

7th International Conference on BIG DATA ANALYSIS AND DATA MINING

July 17-18, 2020

WEBINAR

J Comput Eng Inf Technol 2021, Volume 09

AI-based data analysis for text classification and document summerization

Yuefeng Li

School of Computer Science, Queensland University of Technology (QUT) Brisbane, QLD 4001, Australia

Over the years, businesses have collected very large and complex big data collections, and it has become increasingly difficult to process these big data using the tradition techniques. There is a big challenging issue since the majority of big data is unlabelled in unstructured (information that is not pre-defined) manner. Recently, AI (Artificial Intelligence) based techniques have been used to solve this big issue, e.g., understanding a firm's reputation using on-line customer reviews, or retrieving of training samples from unlabelled tweets and so on. This talk discusses how AI techniques contribute to text classification and document summarization in the case of only obtaining limited user feedback information for relevance. It firstly discusses the principle of a new classification methodology "a three-way decision based binary classification" to understand the hard issue for dealing with the uncertain boundary between the positive class and negative class. It also extended the application of three-way decisions for text classification to document summarization and sentiment analysis. This talk will presents some new experimental results on several popular data collections, such as RCV1, Reuters-21578, Tweets2011 and Tweets2013, DUC 2006 and 2007, and Amazon review data collections. It also discusses many advanced techniques for obtain more knowledge from big data about the relevance in order to help people to create effective machine learning systems for processing big data, and several open issues regarding to AI-based data analysis for text, Web and media data.

Biography

Yuefeng Li is a Professor of Queensland University of Technology (QUT), Australia. Since the award of his PhD in 2001, he has made significant contributions to Data Mining and Text Mining, published over 200 refereed papers, including seven best paper awards. He has demonstrable experience in leading large-scale research projects and has achieved many established research outcomes that have been highly cited in some top Journals (e.g., TKDE, DMKD) and International Conferences (e.g., KDD, ICIS, CIKM, ICDM, WWW and Hypertext). His total number of Google citations = 4729 and h-index = 33. He is the Editor-in-Chief of Web Intelligence journal.