This project aims to compare a performance of two AlexNet models on CIFAR100 dataset using 2 kinds of softmax functions:

1) Normal softmax
2) Gumbel Softmax

Both models were trained for 20 epochs using a learning rate of 1e-2 and SGD with momentum = 0.9.

Some of the important comparisons between them is given below:

Criteria	**Normal Softmax**	**Gumbel Softmax**
Average running time per epoch during training	0.307948 min	0.322585 min
Accuracies on test set	41.13 %	42.77 %
f1 scores	0.2530843183709426	0.2737336973005763
Total time taken to run test images	0.048795 min	0.045898 min

The precision, recall and confusion matrices are present in the Jupyter Notebooks.

Thus, overall for a training period of 20 epochs, using Gumbel Softmax instead of Normal Softmax had the following differences:

- improved accuracy
- took slightly more time per epoch during training
- took slightly less time during testing
- gave a better f1 score.

If trained for a larger number of epochs and on a deeper network, we may have got a bigger/opposite trend in these differences in values.