Ami Ashman

2-12-22

Intro to AI in Python

Linnerud Dataset

The Linnerud Dataset consists of “three exercise (data) and three physiological (target) variables collected from twenty middle-aged men in a fitness club.” The following paragraph explains how to create a Regression model that plots the relationship between waistline and sit-up numbers.

The first step is to import the necessary python libraries, namely matplotlib.pyplot (hereafter referred to plt), numpy (hereafter referred to np), and – from sklearn – datasets, linear\_model, and model\_selection. The next step is to load the dataset, in our case with dataset.laod\_linnerud(return\_X\_y=True). Refactor the dataset using np.newaxis to isolate the particular features you wish to model (in our case waistline and sit-up numbers). This will ensure that the other features do not affect our results. Next, we split the model into our training and testing groups. We will train and hone to model using the training set and reserve the testing group for testing. This will allow us to get a better reflection of our model’s accuracy.

To set up the linear regression, write model = linear\_model.LinearRegression(). This initializes the model that we will train and test. Call .fit() on the model using the X\_train and y\_train group. Our model is now turning the gears and fine-tuning itself. To see how well our model has done, call model.predict(X\_test) and assign it to y\_pred. This is it! Our model has predicted the output but, in a form, unreadable to humans. Using plt, we can plot our data on a graph. Below is an example:

plt.scatter(X\_test, y\_test, color='black')

plt.plot(X\_test, y\_pred, color='blue', linewidth=3)

plt.xlabel('Scaled BMIs')

plt.ylabel('Disease Progression')

plt.title('A Graph Plot Showing Waistline Against Sit-up Ability”)

plt.show()

Python will display a graph inside the Jupyter Notebook. This can be changed based on personal preference.