

Introduction

The Qur'an which literally means "perfect reading" is a name of God's choice that is truly appropriate, because there is not even one reading since mankind knew the five thousand years of reading material that can match the Al-Karim Qur'an, that perfect noble reading¹. The Qur'an is a holy book of Islamic religious people which is the main guideline of religious teachings. The Qur'an itself contains teachings about life given by Allah SWT to the Prophet Muhammad. In the Qur'an there are a lot of verses. With a large number of verses it will be very difficult and take a long time for us to find a name. Manually searching for entities will be very difficult and take a long time to be search. With NER, which is one of the techniques of information extraction that aims to detect entity names, such as people's names, locations, events, and time expected search for entity names in the Qur'an will greatly simplify and shorten the time.

Information extraction is part of Natural Language Processing (NLP). The information extraction system itself is a process of finding information from a collection of documents or natural language texts as input and producing useful information in the form of structured information in a certain format². Named Entity (NE) Recognition (NER) is a component of information extraction that serves to recognize name entities (people's names, locations, organizations), time expressions (date, time, duration) and number expressions (money, percent, numerical, cardinal) in a collection of texts or documents. Based on the technique and discussing it NER has 3 main types, namely: rule-based, machine learning and hybrids³. The expected goal of the process in the NER is to extract and classify names into several categories by referring to the correct meaning. The application of the NER will be very useful to apply to the Qur'an. The application of NER will help us to the search process, especially the search for names & people in the Qur'an. For example, in the translation of the letter Al-kafirun verse 1 says "Katakanlah: Hai orang-orang kafir,"(Say : O disbelievers). In the translation, the names of the people, namely "orang - orang kafir"(disbelievers), will be identified, so the output is "orang - orang kafir"(disbelievers). With NER, it will shorten the search time rather than manually searching.

In NER there are several approaches or algorithms that can be used. The algorithm is Rule Based, Learning Based Machines that utilize Hidden Markov Models (HMM), Maximum Entropy, Decision Tree, Support Vector Machines (SVM), Conditional Random Fields (CRF) and Hybrid approaches⁴.

In this paper the author proposes a Hidden Markov Model method in its application to the Recognition of Named Entities for the Indonesian Language Qur'an. The Hidden Markov Model (HMM) is the development of a statistical model of the Markov model. The HMM in NER itself have functions to combine the combined opportunities into pairs of observations and the order of labels. Parameters are trained to maximize the possibility of a combination of training sets. Theoretically the concept in HMM is easy to implement in the case of NER. The HMM approach will be added with the multiple choice approach. Multiple choice here will provide several possibilities of the entity and be adjusted to the input to be able to detect the entity. The object in this study is only to detect entity names, and name of community in Qur'an. In Table Table 1 shows the example of the Named Entity Recognition surah Al-kafirun verse 1.

Table 1 : Example of Named Entity Recognition

Example	Tags
<i>Katakanlah:</i>	O
<i>Hai</i>	O
<i>Orang – orang</i>	B-PER
<i>kafir</i>	I-PER