**CHAPTER FOUR: SYSTEM ANALYSIS AND REQUIREMENT MODELING**

## ****4.1 Introduction****

This chapter provides a comprehensive analysis of the **current system** used in banking and financial services for churn prediction, sentiment analysis, and personalized recommendations. It explains how the current system works, identifies its limitations, and presents an **AI-driven approach** to enhance automation, efficiency, and accuracy.

I will use **system modeling tools** such as **Flowcharts, Data Flow Diagrams (DFDs), UML diagrams, and Use Cases** to illustrate the existing and proposed systems. Additionally, this chapter discusses **data collection methods, system requirements, and requirement modeling** to ensure a robust AI-driven solution.

**4.2 Current System Analysis**

**4.2.1 Existing System Workflow**

Currently, banks and financial institutions rely on manual or semi-automated processes to handle customer churn analysis, sentiment tracking, and personalized recommendations. Below is a breakdown of the key limitations of the existing system:

1.**Manual Data Processing**: Customer data is collected manually from sources like transaction logs, customer complaints, feedback forms, and surveys.

This results in delayed insights due to manual analysis and report generation.

2.**Limited Churn Prediction:** Basic rule-based or historical analysis is used to estimate customer attrition.

No real-time predictive analytics are in place, making it difficult to anticipate customer churn early.

3.**Sentiment Tracking via Feedback Forms**: Customer feedback is collected using surveys, call center logs, and online forms.

The lack of AI-powered sentiment analysis means banks fail to identify real-time customer dissatisfaction.

4**.Generic Recommendations**: Banks use generalized promotional offers instead of personalized recommendations.

There is no AI-driven personalization based on customer behavior and sentiment.

### ****4.2.2 System Analysis Modeling Tools****

To better understand the **existing system workflow**, we use **system modeling tools** such as **Flowcharts, Data Flow Diagrams (DFDs), Use Case Diagrams, and UML Diagrams**.

#### **(a) Flowchart of the Existing System**

A flowchart visually represents the steps followed in the current system, from data collection to report generation

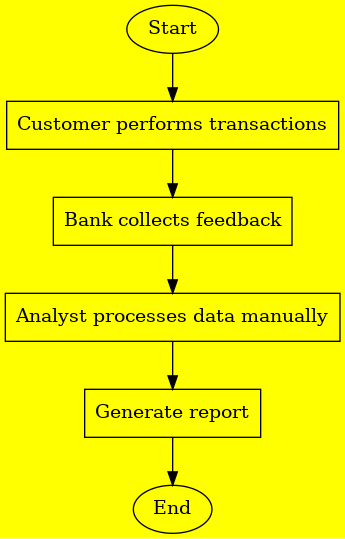


Figure 4.1 flowchart

**(b) Data Flow Diagram (DFD)**

The **DFD** is representing the flow of data in the AI-Driven Churn Prediction, Sentiment Analysis, and Recommendation System for Banking and Financial Services. It consists of two levels namely:

**Level 0:** High-Level Data Flow

Actors: Customers, Banking CRM, Analytics Team

Processes: Data Collection → Processing → Churn & Sentiment Analysis → Reports & Decisions

**Level 1:** Detailed Data Movement

Data Sources: Transactions, Feedback, Support Tickets

Processing Units: Customer Segmentation, Churn Scoring, Sentiment Detection

Outputs: Reports for decision-making

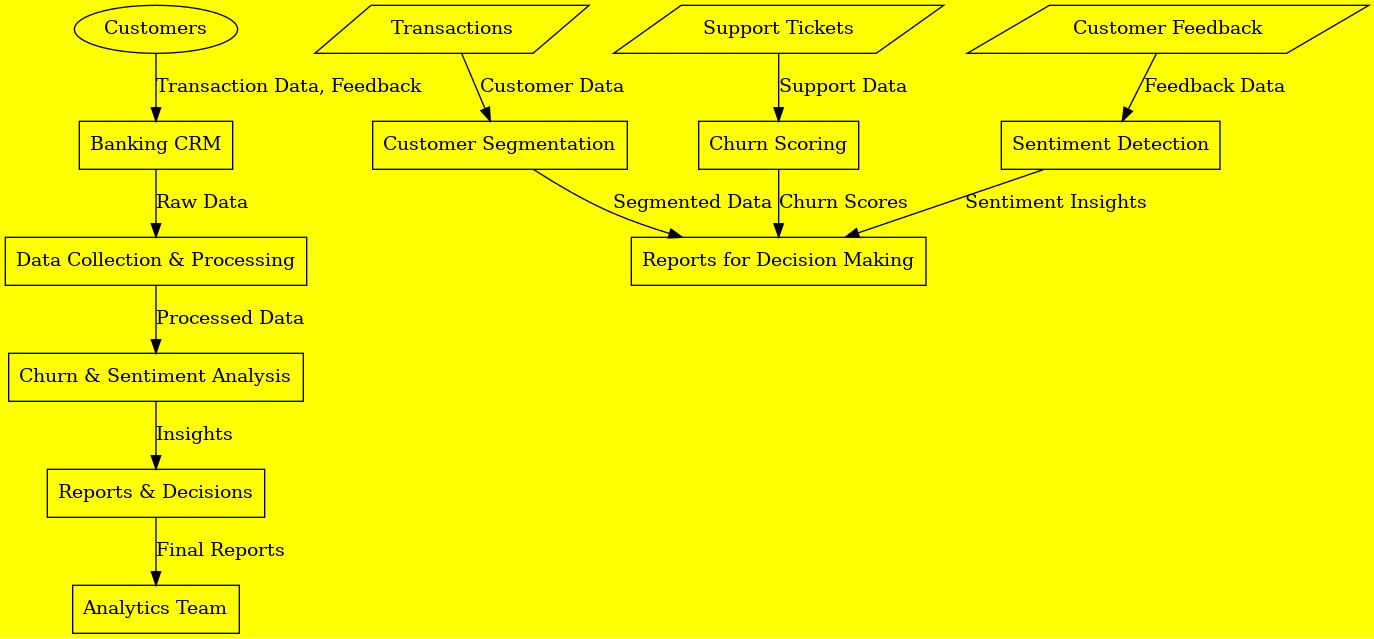


Figure 4.2 DFD Diagram

**(c) Use Case Diagram**

The **Use Case Diagram is** visually representing how different actors interact with the AI-Driven **Churn Prediction, Sentiment Analysis, and Recommendation System** in a banking environment.

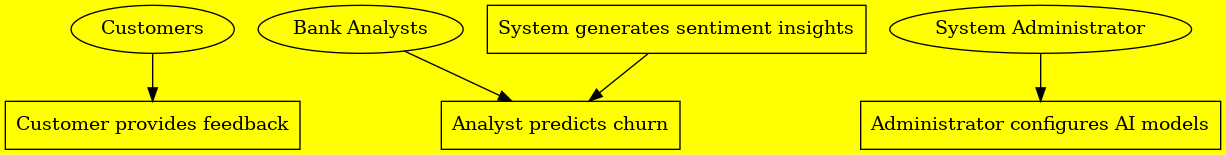
****

Figure 4.3 Use Case Diagram

**(d) UML Diagrams**

The **UML Class Diagram** represents the structure of the AI-Driven **Churn Prediction, Sentiment Analysis, and Recommendation System** by illustrating the key entities, their attributes, and relationships.

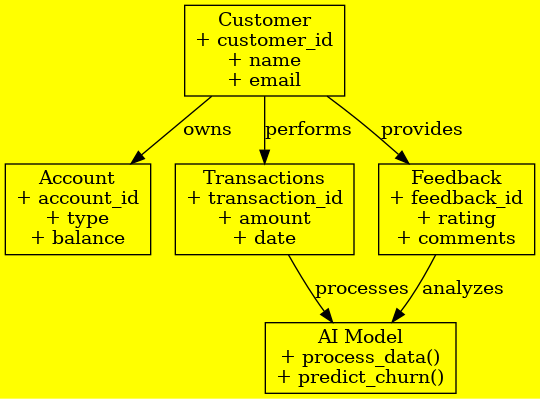


Figure 4.4 UML Diagram

**4.3 Data Collection and Fact Gathering**

**4.3.1 Data Collection Methods**

Data collection is essential for AI models that predict customer churn, analyze sentiment, and provide recommendations in the banking industry. The effectiveness of these models depends on the richness and diversity of the data used. Here are the methods used for data collection that contribute to churn prediction, sentiment analysis, and personalized recommendations:

**1.Surveys & Interviews**:

**Purpose:** Gather direct customer feedback about their experiences with the bank, uncovering reasons for satisfaction or dissatisfaction.

**Data Type**: Qualitative (opinions, experiences) and quantitative (ratings, satisfaction scores).

**Churn Prediction**: Recurring negative feedback or dissatisfaction indicators can suggest a higher risk of churn, helping the AI model predict and intervene early.

**Sentiment Analysis:** Feedback from surveys and interviews provides insight into customer sentiment, categorizing it as positive, negative, or neutral.

**Recommendation:** Insights from feedback can guide recommendations for improving customer experience, such as enhancing specific features or addressing common concerns.

**2.System Logs & Transactions**:

**Purpose:** Analyze transaction patterns and customer interactions with the bank's digital platforms to detect engagement or dissatisfaction.

**Data Type:** Quantitative (transaction amounts, frequency, types of transactions, login behavior).

**Churn Prediction:** Decreased activity or irregular transactions may indicate disengagement, which is a predictor of potential churn.

**Sentiment Analysis:** The frequency and type of customer interactions can help gauge overall sentiment towards banking services—whether positive or negative.

**Recommendation:** Transaction data can inform personalized recommendations, such as tailored banking products or services based on usage patterns and preferences.

**3.Social Media Sentiment Analysis**:

**Purpose:** AI analyzes public sentiment on platforms like Twitter, Facebook, and Instagram to understand customer opinions about the bank and its services.

**Data Type:** Text data (posts, tweets, comments, reviews).

**Churn Prediction:** Negative sentiment or recurring complaints on social media can indicate dissatisfaction, which may signal the potential for customer churn.

**Sentiment Analysis:** Sentiment analysis algorithms categorize social media content into positive, negative, or neutral categories, providing a snapshot of public perception.

**Recommendation**: Social media insights can help the bank refine its offerings, develop marketing strategies, or address customer complaints by recommending actions based on trending sentiments.

**4.Historical Data Analysis:**

**Purpose:** Analyze past customer churn cases to identify behaviors and patterns that may lead to customer attrition.

**Data Type:** Historical transaction data, demographics, service interactions, churn status.

**Churn Prediction:** Historical data helps AI models recognize patterns of behavior associated with previous churn cases, allowing for proactive interventions with at-risk customers.

**Sentiment Analysis:** Historical interactions (e.g., customer service data) help gauge sentiment during the customer journey, identifying periods of dissatisfaction.

**Recommendation:** By examining past interactions and service usage, the AI model can recommend tailored retention strategies and personalized banking services to improve customer engagement and prevent churn.

**4.3.2 Data Processing and Preprocessing**

To improve the accuracy of AI models, data needs to go through preprocessing steps before being used for training:

**1.Data Cleaning:** Ensure data quality by removing duplicates, correcting errors, and handling missing values (e.g., through imputation or removal).

**2.Feature Engineering:** Extract and create relevant features from raw data, such as transaction frequency, sentiment scores, or customer engagement metrics, to provide valuable insights for the model.

**3.Data Labeling:** Categorize the data (e.g., labeling churned vs. active customers, positive vs. negative sentiment) to train AI models, allowing them to learn and make predictions based on labeled examples.

**4.4 Requirement Definitions and Modeling**

**4.4.1 Existing System Requirements**

Analyzing the existing system reveals several limitations:

* No real-time analytics
* No automated churn prediction
* Basic customer segmentation
* Manual sentiment tracking

**4.4.2 Proposed System Requirements**

The AI-driven system will address these limitations through automation.

**(a) Functional Requirements**

* Automated Churn Prediction: AI-based model predicts high-risk customers.
* Real-time Sentiment Analysis: Monitors customer satisfaction in real-time.
* AI-powered Personalized Recommendations: Provides retention strategies.
* User Dashboard & Reports: For analysts to monitor AI insights.

**(b) Non-Functional Requirements**

* Performance: AI models should deliver real-time insights.
* Scalability: Must handle millions of customers in banking systems.
* Security: Ensure data encryption and financial compliance.
* User-Friendliness: Intuitive UI for easy adoption.

**5.12. CHAPTER FIVE: SYSTEM DESIGN**

The system design for the AI-driven Churn Prediction, Sentiment Analysis, and Recommendation System for the banking and financial services industry is the core framework that will handle various data-driven tasks. This chapter outlines the architectural design, database design (conceptual, logical, and physical), and the key components of the system.

**5.1 System Architecture Overview**

The system architecture is designed to handle three main tasks:

**Churn Prediction** – Predicting customer churn based on various data points such as transaction history, customer feedback, and usage patterns.

**Sentiment Analysis** – Analyzing customer sentiments from surveys, social media, and customer interactions to understand public perception and customer satisfaction.

**Recommendation System** – Providing personalized recommendations for banking products, services, or improvements based on user behavior and preferences.

The system will consist of the following components:

**Data Collection Layer**: Gathers data from multiple sources such as transaction logs, surveys, social media, and customer service interactions.

**Data Processing Layer**: Processes raw data through data cleaning, feature engineering, and sentiment analysis.

**Machine Learning Layer**: Implements predictive models (e.g., churn prediction, sentiment analysis) and a recommendation engine using AI/ML algorithms.

**Database Layer**: Manages and stores structured and unstructured data, providing a central repository for all customer-related information.

**User Interface (UI):** Provides dashboards and interactive tools for bank employees and customers to view predictions, sentiments, and recommendations.

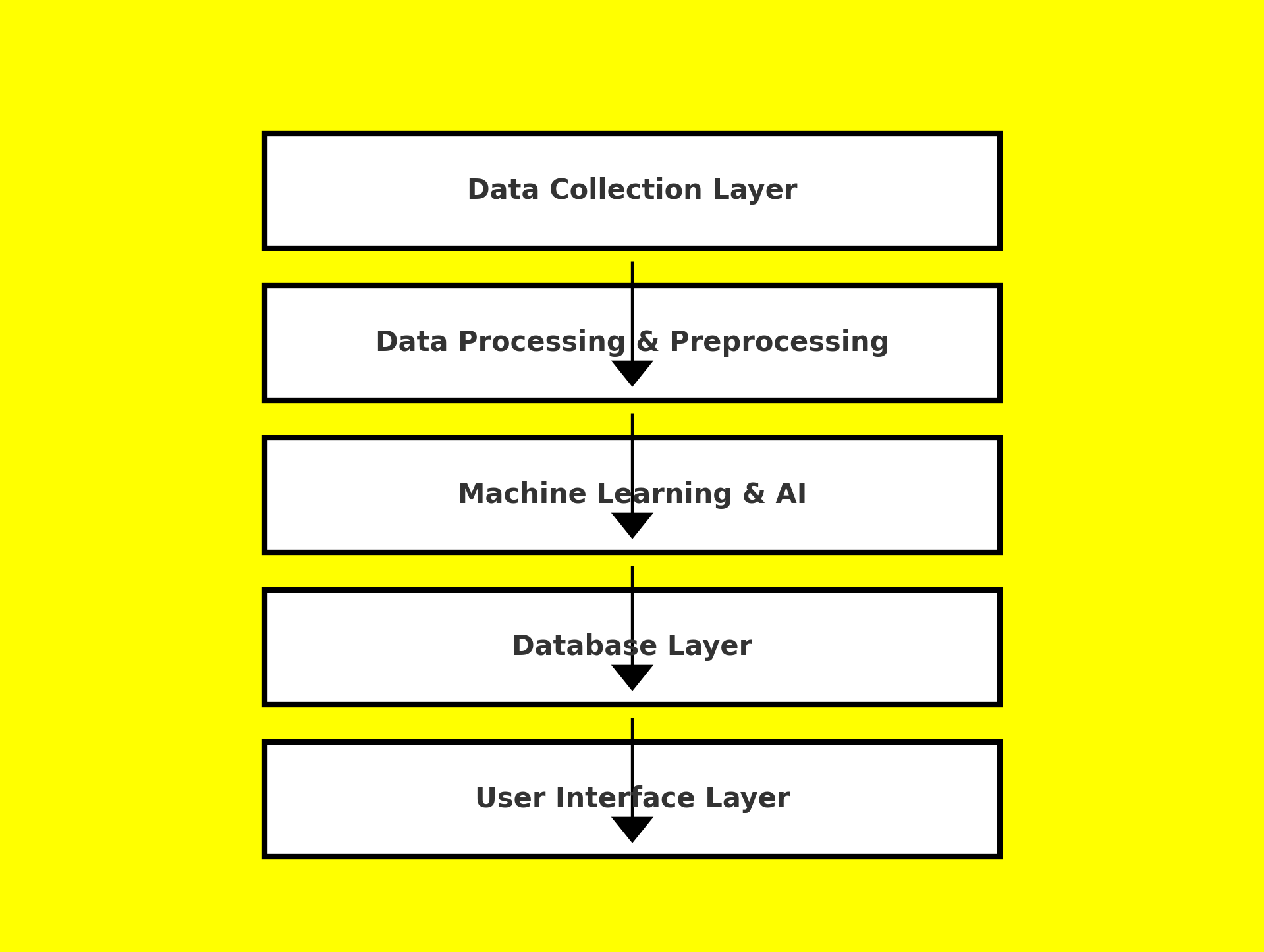


Figure 5.1 System Architectural Diagram

**5.2 Database Design**

The database design encompasses the creation of a logical and physical structure to store, manage, and retrieve the data efficiently. The database consists of multiple tables that store data related to customers, transactions, feedback, and sentiment analysis.

**5.2.1 Conceptual Database Design**

The conceptual database design provides a high-level representation of the data entities and their relationships. The primary entities in the system are:

**1.Customer:**

Attributes: Customer ID, Name, Contact Information, Demographics (age, location, etc.)

Relationship: A customer can have multiple transactions and feedback entries.

**2.Transaction:**

Attributes: Transaction ID, Customer ID, Amount, Type (deposit, withdrawal, loan payment, etc.), Date, Status

Relationship: Each transaction is linked to a customer.

**3.Survey Feedback:**

Attributes: Survey ID, Customer ID, Survey Date, Rating, Comments

Relationship: Each survey feedback entry is associated with a customer.

**4.Social Media Sentiment:**

Attributes: Post ID, Customer ID, Sentiment (positive, negative, neutral), Platform (Twitter, Facebook, etc.), Date

Relationship: Sentiment entries are associated with customers who expressed their opinions on social media.

**5.Recommendation:**

Attributes: Recommendation ID, Customer ID, Product ID, Recommended Service, Date

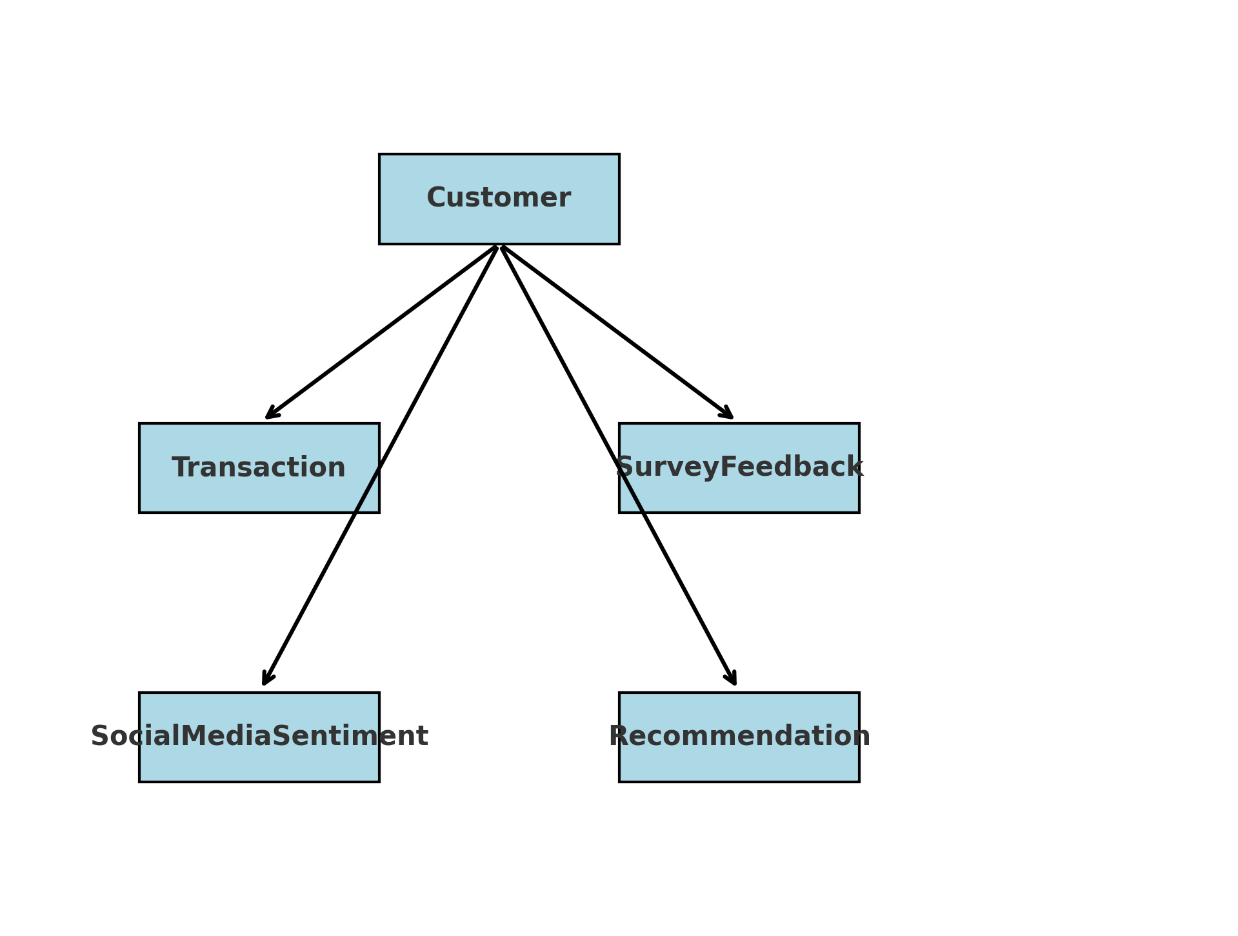
Relationship: Each recommendation is linked to a specific customer

Figure 5.2 Conceptual Database Design Diagram

**5.2.2 Logical Database Design**

The logical database design involves organizing the entities and their attributes into tables, establishing relationships between them, and defining keys. Here’s how the data will be structured in the logical model:

**Customer Table:** Stores all customer-related information.

**Transaction Table:** Contains transaction details and links to the customer.

**Survey Feedback Table**: Holds survey data and connects it to the customer.

**Social Media Sentiment Table:** Stores sentiment data from various platforms, associating them with customers.

**Recommendation Table:** Keeps track of the personalized recommendations generated for customers.

**Relationship Between Tables:**

One-to-Many Relationship: A single customer can have multiple transactions, surveys, and sentiment entries.

Many-to-Many Relationship: A customer can receive multiple recommendations for different products or services, and a product/service can be recommended to many customers.

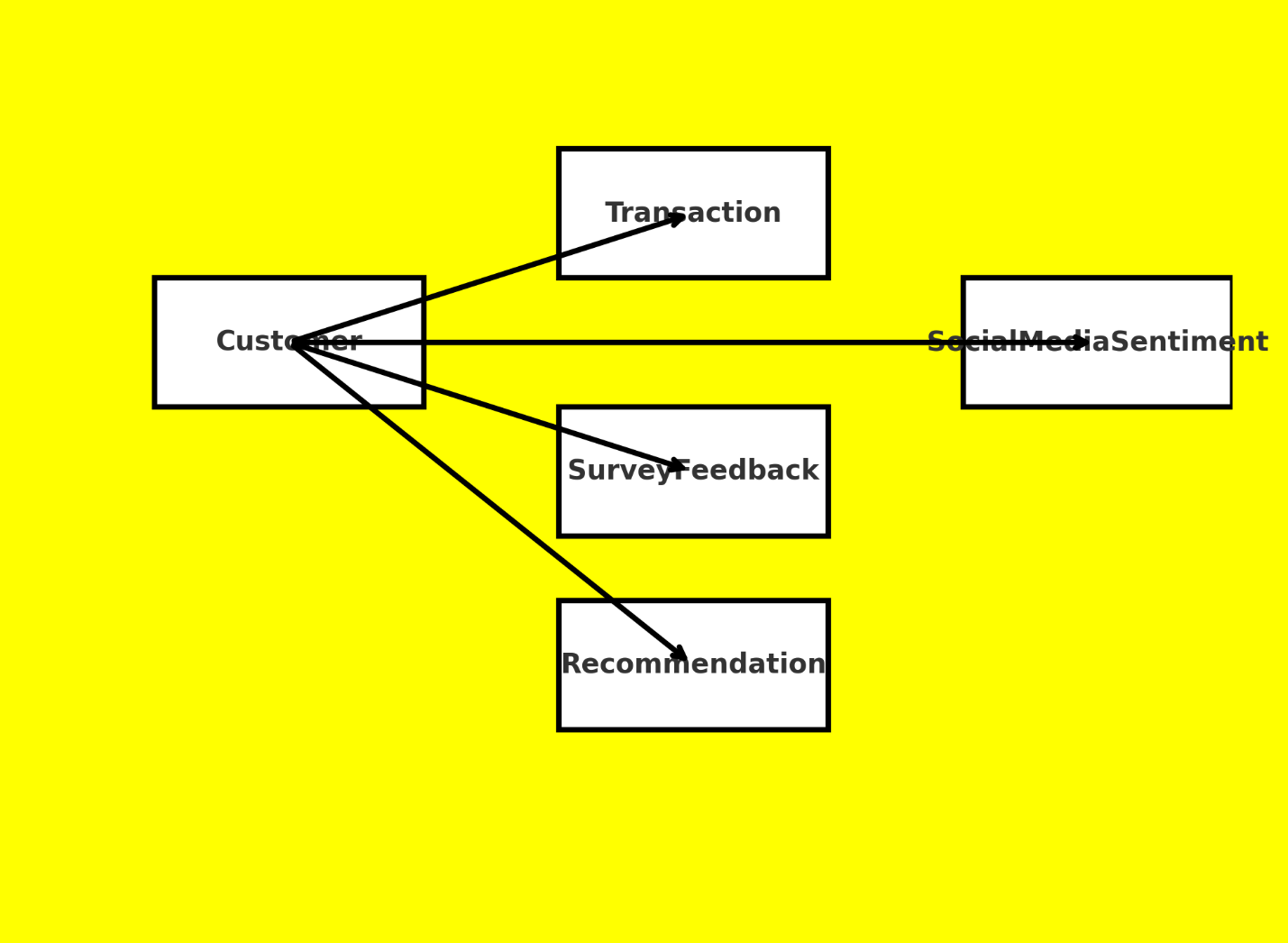


Figure 5.3 Logical Database Design Diagram

**5.2.3 Physical Database Design**

The physical database design focuses on the actual implementation of the system, considering factors like indexing, data storage, and query optimization. The physical design will involve the following:

**1. Data Storage:**

- Data will be stored in relational database management systems (RDBMS) like MySQL or PostgreSQL for structured data.

-For unstructured data (e.g., social media posts, customer feedback), a NoSQL database (e.g., MongoDB) can be used.

**2. Indexes:**

-Indexes will be created on frequently queried fields, such as Customer ID, Transaction ID, and Survey Date, to speed up data retrieval.

**3.Backup and Redundancy**:

-The database will have a robust backup mechanism in place to ensure that data is not lost in case of system failure.

-Redundant copies of the database will be stored in different locations for high availability.

**4.Encryption**:

-Sensitive customer information will be encrypted both at rest and during transmission using SSL/TLS protocols to ensure security and privacy.

**5.3 Machine Learning Models and Algorithms**

The AI-driven system will use various machine learning models to process data and provide insights.

**1. Churn Prediction:**

Algorithms: Random Forest

Input Data: Account Type, Employment Status, Loan Status, Credit Utilization Ratio and Active Products

Output: Predicted likelihood of customer churn (binary classification: churn or not churn).

**2. Sentiment Analysis:**

Algorithms: Logistic Regression and TfidfVectorizer

Input Data: customer reviews/ feedback

Output: Sentiment score (positive, negative).

**3. Recommendation System:**

Algorithms: Collaborative Filtering, Content-Based Filtering, and Hybrid Recommender Systems.

Input Data: Customer\_ID and Customer Feedback Rating

Output: Personalized product and service recommendations.

**5.4 System Integration**

The integration of various modules ensures smooth data flow between different parts of the system.

Data Flow: Customer data is collected through the user interface, which then flows through data processing, sentiment analysis, and churn prediction models. The processed data is stored in the database, which the recommendation engine uses to suggest relevant products or services.

API Integration: APIs can be used for connecting the system to external data sources (e.g., social media platforms for sentiment analysis).

Real-time Processing: The system should be capable of processing real-time transaction data and social media sentiment to update predictions and recommendations promptly.

**5.5 User Interface Design**

The UI will consist of dashboards and interactive tools for bank employees and customers to view:

* **Churn Prediction Dashboard**: Displays predicted churn risk levels for individual customers, enabling proactive engagement.
* **Sentiment Analysis Dashboard**: Provides sentiment trends based on customer feedback and social media posts.
* **Recommendation Dashboard**: Shows personalized recommendations for banking products or services, tailored to the customer's preferences and behavior.

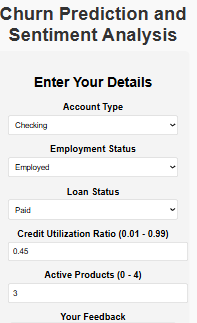
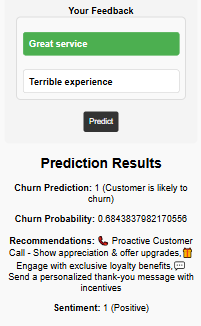
 

Figure 5.4 The Churn Prediction and Sentiment Analysis Dashboard with both inputs and outputs

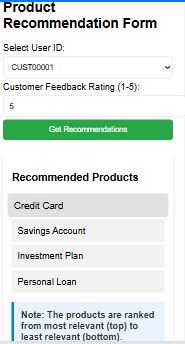


Figure 5.5 Recommendation System Dashboard both inputs and outputs

**5.6 Security and Privacy**

The system will prioritize the security and privacy of customer data by employing:

**User Authentication**: Ensuring that only authorized personnel can access sensitive customer data.

**Data Encryption**: Encrypting customer data at rest and during transmission.

**CHAPTER SIX: SYSTEM IMPLEMENTATION**

## 6.1 Tools Used for Coding and Testing

The implementation of the AI-driven churn prediction, sentiment analysis, and recommendation system for banking and financial services require a combination of various tools and technologies. The selection of these tools was based on their efficiency, compatibility, and ability to handle large datasets effectively. The key tools used include:

### ****Programming Languages:****

|  |  |
| --- | --- |
| Language | Importance |
| Python | Used for data processing, machine learning model development, and API integration. |
| Java script | Utilized for the development of the front-end interface. |

### ****Frameworks and Libraries:****

|  |  |
| --- | --- |
| Frameworks & Libraries | Importance |
| **TensorFlow/Keras** | Used for deep learning-based sentiment analysis and recommendation models. |
| **Scikit-learn** | Applied for traditional machine learning algorithms for churn prediction. |
| **Pandas & NumPy** | Used for data manipulation and preprocessing. |
| **Flask** | Served as the backend framework to integrate the machine learning models with the front-end. |

### ****Testing Tools:****

|  |  |
| --- | --- |
| Testing Tool | Importance |
| **Jupyter Notebook** | Used for model testing and debugging. |
| **Postman** | Used for API testing. |

**6.2 System Test Plan**

The system test plan was designed to ensure the reliability, efficiency, and correctness of the developed system. The testing was carried out in the following phases:

**1.Unit Testing:**

Each module (churn prediction, sentiment analysis, recommendation system) was tested independently.

Focused on validating individual functions, such as data input handling, model inference, and API responses.

**2.Integration Testing:**

Ensured seamless interaction between different system components, such as frontend, backend, and database.

Verified API calls, data retrieval, and processing pipelines.

**3.System Testing:**

Conducted end-to-end testing to validate the system against the original requirements.

Included functional testing, performance testing, and security testing.

**4.User Acceptance Testing (UAT):**

The system was tested by potential users, such as bank analysts and administrators.

Feedback was collected and necessary refinements were made.

**6.3 Testing Approach**

The testing process was data-driven, ensuring that real-world banking and financial datasets were used to validate the accuracy and efficiency of the AI models. The approach consisted of:

**1.Dataset Used for Testing:**

**Customer Transaction Data**: Used for churn prediction by analyzing spending habits and frequency of transactions.

**Customer Feedback & Reviews**: Employed for sentiment analysis to categorize customer satisfaction levels.

**Historical User Interactions**: Used to train and test the recommendation system for personalized service suggestions.

**2.Testing Strategy**:

**Cross-validation**: Applied to ensure model generalization and avoid overfitting.

**Confusion Matrix & Classification Report**: Used to evaluate model performance.

**Load Testing**: Measured system response under varying data loads.

**Security Testing**: Checked for vulnerabilities in API calls and user authentication mechanisms.

**6.4 Proposed Change-over Techniques**

To transition from the existing system to the new AI-driven system, the following change-over techniques were considered:

**1. Parallel Running:**

Both the existing and new systems operate simultaneously for a given period.

Ensures smooth transition while users get accustomed to the new system.

**2. Phased Changeover**:

The system is implemented in stages, starting with one module at a time (e.g., sentiment analysis first, then churn prediction, and finally recommendations).

Reduces the risk of complete system failure.

**3. Pilot Implementation**:

The system is deployed in a single branch or a small user group before full deployment.

Helps in identifying issues before the full-scale rollout.

**4. Direct Changeover**:

The old system is completely replaced with the new one.

Requires thorough testing and user training before implementation.

**CHAPTER SEVEN: LIMITATIONS, CONCLUSIONS, AND RECOMMENDATIONS**

**7.1 Limitations**

During the development and implementation of the AI-Driven Churn Prediction, Sentiment Analysis, and Recommendation System for Banking and Financial Services, several challenges were encountered:

**Time Constraints**: The project had to be completed within a limited timeframe, which restricted extensive testing and optimization.

**Financial Limitations**: Due to budget constraints, access to premium machine learning tools, cloud services, and large-scale computational resources was restricted.

**Data Quality Issues**: Some datasets contained missing or inconsistent values, requiring extensive preprocessing before training the models.

**Limited User Cooperation**: Some stakeholders, particularly financial institutions, were reluctant to share sensitive data, limiting the breadth of real-world testing.

**Computational Resources**: Processing large datasets and training deep learning models required significant computational power, which was a challenge given hardware limitations.

**7.2 Conclusion**

The AI-Driven Churn Prediction, Sentiment Analysis, and Recommendation System was successfully designed and implemented to enhance customer retention and improve service delivery in banking and financial institutions. The system integrates machine learning techniques to analyze customer sentiments, predict potential churn, and provide personalized recommendations to customers. The study demonstrated that AI can significantly enhance decision-making in financial services by leveraging data-driven insights.

The results align with theoretical frameworks in machine learning and predictive analytics, validating the significance of AI in financial institutions. The system’s ability to process real-time customer feedback and provide actionable insights offers a competitive advantage to banks and financial organizations.

**7.3 Recommendations**

To enhance the effectiveness of the system, the following recommendations are suggested:

**Improved Data Collection:** Financial institutions should enhance their data collection strategies to ensure cleaner and more structured datasets for AI models.

**Integration with More Financial Services:** Future improvements should aim to integrate additional financial products such as fraud detection and credit risk analysis.

**Enhanced Computational Resources:** Using cloud-based AI solutions can help overcome computational limitations and improve system efficiency.

**User Training and Adoption**: Bank employees and analysts should be trained to understand and effectively use AI-driven insights for better decision-making.

**Security and Compliance Enhancements**: Ensuring that the system adheres to data protection laws and financial compliance standards will encourage wider adoption among institutions.