Understanding the AI Development Workflow
Course: AI for Software Engineering
Assignment: Predicting 30-day Hospital Readmission Risk
Author: austine makwka
Date: 2025-11-07

--- PART 1: Short Answer Questions ---

1. Problem Definition
Problem: Predict which discharged patients are likely to be readmitted within 30 days.
Objectives:
  i. Accurately predict 30-day readmission risk enabling targeted follow-up.
  ii. Reduce avoidable readmissions and associated costs.
  iii. Ensure fairness and privacy are maintained.
Stakeholders: clinicians, hospital administration, patients.
KPI: AUC-ROC

2. Data Collection & Preprocessing
Data sources:
  i. Electronic Health Records (EHR): admission/discharge notes, vitals, labs, medications.
  ii. Administrative claims / billing & social determinants (if available).
Potential bias:
  - Under-representation of disadvantaged populations leading to worse performance for these groups.
Preprocessing steps:
  i. Missing data imputation (median for numerics, mode for categoricals).
  ii. Normalization/standardization of numeric features.
  iii. Categorical encoding and feature selection; derive time-series aggregates when applicable.

3. Model Development
Model: Multi-Layer Perceptron (Neural Network).
Justification: Flexible non-linear modeling for tabular EHR features; can learn interactions and combine diverse features (after car
Data split:
  - Training: 70% (stratified)
  - Validation: 15%
  - Test: 15%
Hyperparameters to tune:
  i. hidden_layer_sizes (controls model capacity)
  ii. alpha (L2 regularization strength; helps prevent overfitting)

4. Evaluation & Deployment
Metrics:
  - AUC-ROC for overall discriminative ability.
  - Recall (sensitivity) to minimize missed high-risk patients.
Concept drift:
  - Definition: population or practice changes that alter input or label distributions over time.
  - Monitoring: track rolling AUC, calibration, and feature distribution shifts; periodic retraining or incremental updates.
Technical deployment challenge:
  - Integration with hospital EHR systems, ensuring low-latency inference and meeting security/authorization constraints.

--- PART 2: Case Study Application (Hospital Readmission) ---

Problem Scope:
  - Predict 30-day readmission for adult patients at discharge.
  - Objectives: improve post-discharge interventions, allocate resources, evaluate model fairness.
  - Stakeholders: clinicians, care coordinators, hospital IT, patients, compliance officers.

Data Strategy:
  - Sources: EHR clinical data (demographics, labs, vitals), prior admissions, medication lists, discharge disposition, and social det
  - Ethical concerns:

1. Patient privacy and data protection under HIPAA.
2. Algorithmic fairness: avoid systematically disadvantaging subpopulations.

- Preprocessing pipeline (high level):
  - i. Data ingestion & de-identification.
  - ii. Missingness analysis and imputation.
  - iii. Feature engineering: count of prior admissions, Elixhauser/Charlson comorbidity scores, lab trend features (e.g., delta crea
  - iv. Encoding categorical features and scaling numerics.
  - v. Train/val/test stratified splitting.

Model Development:
- Choose MLPClassifier (Neural Network) with 2 hidden layers, ReLU activations, early stopping and dropout (or L2).
- Hypothetical confusion matrix (test) for 1,000 patients:

	Predicted Negative	Predicted Positive
Actual Negative	820	50
Actual Positive	80	50

  - From this CM:
    Precision = TP/(TP+FP) = 50/(50+50) = 0.50
    Recall = TP/(TP+FN) = 50/(50+80) = 0.38
- Interpretation: model has moderate precision and low recall — tune threshold to increase recall for clinical use.

Deployment:
- Steps to integrate:
  1. Wrap model as a RESTful API (FastAPI) with authentication and input validation.
  2. Containerize (Docker) and deploy to hospital on-prem or private cloud.
  3. Integrate via EHR middleware (HL7/FHIR connector) or build an EHR plugin/dashboard for clinicians.
  4. Logging, monitoring, and alerting (performance and data drift).
- HIPAA compliance:
  - Ensure encryption in transit & at rest (TLS, disk encryption).
  - Access controls and audit logs, Business Associate Agreements (BAAs) where needed.
  - De-identification/pseudonymization for data used in model development.

Optimization to address overfitting:
- Use early stopping, L2 regularization (weight decay), and dropout. Consider cross-validation and model ensembling as needed

--- PART 3: Critical Thinking ---

Ethics & Bias:
- Example harm: If the training data under-represents patients from certain socioeconomic groups, the model may under-predict
- Mitigation strategy: stratified sampling, fairness-aware reweighting, subgroup performance audits, and including protected attrib

Trade-offs:
- Interpretability vs accuracy: Clinicians often require interpretable explanations; a complex neural network might perform slightly
- Resource constraints: Limited compute pushes towards smaller models (logistic regression, decision trees) or model compress

--- PART 4: Reflection & Workflow Diagram ---

Reflection:
- Most challenging: ensuring representative, high-quality labeled data while maintaining privacy/legal compliance.
- With more resources: collect multi-site datasets, run prospective validation, and implement clinician-in-the-loop deployment.

Diagram:
- See diagram.png included in this package (CRISP-DM style: Business Understanding → Data Understanding → Data Preparat

References:
- CRISP-DM framework, standard ML texts, HIPAA guidance and healthcare ML papers (detailed references in the full PDF).