Relevance Re-ranking through Proximity based Term Frequency Model

S.Sathya Bama[1,*], M.S.Irfan Ahmed[2], A.Saravanan[1]

[1]Sri Krishna College of Technology, Tamil Nadu, India – 641042
[2]Nehru Institute of Engineering and Technology, Tamil Nadu, India
{ssathya21@gmail.com, a.saravanan21@gmail.com}
msirfan@gmail.com

Abstract. In this internet era, people rely on the most significant tool called search engine for retrieving attractive information from the web. Also, there is a rapid growth in the usage of the web increases the volume of data on the web, due to which most of the documents retrieved by the search engine is overwhelmed with inappropriate and redundant information called outliers. This not only increases the result space, but also roots in wasting the user's time and effort that makes them to surf uninteresting data. Consequently, a method is essential for the web user community to remove uninteresting information and to present the interesting data in an organized manner based on their request. Web content outlier mining is promising research area that serves these features to the web users. In this research work, proximity based term frequency model has been developed for retrieving the appropriate information and for refining the quality of the results offered by the search engine. Experimental results indicate that proximity based term frequency model improves the performance in terms of relevancy re-ranking of the retrieved documents.

Keywords: Relevance Ranking, Search Engine, Term Frequency, Proximity, Web Content Outlier.