

# AI Assignment: Understanding the AI Development Workflow

---

## Part 2: Case Study Application

Scenario: A hospital wants an AI system to predict patient readmission risk within 30 days of discharge.

### Problem Scope

Problem: Develop an AI model that predicts whether a patient is likely to be readmitted within 30 days of discharge from the hospital.

Objectives:

1. Reduce hospital readmission rates to improve patient outcomes.
2. Support clinical decision-making through predictive analytics.
3. Optimize healthcare resource allocation.

Stakeholders:

1. Hospital management and healthcare providers.
2. Patients and their families.

### Data Strategy

Proposed Data Sources:

1. Electronic Health Records (EHRs): patient diagnosis, procedures, medications, length of stay.
2. Demographic data: age, gender, socio-economic status, past readmissions.

Ethical Concerns:

1. Patient privacy: Ensuring data is anonymized and handled in compliance with HIPAA.
2. Bias in data: Historical healthcare disparities may lead to model bias against minority or underserved populations.

Preprocessing Pipeline:

1. Handle missing values through imputation or exclusion.
2. Normalize numerical features (e.g., age, lab results).
3. One-hot encode categorical features (e.g., diagnosis codes).
4. Engineer features such as 'number of prior admissions' or 'comorbidity count'.

## Model Development

Model Chosen: Random Forest Classifier.

Justification: Robust to overfitting, handles both numerical and categorical data, interpretable through feature importance.

Practical Confusion Matrix Results:

	Predicted Readmit	Predicted No Readmit
Actual Readmit	19	0
Actual No Readmit	0	211

Precision: 1.0

Recall: 1.0

Interpretation: The model achieved perfect performance on this test set. This could be due to a small or overly clean dataset and would need real-world validation with a larger, messier dataset.

## Deployment

Steps to Integrate Model into Hospital System:

1. Host the model using Flask or FastAPI.
2. Integrate API with the hospital's EHR system.
3. Create dashboards for clinicians using tools like Streamlit or PowerBI.
4. Regularly retrain the model with updated patient data.

Compliance with Regulations:

- Follow HIPAA guidelines.
- Ensure end-to-end encryption of data.
- Audit trails and logging for traceability.

## Optimization

Method to Address Overfitting:

- Use k-fold cross-validation.
- Prune the Random Forest (limit max depth, minimum samples per split).