CMPE 407
Project Deadline:13.06.2022 23.59

May 10, 2022

RULES

1. Codes without documentation (comments) will not be evaluated.
2. Report and code are both required. In case of missing one, the project will not be evaluated.
3. Plagiarism is strictly prohibited. If it is detected either in codes or reports, involved students will get zero.

SUBMISSION

- One submission from one of group members is enough. But do not forget to write the names of all the group members.
- You are asked to prepare a detailed report about your experiments. Upload PDF version of your report.
- As well as your report, you have to upload your codes. You will find another submission slot dedicated for code submissions in Learn. Codes without report or vice versa will not be evaluated.
- Late submissions will not be accepted. Submission system will be closed after the deadline. Submissions via e-mail will not be accepted.

REPORT

There is no page or time restriction for reports. However, you are asked to convince that you fully understand the subject. Therefore, the report without sufficient descriptions can be penalized. Recommended page length is 5 without cover and references pages.
Provide at least the below sections in your report. You can add more sections if you need.

- Introduction
  What was your aim for this project? What about your dataset, explain it. Which method did you use ?

- Dataset
  Explain your dataset in details. Use plots, statistics or any kind of visualization technique. Be sure that you analyzed the dataset comprehensively.

- Preprocessing
  Explain what you did with your dataset samples to prepare them to feed models (label encoding, feature selection, feature extraction etc.).

- Experiments
  Use at least 3 different algorithm (k-means, Linear Regression, Decision Tree, or sth you want to look into) for your selected method. Be careful that these algorithms has to be suitable which method you selected. For example, you cannot use k-means algorithm for regression. Explain the algorithms brieffly, and justify why you select this algorithms. Provide best hyper-parameters for each algorithm with respect to your experiment results. Explain which metrics that you used in your model evaluations. Be sure that you made cross validation for your experiments.

- Results
  Create a table that includes results of all models. Discuss why the best algorithm provided better results than the others; and worst algorithm got worst scores compared to others. Analyse the errors on the best model in order to enhance the further studies.

- Conclusion
  Summarize the report. Indicate the best model for your aim with respect to your experiments. Provide what else that can be done in order to improve results.

QUESTIONS

1. You can find two different sets of chest xray datasets. You need to combine ChestX-ray8 and Chest-Xray8 (COVID-19) sets and try to predict these xray images into correct classes.

2. Make experiments, and report your experiments.

Note: Ensure that your codes are fully documented, using comments. Codes without report or vice versa will not be evaluated. For details about datasets: https://paperswithcode.com/dataset/chestx-ray8 https://paperswithcode.com/dataset/chest-xray8-covid-19 https://github.com/sharifhasani/coviddetection