# Gebze Technical University
# Department of Computer Engineering
# CSE 654 / 484
# Fall 2022

# Homework 04
# Due date: Jan 23th 2023
# 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 ColabsJupyter 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 develop a simple translator from Ottoman Turkish to modern day Turkish. We provide a training set at the class Teams page. As an example if the input is "itilaf devletleri , mütareke ahkamına riayete lüzum görmüyorlar" the translated output will be "itilaf devletleri , ateşkes hükümlerine uymayı gerekli görmüyorlar ".

There are many good deep learning translation models available at the web. Find one with or two of them based on LSTM methods (no transformer based models allowed), train it with our data set and measure the final BLEU score on your system results. Your performance should be on a separate test set, so do not use all of your data for training.

These are the experiments that will perform
1- An experiment that measures the effects of pretraining of your model with other Turkish set. This will show if pre-training helps.
2- An experiment that measures the effect of data size: run several experiments with different amount of data for the training and see the final results.
3- Change the attention and LSTM parameters of your model and report our results.

What to submit to the Teams page
- Your source code
- Your report that includes your method details and your performance tables
- test run results for at least 20 sentences