

Problem Statement

The goal of this project is to "predict the potentially fraudulent providers" based on the claims filed by them. Along with this, we will also discover important variables helpful in detecting the behaviour of potentially fraudulent providers. Further, we will study fraudulent patterns in the provider's claims to understand the future behaviour of providers.

Introduction to the Dataset

For the purpose of this project, we are considering Inpatient claims, Outpatient claims and Beneficiary details of each provider. Let's see their details:

A) Inpatient Data

This data provides insights about the claims filed for those patients who are admitted in the hospitals. It also provides additional details like their admission and discharge dates and admitted diagnosis code.

B) Outpatient Data

This data provides details about the claims filed for those patients who visit hospitals and not admitted in it.

C) Beneficiary Details Data

This data contains beneficiary KYC details like health conditions, region they belong to etc.