

15th International Conference on
Information Technology, Computer & Telecommunication



An overview of data mining algorithms and information extraction methods

Naser Molla AliAkbari
M.Sc. in Computer Engineering, Software
Orientation, Islamic Azad University, Mahallat
Branch, Markazi, Iran,

Abstract

Data mining means trying to find a specific pattern among the data with the help of algorithms. There are some widely used algorithms in data mining that are approved by more experts. Data mining is a set of operations performed by a computer on a large amount of data to find a specific pattern among the scattered data. This action causes the hidden order to be found in the information that does not seem to have order, and the result is displayed in a way that is understandable to humans so that it can be used for decision making and planning. There are various methods for data mining, such as clustering, classification, and so on, which are used for each of the specific algorithms. Among the data is very useful information to improve the quality of various parts of human life. From detecting the possibility of illness in individuals, to finding better sales patterns in Internet business, and even identifying the face of the offender through CCTV, there are all different aspects that data mining can help a person. Classification is used to find out in which group each data instance is related within a given dataset. It is used for classifying data into different classes according to some constraints. Several major kinds of classification algorithms including C4.5, ID3, k-nearest neighbor classifier, Naive Bayes, SVM, and ANN are used for classification. Generally a classification technique follows three approaches Statistical, Machine Learning and Neural Network for classification. While considering these approaches this paper provides an inclusive survey of different classification algorithms and their features and limitations. In data mining, an algorithm is a set of commands that are defined in computer languages and can be executed by a computer. In data mining, there are many algorithms that analyze large data and extract meaningful patterns from them. Here are some of the most common ones.

Keywords: data mining algorithms, information extraction, Naive Bayes, SVM, k-nearest neighbour.