Logistic Regression run:

```
Opening file titanic_project.csv
Reading line 1
Heading: "","pclass","survived","sex","age"
Closing file titanic_project.csv
Weights:  100.414
-144.164
Accuracy: 0.784553
Sensitivity: 0.695652
Specificity: 0.862595
192ms
192.193ms

Program Terminated.
```

Naïve Bayes run:

```
Opening file titanic_project.csv
Reading line 1
Heading: "","pclass","survived","sex","age"
Closing file titanic_project.csv
Apriori: 0.591778
0.408222
Likelihood values for p(pclass|survived):0.166397 0.235864 0.597738
0.423888 0.269321 0.306792
Likelihood values for p(sex|survived):0.155089 0.844911
0.683841 0.316159
Mean age for survived30.5186
28.8993
Variance age for survived193.868
227.772
Accuracy: 0.784553
Sensitivity: 0.695652
Specificity: 0.862595
8ms
8.5925ms

Program Terminated.
```

Both programs gave almost identical results. The only significant difference was that Naïve Bayes ran significantly faster.

Discriminative classifiers are classifiers that model boundaries between classes.

Generative classifiers are classifiers that model distributions between classes.

Reproducible research is important because it allows you to share your findings with others. In the case of education, it helps students learn how it works and to check that they are not making any mistakes. For others who may want to expand on the research, being able to quickly confirm they have the same results as you allows them to more quickly begin experimenting with it rather than spending time debugging your research.

It also allows for easier comparison between models. If two separately made models use separate data, it can be hard to compare the findings since you would need to show that the data doesn't contain any biases that may be the cause rather than the model itself. By having it be reproducible, you can exchange the datasets to ensure that the differences are in fact between the models and not the datasets.

Reproducibility can be created in a variety of ways. A major factor is any time a random function is called. If you set the seed to the random, it will always be same allowing for someone else to use the same random values. Reproducibility can also come from ensuring any libraries you are using aren't changing. This means noting which version you're using and possibly saving the files separately in case they update so you still have the correct version.

https://arxiv.org/pdf/1709.07267.pdf

https://blog.ml.cmu.edu/2020/08/31/5-reproducibility/