

Data Science Project

Healthcare – Persistency of a Drug

Week 12 works

Team member's details:

Group Name:	DS_SS
Name:	SONIYA SUNNY
Email:	soniyasunny1210@gmail.com
Country:	Canada
College/Company:	Data Glacier
Specialization:	Data Science

Problem Description

To identify the persistency of a drug, a pharmaceutical company approached to develop a model based on data analysis. Factors that affect the persistence of drugs should be identified, along with data insights with predictive analytics, to help the company for their smooth and efficient functioning, with the help of dataset provided by the company.

Github Repo Link

[Final Project DS SS/week 12 at master · Soniyasunny1/Final Project DS SS \(github.com\)](#)

Week 12: Model Selection and Model Building

Tried four different types of models:

1. Logistic Regression
2. Decision Tree Classifier
3. Random Forest Classifier
4. KNN

Random Forest Classifier is the best model in terms of accuracy.

Data Intake Report

Name: Data Science Final Project – ‘Healthcare – Persistency of a Drug’

Report date: August 12, 2022

Internship Batch: LISUM10: 30

Version:<1.0>

Data intake by: Soniya Sunny

Data intake reviewer:<intern who reviewed the report>

Data storage location: [Healthcare dataset.xlsx - Google Drive](#)

Tabular data details:

Total number of observations	3425
Total number of files	1
Total number of features	69
Base format of the file	.xlsx
Size of the data	899 KB