PROJECT GUIDE

SECTION A

WHAT TO DO

- **Understand the problem** - problem statement, aim, and objectives of the project. This should be brief, explain the problem, and how your project intends to solve it.
- **Data collection** - collect data. This can be done by getting data from a data repository, a downloadable link, or by scraping data from the web. The main focus should be on the data source.
- **Data processing** - process the data collected. This may involve checking and removing outliers, data type conversion, checking and removing null values, checking and removing redundant data/entries, etc. Note that different projects require different data processing steps.
- **Exploratory data analysis** - perform initial investigations on data so as to discover patterns, spot anomalies, test hypotheses, and check assumptions with the help of summary statistics and graphical representations. The main focus should be on data visualization. Let's see how you can tell a story using data.
- **Feature engineering** - use domain knowledge to extract features (characteristics, properties, attributes) from raw data, transform raw data into features that better represent the underlying problem, etc. Note that different projects require different feature engineering processes.

PROJECT GUIDE

HAMOYE

WHAT TO DO

- **Model training and evaluation:** provide a machine learning algorithm (that is, the learning algorithm) with training data to learn from. Evaluate the model performance using metrics such as confusion matrix, precision, accuracy, F1 score, MAE, accuracy, specificity, etc. It is advised to try out different models starting with simpler models before moving on to complex models.
- **Model deployment**

TIPS ON IMPROVING YOUR PROJECT

This is a bonus section on how to improve your code.

- **Modularize your code for reusability:** create modules, classes, and functions to avoid redundant code, and allow for code reuse. The Advantages of Modular Software and Programming – Gwentech Embedded
- **Test-driven development:** this approach, used to test development teaches you how to write good code because you have to fix all of your bad code. Here are more advantages, It's Important To Test Your Code - DEV Community
- **Exception handling:** exceptions provide the means to separate the details of what to do when something out of the ordinary happens from the main logic of a program. In traditional programming, error detection, reporting, and handling often lead to confusing spaghetti codes. Check out this link to learn more about exception handling in python Python – Exceptions Handling

SECTION B

WHAT WE WANT TO SEE

This section explains the things that are expected to be included in the presentation slide.

- **About Team:** one slide listing out the team members and their roles.
- **Introduction/Problem statement:** this should be short and straight to the point (1-2 slides stating the problem, and how this project solves it. Also, include any existing solution, their limitations, and how this project solves them).
- **Data collection process:** this should also be short and to the point (one slide).
- **Exploratory data analysis:** show graphical illustrations, charts, plots etc. The images should reinforce and complement your message.
- **Feature engineering:** show final features used in building the final model(short and to the point - One slide).
- **Model training and evaluation:** show all machine learning models, all experiments performed and a comparison between all models using the suitable evaluation metrics. This is a very important part of your presentation.

SECTION B

WHAT WE WANT TO SEE

- **Model deployment:** show the model deployment pipeline, a demo of the deployed model i.e showing the model makes live predictions. This is also an important part of your presentation
- **Summary and conclusion** (short and brief one slide)

PRESENTATION TIPS

- Choose an appropriate font and size.
- Images such as graphical illustrations or charts should be of good quality.
- Limit the number of slides (the lesser, the better).
- Do not read from your slides or speak to them. Use points to get the audience's attention.
- Manage time: it is advisable to always practice your presentation using a timer. Make your points clear and don't waste time on not so important points.