Sleep vs Study Time: Which fares better for results?

INTRODUCTION

What matters more before an exam, more studying or more sleep? In this article, Maksym Morwaski's 320 Exam1 survey dataset is used to create a predictive neural network model for predicting a student's exam letter grade based on various preparation factors leading up to the student's exam. It is found that it is better for a student to sleep more as opposed to studying more, and excessive studying will yield diminishing returns.

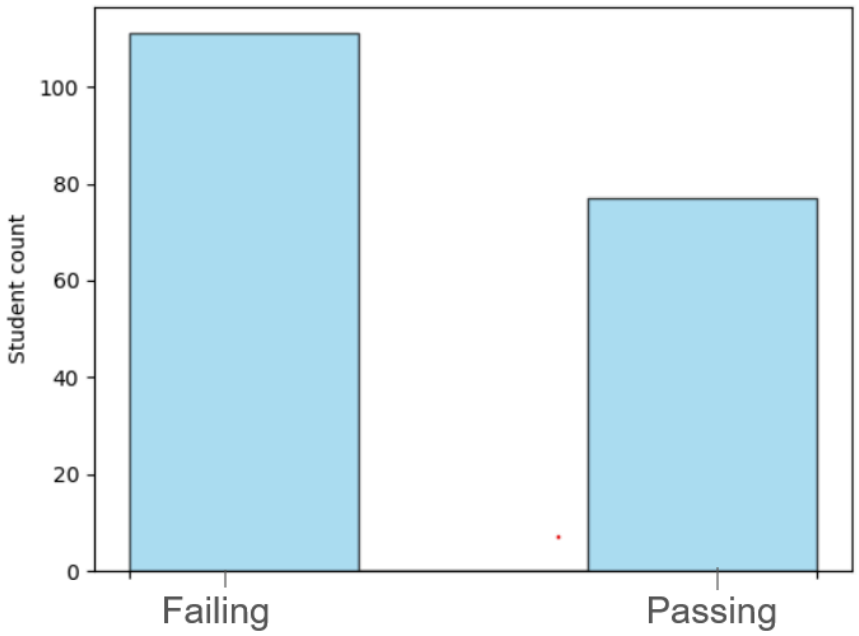
DATA CLEANING, ML DATA PREPARATION, & MODEL ANALYSIS & SELECTION

To prevent leakage, the "Total Score" column was dropped to prevent the model from being given data that already contained characterizing information of the target prediction, and in conjunction the "Max Points" column was dropped due to lacking context.

For model creation, the first 60 data points were selected for training, and the subsequent 20 points were used for evaluation.

Pass Fail Analysis

The given distribution of passings consists of the following



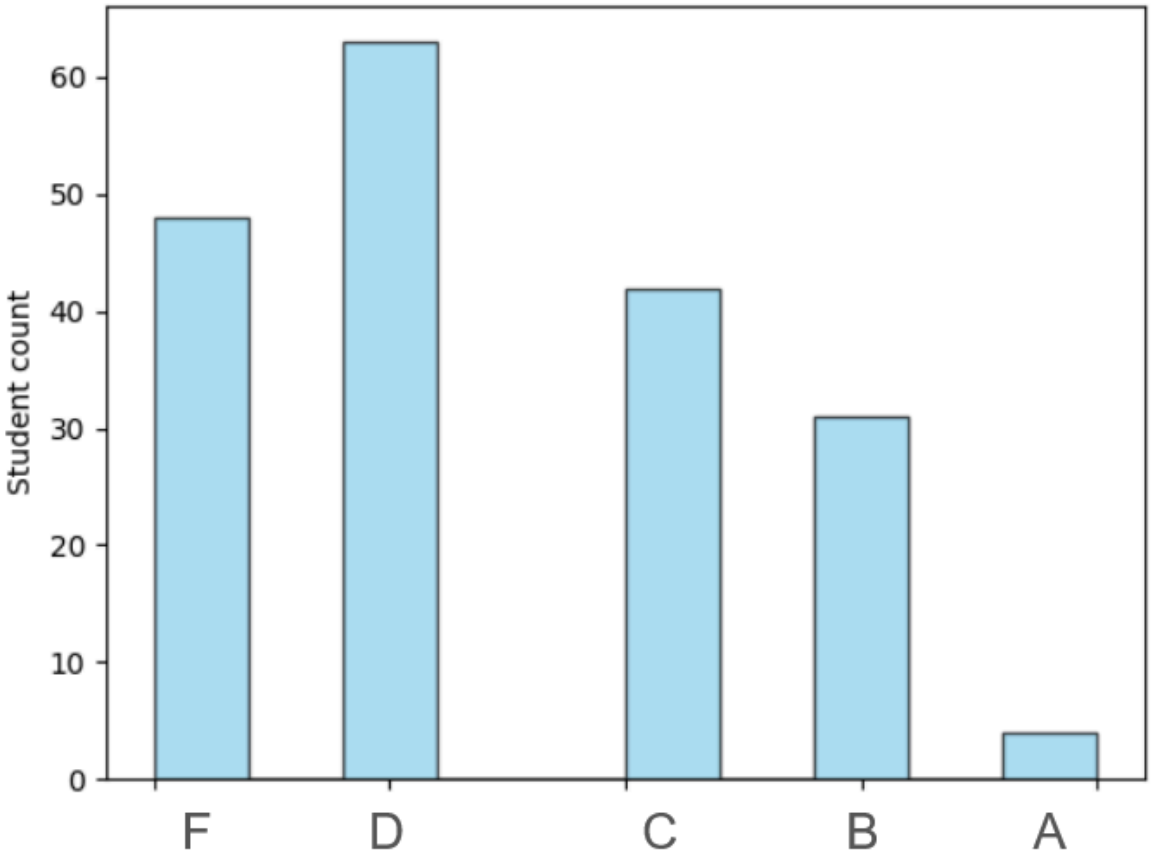
Constructing a model to predict the passing of students results in the following levels of accuracy for each type of model

Accuracy of each model type

- *Decision tree classification: 0.789*
- *Naieve Bayes (K=5): 0.631*
- *Neural Net (1000 iterations): 0.578*

Letter Grade Analysis

Our given distribution of letter grades consists of the following



Constructing a model to predict the letter grade value results in the following levels of accuracy for each type of model

Accuracy of each model type

- *Decision tree classification: 0.368*
- *Naieve Bayes (K=5): 0.368*
- *Neural Net (1000 iterations): 0.526*

Real Valued Score of students Analysis

Constructing a linear regression model to predict the true score of students results in a MSE of 0.029, and an R^2 of 0.250

Assessing Model Options

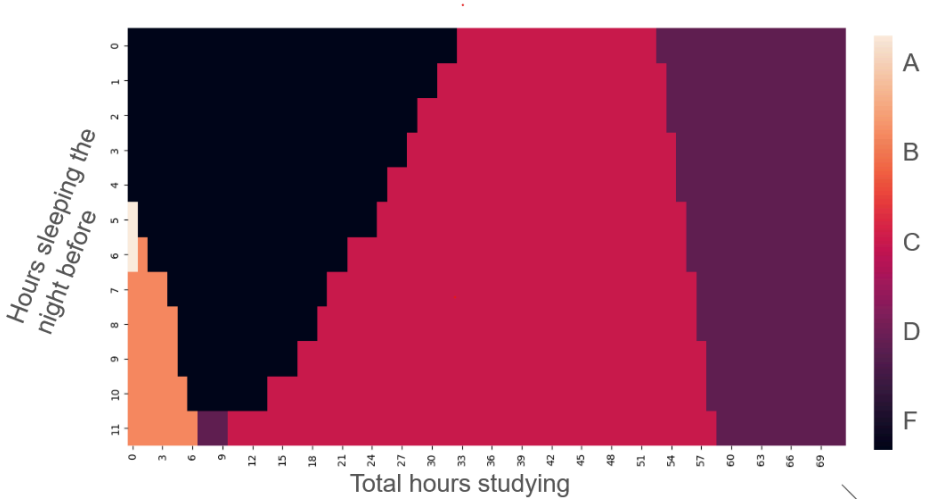
The more simplistic models fared better on the pass fail metric, i.e. the decision tree out performed all models, and provided a reasonable accuracy of 70%+. For the letter grade distribution, more complex labeling was required, and only the neural network was able to produce an acceptable accuracy of 50%+. We can see from an R^2 value of 0.25 that our regression model is a weak representation of the data, so it should only be considered if we really need those exact grade percentages. With this analysis, we will pursue further study in regards to students letter grades, thus we will select the neural network model

EXPERIMENT

To determine if sleep or studying is more critical, we can select a single datum, enumerate every combination of hours studied and hours self from 0,1,...,72 hours and 0,1,...,12 hours respectively, and compare the produced grades of each model output. For the student we are considering for the single datum, we will fare on the side of hardworking but inexperienced, so their remaining traits leading up to the exam will consist of always going to class, having no experience in ML, having a median amount of infinite scroll time, 2pm lecture, junior, read all articles, predicted they got an A, and wants extra credit.

RESULTS

The experiment produced the following results.



From here we can see that the students who did the best surprisingly didn’t study at all and slept a reasonably healthy amount! This is likely due to model underfitting, or a surplus of misleading survey answers. However the data does indicate that a modest amount of studying, roughly 0-6 hours to be exact, results in better results when combined with a healthy amount of sleep. The data does indicate however that if a given student studies for at least 12+ hours it is likely they at least receive a passing score; this could be preferable to a student who wishes to have a reliable but less satisfactory score.

CONCLUSION

We can conclude that if a student wishes to perform exceptionally well on an exam it is more important for them to sleep more as opposed to studying more, and that studying excessively to the detriment of one's sleep leads to decreased academic performance.