Database System Project III

PART-3: Logical Schema Optimization and Machine Learning Model Creation

12.07.2022

Ritika Motwani (N12868512) - rm6383
Sneha Singh (N11049579)-ss15639

Overview for Part III

Enterprise Data Architecture is a discipline designed to simplify, streamline, standardize, and enhance the accessibility of your organization's data. A successful enterprise data architecture plan should cover the policies, procedures, standards, on how data is collected and stored in addition to how data is managed, processed, and used throughout the organization.

Most larger organizations are currently dealing with an extremely fragmented data landscape, with substantial redundancy. In many cases, this landscape has evolved for a variety of reasons, such as rapid growth, limitations with legacy technology, and the lack of investment in managing your organization's data assets.

Continuing from Part-2 of our project, we selected the data:
Indicators for Chronic Disease Surveillance — United States, 2013 for this project.
https://catalog.data.gov/dataset/u-s-chronic-disease-indicators-cdi

The analysis of models performed can be viewed in the Python notebook:
https://colab.research.google.com/drive/1GbjPQevxoNXiF17vScyg9a8Niekx2xM4#scrollTo=Qa9owMNpTbHc