TRAVEL INSURANCE PLAN DETAILS ----------

1. Premium Frequent Traveler Plan - Designed for daily/weekly travelers. - Includes medical emergency coverage. - Trip cancellation & interruption protection. - Lost baggage & passport coverage. - Priority claim processing. - Ideal for business travelers and corporate users.

2. Seasonal Travel Plan - For users who travel 1–3 times per year. - Covers medical emergencies, trip delay, and basic theft. - Budget-friendly option.

3. Business Flyer Plan - Corporate-focused coverage. - Includes laptop + equipment protection. - Flight delay lounge access. - Extra baggage protection.

4. Basic Travel Protection Plan - Entry-level plan for low-frequency travelers. - Covers only essential risks.

TRAVEL INSURANCE CLAIM PROCESS ----------

1. Document Checklist: - Passport copy - Flight tickets - Boarding passes - Medical bills (if applicable) - Police report (for theft/loss) - Proof of travel disruption

2. Steps to File a Claim: Step 1: Notify insurer within 48 hours of incident. Step 2: Submit documents via portal or email. Step 3: Claim gets assigned to an assessor. Step 4: Processing time: 7–14 business days. Step 5: Claim approval or rejection notification sent.

3. Common Claim Rejection Reasons: - Missing documents - Claims filed after deadline - Fraudulent or mismatched information - Pre-existing illness not disclosed

CUSTOMER SEGMENTATION POLICY ----------

Segmentation Objective: Identify travel patterns and cluster users for better insurance pricing and recommendations.

Segments Used:

Segment 0 – Frequent Travelers: - Daily/weekly travelers - Highest exposure risk - Usually business users - Recommended: Premium Frequent Traveler Plan

Segment 1 – Seasonal Travelers: - Only travel during festivals or holidays - Medium exposure - Recommended: Seasonal Plan

Segment 2 – Business Flyers: - Frequent corporate travelers - Carry equipment/laptops - Recommended: Business Flyer Plan

Segment 3 – Low-Frequency Travelers: - Less than 1 trip/quarter - Lowest risk category - Recommended: Basic Travel Plan

Risk Score Usage: - High risk $\rightarrow$ higher premium - Low risk $\rightarrow$ discounts and loyalty rewards

TRAVEL INSURANCE FAQ ----------

Q1: What is travel insurance? A: It protects travelers from unexpected events like delays, medical emergencies, theft, and cancellations.

Q2: Why is segmentation used? A: To group passengers by travel frequency, risk, and loyalty so the insurer can offer better pricing.

Q3: What is covered? A: - Medical emergencies abroad - Trip delays/cancellation - Lost baggage - Passport replacement

Q4: What is NOT covered? A: - Pre-existing diseases - Intentional damage - Illegal travel - War zones

Q5: Who should buy travel insurance? A: - Frequent business travelers - International tourists - Students - Pilgrims - Seasonal vacation travelers

1. Project Overview

This project, "Travel Passenger Segmentation & Loyalty-Based Insurance Recommendation System," is an end-to-end enterprise-level AI, ML, and GenAI solution designed for the Travel Insurance domain.

The objective is to modernize and automate the decision-making pipeline in insurance by integrating Machine Learning for risk modeling and segmentation, as well as a RAG (Retrieval-Augmented Generation) pipeline for document-aware GenAI responses.

This document contains: - BRD (Business Requirements Document) - FRD (Functional Requirements Document) - SRS (Software Requirements Specification) - Segmentation System Design - Risk Prediction System Design - Recommendation System Architecture - RAG Pipeline (GenAI) Architecture - Solution Architecture (End-to-End) - Data Dictionary - Technical Components - APIs - Deployment Pipeline

2. Business Requirements Document (BRD)

BUSINESS REQUIREMENTS DOCUMENT (BRD)

Business Problem: The travel insurance company lacks personalization. All customers receive generic plans despite varying travel frequency patterns and risk profiles. This causes revenue loss, incorrect pricing, claim mismanagement, and reduced customer satisfaction.

Business Goals: 1. Reduce claim losses by identifying high-risk travelers early. 2. Improve customer satisfaction using personalized insurance offerings. 3. Increase policy conversion rate via AI-driven recommendations. 4. Lower operational costs using automation and AI support systems. 5. Enhance risk exposure visibility for underwriting teams.

KPIs: - Increase policy conversion by 12%. - Reduce loss ratio by 8%. - Improve customer retention by 15%. - Automate 40–60% of customer support via GenAI.

3. Functional Requirements Document (FRD)

FUNCTIONAL REQUIREMENTS DOCUMENT (FRD)

Modules: 1. Data Ingestion Module - Reads raw CSV data. - Validates schema. - Stores into raw zone.

2. Preprocessing Module - Cleans missing values. - Removes duplicates. - Standardizes travel frequency.

3. Feature Engineering - Generates RFM features. - Builds frequency_score, value_score, customer_score.

4. Segmentation Module - Uses KMeans clustering. - Outputs 4 traveler segments.

5. Risk Prediction Model - XGBoost regression. - Predicts risk_score.

6. Recommendation Engine - Hybrid rule-based + ML similarity. - Recommends plan per segment.

7. RAG GenAI Pipeline - Loads policy documents. - Generates embeddings. - Uses Pinecone vector DB. - Retrieves relevant context. - LLM produces insurance-aware answers.

8. API Layer - FastAPI endpoints for each model.

9. Deployment - Dockerized services. - Kubernetes for scaling.

4. Software Requirement Specification (SRS)

SOFTWARE REQUIREMENT SPECIFICATION (SRS)

System Features: 1. Segmentation Engine - Input: Cleaned features - Output: Segment (0–3)

2. Risk Model - Input: Segment + other features - Output: Risk score (0–100)

3. Recommendation Engine - Input: Segment - Output: Plan name

4. RAG Assistant - Input: User query - Output: Policy-aware response

Non-Functional Requirements: - Latency < 300 ms for APIs - Horizontal scaling using containers - Data drift monitoring every 24 hours - Robust authentication (JWT/Token)

Hardware Requirements: - 8 CPU, 16GB RAM recommended for training

5. Segmentation System

SEGMENTATION SYSTEM DESIGN

Segmentation Goal: Group customers based on travel behavior and loyalty.

Algorithm Used: - KMeans Clustering - 4 clusters empirically selected

Features Used: - travel_frequency - loyalty_score - past_claims - engineered customer_score

Cluster Definitions: Cluster 0: Premium Frequent Flyers Cluster 1: Seasonal Travelers Cluster 2: Business Travelers Cluster 3: Low-Frequency Travelers

Outputs: - df["segment"]

6. Risk Prediction Model

RISK PREDICTION SYSTEM

Purpose: Predict customer's risk to help underwriting team assign correct premium.

Model: - XGBoost Regressor

Target Variable: - risk_score (continuous)

Features: - Age - Travel Frequency - Loyalty Score - Past Claims - Segment

Model Metrics: - RMSE - MAE - R^2

Usage: risk.predict([[age, travel_frequency, loyalty_score, segment]])

7. Recommendation Engine

RECOMMENDATION SYSTEM

Objective: Recommend the best insurance plan based on the segmentation output.

Method: - Rule-based mapping: Segment 0 → Premium Frequent Traveler Plan Segment 1 → Seasonal Travel Plan Segment 2 → Business Flyer Plan Segment 3 → Basic Travel Protection

Future Enhancements: - Collaborative filtering - Deep learning embeddings - Next-best-offer modeling

8. RAG Pipeline Architecture

RAG PIPELINE (GENAI SYSTEM)

RAG Components: 1. Document Loader 2. Text Chunker 3. Embedding Generator (OpenAI) 4. Vector Store (Pinecone) 5. Retriever 6. LLM Answer Generator

Workflow: 1. Load documents from policy_docs/ 2. Chunk documents into 300–500 token text blocks 3. Generate embeddings using text-embedding-3-large 4. Store embeddings in Pinecone 5. Query embeddings for relevant context 6. LLM generates final insurance-safe reply

Use Cases: - Explain recommended plan - Explain risk score - Compare plans - Answer FAQs

9. Solution Architecture

SOLUTION ARCHITECTURE SUMMARY

DATA → PIPELINE → FEATURE STORE → ML MODELS → RAG PIPELINE → FASTAPI → DOCKER → KUBERNETES

Key Systems: - Data Engineering Layer - ML Layer (Segmentation + Risk + Recommendation) - GenAI Layer (RAG) - API Layer - Model Registry (MLflow) - Monitoring Layer - CI/CD Deployment Layer

API Endpoints: - /segment-customer - /predict-risk - /recommend-plan - /genai-advisor