Week 7: Project Healthcare – Persistency of a Drug

Group Member (Solo):

- *Name: Halit Ayberk Demir*
- *E-Mail: h.ayberk.demir.34@gmail.com*
- *University/Company: Middle East Technical University/Data Analyst at Mega Pharma*
- *Specialization: Data Science*

Problem Description

ABC is a pharmaceutical company and desires to understand the persistency of the drug per physician description. To solve this issue, ABC company reached an analytics company to automate this process of identification.

Business Understanding

The dataset contains 3424 observations about the patients who are prescribed the NTM (non-tuberculous mycobacteria) Rx. Target variable is 'persistency flag' this feature indicates if the patient is persistent or not and my aim is the automate the process of finding out if the patient will be persistent or not. General topics in features are; features about the patient, features about the provider, clinical factors, disease/treatment factor. With this information I will be classifying if a patient is persistent or not.

Project Lifecycle

Week 7	*-Problem Understanding*
Week 8	*-Feature Engineering/Data Cleaning*
Week 9	*-Model Development*
Week 10	*-Model Selection*
Week 11	*-Model Evaluation/ Reports*
Week 12	*-Model Deploying*
Week 13	*-Final Report*

Data Intake Report

Name: HealthCare - Data Science

Report Date: 21/4/2023

Internship Batch: LISUM19

Data intake by: Halit Ayberk DEMIR

Data intake reviewer: Data Glacier

Data storage location: https://github.com/demirayberk/Data-Science-HealthCare/blob/main/Healthcare_dataset.xlsx

Total Number of Observations	*3424*
Total Number of Files	*1*
Total Number of Features	*26*
Base Format of the File	*.xlsx*
Size of The Data	*898 Kb*

Github Repo of the Project: https://github.com/demirayberk/Data-Science-HealthCare

Halit Ayberk DEMIR