**Business Requirements Document (BRD) 1**

**PROJECT NAME:** Bank Customer Churn Analysis and Prediction 2

COMPANY NAME: [Bank Name - e.g., Global Financial Bank] 3

Street Address: [Bank Street Address] 4

City, State and Zip: Thane, Maharashtra, India 5

webaddress.com: [Bank Website Address] 6

**VERSION:** 1.0.0 7

**DATE:** 06/21/2025 8

**VERSION HISTORY 9**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **VERSION** | **APPROVED BY** | **REVISION DATE** | **DESCRIPTION OF CHANGE** | **AUTHOR** |
| 1.0.0 | Mr. George Finos | 06/21/2025 | Initial Draft | Kshitij D Taware |

**TABLE OF CONTENTS 10**

1. EXECUTIVE SUMMARY SNAPSHOT
2. PROJECT DESCRIPTION 12
3. PROJECT SCOPE 13
4. BUSINESS DRIVERS 14
5. CURRENT PROCESS 15
6. PROPOSED PROCESS 16
7. FUNCTIONAL REQUIREMENTS 17
   * PRIORITY 18
   * REQUIREMENTS CATEGORIES (RC1) 19
8. NON-FUNCTIONAL REQUIREMENTS 20
9. FINANCIAL STATEMENTS 21
10. COST AND BENEFIT 22
11. RESOURCES 23
12. SCHEDULE, TIMELINE, AND DEADLINES 24
13. ASSUMPTIONS 25
14. GLOSSARY 26
15. REFERENCES 27
16. APPENDIX 28

**1. EXECUTIVE SUMMARY SNAPSHOT 29**

This document outlines the requirements for the "Bank Customer Churn Analysis and Prediction" project. 30 The project aims to reduce customer churn, which is currently impacting the bank's profitability and growth, by identifying at-risk customers proactively. 31 This initiative will leverage historical customer data and machine learning to predict churn likelihood and provide actionable insights for targeted retention strategies. 32

* Research has shown that customer acquisition costs are significantly higher than customer retention costs, indicating that reducing churn is crucial for sustainable growth and profitability.
* The scope of the proposed project is to analyze historical data, develop a predictive churn model, and generate prioritized lists of at-risk customers for proactive intervention. 33
* The main business drivers behind this effort are to reduce the overall customer churn rate, improve the accuracy of retention efforts, and gain a deeper understanding of customer behavior. 34
* While the current process is reactive, identifying churn only after it occurs, the proposed process will enable proactive identification and targeted engagement. 35
* The functional requirements in this document indicate that the project will involve data integration, feature engineering, model development, and the creation of a reporting dashboard for insights and predictions. 36
* Etc. 37

**2. PROJECT DESCRIPTION 38**

The primary goal of this project is to develop a robust system capable of identifying bank customers who are at a high risk of churning (closing accounts or significantly reducing engagement). 39 By undertaking this project, we aim to reduce our customer churn rate, which currently impacts our profitability and growth, and thereby ensure that we meet our goals of increased customer lifetime value and optimized marketing resource allocation. 40

* While the current process involves reactive identification of churn post-occurrence through manual reports, the proposed process will enable proactive identification through predictive modeling. 41
* The goal of reducing customer churn by 5% and improving retention accuracy by 20% presents challenges related to integrating disparate data sources, defining a precise "churn" metric, and developing an accurate predictive model. 42
* By undertaking this project, we will ensure that we gain a deeper understanding of customer behavior, optimize our retention strategies, and ultimately improve the bank's overall financial health and customer loyalty. 43
* Etc. 44

**3. PROJECT SCOPE 45**

This section provides a high-level description of the project's scope, delineating what is "in" and "out" of scope to guide team members for planning and resourcing. 46

**IN SCOPE** 47

The following are "in scope" for the project: 48

* **Data Sourcing and Integration:** Conduct data extraction, cleaning, and integration from core banking systems, CRM, and online banking usage data to create a unified customer dataset.
* **Feature Engineering:** Develop new, relevant features from raw data (e.g., average monthly balance, transaction frequency, number of complaints) to enhance model performance.
* **Churn Model Development:** Design, train, validate, and test machine learning models to predict customer churn probability. This includes exploring various algorithms (e.g., Logistic Regression, Random Forest, Gradient Boosting).
* **Churn Driver Identification:** Identify and quantify the most significant factors contributing to customer churn based on model insights.
* **Prediction Generation:** Generate periodic churn probability scores and risk segments for all active customers.
* **Reporting and Insights Dashboard:** Develop a user-friendly dashboard to display key churn metrics, at-risk customer lists, and insights into churn drivers.
* **Documentation:** Prepare comprehensive documentation of data sources, methodology, model architecture, and usage guidelines.

**OUT OF SCOPE** 49

The following are "out of scope" for the project: 50

* **Automated Retention Campaigns:** Implementation of automated marketing or customer service campaigns triggered by churn predictions. 51
* **Real-time Churn Prediction:** Development of a real-time system for churn prediction at the transaction or interaction level. 52
* **Product/Service Development:** Direct changes or enhancements to banking products or services based on churn insights in this phase. 53
* **Direct Customer Intervention:** Actual execution of customer retention calls or campaigns; the project will provide the *list* for these actions. 54

**4. BUSINESS DRIVERS 55**

The following are the primary reasons why the bank is initiating the "Bank Customer Churn Analysis and Prediction" project: 56

* **Customer Retention Cost vs. Acquisition Cost:** Because retaining an existing customer is significantly less expensive than acquiring a new one, reducing churn directly improves profitability and optimizes marketing spend.
* **Impact on Profitability:** Because high churn rates erode the customer base and reduce the overall lifetime value of customers, leading to decreased revenue and profitability.
* **Competitive Landscape:** Because in a highly competitive banking sector, proactively addressing customer satisfaction and loyalty is crucial to maintain market share and competitive advantage.
* **Data-Driven Decision Making:** Because leveraging our extensive customer data for predictive insights will enable more informed and effective business strategies.
* **Customer Experience Improvement:** Because understanding the reasons for churn will allow us to address pain points and improve the overall customer experience, leading to higher satisfaction and loyalty.
* Etc. 57

**5. CURRENT PROCESS 58**

Our current process for addressing customer churn is largely reactive and lacks predictive capabilities. 59

* Currently, customer churn is identified primarily *after* an account closure or a significant period of inactivity is observed in the core banking system.
* Monthly or quarterly reports are generated to track historical attrition rates, but these reports are static and do not provide insights into *why* customers are leaving or *who* is likely to leave next.
* Retention efforts, when made, are often ad-hoc, based on general customer feedback or high-value customer segments, rather than precise risk assessments.
* There is no integrated view of customer behavior across different systems (e.g., core banking, CRM, online activity), making it difficult to understand the holistic customer journey and identify early warning signs of churn.

**6. PROPOSED PROCESS 60**

With the implementation of the "Bank Customer Churn Analysis and Prediction" system, the proposed process for addressing customer churn will become proactive and data-driven. 61

* **Data Integration & Preparation:** Data from various sources (core banking, CRM, online activity) will be regularly extracted, cleaned, and integrated into a unified dataset.
* **Feature Engineering & Model Training:** This integrated data will be used to engineer new features and train the churn prediction model on an ongoing basis (e.g., weekly or monthly).
* **Churn Probability Scoring:** The trained model will generate a churn probability score for every active customer in the bank's database.
* **Customer Segmentation & Prioritization:** Based on these scores, customers will be segmented into different risk categories (e.g., High, Medium, Low risk of churn). A prioritized list of high-risk customers will be generated.
* **Actionable Insights & Reporting:** An interactive dashboard will provide insights into churn drivers, model performance, and the characteristics of at-risk segments.
* **Targeted Retention Strategies:** The prioritized list of at-risk customers and churn driver insights will be disseminated to relevant teams (e.g., Marketing, Customer Relationship Managers) to enable targeted and personalized retention campaigns and interventions.
* **Continuous Improvement:** The model will be periodically re-evaluated and retrained with new data to maintain its accuracy and adapt to changing customer behaviors.

**7. FUNCTIONAL REQUIREMENTS 62**

This section details the functional requirements necessary for the project's success. 63

**PRIORITY 64**

|  |  |  |
| --- | --- | --- |
| **Value** | **Rating** | **Description** |
| 1 | Critical | The requirement is critical to the project's success. Without fulfilling this requirement, the project is not possible. 65 |
| 2 | High | The requirement is high priority re the project's success, but the project could still be implemented in a minimum viable product (MVP) scenario. 66 |
| 3 | Medium | The requirement is important to the project's success, as it provides value, but the project could still be implemented in an MVP scenario. 67 |
| 4 | Low | The requirement is low priority (i.e., it would be nice to have), but the project's success is not dependent upon it. 68 |
| 5 | Future | The requirement is outside of the project's scope and is included as a possible component of a prospective release and/or feature. 69 |

**REQUIREMENTS CATEGORIES (RC1) - Data & Model Development 70**

|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **Requirement** | **Priority** | **Raised By** |
| FR-DM-001 | The system shall integrate customer demographic data from the core banking system. | 1 | Business Analyst |
| FR-DM-002 | The system shall integrate account details (type, balance, tenure) from the core banking system. | 1 | Business Analyst |
| FR-DM-003 | The system shall integrate transaction history (frequency, amount) data. | 1 | Business Analyst |
| FR-DM-004 | The system shall integrate CRM interaction history (complaints, service requests). | 2 | Business Analyst |
| FR-DM-005 | The system shall be able to handle and clean missing values in the integrated dataset. | 1 | Data Scientist |
| FR-DM-006 | The system shall be able to create new features (e.g., average monthly balance, number of complaints in last 6 months). | 1 | Data Scientist |
| FR-DM-007 | The system shall define "churn" based on agreed-upon criteria (e.g., account closure or no activity for X months). | 1 | Business Stakeholder |
| FR-DM-008 | The system shall train a machine learning model to predict churn probability. | 1 | Data Scientist |
| FR-DM-009 | The model shall output a churn probability score for each customer between 0 and 1. | 1 | Data Scientist |
| FR-DM-010 | The system shall evaluate the model using metrics such as Accuracy, Precision, Recall, F1-Score, and AUC-ROC. | 1 | Data Scientist |
| FR-DM-011 | The system shall identify and display the key drivers of customer churn. | 2 | Business Analyst |
| FR-DM-012 | The system shall provide a mechanism for regular model retraining. | 2 | Data Scientist |

**REQUIREMENTS CATEGORIES (RC2) - Reporting & Insights 71**

|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **Requirement** | **Priority** | **Raised By** |
| FR-RI-001 | The system shall generate a prioritized list of high-risk customers based on churn probability scores. | 1 | Business Analyst |
| FR-RI-002 | The system shall categorize customers into churn risk segments (High, Medium, Low). | 2 | Business Analyst |
| FR-RI-003 | The system shall provide an interactive dashboard displaying overall churn rates and trends. | 2 | Business Analyst |
| FR-RI-004 | The dashboard shall visualize key churn drivers. | 2 | Business Analyst |
| FR-RI-005 | The dashboard shall display model performance metrics (Accuracy, Recall, etc.). | 2 | Data Scientist |
| FR-RI-006 | The system shall allow export of the high-risk customer list in CSV format. | 3 | Business Stakeholder |
| FR-RI-007 | The system shall allow drill-down into customer segments to understand their characteristics. | 3 | Business Analyst |

**8. NON-FUNCTIONAL REQUIREMENTS 72**

|  |  |  |
| --- | --- | --- |
| **ID** | **Requirement** | **Priority** |
| NFR-PERF-001 | The model must generate churn predictions for the entire customer base within 6 hours for weekly updates. | 1 |
| NFR-PERF-002 | Data extraction and preparation processes must complete within 4 hours daily. | 2 |
| NFR-PERF-003 | The insights dashboard should load within 5 seconds for any query. | 2 |
| NFR-SEC-001 | All customer data used for analysis must be pseudonymized. | 1 |
| NFR-SEC-002 | Access to the prediction system and reports must be role-based and secured. | 1 |
| NFR-SEC-003 | The system must comply with GDPR and local banking data privacy regulations. | 1 |
| NFR-USAB-001 | The insights dashboard must be intuitive and easy to navigate for business users. | 2 |
| NFR-USAB-002 | The output format for at-risk customers must be compatible with existing CRM systems. | 2 |
| NFR-SCAL-001 | The system should be scalable to accommodate a 50% increase in customer data volume over the next 2 years. | 2 |
| NFR-RELI-001 | The data pipelines and model inference process must be automated and reliable. | 1 |
| NFR-RELI-002 | Automated error logging and alerting mechanisms must be in place. | 2 |
| NFR-MAIN-001 | All code for data processing and model development must be well-documented. | 2 |
| NFR-MAINT-002 | The model architecture should facilitate easy retraining and updates. | 2 |
| NFR-INTEG-001 | Seamless data integration with the bank's core banking system. | 1 |
| NFR-INTEG-002 | Potential for future API integration with CRM/marketing automation systems. | 3 |

**9. FINANCIAL STATEMENTS 73**

*This section would typically contain details regarding the financial health of the project, including projected revenue impacts from reduced churn and any direct costs. For a student project, this might focus more on the estimated cost of resources, software licenses (if any), and potential savings for the hypothetical bank.*

* **Estimated Cost for Project Execution:** [e.g., Staff time, cloud computing resources, software licenses (if any)]
* **Projected Financial Benefit:** [e.g., Estimated savings from reduced churn (e.g., 5% reduction in churn on a customer base of X customers, each with an average LTV of Y), increased customer lifetime value.]

**10. COST AND BENEFIT 74**

**Costs:**

* **Development Cost:** Estimated person-hours for data scientists, data engineers, and business analysts.
* **Technology Cost:** Cost of cloud computing resources (e.g., AWS, Azure, GCP) for data storage, processing, and model training/inference. Software licenses for any specific tools (e.g., data visualization, specialized ML libraries).
* **Data Acquisition/Integration Cost:** Effort involved in setting up data pipelines and ensuring data quality from various internal systems.
* **Training Cost:** Time and resources for training business users on the new dashboards and insights.

**Benefits:**

* **Reduced Customer Churn:** Directly translates to increased revenue and higher customer lifetime value.
* **Optimized Marketing Spend:** More efficient allocation of retention budgets by targeting high-risk customers.
* **Improved Customer Satisfaction:** By understanding and addressing churn drivers, overall customer experience can be enhanced.
* **Competitive Advantage:** Proactive churn management positions the bank as a customer-centric institution.
* **Data-Driven Insights:** Provides valuable intelligence about customer behavior that can inform product development and service improvements.

**11. RESOURCES 75**

* **Project Team:**
  + Business Analyst (Primary Author)
  + Data Scientist(s)
  + Data Engineer(s)
  + Project Manager
  + Subject Matter Experts from Retail Banking, Marketing, Customer Service
* **Technology Resources:**
  + Data Storage (e.g., Data Lake, Data Warehouse)
  + Data Processing Environment (e.g., Python/R, Spark, Cloud ML Platforms)
  + Visualization Tools (e.g., Tableau, Power BI, custom dashboards)
  + Version Control System (e.g., Git)
* **Data Access:** Secure access to historical customer data from various internal systems.
* **Budget:** Allocated budget for cloud services, tools, and personnel.

**12. SCHEDULE, TIMELINE, AND DEADLINES 76**

* **Project Kick-off:** [Date]
* **Data Collection & Integration Phase:** [Start Date] - [End Date] (e.g., 4 weeks)
* **Feature Engineering & Data Pre-processing Phase:** [Start Date] - [End Date] (e.g., 3 weeks)
* **Model Development & Training Phase:** [Start Date] - [End Date] (e.g., 6 weeks)
* **Model Evaluation & Refinement Phase:** [Start Date] - [End Date] (e.g., 2 weeks)
* **Dashboard Development & Reporting Integration:** [Start Date] - [End Date] (e.g., 3 weeks)
* **UAT (User Acceptance Testing):** [Start Date] - [End Date] (e.g., 1 week)
* **Deployment:** [Date]
* **Post-Implementation Review:** [Date, e.g., 3 months after deployment]
* **Overall Project Completion Deadline:** [Specific Date - e.g., End of Q2 2026]

**13. ASSUMPTIONS 77**

* Adequate access to historical customer data (transactional, demographic, interaction) will be granted for analysis and model training.
* The definition of "churn" (e.g., inactivity period, account closure) will be mutually agreed upon by business stakeholders and the data science team.
* Sufficient computing resources (e.g., cloud infrastructure, processing power) will be available for data processing and model training.
* Key business stakeholders will be available for regular feedback, data validation, and decision-making throughout the project lifecycle.
* The data quality of the source systems is sufficient, or resources will be allocated for data cleaning and reconciliation.
* The bank's internal IT department has the capacity to support the integration and eventual deployment of the solution.
* Business users (Marketing, Customer Service) are willing to adopt and act upon the insights provided by the churn prediction model.

**14. GLOSSARY 78**

|  |  |
| --- | --- |
| **Term / Abbreviation** | **Explanation** |
| **BRD** | Business Requirements Document |
| **Churn** | A customer ceasing to do business with the bank (e.g., closing all accounts, significant inactivity). |
| **CSAT** | Customer Satisfaction Score |
| **LTV** | Customer Lifetime Value |
| **FCR** | First Contact Resolution |
| **KPI** | Key Performance Indicator |
| **MVP** | Minimum Viable Product 79 |
| **AI** | Artificial Intelligence |
| **ML** | Machine Learning |
| **ETL** | Extract, Transform, Load (Data process) |
| **CRM** | Customer Relationship Management |
| **AUC-ROC** | Area Under the Receiver Operating Characteristic Curve (a model evaluation metric) |
| **PII** | Personally Identifiable Information |
| **UAT** | User Acceptance Testing |

**15. REFERENCES 80**

* [Internal Bank Data Standards Document Name], Location: [Internal Shared Drive/SharePoint Link]
* [Relevant Industry Reports on Customer Churn], Location: [Web Link/Document Name]
* [Any specific academic papers or benchmarks referenced], Location: [Web Link/Document Name]

**16. APPENDIX 81**

* **High-Level Data Flow Diagram:** (Diagram illustrating data movement from source systems to the prediction model and reporting dashboard)
* **Detailed Churn Definition:** (Specific criteria and timeframes for defining a "churned" customer for the model)
* **Initial Feature Set Brainstorm:** (List of potential features derived from customer data for model training)