**2.1 Project Background**

**Problem/Opportunity Description**

Losing clients is more than just a small setback in the quick-paced world of subscription-based businesses; it directly affects the bottom line. Imagine being able to identify which clients are about to leave and be able to entice them back before they do. Our Churn Prediction System is designed to do just that. We're developing a platform that uses state-of-the-art machine learning and data analytics to forecast customer attrition and offer useful insights to maintain and expand your clientele.

**Benefits**

**Enhanced Customer Retention**:  
By employing churn prediction models to detect customers at risk ahead of time, this method assists businesses in taking preventative action to create focused retaining initiatives. In order to maintain company development, these strategies will increase for a long time client retention, reduce loss, and retain reputable customers.

**Increased Revenue**:  
When churn is reduced, an increasingly steady consumer results in greater regular revenue. By reducing consumer loss, the technology helps maintain a higher customer lifetime value (CLV), that ultimately boosts sales and fortifies the company's financial health.

**Data-Driven Decision Making**:  
This application uses predicted analytics along with machine learning to provide businesses with based on data, practical recommendations. These findings will help businesses achieve more strategic, well-informed decisions about marketing campaigns, product development, and customer contact, ensuring that all choices are based on data as opposed to speculation or intuition. This will improve operational efficiency and allow businesses to quickly adapt to changing customer demands.

**Goals**

The goal of the project seeks to develop a sophisticated prediction of churn system that will assist businesses in anticipating at-risk customers and implementing tailored retention methods. This solution will leverage data machine learning and data analytics to provide enterprises with concrete findings that will enhance customer satisfaction, profits, and retention. This includes the development of prediction churn models, data collection and preprocessing, feature engineering and selection, model construction, learning, and evaluation, dashboards graphical representation, integration as an API, and customized retention tactics.

**Stakeholders and Clients**

Internal Stakeholders: The university, academic advisers, and our committed project team.

External Stakeholders: Subscription-based organizations wanting to increase their client retention methods.

Clients: Media streaming services, SaaS firms, and e-commerce sites are keen to increase customer loyalty and lower attrition.

People that are going to be affected: Businesses that compete with their customers using this project.

**2.2 Project Scope**

**Objectives:**

**Objective 1: Data Collection & Preparation**  
To be prepared for training a machine learning model, collect and preprocess pertinent consumer data, such as demographics, subscription activity, and use statistics. This covers feature selection, data cleansing, and dataset preparation for modeling.

**Objective 2: Machine Learning Model Development**  
Using the supplied dataset, create a predictive churn model that uses suitable machine learning methods (such logistic regression, random forest, or neural networks) to identify consumers who are at risk of churning. The model will undergo training, validation, and performance evaluation.

**Objective 3: Web Application Development**  
Create a web-based dashboard that visualizes customer behavior, churn projections, and other pertinent data using Flask/Django for the server and React for the frontend. Users will be able to engage with the churn estimates and take action on clients who are at danger thanks to this application.

**Objective 4: Alerts & Notifications**  
Establish in place a system that will send out notifications when a client is determined to be at high risk of leaving. Business users will get alerts through the application interface, and they may be set up to receive other notifications (like emails).

**Objective 5: Documentation**  
To ensure a smooth handoff and future application maintenance, provide thorough project documentation that includes system architecture, model information, user instructions, and deployment requirements.

**Deliverables:**

**Objective 1**:

|  |  |
| --- | --- |
| **Project Deliverable** | **Work Products/Description** |
| Dataset | Customer data that has been preprocessed and prepared for training machine learning models, such as demographics, subscription activity, and use statistics. |

**Objective 2:**

|  |  |
| --- | --- |
| **Project Deliverable** | **Work Products/Description** |
| ML Model | Built utilizing suitable techniques like logistic regression, random forest, or neural networks, trained and validated the churn prediction model, and assessed its performance. |

**Objective 3:**

|  |  |
| --- | --- |
| **Project Deliverable** | **Work Products/Description** |
| API | REST API for model inference based on Flask/Django, which allows for predictive integration with the web application. |
| Dashboard | web-based visualization produced using Streamlit, Tableau, or Power BI that provides information about churn forecasts and customer behavior for interactive analysis. |

**Objective 4:**

|  |  |
| --- | --- |
| **Project Deliverable** | **Work Products/Description** |
| Alert System | When a client is determined to be at high risk of churn, a system is in place to generate alerts. Notifications are sent by email or the application interface. |

**Objective 5:**

|  |  |
| --- | --- |
| **Project Deliverable** | **Work Products/Description** |
| Final Report | Thorough project documentation that includes user manuals, deployment directions, system architecture, model information, methodology, and outcomes. |

**Out of Scope:**

1. **Universal Integration and CRM Integration**

The project does not include integration with any CRM or subscription platforms. Creating a churn prediction system for a certain, predetermined set of platforms will be the main goal.

1. **Enterprise-Level Real-Time Processing**

Large-scale enterprise-level data handling and real-time data processing are outside the purview. Batch processing and periodic churn prediction will be the project's main objectives.

1. **Advanced Features Beyond Churn Prediction**

This phase will not contain sophisticated features like consumer segmentation or in-depth behavioral analytics. With the possibility of future improvements, the emphasis will continue to be on churn prediction.

**2.3 Project Plan**

**Approach and Methodology:**

**Agile Methodology:** Agile project management will be used to guarantee adaptability and incremental delivery. To enable constant progress, frequent feedback, and flexibility in response to changes, the development will be divided into sprints that are scheduled every two weeks.

Key Phases:

Planning & Requirements Gathering: The first stage to collect stakeholders' specific project needs.

Agile Sprints:  
 Biweekly sprints that are agile and include:

* Sprint Planning: Using sprint planning, rank tasks according to business requirements.
* Daily Standup: Hold daily standups to review progress and resolve obstacles.
* Sprint Review: Present finished products and get input.
* Sprint Retrospective: Conduct a sprint retrospective to evaluate the sprint and enhance procedures.

Iterative Development: Every sprint will concentrate on particular deliverables such warning and notification integration, dashboard design, churn model development, and data pretreatment.

Testing and Quality Assurance: User Acceptance Testing (UAT), integration testing, and unit testing will all be conducted continuously to ensure functionality, security, and performance.

Feedback from Stakeholders and Adjustments: Consistent evaluations with stakeholders guarantee that the project remains in line with company goals and that any required modifications are made.

**Project Timeline:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ID** | **Task Name** | **Start** | **Finish** | **Duration** |
| 1 | Requirements Gathering | 1/29/25 | 2/2/25 | 1 |
| 2 | Data Collection & Integration | 2/3/25 | 2/9/25 | 1 |
| 3 | Data Preprocessing & Feature Engineering | 2/10/25 | 2/16/25 | 1 |
| 4 | Machine Learning Model Development | 2/17/25 | 3/1/25 | 2 |
| 5 | Backend Development | 2/17/25 | 3/1/25 | 2 |
| 6 | Frontend Dashboard Design | 2/24/25 | 3/7/25 | 2 |
| 7 | Model Integration with Backend | 3/2/25 | 3/7/25 | 1 |
| 8 | Frontend Integration & Testing | 3/8/25 | 3/14/25 | 1 |
| 9 | Retention Strategy Implementation | 3/15/25 | 3/21/25 | 1 |
| 10 | Alerts and Notifications | 3/22/25 | 3/28/25 | 1 |
| 11 | System Testing & Bug Fixes | 3/29/25 | 4/4/25 | 1 |
| 12 | User Acceptance Testing (UAT) | 4/5/25 | 4/11/25 | 1 |
| 13 | Deployment & Final Documentation | 4/12/25 | 4/18/25 | 1 |
| 14 | Final Presentation & Review | 4/19/25 | 4/25/25 | 1 |
| 15 | Project Handover & Submission | 4/26/25 | 4/30/25 | 1 |

**Success Criteria:**

A churn prediction model with a high F1 score, which guarantees high accuracy in identifying at-risk consumers, will be used to gauge the project's effectiveness.

The project will also be considered successful if a functioning dashboard that is easy to use and gives business users clear, actionable insights is supplied.

Lastly, by lowering customer attrition by 10–15% during the first quarter of implementation, the project will have a quantifiable commercial benefit, improving customer retention and raising the organization's lifetime value.

**Issues and Policy Implications:**

Data Privacy: Verify that all data management conforms with privacy laws.

Ethical Use: Make a commitment to using predictive analytics in an ethical manner.

**Risk Management Plan:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk Factor** | **Probability  (H-M-L)** | **Impact  (H-M-L)** | **Risk Management Action** |
| Data Quality Issues | High | High | Make use of a variety of data sources and thorough preparation methods. routine audits of data. |
| Model Inaccuracy | Medium | High | Routine testing and validation. If the primary models don't work as well, use fallback models. |
| User Resistance to New System | Medium | Medium | Organize training sessions and offer thorough documentation. |
| System Downtime | Low | High | Have a catastrophe recovery strategy in place and put in place a strong monitoring system. |
| Support Center Overload | Medium | Medium | During the initial deployment, increase headcount and provide clear escalation routes. |

**Service Transition:**

**Role of the Support Center in Ongoing Support:**

After the service is deployed, the Support Center will keep offering assistance. Depending on the deliverables and the requirements of the service, modifications to direct assistance, escalation routes, and FAQs will be made.

Implementation: The Support Center's contents will be expanded to include new frequently asked questions and escalation protocols. Support Center employees will receive training to acquaint them with the new service elements.

**Impact Mitigation:**

End User Impact: During the transition period, especially during the "go live" events, any effects on end users are expected.

Support Plan: Any end-user problems will be handled by a specialized staff. Change communication will be controlled via email, intranet updates, and service notifications.

Tools/References/Checklists: The support staff will get tools, references, and checklists to guarantee uniformity and effective communication. When necessary, the IT staff will provide direct support for end-user computers.

**Transitional Support:**

Additional Support: To handle any urgent problems during the go-live phase, more help will be needed.

Coordination: The extra workload has been communicated to the Support Center, CRM, and Service Team. During this time, a project manager will oversee the support activities to guarantee seamless transitions and prompt problem solving.

**Service Management:**

Changes to Service Agreements: Operating level agreements (OLAs), service level agreements (SLAs), and other service procedures might need to be updated.

Coordination and Communication: The service management team will arrange any required modifications to these agreements. They will make sure the required changes are made to websites that are visible to the public and notify the appropriate parties of the revisions, if applicable.

Service Team: Following the transfer, a specialized service team will be in place to manage the service's continued operation and support.

**Handover and Ongoing Support:**

Documentation and Training: Stakeholders will get thorough documentation, and training sessions will be held to guarantee appropriate knowledge transfer.

Ongoing Support: For a little time after the project is finished, continuing help will be offered to handle any unanticipated problems or requests.

**Options Analysis:**

**Collaboration Across the Clark University System:**

Approach Considered: Using procurement procedures or working with partner universities to share resources or cut expenses.

Analysis: The particular requirements of this project (such as the development of machine learning models, web applications, and dashboards) call for specialized tools and expertise not currently offered by other campuses or existing services, even though working together across the Clark University system may offer opportunities for cost savings or shared resources.

Show-Stoppers: It is less practical to use current Clark services or work with other campuses because of the technological needs of this project, which include creating unique machine learning models and implementing a customized interface.

**Open Source Options:**

Approach Considered: Seeking free or inexpensive open source tools for developing machine learning models, visualizing dashboards, or providing API services.

Analysis: For several aspects of the project, open source tools like TensorFlow, Scikit-learn, Flask, or Streamlit were taken into consideration. These tools are appealing choices since they are flexible and don't require any license fees.

Show-Stoppers: Open-source solutions can save money, but they can also present problems with scalability, interaction with current infrastructure, and continuing maintenance. Long-term maintenance and troubleshooting would need more work from the development team, thereby raising the total cost of ownership.

**Extending or Enhancing Existing Tools or Services:**

Approach Considered: Determining whether the project's needs may be satisfied by extending the use of current technologies (such as customer relationship management (CRM) systems and analytics platforms).

Analysis: In order to possibly integrate churn prediction or customer data straight into current dashboards, tools such as Power BI, Tableau, and CRM systems were examined.

Show-Stoppers: The sophisticated machine learning skills required to precisely forecast churn are absent from current technologies. Furthermore, expanding these capabilities would need extensive customisation, which might be expensive and time-consuming. Furthermore, compatibility problems and a lack of flexibility may make it challenging to integrate machine learning models into current systems.

**Product Comparisons:**

Approach Considered: Evaluating several dashboard programs, APIs, and machine learning platforms in order to choose a product.

Analysis: Platforms including Flask/Django-based APIs, Tableau, Power BI, and Streamlit were taken into consideration. The accuracy and scalability of Scikit-learn and TensorFlow were assessed for model construction.

Findings: Although Power BI/Tableau has strong dashboard visualization, it does not have built-in machine learning capabilities. The best options for the project are Flask/Django and Scikit-learn as they offer the flexibility required for the creation of the custom machine learning model and API.

Compelling Reasons for Selection: Power BI/Tableau guarantees efficient presentation of the churn predictions, while Flask/Django for API development and Scikit-learn for model creation offer the most flexibility and scalability.

**Alternative Project Approaches:**

Approach Considered: Putting off the project or choosing a less complicated approach.

Analysis: If time or resources are limited, postponing the project could be a possibility, but considering the possible effect on client retention, it might not be advantageous in the long term.

Show-Stoppers: Actionable insights are necessary to boost retention and decrease churn immediately. Delaying the project will probably result in ongoing clientele loss, which might affect long-term earnings and the viability of the company.

**2.4 Technical Features**

To enhance churn prediction and client retention for the company, the suggested application will have a number of essential features:

**Machine Learning Model for Churn Prediction**

Created with the aid of sophisticated algorithms like neural networks, logistic regression, and random forests.  
Trained and verified using past customer data to make precise predictions about customer attrition.   
For smooth, real-time predictions, it is integrated into the program through an API.

**User-Friendly Dashboard**

An online visualization tool that makes use of Streamlit, Tableau, or Power BI.   
Shows important business KPIs, consumer behavior insights, and churn forecasts.   
Enables business users to engage with the data and take action on customers who are at risk.

**Backend Services and API**

Constructed using Flask or Django to manage data processing and model inference.   
Offers endpoints for managing data integrations and obtaining forecasts.

**Quality Assurance and Testing**

The application, model performance, and integration points are continuously tested and validated.   
QA procedures guarantee that the application is reliable, safe, and error-free.