



# AI Claims Processing Chatbot

---

## 1. Project Overview

The **AI Claims Processing Chatbot** is an intelligent insurance assistant designed to automate policy understanding and claim approval prediction.

This system integrates:

- Policy document compression
- Retrieval-Augmented Generation (RAG)
- Machine Learning-based claim prediction
- Interactive web interface

The objective is to accelerate claim submissions, improve approval accuracy, and reduce operational costs through intelligent automation.

---

## 2. Key Features

### 2.1 Policy Question Answering (RAG-Based)

- Upload insurance policy PDF
- Automatic text extraction and chunking
- Embedding generation using Sentence Transformers
- Storage of vectors in FAISS
- Semantic retrieval of relevant clauses
- Context-aware response generation using LLM

---

### 2.2 Claim Approval Prediction

The system predicts whether a claim is:

- Likely Approved

- Risk of Rejection

**Prediction is based on:**

- Claim amount
- Policy duration (years active)
- Network hospital status

**Model Used:**

RandomForestClassifier (Scikit-learn)

---

## 2.3 Modern Web Interface

- Chat-style user interface
  - Claim probability checker
  - Responsive design
  - Real-time backend communication
- 

## 3. System Architecture & Workflow

### Step 1: Policy Compression

1. Policy PDF is uploaded
2. Text is extracted and split into chunks
3. Embeddings are generated
4. Stored in FAISS vector database

This enables fast semantic search instead of scanning the entire document.

---

### Step 2: Retrieval-Augmented Generation

1. User submits a question
2. Relevant policy chunks are retrieved

- 
3. LLM generates a context-aware answer

---

### Step 3: Machine Learning Prediction

The model is trained using historical claim data.

#### Features:

- claim\_amount
- policy\_years
- hospital\_network

#### Target Variable:

- approved (1 = Approved, 0 = Rejected)

The RandomForest model learns patterns from past claims and predicts outcomes for new claims.

---

## 4. Technology Stack

### Backend

- FastAPI
- Uvicorn

### Machine Learning

- Scikit-learn
- Pandas
- NumPy

### NLP & Vector Search

- Sentence Transformers
- FAISS

## **Frontend**

- HTML
- CSS
- JavaScript

## **LLM Integration**

- OpenAI API (for policy Q&A)
- 

## **5. Business Impact**

- Faster claim processing
  - Reduced manual review costs
  - Improved prediction accuracy
  - Intelligent automation
  - Efficient policy understanding
- 

## **6. Security Recommendations (Production Deployment)**

- Store API keys in environment variables
  - Enable HTTPS
  - Implement authentication
  - Validate all user inputs
  - Mask sensitive data
- 

## **7. Future Enhancements**

- Display approval probability percentage
- Fraud detection scoring

- Explainable AI (feature importance visualization)
  - User authentication system
  - Cloud deployment (AWS / Azure)
  - Analytics dashboard
  - Voice-enabled chatbot
  - Multi-language support
- 

## 8. Author

Developed as an AI and Machine Learning-based insurance automation system.

---

## 9. Project Summary

This project demonstrates:

- Real-world implementation of Retrieval-Augmented Generation (RAG)
- End-to-end Machine Learning pipeline
- Backend and frontend integration
- Vector database usage (FAISS)
- Practical application of Generative AI in the insurance domain

It serves as a strong example of applying Artificial Intelligence to solve real-world business challenges efficiently.