**Required Capstone Assignment 6.1**

**Draft the Problem Statement [02:00:00]**

Professional Certificate in Machine Learning and Artificial Intelligence

Karen Chang

Prof. Arya Roy

August 7, 2025

**Problem Statement:**

As a clinically experienced registered nurse from Taiwan with a license, I have frequently observed patients with chronic conditions return to the hospital shortly after discharge. My objective for my capstone project is to develop a predictive model that can identify patients who are at a high risk of being readmitted within 30 days of discharge. Finding high-risk patients early on can help nurses and care teams provide focused teaching and follow-up care after a patient leaves the hospital, which will eventually improve patient results and reduce the hospital's workload.

**Expected Data:**

In order to build the model, I would need information about the patient's age, gender, primary diagnosis, and any other conditions they may have, like diabetes or heart failure. I would also need information about their hospital stays and follow-up appointments, as well as information about their discharge.

The UCI Diabetic Readmission dataset or a de-identified EHR dataset with comparable features could serve as an appropriate open-source dataset.  
(<https://archive.ics.uci.edu/dataset/296/diabetes+130-us+hospitals+for+years+1999-2008>)

**Possible Techniques:**

Given that this is a classification problem (predicting whether a patient will be readmitted), I intend to investigate supervised learning techniques, including logistic regression, decision trees, and random forest classifiers. I may also employ techniques such as feature selection and oversampling (e.g., SMOTE) to address class imbalance, contingent upon the dataset.