88] Al-Gashiyah 4 [89] Al-Fajr 30 [89] Al-Fajr 29 [98] Al-Bayyinah 8 [101] Al-Qoriah 7 [108] Al-Kawthar 1 [110] An-Nasr 2 [111] Al-Lahab 3	[78] An-Naba 12 [78] An-Naba 19 [78] An-Naba 37 [81] At-Takweer 11 [81] At-Takweer 21 [81] At-Takweer 25 [82] Al-Infitar 1 [82] Al-Infitar 15 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12 [85] Al-Burooj 9 [86] At-Tariq 1 [87] Al-Alaa 12 [88] Al-Gashiyah 4 [89] Al-Fajr 29 [89] Al-Fajr 30 [92] Al-Layl 21 [110] An-Nasr 2 [111] Al-Lahab 3	10	9	6	539
19.	the greatest punishment	[78] An-Naba 30 [78] An-Naba 2 [78] An-Naba 26 [78] An-Naba 40 [79] An-Nazi'at 25 [79] An-Nazi'at 20 [84] Al-Inshiqaq 24 [85] Al-Burooj 10 [85] Al-Burooj 11 [87] Al-Alaa 12 [88] Al-Gashiyah24 [89] Al-Fajr 25 [89] Al-Fajr 13	[78] An-Naba 2 [79] An-Nazi'at 9 [79] An-Nazi'at 20 [79] An-Nazi'at 25 [85] Al-Burooj 10 [85] Al-Burooj 11 [87] Al-Alaa 12 [84] Al-Inshiqaq 24 [88] Al-Gashiyah 24 [89] Al-Fajr 25 [106] Al-Quraish 1	9	2	4	549
20.	give the book	[78] An-Naba 29 [80] Abasa 6 [80] Abasa 13 [82] Al-Infitar 11 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen 9 [83] Al-Mutaffifeen 20 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen 3	[78] An-Naba 29 [80] Abasa 6 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 20 [83] Al-Mutaffifeen 3 [83] Al-Mutaffifeen 9 [84] Al-Inshiqaq 24 [86] At-Tariq 17 [89] Al-Fajr 15 [92] Al-Layl 12	15	2	8	539

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.30 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 7 [84] Al-Inshiqaq 24 [86] At-Tariq 17 [89] Al-Fajr 15 [92] Al-Layl 12 [92] Al-Layl 5 [92] Al-Layl 18 [93] Ad-Duha 5 [98] Al-Bayyinah 5 [108] Al-Kawthar 1 [112] Al-Ikhlas 3	[92] Al-Layl 18 [92] Al-Layl 5 [93] Ad-Duha 5 [98] Al-Bayyinah 5 [104] Al-Humazah5 [112] Al-Ikhlas 3				
21.	prefer the present live	[78] An-Naba 39 [79] An-Nazi'at 38 [79] An-Nazi'at 10 [81] At-Takweer 27 [83] Al-Mutaffifeen 6 [87] Al-Alaa 16 [89] Al-Fajr 24 [101] Al-Qoriah7 [102] At-Takathur 1	[79] An-Nazi'at 10 [79] An-Nazi'at 38 [87] Al-Alaa 16 [101] Al-Qoriah 7	4	2	7	551
22.	there is no power	[81] At-Takweer 20 [82] Al-Infitar 19 [86] At-Tariq 10 [90] Al-Fajr 5 [110] An-Nasr 1	[81] At-Takweer 20 [82] Al-Infitar 19 [86] At-Tariq 10 [89] Al-Fajr 5	3	1	2	558
23.	hell fuel	[78] An-Naba 21 [79] An-Nazi'at 36 [81] At-Takweer 6 [81] At-Takweer 12 [82] Al-Infitar 14 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12 [85] Al-Burooj 5 [85] Al-Burooj 10 [87] Al-Alaa 12 [88] Al-Gashiyah 4 [89] Al-Fajr 23 [90] Al-Balad 20 [92] Al-Layl 14 [98] Al-Bayyinah 6	[78] An-Naba 21 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [81] At-Takweer 12 [85] Al-Burooj 10 [85] Al-Burooj 5 [92] Al-Layl 11 [98] Al-Bayyinah 6 [101] Al-Qoriah 9 [102] At-Takathur 6	7	3	16	538

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.31 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[101] Al-Qoriah11 [101] Al-Qoriah9 [102] At-Takathur 6 [104] Al-Humazah6 [111] Al-Lahab 3					
24.	back to family	[79] An-Nazi'at 12 [79] An-Nazi'at 22 [79] An-Nazi'at 44 [79] An-Nazi'at 10 [82] Al-Infitar 5 [84] Al-Inshiqaq 9 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 14 [83] Al-Mutaffifeen 31 [86] At-Tariq 11 [86] At-Tariq 8 [88] Al-Gashiyah 25 [89] Al-Fajr 28 [92] Al-Layl 19 [94] Al-Inshirah 3 [95] At-Teen 5 [96] Al-Alaq 8	[82] Al-Infitar 5 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 14 [84] Al-Inshiqaq 9 [94] Al-Inshirah 2 [94] Al-Inshirah 3	6	2	12	544
25.	do you know	[78] An-Naba 4 [78] An-Naba 5 [80] Abasa 3 [82] Al-Infitar 5 [81] At-Takweer 14 [82] Al-Infitar 12 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 5 [83] Al-Mutaffifeen 24 [83] Al-Mutaffifeen 8 [83] Al-Mutaffifeen 19 [84] Al-Inshiqaq 23 [87] Al-Alaa 7 [90] Al-Balad 12 [96] Al-Alaq 5 [96] Al-Alaq 14 [97] Al-Qadr 2 [100] Al-Adiyaat 9	[78] An-Naba 4 [78] An-Naba 5 [81] At-Takweer 14 [82] Al-Infitar 12 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 5 [83] Al-Mutaffifeen 19 [83] Al-Mutaffifeen 8 [84] Al-Inshiqaq 23 [87] Al-Alaa 7 [90] Al-Balad 12 [96] Al-Alaq 14 [96] Al-Alaq 5 [97] Al-Qadr 2 [100] Al-Adiyaat 9 [101] Al-Qoriah 10 [101] Al-Qoriah 3 [102] At-Takathur 3	20	0	4	540

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.32 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[101] Al-Qoriah10 [101] Al-Qoriah 3 [102] At-Takathur 3 [102] At-Takathur 4 [102] At-Takathur 5	[102] At-Takathur 4				
26.	refuse to believe	[78] An-Naba 28 [79] An-Nazi'at 21 [80] Abasa 42 [82] Al-Infitar 9 [83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 17 [84] Al-Inshiqaq 20 [84] Al-Inshiqaq 25 [91] Asy-Shams11 [91] Asy-Shams14 [95] At-Teen 7 [95] At-Teen 6 [96] Al-Alaq 13 [98] Al-Bayyinah 7 [107] Al-Maun 1 [107] Al-Maun 7	[84] Al-Inshiqaq 20 [84] Al-Inshiqaq 25 [95] At-Teen 6 [98] Al-Bayyinah 7 [107] Al-Maun 7	5	0	11	548
27.	righteous are in pleasure	[78] An-Naba 31 [78] An-Naba 38 [82] Al-Infitar 13 [83] Al-Mutaffifeen 22 [83] Al-Mutaffifeen 18 [83] Al-Mutaffifeen 24 [83] Al-Mutaffifeen 12 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 9 [84] Al-Inshiqaq 7 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 8 [90] Al-Balad 18 [95] At-Teen 6 [103] Al-Asr 3	[78] An-Naba 31 [78] An-Naba 38 [82] Al-Infitar 13 [83] Al-Mutaffifeen 12 [83] Al-Mutaffifeen 22 [84] Al-Inshiqaq 25 [84] Al-Inshiqaq 7 [85] Al-Burooj 11 [88] Al-Gashiyah 8 [90] Al-Balad 18 [95] At-Teen 6 [103] Al-Asr 3	12	0	4	548
28.	moses and pharaoh	[79] An-Nazi'at 15 [79] An-Nazi'at 21 [79] An-Nazi'at 17 [87] Al-Alaa 19 [89] Al-Fajr 10	[79] An-Nazi'at 15 [79] An-Nazi'at 21 [79] An-Nazi'at 17 [87] Al-Alaa 19	4	0	1	559

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.33 Reference Data, Output of Cosine Similarity Method and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
29.	Destruction	[78] An-Naba 40 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 10 [84] Al-Inshiqaq 11 [91] Asy-Shams14 [104] Al-Humazah 1 [107] Al-Maun 4	[91] Asy-Shams 14 [84] Al-Inshiqaq 11 [83] Al-Mutaffifeen 10 [78] An-Naba 40 [107] Al-Maun 4	5	0	2	557
30.	warning for whoever fear	[78] An-Naba 40 [79] An-Nazi'at 45 [79] An-Nazi'at 26 [79] An-Nazi'at 19 [79] An-Nazi'at 8 [79] An-Nazi'at 9 [79] An-Nazi'at 40 [80] Abasa 9 [80] Abasa 11 [83] Al-Mutaffifeen 18 [87] Al-Alaa 10 [88] Al-Gashiyah 2 [91] Asy-Shams15 [92] Al-Layl 14 [92] Al-Layl 5 [92] Al-Layl 17 [96] Al-Alaq 12 [98] Al-Bayyinah 8 [106] Al-Quraish 4	[78] An-Naba 40 [79] An-Nazi'at 45 [79] An-Nazi'at 26 [79] An-Nazi'at 19 [79] An-Nazi'at 8 [79] An-Nazi'at 9 [79] An-Nazi'at 40 [80] Abasa 9 [83] Al-Mutaffifeen 18 [87] Al-Alaa 10 [88] Al-Gashiyah 2 [91] Asy-Shams15 [92] Al-Layl 14 [92] Al-Layl 5 [92] Al-Layl 17 [96] Al-Alaq 12 [98] Al-Bayyinah 8 [106] Al-Quraish 4 [106] Al-Quraish 2	18	1	1	544

Furthermore, the comparison between the reference data with the Cosine Similarity method using a thesaurus. The comparison can be seen in table 4:27 and the values TP, FP, FN, TN are also determined.

Table 4.34 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
1.	Day	[78] An-Naba 11 [78] An-Naba 17 [78] An-Naba 18 [78] An-Naba 39 [78] An-Naba 40	[78] An-Naba 11 [78] An-Naba 17 [78] An-Naba 18 [78] An-Naba 39 [78] An-Naba 40	38	2	1	523

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.35 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[78] An-Naba 8 [79] An-Nazi'at 46 [79] An-Nazi'at 8 [79] An-Nazi'at 35 [79] An-Nazi'at 6 [80] Abasa 40 [80] Abasa 38 [80] Abasa 37 [80] Abasa 34 [81] At-Takweer 15 [82] Al-Infitar 19 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 15 [83] Al-Mutaffifeen 15 [83] Al-Mutaffifeen 5 [83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 11 [83] Al-Mutaffifeen 34 [85] Al-Burooj 2 [86] At-Tariq 9 [88] Al-Gashiyah 2 [88] Al-Gashiyah 8 [89] Al-Fajr 23 [89] Al-Fajr 25 [90] Al-Balad 14 [91] Asy-Shams3 [92] Al-Layl 2 [99] Al-Zalzalah 4 [99] Al-Zalzalah 6 [100] Al-Adiyaat 11 [101] Al-Qoriah 4 [102] At-Takathur 8 [108] Al-Kawthar 1	[78] An-Naba 8 [79] An-Nazi'at 46 [79] An-Nazi'at 8 [79] An-Nazi'at 35 [79] An-Nazi'at 6 [80] Abasa 40 [80] Abasa 38 [80] Abasa 37 [80] Abasa 34 [81] At-Takweer 15 [82] Al-Infitar 19 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 15 [83] Al-Mutaffifeen 5 [83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 11 [83] Al-Mutaffifeen 34 [83] Al-Mutaffifeen 6 [85] Al-Burooj 2 [86] At-Tariq 9 [88] Al-Gashiyah 2 [88] Al-Gashiyah 8 [89] Al-Fajr 1 [89] Al-Fajr 23 [89] Al-Fajr 25 [90] Al-Balad 14 [91] Asy-Shams3 [92] Al-Layl 2 [99] Al-Zalzalah 4 [99] Al-Zalzalah 6 [100] Al-Adiyaat 11 [101] Al-Qoriah4 [102] At-Takathur 8 [108] Al-Kawthar 1				
2.	Disbelive	[78] An-Naba 40 [79] An-Nazi'at 17 [80] Abasa 42 [83] Al-Mutaffifeen36 [86] At-Tariq 17 [88] Al-Gashiyah 23 [109] Al-Kafiroon 1	[78] An-Naba 40 [79] An-Nazi'at 17 [80] Abasa 42 [83] Al-Mutaffifeen36 [83] Al-Mutaffifeen 30 [86] At-Tariq 17 [88] Al-Gashiyah 23 [109] Al-Kafiroon 1	7	1	0	556

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.36 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
3.	Sin	[81] At-Takweer 9 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen12 [91] Asy-Shams14 [96] Al-Alaq 16 [82] Al-Infitar 14	[79] An-Nazi'at 17 [81] At-Takweer 9 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen12 [91] Asy-Shams14 [96] Al-Alaq 16 [82] Al-Infitar 14	6	1	0	557
4.	Rest	[78] An-Naba 9 [89] Al-Fajr 27	[78] An-Naba 9 [89] Al-Fajr 27	2	0	0	562
5.	Glory	[80] Abasa 13 [85] Al-Burooj 21 [85] Al-Burooj 15 [89] Al-Fajr 15 [92] Al-Layl 2	[80] Abasa 13 [85] Al-Burooj 21 [85] Al-Burooj 15 [89] Al-Fajr 17 [89] Al-Fajr 15 [92] Al-Layl 2	5	1	0	558
6.	Punishment	[78] An-Naba 30 [78] An-Naba 40 [79] An-Nazi'at 25 [84] Al-Inshiqaq 24 [85] Al-Burooj 10 [88] Al-Gashiyah 24 [89] Al-Fajr 25 [89] Al-Fajr 13	[78] An-Naba 30 [78] An-Naba 40 [79] An-Nazi'at 25 [84] Al-Inshiqaq 24 [85] Al-Burooj 10 [88] Al-Gashiyah 24 [89] Al-Fajr 25 [89] Al-Fajr 13	8	0	0	556
7.	Say	[78] An-Naba 37 [78] An-Naba 38 [78] An-Naba 40 [79] An-Nazi'at 18 [79] An-Nazi'at 24 [79] An-Nazi'at 43 [79] An-Nazi'at 12 [79] An-Nazi'at 10 [83] Al-Mutaffifeen 17 [83] Al-Mutaffifeen 32 [83] Al-Mutaffifeen 13 [89] Al-Fajr 27 [89] Al-Fajr 24 [89] Al-Fajr 16 [89] Al-Fajr 15 [90] Al-Balad 6 [91] Asy-Shams 13 [99] Al-Zalzalah 3 [109] Al-Kafiroon 1	[78] An-Naba 38 [78] An-Naba 40 [79] An-Nazi'at 43 [79] An-Nazi'at 12 [79] An-Nazi'at 10 [79] An-Nazi'at 46 [83] Al-Mutaffifeen 17 [83] Al-Mutaffifeen 32 [83] Al-Mutaffifeen 13 [89] Al-Fajr 27 [89] Al-Fajr 24 [89] Al-Fajr 16 [89] Al-Fajr 15 [90] Al-Balad 6 [91] Asy-Shams 13 [99] Al-Zalzalah 3 [109] Al-Kafiroon 1 [112] Al-Ikhlas 1 [113] Al-Falaq 1	19	1	3	541

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.37 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[112] Al-Ikhlas 1 [113] Al-Falaq 1 [114] An-Naas 1	[114] An-Naas 1				
8.	bad deeds	[78] An-Naba 40 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 9 [95] At-Teen 6 [99] Al-Zalzalah 8 [101] Al-Qoriah 8 [101] Al-Qoriah 6 [103] Al-Asr 3 [107] Al-Maun 6 [113] Al-Falaq 2 [113] Al-Falaq 5 [113] Al-Falaq 3 [113] Al-Falaq 4	[78] An-Naba 26 [78] An-Naba 8 [78] An-Naba 36 [78] An-Naba 40 [79] An-Nazi'at 18 [81] At-Takweer 10 [81] At-Takweer 14 [83] Al-Mutaffifeen 18 [95] At-Teen 6 [99] Al-Zalzalah 8 [99] Al-Zalzalah 6 [101] Al-Qoriah 8 [101] Al-Qoriah 6 [103] Al-Asr 3 [107] Al-Maun 6 [113] Al-Falaq 2	18	7	0	539
		[99] Al-Zalzalah 6 [114] An-Naas 4	[84] Al-Inshiqaq 23 [84] Al-Inshiqaq 6 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 9 [113] Al-Falaq 5 [113] Al-Falaq 3 [113] Al-Falaq 4 [114] An-Naas 4				
9.	allah swears	[78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 25 [80] Abasa 9 [81] At-Takweer 15 [81] At-Takweer 29 [82] Al-Infitar 19 [83] Al-Mutaffifeen 28 [84] Al-Inshiqaq 16 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [85] Al-Burooj 3 [85] Al-Burooj 20 [85] Al-Burooj	[78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 25 [80] Abasa 9 [81] At-Takweer 15 [81] At-Takweer 29 [82] Al-Infitar 19 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 16 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [85] Al-Burooj 3	50	4	0	510

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.38 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[85] Al-Burooj 8	[85] Al-Burooj 20				
		[85] Al-Burooj 9	[85] Al-Burooj 1				
		[85] Al-Burooj 15	[85] Al-Burooj 8				
		[87] Al-Alaa 10	[85] Al-Burooj 9				
		[87] Al-Alaa 7	[85] Al-Burooj 15				
		[88] Al-Gashiyah 24	[87] Al-Alaa 10				
		[89] Al-Fajr 5	[87] Al-Alaa 7				
		[89] Al-Fajr 2	[88] Al-Gashiyah 24				
		[89] Al-Fajr 3	[88] Al-Gashiyah 3				
		[89] Al-Fajr 11	[89] Al-Fajr 4				
		[90] Al-Balad 1	[89] Al-Fajr 11				
		[91] Asy-Shams 5	[90] Al-Balad 1				
		[91] Asy-Shams 6	[91] Asy-Shams 5				
		[91] Asy-Shams 2	[91] Asy-Shams 6				
		[91] Asy-Shams 15	[91] Asy-Shams 2				
		[91] Asy-Shams 4	[91] Asy-Shams 15				
		[91] Asy-Shams 13	[91] Asy-Shams 4				
		[92] Al-Layl 1	[91] Asy-Shams 13				
		[92] Al-Layl 5	[92] Al-Layl 1				
		[95] At-Teen 1	[92] Al-Layl 5				
		[95] At-Teen 8	[95] At-Teen 1				
		[95] At-Teen 2	[95] At-Teen 8				
		[96] Al-Alaq 14	[95] At-Teen 2				
		[96] Al-Alaq 19	[96] Al-Alaq 14				
		[96] Al-Alaq 12	[96] Al-Alaq 19				
		[103] Al-Asr 1	[96] Al-Alaq 12				
		[104] Al-Humazah	[103] Al-Asr 1				
		[106] Al-Quraish 3	[104] Al-Humazah 6				
		[106] Al-Quraish 2	[106] Al-Quraish 1				
		[110] An-Nasr 1	[106] Al-Quraish 3				
		[110] An-Nasr 2	[106] Al-Quraish 2				
		[112] Al-Ikhlas 2	[110] An-Nasr 1				
		[89] Al-Fajr 1	[89] Al-Fajr 5				
		[89] Al-Fajr 4	[89] Al-Fajr 2				
		[112] Al-Ikhlas 1	[89] Al-Fajr 3				
		[113] Al-Falaq 1	[89] Al-Fajr 1				
			[110] An-Nasr 2				
			[112] Al-Ikhlas 2				
			[112] Al-Ikhlas 1				
			[113] Al-Falaq 1				

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.39 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
10.	messenger from god	[78] An-Naba 36 [78] An-Naba 37 [79] An-Nazi'at 24 [79] An-Nazi'at 19 [79] An-Nazi'at 5 [79] An-Nazi'at 44 [79] An-Nazi'at 40 [80] Abasa 15 [81] At-Takweer 20 [81] At-Takweer 19 [81] At-Takweer 29 [82] Al-Infitar 6 [83] Al-Mutaffifeen 15 [83] Al-Mutaffifeen 6 [84] Al-Inshiqaq 15 [84] Al-Inshiqaq 2 [84] Al-Inshiqaq 5 [84] Al-Inshiqaq 6 [85] Al-Burooj 12 [87] Al-Alaa 1 [87] Al-Alaa 15 [89] Al-Fajr 15 [89] Al-Fajr 14 [89] Al-Fajr 13 [89] Al-Fajr 16 [89] Al-Fajr 22 [89] Al-Fajr 6 [89] Al-Fajr 28 [91] Asy-Shams13 [91] Asy-Shams14 [92] Al-Layl 20 [93] Ad-Duha 5 [93] Ad-Duha 3 [94] Al-Inshirah8 [96] Al-Alaq 1 [97] Al-Qadr 4 [98] Al-Bayyinah 2 [98] Al-Bayyinah 8 [100] Al-Adiyaat 11 [105] Al-Feel 1 [106] Al-Quraish 3 [108] Al-Kawthar 2	[78] An-Naba 36 [78] An-Naba 37 [79] An-Nazi'at 24 [79] An-Nazi'at 19 [79] An-Nazi'at 5 [79] An-Nazi'at 44 [79] An-Nazi'at 40 [80] Abasa 15 [81] At-Takweer 19 [81] At-Takweer 29 [82] Al-Infitar 6 [83] Al-Mutaffifeen 15 [83] Al-Mutaffifeen 6 [84] Al-Inshiqaq 15 [84] Al-Inshiqaq 2 [84] Al-Inshiqaq 5 [84] Al-Inshiqaq 6 [87] Al-Alaa 1 [87] Al-Alaa 15 [89] Al-Fajr 15 [89] Al-Fajr 14 [89] Al-Fajr 13 [89] Al-Fajr 16 [89] Al-Fajr 22 [91] Asy-Shams13 [91] Asy-Shams14 [92] Al-Layl 20 [93] Ad-Duha 5 [93] Ad-Duha 3 [94] Al-Inshirah8 [96] Al-Alaq 12 [96] Al-Alaq 1 [97] Al-Qadr 4 [98] Al-Bayyinah 2 [98] Al-Bayyinah 8 [98] Al-Bayyinah 5 [100] Al-Adiyaat 11 [105] Al-Feel 1 [106] Al-Quraish 3 [108] Al-Kawthar 2 [113] Al-Falaq 1 [114] An-Naas 3	43	2	4	515

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.40 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[113] Al-Falaq 1 [114] An-Naas 3 [114] An-Naas 1 [79] An-Nazi'at 16 [78] An-Naba 39	[114] An-Naas 1 [79] An-Nazi'at 16 [78] An-Naba 39				
11.	night of decree	[78] An-Naba 10 [79] An-Nazi'at 29 [80] Abasa 41 [81] At-Takweer 17 [84] Al-Inshiqaq 17 [89] Al-Fajr 2 [89] Al-Fajr 4 [91] Asy-Shams4 [92] Al-Layl 1 [93] Ad-Duha 2 [97] Al-Qadr 1 [97] Al-Qadr 2 [97] Al-Qadr 3 [113] Al-Falaq 3	[78] An-Naba 10 [79] An-Nazi'at 29 [80] Abasa 41 [81] At-Takweer 17 [81] At-Takweer 15 [84] Al-Inshiqaq 17 [89] Al-Fajr 2 [89] Al-Fajr 4 [91] Asy-Shams4 [92] Al-Layl 1 [93] Ad-Duha 2 [97] Al-Qadr 1 [97] Al-Qadr 2 [97] Al-Qadr 3 [113] Al-Falaq 3	14	1	1	548
12.	allah taught man	[78] An-Naba 40 [78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 35 [80] Abasa 24 [80] Abasa 23 [80] Abasa 2 [80] Abasa 17 [80] Abasa 34 [81] At-Takweer 29 [81] At-Takweer 27 [82] Al-Infitar 6 [82] Al-Infitar 19 [83] Al-Mutaffifeen 6 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [84] Al-Inshiqaq 23 [84] Al-Inshiqaq 6 [85] Al-Burooj 20 [85] Al-Burooj 8 [90] Al-Balad 4 [91] Asy-Shams15	[78] An-Naba 40 [78] An-Naba 38 [79] An-Nazi'at 26 [79] An-Nazi'at 35 [79] An-Nazi'at 25 [80] Abasa 24 [80] Abasa 9 [80] Abasa 23 [80] Abasa 2 [80] Abasa 17 [80] Abasa 34 [81] At-Takweer 29 [82] Al-Infitar 6 [82] Al-Infitar 19 [83] Al-Mutaffifeen 6 [83] Al-Mutaffifeen 21 [83] Al-Mutaffifeen 28 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 21 [84] Al-Inshiqaq 23 [89] Al-Fajr 15 [90] Al-Balad 4	54	9	1	500

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.41 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[91] Asy-Shams13	[91] Asy-Shams15				
		[92] Al-Layl 5	[91] Asy-Shams13				
		[95] At-Teen 4	[92] Al-Layl 5				
		[95] At-Teen 6	[95] At-Teen 4				
		[96] Al-Alaq 5	[95] At-Teen 8				
		[96] Al-Alaq 4	[96] Al-Alaq 5				
		[96] Al-Alaq 6	[96] Al-Alaq 4				
		[96] Al-Alaq 2	[96] Al-Alaq 6				
		[96] Al-Alaq 14	[96] Al-Alaq 2				
		[96] Al-Alaq 19	[96] Al-Alaq 14				
		[96] Al-Alaq 12	[96] Al-Alaq 19				
		[99] Al-Zalzalah 6	[96] Al-Alaq 12				
		[99] Al-Zalzalah 3	[99] Al-Zalzalah 6				
		[100] Al-Adiyaat 6	[99] Al-Zalzalah 3				
		[103] Al-Asr 2	[100] Al-Adiyaat 6				
		[104] Al-Humazah6	[103] Al-Asr 2				
		[110] An-Nasr 1	[104] Al-Humazah6				
		[110] An-Nasr 2	[106] Al-Quraish 1				
		[85] Al-Burooj 9	[84] Al-Inshiqaq 6				
		[85] Al-Burooj 15	[85] Al-Burooj 20				
		[86] At-Tariq 5	[85] Al-Burooj 8				
		[86] At-Tariq 10	[85] Al-Burooj 9				
		[87] Al-Alaa 10	[85] Al-Burooj 15				
		[87] Al-Alaa 7	[86] At-Tariq 5				
		[88] Al-Gashiyah 24	[86] At-Tariq 10				
		[89] Al-Fajr 15	[87] Al-Alaa 10				
		[112] Al-Ikhlas 2	[87] Al-Alaa 7				
		[112] Al-Ikhlas 1	[88] Al-Gashiyah 24				
		[113] Al-Falaq 1	[88] Al-Gashiyah 3				
		[114] An-Naas 3	[89] Al-Fajr 11				
		[114] An-Naas 2	[106] Al-Quraish 3				
		[114] An-Naas 6	[106] Al-Quraish 2				
		[114] An-Naas 1	[110] An-Nasr 1				
			[110] An-Nasr 2				
			[112] Al-Ikhlas 2				
			[112] Al-Ikhlas 1				
			[113] Al-Falaq 1				
			[114] An-Naas 3				
			[114] An-Naas 2				
			[114] An-Naas 6				
			[114] An-Naas 1				

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.42 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
13.	see creation	[78] An-Naba 8 [78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [80] Abasa 24 [84] Al-Inshiqaq 15 [85] Al-Burooj 13 [85] Al-Burooj 15 [86] At-Tariq 5 [88] Al-Gashiyah 17 [89] Al-Fajr 6 [96] Al-Alaq 7 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8 [99] Al-Zalzalah 6 [102] At-Takathur 7 [102] At-Takathur 6 [110] An-Nasr 2	[78] An-Naba 40 [79] An-Nazi'at 27 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [80] Abasa 24 [84] Al-Inshiqaq 15 [85] Al-Burooj 13 [85] Al-Burooj 15 [86] At-Tariq 5 [88] Al-Gashiyah 17 [89] Al-Fajr 6 [96] Al-Alaq 7 [99] Al-Zalzalah 7 [99] Al-Zalzalah 8 [99] Al-Zalzalah 6 [102] At-Takathur 7 [102] At-Takathur 6 [110] An-Nasr 2	16	2	13	533
14.	when quran recited	[81] At-Takweer 25 [83] Al-Mutaffifeen 13 [84] Al-Inshiqaq 21 [85] Al-Burooj 21 [87] Al-Alaa 6 [96] Al-Alaq 1 [96] Al-Alaq [98] Al-Bayyinah	[81] At-Takweer [84] Al-Inshiqaq [84] Al-Inshiqaq [85] Al-Burooj [87] Al-Alaa [96] Al-Alaq [96] Al-Alaq [98] Al-Bayyinah	7	1	1	555
15.	wealth is not eternal	[78] An-Naba 23 [78] An-Naba 14 [89] Al-Fajr 20 [90] Al-Balad 6 [92] Al-Layl 18	[78] An-Naba 23 [78] An-Naba 14 [89] Al-Fajr 20 [90] Al-Balad 6 [92] Al-Layl 18	10	0	1	553

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.43 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[92] Al-Layl 11 [95] At-Teen 6 [100] Al-Adiyaat 8 [104] Al-Humazah 2 [108] Al-Kawthar 1 [111] Al-Lahab 2 [104] Al-Humazah3	[92] Al-Layl 11 [100] Al-Adiyaat 8 [104] Al-Humazah 2 [104] Al-Humazah 3 [108] Al-Kawthar 1 [111] Al-Lahab 2				
16.	moon follows sun	[79] An-Nazi'at 7 [81] At-Takweer 1 [84] Al-Inshiqaq 18 [91] Asy-Shams 12 [91] Asy-Shams 1 [91] Asy-Shams 12 [91] Asy-Shams 4	[79] An-Nazi'at 7 [81] At-Takweer 1 [84] Al-Inshiqaq 18 [91] Asy-Shams 1 [91] Asy-Shams 12 [91] Asy-Shams 4	6	0	1	557
17.	advice patience and kindness	[90] Al-Balad 17 [103] Al-Asr 3	[89] Al-Fajr [90] Al-Balad [103] Al-Asr [103] Al-Asr [104] Al-Humazah	2	3	0	559
18.	enter my heaven	[78] An-Naba 31 [79] An-Nazi'at 41 [81] At-Takweer 13 [81] At-Takweer 21 [82] Al-Infitar 15 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12 [87] Al-Alaa 12 [88] Al-Gashiyah 4 [89] Al-Fajr 30 [89] Al-Fajr 29 [98] Al-Bayyinah 8 [101] Al-Qoriah 7 [108] Al-Kawthar 1 [110] An-Nasr 2 [111] Al-Lahab 3	[78] An-Naba 31 [78] An-Naba 12 [78] An-Naba 19 [78] An-Naba 37 [79] An-Nazi'at 41 [81] At-Takweer 11 [81] At-Takweer 13 [81] At-Takweer 21 [81] At-Takweer 25 [82] Al-Infitar 15 [85] Al-Burooj 9 [86] At-Tariq 1 [87] Al-Alaa 12 [88] Al-Gashiyah 4 [89] Al-Fajr 30 [89] Al-Fajr 29 [92] Al-Layl 21 [98] Al-Bayyinah 8 [101] Al-Qoriah 7 [108] Al-Kawthar 1 [82] Al-Infitar 1 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12 [110] An-Nasr 2 [111] Al-Lahab 3	16	9	0	539

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.44 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
19.	the greatest punishment	[78] An-Naba 30 [78] An-Naba 2 [78] An-Naba 26 [78] An-Naba 40 [79] An-Nazi'at 25 [79] An-Nazi'at 20 [84] Al-Inshiqaq 24 [85] Al-Burooj 10 [85] Al-Burooj 11 [87] Al-Alaa 12 [88] Al-Gashiyah24 [89] Al-Fajr 25 [89] Al-Fajr 13	[78] An-Naba 30 [78] An-Naba 2 [78] An-Naba 40 [79] An-Nazi'at 25 [79] An-Nazi'at 20 [79] An-Nazi'at 9 [84] Al-Inshiqaq 24 [85] Al-Burooj 10 [85] Al-Burooj 11 [87] Al-Alaa 12 [88] Al-Gashiyah24 [89] Al-Fajr 25 [89] Al-Fajr 13 [106] Al-Quraish 1	12	2	1	549
20.	give the book	[78] An-Naba 29 [80] Abasa 6 [80] Abasa 13 [82] Al-Infitar 11 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen 9 [83] Al-Mutaffifeen20 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen 3 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 7 [84] Al-Inshiqaq 24 [86] At-Tariq 17 [89] Al-Fajr 15 [92] Al-Layl 12 [92] Al-Layl 5 [92] Al-Layl 18 [93] Ad-Duha 5 [98] Al-Bayyinah 5 [108] Al-Kawthar 1 [112] Al-Ikhlas 3	[78] An-Naba 29 [80] Abasa 6 [80] Abasa 13 [82] Al-Infitar 11 [83] Al-Mutaffifeen 9 [83] Al-Mutaffifeen 20 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 7 [83] Al-Mutaffifeen 3 [83] Al-Mutaffifeen 18 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 7 [84] Al-Inshiqaq 24 [86] At-Tariq 17 [89] Al-Fajr 15 [92] Al-Layl 15 [92] Al-Layl 12 [92] Al-Layl 5 [92] Al-Layl 18 [93] Ad-Duha 5 [98] Al-Bayyinah 5 [104] Al-Humazah5 [112] Al-Ikhlas 3	21	2	2	539
21.	prefer the present live	[78] An-Naba 39 [79] An-Nazi'at 38 [79] An-Nazi'at 10[81] At-Takweer 27	[78] An-Naba 39 [79] An-Nazi'at 38 [79] An-Nazi'at 10 [79] An-Nazi'at 46	8	6	1	549

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.45 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[83] Al-Mutaffifeen 6 [87] Al-Alaa 16 [89] Al-Fajr 24 [101] Al-Qoriah7 [102] At-Takathur 1	[81] At-Takweer 29 [81] At-Takweer 19 [81] At-Takweer 25 [81] At-Takweer 27 [83] Al-Mutaffifeen 6 [86] At-Tariq 13 [87] Al-Alaa 16 [88] Al-Gashiyah 3 [101] Al-Qoriah7 [102] At-Takathur 1				
22.	there is no power	[81] At-Takweer 20 [82] Al-Infitar 19 [86] At-Tariq 10 [90] Al-Fajr 5 [110] An-Nasr 1	[86] At-Tariq 10 [110] An-Nasr 1 [82] Al-Infitar 19 [89] Al-Fajr 5 [81] At-Takweer 20	4	1	1	558
23.	hell fuel	[78] An-Naba 21 [79] An-Nazi'at 36 [81] At-Takweer 6 [81] At-Takweer 12 [82] Al-Infitar 14 [83] Al-Mutaffifeen 16 [84] Al-Inshiqaq 12 [85] Al-Burooj 5 [85] Al-Burooj 10 [87] Al-Alaa 12 [88] Al-Gashiyah 4 [89] Al-Fajr 23 [90] Al-Balad 20 [92] Al-Layl 14 [98] Al-Bayyinah 6 [101] Al-Qoriah11 [101] Al-Qoriah9 [102] At-Takathur 6 [104] Al-Humazah6 [111] Al-Lahab 3	[78] An-Naba 21 [79] An-Nazi'at 36 [79] An-Nazi'at 46 [81] At-Takweer 6 [81] At-Takweer 12 [84] Al-Inshiqaq 12 [85] Al-Burooj 5 [85] Al-Burooj 10 [86] At-Tariq 3 [87] Al-Alaa 12 [88] Al-Gashiyah 4 [89] Al-Fajr 23 [90] Al-Balad 20 [92] Al-Layl 14 [92] Al-Layl 11 [98] Al-Bayyinah 6 [101] Al-Qoriah11 [101] Al-Qoriah9 [102] At-Takathur 6 [104] Al-Humazah6 [111] Al-Lahab 3	21	3	2	538
24.	back to family	[79] An-Nazi'at 12 [79] An-Nazi'at 22 [79] An-Nazi'at 44 [79] An-Nazi'at 10 [82] Al-Infitar 5	[79] An-Nazi'at 12 [79] An-Nazi'at 44 [79] An-Nazi'at 10 [82] Al-Infitar 5 [83] Al-Mutaffifeen 3	17	2	1	544

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.46 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[84] Al-Inshiqaq 9 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 14 [83] Al-Mutaffifeen 31 [86] At-Tariq 11 [86] At-Tariq 8 [88] Al-Gashiyah 25 [89] Al-Fajr 28 [92] Al-Layl 19 [94] Al-Inshirah 3 [95] At-Teen 5 [96] Al-Alaq 8	[84] Al-Inshiqaq 9 [84] Al-Inshiqaq 10 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 14 [84] Al-Inshiqaq 6 [86] At-Tariq 11 [86] At-Tariq 8 [88] Al-Gashiyah 25 [89] Al-Fajr 28 [92] Al-Layl 19 [94] Al-Inshirah3 [94] Al-Inshirah2 [95] At-Teen 5 [96] Al-Alaq 8				
25.	do you know	[78] An-Naba 4 [78] An-Naba 5 [80] Abasa 3 [82] Al-Infitar 5 [81] At-Takweer 14 [82] Al-Infitar 12 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 5 [83] Al-Mutaffifeen 24 [83] Al-Mutaffifeen 8 [83] Al-Mutaffifeen 19 [84] Al-Inshiqaq 23 [87] Al-Alaa 7 [90] Al-Balad 12 [96] Al-Alaq 5 [96] Al-Alaq 14 [97] Al-Qadr 2 [100] Al-Adiyaat 9 [101] Al-Qoriah 10 [101] Al-Qoriah 3 [102] At-Takathur 3 [102] At-Takathur 4 [102] At-Takathur 5	[78] An-Naba 4 [78] An-Naba 5 [81] At-Takweer 14 [82] Al-Infitar 12 [82] Al-Infitar 17 [82] Al-Infitar 18 [82] Al-Infitar 5 [83] Al-Mutaffifeen 24 [83] Al-Mutaffifeen 8 [83] Al-Mutaffifeen 19 [84] Al-Inshiqaq 23 [87] Al-Alaa 7 [90] Al-Balad 12 [96] Al-Alaq 5 [96] Al-Alaq 14 [97] Al-Qadr 2 [100] Al-Adiyaat 9 [101] Al-Qoriah10 [101] Al-Qoriah 3 [102] At-Takathur 3 [102] At-Takathur 4	21	0	3	540
26.	refuse to believe	[78] An-Naba 28 [79] An-Nazi'at 21 [80] Abasa 42 [82] Al-Infitar 9	[80] Abasa 42 [82] Al-Infitar 9 [83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 17 [84] Al-Inshiqaq 20	12	0	4	548

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.47 Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
		[83] Al-Mutaffifeen 10 [83] Al-Mutaffifeen 17 [84] Al-Inshiqaq 20 [84] Al-Inshiqaq 25 [91] Asy-Shams11 [91] Asy-Shams14 [95] At-Teen 7 [95] At-Teen 6 [96] Al-Alaq 13 [98] Al-Bayyinah 7 [107] Al-Maun 1 [107] Al-Maun 7	[84] Al-Inshiqaq 25 [91] Asy-Shams11 [91] Asy-Shams14 [95] At-Teen 7 [95] At-Teen 6 [98] Al-Bayyinah 7 [107] Al-Maun 7				
27.	righteous are in pleasure	[78] An-Naba 31 [78] An-Naba 38 [82] Al-Infitar 13 [83] Al-Mutaffifeen 22 [83] Al-Mutaffifeen 18 [83] Al-Mutaffifeen 24 [83] Al-Mutaffifeen 12 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 9 [84] Al-Inshiqaq 7 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 8 [90] Al-Balad 18 [95] At-Teen 6 [103] Al-Asr 3	[78] An-Naba 31 [78] An-Naba 38 [82] Al-Infitar 13 [83] Al-Mutaffifeen 22 [83] Al-Mutaffifeen 12 [84] Al-Inshiqaq 13 [84] Al-Inshiqaq 9 [84] Al-Inshiqaq 7 [84] Al-Inshiqaq 25 [85] Al-Burooj 11 [88] Al-Gashiyah 8 [90] Al-Balad 18 [95] At-Teen 6 [103] Al-Asr 3	14	0	2	548
28.	moses and pharaoh	[79] An-Nazi'at 15 [79] An-Nazi'at 21 [79] An-Nazi'at 17 [87] Al-Alaa 19 [89] Al-Fajr 10	[79] An-Nazi'at 15 [79] An-Nazi'at 21 [79] An-Nazi'at 17 [87] Al-Alaa 19	4	0	1	559
29.	destruction	[78] An-Naba 40 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 10 [84] Al-Inshiqaq 11 [91] Asy-Shams14 [104] Al-Humazah 1 [107] Al-Maun 4	[78] An-Naba 40 [83] Al-Mutaffifeen 1 [83] Al-Mutaffifeen 10 [84] Al-Inshiqaq 11 [91] Asy-Shams14 [104] Al-Humazah 1 [107] Al-Maun 4	7	0	0	557
30.	warning for	[78] An-Naba 40 [79] An-Nazi'at 45	[78] An-Naba 40 [79] An-Nazi'at 45	18	1	1	544

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Reference Data, Output of Cosine Similarity Method using Query Expansion and Value of TP, FP, FN, TN (more)

No.	Queries	Reference Data	System Output	TP	FP	FN	TN
	whoever fear	[79] An-Nazi'at 26 [79] An-Nazi'at 19 [79] An-Nazi'at 8 [79] An-Nazi'at 9 [79] An-Nazi'at 40 [80] Abasa 9 [80] Abasa 11 [83] Al-Mutaffifeen 18 [87] Al-Alaa 10 [88] Al-Gashiyah 2 [91] Asy-Shams15 [92] Al-Layl 14 [92] Al-Layl 5 [92] Al-Layl 17 [96] Al-Alaq 12 [98] Al-Bayyinah 8 [106] Al-Quraish 4	[79] An-Nazi'at 26 [79] An-Nazi'at 19 [79] An-Nazi'at 8 [79] An-Nazi'at 9 [79] An-Nazi'at 40 [80] Abasa 9 [83] Al-Mutaffifeen 18 [87] Al-Alaa 10 [88] Al-Gashiyah 2 [91] Asy-Shams15 [92] Al-Layl 14 [92] Al-Layl 5 [92] Al-Layl 17 [96] Al-Alaq 12 [98] Al-Bayyinah 8 [106] Al-Quraish 4 [106] Al-Quraish 2				

from the two configurations above, the first configuration only with Cosine Similarity produces precision, recall, F-Measure values for each query as follows:

Table 4.48 Value of Precision, Recall, F-Measure and Accuracy of Cosine Similarity Experiment without Query Expansion

No.	Queries	Precision	Recall	F-Measure	Accuracy
1	Day	95,00	97,44	96,20	99,47
2	Disbeliever	87,50	100,00	93,33	99,82
3	Sin	80,00	66,67	72,73	99,47
4	Rest	100,00	100,00	100,00	100,00
5	Glory	66,67	33,33	44,44	99,11
6	Punishment	100,00	62,50	76,92	99,47
7	Say	92,86	59,09	72,22	98,23
8	bad deeds	63,16	66,67	64,86	97,70
9	allah swears	92,45	98,00	95,15	99,11
10	messenger from god	62,50	10,87	18,52	92,20
11	night of decree	92,86	92,86	92,86	99,65
12	allah taught man	83,64	83,64	83,64	96,81
13	see creation	87,50	48,28	62,22	96,99
14	when quran recited	87,50	87,50	87,50	99,65

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.49 Value of Precision, Recall, F-Measure and Accuracy of Cosine Similarity Experiment without Query Expansion (more)

15	wealth is not eternal	100,00	72,73	84,21	99,47
16	moon follows sun	100,00	75,00	85,71	99,65
17	advice patience and kindness	40,00	100,00	57,14	99,47
18	enter my heaven	52,63	62,50	57,14	97,34
19	the greatest punishment	81,82	69,23	75,00	98,94
20	give the book	88,24	65,22	75,00	98,23
21	prefer the present live	66,67	36,36	47,06	98,40
22	there is no power	75,00	60,00	66,67	99,47
23	hell fuel	70,00	30,43	42,42	96,63
24	back to family	75,00	33,33	46,15	97,52
25	do you know	100,00	83,33	90,91	99,29
26	refuse to believe	100,00	31,25	47,62	98,05
27	righteous are in pleasure	100,00	75,00	85,71	99,29
28	moses and pharaoh	100,00	80,00	88,89	99,82
29	Destruction	100,00	71,43	83,33	99,65
30	warning for whoever fear	94,74	94,74	94,74	99,65
	AVERAGE	**84,52**	**68,25**	**72,94**	**98,62**

The results of the value of precision, recall, F-Measure with Cosine Similarity that apply query expansion are as follows:

Table 4.50 Value of Precision, Recall, F-Measure and Accuracy of Cosine Similarity Experiment using Query Expansion

No.	Queries	Precision	Recall	F-Measure	Accuracy
1	Day	95,00	97,44	96,20	99,47
2	Disbeliever	87,50	100,00	93,33	99,82
3	Sin	85,71	100,00	92,31	99,82
4	Rest	100,00	100,00	100,00	100,00
5	Glory	83,33	100,00	90,91	99,82
6	Punishment	100,00	100,00	100,00	100,00
7	Say	95,00	86,36	90,48	99,29
8	bad deeds	72,00	100,00	83,72	98,76
9	allah swears	92,59	100,00	96,15	99,29
10	messenger from god	95,56	91,49	93,48	98,94
11	night of decree	93,33	93,33	93,33	99,65
12	allah taught man	85,71	98,18	91,53	98,23
13	see creation	88,89	55,17	68,09	97,34
14	when quran recited	87,50	87,50	87,50	99,65

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Table 4.51 Value of Precision, Recall, F-Measure and Accuracy of Cosine Similarity Experiment without Query Expansion (more)

15	wealth is not eternal	100,00	90,91	95,24	99,82
16	moon follows sun	100,00	85,71	92,31	99,82
17	advice patience and kindness	40,00	100,00	57,14	99,47
18	enter my heaven	64,00	100,00	78,05	98,40
19	the greatest punishment	85,71	92,31	88,89	99,47
20	give the book	91,30	91,30	91,30	99,29
21	prefer the present live	57,14	88,89	69,57	98,76
22	there is no power	80,00	80,00	80,00	99,65
23	hell fuel	87,50	91,30	89,36	99,11
24	back to family	89,47	94,44	91,89	99,47
25	do you know	100,00	87,50	93,33	99,47
26	refuse to believe	100,00	75,00	85,71	99,29
27	righteous are in pleasure	100,00	87,50	93,33	99,65
28	moses and pharaoh	100,00	80,00	88,89	99,82
29	Destruction	100,00	100,00	100,00	100,00
30	warning for whoever fear	94,74	94,74	94,74	99,65
	AVERAGE	**91,64**	**88,40**	**88,89**	**99,37**

4.2 Discussion

Based on sub-chapter 4.2, we can see the Experiment Results in table 4.38 for the application of the Cosine Similarity method without Query Expansion and in table 4.39 for the application of the Cosine Similarity method using Query Expansion. The experiment results are in the form of precision, recall, F-Measure and accuracy values. The experiment results show that using Query Expansion in the form of a thesaurus dictionary gives more results in each experiment, it can be seen without using Query Expansion, Cosine Similarity is able to produce a precision value of 84.52%, recall of 68.25%, weight f-measure of 72.94% and an accuracy of 98.62%. Even though Cosine Similarity itself has produced a high value, using Query Expansion can give a higher value, it can be seen using Query Expansion, Cosine Similarity can produce a precision

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

value of 91.64%, a recall of 88.40%, weight f-measure of 88.89% and an accuracy of 99.37%. The following is a graph of the comparison of the F-Measure values for each query.

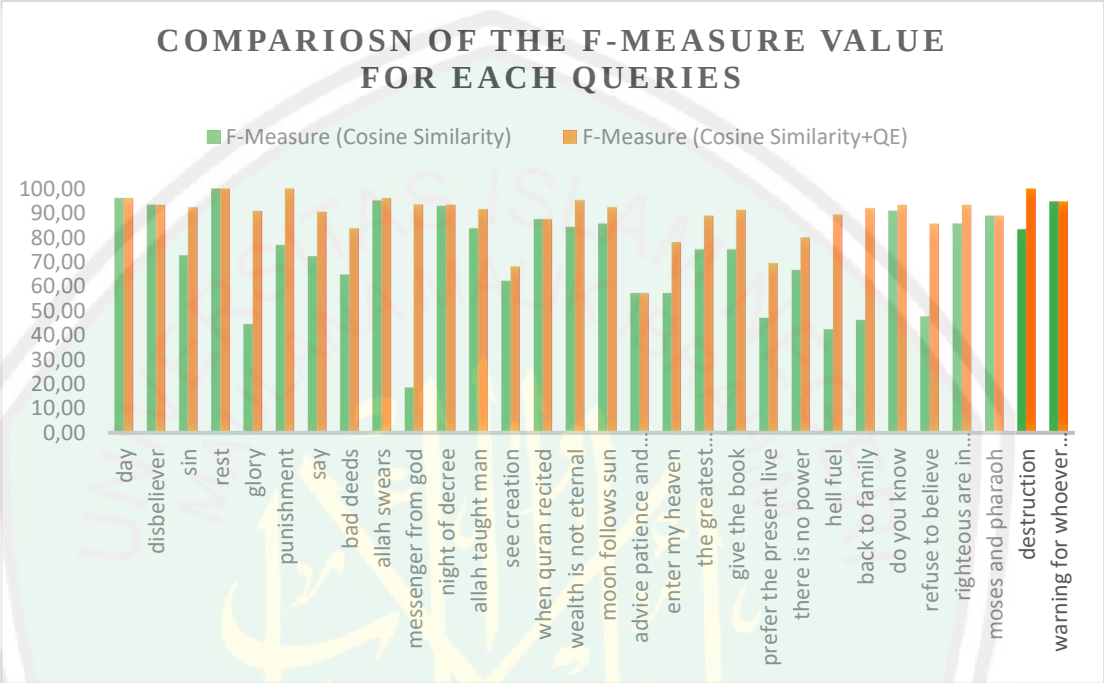

Figure 4.1 Comparison of the F-Measure Value for each Queries

In the graph above, it can be seen that the height of the F-Measure value using Cosine similarity + Query Expansion which is red tends to be higher than the use of Cosine similarity only (green color). In the graph, we can also see that some queries have significant differences with the use of the Expansion Query, as we can see in the "messanger from god" query, this is because the word "god" thesaurus is "lord" so widely used in research documents (juz. 30) so that with the addition of a thesaurus, the TP value is increasing and the FN is decreasing, as well as the "hell fuel" and "glory" queries. This causes the recall value to increase, so that as the recall value increases, the f-measure also increases. In addition, there are those that do not have a

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

difference between using Query Expansion and without using it because the test query does not have a thesaurus in the WordNet dictionary or has a thesaurus whose terms are not used in the research document. From the visualization and comparison of the table above, it can be said that the use of Query Expansion can increase the f-measure value. And here is a comparison of the accuracy value of each queries.

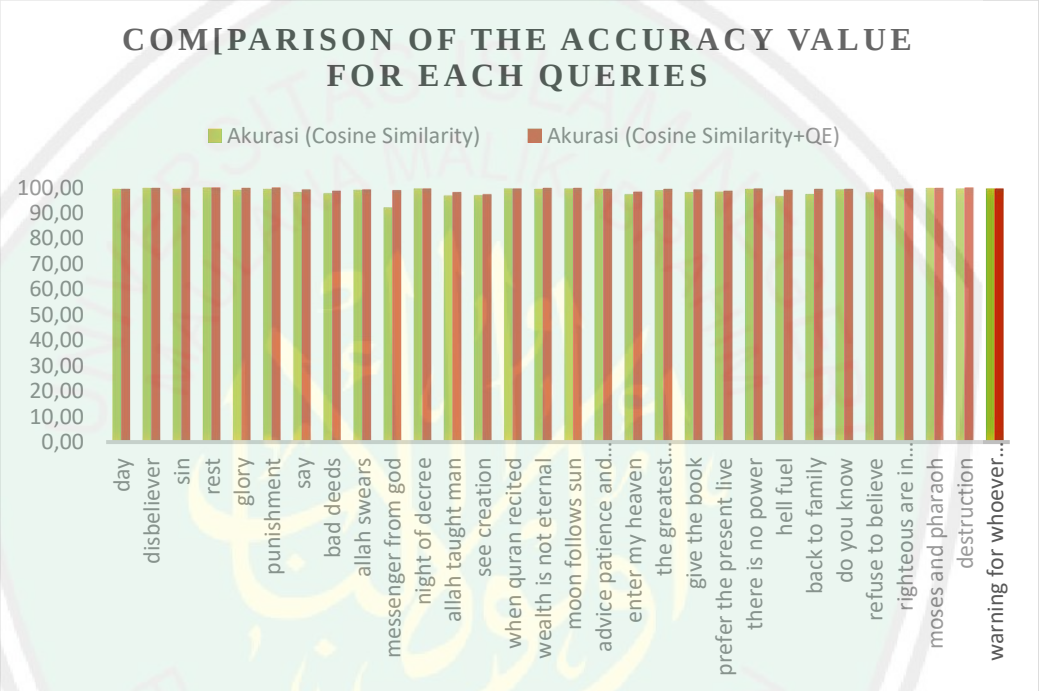

Figure 4.2 Comparison of the Accuracy Value for each Queries

In the graph above, it can also be seen that the height of the accuracy value using Cosine Similarity + Query Expansion which is red tends to be higher than the use of Cosine Similarity alone (green color). The comparison is not too significant, because using Cosine Similarity alone has resulted in high accuracy values, but using Query Expansion is able to produce higher values even though there is a slight difference. So it can be said that the use of Query Expansion can also increase the accuracy value.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Based on the two graphs in Figures 4.1 and 4.2, it can be seen which experiment values have increased or not, and the following is an analysis for each query related to an increase in the value of the evaluation test.

Table 4.52 Analysis of Testing Value Comparison

No.	Query	Analysis of Testing Value Comparison
1	Day	Equal, because there is no addition of a thesaurus
2	Disbeliever	Equal, because there are no documents that have the same meaning as the thesaurus
3	Sin	Increases, by the addition of a thesaurus
4	Rest	Equal, because there are no documents that have the same meaning as the thesaurus
5	Glory	Increases, by the addition of a thesaurus
6	Punishment	Increases, by the addition of a thesaurus
7	Say	Increases, by the addition of a thesaurus
8	bad deeds	Increases, by the addition of a thesaurus
9	allah swears	Increases, by the addition of a thesaurus
10	messenger from god	Increases, by the addition of a thesaurus
11	night of decree	Equal, because there are no documents that have the same meaning as the thesaurus
12	allah taught man	Increases, by the addition of a thesaurus
13	see creation	Increases, by the addition of a thesaurus
14	when quran recited	Equal, because there are no documents that have the same meaning as the thesaurus
15	wealth is not eternal	Increases, by the addition of a thesaurus
16	moon follows sun	Equal, because there are no documents that have the same meaning as the thesaurus
17	advice patience and kindness	Equal, because there are no documents that have the same meaning as the thesaurus
18	enter my heaven	Increases, by the addition of a thesaurus
19	the greatest punishment	Increases, by the addition of a thesaurus
20	give the book	Increases, by the addition of a thesaurus
21	prefer the present live	Increases, by the addition of a thesaurus
22	there is no power	Increases, by the addition of a thesaurus
23	hell fuel	Increases, by the addition of a thesaurus
24	back to family	Increases, by the addition of a thesaurus
25	do you know	Increases, by the addition of a thesaurus
26	refuse to believe	Increases, by the addition of a thesaurus
27	righteous are in pleasure	Increases, by the addition of a thesaurus
28	moses and pharaoh	Equal, because there is no addition of a thesaurus
29	Destruction	Increases, by the addition of a thesaurus
30	warning for whoever fear	Equal, because there are no documents that have the same meaning as the thesaurus

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

From the results of the analysis in the table above, a graph can be visualized to compare the two configurations. The following is a graph of the value of precision, recall, F-Measure and accuracy of each configuration.

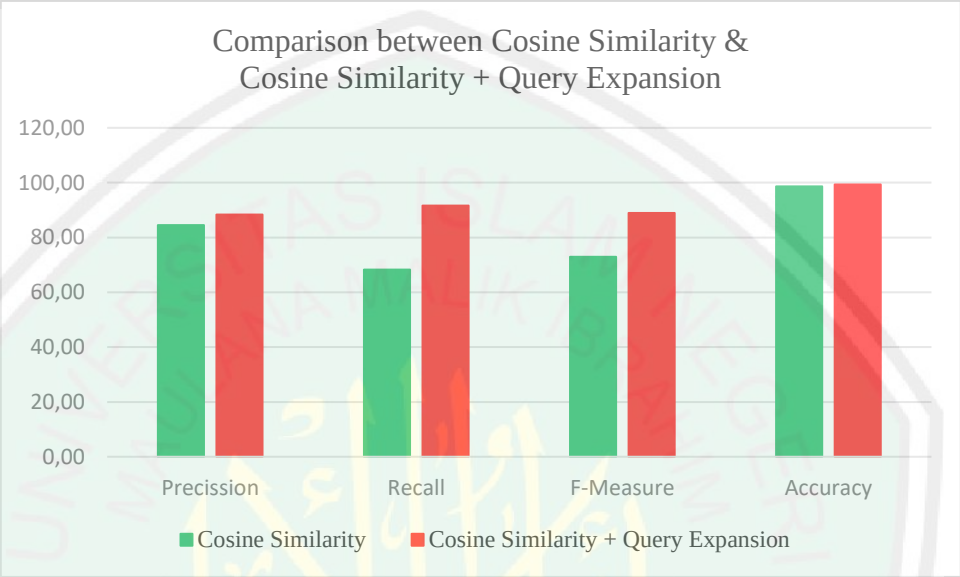

Figure 4.3 Comparison between Cosine Similarity & Cosine Similarity + Query Expansion

From this graph, it can be concluded that Cosine Similarity using Query Expansion in the form of a thesaurus is more effective in the search application for the English Al-Quran translation in this study. We can see the difference in evaluation test values in the table below.

Table 4.53 Difference of Testing Value

Precision	Recall	F-Measure	Accuracy
4,10%	23,35%	16,20%	0,73%

We can see that the increase in the recall value is the highest, this is because the use of Query Expansion in the form of a thesaurus finds more lexically and semantically relevant documents, in text and meaning or synonym, so that the TP value (i.e. the

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

correct document) increases and the value FN (i.e. relevant documents not found) are decrease, so that the difference in recall is the highest. Finally, it can be concluded that the use of Query Expansion in Cosine Similarity can improve performance in retrieving relevant documents.

Al-Quran is the main source for the Follower of Islam. This source includes all universal things, from worship, life, history, law, economy, social etc. Therefore, as a Muslim, we must stick to the Al-Quran and practice it in life. Al-Quran itself was revealed in Arabic, the language used by Arab countries and some Middle East. With the Al-Quran revealed in Arabic, a Muslim who does not understand Arabic will have difficulty in understanding it. One way to understand the Al-Quran for non-Arab people is to ask the experts. Allah SWT said:

وَمَا أَرْسَلْنَا مِنْ قَبْلِكَ إِلَّا رِجَالًا نُوحِي إِلَيْهِمْ ۚ فَاسْأَلُوا أَهْلَ الذِّكْرِ إِنْ كُنْتُمْ لَا تَعْلَمُونَ

Mean: And We did not send before you, except men whom We gave revelations to them; then ask people who know if you don't know. (Surah An-Nahl: 43).

In the Tafseer Al-Mukhtashar, this verse is generally applicable in all religious matters. If we don't know something about religion, then let's ask knowledgeable people. And if we cannot ask the experts, then we can learn a little translation or Tafseer of the *Mu'tabarah* (can be justified). With the development of translation and interpretation applications, we can learn the Al-Quran easily, and of course we have to stick with the teacher's guidance if we encounter incomprehension. This is also intended so that a Muslim does not just receive information in any case that comes from an expert. A

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Muslim must clarify the truth and its accuracy to an expert or an accountable source. Allah SWT said:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِنْ جَاءَكُمْ فَاسِقٌ بِنَبَإٍ فَتَبَيَّنُوا أَنْ تُصِيبُوا قَوْمًا بِجَهَالَةٍ فَتُصْبِحُوا عَلَىٰ مَا فَعَلْتُمْ نَادِمِينَ

Mean: O you who believe, if a wicked person comes to you with news, then check it carefully so that you do not impose a disaster on a people without knowing the circumstances that cause you to regret your actions. (Surah Al Hujurat: 6).

By building the Al-Quran translation application, it will make it easier to learn the Al-Quran. And the person who studies the Al-Quran and useful knowledge, that person has traveled the road to heaven. In the book Riyadhus Shalihin hadith number 618 Rasulullah SAW said:

وعن أَبي هريرة - رضي الله عنه - : أَنَّ رسول الله - صلى الله عليه وسلم - ، قَالَ : ((وَمَنْ سَلَكَ طَرِيقاً يَلْتَمِسُ فِيهِ عِلْماً ، سَهَّلَ اللهُ لَهُ طَرِيقاً إِلَى الجَنَّةِ)) . رواه مسلم

Mean: narrated From Abu Hurairah RA. Indeed, Rasulullah SAW Said. "Whoever takes a path to study, Allah will make it easier for him to go to heaven". (HR. Muslim).

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

CHAPTER IV

CONCLUSION AND SUGGESTION

5.1 Conclusion

Experiment that has been carried out can show 2 points of conclusion based on the research question as follows:

1. Value of precision, recall, f-measure and accuracy of the application of the Cosine Similarity method without using Query Expansion and by using Query Expansion. The experiments that have been conducted have concluded that the Cosine Similarity method has a high evaluation test value in the search application of English translation, namely with a precision value of 84.52%, a recall of 68.25%, a f-measure weight of 72.94. % and an accuracy of 98.62%. Even though Cosine Similarity itself has resulted in a high evaluation value, using Query Expansion can give a higher value, can be seen using Query Expansion, Cosine Similarity can produce a precision value of 91.64%, a recall of 88.40%, a f-measure weight of 88.89% and an accuracy of 99.37%.
2. The Comparison of the two methods shows an increase in the Cosine Similarity method using Query Expansion on 4 evaluation test values. The increase in recall value is the highest, this is because the use of Query Expansion in the form of a thesaurus finds more lexically and semantically relevant documentsin the text and meaning or synonym in text, so that the value of TP (i.e. the correct document increases) and the value of FN (i.e. documents which is not found) decrease, so that the recall difference is the highest. Finally, it can be concluded that the use of

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Query Expansion in Cosine Similarity can improve performance in finding back relevant documents.

5.2 Suggestion

From a series of experiments that have been carried out in this study, of course there are deficiencies that still need to be corrected for further research. The suggestions for further research based on this research experiment are as follows:

1. The database is not only juz 30, but the entire contents of the Al-Quran to make it more helpful in understanding it.
2. Use of a thesaurus that more complete, not just the words found in Juz 30.
3. Use of other similarity methods and with programming languages that are simpler and newest, because of their more complete features.
4. Use of stemming which can detect more affixed words.
5. Selection of the thesaurus that can be performed automatically by the system at the Query Expansion step.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

REFERENCES

Alodadi, M., & Vandana, J. p. (2015). Similarity in Patient Support Forums Using TF-IDF and Cosine Similarity Metrics. *International Conference on Healthcare Informatics (ICHI)* (hal. 521-522). Dallas: IEEE, DOI: 10.1109/ICHI.2015.99.

Qiu, Y, D. (1993). *Concept-based query expansion.* SIGIR.

E, S. (1997). *Information retrieval advances using relevance feedback.* . Washington: University of Washington Department of Computer Science and Engineering General Exam.

Lancaster, F. (1979). *Information Retrieval Systems: Characteristics, Testing, and Evaluation, 2nd Edition.* New York: John Willey.

http://carihadis.com/Musnad_Ahmad/22845. diakses pada 17April 2020.

https://quran.kemenag.go.id/index.php/sura/38. diakses pada 17 April 2020.

https://www.kompas.com/skola/read/2020/02/29/140000369/kenapa-bahasa-inggris-jadi-bahasa-internasional. diakses pada 18 April 2020.

Iskandar, N. (2019). Jihad dan Terorisme dalam Tinjauan Al-Quran, Hadis, dan Fikih. *Al-Qisthu*, Vol 1, DOI : 10.32694/010650.

Khatri, S. K., Batham, T., & Johri, P. (2017). Optimizing search query using feature thesaurus and optimizaton tool. *International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS)* (hal. 171-174). Dubai: IEEE, DOI: 10.1109/ICTUS.2017.8285999.

Manning Christopher D, Prabhakar, R., & Hinrich, S. (2008). *Introduction to Information Retrieval,*. Cambridgeshire: Cambridge University Press.

Mohsen, G., Al-Ayyoub, M., Hmeidi, I., & Al-Aiad, A. (2018). On the automatic construction of an Arabic thesaurus. *9th International Conference on Information and Communication Systems (ICICS)* (hal. 243-247). Irbid: IEEE, DOI: 10.1109/IACS.2018.8355431.

Pudaruth, S., Ankiah, Y., & Sembhoo, K. (2014). Using a thesaurus-based approach for the categorisation of web sites. *Seventh International Conference on Contemporary Computing (IC3)* (hal. 624-628). Noida: IEEE, DOI: 10.1109/IC3.2014.6897245.

Rahman, N. A., Bakar, Z. A., & Sembok, T. M. (2010). Query expansion using thesaurus in improving Malay Hadith retrieval system. *International*

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG

Symposium on Information Technology (pp. 1404-1409). Kuala Lumpur: IEEE, DOI: 10.1109/ITSIM.2010.5561518.

Rasyidi, I., Romadhony, A., & Wibowo, A. T. (2013). Indonesian Hadith Retrieval System using thesaurus. *International Conference on Computer, Control, Informatics and Its Applications (IC3INA)* (pp. 285-288). Jakarta: IEEE, DOI: 10.1109/IC3INA.2013.6819189.

Shchitov, I., Lagutina, K., Lagutina, N., Paramonov, I., & Vasilyev, A. (2017). A survey on thesauri application in automatic natural language processing. *21st Conference of Open Innovations Association (FRUCT)* (hal. 296-303). Helsinki: IEEE, DOI: 10.23919/FRUCT.2017.8250196.

Soyusiawaty, D., & Zakaria, Y. (2018). Book Data Content Similarity Detector With Cosine Similarity. *12th International Conference on Telecommunication Systems, Services, and Applications (TSSA)* (hal. 1-6). Yogyakarta: IEEE, DOI: 10.1109/TSSA.2018.8708758.

Warraq, I. (1995). *Why I am Not A Muslim.* New York: Prometheus Books.

Zafikri, A. (2010). *Implementasi Metode Term Frequency Inverse Document.* Medan: Universitas Sumatra.

CENTRAL LIBRARY OF MAULANA MALIK IBRAHIM STATE ISLAMIC UNIVERSITY OF MALANG