

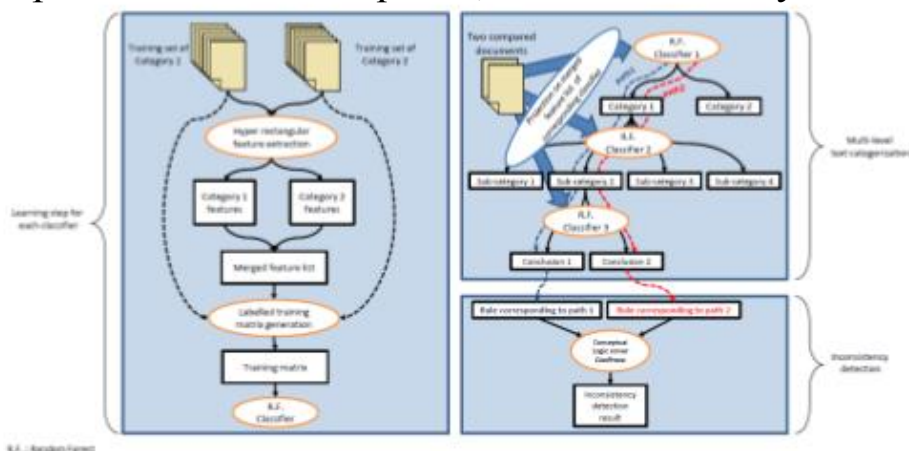
Student corner

Information extraction and detecting inconsistency

Working with Islamic domain to extract information and discover any inconsistency is a very challenging research. The domain is sensitive and risky because of the important and critical reliance on these data that exist in the web.

The attempt is done by avoiding deep linguistic analysis for discovering named entities. Instead, classification approach was used to label texts into different targeted categories as defined in the initial domain ontology. For each categorization level, discriminative keywords were extracted using the hyper-rectangular decomposition method. Then, obtained keywords were fed into the random forest classifier, which automatically detects the category of each advisory opinion. Here, in the case of Fatwas, all labels are considered as premises except judgements considered as consequents.

The results are interesting and show that the proposed method is promising in the field. The proposed method is applicable for a general domain. However, it requires the definition of an ontology, the discrimination of labels into premises and consequents, and contradictory labels in general.



Dr. Jameela Al-Otaibi
(CSE Alumni)