**University of Canberra  
Faculty of Science and Technology**

**Pattern Recognition and Machine Learning**

**Assignment 1 Part B**

**Group No.:**

**Project Team Members (maximum of 2):**

In Assignment 1-Part B of the assessment you (in your team) are asked to study the MNIST & Fashion-MNIST pattern recognition modeling using logistic regression and answer the given questions.

The Logistic Regression model of learning and its application to the MNIST dataset has been covered in Tutorial Week 3 – the code, data and documentation is available under this section on assignment information.

The exercise is to apply the given code to a new dataset for the Fashion-MNIST problem - <https://www.kaggle.com/datasets/zalando-research/fashionmnist>).

You are asked to go through the process used in applying Logistic Regression to MNIST and explaining your understanding and the outcomes for the given Fashion-MNIST dataset. You are expected to generate a similar report to the one generated for the MNIST dataset along the lines of the provided Jupyter Notebook report (as ipynb file) for MNIST. (You may use Spyder to run the code also but convert to ipynb for submission.)

Your report should cover the following.

1. Introduce the problem you wish to solve. Explain what questions would you want to ask from the dataset?
2. Describe the dataset. What do the rows and columns mean and are there any special characteristics in the data to use in modelling the learning?
3. Explain why Logistic Regression would suit as the learning model to answer the questions.

*Python constructs – explain each step below*

1. Retrieving data in the program
2. Exploring data – explain any special adaptation and characteristics
3. Show image & data
4. Show corresponding matrix
5. Describe & understand the data
6. Build a Logistic Regression model
7. Load packages
8. Select target variable
9. Prepare data
10. Split data into training and validation set
11. Select a linear regression classifier
12. Fit the data
13. Prediction of unseen data
14. Analysis of results – classification report
15. Create confusion matrix
16. Visualise correct prediction
17. Visualise misclassified images
18. Visualise corrected images
19. Explain the concept of regularization in logistic regression and how does it prevent overfitting? Explain how can it be used in Fashion-MNIST modelling?
20. Save trained model and explain how you would use it for new unseen data.

**Submission:**

For Part B you need to submit a report answering the above questions in 2-3 pages as a pdf file or ipynb file but compiled and zipped as a single file and uploaded through canvas by the submission deadline.