Recruitment of appropriate employees and their retention are the major concerns towards creating the competitive strength in the knowledge economy. Every year IT companies recruit fresh graduates through their campus selection programs after examining their skills by conducting tests, group discussion and a number of interviews. The recruitment process requires enormous amount of effort and investment. During each phase of the recruitment process, candidates are filtered based on some performance criteria. Intense analysis on the system indicates that a pattern exists among the candidates selected for an industry. The problem domain is complex and the aspects of candidates that impact the recruitment process is not explicit. In this research, the domain knowledge is extracted through knowledge acquisition techniques. Data mining techniques that fit the problem are determined. A study has been made by applying K-means and fuzzy C-means clustering and decision tree classification algorithms to the recruitment data of an industry. Experiments were conducted with the data collected from an IT industry to support their hiring decisions. Pruned and
unpruned trees were constructed using ID3, C4.5 and CART algorithms. From the comparative study, it has been observed that the clustering algorithms are not much suitable for the problem and performance of the C4.5 decision tree algorithm is high. Using the constructed decision trees, discussions were made with the domain experts to deduce viable decision rules.

Reference

- Jiawei Han, Micheline Kamber. 2006. Data Mining Concepts and Techniques”, Second Edition Morgan Kaufmann Publishers, San Francisco.
- R. Agrawal, A. Arning, T. Bollinger, M. Mehta, J. Shafer, R. Srikant. 1996. The Quest Data Mining System, in Proc. 2nd Int'l Conference on Knowledge Discovery in Databases and Data Mining, Portland, Oregon, 244-249.
- Ron Kohavi, Ross Quinlan. 1999. Decision Tree Discovery, in Proc. International Conference on Data Mining and Knowledge Discovery, Vol. 6, CA, USA.
Applicability of Clustering and Classification Algorithms for Recruitment Data Mining

Index Terms

Computer Science
Data Mining

Key words

Recruitment process
Knowledge

Engineering
Classification

Decision trees
Fuzzy C-means algorithm
K-means algorithm