

BAB I

INTRODUCTION

Hadith is one of the guidelines for Muslims in the world besides the Qur'an. There are important figures take effects in history of Islam and the journey of the Prophet Muhammad for completing his religious mission since getting the first revelation [8]. The important figures mentioned in hadith are often ignored because so many names not sounds familiar to people.

Named Entity Recognition (NER) is one of the main components of information extraction that aims to classify named entities from a document. In general, NER can detect a person name, place name, organization name, time, month, date, etc from a document. In a short context use the application using Natural Language Processing (NLP) and Text Mining. Examples of using NER are internet search engine, indexing of a document, question answering, etc [5]. So NER is the most important step when you want to extract an information from a document.

There are many methods that can be used to build a NER system, one of them is to use the Hidden Markov Model (HMM) calculation method. The HMM method is very popular in NLP, because the performance using HMM evaluations is higher than using other evaluation methods. The main reason is that HMM can determine the locality of the phenomenon, especially the name in the document [10].

In this research will make an indexing person's names from the hadith collection which aims to extract information, classifier and make it easier for searching the person's names who appear in the history of hadith collections. The data will be used in indexing names from sentences of nine narrators's hadith books. The algorithm that will be used is HMM, because using HMM's method for Indonesian translation is still rarely used. It is expected to use HMM and combination features can improve accuracy in this research