

Group project: Week 7 report

Team: FeatureTransformers

Members: Wangu Ndungu, Nikola Andrejić

Batch code: LISUM10

Contents

1	Team and team members details	3
2	Project details	4
2.1	Introduction	4
2.2	Problem statement	4
2.3	Business understanding	4
3	Project Lifecycle	5
4	Data Intake report	6

1 TEAM AND TEAM MEMBERS DETAILS

Team name: **FeatureTransformers**

Team members:

name: **Wangu Ndungu**

email: **nwangu349@gmail.com**

country: **Kenya**

college/Company: **Kenyatta University**

specialization: **Data Science**

name: **Nikola Andrejić**

email: **nikola.ing.nl@gmail.com**

country: **Serbia**

college/Company: **University of Niš**

specialization: **Data Science**

2 PROJECT DETAILS

Project title: Drug persistence and Medical Adherence

2.1 INTRODUCTION

Let's start things off by defining the terms medical adherence and drug persistence.

The U.S. Food and Drug Administration (FDA) terms medical adherence as: *"The extent to which patients take medication as prescribed by their doctors. This involves factors such as getting prescriptions filled, remembering to take medication on time, and understanding the directions."*

Drug persistence can be defined as the extent to which a patient acts in compliance to the prescribed interval, and dose of a dosing regimen.

What is the difference between these two terms? Adherence refers to the proportion of pills taken within a specific time interval and persistence refers to the continuing use (in time) of the prescribed therapy.

2.2 PROBLEM STATEMENT

According to the World Health Organisation, only 50-70% of patients adhere properly to prescribed drugs during therapy. This is especially true among those with long term medication. This worrying statistic is caused by various factors, for example: patient's condition or disease, their socio-economic status, confusion by the schedule, forgetting, discontinuing because they feel better, just to name a few. Medical non-adherence can lead to devastating consequences on one's health, especially those with chronic illnesses.

The purpose of this project is to study trends among patients in a sample and build a model that'll classify a new patient as Persistent or Non-Persistent.

This project will give medical practitioners(especially pharmaceuticals) insight on which patients might require more rigorous follow-ups to ensure they will adhere to their prescriptions.

2.3 BUSINESS UNDERSTANDING

According to a study carried out by LexisNexis in 2020, Medical Non-adherence is one of the biggest issues faced by the pharmacy industry. Pharmacists stated that they would want to put more effort into educating patients on the importance of medical adherence, send reminders to them, improve drug packaging and much more in order to deal with the issue.

Technology and Machine Learning being incorporated in the industry would help pharmacists identify patients who are likely to be non-adherent and place additional monitoring and follow-ups on them.

3 PROJECT LIFECYCLE

1. Problem statement and business understanding (19.7.2022)
2. Data understanding (26.7.2022)
3. Data cleaning and preparation for analysis (2.8.2022)
4. Performing EDA on the data (9.8.2022)
5. Presenting the results of EDA in the form of business presentation (16.8.2022)
6. Evaluating models from different families on our problem (23.8.2022)
7. Final report (30.8.2022)

4 DATA INTAKE REPORT

The data intake report is attached as the next page to this document.

Data Intake Report

Name: Drug persistence and Medical Adherence

Report date: 16/07/2022

Internship Batch: LISUM10:30

Version:1.0

Data intake by: Wangu Ndungu

Data intake reviewer: Nikola Andrejić

Data storage location: <https://www.kaggle.com/datasets/harbhajansingh21/persistent-vs-nonpersistent>

Tabular data details:

Persistent vs NonPersistent

Total number of observations	3425
Total number of files	1
Total number of features	69
Base format of the file	.csv
Size of the data	889KB

Proposed Approach:

- Exploratory data analysis to discover outliers, skewness, trends and correlation among various variables in the dataset.
- Building a classification model that'll predict a patient's chances of being persistent or non-persistent with their prescribed drugs.