An improved Id3 algorithm for clinical data classification

119003080	Korrapati Gangadhar	Team No. :
119003107	Podugu Rohith Surya	Place: SASTRA
119003228	Nanisetti Lakshman Veerendra	

Data mining has seen a wide adoption in recent years largely because of the ability of data mining techniques to rapidly yield answers for large variety of problems in a short amount of time and the availability of large quantities of data to exploit.

This project proposes an improved classical Iterative Dichotomiser 3 (Id3) algorithm for disease prediction[1].

The ID3 algorithm is used by training on a dataset to produce a decision tree which is stored in memory. At runtime, this decision tree is used to classify new unseen test cases by working down the decision tree using the values of this test case to arrive at a terminal node that tells you what class this test case belongs to.

The Id3 algorithm has the following problems:

- multi-value bias when selecting test or split attributes
- Numeric attribute discretization

To solve these problems a improved Id3 algorithm is proposed which uses rule-based classification model to reduce average memory cost.

Improvement in terms of accuracy, stability and minor error rate when compared to existing systems.

References:

1. An improved Id3 algorithm for medical data classification https://www.sciencedirect.com/science/article/pii/S004579061732517X

Smt. S. Kamakshi