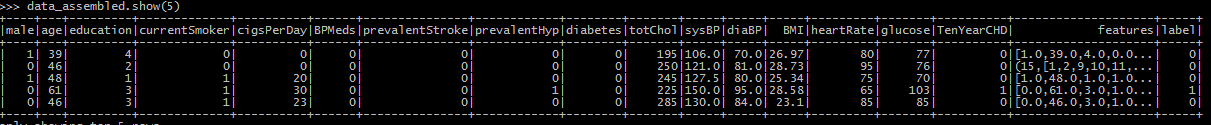
P2.

Preprocess:

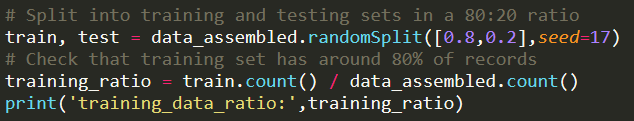
Drop the data rows with missing value



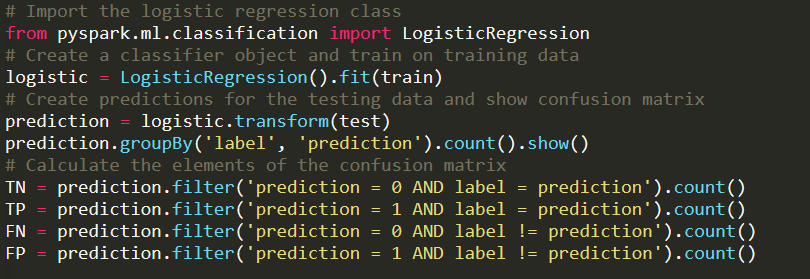
Add features vectors and label columns as the model input require form



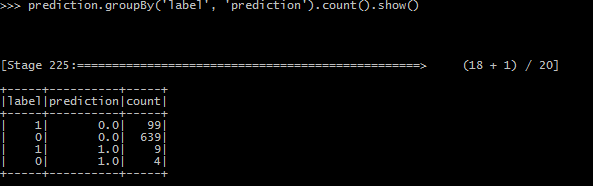
Train\_test split

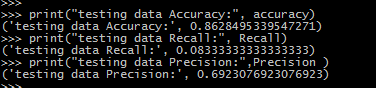


Training :



Testing result evaluation:





Result Analysis :

Although the accuracy of the model is pretty high(0.86), it mainly comes from the high TrueNegative. From the extremely low recall rate, we can find that the model can’t really identify tshe True positive cases. And from the precision, we can also see that only 0.69 of the prediction are truly positive