Project Guidelines

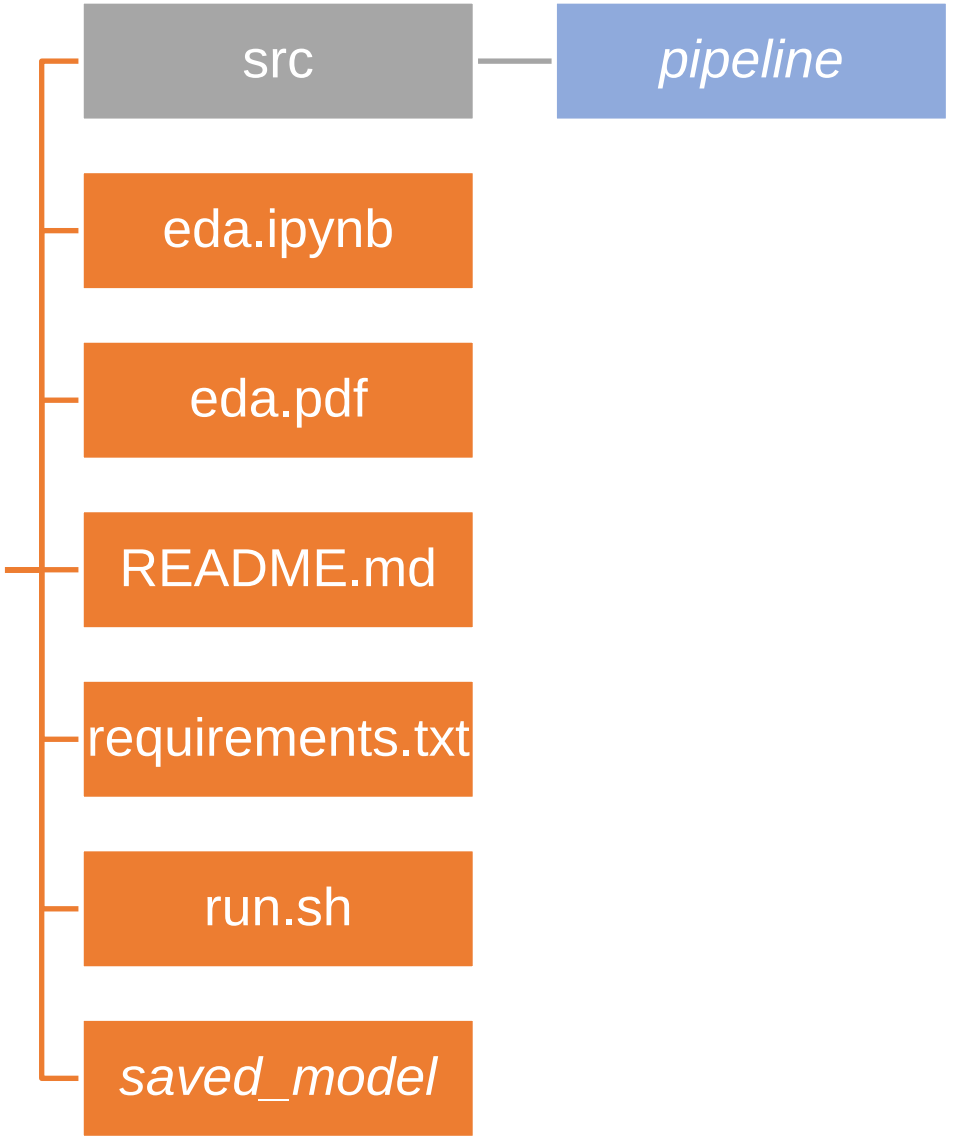
Project – 40% of the Learning Unit

PROJECT	Project presentation (team)	Week 7
	Project work submission (team)	Week 7

Objective

- Perform EDA and document insights
- Document choice of machine learning models
- Build ML pipeline
- Practise presentation skills

Project Requirement



Project Requirement (Bonus)

Container – jupyter lab server

- Accessible via http://localhost:888x
- Able to run individual cells of eda.ipynb
- Optional: pipeline execution by run.sh

Container – executes run.sh

- Execute pipeline and outputs results

Project Requirement

- At least n tasks
  (n is the number of members in the team, min 3 pax in the team)
  - EDA (Data Cleaning / Data Analysis / Feature Engineering)
  - Model Training
  - Model Inference & Reporting
- Each member to present and do code walk through on key parts of their contributions
- Version Control
  - Code changes should be consistently tracked and uploaded to a version control repository with clear commit messages.

Group Formation

- 4 students per group
- Submit students name list (with a team name) on BrightSpace under Assignment -> Group Name List Submission
- One submission per group
- Group information submit by end of week 4 (16 May 2025)

Deliverables – Group Presentation (Week 7)

- Each team will present your project work during class in a 20-minute presentation
  - 5-minute presentation, 5-minute code walk through, 10-minute Q&A
  - Each team member will be asked to briefly walk through the code he/she manages.
  - Indicate contributions by each team member.
- Presentation should include (but not limited to) the following:
  - Highlight key parts of your EDA (null, erroneous value handling, insights, feature engineering)
  - Machine Learning Pipeline (e.g ingestion, hyperparameter tuning, model selection)
  - Comment on Model's Accuracy
  - Further Enhancement or improvements (if any)
  - Others

Deliverables – Group Reflection & Submission (Week 7)

- ~~Reflection:~~ (No reflection)
  - ~~Each team member writes their reflection in the same document. Indicate your name on the section that you have written.~~
  - ~~Reflection to focus on knowledge, skills and attitudes that you have personally changed, acquired or action plans related to KSA.~~
  - ~~Not more than 400 words per person.~~
- Please put all work files and reflection document in ONE Zip folder and submit to BrightSpace, following naming convention below.

  Zip Folder Name: EGT309_T1/2_<project team name>.zip.
- All codes files, presentation slides and reflection document should be compressed into the folder for submission

Submission Deadline

6 June 2025, 2359

- Submission time will be captured based on BrightSpace record
- Refer to Extension and Late Submission Policy 2024
- Zero mark will be given for late submission after 5 calendar days from due date.

Sharing yours or your team's work with other learners or teams is considered an act of plagiarism.
ALL parties involved, including those who share their work, will face discipline under NYP's Academic Integrity Policy.

Rubrics [Group – Technical Skills and Knowledge]

	Excellent	Good	Satisfactory	Need Improvement
	8 – 10 points	5 - 7 points	3 - 4 points	0 - 2 points
Overall Analysis (10 points)	• In-depth Data Analysis and insights • In-depth analysis of models' performance and possible follow-up actions	• Some Data Analysis and insights • Some analysis of models' performance and possible follow-up actions	• Limited data analysis and insights • Some incorrect analysis of models' performance and possible follow-up actions	• Minimal or incorrect data analysis and insights • Incorrect or no analysis of models' performance or possible follow-up actions
Overall Design (10 points)	• Sound justification of choice of models • Pipeline handles all aspects from data ingestion to model evaluation	• Some valid justification of choice of models • Pipeline handles most aspects from data ingestion to model evaluation	• Some invalid justification of choice of models • Pipeline handles some aspects from data ingestion to model evaluation	• Little or invalid justification of choice of models • Pipeline handles minimal aspects from data ingestion to model evaluation
Use of tools and libraries (10 points)	Seamlessly applies docker and kedro in building and deploying the machine learning pipeline	Some use of both docker and kedro in building and deploying the machine learning pipeline	Some use of either docker or kedro but not both in building and deploying the machine learning pipeline	Did not apply docker nor kedro in building and deploying machine learning pipeline
Version Control (10 points)	Code changes are consistently tracked and uploaded to a version control repository, with detailed commit messages and consistent version history.	Code changes are regularly tracked and uploaded to a version control repository with unclear commit messages.	Code changes tracked with inconsistent commits and unclear commit messages	Little to no evidence of version control. Repository may exist but is not utilized effectively to track changes

Rubrics [Group – Non-Technical]

	Excellent	Good	Satisfactory	Need Improvement
	8 – 10 points	5 - 7 points	3 - 4 points	0 - 2 points
Presentation Content Organisation (10 points)	• The presentation is organised in a logical sequence that flows naturally. • Presentation load is well distributed.	• The presentation is organised in a sequence that can be followed. • Presentation load is evenly distributed.	• The presentation is organised. • Imbalanced presentation (some members presented much more than others).	• The presentation is difficult to follow • Presentation done by less than 50% of the members
Work Independence (10 points)	Consistently works independently, requiring minimal guidance, and proactively solves problems.	Works independently on most tasks but occasionally seeks clarification or assistance.	Requires frequent guidance and supervision to complete tasks. Struggles to work independently.	Requires consistent guidance and supervision to continue with the project

Rubrics [Individual – Technical Skills and Knowledge]

	Excellent	Good	Satisfactory	Need Improvement
	12 – 15 points	8 - 11 points	4 - 7 points	0-3 points
Code Walk-through (15 points)	Delivers a clear, thorough, and well-structured code walk-through, explaining logic, decisions, and functionality in detail.	Provides a solid code walk-through that covers key components and logic, though some details may be less clear.	Presents a basic walk-through with limited explanation of the code and its logic.	Code walk-through is incomprehensible and sometimes illogical
Code Quality (15 points)	• Code is always structured to minimise repetition. • Code is always organized in a readable format • Intent of all functions and classes are stated clearly in the code • Extensive use of configuration files	• Code is mostly structured to minimise repetition. • Code is mostly organized in a readable format • Intent of most functions and classes are stated clearly in the code • Some use of configuration files	• Code is rarely structured to minimise repetition. • Code is rarely organized in a readable format • Intent of some functions and classes are stated clearly in the code • Minimal use of configuration files	• Code is not structured to minimise repetition. • Code is unorganized • Intent of functions are not stated • No use of configuration files

Rubrics [Individual – Non-Techincal]

	Excellent	Good	Satisfactory	Need Improvement
	8-10 points	5-7 points	3-4 points	0-2 points
Handling Q&A (10 points)	Answers questions clearly and confidently. Handles challenging or unexpected questions with ease, providing thorough and insightful responses.	Responds to most questions effectively. May need some time to think or seek clarification but provides reasonable answers overall.	Struggles to answer questions fully or clearly, often providing vague or incomplete responses. Has difficulty handling challenging questions.	Unable to answer most of the questions. Answers may not be correct.