

```
Microsoft Visual Studio Debug Console
Opening file titanic_project.csv.
Reading line 1
","pclass","survived","sex","age"
Closing file titanic_project.csv.
3,0,1,19
3,1,0,22
3,1,1,20
3,0,0,1
2,0,1,63
1,0,1,38
3,0,1,19
2,0,1,39
3,1,0,17
3,0,0,3
2,0,1,19
1,0,1,28
2,0,1,34
3,0,1,20
3,0,1,32
2,0,1,42
1,1,1,30
2,1,0,54
3,0,1,8
3,0,1,28
2,1,0,42
3,1,0,15
2,1,0,14
3,1,1,25
3,1,0,22

Number of records: 1046
```

c)

To compare generative and descriptive classifiers, both perform classification and are beneficial depending on the situation. Metrics also affect these since depending on the situation the accuracy can depend more heavily for generative models due to specific variables.

To contrast generative and descriptive classifiers, generative usually require more data but are prone to not working well with outliers, but for descriptive this is different because they do work well with outliers. With generative depending on more data, it is also more expensive due to this.

d)

Reproducibility means to be replicated or reproduced (as evident by its name). For machine learning this translates to replicating the results of the work that is being referenced. The importance of this lies in the fact that without the ability to reproduce outcomes there is a lack of understanding on how to implement and follow the intended systems. This can be implemented by first attempting to replicate the environment rather than the work itself.

If the environment differs then it is obvious why results can differ. To go even beyond the environment, we can make sure that the environment's machine is replicated as well. Perhaps you are not testing correctly, meaning there could be something logical wrong with an implementation rather than something with the hardware or environment.

This doesn't even mention that there are also random variables that could be affecting the work, such as hidden data or accidentally implementing more than you need to. This could even mean that you introduced some unintended behavior by not implementing correctly.

Sources for part c) of document:

<https://towardsdatascience.com/generative-vs-discriminative-classifiers-in-machine-learning-9ee265be859e>

<https://www.analyticsvidhya.com/blog/2021/07/deep-understanding-of-discriminative-and-generative-models-in-machine-learning/>

Sources for part d) of document:

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

<https://towardsdatascience.com/reproducible-machine-learning-cf1841606805>

<https://neptune.ai/blog/how-to-solve-reproducibility-in-ml>