Problem’s overview

- Provide a conceptual figure to explain the problem in hand
  - The figure should show an end-to-end dataflow, and provide insights on the problem
  - You can write a few sentences to further explain the problem, if needed
  - Copying and pasting the figure from online resources is not allowed

Dataset’s overview (EDA)

- Briefly study the dataset and provide your insights and any relevant information. That includes, but not limited to:
  - Number of instances per class
  - Number of null values per feature
  - Number of possible outliers
  - Basic statistical analysis for every feature (mean, std, min, max)

Flowchart

- Provide an end-to-end flowchart, where you show every step in the process of the project’s implementation.
  - The flowchart should be clear and the font’s size is visible

Q1) Obtain a baseline performance

- Build an MLP classifier using the provided dataset and the following parameters:
  - 1 hidden layers with 10 neurons
  - Relu activation function
  - Use AdamW optimizer and set the learning rate as 0.001
  - # of epochs=500
  - Batch size =1
- Plot the training and testing losses vs. the number of epochs
- Complete the table below, and provide the confusion matrix based on the average test accuracy
- Provide 2D TSNE plots, one for the training set, one for the test set, and one for the validation

Max training acc	Max test acc	Min training acc	Min test acc	Avg training acc	Avg test acc

Q2) Compre dimensionality reduction to feature selection

Q2.1) Dimensionality reduction

- Find the best value for n_components based on the test accuracy of the MLP classifier, using Principal Component Analysis,
  - PCA(n_components=n, random_state=0)
- Plot the (number of component-test accuracy) graph with the baseline performance.
- Apply TSNE(n_components=2, random_state=0) to visualise the training and test datasets after DR

Q2.2) Feature selection

- Find the best number of features based on the MLP test accuracy, using the feature selection methods:
  - ANOVA
  - Mutual Information
- Choose the method that achieves the best test accuracy results (either the ANOVA or mutual information), and provide 2D TSNE plots, one for the training set and one for the test set.

Update your dataset, **to be used in the next steps**, based on the technique that provides you with highest validation accuracy (either dimensionality reduction or feature selection), **and provide the confusion matrix**.

Q3) Vary the MLP parameters [1/5]

Q3.1) Batch size

- Use the hyper-parameters of the baseline model, unless otherwise specified.
- Examine the impact of batch size on the MLP classifier's **average test accuracy**, which should be taken **over 5 different runs**.

Highlight the combination that achieves the highest average accuracy for **both training and test accuracies at the same time**. (if you cannot find a combination that achieves both at the same time, then highlight the combination that achieves the highest average **test accuracy**).

Batch size= 32						**Batch size= 64**						**Batch size= 128**					
Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc

Q3) Vary the MLP parameters [2/5]

Q3.2) Hidden layers vs. neurons/layer

- Examine the training and test accuracies under different numbers of hidden layers and nodes/layer. Use the best batch-size from Q3.1 that achieves the **highest average validation accuracy.**
- Plot the number of neurons vs the avg test accuracy. (You will plot only one figure that shows 4 different lines, each one for a different number of hidden layers)

Highlight the combination that achieves the highest average accuracy for **both training and test accuracies at the same time**. (if you cannot find a combination that achieves both at the same time, then highlight the combination that achieves the highest average **test accuracy**)

	1 hidden layer						2 hidden layer						4 hidden layer						8 hidden layer					
# of neurons for each layer	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test a
10																								
20																								
30																								
40																								

Q3) Vary the MLP parameters [3/5]

Q3.3) Learning rate and different optimizers

- Study the impact of different optimizers and learning rate on the accuracy results. Use the best hyper-parameters from each of the previous questions that achieves the **highest average validation accuracy.**
- Highlight the combination that achieves the highest average test accuracy for **both training and test accuracies**. (if you cannot find a combination that achieves both at the same time, then highlight the combination that achieves the highest average **test accuracy**)

	Learning rate=0.1						**Learning rate=0.01**						**Learning rate=0.001**					
optimizer	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc
AdamW																		
SGD (momentum =0.1)																		
SGD (momentum =0.5)																		
SGD (momentum =0.9)																		
Rprop																		

Q3) Vary the MLP parameters [4/5]

Q3.4) Activation functions

- Examine the impact of different activation functions on the MLP classifier's test accuracy. Use the best obtained combination from Q3.3, and try them with the following activation functions:

Relu						Leaky Relu						Sigmoid						Tanh					
Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc	Max train acc	Max test acc	Min train acc	Min test acc	Avg train acc	Avg test acc

- Highlight the activation function that provides you with the highest average test accuracy.
- Provide the confusion matrix for the highest average accuracy.

Q3) Vary the MLP parameters [5/5]

Q3.5) Bonus question

- Try to optimise the parameter tuning process, and try anything you can to improve the average test accuracy even further. (You are allowed to tune any hyper-parameters, even the previous ones. You have to use an **MLP model**)

Conclusion

- Include a few sentences, in a qualitative manner, to summarize the results and the outcomes of the project.