

**Project Design Phase-I**  
**Solution Architecture**

Date	27 October 2023
Team ID	PNT2022TMID591995
Project Name	Project - Car Purchase Prediction Using ML
Maximum Marks	4 Marks

**Solution Architecture:**

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered.

**Components:**

User Interface (UI):

Allows users to input information for prediction.

Displays prediction results.

Frontend:

Handles user interactions and communicates with the backend.

Sends user inputs to the backend for processing.

### Backend:

Receives input from the frontend.

Processes the input and prepares it for the machine learning model.

Interfaces with the machine learning model for prediction.

Sends the prediction results back to the frontend.

### Machine Learning Model:

Trained model for predicting the likelihood of a user purchasing a car.

Takes features such as customer demographics, browsing history, and financial information as input.

Outputs a prediction score or probability.

### Data Storage:

Stores historical data used for training and validating the machine learning model.

May include databases or data warehouses.

### APIs:

Connects different components of the system.

Exposes endpoints for communication between the frontend, backend, and machine learning model.

### Data Preprocessing:

Cleans and preprocesses incoming data before it is fed into the machine learning model.

Handles missing values, scales features, and ensures data quality.

### Training Pipeline:

In a separate environment, this pipeline periodically retrains the machine learning model using new data.

The updated model is then deployed to the production environment.

#### Monitoring and Logging:

Monitors the system for anomalies, performance issues, and model degradation.

Logs activities for debugging, auditing, and performance analysis.

#### Authentication and Authorization:

Ensures that only authorized users can access the system.

Manages user authentication and permissions.

#### External Data Sources:

Integrates with external sources for additional data, such as market trends, economic indicators, or industry news.

#### Scalability and Load Balancing:

Ensures the system can handle varying loads and scales horizontally if needed.

Distributes incoming requests evenly to different instances of the application.

#### Security:

Implements security measures to protect sensitive user data and the integrity of the system.

#### **Flow:**

1. The user interacts with the UI to provide input.
2. The frontend sends the input to the backend through APIs.
3. The backend preprocesses the data, sends it to the machine learning model, and receives the prediction.
4. The prediction is sent back to the frontend and displayed to the user.
5. The system periodically retrains the machine learning model using new data.

#### **Example - Solution Architecture Diagram:**

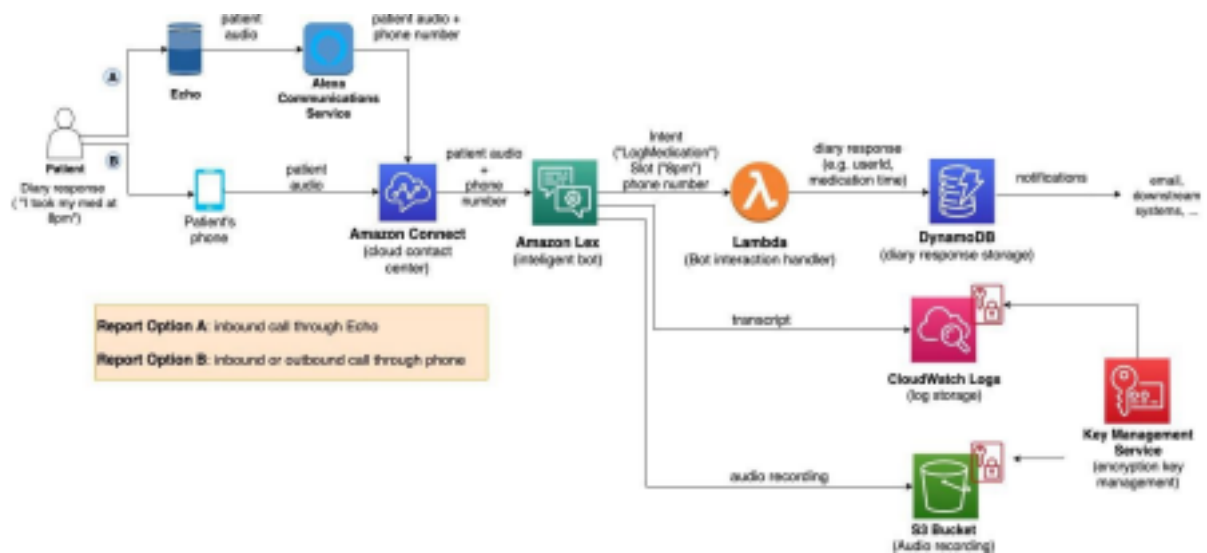


Figure 1: Architecture and data flow of the voice patient diary sample application

Reference: <https://aws.amazon.com/blogs/industries/voice-applications-in-clinical-research-powered-by-ai-on-aws-part-1-architecture-and-design-considerations/>