

# **Performance of Naïve-Bayes Classifier in Estimating Writing Portfolio Grades**

Daniel Alabi and Cody Wang

In order to estimate the writing portfolio grades for each student in the dataset, we used Naïve-Bayes Classifier to generate the probabilities that a student would pass.

For each data field (all numerical), we used two methods to generate the probabilities: binning and calculating the probability densities based on the Gaussian distributions.

By using different methods on each of the numeric fields, we obtained different accuracies. Based on our experiments, the following choices generated the best accuracy: 85.11%.

Binned fields:

Minnesota

International

Birth Year

# essays in portfolios

AP Credits

CS Credits

Science Credits

Writing Credits

Fields assumed to have Normal distribution:

Verbal SAT

Math SAT

Abroad Credits

Cumulative GPA

English Credits

For each of the binned fields, we decided that by having one bin for each distinct value, the results had the best accuracy. This is the same as treating them as nominal values. We think that the reason is that these fields are highly reliable for estimating the writing portfolio grades for the students.

Comparing with the k-nearest method, which generated a highest accuracy of 76.60% with k=39, Naïve-Bayes Classifier outperforms it by about 8%.