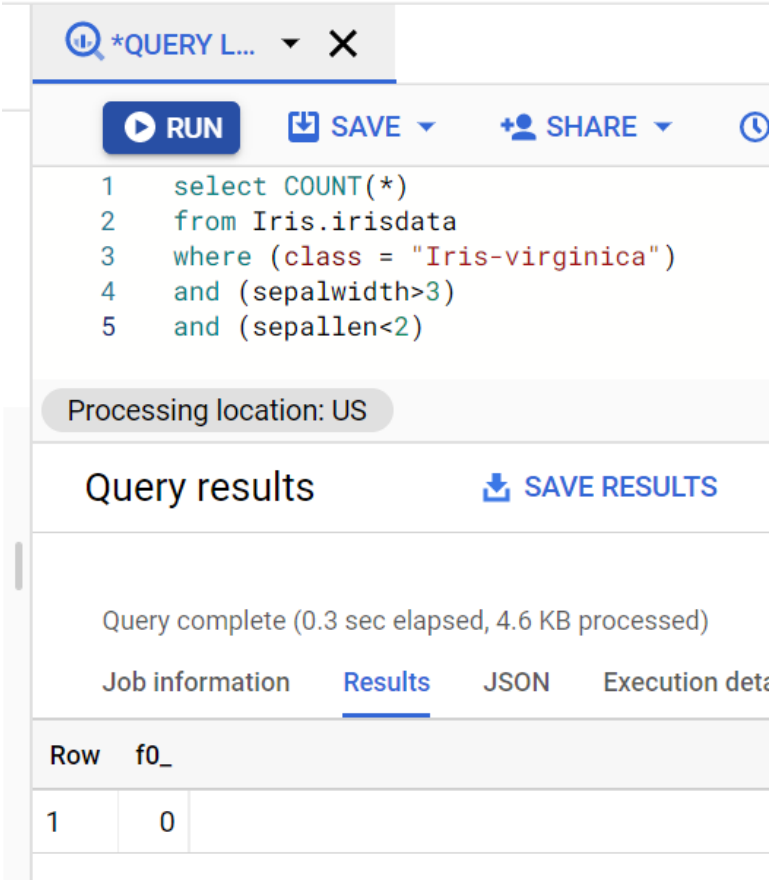
CS4830 Assignment 3

Shashank H S BE18B006

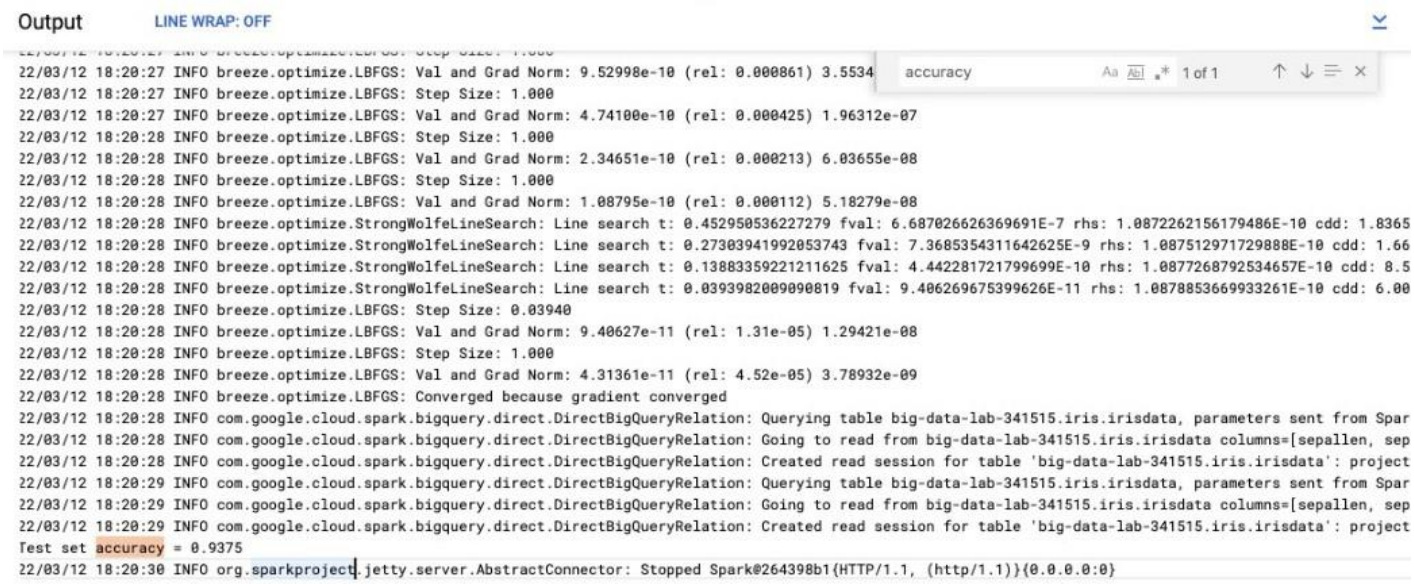
2. Query and query result are shown in the picture below. No Iris virginica flowers fit the given constraints. Full picture in zipped folder.



3. 80% of the data was used for training and 20% was used for testing.

Preprocessing	Model	Test set accuracy
None	Random Forest classifier	0.92
None	Logistic regression	0.805556
None	Multilayer Perceptron	0.9375
Standard scaler	Multilayer Perceptron	0.933333
Standard scaler and PCA	Multilayer Perceptron	0.936

A screenshot for the Multilayer Perceptron training job in Dataproc.



Multilayer Perceptron performs the best. PCA reduces its accuracy since compression is not lossless.