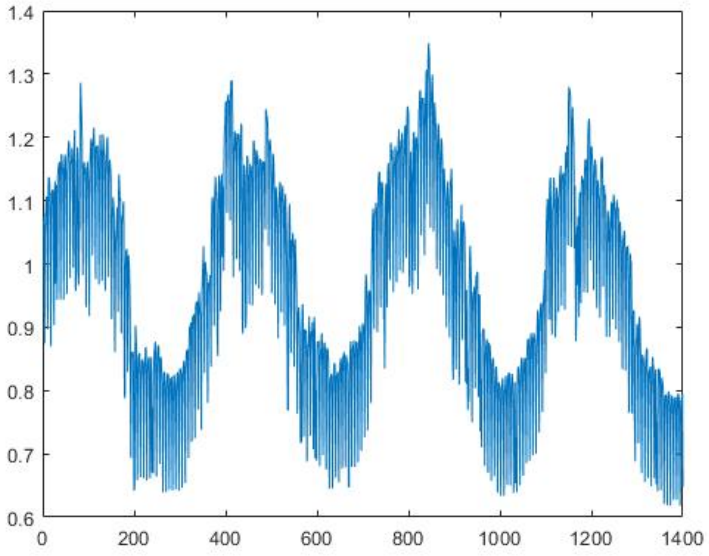
CMPG-765/CMPT- 465 Neural Networks and Learning Systems
Homework-8

Time series prediction (You may get 110 points)

"**Power consumption in Poland**" is one of the benchmark datasets used for testing time series prediction tools.

This dataset consists of the data measured daily in 1990s during a period of 1,400 days, that is almost 4 years. It is pretty clear that power consumption is a quasi-periodic process – it increases from mid fall to late winter and decreases from early spring to early fall. These trends are well visible from the following plot:



The **Poland_Power_Consumption_Load_1990s.mat** file, which was shared with you, contains 3 arrays: Data (original time series data) Data_MLP (original data transformed for the use with MLP by rescaling the data to fit a series in the interval [-1.5, 1.5]), and Data_MVN (original data transformed for the use with MLMVN by rescaling the data to fit a series in the interval [π/4, 7π/4] – in both of these rescaled arrays "gaps" (0.5 and π/4, respectively) were left to make it possible that smaller/larger values than the original min and max should be potentially predicted)

1. (**15 points**) Design a function to create a learning set. This function should accept the following arguments: an array containing a time series, n – the number of members in the series, on which the following member depends $x_{n+1} = f\left(x_1,...,x_n\right)$, k – the number of samples to be included in the learning set.

2. (**25 points**) Using the function designed in Task 1 create a learning set for MLP with n = 183, k=365 (that is, power consumption tomorrow depends on its consumption for the preceding 6 months, and the learning process should be based on the historical data for 1 year).
   1) Train MLP with 2 hidden layers using this learning set (a software simulator was shared with you).

      Undergraduate students. You should use 6 or 8 hidden neurons in the first hidden layer and significantly larger number of neurons in the second hidden layer (significantly larger means from dozens to hundreds, and to thousands of neurons).
      Graduate students (**15 extra credit points for undergraduate students**) You should make 4 experiments with 2, 4, 6 and 8 hidden neurons in the first hidden layer and significantly larger number of neurons in the second hidden layer (significantly larger means from dozens to hundreds, and to thousands of neurons).
3. (**20 points**) Design a function/script for making predictions based on the results of learning. This function should use the TestingMLP function in a for loop, which should be repeated as many times as the number of predictions to be made. Prediction should start from predicting a member of the series following the last desired output in the learning set based on the actual series members. To predict the next member, a value predicted on the preceding step should be used as it is shown in slides 7-8 of Lecture-12:

Inputs			Actual output	Desired output	
x_{k+1-n}	x_{k+2-n}	...	x_k	y_{k+1}	x_{k+1}
x_{k+2-n}	x_{k+3-n}	...	y_{k+1}	y_{k+2}	x_{k+2}
...	...	...	...	...	
...	...	...	y_{m-1}	y_m	x_m

   Predict 90 values (that is, make your prediction for the next 3 months) and find RMSE between ***y*** values (predicted values) and ***x*** values (actual series members).

4. (**45 points**) Repeat what was done in Tasks 2-3 with MLP, but with MLMVN. Use Net_Learn function for learning and Net_Test for function prediction.
5. (**5 points**). Summarize your results in a brief report and turn it in along with a source code, which you designed.