

Healthcare - Persistency of a drug

Data Science Final Project Report

Group Name: Data Science Warrior

Name: Ugur Selim Ozen

Email: ugur_ozen58@hotmail.com

Country: Turkey

Collage/Company : Yıldız Technical University

Specialization : Data Science

Submission date: 11.08.2021

Github Repo Link:

https://github.com/UGURSELIMOZEN/Data_Glacier_DS_Internship/tree/main/DataScience_Healthcare_Final_Project

1.Problem Description

With an objective to gather insights on the factors that are impacting the persistency, build a classification for the given dataset.

2.Business Understanding

One of the challenge for all Pharmaceutical companies is to understand the persistency of drug as per the physician prescription. To solve this problem ABC pharma company approached an analytics company to automate this process of identification.

3.Project Lifecycle along with deadline

- Problem Understanding (Week 7) – November 8 2021
- Data Understanding (Week 8) – November 8 2021
- Data Cleaning and Feature engineering (Week 9) – November 15 2021
- Exploratory Data Analysis (Week 10) – November 22 2021
- Model Development (Week 11) – November 29 2021
- Model Selection (Week 12) – December 06 2021
- Model Evaluation (Week 12) – December 06 2021

Final Week 13 – December 18 2021

- Report the accuracy, precision and recall of both the class of target variable
- Report ROC-AUC as well
- Deploy the model
- Explain the challenges and model selection

4.Data Intake Report

Tabular data details:

Total number of observations	3424
Total number of files	1
Total number of features	69
Base format of the file	.xlsx
Size of the data	1.8 MB

My Considerations:

.xlsx dataset has no missing values for all features.