Logistic Regression Output:

```
titanic_project.csv
Reading line 1
heading: "","pclass","survived","sex","age"
Closing file titanic_project.csv.
Number of records: 1046

Training Time: 104.555s
w1: 0.999877
w2: -2.41086
accuracy: 0.784553
specificity: 0.862595
sensitivity: 0.695652

Program Terminated.
```

Naïve Bayes Output:

```
titanic_project.csv
Reading line 1
heading: "","pclass","survive
Closing file titanic_project.
Number of records: 1046

Prior Probs: 0.61  0.39

likelihood p(survived|pclass)
0.17213 0.22541 0.60246
0.41667 0.26282 0.32051

likelihood p(survived|sex)
0.15984 0.84016
0.67949 0.32051

Age Means: 30.418 28.826
Age Variances: 205.15 209.16

Training Time: 0.0054876s

accuracy: 0.78455
specificity: 0.8626
sensitivity: 0.69565

Program Terminated.
```

Analysis of Results:

We see by logistic regression, the training time took about 104 seconds. Naïve Bayes took about 0.005 seconds of training time. The accuracy of both algorithms were the same, 0.7845. Because of this, we can assume that both algorithms can predict the data decently. The specificity and sensitivity were also identical between both algorithms.

Generative classifiers vs.. Discriminative classifiers:

Generative classifiers are classifiers that try to model how a class would generate input (Malhotra). In other words, a generative model would try to explain how data was generated. Generative models will create new data instances to distinguish classes, using prior, likelihood, and posterior probabilities to get the conditional probabilities. (Goyal). An example of a generative classifier is Naïve Bayes. Discriminative classifiers are classifiers that try to find specific features in given input and create a conclusion that makes different classes distinguishable from each other (Malhotra). Discriminative models will directly use the functional form of conditional probability and estimate its parameters using the training data (Goyal). An example of a discriminative classifier is logistic regression.

They both learn different types of probabilities to generate conditional probabilities. Discriminative classifiers are typically more used than generative classifiers because they tend to be more accurate. While the generative classifiers would solve a general problem as an intermediate step, the discriminative classifiers differ as they will take a more direct approach in classifying (Malhotra). Generative models will typically have more bias than discriminative models and therefore, can require less data to train comparatively (Goyal).

Reproducible research in machine learning:

In short, reproducible research is when a researcher can reproduce a past computational experiment done by another researcher. Furthermore, it constitutes using the same data and methods by the original researcher as well given that the researcher reproducing it has the necessary computing power. Reproducing a past computational experiment would require that the quantitative results would be replicated as well (Shenouda and Bajwa).

Reproducible research is incredibly important in machine learning as it tests and can further prove the validity of hypotheses and theories. Being able to reproduce a computational experiment and output the same results is integral in backing up research claims (Shenouda and Bajwa). To be able to replicate data results by reproducing such experiments would lessen the risk of errors and therefore support the claims the original research provided (Ding).

Reproducibility can be implemented specifically when the original researcher has recorded the steps of the experiment to an incredibly specific degree, as the goal is for future researchers to replicate the experiment and get the same results. The most important requirement is that the data, code, and overall platform must be recorded accurately (Ding). This documentation would require that links to source codes and datasets are provided, the description of the algorithm be to adequate detail including

sample size, space complexity, and time complexity, and metrics be clearly defined and explained amongst several other things (Ding).

Works Cited

Ding, Zihao. “5 - Reproducibility.” *Machine Learning Blog | ML@CMU | Carnegie Mellon University*, 24 Aug. 2020, https://blog.ml.cmu.edu/2020/08/31/5-reproducibility/.

Goyal, Chirag. “Deep Understanding of Discriminative and Generative Models.” *Analytics Vidhya*, 19 July 2021, https://www.analyticsvidhya.com/blog/2021/07/deep-understanding-of-discriminative-and-generative-models-in-machine-learning/.

Malhotra, Akanksha. “Generative Classifiers v/s Discriminative Classifiers.” *Medium*, Medium, 16 Oct. 2019, https://medium.com/@akankshamalhotra24/generative-classifiers-v-s-discriminative-classifiers-1045f499d8cc.

Shenouda, Joseph, and Waheed U. Bajwa. “A Guide to Computational Reproducibility in Signal Processing and Machine Learning.” *ArXiv*, 27 Aug. 2021, https://arxiv.org/abs/2108.12383.