

CSCI 5521: Machine Learning Fundamentals (Fall 2022)¹

Extra Credit Homework -2 (5%) (Monday, Nov. 21)

Due on Gradescope at 11:59 PM, Tuesday, Nov 29

1. **(100 points)** In this programming exercise, you will operate on PyTorch to implement a neural network for natural language processing. In the previous assignments, you learned about feed-forward and convolution networks. This extra credit assignment aims to get you exposure to Recurrent Networks (RNN).
 - (a) **(50 points)** In this question, you will build and train a character-level RNN to classify words. A character-level RNN reads words as a series of characters - outputting a prediction and “hidden state” at each step, feeding its previous hidden state into each next step. We take the final forecast as the output, i.e., which class the word belongs to for the final prediction. The dataset and the corresponding code template are available at [Link](#). Specifically, you will be training on a few thousand surnames from 18 languages of origin and predict which language a name belongs to based on the spelling. Build over the code, understand it thoroughly, try different parameters and see if you can improve the existing model.
 - (b) **(50 points)** In this question, you will implement an RNN to generate names from character level RNN. This problem differs from the above problem in the sense that instead of predicting a category after reading all the letters of a word, we input a category and output one letter at a time. E.g., If the input is Russian, then your model will output a Russian name. The dataset and the corresponding code template are available at [Link](#). Build over the code, understand it thoroughly, try different parameters and see if you can enhance the existing model.

For grading, you must submit a 5-10 page report describing your model in detail and your understanding of the model and the dataset. Since the code for the problem is available, you will be graded based on your understanding of the model and how rigorously you tried different hyperparameters, the accuracy you got, and what you learned from these experiments. Also, for every model run, in addition to monitoring loss/accuracy, try to come up with different examples where your model did a phenomenal job and where your model failed miserably. Also, what other potential application can you see where you can use your model architecture? Overall, the report will be graded based on your thoroughness and not on how much improvement you achieved.

Submission

- Things to submit: You need to submit the materials electronically to the Gradescope. **For this assignment, you need to submit the materials [ExtraCredit2-Written (project report) and the ExtraCredit2 programming (the compressed zip file which include all the Python codes you wrote)].**

¹Instructor: Kshitij Tayal (tayal@umn.edu). TA: Xianyu Chen, and Yoshitaka Inoue (csci5521@umn.edu).