**ABSTRACT**

Document clustering is the process of segmenting a particular collection of texts into subgroups including content based similar ones. The purpose of document clustering is to meet human interests in information searching and understanding. Nowadays all paper documents are in electronic form, because of quick access and smaller storage. So, it is a major issue to retrieve relevant documents from the larger database. Text mining is not a standalone task that human analysts typically engage in. The goal is to transform text composed of everyday language in a structured, database format. In this way, heterogeneous documents are summarized and presented in a uniform manner. Among others, the challenging problems of document clustering are big volume, high dimensionality and complex semantics