

This week, I focused on studying how to build a language model from scratch using PyTorch. I began by understanding the importance of good data and how to preprocess it properly to train an effective model. Then, I learned about recurrent neural networks (RNNs) and how they can be used to process sequential data. I explored different techniques for improving the performance of RNNs, such as using multilayer RNNs, LSTMs, and regularization techniques like dropout and activation regularization. Finally, I trained a weight-tied regularized LSTM language model and experimented with different hyperparameters to achieve the best performance. Overall, I gained a deeper understanding of how to build and train effective language models and the various techniques used to improve their performance.