GPT Mastery: A Comprehensive Guide To Crafting Your Own ChatGPT

Assignment

Week 2

Complete the following task in a Colab notebook:-

1.) Train a trigram language model, i.e. take two characters as an input to predict the 3rd one. Feel free to use either counting or a neural net. Evaluate the loss; Did it improve over a bigram model?

2.) Split up the dataset randomly into 80% train set, 10% dev set, 10% test set. Train the bigram and trigram models only on the training set. Evaluate them on dev and test splits. What can you see?

3.) Use the dev set to tune the strength of smoothing (or regularization) for the trigram model - i.e. try many possibilities and see which one works best based on the dev set loss. What patterns can you see in the train and dev set loss as you tune this strength? Take the best setting of the smoothing and evaluate on the test set once and at the end. How good of a loss do you achieve?

4.) We saw that our 1-hot vectors merely select a row of W, so producing these vectors explicitly feels wasteful. Can you delete our use of F.one_hot in favor of simply indexing into rows of W?

5.) Look up and use F.cross_entropy instead. You should achieve the same result. Can you think of why we'd prefer to use F.cross_entropy instead?

6.) Meta-exercise! Think of a fun/interesting exercise and complete it.

SUBMISSION LINK: https://forms.gle/mfZ6K1DSExjYUWNs6

NOTE: Please ensure you submit the assignments. While there are no specific deadlines, you must submit at least 4 assignments by the end of this project. Failure to do so will result in not receiving the certificates. This is important as we need to present your work to WnCC for approval of your resume point. Please take this seriously.