This project aims to compare a performance of 2 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

Criteria	Normal Softmax	Gumbel Softmax
Average running time per epoch during training	0.320305 minutes	0.295297 minutes
Accuracies on test set	42.09%	42.21%
f1 scores	0.2604820515648541	0.23373661016973118

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 improved accuracy slightly and took slightly less time per epoch but gave a slightly poorer f1 score.

If trained for a larger number of epochs and on a deeper network, we may expect a bigger difference.