

In the first stage of this assignment, I created a dataset which had two classes of students: passing (1) and failing (0). The dataset was linearly separable since the Perceptron was able to converge and separate the two classes. I double checked this by comparing its predicted values with the ground truth class label.

I moved some of the samples in class (1) to that of class (0) and vice versa. This dataset was not linearly separable. I fit the perceptron and compared its predictions with the ground truth, but this time there were faults in its comparison, which proves that the dataset was not linearly separable. After 30 epochs the model reached 60% accuracy and after 100 epochs the model reached 80% accuracy.

In developing a perceptron model, I found that mine tended to be full of errors so I implemented the model described in the book with the exception that their initial size parameter for `self.w_` was `1+X.shape(X)[1]` which caused problems when it received an ordinary list. I modified the parameter as `1+np.shape(X)[1]`. I also used list comprehension to fix another type error which arose when updating weights.

I fitted the titanic dataset with the batch AdalineGD class described in the book using the following passenger features: {pClass, sex, age, sibSp}. The algorithm did not converge, but I ended training after 30 epochs with a `cost_` value of 6.34. The AdalineGD algorithm fitted the dataset with 61.39% accuracy. I assume that a better activation function would allow a closer fit to the data.

The most significant predictive feature of the dataset was a passenger's age. The weight corresponding to age -0.0107 was double the absolute value of the weight corresponding to sex. The only weight more significant than age was `w_[0]` or the bias. The bias was -0.65 which presupposes that the passenger will die most of the time regardless of their age or sex.

The baseline model I created randomly guesses by using `random.random()` which guesses a number between 0 and 1. If the guess is greater than 0.5 it selects (1) life. Values less than 0.5 become (-1) death. The baseline model changes accuracy every time, but is approximately 50% accurate. At 61.39% accuracy, AdalineGD is only slightly better at guessing.