

Problem Statement:-

As we are well aware of the fraudulent practices in the health care fields like:

- Using insurance that belongs to someone else.
- Providing false information in order to receive medical coverage.
- Duplicating claims for the same or different patients.
- Providing a service that is not covered under a patient's insurance policy, then billing for a service that is covered.

How we came up with the idea:-

During the Covid-19 pandemic, we have seen cases of fraud in healthcare sector.

Recognizing the prevalence of healthcare fraud and its detrimental impact on patients and resources, we were compelled to find a solution. The potential to leverage technology for fraud detection, improve resource allocation, and enhance patient care became our driving motivation. With a vision of a more transparent and efficient healthcare system, we embarked on the journey to develop a predictive model that would contribute to a healthier and more trustworthy medical environment.

Goal:-

The goal of this project is to understand the behavior of potential fraud providers and discover the important parameters that can help us, and finally build a predictive model that can flag the potential fraud providers based on the claims fields.

Datasets used:-

1. Beneficiary like Age, Race, gender, DOB, etc.
2. Inpatient/outpatient claims information such as the amount, Medical codes, Physician ID
3. Labels (Fraud or Not Fraud)

Outline:-

1. Firstly, we merged all the datasets together and then we removed all the null values from the dataset and dropped the columns with missing values. After that, we explored the relations between the features and the target "Possible Fraud" variable.
2. Next, we split the data into train and test and then used logistic regression to classify the trained model on the basic or potential fraud.
3. We cross-checked the predicted values against the tested "Possible Fraud" variable and evaluated the accuracy of the model using plots.

Conclusion:-

Improving this model, will help in predicting Provider fraud, help Insurance companies and the Government to take decisions against fraudulent health providers, and also improve the health of the economy.

Future Scope:-

1. We can improve the accuracy of the model by using more complex machine learning techniques like Support Vector Machine (SVM), Naive Bayes, etc.
2. We can build a model which would predict values based on real-time data which would help provide a broader scope.
3. Integration with Claims Processing Systems: Integrate your model with healthcare claims processing systems. This integration could automatically flag potential fraud cases and provide recommendations for further investigation.
4. We can also implement our model in already existing fraud detection software which would further enhance their performance.
5. Continuous benchmarking of the model against other fraud detection systems in order to constantly keep track of its efficiency.