Background & Introduction

Accurate estimates of surgery duration are critical to the optimal scheduling of Operating Theatres (OT). This can lead to hospitals being better able to provide better patient care and help to reduce operation cost. Previous work has already been done to predict surgical durations through a Gradient Boosting Model (gbm) fitted using data from historical surgeries. Despite so, these predictions (i.e., estimations) are not made accessible for Operating Theatre Management Unit (OTMU) staff to refer to when they make bookings for OT.

Furthermore, the current prediction generation process can also be made more scalable through the use of a database for data storage. Having a database would also help to address issues such as data integrity and security, and make it much easier to accommodate more (sources of) data in the future.

Methodology

We initialized a MySQL database and created multiple tables according to the following steps:

**Step 1:** Cleaned historical surgery data which has been extracted as a csv file from the data lake is imported into a newly created table named ‘cleaned’.

**Step 2:** We performed feature engineering to create new variables that will be used in the machine learning model, and store the augmented data as ‘prepared’.

**Step 3:** The prepared data is used as input to fit the machine learning model (gbm algorithm) that is coded using R software.

**Step 4:** The fitted model is used to generate surgical duration predictions for historical surgeries and this information is then stored in another table called ‘predicted’.

**Step 5:** The dashboard then reads, and selectively displays data from ‘predicted’.

Aim

- To develop an operational dashboard to facilitate surgical duration estimates made by OTMU staff for OT scheduling.
- To create a database that allows easy data storage and access for predictive model and dashboard.

Workflow

![Workflow Diagram]

- **cleaned** (extracted data)
- **prepared**
- **predicted**

Results

We were successful in setting up the MySQL database to store the required data, and replicate data processing steps in MySQL were previously done in R to produce equivalent results.

Using free open-source software (Python), we were also able to code a dashboard that is able to display key information relevant for estimating a certain surgery’s expected surgical duration.

Conclusion & Future Work

- Setting up of an intermediate database helps to store data in an accessible manner. This facilitates both predictive machine learning modeling work that requires such data as input, and the recording and consequent display of predictive model output results and key surgery information to stakeholders.
- Future work can include: (1) Automating the whole dashboard’s data preparation (i.e., database’s tables creation) process, (2) Compare with order database options and choose the most efficient database, (3) Improving the dashboard’s visualizations and user-intuitiveness.