Healthcare Fraud Detection - Complete Column List

Introduction

This document contains a comprehensive list of structured and text columns that can be used for healthcare fraud detection models. The structured columns provide numerical and categorical data, while the text fields are designed for NLP-based analysis to detect fraud in healthcare claims.

1. Structured Columns

- Claim ID
- Provider ID
- Patient ID
- Claim Amount
- Diagnosis Code
- Procedure Code
- Patient Age
- Patient Gender
- Date of Service
- Service Count
- Total Submitted Charges
- Total Payments
- Bene Count
- Fraudulent Flag (0/1)

2. Text Fields for NLP

The following text fields are suitable for applying Natural Language Processing (NLP) techniques to detect patterns and anomalies that might indicate fraud in healthcare claims.

- Billing Description: Describes the services or procedures billed to the insurance company.

- Diagnosis Description: A textual description of the diagnosis related to the claim.
- Procedure Notes: Provider's notes detailing the procedures performed during the service.
- Provider Notes: Healthcare provider's notes from the patient's visit.
- Prescription Details: Describes medications or prescriptions given to the patient.
- Treatment Plan or Recommendations: Contains the treatment plan or recommendations from the healthca
- Discharge Summary: A summary of the patient's condition at discharge.
- Patient Complaints: Text that describes the patient's reason for visiting the healthcare provider.
- Referral Letter: Describes the reason for referring the patient to a specialist.
- Medical History: Describes the patient's medical history.
- Insurance Claim Explanation: Explanation or justification provided by the healthcare provider for a specifi
- Authorization Letters: Pre-authorization letters from the insurance company.
- Appeal or Denial Letters: Letters explaining the reason for the denial or appeal of an insurance claim.

Conclusion

By combining structured columns with text fields and applying NLP techniques like text classification, named entity recognition, and topic modeling, a healthcare fraud detection system can effectively identify suspicious claims and fraudulent activities.