PHASE 1. PROBLEM DEFINITION AND RESEARCH QUESTION

1. Working Title

- A clear and concise title (can be tentative at early stages)

2. Research Problem / Motivation

- What is the broad question?
- Why is it important? (real-world relevance, academic gap)

3. Research Objectives / Hypotheses

- What are the specific aims?
- Any hypotheses or expected patterns?

Checklist:

- Identified 2-3 potential research domains
- Written a brief explanation of why each domain interests you
- Considered your background knowledge in each area
- Selected one primary domain to focus on
- Research question is specific and measurable
- Question addresses a real problem with practical implications
- Question is neither too broad nor too narrow for MSc level
- Question can realistically be answered with available resources
- Question contributes something new to the field

PHASE 2. DATA STRATEGY AND ACCESS

4. Data Source

- What data do they intend to use?
    - Public dataset?
    - Self-collected?
    - Institutional?
- Availability and accessibility
- Ethics or permissions required

5. Variables

- Define outcome variable(s)

- Identify key predictor variables
- Any additional covariates or control variables?

Checklist:

- Listed all required variables with clear definitions
- Identified target variable(s) clearly
- Estimated minimum sample size needed
- Considered data quality requirements
- Thought about temporal aspects of data
- Researched at least 3 potential data sources
- Verified data availability and access requirements
- Checked data licensing and usage rights
- Assessed data quality and completeness
- Identified backup data sources
- Considered ethical implications of data use
- Secured access to primary data source
- Obtained necessary permissions/approvals
- Downloaded or collected initial data sample
- Verified data format and structure
- Established data backup procedures

PHASE 3. LITERATURE REVIEW

6. Literature Review Summary

- 3–5 relevant studies
- Gaps they aim to address
- Methods commonly used in this area

Checklist:

- Found and read 10-15 relevant research papers
- Summarized key findings and methods from each paper
- Identified gaps or limitations in existing work

- Noted common datasets and evaluation metrics used
- Created a reference management system (Zotero, Mendeley, etc.)

PHASE 4. METHODOLOGY AND ANALYSIS

7. Methodology

- Data cleaning/preparation
- Feature engineering
- Exploratory data analysis
- Chosen models/algorithms (e.g., regression, classification, clustering)
- Evaluation metrics

8. Tools

- Programming language(s) and libraries
- Cloud or local tools (e.g., Jupyter, RStudio, GitHub, etc.)

Checklist:

- Chosen appropriate methodology for your research question
- Identified 2-3 methods to compare
- Defined clear evaluation metrics
- Established baseline for comparison
- Considered computational requirements
- Designed appropriate train/validation/test splits
- Selected suitable cross-validation strategy
- Identified key hyperparameters to tune
- Chosen appropriate evaluation metrics
- Created evaluation framework
- Performed exploratory data analysis (EDA)
- Handled missing values appropriately
- Detected and addressed outliers
- Applied feature engineering techniques
- Normalized/standardized features if needed
- Created data preprocessing pipeline

- Documented all preprocessing steps
- Implemented baseline model(s)
- Implemented main methodologies
- Created reproducible code with version control
- Implemented proper error handling
- Added logging and progress tracking
- Created model training pipeline
- Implemented hyperparameter tuning
- Evaluated all models using defined metrics
- Performed statistical significance tests
- Created visualizations of results
- Analyzed feature importance/model interpretability
- Identified best performing approach
- Documented key findings and insights
- Acknowledged limitations

PHASE 5. OUTPUTS

9. Expected Outputs

- Figures, tables, performance metrics, model explanations
- Any deployment or dashboard?
- Thesis writing

Writing Checklist:

- Created detailed outline for each chapter
- Written first draft of all chapters
- Created all figures and tables
- Ensured proper citations throughout
- Checked for logical flow between sections
- Proofread for grammar and clarity
- Verified all references are complete

Timeline / Milestones

- Week-by-week goals (e.g., literature review, data cleaning, modeling, results)