

Title of the R&D Project
**Teacher Recruitment Modeling Using Partitioning Based
Data Mining Techniques**

(F.No. 42-132/2013(SR) dated 14-10-2014)

Funding Agency:	UGC, New Delhi
Total Grants Received:	Rs. 5,85,800/-
Duration of the Project:	22 nd Oct. 2014 to 31 st Mar. 2016
Principal Investigator:	Dr. K. Vijayakumar, BMSCE
Co-Investigator:	Dr. AtulNegi, University of Hyderabad
Project Fellow:	Mr. Shankru Guggari
Data Sets used:	Teachers data from 10 Engineering Colleges from South India (Source: AICTE Mandatory Disclosures)

Summary:

The recruitment of good teachers is crucial for educational institutions to provide quality education, thereby moulding the students to face the challenges of tomorrow. The identification of high quality teachers is a challenging task. Traditional approaches (e.g. manual selection of teachers) may fail to choose the right teachers for the right job due to their subjective nature. Therefore, a promising technology is required to support the college administration during the recruitment of the teachers. In this project, we develop and apply data mining techniques on engineering college teachers data to elicit hidden patterns or relationships among the characteristics of the teachers. We use the data of 1992 teachers collected from AICTE mandatory disclosure documents of 10 engineering colleges from South India for our investigation. In this work, we propose a Partitioning based Decision Tree method (PDT) which creates sub-objects for each data object based on themes, constructs multiple local decision trees using the sub-objects, and combines the decisions based on nearest neighbour rule. The PDT method shows improved classification over classical decision tree classifiers (CART, C4.5, C5.0) since it overcomes the issue of curse of dimensionality by constructing decision trees from feature subsets (i.e. sub-objects) instead of original data objects. The empirical results on Teacher data confirm the superiority of PDT method over CART, C4.5, and C5.0 decision tree methods. Subsequently, we identify relationships between characteristics (features) of the

Teachers using Association Rule Mining. Association Rules help the college administration to develop recruitment and HR policies to enhance quality of research, teaching and learning. The study brings out various recommendations for improved research output, patents, awards, R & D grants and book publications. We also propose an Itemset partitioning approach to frequent item mining for computation of frequent itemsets more efficiently. Lastly, we use an outlier method to identify the teachers with good performance. Using the outlier knowledge, the administration can initiate suitable steps for recognizing the efforts of good teachers.

Distribution of Teachers data:

Designation	No. of Teachers
Assistant Professor	1203
Associate Professor	458
Professor	331
Total	1992

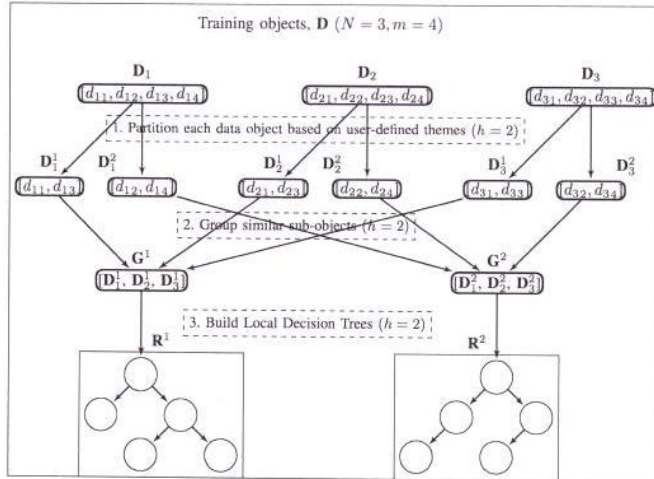
Characteristics of the Teachers data used:

Designation (Nominal), Department (Nominal), PG-Degree (Binary), PhD-Degree (Binary), UG-Class (Ordinal), PG-Class (Ordinal), Teaching-Experience (Ratio), Industry-Experience (Ratio), Research-Experience (Ratio), National-Journal-Papers (Ratio), International-Journal-Papers (Ratio), National-Conference-Papers (Ratio), International-Conference-Papers (Ratio), PhD-Institute (Ordinal), No.-of-PhD-Students (Ratio), UG-Projects-Guided (Ratio), PG-Projects-Guided (Ratio), Books-Published (Ratio), Number-of-Patents (Ratio), Professional-Body-Membership (Binary), Consultancy (Binary), Awards (Binary), Grants-Fetched (Binary), Interaction-With-Professional-Institutions (Binary).

Outcomes of the Project:

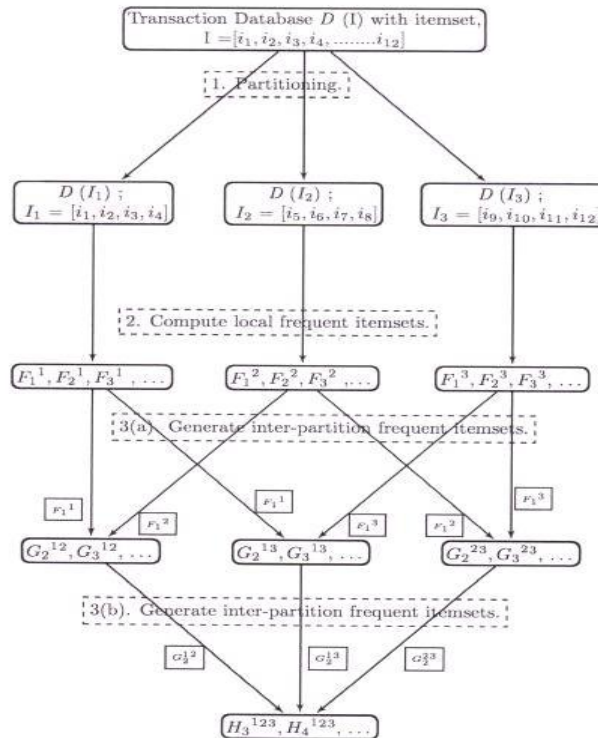
1. Decision Tree Classifier based on Partitioning of Feature set (PDT):

- (i) The method creates sub-objects for each data object based on themes, constructs multiple local decision trees using the sub-objects and combines the decisions based on a combiner.



- (ii) The PDT method shows improved classification over classical decision tree classifiers (CART, C4.5, C5.0) since it overcomes the issue of curse of dimensionality by constructing decision trees from feature subsets (i.e. sub-objects) instead of original data objects.

2. Frequent Item Mining method based on Partitioning to facilitate parallel computation of frequent itemsets.:



3. Recommendations for Recruitment of Teachers:

The study brings out various recommendations for improved research output, patents, awards, R & D grants and book publications. Some of them are as follows:

- *Characteristics for Recruitment of Assistant Professors:* Professional body membership, a Research publication, Minimum of 5 years of Teaching experience.
- *Characteristics for Recruitment of Associate Professors:* Professional body membership, Interaction with professional institutions, Research experience, Ph.D., Research publications, Consultancy, Minimum of 10 years of Teaching experience.
- *Characteristics for Recruitment of Professors:* Interaction with professional institutions, Consultancy, Research publications, Research experience, Guiding research scholars, Grants, Awards, Ph.D., or Book publications.

4. Outlier Identification in Teachers Data:

- a. Useful to identify the teachers with good performance.
- b. Using the outlier knowledge, the administration can initiate suitable steps for recognizing the efforts of good teachers.
- c. We identified outliers using Clustering based method.

5. List of Publications/Accepted Manuscripts:

1. Vijayakumar Kadappa, Shankru Guggari, and Atul Negi, “Decision tree classifier using theme based partitioning”, In Proceedings of *IEEE International Conference on Computing and Network Communications (CoCoNet-2015)*, Dec. 16-19, 2015, Trivendrum, India.
2. Vijayakumar Kadappa, Shankru Guggari, and Atul Negi, “Teacher Recruitment Modelling based on Association Rule Mining”, *IEEE International Conference on Data Science and Engineering (ICDSE-2016)*, Aug. 23-25, 2016, Cochin, India (*Accepted for Publication*).