

Universiti Teknologi MARA

**SEMANTIC ANALYSIS SYSTEM BY USING LATENT
SEMANTIC ANALYSIS (LSA) ALGORITHM ON KHUTBAH
TEXTS**

Ammar Bin Zuber

**Thesis submitted in fulfillment of the requirements for
Bachelor of Computer Science (Hons.)
Faculty of Computer and Mathematical Sciences**

JANUARY 2019

ACKNOWLEDGEMENT

Alhamdulillah, praises and thanks to Allah because of His Almighty and His utmost blessing, I was able to finish this research within the time duration given. Firstly, my special thanks goes to my supervisor, PM Dr Nasiroh Omar for her guidance for the past semester.

Special appreciation also goes to my beloved parents for her endless support. Last but not least, I would like to give my gratitude to my dearest friend who has been a wonderful help to me in finishing this project

ABSTRACT

Friday Sermons is a pivotal part of the Friday prayers in the muslim's community, as its purpose is to send out knowledge about the islam religion to its adherent. As nowadays, some muslims attend the Friday Prayers just to obey God instruction instead of willingly come to gain new knowledge. This is where the *khutbah* texts need to have a certain level of quality and also new knowledge to ensure muslims do not go back home empty-handed. With the technology nowadays that comes with text analyser, JAIS should take advantage on this technology to use it on *khutbah* texts which is way more efficient than analysing it manually. In order to make it happen, this project aims to create a system to analyse *khutbah* texts semantically which analyse the documents meaning and its relationship to other documents compared to the standard text comparison. This project will be using Latent Semantic Analysis Algorithm to cope with comparisons of *khutbah* texts as it's cover most of the requirement of the system. This project is developed by using Waterfall Model methodology as it's the most direct model that is suitable for developing a system which consist of 5 phases which are Requirement Analysis, Design, Implementation, Testing and Maintenance. In the Requirement Analysis we managed to create the use case diagram and the system specification. Furthermore, in the design phase we constructed the system architecture, system flowchart, and the activity diagram. The implementation phase are done by developing the project based on the Python language and also the gensim package tool that are available for use. The testing phase is where two testing method were conducted which were the functionality testing and the benchmark testing. The key results that were found during the project is that it is possible to find the semantic features with the system can produce results based on the text meaning such as the system. With this information the system is able to produce results that can gives the user some useful information which can contribute in the later construction of *khutbah* texts. Although the system has some few limitation which it can hold back the user experience of the system, there are few recommendations which can be done in the future which can improve and enhance the system.

TABLE OF CONTENTS

CONTENTS	PAGE
SUPERVISOR'S APPROVAL	I
DECLARATION	II
ACKNOWLEDGEMENT	III
ABSTRACT	IV
TABLE OF CONTENTS	V
LIST OF TABLES	VII
LIST OF FIGURES	IX
LIST OF ABBREVIATIONS	XI
CHAPTER 1 : INTRODUCTION	
1.1 Background Study	1
1.2 Problem Statement	3
1.3 Project Aim and Objective	4
1.4 Scope of Project	4
1.5 Significant of Project	4
CHAPTER 2 : LITERATURE REVIEW	
2.1 Introduction	5
2.2 Text Analysis Techniques	5
2.2.1 Topic Modelling	5
2.2.2 Term frequency – inverse document frequency	7
2.2.3 Named Entity Recognition	8
2.2.4 Event Extraction	10
2.2.5 Semantic Analysis	12
i) Latent Semantic Indexing	13
ii) Latent Semantic Analysis	14
2.3 Khutbah Texts	16
2.4 Comparison of Existing Similar Work	18
2.4.1 Comparison of the Themes of Malaysian Friday Sermons Between the Year 2010 and 2015.	18

CHAPTER 1

INTRODUCTION

This chapter provides the background and rationale as the basis to support this project. It also gives details of the significance of *Khutbah*, the issues and problems that led to this work.

1.1 Background of Study

In Islam, men are required to assemble on Fridays at noon, in place of the normal daily Zuhr prayers. This activity take place at a mosque at any area that can be able to accommodate to a special congressional prayer called “*Salat Al-Jumua*”. There are two main parts that consists in this special Friday congregational prayer, *Salat Al-Juma*’ which are a sermon and a prayer (Khader 2017). The Friday *Khutbah* has a large responsibility on improving the Muslim *Ummah* spiritually and physically as well as unifying them in their stand over the challenges facing them. According to Alyan (2016), These are one of the many cultural symbols that is linked to the act of sequence of the Friday khutbah, which are -:

- Taqwa (piety)

Taqwa which means fear of God, God consciousness, and mindfulness of God. It was that in the introduction of each *khutbah* the Imam will recite some verses from al-Quran that urge Muslims to have taqwa. Alyan (2016).

- Freedom

Khutbah texts has shown that the term freedom has special conceptual meanings. For example, humans are free to choose how they may live, to be thankful to God, who has the guidance of the right path, to practice its own way of living openly and freely, to choose within God’s boundaries, and to experience the freedom of