ES667: Project Proposal: The Lottery Ticket Hypothesis

Aditya Mehta - 22110017 | Nikhil Goyal - 23110218 | Shardul Junagade - 23110297

Project Overview

The Lottery Ticket Hypothesis (LTH) proposes that dense neural networks contain sparse subnetworks ("winning tickets") that can achieve comparable accuracy to the full network when trained from their original initialisation. This project aims to empirically validate this hypothesis and investigate the factors that determine winning tickets.

We will implement the Iterative Magnitude Pruning (IMP) algorithm: train a network, prune the weights, reset the weights to their original initialisation, and retrain. This process repeats until reaching 80-90% sparsity. The key insight is that initialisation matters—randomly reinitialised sparse networks fail to match performance.

Reference Paper: "The Lottery Ticket Hypothesis: Finding Sparse, Trainable Neural Networks" Jonathan Frankle and Michael Carbin, ICLR 2019 | LINK

Methodology

- **Architectures:** LeNet-5 for MNIST, small CNN for CIFAR-10
- **Pruning Strategy:** Iterative magnitude pruning (20% per round)
- **Comparison Baselines:**
    - Magnitude pruning + original initialisation (winning tickets)
    - Magnitude pruning + random reinitialization
    - Random pruning + original initialisation
- Validate that winning tickets achieve comparable accuracy at sparsity (80%+)
- Demonstrate that reinitialization destroys the winning ticket property
- Test "late rewinding" (resetting to epoch k > 0)
- Visualise surviving weight patterns and pruning masks

Expected Outcomes

1. **Primary Validation:** Demonstrate sparse subnetworks (80-90% pruned) with original initialisation match dense network accuracy within 2%
2. **Initialisation Dependency:** Show winning tickets outperform randomly reinitialised sparse networks by 10-15% accuracy
3. **Practical Insights:**
    - Identify optimal pruning schedules and rates
    - Quantify computational savings (5-10× reduction in parameters)
    - Determine transferability across architectures
4. **Architectural Understanding:**
    - Visualise which weight structures survive pruning
    - Analyse loss landscape differences between winning tickets and random sparse networks after pruning and retraining
    - Investigate the predictability of winning initialisations