REPOSITORY INFORMATION

I've uploaded the jupyter notebook file, there is no other python scripts or any of the data for the reason of repository being public, having the dataset folder and scores.csv file put in the same folder as the project_code.ipynb file should be enogh to run the code without any issues. The detailed information on all functions used in the project are available as comments above the respective functions.

METHODOLOGY

I've used the neural network implementation from our previous assignment in this project however naturally a lot of changes had to be implemented. There were several issues and problems I had to overcome during the implementation of the project; I've generalized the scores based on the range of the most frequent scores distribution since it would make it easier for the model to bag individual scores into 5 categories and train and predict using that instead of with the whole range of 0-100 which could easily overfit. Overall my approach was to achieve a model that did not overfit and to achieve this, I tried to generalize the parameters, outputs and adjustedhyperparameters accordingly to achieve the desired effect.

RESULTS

The resulting model had a 58% accuracy in guessing the score of a student's assignment based on prompts alone HOWEVER the guesses made by my model are not pinpoint scores, rather than that, I've used generalized 5 point gaps until 80 and 0 for all others to generalize the y value and avoid overfitting since there were very limited data compared to the wide range of possible outcomes (being 0 to 100) and majority of the data's score were in between 80 to 100. Visual representation of approximate estimation on test data is available in the ipynb file.

TEAM CONTRIBUTIONS

Berkem Elgül - 29363: Did the project by myself so all credit goes to me I guess :)