Task - Recommendations to Improve Hospital Operations

Request

You are a data scientist at a hospital which specializes in diabetes care. In order to improve hospital operations, an estimate of a patient's cost of care at various stages in the patient's journey would be necessary. You are given the task of creating a recommendation report that is based on a predictive model you are developing. In the process, you are making recommendations about which patients should be tended to first in order to lower future expenditure in an Excel workbook called Patient_Recommendations_Cost_Insights.xlsx. The first spreadsheet should be called Patient_Recommendations, which will contain the columns "patient_id", "priority" (determines whether the patient should be prioritized first), "estimated_cost_of_care" which is the prediction from the model as a cost, Explanation, which describes the drivers of the cost prediction for that patient, and Recommendation, which will provide a suggestion to the care team as to what to do with the patient. The second spreadsheet must be called Priority_Summary which will have two tables. The first table will describe the patient priority level where the levels include: "Critical", "High", "Moderate", and "Standard". The table should have the columns: "Priority Level", "Patient Count", "Total Estimated Cost", and "Average Cost per Patient". The second table should be "Distribution by Journey Level". This table should have the columns: "Journey Level", "Patient Count", and "Total Estimated Cost".

Assume that the patient's journey can fall under various categories such as "inpatient admissions", "emergency department visits", "getting medications from a pharmacy", or "going towards hospice care". The hospital you work at has cost accounting completed in the form of Excel spreadsheets which are updated once per quarter. The workbook contains three tabs: "Monthly_Expenditure" which shows details about the various transactions the hospital has undergone, "Patient_Transaction_Mapping" which maps the transaction identifiers (IDs) to specific patient IDs that can be linked back to the semantic layer, and the "Summary_Statistics" tab, which goes into the total transaction breakdown and expenditure by various categories. A consolidated spreadsheet has been provided to you in the Excel file diabetic_patient_expenditure. The rest of the data collected by the hospital comes from an internal relational database. The schema and tables within the database have been provided to you in the file semantic_database_schema.pdf.

Create the predictive model using the PyTorch framework in Python and build the recommendation system using any learning to rank capabilities using the torchrec library which specializes in building recommendation systems using PyTorch. Include any written explanations of model explainability

metrics to understand drivers of cost such as "feature importance", "Shaply Additive Explanations (SHAP) values", "integrated gradients", "attention maps", etc., to make the analysis more transparent to clinical staff. For example, a SHAP value of num_procedures being the highest can translate to "The patient has undergone a lot of procedures during care.""

Request - Resources

- Semantic Database
- Diabetic Patient Expenditure - Excel

Deliverable - Introduction

Here is the attached Excel file, Patient_Recommendations_Cost_Insights, you requested. The spreadsheet titled Patient_Recommendations shows cost estimates of patients along with the "Explanation" column which delves deeper into cost drivers. The"Recommendation" column provides a natural language explanation as to why the patient should be prioritized by the hospital staff as a means of lowering future expenditure. The "Recommendation" column always starts with with one of the following options: "IMMEDIATE ACTION", "URGENT", "CRITICAL PRIORITY", "RECOMMEND", "PRIORITIZE", "MODERATE PRIORITY", "STANDARD FOLLOW-UP", "MONITOR", "STABLE", "MAINTENANCE CARE", and "PREVENTIVE FOCUS". The second Excel spreadsheet titled Priority_Summary shows a breakdown by patient priority as well as the distribution by journey level.

Deliverable Files

- Patient_Recommendations_Cost_Insights.xlsx