ECE 493/579: Machine Learning for IoT
Department of Electrical and Computer Engineering
Rutgers University

RUTGERS
School of Engineering
ELECTRICAL AND COMPUTER ENGINEERING

ECE 493/579 Machine Learning for the Internet of Things (Fall 2022)
Homework #3 (225 points)

Rationale and learning expectations: This homework covers some of the basic topics from networking and clustering, which are necessary for understanding of most concepts in machine learning for IoT. All students, regardless of their targeted grade in this course, are expected to have complete mastery of the concepts being tested in this homework.

1. **Networking Principles [30 points]**

   a) What are the 7 layers of the Open Systems Interconnection model (OSI model)? **[5 points]**

   b) How many of the 7 layers are actually implemented? **[5 points]**

   c) If we have a sensor deployment on the moon, what general properties would you need to consider in your design of a networking stack? A one paragraph description is sufficient. **[10 points]**

   d) What about a sensor network with lots of transmission losses? Which layer in the stack would you need to modify or redesign to address these challenges? **[10 points]**

2. **End-to-end Argument [45 points]** Consider the two-hop network shown above the nodes denoted by the numerals $1, 2$ and 3. The probability of packet loss at each hop is p.

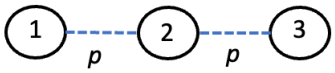

Figura 1: Three nodes with packet loss p at each hop.

   a) Derive the formula for the expected number of transmissions at each hop as a function of the p and the number of packets sent N. **[10 points]**

   b) If we perform hop-by-hop retransmissions and send 100 packets and $p = 0{,}95$. How many packets are sent over the network in total? **[10 points]**

   c) Now, if retransmissions are only performed at the ends (i.e., node 1 only re-transmits packets that do not make it to node 3. How many packets are sent over the network? **[10 points]**

   d) What would p have to be to choose end-to-end retransmissions over hop-by-hop? What general statement can you make about reliability that would justify the end-to-end argument? **[15 points]**

3. **Socket Programming: RESTFul API [50 points]** Download the sensor logger app for android or iPhone.

   a) Write a socket program that shows the data received on the screen. Follow the python using a TCP socket server connection. **[35 points]**

   b) Parse out the JSON object and store it for storage to a CSV or txt file. **[15 points]**

4. **Clustering: Hierarchical Clustering from Scratch [50 points]** Perform different activities with your new measurement tool and data collection code. Perform 3 distinguishable activities. Record the data in separate CSV files and use it to perform clustering analysis. In this problem, use your data of each activity and define a distance function to compare each sample. Using a bottom-up merging approach, design a hierarchical clustering scheme that merges samples, sample-group pairs, and group-group pairs.

   For extra credit (**[15 points]**), display it using anytree.

5. **Clustering: Partitional Clustering w/K-means [50 points]** Write the K-means algorithm from scratch and use it to cluster your data. Generate a 2D plot using PCA or MDS after K-means converges.

   For extra credit (**[10 points]**), generate a Elbow plot of your data. Is the optimal value for k easy to determine? Does it match the number of activities your recorded?