

Abstract

The quantity of textual data is increasing in online realm with the blink of eye and it has become very difficult to extract useful information from such enormous bundle of information. Automated text summarization is an area of NLP which provides high quality and meaningful gists, abstracts and summaries of text written in different human languages. The available summarising approaches are extractive and abstractive summarization. A lot of research is being conducted especially in extractive summary. In Urdu language, there is no research efforts in abstractive summarization up till now, so it is much needed of having research works to be done in this domain. Urdu language is mainly spoken in South Asia. In the proposed research work we use amalgam of extractive and abstractive algorithms to generate summaries. Sentence weight, TF-IDF, word frequency algorithms are used for extractive summaries. To refine results of extractive summaries, a hybrid approach is used. The summaries generated by hybrid approach will be processed by BERT model to produce abstractive summaries. The help of some Urdu language professionals is taken to evaluate the system generated summaries.