Training with Pruning

[TOC]

⚠ Disclaimer: All datasets hyperlinked from this page are not owned or distributed by Google. The dataset is made available by third parties. Please review the terms and conditions made available by the third parties before using the data.

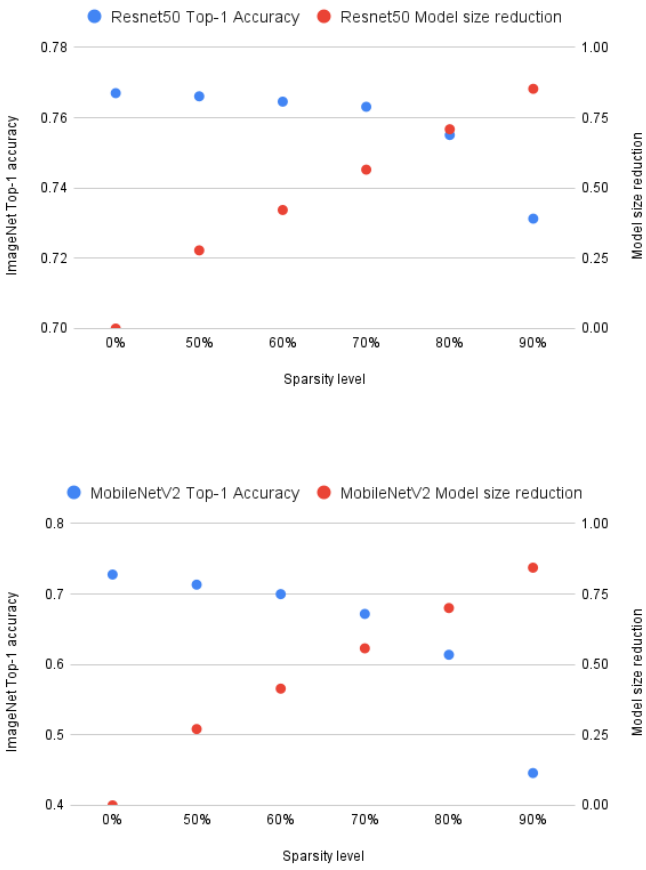
Overview

This project includes pruning codes for TensorFlow models. These are examples to show how to apply the Model Optimization Toolkit's pruning API.

How to train a model

```
EXPERIMENT=xxx  # Change this for your run, for example, 'resnet_imagenet_pruning'
CONFIG_FILE=xxx  # Change this for your run, for example, path of
imagenet_resnet50_pruning_gpu.yaml
MODEL_DIR=xxx  #  Change this for your run, for example, /tmp/model_dir
python3 train.py \
  --experiment=${EXPERIMENT} \
  --config_file=${CONFIG_FILE} \
  --model_dir=${MODEL_DIR} \
  --mode=train_and_eval
```

Accuracy



Comparison of Imagenet top-1 accuracy for the classification models

Note: The Top-1 model accuracy is measured on the validation set of ImageNet.

Pre-trained Models

Image Classification

Model	Resolution	Top-1 Accuracy (Dense)	Top-1 Accuracy (50% sparsity)	Top-1 Accuracy (80% sparsity)	Config	Download
MobileNetV2	224x224	72.768%	71.334%	61.378%	config	TFLite(50% sparsity),
ResNet50	224x224	76.704%	76.61%	75.508%	config	TFLite(80% sparsity)