Abstract No: SO-04   **Software Intensive Systems**

Natural language processing framework: *WordNet* based sentimental analyzer

J. R. K. C. Jayakody

Department of Computing and Information System,
Wayamba University of Sri Lanka, Dandagamuwa, Sri Lanka
kithsirij@wyb.ac.lk

Sentimental analysis is a technique which is used to classify different types of documents as positive, negative or neutral. Hand written form, mails, telephone surveys or online feedback forms are used to get customer feedbacks about products and services. In fact, sentimental analysis is the technique which is used to mine online and offline customer feedback data to gain insight of product and services automatically. Since business types are different, it is quite challenging to develop a generic sentimental analyzer. Therefore, this ongoing research focused on developing a generic framework that can be extended further in future to develop the best generic sentiment analyzer. Several online customer feedback forms were used as the dataset. Webpage scraping module was developed to extract the reviews from web pages and chunk and chink rules were developed to extract the comparative and superlative adverbs to build the knowledge base. The web site (Thesaurus.com) was used to build the test data with synonyms of good, bad and neutrals. Next WordNet database was used with different semantic similarity algorithms such as path similarity, *Leacock-Chodorow*-similarity, *Wu- Palmer*-Similarity and *Jiang-conrath* similarity to test the sentiments. Accuracy of this framework was improved further with the vector model built with natural language processing techniques. Label dataset of amazon product reviews provided by University of Pennsylvania were used to test the accuracy. Framework was developed to change the multiplied value based on the domain. The accuracy of the final sentiment value was given as a percentage of the positive or negative type. This framework gave fairly accurate results which are useful to generate good insights with user reviews.

Keywords: Sentimental analysis, Data mining, Natural language processing