

## Capstone Project Planning Document

**Data:** Healthcare Utilization Project National Readmission Database (HCUP NRD)

<https://www.hcup-us.ahrq.gov/nrdoverview.jsp>

<https://www.hcup-us.ahrq.gov/db/nation/nrd/nrddbdocumentation.jsp>

This database contains 17,686,511 encounter-level records and over 100 variables. These variables include those necessary for readmission analyses (e.g., patient linkage numbers, length of stay, time between stay), patient demographics (e.g., sex, age, income quartile), hospital characteristics (e.g., number of beds, ownership, teaching status), financial information (e.g., total charges, payment source, cost-to-charge ratio), ICD-10-CM/PCS codes, medication codes, severity information (i.e., mortality, risk of mortality, severity of illness), and sample weights for generating national estimates.

**Track:** Modeling-Focused Engineer

Analyzing the HCUP NRD will require a substantial amount of data processing and feature engineering; for example, I will need to reshape data to link readmissions, aggregate diagnosis and procedure/medication codes, generate comorbidity and interaction terms, and conduct typical data preparation tasks (such as cleaning and scaling variables). However, I am particularly interested in building different models, tuning hyperparameters, employing advanced techniques like blending or stacking, and comparing different solutions with visualizations.

**Problem:** Classification of readmission status for ICD-10-CM/PCS Codes representing Mental, Behavioral and Neurodevelopmental Disorders. A solution I have considered is using a random forest decision tree with boosting to solve this classification problem.

F01-F09	Mental Disorders due to known Physiological condition
F10-F19	Mental and Behavioral Disorders due to Psychoactive Substance Abuse
F20-F29	Schizophrenia, Schizotypal, Delusional, and Other Non-Mood Disorders
F30-F39	Mood Affective Disorders
F40-F48	Anxiety, Dissociative, Stress, Somatoform and Nonpsychotic Mental Disorders
F50-F59	Behavioral Syndrome Associated with Physiological Disturbances and Factors
F60-F69	Disorders of Adult Personality and Behavior
F70-F79	Intellectual Disabilities
F80-F89	Pervasive and Specific Developmental Disorders
F90-F98	Behavioral and Emotional Disorders with Childhood and Adolescence Onset
F99	Unspecified Mental Disorder