

GPT Mastery: A Comprehensive Guide To Crafting Your Own ChatGPT

Week 3

Let's move on to week 3 and continue building Makemore, delving into more complex concepts of training a neural network.

1.) Building makemore Part 2: MLP

This video covers implementing a multilayer perceptron (MLP) character-level language model. It also introduces machine learning basics, such as model training, learning rate tuning, hyperparameters, evaluation, train/dev/test splits, and addressing under/overfitting.

Youtube:  Building makemore Part 2: MLP

2.) Building makemore Part 3: Activations & Gradients, BatchNorm

In this video, we explore MLP internals, focusing on forward pass activations and backward pass gradients, along with their scaling issues. Diagnostic tools and visualizations for deep network health are discussed. The video also introduces Batch Normalization, a key innovation for stabilizing deep neural network training. Residual connections and the Adam optimizer will be covered in future videos.

Youtube:  Building makemore Part 3: Activations & Gradients, BatchNorm

Github: <https://github.com/karpathy/makemore>