HY673 - Spring 2024 Problem Set #2
Yannis Pantazis (17/04/2024)

Instructions

- **Due date:** Sunday April 28th, 2024
- Submission via e-mail to the class account: hy673@csd.uoc.gr
- Provide one file with the written solutions.
- Provide one folder with code.
    - The name of each file in the folder should indicate the respective exercise.
    - Your code should run on colab.
- All assignments in this course are individual, not group, assignments. You may freely discuss homework assignments with your fellow classmates. The final solutions, however, must be written entirely on your own. This includes programming assignments.
- You are allowed to use Generative Al Tools such as ChatGPT for help with homework assignments for grammatical corrections only. To maintain academic integrity, students must disclose any use of Al-generated material.

Problem 1 (Conditional generation of names)**.** *In this exercise, you will extend the AR model presented in Tutorial 5 to the case of conditional generation of names. For each country, you will assign a one-hot encoding and pass it as an extra argument -implemented via vector concatenation- to the RNN. Write a program that takes as input a country name and returns a list of 10 names. Repeat the same for the LSTM and GRU architectures as well as for the Transformer architecture (from Tutorial 6).*

Problem 2 (From numbers to words)**.** *(a) Using num2word function (`https://pypi.org/project/num2words/`) create a dataset of numeral - sentence pairs. For instance, a sample is (1033, "one thousand thirty three"). Use numbers up to 10^9 while the size of the dataset will be 10^5.*

(b) Implement and then train an GRU that converts a number to its sentence analog. Plot the training loss as a function of iterations. Report the accuracy on unseen numbers. Report also what is the result when a number outside the training range is given (e.g, 101001001001 or 1001001001001).

(c) Repeat (b) but now using the Transformer architecture.

Problem 3 (Music Transformer)**.** *Go to `https://github.com/gwinndr/MusicTransformer-Pytorch` and download the Music Transformer. Train on MAESTRO dataset and then generate new melodies for various instruments (eg, piano, base, etc.)*