COMPSYS 306 Artificial Intelligence and Machine Learning

Project phase 1 (20%), due 26th Sept 5pm

Kitty Li, Xiaozhou (Jason) Ye, and Kevin I-Kai Wang

The project is designed for the purpose of engaging you in the process of designing a machine learning application with a real-world dataset. The first part of the project (i.e. phase 1) is an individual assessment that requires you to design machine learning models for a real-world traffic sign recognition application using popular neural network (multi-layer perceptron, MLP) and support vector machine (SVM) models that you have learned during the lectures and the laboratories (lab 3 and 4). You will use the following public dataset to train, validate, and evaluate your models.
Kaggle dataset link: https://www.kaggle.com/datasets/flo2607/traffic-signs-classification

The first phase of the project is an individual assessment. In this part, you will follow the steps that you have been introduced during the lecture and lab exercises to import the data, design models, do the necessary pre-processing and parameter tuning as you see fit.

With the designed models, you will then need to conduct performance evaluation (as introduced in the lecture, Topic 8) to compare the performances between the two models (i.e. MLP and SVM).

The deliverables of this part of the project include:
1. A report (16%) following the template on Canvas. The report should include:
   - Dataset and application description.
   - Data pre-processing (e.g. standardization/normalization, feature extraction, data split) and why.
   - Model designs and configurations.
   - Experimental setup (e.g. how do you prepare the training, validation, and testing data set).
   - The comparative analysis (e.g. accuracy, precision, recall, F1-score) of the model performance. The performance comparison results should be presented in a **tabular form**.
2. The code (2%) should be submitted as Python (.py) files. **Your code should be well commented** so that the markers can easily follow and understand your code.
3. The constructed model and the testing dataset should be saved and submitted as .joblib files, (2%). You also need to describe how you have prepared the test data clearly in your report.

Note: The report, code, model, and test dataset should be packaged into a .zip file with a README instruction explaining the included files. The zip file should be named as project_username.zip (e.g. project_khan029.zip) for Canvas submission.

Late submission: Please refer to the "COMPSYS 306 2023 Course Info.pdf" in the course Modules.