Instant Personalized Quote Generator Using ML & LLM

Project Overview

The Instant Personalized Quote Generator is a fully automated system that provides instant, personalized insurance quotes to customers. By combining Machine Learning (ML) and Large Language Models (LLM), this solution reduces human errors, eliminates long wait times, and improves customer engagement.

Key Features

- **Risk Prediction:** Uses historical customer data to predict insurance risk.
- **Premium Calculation:** Automatically calculates personalized insurance premiums.
- **LLM Explanations:** Generates customer-friendly textual explanations for quotes.
- **End-to-End Automation:** Input ➜ Risk ➜ Premium ➜ Explanation.
- **Scalable:** Can handle thousands of customer requests with minimal manual intervention.

Use Case

Use Case	Customer Problem	Solution
Instant Personalized Quote	Customers find traditional insurance quoting slow and generic	Automatically predict risk, calculate premiums, and generate an easy-to-understand explanation using ML and LLM

Dataset

Sample Columns: `customer_id, age, vehicle_type, location, claim_history, base_rate, risk_multiplier`

Features Used for ML: - `age` - `vehicle_type` - `location` - `claim_history`

Target Variable: - `risk_level` (`low` , `medium` , `high`)

Project Pipeline

1. **Load and Preprocess Data:**
2. Load CSV dataset with pandas.

3. Encode categorical variables (`vehicle_type` , `location`) using LabelEncoder.

4. **ML Model Training:**

5. Train Random Forest Classifier to predict `risk_level` .
6. Features: `age` , `vehicle_type` , `location` , `claim_history` .

7. Evaluate model using accuracy and classification report.

8. **Premium Calculation:**

9. Formula: `Premium = base_rate * risk_multiplier * coverage_factor`

10. Coverage factors: Liability=1, Collision=1.2, Comprehensive=1.5.

11. **LLM Explanation:**

12. Integrated Falcon 7B Instruct LLM.

13. Generates textual explanations for predicted premiums in plain language.

14. **Full Pipeline:**

15. Input new customer data ➜ Predict risk ➜ Calculate premium ➜ Generate explanation.
16. Fully automated workflow for instant quotes.

Sample Output

Input: - Age: 40 - Vehicle Type: Car - Location: Urban - Claim History: 2 - Base Rate: 5500 - Risk Multiplier: 1.4

Output: - Predicted Risk: Medium - Final Premium: ₹7700 - LLM Explanation: "The insurance premium for this customer is ₹7700 because their age, claim history, and location classify them as medium risk. This ensures they are covered adequately while keeping the premium fair."

Tools & Libraries

- Python 3.12
- Pandas: Data processing
- Scikit-learn: Random Forest ML model
- Transformers (HuggingFace): Falcon LLM for explanations
- Joblib: Save/load ML models

How to Run

1. Load dataset: `historical_applications.csv`
2. Train the Random Forest model or load the saved model.
3. Run the full quote pipeline for new customer input.
4. Generate risk, premium, and explanation automatically.

Key Takeaways

- **Data-Driven:** ML predicts risk from historical applications.
- **Customer-Friendly:** LLM explains premiums in plain language.
- **Automated & Scalable:** Handles thousands of requests instantly.
- **Business Impact:** Increases conversion by providing personalized quotes.