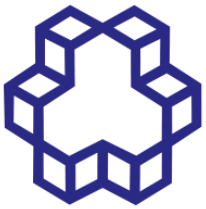

K.N. Toosi University of Technology

Faculty of Electrical and Computer Engineering

Machine Learning

Bonus Project

Table of content

Question 1

We know that transformer networks, unlike recurrent networks, are non-sequential, which makes them somewhat weaker in time series problems. Various methods exist to address this challenge through modifications in positional encoding, the attention module, and at the architecture level. In this task, you will become familiar with one of the most recent proposals, called the iTransformer network, and implement it yourself. The related paper is indexed in ICLR 2024, and the corresponding model code is available on their GitHub.

Since the model is new and there aren't many resources available about it, it is essential to read both the paper and the GitHub repository to understand the concept of the iTransformer. You are required to implement the predictive model using both the transformer and iTransformer networks, so you can compare their performance. The training data is related to the exchange rates of 8 countries over time, and since there are 8 variables to predict, it is considered a multivariate problem.

The strength of the iTransformer will particularly show in problems with high-dimensional variables over time. Note that both the transformer and iTransformer models must be written by you without using predefined functions such as nn.transformer (you must write the attention and FeedForward blocks yourself, but you may use nn.layernorm for layer normalization). After implementing both models, you will notice the differences between them

Question 2

1. Article Introduction

In the **article** used for this exercise, an X-ray image of the lungs of COVID-19 patients and healthy individuals is used to design a Convolutional Neural Network (CNN). Given the life-threatening nature of this disease, training a network with high accuracy is of great importance.

You can download the article from this link:
https://onlinelibrary.wiley.com/doi/epdf/10.1155/2021/6621607

2. Data Collection and Image Preprocessing (25 points)

In the data collection section, the method of collecting data is explained. This dataset has been provided for you to get started. Therefore, for this part, you will begin with preprocessing the data.

To preprocess the data, according to the explanation in the article, perform the following: Implement four types of **data augmentation** step by step. It is recommended to do this in such a way as to prevent overfitting in the network. This means you should first apply the first type to the dataset, then train the network and analyze the results. Repeat the process for the second type, and so on. Also, perform **normalization** as described.

3. Network Training (35 points)

Train the CNN based on the provided descriptions. Implement the network and train it on the training data. Plot the Accuracy and Loss charts for the training and validation data.

4. Network Evaluation (30 points)

Evaluate the trained network on the test data and report the evaluation metrics, including Accuracy, Precision, Sensitivity, Specificity, and F1 score. Also, plot the confusion matrix. Do not forget that, in the article, the preprocessing stage was performed under the supervision of a medical expert for data selection. Since we cannot perform this step, your network may perform differently compared to the article

5. Extra Network Evaluation (10 points)

Report the accuracy of the network based on the first column of Table 6 for tests with different numbers of layers for the network. Keep in mind that it is necessary to train and test the network each time.

Note:

Students with odd student numbers must implement the first dataset (**dataset1**) focusing on COVID-19, while students with even student numbers must implement the second dataset (**dataset2**) and follow the same procedure for Pneumonia instead of COVID-19.

Students whose second-to-last student number digit is odd must implement the first dataset (**dataset1**) focusing on COVID-19, while students whose second-to-last digit is even must implement the second dataset (**dataset2**) and follow the same procedure for Pneumonia instead of COVID-19.

Datasets are available at :

https://drive.google.com/file/d/1wM1NufVrRtbHuLmeBjiOls_e80Qrsoy-/view?usp=drive_link

The use of pre-written or readily available codes from the internet is strictly prohibited. Any similarity or correlation between the submitted codes of different students or between different parts of an individual assignment will result in a penalty of **-50 points** and will be considered academic misconduct.

The total points will be allocated after the presentation and submission of a well-prepared report. The individual responsible for writing the code must be clearly identified in the report. Failure to comply with these requirements may result in a deduction of points.