

assignment1

October 2, 2025

Assignment 1 - Deep Learning-E25

2nd October 2025

1 Part 1:

1.0.1 a: In your opinion, what were the most important turning points in the history of deep learning?

From my perspective, the history of deep learning has several key turning points. Early neural networks and perceptrons introduced the idea of learning from data, but training deep networks became practical only after backpropagation was developed. The 2006 paper on deep belief networks and the term “deep learning” marked a significant turning point, separating it from traditional neural networks.

Later, the success of GPU-accelerated CNNs could achieve better results in image recognition. And after few years, plenty of architectures like RNNs, LSTMs, and Transformers were developed, which revolutionized natural language processing and sequence modeling. Now we Have ChatGPT and other large language models that have transformed how we interact with AI. Personally, OpenAI Five was the also an impressive shift, because now I might get to play against a bot that has 99.4% winrate.

1.0.2 b: Explain the ADAM optimizer.

The ADAM (Adaptive Moment Estimation) optimizer is an advanced optimization algorithm that combines the benefits of two other popular optimizers: AdaGrad and RMSProp. It computes adaptive learning rates for each parameter by maintaining two moving averages: the first moment (mean) and the second moment (uncentered variance) of the gradients. This allows ADAM to adjust the learning rate for each parameter individually, improving convergence speed and stability. The algorithm is particularly well-suited for training deep learning models and has become a default choice for many practitioners due to its effectiveness and ease of use.