

## Abstract

Text mining is a part of the data mining field which focuses its work on getting valuable information from the text database. If we pay attention, most of the data mining process uses two methods, categorization and clusterization. Categorization is a technique that aims to determine the topic of an article or text document based on the class attribute (supervised). While clusterization groups the article into topics based on the similarity with the other articles (unsupervised).

Most of the categorization techniques are based on analysis of the frequency of term occurrences in the document without considering the semantics or meaning of the word. However, two terms can have the same frequency in their documents, but one term contributes more to the meaning of its sentences than the other term. Therefore, we need to apply a new approach that involves the terms that have a role in the semantics to achieve more accurate results.

Graphical concept-based mining model provides a new approach in the separation of the terms which are important and not important for determining the topic of a document based on the role of semantics. This graphical concept-based mining model provides more accurate results.

**Keywords:** TF IDF, graphical concept-based mining model, conceptual ontological graph, concept, semantic.