



Predict Adverse Drug Reaction with Pharmicovigilance

16.04.2021

19CE018-VANSH DESAI

19CE129-AMAN SHAH

Overview

This is the research project about predicting adverse reactions of covid-19(coronavirus) vaccines. We have a dataset from the Vaccine Adverse Event Reporting System (VAERS) and with the help of that we can predict several outcomes and a particular person can get its vaccine according to its need.

Goals

1. In this pandemic situation many research work has been done and our goal is to share every work with the world so that it saves much time and energy.
2. This project predicts ADRs of covid-19 vaccines so that particular person can get a vaccine according to their body type and body health.
3. Despite this project, this system can also predict ADRs of other medicines and through that we can save many lives and also we can save money & time.

Specifications

This project is all about prediction of adverse reaction of covid-19 vaccine. We downloaded three excel files from VARS for data symptoms and vaccines and we generated a relation between them through primary key patient_id. Our project has all the records of covid-19 vaccines till date which are accurate and are recorded by world's best doctors.

We used Machine learning to build this project our data was stored in microsoft excel later on through panda we called every data in google colab and this is how using above softwares we build this project.

Our main space for working was google colab where we used python as our language and with this it makes our project a perfect example of python+machine learning projects.

Milestones

I. Challenges faced by us:

- As there were thousands of data each data came with many null values and therefore it was very difficult to remove those values.

- After removing those values it was difficult to replace those values with the new values as this is a project related to the medical field now we have to choose our data wisely.
- The input to this transformer should be an array-like of integers or strings, denoting the values taken on by categorical (discrete) features. The features are encoded using a one-hot encoding scheme. This creates a binary column for each category and returns a sparse matrix or dense array depending on the sparse parameter.
- Correlation is an indication about the changes between two variables. We have discussed Pearson's Correlation coefficients and the importance of Correlation too. We can plot correlation matrix to show which variable is having a high or low correlation in respect to another variable.

II. Motivation behind this project:

- ADRs are a huge burden of financial resources and labor.
- It was estimated by Lazarou et al that more than 100,000 die from ADRs annually. Moreover, Sultana et al found that around \$30.1 billion are spent annually on ADRs in the U.S., and nearly half of these costs can be prevented based on a study by Bates et al.
- The unnecessary high financial costs and labor spent on ADRs provide strong motivation to learn more about ADRs and be able to predict the probability of ADRs accurately.