Title: Mayo TCU STR Placement Need Predictor

The Mayo Clinic Health System (MCHS) is a community-based health care provider that brings the resources and expertise of the Mayo Clinic to communities in the state's broader regions, including our city – Mankato.

When patients are admitted to the hospital, MCHS is not currently immediately evaluating whether the patient will be able to immediately return home (or to whatever location that arrived from) upon release. In cases where patients need to be placed in short term rehabilitation (STR) care at release, this often leads to delays at discharge. These delays cause a variety of medical, operational, and financial problems for both patients and caregivers – both at the hospital ("the acute care setting") and at receiving STR sites ("transitional care settings").

The goal of this project is to build a prediction tool that can be run by MCHS staff at the time a patient is admitted. An effective tool will give the MCHS care team a "heads up" to start working on post-release plans earlier (in cases where this is needed), thereby eliminating (or at least majorly minimizing) discharge delays.

The project team will work with MCHS leaders and staff members to design and build a system that analyzes a variety of data sources to automatically predict patient placement needs at discharge. The first step will be organizing and preparing data drawn from patient medical histories, patient home setting factors, planned in-hospital treatments after admission, patient intake interviews, and possibly other sources drawn from existing Mayo social work processes (*ref.* Rochester). For this data engineering phase of the project, the team will work under the guidance of the MCHS client to assure that proper precautions are taken to de-identify patient data and protect data privacy per HIPAA and other regulatory frameworks.

(Depending on various client-side conditions, the initial/prototype data used for this project may need to be experimentally synthesized after the data fields and structure have been agreed with the client.)

The project team will then use data science and machine learning techniques to build a software-based predictor that produces the likelihood that an admitted patient will need STR/transition care upon release. The predictor's output will be an *a posteriori* value between 0 and 1. This predictor output will be presented to MCHS caregivers via an additional, simple "wrapper" application (details TBD) that indicates red/yellow/green based on thresholds on the predictor's output. These thresholds will be configurable by the MCHS care team in the hospital.

The MCHS client lead for this project will be April Lanz (Lanz.April@mayo.edu).

Deliverables	Type of work	Activities	Resources	Tech Skills	Priority
Requirements Analysis Document (RAD)	Requirements analysis	Work with main client and her colleagues/SMEs to understand the details of what the predictor needs to do, and why, and how. Includes requirements for the red/yellow/green "wrapper" app around the ML predictor.	Client lead and SMEs	System and requirements analysis and documentation	High
Data Sources and Data Model Design and Plan Document (DSDM)	Data engineering and data model planning – what are the fields and structures of the data to be used?	Work with main client and her colleagues/SMEs to identify all useful and reasonably accessible sources of data to be used in construction of the predictor. Define all data fields and structures. Define all data pull/organization/ETL pathways. Assure compliance with HIPAA and other regulatory frameworks.	Client lead and SMEs. Project coach and online materials about data engineering, HIPAA, and other regulatory constraints.	Data engineering, data model design and documentation	High
Data Gathering and Aggregation – data set to be used for training and testing	Creating the data set to be used for training and testing	If the client is able to support the needed data pulls (ETL), build the data set from actual client data. If not, create a synthesized data set to support the next phases of the project.	Client lead and SMEs, project coach, and – if needed – online information about how to synthesize data sets against a data design plan	Data engineering, ETL tools	High
Machine learning predictor – working system, notebooks, and documentation	Machine learning modeling and documentation	Use machine learning techniques to build a predictor. Create a work system, document design choices in digital notebooks, provide thorough documentation of design, test, and execution environments and choices. Important to include test performance results.	Project coach, Freely available ML tools	Data science and machine learning, Python, SciKit learn, Jupyter, etc.	High
A red/yellow/green application for use by MCHS hospital staff	Application development	Prototype a "field use" (hospital use) app that wraps the predictor into an easy-to-use format for hospital staff (platform(s) TBD).	Client lead and colleagues	App dev (platforms TBD)	Medium
Handover document	Project next-phase planning and documentation	Descriptions and plans for what to tackle in the next phase of the project. An example might be: if this project team used "only" synthesized data, how to move to "real" data in the next phase. Another example might be: if more data or further ML experimentation is required to get good-enough performance, focus there. Yet another example might be: if the red/yellow/green app is far too simple or only runs on one platform, make further app extensions.			High