

A Strategic Approach to customer retention/Project: Al X

July 2025

Version 1.0

# **1 Document Revisions**

| Date       | Version<br>Number | Document Changes |
|------------|-------------------|------------------|
| 02/07/2025 | 1.0               | Final edit       |
|            |                   |                  |
|            |                   |                  |
|            |                   |                  |
|            |                   |                  |
|            |                   |                  |
|            |                   |                  |
|            |                   |                  |
|            |                   |                  |

# 2 Approvals

| Role            | Name            | Title                         | Signature | Date |
|-----------------|-----------------|-------------------------------|-----------|------|
| Project Sponsor | Quỳnh Trang     | Phê duyệt dự án               |           |      |
| Product Owner   | Lê Văn Túc      | Ra chiến lược cho<br>dự án    |           |      |
| Project Manager | Dương Minh Dũng | Quản lý và thực<br>hiện dự án |           |      |
| Client          | Nguyễn Văn A    | Người dùng cuối               |           |      |
|                 |                 |                               |           |      |
|                 |                 |                               |           |      |
|                 |                 |                               |           |      |
|                 |                 |                               |           |      |

# 3 Introduction

## 3.1 Project Summary

#### 3.1.1 Objectives

• Our current customer retention efforts are reactive and inefficient. We lack a centralized, data-driven way to identify which customers are likely to churn, why they are at risk, and what the financial impact is. This leads to wasted marketing spend, low team productivity, and preventable revenue loss.

#### 3.1.2 Background

Understanding the current demand for customer retention, we developed a solution that allows businesses to easily monitor customer purchasing behavior, including purchase rates, customer feedback, and customer segmentation. Building on this, we also provide an integrated AI solution for businesses to predict customer churn rates with greater precision and visual clarity.

#### 3.1.3 Business Drivers

#### Financial Drivers:

- Increased Revenue/Sales:
  - **Enhanced Customer Engagement:** A more intelligent chatbot can lead to better product recommendations and a smoother shopping experience, driving higher conversion rates.
  - **Upselling/Cross-selling Opportunities:** Al can identify opportunities to suggest complementary or higher-value products based on customer behavior and preferences.
- Cost Reduction:
  - Reduced Customer Service Costs: Automating routine inquiries and common support issues through the chatbot significantly lowers the need for human agents, leading to substantial savings in labor costs.
  - **Optimized Resource Allocation:** Freeing up human agents to focus on complex issues rather than repetitive tasks.
- Improved ROI (Return on Investment):
  - Directly measurable impact on sales growth and cost savings contributes to a strong ROI for the technology investment.

## 3.2 Project Scope

This feature is the core of the Minimum Viable Product (MVP). Its purpose is to validate our core hypothesis: that we can use an AI model to accurately identify high-risk customers and present this information in a way that is valuable and actionable for the Marketing and Customer Success teams.

#### 3.2.1 In Scope Functionality

- A single, read-only dashboard view.
- User login, log out, sign up
- KPI cards for High-Risk Count and Revenue At-Risk.
- A prioritized "Actionable Customer List" for the High-Risk segment only.
- Static, rule-based recommendations (e.g., if driver is X, suggest Y).
- Manual data refresh via CSV upload or a database table update by the technical team.
- Basic user login/authentication.

### 3.2.2 Out of Scope Functionality

- Medium and Low-Risk customer segments.
- Live API integrations for actions (e.g., no "Take Action" buttons).
- Automated MLOps pipelines or self-service model training.
- Advanced filtering or sorting beyond basic column sorting.
- Performance tracking measurement dashboards.

### 3.2.3 Assumptions

- Validate the core hypothesis:
  - -> Can our predictive model accurately identify high-risk customers?
  - -> Will the marketing team find the list valuable enough to take action?

#### **3.2.4** Risks

#### **Technical Risk:**

#### • Inaccurate Model Predictions:

Al models may misidentify at-risk customers, resulting in misleading or low-value outputs.

# **Adoption Risk:**

#### Low User Adoption:

Marketing and CS teams may distrust AI recommendations or find the platform hard to use.

# **Project Risk:**

#### Scope Creep & Delays :

The project may grow beyond MVP scope, risking timelines and budget.

## 3.3 Project Timeline

• Phase 1: Complete The AI Engine by 25/06/2025 - 28/06/2025

#### A. Data Sources

- CRM (Salesforce, HubSpot),
- Product usage logs,
- Support tickets,
- Billing/subscription system,
- Campaign history.

### **B. Data Pipeline & Storage**

- ETL tools (Airflow, dbt, or custom scripts).
- Data lake / warehouse (BigQuery, Snowflake, Redshift).
- Feature Store (e.g. Feast, Vertex AI Feature Store).

#### C. Al Model Layer

- Trained churn prediction model (e.g. XGBoost, LightGBM, Neural Network).
- **Input:** engineered features.
- Output: churn probability, risk level, drivers.

## **D.** Recommendation Engine

- Uses LLMs to explain "why this user may churn".
- Suggests actions based on churn drivers + persona profile.

## Phase 2: Complete The action platform by 29/06/2025 - 1/07/2025

## **E. Action Trigger Engine**

- Logic to map:
- Risk level → Playbook
- Suggested action → Delivery channel
- Rules-based or ML-based
- Could be custom microservice, Zapier, or orchestration tool

# F. Engagement Tools

- Email/SMS tools (e.g. Mailchimp, Twilio, Braze)
- CRM (Salesforce, Zoho, etc.)
- In-app messaging SDK

- Customer Success tools (e.g. Gainsight, Intercom)
- Phase 3: Complete The Learning Loop by date 02/07/2025 -0 4/07/2025
   G. Monitoring + Feedback
  - Track if user responded, stayed, or churned
  - Dashboards (e.g. Looker, Power BI)
  - Model retraining schedule
  - Business KPIs (e.g. retention uplift, ROI)

# 3.4 Key Stakeholders (Optional)

[Identify key stakeholders and outline their roles and responsibilities.]

- Project Manager: responsible for holding all parties accountable to the project timeline
- Department Heads: share desired needs with BA for a comprehensive list of requirements
- Product owner: Define the long-term vision for the product, what it will look like, who it will serve, and what problem it will solve.
- Business analyst:Requirements Elicitation & Analysis, Requirements Documentation, Testing & Validation Đặc biệt là UAT
- Tester: A tester is not a "bug catcher" but a comprehensive quality assurance person who helps shape the product, detects problems early, and contributes significantly to project success and end-user satisfaction.
- Developer: Through the requirements document from the BA, the developers implement the final product code
- End User: include: Marketing manager, customers service. The customer is the person who uses the product and implements the features that the user has requested in advance.

# Business Requirements

[The specific business requirements elicited from stakeholders should be listed, categorized by both priority and area of functionality to smooth the process of reading and tracking them. Include links to use case documentation, and other key reference material as needed to make the requirements as complete and understandable as possible. You may wish to incorporate the functional and non-functional requirements into a traceability matrix that can be followed throughout the project.]

| Req. ID | Requirement  | Business Value   |
|---------|--|--|
| BR-01   | Provide a high-level summary of overall customer churn risk.   | Delivers an immediate, quantitative snapshot of churn exposure to guide strategic attention.             |
| BR-02   | Surface the key drivers contributing to customer churn.        | Equips teams with insight into root causes, enabling more targeted and effective retention initiatives.  |
| BR-03   | Display a prioritized list of customers at highest risk.       | Enables focused outreach by concentrating efforts on the most vulnerable and valuable accounts.          |
| BR-04   | Enable export of the high-risk customer list for external use. | Supports integration with existing marketing and CRM workflows, facilitating timely follow-up actions.   |
| BR-05   | Estimate the potential revenue at risk due to churn.           | Makes churn financially tangible, helping stakeholders evaluate impact and prioritize retention efforts. |

The requirements in this document are prioritized as follows:

| Value | Rating   | Description  |
|-------|----------|--|
| 1     | Critical | This requirement is critical to the success of the project. The project will not be possible without this requirement. |
| 2     | High     | This requirement is high priority, but the project can be implemented at a bare minimum without this requirement.      |
| 3     | Medium   | This requirement is somewhat important, as it provides some value but the project can proceed without it.              |
| 4     | Low      | This is a low priority requirement, or a "nice to have" feature, if time and cost allow it.                            |
| 5     | Future   | This requirement is out of scope for this project, and has been included here for a possible future release.           |

4.1 Functional Requirements

| Req#          | Priority                 | Description  | Use Case Reference (Optional) | Impacted Stakeholders |  |  |
|---------------|--------------------------|--|-------------------------------|-----------------------|--|--|
| The Predictiv | The Predictive Dashboard |  |                               |                       |  |  |
| FR-001        | 1                        | The main landing page after a user logs in. Provides a high-level, aggregated view of customer risk.  UI Components & Behavior:  KPI Cards (Top Section):  High-Risk Customers: Displays a single integer count.  (Data: count_high_risk_customers)  Revenue At-Risk: Displays a single currency value. (Data: sum_revenue_at_risk_high_risk)  Top Churn Driving Factors (Chart Section):  Type: A simple bar chart.  Behavior: Displays the top 4-5 churn drivers ranked by the number of customers affected. Each bar represents a driver (e.g., "Purchase Recency," "High Support |                               | Marketing manager     |  |  |

|              |                   | Interaction"). The height of the bar corresponds to the percentage of high-risk customers exhibiting that driver. (Data: Aggregated count of primary churn drivers for all high-risk customers).  High-Risk Customer Action List (Preview Table): Behavior: Displays a preview of the top 3-5 customers from the full Actionable List. Columns should match the full list.  Includes a prominent button/link titled "View Full Action List".   |                  |
|--------------|-------------------|--|------------------|
| FR-002       |                   |  | _                |
| The Actional | ble Customer List |  |                  |
| FR-003       | 1                 | A detailed, tabular view of all customers in the "High-Risk" segment. Reached by clicking "View Full Action List" from the dashboard.  UI Components & Behavior: Page Title: "High-Risk Customer Action List". Action Button: An "Export to CSV" button must be present. Logic: This button is disabled by default. It becomes enabled only when at least one customer row is selected via the checkbox. Customer Table: A sortable table containing all customers with a churn probability above the "High-Risk" threshold. Columns: Checkbox: Allows for multi-select. Customer Name / ID: Text. | Customer service |

|        | Churn Probability: Percentage. Should be color-coded. Top Churn Driver: Text (e.g., "Low Product Usage"). Action taken: Action already taken. Recommendation Action:Recommended action suggested by AI. Sorting: The table must be sortable by each column header (ascending/descending). The default sort order is by Churn Probability (descending). |  |
|--------|--|--|
| FR-004 |  |  |

# **4.2** Non-Functional Requirements

[Include technical and operational requirements that are not specific to a function. This typically includes requirements such as processing time, concurrent users, availability, etc.]

| ID      | Requirement   |
|---------|---|
| NFR-001 | Performance: The dashboard page must load in under 3 seconds. The Actionable List must load in under 5 seconds for up to 1,000 records.                                     |
| NFR-002 | Security: Access to all pages must be restricted to authenticated users.  |
| NFR-003 | Usability: The interface must be clean and intuitive, requiring minimal training for Maria and David.   |
| NFR-004 | The web interface must display and function correctly on the latest 2 versions of the following browsers: Google Chrome, Mozilla Firefox, Apple Safari, and Microsoft Edge. |
| NFR-005 | Responsive design: The interface should automatically adjust to display well on popular screen sizes: Desktop (1920x1080), Tablet (768x1024), and Mobile (375x667).         |
| NFR-006 | Personal Data Protection: The system must comply with Vietnamese regulations on personal data protection, especially Decree 13/2023/ND-CP.                                  |
| NFR-007 | 99% of core API calls must complete in under 500ms.   |

# Appendices

# **5.1** List of Acronyms

| ID | Technical terms | Meaning  |
|----|-----------------|--|
| 1  | ROI             | Stands for Application Programming Interface. It is a set of rules and tools that allows different software applications to communicate with each other. It acts as an intermediary, enabling one program to request services or data from another.        |
| 2  | API             | Stands for Machine Learning Operations. It's a set of practices for deploying and maintaining machine learning (ML) models in production reliably and efficiently. It combines ML with DevOps to automate the model lifecycle                              |
| 3  | MLops           | Stands for Customer Relationship<br>Management. It refers to a system<br>or technology for managing all of a<br>company's relationships and<br>interactions with its current and<br>potential customers. Examples:<br>Salesforce, HubSpot.                 |
| 4  | CRM             | Stands for Customer Relationship<br>Management. It refers to a system<br>or technology for managing all of a<br>company's relationships and<br>interactions with its current and<br>potential customers. Examples:<br>Salesforce, HubSpot.                 |
| 5  | ETL             | Stands for Extract, Transform, Load. It is a data integration process with three steps: Extracting data from various sources, Transforming the data (e.g., cleaning, reformatting), and Loading the processed data into a destination system (often a Data |

|   |                      | Warehouse).   |
|---|----------------------|---|
| 6 | Data lake/ warehouse | Data Warehouse: A central repository that stores structured, processed data to support business intelligence (BI) and reporting.<br>br>- Data Lake: A storage repository that holds a vast amount of raw, unstructured data in its native format. It offers more flexibility but requires further processing before analysis. |

## 5.2 Glossary of Terms

[If needed, identify and define any terms that may be unfamiliar to readers, including terms that are unique to the organization, the technology to be employed, or the standards in use.]

#### **System Design & Architecture Concepts**

**Use Case Diagram:** A visual diagram that shows the different ways users (called "actors" like Customer, Admin) can interact with the system. It helps to clarify the system's intended functions from a user's perspective (e.g., a "Customer" can "Place Order").

**Three-Tier Architecture:** An established software architecture pattern that separates the application into three logical and physical computing tiers:

**Presentation Tier (Frontend):** The user interface that customers and restaurant owners see and interact with in their web browser (built with React.js).

**Logic/Application Tier (Backend):** The "brain" of the system. It processes business logic, handles user requests, and communicates with the database (built with Node.is/Express.is).

**Data Tier (Database):** Where all the information (user data, menus, orders) is stored and managed (using MongoDB).

**RESTful API:** An architectural style for designing networked applications. In this project, it acts as the "contract" or communication channel between the Frontend (the user's app) and the Backend (the server). For example, when you browse a menu, the Frontend sends a "GET menu" request to the Backend's API.

**ERD (Entity-Relationship Diagram):** A type of flowchart that illustrates how "entities" such as Users, Restaurants, Orders, and MenuItems relate to each other within the database. It is the blueprint for the database design.

### Technologies to be Employed (The "Tech Stack")

**React.js (Frontend)**: A popular JavaScript library for building fast, modern, and interactive user interfaces. It will be used to create the web application that users see.

**Node.js (Backend):** A JavaScript runtime environment that allows developers to write the server-side (backend) code in JavaScript. It's the foundation upon which the application's logic is built.

**Express.js (Backend):** A framework that runs on top of Node.js. It simplifies the process of building the backend and creating the RESTful API, providing tools for handling web traffic and requests.

**MongoDB (Database)**: A type of NoSQL database that stores data in flexible, JSON-like documents. This is different from traditional SQL databases (like tables with rows and columns) and is well-suited for applications where the data structure might evolve.

**Docker (Deployment):** A platform for creating "containers." A container is a package that bundles the application's code with all the libraries and dependencies it needs to run. This ensures the application runs consistently and reliably, whether on a developer's laptop or on the production servers.

**AWS (Amazon Web Services):** A cloud computing platform provided by Amazon. It will be used to host the entire application (the backend, database, and frontend files), making it accessible to users over the internet.

**Stripe (Payment Gateway)**: A third-party service that securely processes online payments. By integrating Stripe, the food ordering system can accept credit/debit card payments without having to handle and store sensitive financial information itself, which greatly improves security.