## TRANING LARGE MODEL BY SMALL MODELS BY ENSEMBLE KNOWLEDGE TRANSFER– ARULNITHI P(12210280)

Ensemble learning is a powerful technique that involves combining the predictions of multiple models to improve overall accuracy and robustness. Bidirectional ensemble knowledge transfer refers to the process of transferring knowledge between two sets of models in both directions.

**BIDIRECTIONAL ENSEMBLE KNOWLEDGE TRANSFER EXPERIMENT 1 – HOMOGENOUS DATASET FOR TRANING**

Objective is to train the large model VGG19 by ensemble small model such as simple CNN, Resnet_8,Resnet_50 and Squeeze net. I have used CIFAR10 data set for the classification. 70% of the data is provided to small models for traning and 30% for testing the target model and update the parameters by bidirectional ensemble knowledge transfer between the target model VGG19 and small ensemble models. Bidirectional ensemble knowledge transfer refers to the process of transferring knowledge between a target model (VGG19 in your case) and a set of small ensemble models (ResNet8, ResNet50, and Squeeze Net in both directions.

The first step in this process is to train the small ensemble models on a subset of the CIFAR10 dataset. Once the small ensembles are trained, they can be used to transfer knowledge to the target model by using a technique called knowledge distillation. This involves training the target model to mimic the behavior of the small ensembles by minimizing the difference between their predictions on the same inputs.

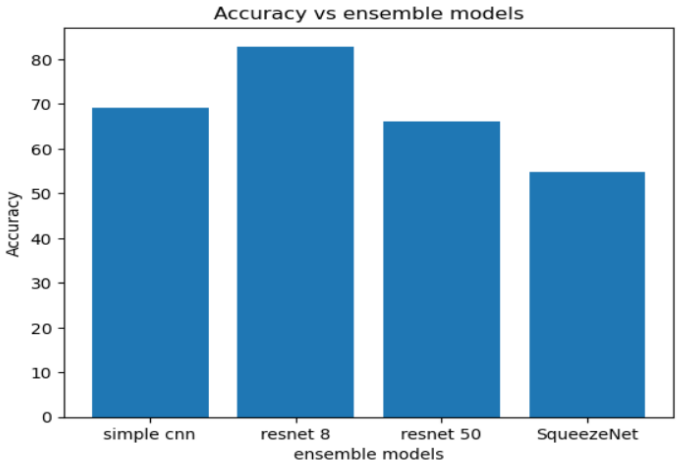
In the code, you are using a bidirectional approach to knowledge transfer, which means that you are not only transferring knowledge from the small ensembles to the target model, but also from the target model back to the small ensembles. This is done by training the small ensembles again using the outputs of the target model as additional training data.

The idea behind bidirectional knowledge transfer is to improve the performance of both the small ensembles and the target model by leveraging the strengths of each model. The small ensembles can provide valuable information to the target model about how to make more accurate predictions, while the target model can provide feedback to the small ensembles to help them improve their performance. By exchanging knowledge in both directions, the models can learn from each other and become more accurate over time.

RESULTS

In Experiment 1 all the ensemble models are trained on the same dataset CIFAR-10 (The CIFAR-10 dataset is an image dataset that has 10 classes. It consists of 60,000 32x32 color images in 10 classes, with 6,000 images per class. The dataset is split into 50,000 training images and 10,000 testing images. Each class represents a different object or animal, such as airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck).

There are four GPUs in the setup and each GPU is given a specific ensemble model to train. Training is done for 50 epoch.




**Experiment setup 2 – Heterogenous dataset**

Here we are training the large model by heterogenous data set such as simple CNN is trained on MNIST, Resnet_8 is trained on CIFAR10, Resnet 50 is trained on fashion MNIST and Squeeze net is trained on The German Traffic Sign Recognition Benchmark (GTSRB). Then in the final target VGG19 is evaluated on random subset of the above datasets.