

ABSTRACT

The rapid development of information technology has an impact on the availability of information that can be easily accessed by anyone and anywhere. However, in between these conveniences, sometimes people also have difficulty obtaining the information they actually need. Therefore, a system called the Question Answering System (QAS) was developed, which is a system that can process questions given by users and then can provide results or answers in the form of information that is expected to be in accordance with what the user wants. The processing is inseparable from Natural Language Processing (NLP) because the questions given by users are natural language that must be processed first so that they can be processed to conduct Information Retrieval (IR) and produce answers in the form of information that users want. To optimize the QAS process, a clustering of the questions given by the user can be carried out. This is done so that the information search process can be carried out only on data related to the question. This clustering process can then determine which questions are given into which cluster, and can then be classified according to the 5W2H rule. K-Means Clustering is one of the most efficient algorithms used in the clustering process because it can produce clusters that are divided well. The Elbow method can optimize the process of K-Means Clustering by calculating in advance the Sum of Squared Errors (SSE) value of each cluster value so that the most optimal number of clusters is obtained. The results of this study show that K-Means Clustering optimized with the Elbow method can produce clusters that are in accordance with the 5W2H rule from automotive question data, as evidenced by the results of the Confusion Matrix in the form of a precision value of 90,85%, recall 88,67%, and f1-score of 88,75%. It is hoped that the results of this research can be redeveloped to build a QAS system in the automotive world.

Keywords: K-Means Clustering, Elbow Method, Natural Language Processing, Automotive, Question Answering System