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 40 epochs using a learning rate of 1e-2

Criteria	Normal Softmax	Gumbel Softmax
Average running time per epoch during training	0.307636 min	0.291476 min
Accuracies on test set	41.06%	42.57%
f1 scores	0.21930037407744415	0.2514361081240062

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

Thus, overall for a training period of 40 epochs, using Gumbel Softmax instead of Normal Softmax improved accuracy slightly and took slightly less time per epoch and gave a better f1 score.

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