Aim:

The study aims to use ML algorithm to predict the GDP output of a student. The goal is to justify one of the factors in the financial outcome of a student, enabling one to justify the success or failure of the student as the final result of the report.

Model Selection:

The GDP class is optimized to -1,0,1 during the former section to help justify the success of a student. Thus instead of a numerical variable, the variable is classified. For these values, classifiers can have much more precision and accuracy than regression methods. A popular choice of classifiers are the decision tree models, and one of the best among them is the random forest. Firstly, for a high-dementional data, random forest can perform better than the traditional decision tree model. 他相比于传统的决策树具有更高的鲁棒性，意味着可以有效地减少过拟合的现象。同时，他降低了对训练数据的敏感程度，以此得到相对低的generalization error。

Data preparation: