Author: Denghao Sun

Date: December 22, 2023

NetID: ds6963

Section: 001

**Database Systems Project Part IV**

**End-to-End Solution Integration and Data-Driven / Database Programming**

**Affirmation of our independent efforts: \_\_\_\_Denghao Sun\_\_\_\_\_\_\_\_\_\_\_**

**GitHub Repository of The Project:** [**https://github.com/SimonSun22/db\_final\_project.git**](https://github.com/SimonSun22/db_final_project.git)

**Project Overview**

This project introduces an innovative web application that provides instant health insurance quotes. Its ability to retain customer data for ongoing engagement and subsequent consultations stands out. Using a well-trained ML model, the application uniquely predicts health insurance costs based on personal user data like age and residence. This functionality showcases the integration of machine learning in providing real-time, personalized financial services.

**Key Features:**

Instant Health Insurance Quotes: The application provides immediate, accurate insurance estimations by leveraging a state-of-the-art machine learning model.

Information Retention: Customers have the option to save their personal details for future interactions and queries securely.

**Technological Stack and Methodologies:**

**Application Architecture**

Structured into three main components: server, frontend, and database, the application's architecture is designed for efficiency and scalability.

Backend Server: Developed with Python and Django, it establishes a secure and responsive API layer.

Database Management: Using MySQL on AWS RDS, the application enjoys a robust, cloud-based database solution.

Frontend Development: The combination of HTML, CSS, and JavaScript ensures an engaging and intuitive user interface.

**Backend: Python with Django Framework**

Python: Renowned for its simplicity and efficiency, it is a versatile programming language for backend development. Its extensive libraries and frameworks make it a preferred choice for web applications.

Django: A high-level Python web framework, Django is chosen for its robustness and scalability. It excels in rapid development and pragmatic design, making it perfect for building secure and maintainable web applications.

**Frontend: JavaScript and HTML**

JavaScript: As the backbone of web development, JavaScript ensures interactive and dynamic user interfaces. Its universal compatibility with web browsers makes it indispensable for front-end development.

HTML: The standard markup language for creating web pages, HTML is used for structuring and presenting content on the Internet, providing the foundational layout for the application.

**Machine Learning Model:**

Linear Regression Model: This model is a fundamental statistical and machine-learning technique used to predict numerical values. It’s chosen for its efficiency in forecasting insurance costs based on user input, exemplifying the application of ML in practical, real-world scenarios.

**Database Interaction: Django’s ORM and MySQL**

Django’s ORM (Object Relational Mapping): This feature of Django simplifies database interactions by converting data between incompatible systems, streamlining the process of database schema design.

MySQL Database on AWS RDS: MySQL offers a reliable and scalable database solution. When hosted on AWS RDS, it ensures enhanced performance, backup capabilities, and scalability, which are crucial for handling potentially large volumes of customer data.

**Database Design**

Type: A cloud-hosted relational database on AWS RDS, chosen for its robustness and scalability.

Schema: Designed for versatility, it efficiently handles structured and potential future unstructured data, with indexes to enhance query performance.

ML Model Integration: Trained on a comprehensive dataset, the model demonstrates an accuracy above 73.4%, validating its effectiveness in insurance quote estimation.

**Reference Architecture (RA) Documentation**

**Foundational Principles**

Business Principles

* Ensuring robust operations and minimal system interruptions.
* Adherence to information management regulations.
* Promoting shared understanding and decision-making processes.
* Maximizing enterprise benefits through strategic technology and data choices.

Data Principles

* Recognizing and carefully managing data is a crucial asset.
* Facilitating secure data sharing to foster collaboration.
* Ensuring high data quality and integrity with designated trustees.
* Uniform data definition and documentation practices.
* Rigorous protection against unauthorized data access.

Application and Technology Principles

* Prioritizing clear documentation and convenience for end-users.
* Focusing on application portability and middleware decoupling.
* Aligning technologies for optimal efficiency and ease of use.
* Ensuring compatibility across different software applications.

**Organizing Framework**

* Harnessing structured, semi-structured, and unstructured data with a focus on providing actionable business insights through data analysis and machine learning models.
* Automating processes and delivering functionalities crucial for insurance quote generation and user data management.
* Collaborating with various external and internal applications and integrating with different system partners and software vendors.

**Business Solutions**

* Creating solutions with clear, defined goals to address user needs for insurance quotes.
* Employing iterative development with feedback loops for continuous improvement.
* Designing applications to be scalable, modular, and efficient.
* Ensuring applications meet regulatory requirements and pass comprehensive testing.

**Data Governance**

* Standardizing API use for efficient, secure customer data handling, with access controlled by designated authorities.
* Implementing strong encryption and privacy measures for sensitive customer data.
* Regular assessment and improvement of data quality, especially for customer-provided information.