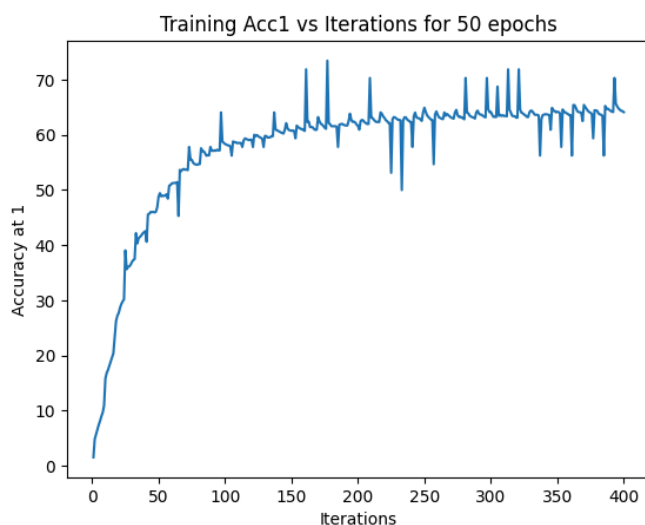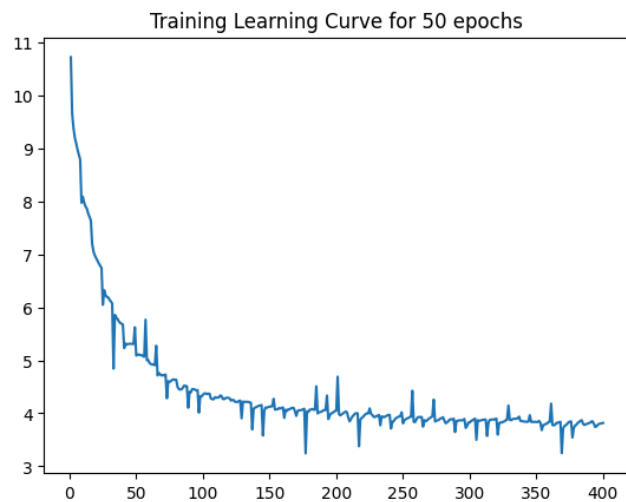# Hinton's Knowledge Distillation Implementation on CIFAR
## Final 240th Epoch

```
 * Acc@1 88.564 Acc@5 98.336
epoch 239, total time 39.13
Test: [0/313]   Time 0.287 (0.287)     Loss 0.7728 (0.7728)    Acc@1 78.125 (78.125)   Acc@5 90.625 (90.625)
Test: [100/313] Time 0.011 (0.013)     Loss 0.8055 (1.1366)    Acc@1 78.125 (72.803)   Acc@5 93.750 (91.955)
Test: [200/313] Time 0.012 (0.012)     Loss 0.9593 (1.1158)    Acc@1 68.750 (73.010)   Acc@5 96.875 (92.304)
Test: [300/313] Time 0.017 (0.011)     Loss 1.5038 (1.1126)    Acc@1 75.000 (73.059)   Acc@5 93.750 (92.483)
 * Acc@1 73.120 Acc@5 92.500
==> training...
Epoch: [240][0/782]     Time 0.660 (0.660)      Data 0.526 (0.526)      Loss 1.7971 (1.7971)    Acc@1 84.375 (84.375)   Acc@5 98.438 (9
8.438)
Epoch: [240][100/782]   Time 0.048 (0.055)      Data 0.003 (0.008)      Loss 1.5487 (1.7557)    Acc@1 89.062 (88.165)   Acc@5 98.438 (9
8.267)
Epoch: [240][200/782]   Time 0.047 (0.051)      Data 0.002 (0.006)      Loss 2.0024 (1.7411)    Acc@1 81.250 (88.697)   Acc@5 98.438 (9
8.422)
Epoch: [240][300/782]   Time 0.047 (0.050)      Data 0.002 (0.005)      Loss 1.9450 (1.7407)    Acc@1 82.812 (88.741)   Acc@5 98.438 (9
8.417)
Epoch: [240][400/782]   Time 0.049 (0.049)      Data 0.002 (0.004)      Loss 1.9100 (1.7356)    Acc@1 89.062 (88.794)   Acc@5 98.438 (9
8.391)
Epoch: [240][500/782]   Time 0.048 (0.049)      Data 0.002 (0.004)      Loss 1.7911 (1.7464)    Acc@1 89.062 (88.526)   Acc@5 96.875 (9
8.356)
Epoch: [240][600/782]   Time 0.048 (0.049)      Data 0.003 (0.004)      Loss 1.8810 (1.7430)    Acc@1 87.500 (88.597)   Acc@5 98.438 (9
8.398)
Epoch: [240][700/782]   Time 0.048 (0.049)      Data 0.003 (0.004)      Loss 1.7189 (1.7457)    Acc@1 89.062 (88.608)   Acc@5 100.000 (9
8.404)
 * Acc@1 88.620 Acc@5 98.376
epoch 240, total time 38.55
Test: [0/313]   Time 0.255 (0.255)     Loss 0.7968 (0.7968)    Acc@1 78.125 (78.125)   Acc@5 90.625 (90.625)
Test: [100/313] Time 0.009 (0.013)     Loss 0.8443 (1.1435)    Acc@1 81.250 (72.803)   Acc@5 93.750 (91.863)
Test: [200/313] Time 0.009 (0.011)     Loss 0.9344 (1.1228)    Acc@1 71.875 (72.994)   Acc@5 96.875 (92.226)
Test: [300/313] Time 0.006 (0.011)     Loss 1.5099 (1.1202)    Acc@1 75.000 (73.038)   Acc@5 93.750 (92.442)
 * Acc@1 73.150 Acc@5 92.470
==> Saving...
best accuracy: tensor(73.3100, device='cuda:0')
```

## Training curve for 50 epochs



Training Learning Curve for 50 epochs



Training Acc1 vs Iterations for 50 epochs

# Test Curves for 50 epochs

## Test loss vs Iterations for 50 epochs



## Test Acc1 vs Iterations for 50 epochs