**Lab week 9: Neural Networks**          **Name: Kassi Winter**

**Jupyter Notebooks**

A quick introduction
https://colab.research.google.com/?utm_source=scs-index#scrollTo=GJBs_flRovLc


**Classification: Banknotes**

Link to Colab for this lab:
 https://colab.research.google.com/drive/1JO4VCEb5ApZUaSj5JQtyNxJa-njOEmlT?usp=sharing

1. Print screenshots of results from Banknotes0 runs in Colab

**Perceptron**
```
Results for model Perceptron
Correct: 541
Incorrect: 7
Accuracy: 98.72%
```

**SVC**
```
Results for model SVC
Correct: 545
Incorrect: 3
Accuracy: 99.45%
```

**K-NN**
```
Results for model KNeighborsClassifier
Correct: 547
Incorrect: 1
Accuracy: 99.82%
```

**GaussianNB**
```
Results for model GaussianNB
Correct: 463
Incorrect: 85
Accuracy: 84.49%
```

2.  Are there any conclusions you can draw from these results?
    From the results we can see that the most accurate model is 'KNN' and the worst model is 'GaussianNB'. Due to the dataset having multiple labels (variance, skewness, curtosis and entropy), these conclusions make sense because the KNN algorithm works really well for classification of clusters, captures nonlinear relationships between features and does not typically take into account outliers. Further, the 'GaussianNB' algorithm is not as accurate because it assumes the features are independent from the label and is sensitive to data outliers.

3.  Print screenshots of results from Banknotes run in Colab including how the accuracy improves over the different epochs.

```
Results for model Perceptron
Correct: 545
Incorrect: 4
Accuracy: 99.27%


Results for model SVC
Correct: 546
Incorrect: 3
Accuracy: 99.45%


Results for model KNeighborsClassifier
Correct: 549
Incorrect: 0
Accuracy: 100.00%


Results for model GaussianNB
Correct: 449
Incorrect: 100
Accuracy: 81.79%
```

Questions from the video from Week 5 Neural Networks from 45:50 to 53:00 where the code is explained and executed

4.  Why is the sigmoid function used in the output layer?
    The sigmoid function is set as the activation function when adding an output layer to the neural network. Sigmoid is used because the model is trying to predict probability, and because the sigmoid function fits the data to an s-shaped curve that results in a probability value of whether the given bill is real or counterfeit, it is a good choice for accurate results.

5. What is an epoch?

An epoch is when the training data is used all at once through an algorithm, allowing each sample to be exposed to the parameters / weights and updated accordingly.The number of passes, specified by the epoch number, acts as a hyperparameter for the algorithm, specifying the number of complete passes through the training dataset / learning process of the algorithm. For example, if the epochs was set to 20, it means "train the model to go through the data points 20 times to update the weights".

6. Run the program 3 times with the given parameters?  How consistent are the different runs?

After running the program three times with the parameters, the results are not very consistent (all differ around 0.2 at most from each other).

- Round #1 Accuracy: 0.9672
- Round #2 Accuracy: 0.8160
- Round #3 Accuracy: 0.9126

7. Increase the number of units in the hidden layer to 10?  Conclusion?.

After running the program three times with the parameters, the results are pretty consistent. This, however, seems to be overfitting the data.

- Round #1 Accuracy: 0.9909
- Round #2 Accuracy: 0.9927
- Round #3 Accuracy: 0.9872

8. Decrease the number of units in the hidden layer to 6?  Conclusion?

After running the program three times with the parameters, the results are more accurate.. however are not very consistent.

- Round #1 Accuracy: 0.9380
- Round #2 Accuracy: 1.0000
- Round #3 Accuracy: 0.9290