

Milestone 3 & 4

M3: Writing the Methodology Section

M4: Descriptive Progress Report

By the end of this lesson, students should be able to:

- Write a clear and concise Methodology section.
- Structure and organize experimental procedures with clarity and precision.
- Track and evaluate their progress quantitatively using rubrics and progress bars.
- Describe how data is collected, processed, and analyzed in ML/DS experiments.
- Demonstrate reproducibility by clearly documenting the experimentation stage.

Importance of the Methodology Section

- The Methodology is the backbone of any scientific research.
- This section answers the question: How did you conduct your research?
- It must be clear enough that another researcher could replicate the experiment based on your descriptions.

Key Components: Research Design

- **Definition:** A high-level blueprint of how you approach your research.
- **Types of Research Design:**
 - **Experimental Design:** Typically used in ML/Data Science projects, where models are trained and tested under various conditions.
 - **Comparative Design:** Comparing different algorithms (e.g., SVM, Random Forest, Neural Networks) or different datasets.
 - **Exploratory Design:** May involve an exploratory analysis of new data using unsupervised techniques like clustering.

Key Components: Experiment Procedure (Sample)

- **Step 1: Data Collection:**

- Explain how you collected or accessed the dataset.
- **Example:** "We scraped 10,000 news articles using a Python-based web scraper and filtered them based on their publication date."
- **Dataset Characteristics:** Size, structure, source (e.g., Kaggle, UCI Machine Learning Repository).

- **Step 2: Data Preprocessing:**

- Describe how the data was cleaned and prepared for modeling.
- **Example:** Removing missing values, normalizing data, feature selection, data augmentation for images, etc.
- **Tip:** Use tables/diagrams to represent the stages of data transformation.

- **Step 3: Feature Engineering:**

- Define how features were selected or created to improve model performance.
- **Example:** "We applied TF-IDF vectorization to convert text data into numerical format for machine learning models."

- **Step 4: Model Selection:**

- Describe the algorithms or techniques used (e.g., decision trees, neural networks, support vector machines).
- **Provide justification** for model choice.
- **Example:** "A CNN was chosen for image classification due to its superior ability to capture spatial hierarchies in images."

- **Step 5: Training the Model:**

- Explain how you split the dataset into training, validation, and test sets.
- Mention the tools used for model development (e.g., Python, TensorFlow, Scikit-learn).
- **Include details** like hyperparameters (e.g., learning rate, batch size, number of epochs) and optimization techniques (e.g., gradient descent).

- **Step 6: Model Evaluation:**

- Outline the metrics used to evaluate model performance (e.g., accuracy, precision, recall, F1-score, ROC-AUC curve).
- Justify why you chose these metrics.
- **Example:** "We selected F1-score to balance the trade-off between precision and recall for our imbalanced dataset."

- **Step 7: Experimentation with Variants:**

- Discuss any variations or tweaks made to the models (e.g., changing hyperparameters, adding dropout layers).
- Mention if you experimented with ensemble methods or different algorithms to improve performance.

- **Step 8: Testing and Final Evaluation:**

- Explain how the final model was tested on a separate test set and any validation techniques (e.g., cross-validation) applied.

- **Step 9: Technical Output Model Integration**

- Discuss what Technical output (Desktop, Web, Mobile app, etc) design, processes, procedures to show your model is working

Guidelines and Tips:

- **Be Clear and Concise:**

- Avoid unnecessary jargon or overly technical language. Explain terms where needed.
- Keep sentences focused on one aspect of the methodology to avoid confusion.

- **Use Diagrams and Flowcharts:**

- Visual aids like **process diagrams** or **flowcharts** help clarify the steps and make the methodology easier to follow.
- **Example:** A diagram showing the data pipeline from data collection to final model evaluation.
- Example: Dissect conceptual framework parts into small understandable portion of your methods.

Guidelines and Tips:

- **Provide Justifications:**

- Always explain **why** a particular method, tool, algorithm, or process was used.
- **Example:** "We chose Random Forest for its ability to handle overfitting by averaging multiple decision trees."

- **Include Versions of Tools and Libraries:**

- For reproducibility, include the versions of software, libraries, and tools used (e.g., TensorFlow 2.4, Scikit-learn 0.24).

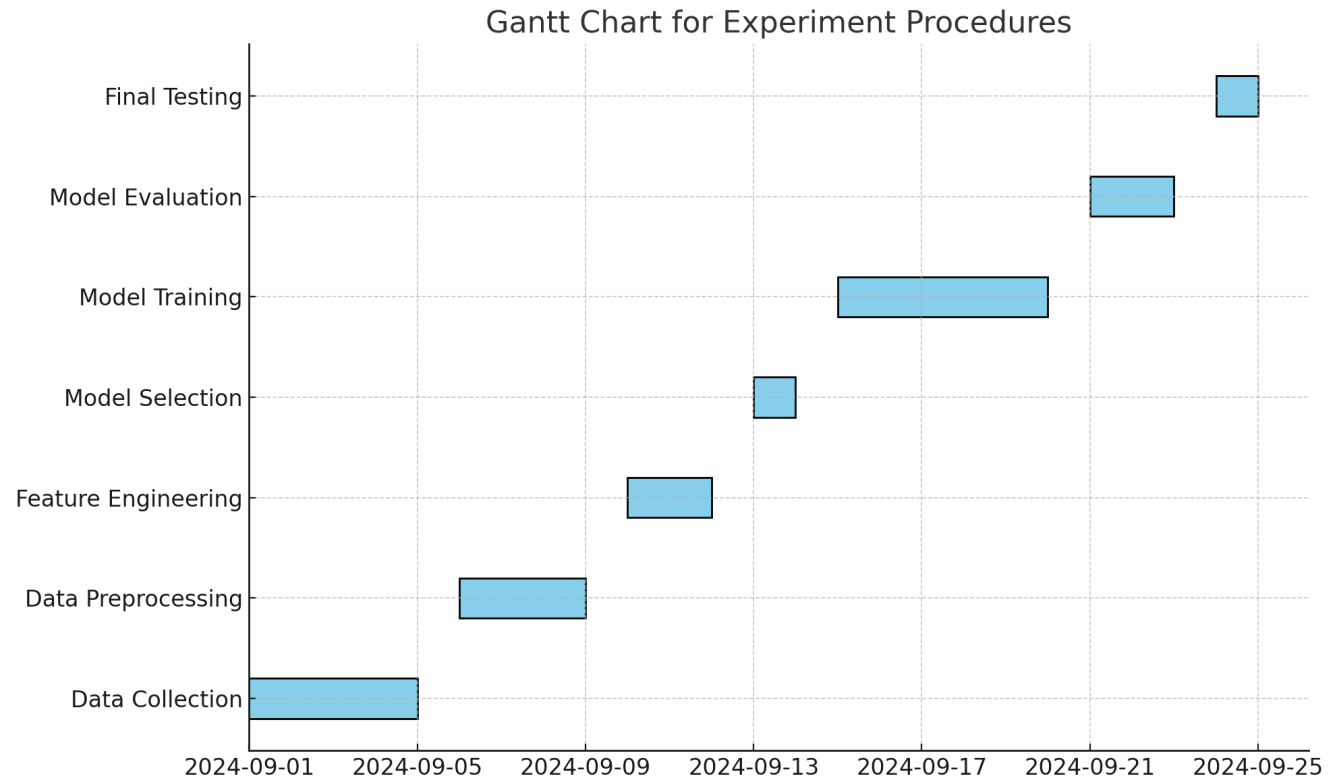
Progress Tracking Using Progress Rubrics

- Create a grading rubric for each part of the Methodology section.
- Do self-evaluate or peer-review based on this rubric to monitor their writing progress.

	1	2	3	4	5
Clarity of Experimental Procedure					
Justification of Tools/Methods					
Consistency with Hypothesis					
Use of Visual Aids					

Progress Tracking Using Gantt Chart

- Gantt chart visualizing the progress for a thesis experiment's procedures. Each task, is displayed with its corresponding start and end dates.
- This type of chart helps track the timeline for different stages in a project, providing a clear overview of how time is allocated for each step in the experiment.
- By updating regularly, can monitor progress effectively and ensure they meet their project deadlines.



Gantt Chart in Google Sheet

- Follow the link below to create the Google Sheet
- <https://www.youtube.com/watch?v=8eKk0M2zGIk>
- Start August 12, 2024 (Week 1) until November 15, 2024 (Week 14)

Procedures for Documenting Experiment Outcomes

- **Record Model Performance:**

- After each training session, record metrics like accuracy, loss, F1-score in a table.
- Track how changing hyperparameters affects performance.

- **Document Failures and Solutions:**

- It's important to document what didn't work as well as what did.
- **Example:** "Training with a high learning rate led to unstable results, so we lowered the learning rate to 0.001, which stabilized training."

- **Use Version Control:**

- Encourage students to use tools like GitHub to track code changes and version experiments.

Milestone 3

- **1:** Create Google docs for Methodology Write ups
- Include Frameworks: Conceptual and Operational; and Methodology
- Methodology parts
 - Sample: Research Design, Data Collection, Data Analysis, Tools and Technologies, Ethical, Limitations
 - Experiment Procedure (equated with Operational Framework)
 - Designs and plans, procedures, processes, results, and outcomes
- **2:** Edit Gantt chart in Google sheets
- **3:** add Progress Rubric in Google docs and do self-assessment using the rubric
- Due: August 29, 2025, Friday, share the Google Docs link in submission link
- Present Methodology write ups highlights

Milestone 3 presentation schedule

27-Aug-25		26-Aug-25	26-Aug-25
Baltazar et al.		Mendrico et al.	Guevarra et al.
Al Adwani et al.		De Los Santos et al.	Lumanog et al.
Ofamin et al.			Co et al.
Lo et al.			Jala et al.
Dicipulo et al.		28-Aug-25	
		Canacan et al.	28-Aug-25
1-Sep-25		De Jesus et al.	Cameguing et al.
Tabudlong et al.			Juarbal et al.
Macabangon et al.			Dable et al.
Garcia et al.			Agreda et al.
Casiple et al			
Digamon et al			2-Sep-25
			Agawin et al.
3-Sep-25			Ipdan et al.
Advincula et al.			Lampa et al.
Cruz et al			
Aguado et al.			
Gempesaw et al			
Aliño et al.			

Milestone 4

- Progress Report Presentation
- Minimum requirements finish
 - CS/DS: all data gathering, data preprocessing, initial Model development
 - Suggestion use the experimentation dataset or your final dataset as for your data for model development.
 - IS/IT: all system requirements, data model, process model, UI/UX Designs
- Updated Google Docs and Sheets, deadline: Presentation schedule
- Slides (pdf): updated progress from experimentation, Progress highlights, updated Gantt Chart,
- Descriptive Progress Report (pdf)

Milestone 4 presentation schedule

8-Sep-25		2-Sep-25	2-Sep-25
Baltazar et al.		Mendrico et al.	Guevarra et al.
Al Adwani et al.		De Los Santos et al.	Lumanog et al.
Ofamin et al.			Co et al.
Lo et al.			Jala et al.
Dicipulo et al.		4-Sep-25	
		Canacan et al.	4-Sep-25
10-Sep-25		De Jesus et al.	Cameguing et al.
Tabudlong et al.			Juarbal et al.
Macabangon et al.			Dable et al.
Garcia et al.			Agreda et al.
Casiple et al			
Digamon et al			9-Sep-25
			Agawin et al.
15-Sep-25			Ipdan et al.
Advincula et al.			Lampa et al.
Cruz et al			
Aguado et al.			
Gempesaw et al			
Aliño et al.			