Predicting Student Performance

Questions to ask for paper

Abstract ....

Introduction ...

Dataset

- How many features are there? What were the types of features? How did we handle curse of dimensionality? What transformations did we perform?
- How did the histograms between STEM and Non-Stem subject differ? E.g grades in first term in non-stem seem to be better average than stem. Seems to be more improvement in stem grades. Seem to be less absences in non-stem. Seem to be more absences in stem ranging from 0 – 20.
- From seaborn plot of pairwise correlations, which features showed to be more correlated with each other? How would this affect the models? E.g. greater co correlation can impact the performance of the model as more weight is placed on these features.
- Was there any abnormalities in the dataset? Missing values, outliers, duplicates, inconsistent data? How did we correct for this?

Methodology

- Why did we choose the models we chose?
- How did we approach the problem of predicting student performance?
- What did we do differently?

Models

- Based on the regression models, what features seemed to be the most predictive?
- Examine all of the plots and numbers, do you have any comments on them?

- Do you have any idea what might be causing the trend in data?
- What did you notice about the models accuracy as it performed on the test sets? Did they drop? What about bias vs variance, how did we handle this?
- How balanced was the target variables? How did we handle imbalance in predicting pass vs no pass?
- In a comparison of models, which seemed to perform better in labelling classes? What did the confusion matrix show?
- In the future how can the models be improved to handle new data?

Results

- How did students in a stem class compete against students in a non-stem class?
- What do predictions look like, do the features we chose help determine whether a student would pass the class or not?
- What features showed that they were less useful for predicting student performance?
- What can students do to improve performance in school?
- Are there discrepancies between what students can do in a stem field and a non-stem field? Which features showed that performance improved in stem and which showed improvement in non-stem?