CS522 Project Proposal - Predict Grant Application

Priyaranjan Behera
Computer Science Department
North Carolina State University
pbehera@ncsu.edu

Sai Sri Harsha Kunapareddy
Computer Science Department
North Carolina State University
skunapa@ncsu.edu

1 Project description

This project involves the prediction of the outcome of grant applications in a university. The project will be to incorporate most of the methodologies that has been covered in the class including data preprocessing, feature selection, data transformation and implementaion & evaluation of different classifiers.

2 Dataset

The dataset is obtained from a Kaggle competition hosted by the University of Melbourne.

- 249 multivariate, time-series attributes
- 8K data objects

3 Software requirements

The data processing and classifications will be done in Matlab with extensive use of the Statistics and Machine Learning Toolbox.

4 Papers to be referred

- Top 10 algorithms in data mining. Knowl. Inf. Syst. 14, 1 (December 2007), 1-37. DOI=http://dx.doi.org/10.1007/s10115-007-0114-2
- 8K data objects

5 Work division

Priyaranjan Behera and Sai Sri Harsha Kunapareddy will work together in this project and most of the work will be done collaboratively. The number of attributes to be processed, preprocessing methodologies and classifier implementations will be split evenly.

6 Midterm milestone

By November 12, we aim to complete the preprocessing steps of the data and start with the classification implementations (Viz. - Decision tree, SVC, Naive-Bayes, Neural Network)

asd

asd

asd

asd

asd