

## Venkata Surya Abburi

### Resume\_proejct\_1(health insurance claim predictor)

#### □ **Project Name & Purpose** –

“My project is called *AI-Powered Health Insurance Claim Predictor*. I built it to automate claim verification and detect fraud in the health insurance sector.”

#### □ **Real-life Use Cases** –

Used by insurance companies to flag suspicious claims, speed up claim settlement, reduce manual verification, and improve customer trust.

#### □ **Impact in the Industry** –

Reduces claim processing time, saves operational costs, improves fraud detection accuracy, and supports faster decision-making for insurers.

#### □ **Why These Technologies** –

- **Python & Pandas** – Data cleaning and processing.
- **Scikit-learn** – Fraud detection model building.
- **LangChain** – Summarizing complex medical reports and enabling claim query chatbot.
- **Power BI** – Real-time claim analytics dashboard.
- **Flask** – Lightweight deployment for end-to-end accessibility.

#### □ **Result** –

Achieved high fraud detection accuracy, automated medical report understanding, and delivered a unified solution for insurers.

#### □ **Q:** Which algorithm did you use for fraud detection and why?

**A:** I used **XGBoost** because it handles imbalanced data effectively, is fast, and delivers high accuracy on structured datasets like claim records.

#### □ **Q:** How did you handle class imbalance in your dataset?

**A:** I applied **SMOTE oversampling** for the minority class and adjusted class weights in the model to improve recall on fraudulent claims.

□ **Q:** Why did you use LangChain in this project?

**A:** LangChain allowed me to **summarize lengthy medical reports** into short insights and power a chatbot for quick claim-related queries.

□ **Q:** Why did you choose Flask over Django?

**A:** Flask is **lightweight and easy to integrate** with ML models, making it ideal for quick deployment without the overhead of a full framework.

□ **Q:** Which evaluation metric was most important and why?

**A:** I focused on **Recall** and **F1-score** because in fraud detection, catching fraudulent claims is more critical than overall accuracy.