

## Table of Contents

Page 1: Project Overview .....	2
Page 2: Data Preparation .....	3
Page 3: Model Development.....	4
Page 4: Evaluation and Optimization .....	5
Page 5: Deployment and Monitoring.....	6

**1. Define the Problem Statement**

- Clearly articulate the problem you're solving.
- Example: "Build a sentiment analysis tool for customer reviews using Transformers."

**2. Specify the Objective**

- Define measurable outcomes (e.g., accuracy, latency, or user satisfaction).
- Example: "Achieve a sentiment classification accuracy of 90% within a latency of 200ms per inference."

**3. Dataset Selection**

- List datasets being used (public or proprietary).
  - Mention dataset sources, size, and format.
  - Example: "We use the IMDB dataset with 50,000 labeled movie reviews."
-

#### 4. Data Preprocessing

- Describe preprocessing steps:
  - Text cleaning (e.g., removing stop words, lemmatization).
  - Data augmentation (e.g., backtranslation for text).
- Tools: Python libraries like NLTK, Spacy, pandas.

#### 5. Exploratory Data Analysis (EDA)

- Visualize data distribution and patterns.
  - Example tools:
    - **matplotlib** for distributions.
    - **seaborn** for heatmaps.
-

## 6. Model Selection

- Choose a baseline model (e.g., Random Forest for ML).
- Identify advanced models:
  - **Transformers:** BERT, GPT, or T5.
  - **LLMs:** OpenAI, Hugging Face models.
  - **Agents:** LangChain, custom reinforcement learning agents.

## 7. Training and Hyperparameter Tuning

- Key parameters: learning rate, batch size, epochs.
  - Tools: grid search, random search, or automated tools like Optuna.
-

## **8. Model Evaluation**

- Define metrics:
  - ML: Precision, Recall, F1-Score.
  - LLM: BLEU, ROUGE.
  - Agent: Reward scores.
- Use confusion matrices, ROC curves for visualization.

## **9. Optimization**

- Methods:
    - Pruning for overfitting.
    - Quantization for deployment optimization.
-

## **10. Deployment**

- Deploy using frameworks:
  - Flask, FastAPI for APIs.
  - Docker and Kubernetes for scalability.
- Edge devices: Convert models to ONNX/TensorFlow Lite for mobile deployment.

## **11. Monitoring and Iteration**

- Tools: Prometheus for logging, Grafana for visualization.
  - Regularly update the model using new data.
-