

CMSC 409: Artificial Intelligence

Fall 2017, Instructor: Dr. Milos Manic, <http://www.people.vcu.edu/~mmanic>

Project 3

CMSC 409: Artificial Intelligence

Project No. 3

Due Nov. 20, 2017, noon

Pr.3.

You are given hourly temperature data from 5:00AM to 1:00PM for 3 days. Using this data, predict the temperature of the 4th day (hint: you can use your code from project 2).

The data is given in hourly increments in four text files (please download “Pr3_data.zip”). The training data for the first 3 days are named “*train_data_X.txt*” where “X” is the date. Use this data for training.

Once the decision unit is trained, predict the temperatures for the 4th date for hourly increments from 6:00AM to 1:00PM. Use the testing data to calculate the error of the prediction. The testing data is in the file “*test_data_4.txt*”. (Note: do not use this data for training, the testing data is only for testing.)

- 1) What are the inputs and outputs for this problem?
- 2) What should be the activation function of your decision unit? Why did you choose it? (Hint: it should output continuous real values **linearly** from $-\infty$ to $+\infty$, not asymptotically ending in 0 and 1;)
- 3) Train the decision unit on the data from the first 3 days. Report training error. (Clearly show in a table the desired output and the output of your decision unit and the errors)
- 4) Predict the temperature of the 4th date. Calculate the error of your prediction using the data for the 4th date. Report testing error. (Clearly show in a table the desired output and the output of your decision unit and the errors)
- 5) Report the number of iterations, the learning rate, and any data pre-processing steps you chose (hint: consider normalization of input data). Clearly explain why you selected these values.
- 6) Based on the learning trend of the decision unit, would you be able to predict if the decision unit would improve its prediction if more iterations were allowed?
- 7) Could the error be further reduced using a neural network (opposed to a single decision unit)? If so, discuss how and why these methods would reduce the error.

Compile your answers into a single PDF file and submit along with your code.

Note:

1. Your software must be user friendly. The TA must be able to test it simply by executing the code.
2. Hint: you can consider normalization of input data (denormalization of output data).
3. Project deliverable should be a zip file containing:
 - a. Written report with answers to the questions above in word, pdf, ps, or txt format
 - b. The data and separation lines in format as specified by Project1_data.zip
 - c. The source code.
4. Submit your zip file to Instructor mmanic@vcu.edu and cc TA Darshini (Samantha) Mahendran, mahendrand@vcu.edu. Please use the subject line [CMSC 409] Family name, Project 3.