Association rules of data mining application for respiratory illness by air pollution

ABSTRACT

Exposure to air pollution has been related with vary adverse health effects. This study aims to assess the impact of air pollution to the number of hospitalization for respiratory illness in Kuala Lumpur as the case study. Kuala Lumpur, the capital city of Malaysia, is an urban and industrialized city in the tropical climate of Malaysia that often experiencing has highest record of severe respiratory illness due to air pollution. The effects of air pollution on health triggers oxidative stress and inflammation, and it is plausible that high levels of air pollutants causing the high number of hospitalizations. In this study, an intelligent approach in data mining called association rules has been used based on its capability to search for an interesting relationship among attributes in a larger database and to its ability to handle uncertain database that often occurs in the real world problem. Association rules mining is a discovery of association relationships, frequent patterns or correlations among sets of items or elements in databases. In air pollution and healthcare database, association rules are useful as they offer the possibility to conduct intelligent diagnosis and extract invaluable information and build important knowledge bases quickly and automatically, in order to develop effective strategies to minimize the health exposure to the air pollution. A total of 2102 data were obtained from the Department of Environment Malaysia and Malaysian Ministry of Health. There were six attributes used as input and one attribute as an output for the association rule mining. Data has been through a pre-processing stage to facilitate the requirement of the modeling process. As for conclusion, association rules mining has given a promising result with more than 90% accuracy and the rules obtained have contributing to knowledge for the respiratory illness.