Description of final project of ESE417 SP2023

1. ***This project is due in 05/05/2023 (Friday) before 5:00PM.***

2. Students in each group should design, implement and test classification algorithms to achieve the best classification performance on the given data set.

3. Every group should work on this project independently and submit a project report online through Canvas before the due date. The report should include at least the following sections:
   Section 1: Introduction (provide a brief description of the project including background on machine learning, the data set description, the goals to achieve, a summary of the methods used and what you have achieved).
   Section 2: Methods (provide a detailed description of the methods used and how they are implemented).
   Section 3*:* Results and Analysis (describe the performance indices used, the performance evaluation methods and the classification results in the forms of charts, graphs and tables)
   Section 4: Conclusions (present the conclusions based on the results)
   Section 5: Appendix (include your Python code and running results here)

   Group members should work together in this project. Please specify the contribution of each member to the project at the end of Section 4.

4. At least two different classification methods covered in this course should be used. One of the following three methods must be used: **Support Vector Machine** method, **Artificial Neural Networks** method and **Random Forest** method. The performance of the methods used should be compared and results should be presented in the project report. On top of the basic version of the methods, if you implement any additional measures (such as data cleaning, feature normalization, hyperparameter optimization, etc.) to improve the performance of the classification methods, please provide descriptions of those measures in your project report.

5. The programming language should be **Python**. The machine learning package that can be used is **sklearn**. No other machine learning packages are allowed. But other supporting packages in Python such as NumPy, SciPy, Pandas and matplotlib can be used.

6. Submit runnable Python code of the project with your report. The instructor may run your program to verify the results.

7. The final project will be graded based on the submitted report (the presentation quality of the report matters!).

8. The number of pages of the report should be limited to 7 pages excluding the Appendix section. The report must be in pdf format.

9. The data set is the **red wine quality** data set from UCI Machine Learning Repository: https://archive.ics.uci.edu/ml/datasets/wine+quality. Please download winequality-red.csv from the data folder of the site. The data set and description are also available on Canvas.