CISC/CMPE452/CISC874/COGS 400

Assignment 3

Implement Competitive Learning Networks

Farhana Zulkernine

10 marks

General Instructions for Code and Submission (for all assignments)

1. You can use any programming language (preferred python, C, C++, Java, or Matlab)
2. Make **one zip file named as Asg3_studentID**.
3. **Upload zip file** to the OnQ site. Multiple uploads are allowed. Only the most recent version will be kept.

Assignment 3 : The Data Set

- The data set you are going to use for this assignment is a synthetic dataset.
- There are 3 features. Use these features to find 2 clusters in the dataset.
- There are 1000 rows of data in the CSV file.

Assignment 3 : Tasks (10 marks)

Part I: ANN and Kohonen

1. Design a simple two layer competitive learning clustering ANN model to find clusters in the given dataset using unsupervised learning. Each output node should implement a linear output function (sum of weighted inputs) and should have inhibitory connection with all other output nodes to implement a Kohonen layer as discussed in the class. Multiple iterations may be needed at the Kohonen layer to find the winning node. **(3 marks)**
2. Next train the model weights based on Kohonen learning algorithm to find two clusters in the dataset such that the clustering error is minimum. Use appropriate terminating conditions. **(2 marks)**
3. Marks will be deducted for inaccurate coding and algorithm. Code should be easy to understand and commented. You will also lose mark if code cannot be executed.

Assignment 3 : Tasks (cont...)

Part 2: K-means

1. Use the same network but implement the K-means algorithm this time (replace the learning algorithm in the previous part) as discussed in the class. **(2 marks)**

Tip: If you have multiple winning nodes for a data point, you can ignore updating the weight vector for that data point in that iteration – state your design choice in the text file.

Part 3: Report

- You must submit the following in the text file for each part of the assignment.
  - Explain ***your assumptions and design choices in a text file*** for both parts 1 and 2, and analyze your result. **(1 mark)**
  - Weight vectors of the cluster centres, the sum squared error for each cluster centre (and the total for both centres) for all data points in the cluster) for part 1 and 2. **(2 marks)**

Submit

- **Deliverables**
  - The program code with comments, executable code (if applicable). (#1 last slide)
  - A text file with all design choices and performance statistics.
  - A text/csv file with the output – cluster number – for each data point in the given order of data in the dataset.
  - Combine all the above in one zip file and upload that on OnQ.
  - Missing submission will be marked as 0.
  - The penalty for late submission past the due date will be -0.5 mark per day.