Chris Donaton, Data Science specialization

Problem description

Given an array of features that describe a patient's medical history, calculate that patient persistency likelihood.

Business understanding

In the medical field, persistence is defined as continuing a course of therapy prescribed by a medical professional. A pharmaceutical company that can't keep its customers using its product loses money. It is a well-regarded fact that obtaining new customers is more than double the cost of keeping current ones.

As a society, patients that do not follow long term medical advice drain additional resources from an already overloaded medical system. Low levels of persistency lead to 125,000 deaths per year and cost the U.S. health care system $100 billion annually. (Rubin, 2006).

Project lifecycle along with deadline

26 Jan 2023:   Initial analysis

2 Feb 2023:   Establish Data pipeline (ETL)

9 Feb 2023:   EDA and Analysis Code

16 Feb 2023:   EDA Presentation and proposed modeling technique

23 Feb 2023:   Model Selection and Model Building/Dashboard

28 Feb 2023:   Final Project Report and Code

Github Repo: https://github.com/rambles-tech/DG_virtual_internship/week07

Rubin, D. (2006). Patient retention strategies, DTC Perspectives, December, 26–30.