A green and white logo

Description automatically generated

**Database Design**

A black background with white text and red squares

Description automatically generated

Table of Contents

[**Introduction** 3](#_Toc151157513)

[**Solution 1: AI based solution.** 3](#_Toc151157514)

[**Solution 2: External AML vendor.** 7](#_Toc151157515)

# **Introduction**

The solution design plans to improve compliance at TD bank demand leveraging the use of data. Both the solutions, including the AI based solution as well as the external vendor ideas have been considered. This document explains the database design for the respective solutions.

# **Solution 1: AI based solution.**

The AI based solution will use a combination of Machine Learning Algorithms for various use cases across different aspects of Anti Money Laundering. The ER diagram primarily tackles the major aspects such as Transaction Monitoring, Alerts and Rapid Reporting and using generative AI technology for generating training data for identifying new problem statements and generating Training Simulations to train the employees to make accurate decisions and also reduce false positives. (Refer ER Diagram – Solution 1 visio file)

1. Primary Keys:

|  |  |
| --- | --- |
| **Table Name** | **Primary Key** |
| Customer | CustomerID |
| Transactions | TransactionID |
| Training | TrainingID |
| Alerts | AlertID |
| Investigation | InvestigationID |
| AuditLog | ReportID |
| AIModel | ModelID |
| SanctionList | SanctionID |

These primary keys uniquely identify each record in their respective tables. In relational database design, primary keys are crucial for maintaining data integrity and ensuring each record is uniquely identifiable within a table. The ranges are not specified yet in order to accommodate scalability after consultation with the Data Engineering Team which will happen during implementation.

2. How it is in Normalized Form?

Normalization is a process used in database design to organize data and minimize redundancy while maintaining data integrity. The normalization rules are typically applied to ensure that a database schema is in a specific normal form, such as the third normal form (3NF). Below is a documentation of how the provided tables have been normalized:

1. Alerts

Normalization Steps:

No repeated groups or arrays of data.

Each column stores a single piece of information.

The primary key (AlertID) uniquely identifies each record.

Normal Form: Already in 3rd Normal Form (3NF).

2. AI Model

Normalization Steps:

Removed redundant attributes and grouped related information together.

Each column stores a single piece of information.

The primary key (ModelID) uniquely identifies each record.

Foreign keys reference the primary keys in related tables.

Normal Form: Already in 3rd Normal Form (3NF).

3. Employee

Normalization Steps:

No repeated groups or arrays of data.

Each column stores a single piece of information.

The primary key (EmployeeID) uniquely identifies each employee record.

Normal Form: Already in 3rd Normal Form (3NF).

4. AuditLog

Normalization Steps:

Removed redundant attributes and grouped related information together.

Each column stores a single piece of information.

The primary key (ReportID) uniquely identifies each record.

Foreign keys reference the primary keys in related tables.

Normal Form: Already in 3rd Normal Form (3NF).

5. SanctionList

Normalization Steps:

No repeated groups or arrays of data.

Each column stores a single piece of information.

The primary key (SanctionID) uniquely identifies each record.

Foreign keys reference the primary keys in related tables.

Normal Form: Already in 3rd Normal Form (3NF).

6. Customer

Normalization Steps:

No repeated groups or arrays of data.

Each column stores a single piece of information.

The primary key (CustomerID) uniquely identifies each record.

Normal Form: Already in 3rd Normal Form (3NF).

7. Training

Normalization Steps:

No repeated groups or arrays of data.

Each column stores a single piece of information.

The primary key (TrainingID) uniquely identifies each record.

Foreign keys reference the primary keys in related tables.

Normal Form: Already in 3rd Normal Form (3NF).

8. Transactions

Normalization Steps:

No repeated groups or arrays of data.

Each column stores a single piece of information.

The primary key (TransactionID) uniquely identifies each record.

Foreign keys reference the primary keys in related tables.

Normal Form: Already in 3rd Normal Form (3NF).

# **Solution 2: External AML vendor.**

This solution will involve collaboration with external vendors who are leading in the industry space. A combination of vendor softwares will be used in conjunction with the bank’s data systems. This will require implementation and integration of the AML software. The major aspects covered will be Transaction monitoring and Regulatory reporting. (Refer ER Diagram – Solution 2 visio file)

1. **Purpose of each of the elements as part of the solution design**

The below listed tables and the associated data will help us in identifying any suspicious activity monitoring, alerting, investigating, and assessing the risk of the threat by cross checking the activity against the policies to be adhered.

|  |  |
| --- | --- |
| **Table Name** | **Description** |
| Customer | This table holds up all the **customer information** including the contact details and SIN number. |
| Accounts | This table gives the details about the **type of account and status of the account** i.e. Active / Inactive. |
| Transactions | This table investigates the transaction details of the customer and the **source of the transaction**. |
| Users | This table will help in identifying different roles of the user and to **implement authorization levels or access**. |
| Alerts | Alert table is used to **identify the type, severity, and the priority of the threat alert** to deal with a suspicious activity. |
| Investigation | The investigation table is used **to cross check the suspicious activity against the policies to identify the investigation status.** |
| Customer Alerts | CustomerAlerts table is a **reference table** that stores the primary keys of the customer table, transaction table and alert table so that these **accounts can be monitored regularly** for any suspicious activity. |
| Policies | The policies table has all the **policies and regulation to be covered under money laundering activities** according to the U.S and Canadian government. |
| Risk Assessment | The risk assessment table will **calculate the risk score** based with the help of alert severity and priority. |

1. **What are the range of values associated with each of the data elements?**

|  |  |
| --- | --- |
| **Data Fields** | **Range** |
| Alert Severity | Low, Medium, High, Critical |
| Alert Priority | Low, Medium, High, Immediate |
| Risk Score | 0 - 100 |

1. **What are the primary keys associated with each of these tables?**

|  |  |
| --- | --- |
| **Table Name** | **Primary Key** |
| Customer | Customer\_ID |
| Accounts | Account\_ID |
| Transactions | Transaction\_ID |
| Users | User\_ID |
| Alerts | Alert\_ID |
| Investigation | Investigation\_ID |
| Customer Alerts | CustomerAlerts\_ID |
| Policies | Policy\_ID |
| Risk Assessment | Assessment\_ID |

1. **Document how your tables have had normalization rules applied.**

Normalization is implemented in a database to reduce the redundancy and inconsistencies within the database**.** Normalizing databases helps in improving efficiency and preventing issues with dependencies**.** Some of the most common normalization forms are:

* First Normal Form (1NF)
* Second Normal Form (2NF)
* Third Normal Form (3NF)
* Boyce-Codd Normal Form (BCNF)
* Fourth Normal Form (4NF)
* Fifth Normal Form (5NF)

In the above-mentioned tables, we have implemented the normalization as follows:

1. All the tables have primary key to uniquely identify each row.
2. The tables don’t have any redundant fields.
3. None of the tables have transitive or multivalued dependencies.
4. All the tables meet the requirement for 1NF,2NF and 3NF.
5. **How will these data elements interrelate?**

Foreign keys are used within the database to establish relationships between the tables. A foreign key in a table is pointing to a primary key in another table. These relationships are used to increase the integrity of the tables. The foreign keys of the tables in this database are as follows:

|  |  |
| --- | --- |
| **Table Name** | **Foreign Key** |
| Accounts | Customer\_ID |
| Transactions | Account\_ID |
| Investigation | Alert\_ID, User\_ID |
| Customer Alerts | Customer\_ID, Transaction\_ID, Alert\_ID |
| Risk Assessment | Customer\_ID, Alert\_ID |