CORE MACHINE LEARNING

The CIFAR-100 dataset has 600 images of 100 classes.Out of which 500 are training and 100 are testing images.
The training datasets were transformed(normalized the pixel value between 0 and 1) and also flipped to ensure unbias training.
Subsequently the dataloaders were created.
Architecture

I have trained the model using Resnet18 architecture.The model was also trained on Resnet50 but couldn't get substantially different results.The architecture was implemented from scratch.
Resnet-18 consists of convolutional layers,residual blocks and fully connected layers.
The last fully connected layer will have 100 units to match the 100 classes in CIFAR-100

Standard Softmax:

It computes the probability distribution over the classes.
The time complexity of the softmax function is O(k) where K is the number of classes.

Training:

Cross entropy loss was used since we have a multi-class classification problem.
SGD was used as a optimizer and other learning parameters were set.
The images were trained in batches
The hyper parameters were adjusted ,I saw a trend of accuracy increase till the half way of batch instance and in the second half the accuracy decrease.This might be because of batch of training input fed into the model.

Second Model:

The alternative like Gumbel softmax is being implemented in the last layer of Resnet-18.Gumbel is basically sampling from Gumbel distribution and performing a softmax .Hence overall time complexity is O(k*logk).