Gebze Technical University
Department of Computer Engineering
CSE 654 / 484
Fall 2023

Homework 03
Due date: Jan 29th 2024
No Late Submissions

You are expected to use the Python language and the Keras library on Google Colabs in this homework. You will prepare a report including your code and results (in a Google Colabs Jupyter Notebook). The report format is given at the end of this document.

Go to https://colab.research.google.com/notebooks/welcome.ipynb and familiarize yourself with the Google Colaboratory which is a free Jupyter notebook environment that requires no setup and runs entirely in the cloud. With Colaboratory you can "write and execute code, save and share your analysis, and access powerful computing resources, all for free from your browser".

Then read and understand the deep learning example at https://colab.research.google.com/github/lmoroney/mlday-tokyo/blob/master/Lab1-Hello-ML-World.ipynb . Similar examples can be found at

https://colab.research.google.com/github/lmoroney/mlday-tokyo/blob/master

Also, you may want to watch associated videos such as https://youtu.be/KNAWp2S3w94

In this homework, you will write a classifier that test if the Turkish "de" and "ki" suffixes should be separated or not. For example, in "Öğrenciler de geldi" it is separated, but in "Öğrencilerde gelişme var." it is not separated.

Your network will take a sentence as an input and it will produce a true (separated) or false for a single word in the sentence that can have this ambiguity. The words should be given as not separated as input.

You will get your training set yourself from web sources. Wikipedia dump of Turkish is a nice source. Turkish Ministry of Education textbooks are other good sources.

What to submit to the Teams page

- Your source code
- The data that you have used for training.
- The model details.
- Your report that includes your method details and your performance tables including how you measured your performance.
- test run results for at least 20 sentences