Data Science Labs

Lab 1
Baseline Models

GOAL

To train models from raw data, to serve as a baseline comparison.
To assess the model's performance.

WORK TO BE DONE

To clean the data by:

Dropping any missing values, eliminating:

All variables totally empty.

All records having some missing values.

To discard all non-numeric data, eliminating all non-numeric variables.

To train models with the following techniques on the cleaned data:

Naïve Bayes

Logistic regression

KNN

Decision trees

Multi-layer perceptron.

REPORT TO DELIVER

PDF file only with charts and tables. No analysis or justification is needed at this point.
Description of the best model trained for each technique – hyperparameters found.
Performance results of each one of the best models:

Accuracy; Precision; Recall; Others, if needed.

Suggested charts per technique:

Hyperparameters study;

Best model's performance.

Good work!!!