Introduction to Data, Text, and Web Mining for Managerial Decision Support
Mini-track

Dursun Delen
Oklahoma State University
dursun.delen@okstate.edu

Asil Oztekin
University of Massachusetts Lowell
asil_oztekin@uml.edu

This mini-track has five papers that are about developing systems for decision support by means of data, text, or web mining. These five papers focus on a wide range of application areas from healthcare to social media, reinforcing the fact that data, text, and web mining are effective and recently popularized tools to develop decision support systems in various domains.

In the first paper by Erranguntla et al., data mining is used as a means of determining the root causes of reactions that blood donors may have and in turn may suffer from serious injuries. Their study employs multivariate logistic regression and odds ratio calculation in order to reveal factors affecting the reactions. DonorHART data is utilized as the dataset. Findings of the study outline the interventions to be undertaken so as to improve the donor safety for blood donations.

The second paper by Sul et al. talks about a study where they seek the answer to the question of whether or not postings about S&P 500 in Tweeter has an impact on the abnormal returns. By calculating the emotional valence of tweets, they revealed that tweets associated with many followers have an influence on the same day returns whereas tweets with fewer followers have an effect on the future return predictability.

The third study by Lak and Turetken is another form of utilizing text analytic techniques. They compare and contrast the effectiveness of sentiment analysis versus star ratings. This study exemplifies the suggested methodologies in three domains: four amazon.com-based products, hotels, and reviews of doctors about their patients. The results indicate that there really is a significant difference between sentiment analysis and star ratings. This finding is mainly attributed to the tone of the star ratings ranging from being neutral zone to explicit ratings.

The fourth study of this mini-track is conducted by Tao et al. as a text mining case which is focused on an intelligent agent as a decision support tool to analyze Initial Public Offering (IPO) processes. An ontology-based information extraction framework is developed with three modules i.e. information extraction module, reasoning & learning module, and analytics module. The proposed framework can be used by both researchers and practitioners in order to determine the price changes retrieved from the texts of the IPO prospectus.

The fifth paper in this mini-track reports on a study conducted by Best et al. in which they develop a data mining based decision support tool to determine future net share inflows and outflows via the use of Exchange Traded Funds (ETFs). CRISP-DM methodology along with logistic regression, decision trees, and artificial neural network is utilized to predict the flow shares in ETF mutual fund markets. Information fusion-based sensitivity analysis is deployed in order to rank order the predictor variables in terms of their importance in prediction.