

## Project Guidelines

**Deadline: 12 May 2024, 11:59 PM**

### For 10 marks

- 1) Apply data preprocessing (such as Handling Null Values, Standardization (standard scaler, minmax scaler, etc), Handling Categorical Variables, One-Hot Encoding, and Multicollinearity)
- 2) Data Visualization (At least 5 types of graphs, such as scatter plots, bar plots, histograms, etc) try to explain the trends in the graph.
- 3) Check for class imbalance, if present apply oversampling/undersampling techniques. <https://imbalanced-learn.org/stable/>
- 4) If it's a simple Classification/Regression problem. Apply at least 8 different models.
- 5) If you are working with images. Apply at least 3 models (e.g. MLP, CNN, and transfer learning)
- 6) Use cross-validation and grid search CV module to choose hyperparameters.
- 7) Use different hyperparameters to achieve optimum results.
- 8) Report precision, recall, sensitivity, specificity, etc in addition to accuracy.
- 9) Your project grade will depend on multiple factors. For example how complex your problem was ((obviously a binary classification problem is easier than multiclass classification, dealing when images is harder than working with quantitative data), how much work had already been done on that data set, how much work you have done for the project and your team size (obviously more work is expected from a team of four students than a team of three students))

For 5 marks (bonus points (if you are working on a quantitative data set))

- 1) Use feature extraction techniques such as filter-based methods, wrapper-based methods etc (<https://www.analyticsvidhya.com/blog/2020/10/feature-selection-techniques-in-machine-learning/>)
- 2) Use explainable AI models, LIME and Shap to explain how different features impact final results (<https://www.datacamp.com/tutorial/explainable-ai-understanding-and-trusting-machine-learning-models>)

For 5 marks (bonus points (if you are working on an image data set))

- 1) Explore and apply advanced CV algorithms such as YOLO <https://www.v7labs.com/blog/yolo-object-detection>
- 2) You can also try to experiment with RNNs, LSTMs, and Transformers

**Deliverables:**

- 1) The final deliverable will be your well-documented and commented code file.
- 2) A well-written project report. The project report template will be posted on GCR.