**Chronic Disease Risk Prediction and Product Recommendation System**

**Introduction**

This project aims to develop a comprehensive system for chronic disease risk prediction and personalized insurance product recommendations. It combines a well-structured relational database, machine learning algorithms, and data-driven logic to address key business needs in the insurance domain.

**Business Usage**

The primary goal of the system is to enhance the customer experience and streamline product recommendations for clients. By analyzing customer health data, the system identifies risk levels and tailors product offerings to meet individual needs. This ensures:

* Personalized Recommendations: Clients receive products aligned with their health risk levels.
* Data-Driven Insights: The system provides actionable insights to insurance providers for better product positioning.
* Operational Efficiency: Automating data processing, risk prediction, and product recommendations reduces manual effort and improves decision-making.

**System Design**

Database Schema:

* The schema consists of normalized tables for customer information, health metrics, chronic disease history, and products.
* The Product table includes a risklevel column, which is mapped to predicted risk levels to recommend suitable products.

Data Generation:

* Randomized data generation simulates real-world scenarios:
* Clients are categorized into diagnosed, undiagnosed, and undetermined groups.
* Health metrics and product details are randomly generated for testing and demonstration.

Machine Learning Logic:

* A Random Forest model predicts the chronic disease risk of undetermined clients.
* Risk levels are scaled from 0-100 to 0-5 for easier mapping to insurance products.

Recommendation Logic:

* The system fetches a random undetermined client using the get\_random\_undetermined\_client function.
* Based on the client’s risk level, a product is recommended via the recommend\_products function.
* Personalized messages are generated, including the client's name and the recommended product.

**Optimization Process**

Database Refinements:

* Introduced risklevel in the Product table to align product offerings with client risk profiles.
* Optimized queries by adding indices and ensuring proper normalization.

Machine Learning Model:

* Tuned the Random Forest model with appropriate hyperparameters for better accuracy.
* Ensured the training dataset excluded undetermined clients to avoid noisy data.

Integration Improvements:

* Consolidated database and ML logic to maintain modularity and clear separation of concerns.
* Enhanced recommend\_products to dynamically fetch risk levels and generate human-readable outputs.

Error Handling and Debugging:

* Addressed edge cases, such as missing data or undetermined clients, to ensure smooth execution.
* Refined the product recommendation function to avoid duplicate outputs and align with business logic.

**Logic Summary**

Risk Prediction:

* Health metrics from HealthMetrics are used to predict chronic disease risk for undetermined clients.
* Predicted risk levels are stored in the ChronicDiseaseRisk table.

Product Recommendation:

* The risklevel in Product is mapped to a client’s predicted risk level to recommend an appropriate product.
* A personalized message is generated: "Hello [Name]! The recommendation for you is [Product] under the plan [PlanName]."

Business Value:

* High-risk clients are directed toward comprehensive plans, while low-risk clients are offered cost-effective options.
* Undetermined clients are engaged through proactive product suggestions, enhancing client retention.

**Conclusion**

This system demonstrates a robust solution for integrating data analysis and machine learning into insurance operations. By leveraging health metrics, it predicts chronic disease risks and aligns product offerings to individual client needs. The project highlights the importance of optimization and logic refinement in building scalable, impactful systems. Future improvements could include implementing a user-friendly interface and expanding the ML model to include additional health metrics for more accurate predictions.

**Example Usage:**

A screen shot of a computer

Description automatically generated