Summary

October 22, 2023

[ ]:
```
Riker Wachtler
22 October 2023
Wine Classification Midterm Project
https://github.com/RikerW/midterm

This is a classification attempt on a very unoriginal dataset (for the midterm,␣
 ↪I mean), the wine dataset from UCI, used via the sklearn databases.

The data is composed of 13 features; Alcohol, Malic acid, Ash, Alcalinity of␣
 ↪ash, Magnesium, Total phenols, Flavanoids, Nonflavanoid phenols,␣
 ↪Proanthocyanins, Color intensity, Hue, OD280/OD315 of diluted wines, and␣
 ↪Proline. They are not measured on any uniform scale. The samples are in 3␣
 ↪classes, so this is a classification problem on those features. It has no␣
 ↪missing values and has been cleaned up already.

I used a very simple ANN with one hidden layer, and used hyperparameter tuning␣
 ↪to decide the number of nodes in that layer. The input layer was obviously␣
 ↪based on the input dimension & the output the number of classes, so they␣
 ↪were not tuned. I used a Hyperband tuner to do this, with the normal epoch/
 ↪factor of 100/3. After that search, I fit the model on the training data␣
 ↪(randomly selected as 75% of the dataset) over 300 epochs, validated with␣
 ↪the remaining 25% test data from the dataset, and then plotted the accuracy/
 ↪loss & confusion matrix.

The precision of the model seems to be pretty great, the recall is pretty␣
 ↪solid, and the f1-score is quite solid overall. The accuracy is good as well.
 ↪ Overall, very few issues in the predictions on the test set, so this model␣
 ↪is probably not overfit. The accuracy vs. val accuracy was around 99% vs.␣
 ↪97%, so that's better obviously but really not a huge difference. All in␣
 ↪all, I think the model is well fit.
```