Expectation Maximization (EM) Algorithm for Classification in Medical Disease: Tuberculosis (TB)

May Thazin Chain, Zaw Tun

maythuchain@gmail.com
University of Computer Studies,Yangon

Abstract

Data Mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. Classification is one of the most popular data mining tasks with a wide range of applications and lot of algorithms have been proposed to build accurate and scalable classifiers. This paper uses Expectation-Maximization (EM) algorithm to classify both labeled and unlabeled data. EM is a class of iterative algorithms for maximum likelihood or maximum a posterior estimation in problems with incomplete data. This system implements EM algorithm along with the extended definition for identifying classification of medical disease provided good classification accuracy based model.