

Abstract

Named entities are proper nouns or objects contained in a text, such as a person's name, country name, and others. Names of persons in some text are often the same or similar, which makes it difficult for people to distinguish or find out these same names are the same person or not. An ambiguity of names also found in hadith, like the name Abdullah in hadith number 86 and 2411, this causes ambiguity between names because both names are the same and not yet proved to be the same person. Based on this problem, then this study focuses on named entity disambiguation, which considered further context and coherent relation between a named entity. Expected in the future, it would help people to understand the ambiguity of the name or minimize the misinterpretation of names. The method used in this research was Robust Disambiguation because, in this method, the context of the named entity considered. The resulted output obtained was in the form of named entity that grouped based on the same person or different person processed with Density-based Spatial Clustering of Applications with Noise. This research resulted in an accuracy value of 90%, a precision value of 97%, and a recall value of 89% obtained from actual value and predicted value.

Keywords—Density-based Clustering, Disambiguation, Hadith Sahih Bukhari, Jaccard Similarity, Robust Disambiguation.