CH6870 : Assignment

Submission Date: 14.10.2024 **TOTAL MARKS: 70**

General instruction:

1. Solve the assignment mentioned below, submit a detailed report by 14th October 2024, 11.59PM. There would be no extension for the submission date.
2. A test function is given to you, and you should generate the data as per the function given.
3. For every student, the test function is different. Kindly check the excel sheet provided to find the test function given to you.
4. The coding language should be PYTHON.

Write a PYTHON code for the following problem.

1. Using the test function given to you, generate your own nonlinear dataset. **(10M)**
   a. Ensure that you have enough number of samples before training the ANN.
   b. Write the description of the data (as comments in the code and in the report).
   c. Save the data into an excel sheet.
   d. Read the data from the saved excel file and split the data into inputs and outputs. (The function values must be taken as the output and the rest columns as inputs)
   e. Normalize the data between 0 and 1 before training the ANN.
   f. Divide the data into 3 parts for training, validation, and testing (70%:15%:15%)
2. Create an artificial neural network in PYTHON with **any number** of nodes, **any number of** hidden layers and **any** activation function. **(5M)**
3. Train the neural network using **Adam** optimizer and **MSE** loss function. **(5M)**
4. Report the following: **(5M)**
   a. Mean squared error for training data.
   b. R^2 values for training data.
5. Predict the outputs for testing dataset using the trained network. Report R^2 values for the testing dataset. **(5M)**

6. Now train the same network using **RMSProp** optimizer and draw comparisons between Adam and RMSProp (report relevant metrics to justify your observations). **(5M)**
7. Perform an analysis by varying the number of hidden layers, number of hidden nodes, activation functions, number of epochs and sample size for training. (At least three variations in each of the parameters must be done). Plot a graph between the predicted outputs and the true outputs (for the training data and testing data) using the MATPLOTLIB package for each of the analysis performed. **(15M)**
8. **Write a detailed report with reference to the data considered for doing the assignment. Include in this report the following:**
   a) **How to read data from an excel sheet, normalize the data, split the data into train and test dataset and generate an ANN model for the given dataset in PYTHON. (5M)**
   b) **Report the metric values, inferences and plot the original vs predicted values (wherever needed) for Q2-Q6. (5M)**
   c) **Include all possible inferences, observations, and comparisons in the report with relevant figures for Q7. (10M)**
   d) **Submit all the codes and datafiles including the report in a folder and name the folder as Your name_ML and zip the folder while attaching in the google classroom.**