Abstract

Identifying the author of an anonymous text based on a predefined collection of sample texts for candidate authors is called Authorship identification. Due to the huge increase in the volume of digital texts, authorship identification becomes a need. This paper proposed a novel set of features in addition to the known features in the field in order to improve the authorship identification process in Arabic texts. The proposed framework in this paper was evaluated using a dataset that is a part of the Authorship attribution of Ancient Arabic Texts dataset. The used classifier in this research is the Logistic Model Tree classifier, which works fine with small datasets. The experimental results showed an accuracy of 88.73%. Such results support the efficacy of the proposed framework in order to identify the author of an anonymous text.