Abstract

This paper describes a strategy of extracting information from raw data and visualizing them in web browser. Raw data are collected from newspaper. These raw data are in English language. By implementing text mining process specific information extracted and this process explained clearly. Derived information is specifically on road accident related news but raw data contains all kind of news. One of the significant parts of this process is to visualize retrieved information in an interactive way in web browser. Interactive geographical map, bar chart, bubble chart have used for visualizing data. Interactive data visualization helps people to understand data easily, let them take actions on the data, give comfort to their eyes as data visualization is a combination of art and science. The whole process contains two major parts – information extraction, interactive visual representation of extracted information. People can understand easily by interactive visualization at a glance. Interactive data visualization can make people more concern about road accident. There is also an evaluation part which performed in two ways, one is by measuring the accuracy of information by standard information extraction process and another way to evaluate is by taking feedback from user. It is done by reading
some of the raw data files of some days and matching the similarity with the retrieved data of the same days. Some limitations of this process have been described and some improvement which can be take to extend the process and will help to go deeper level have been described in the final part.

References


Index Terms

Computer Science Information Sciences

Keywords

Text mining, Information Extraction, Natural Language Processing, Visualization, Accident.