**1.0 Introduction**

**Project Overview**

A database design on a real-time predictive model for energy consumption in industrial

environment using machine learning with the following objectives:To develop and implement machine learning models for predicting energy consumption, evaluate the accuracy and efficiency of the models using real-world data while recommending strategies for optimizing energy consumption based on predictive model outputs. Also, developing and implementing a LSTM-based machine learning model for predicting industrial energy consumption using historical and real-time data.

**1.1 Purpose and scope of the database**

* Predict energy demand based on historical and real-time data, allowing for precise, data-backed decisions.
* Analyze patterns and factors that impact energy consumption, allowing managers to reduce unnecessary energy use, lower utility costs, and optimize equipment schedules.
* By minimizing energy waste, the model helps facilities reduce greenhouse gas emissions while complying with environmental regulations and sustainability goals.

**Scope of database includes:**

* Core Functionality: Support real-time data ingestion, historical storage, and quick access to predictions.
* Data Volume Handling: Implement scalable strategies for high-frequency data retention.
* Real-Time Performance: Optimize for low-latency queries for immediate insights.
* Security and Monitoring: Maintain data integrity with robust tracking and logging.
* UI Support:Integrate with dashboards and visualization tools for accessible insights.

**1.2 Problem Statement**

Energy consumption is a major cost and environmental issue in industrial settings. Predicting energy use is challenging due to factors like machine schedules and weather, leading to inefficiencies and higher costs.

The goal is to create a predictive model using historical and real-time data to enhance energy consumption forecasts. This will help facility managers make informed decisions, optimize energy use, and reduce costs.

Major Challenges faced include:

* Data quality and availability
* Data Integration and scalability
* Model selection and complexity
* Real time prediction and low latency requirement
* Evaluation and accuracy prediction

**1.3 Objectives**

The Objectives of designing a predictive model of energy consumption is as follows;

* Develop and Implement Predictive Models Utilize machine learning, particularly Long Short-Term Memory (LSTM) networks, to effectively forecast both short- and long-term energy consumption using historical and real-time data for improved energy management.
* Integrate External and Contextual Data : Boost model accuracy by incorporating essential data such as production rates, weather conditions, and equipment usage, creating a comprehensive view of energy consumption influences.
* Evaluate Model Performance: Assess model accuracy and efficiency using metrics like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to ensure robust performance in real-world conditions, even with large data streams.
* Real-Time Data Processing and Prediction: Implement a system that processes real-time data from industrial sensors, enabling immediate predictions and dynamic updates as new information arrives.
* Optimize Energy Consumption Through Actionable Recommendations: Generate actionable recommendations from predictive outputs, suggesting adjustments to equipment schedules or production planning to reduce energy use during peak demand.
* Scalable and Efficient Database Design: Create a scalable database architecture that manages high-frequency data ingestion, efficient storage, and real-time querying of sensor data and predictions.
* Develop a User Interface for Visualization and Insights: Design an intuitive dashboard for visualizing energy usage, predictions, and optimization suggestions, empowering facility managers to monitor trends and make informed decisions efficiently.

**2.0 Requirement Analysis**

Here are the key entities and attributes defined as follows;

* Facility ( A place where the energy is been consumed)
  + Facility ID ( Primary key) a unique identifier
  + Facility Type ( type of activities performed in the facility)
  + Facility Location
  + Date of establishment
* Sensor(Any device that receives signal)
  + Sensor ID(Primary Key) a unique identifier
  + Facility ID (foreign key referencing facility)
  + Sensor type (The type of data it collects e.g power, temperature etc)
  + Installation date (The actual date of installation)
  + Unit (unit of measurement e.g kWh, oC etc)
  + Status ( Current operational status either active or inactive)
* Sensor data ( Stores time-oriented data from sensor for training as well as predictions)
  + Sensor Data ID (Primary key)
  + Sensor ID (Foreign key references sensor)
  + Timestamp (date and time of reading)
  + Value (Recorded sensor value)
  + Processed ( collects a boolean value True/false)
* Energy Consumption History
  + Energy Consumption ID (Primary Key)
  + Facility ID (foreign key referencing facility)
  + Timestamp (Date and time of energy recorded)
  + Energy consumed( Amount of energy used)
  + Section (Facility section where energy was consumed)
  + Source Type (Type of equipment used)
* Production Log
  + Production Log ID (Primary Key)
  + Facility ID (foreign key referencing facility)
  + Timestamp ( Date and time of production log entry)
  + Machine ID ( Identifier of machine)
  + Status ( Current operational status either active or inactive )
* Weather Data
  + Weather ID ( Primary Key)
  + FacilityID (Foreign key references Facility)
  + Timestamp (Date and time of weather entry)
  + Temperature (Temperature in oc)
  + Humidity (Humidity in percentages)
  + Condition ( like Snowfall, rain, heat etc)
* Model
  + Model ID ( Primary key) a unique identifier
  + Model type ( e.g LSTM, SVM algorithm, random forest etc)
  + Training Date ( The date the model was trained)
  + Accuracy ( Accuracy measured in e.g MAE, RMSE)
  + Additional Notes ( An additional note on model performance)
* Prediction
  + Prediction ID (Primary Key)
  + Model ID (Foreign key referencing model
  + Timestamp ( Date and time of prediction)
  + Predicted energy( Forecasted energy consumption)
  + Prediction duration( Forecasted energy consumption time)
* Optimized Recommendation
  + Recommendation ID ( Primary Key)
  + Prediction ID(Foreign key references prediction)
  + Timestamp ( date and time of recommendation)
  + Recommendation Note ( Suggested Action to be taken)
  + Predicted savings ( Expected energy to be saved)
  + Priority level (Level of importance)

**3.0 Conceptual Design**

**Entities**

* Facility (A place where the energy is been consumed)
* Sensor ( A device that receives and keep track of signal
* Sensor Data( Data gotten from the sensor device)
* Energy consumption history
* Production Log
* Weather data
* Model
* Prediction
* Optimized Recommendation

**Relationships**

Facility and sensor: A facility can have one or many sensor devices all with different identification. One facility to many sensors which makes it a (1:N) one-to-many relationship.

Sensors and Sensor Data: Each sensor generates multiple data records over time. This one-to-many relationship allows for capturing detailed, time-stamped readings, which provide the basis for training and real-time model inputs.

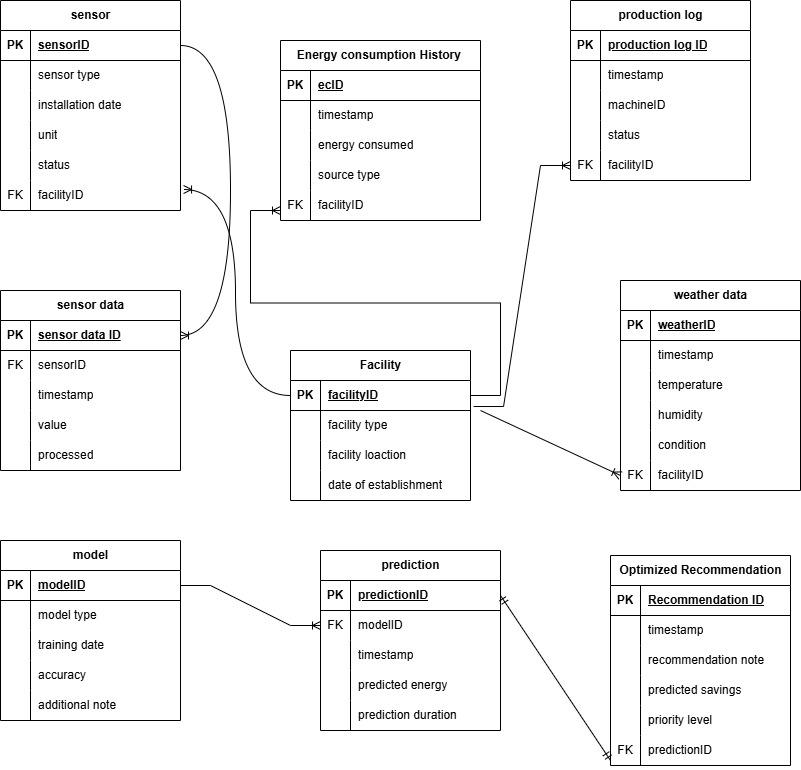
Energy Consumption History and Prediction: By comparing predictions to historical data, the model can evaluate its accuracy and refine future predictions.

Prediction and Model: Linking predictions to specific model versions allows analysts to trace prediction results back to the model version used, facilitating version control and tracking model evolution over time.

Production Log and Energy Consumption History: Production logs often correlate directly with energy usage, as changes in production status or output can lead to variations in consumption.

Optimized Recommendation and Prediction: Recommendations are based on predictions. A strong relationship between these entities enables effective follow-up actions and the potential for continuous feedback based on results.

**ER Diagram for a real-time predictive model for energy consumption**

****

**4.0 Logical Design**

The logical design phase involves converting the conceptual design into a detailed relational schema. This schema defines the tables, their fields, data types, and relationships between them. We’ll also normalize the tables to ensure the database is efficient, avoids redundancy, and maintains data integrity.

Relational Schemas

The relational schema outlines how data will be stored in tables. Here are the main tables derived from the entities in the conceptual design:

* Facility Table
  + FacilityID: INT (Primary Key)
  + Facility Type: VARCHAR(50)
  + Facility Location: VARCHAR(255)
  + Date of Establishment: DATETIME
* Sensor Table
  + SensorID: INT (Primary key)
  + FacilityID: INT (Foreign Key)
  + Installation Date: DATETIME
  + Unit: VARCHAR(5)
  + Status: ENUM(‘Active’, ‘Inactive’)
* Sensor Data Table
  + Sensor Data ID: INT (Primary Key)
  + SensorID: INT(Foreign Key)
  + Timestamp: DATETIME
  + Value: VARCHAR(20)
  + Processed: BOOLEAN
* Energy Consumption history Table
  + ecID: INT (Primary Key)
  + FacilityID: INT (Foreign Key)
  + Timestamp: DATETIME
  + Energy Consumed: VARCHAR(20)
  + Source Type: VARCHAR(50)
  + Session: VARCHAR(255)
* Production Log Table
  + Production Log ID: INT(Primary Key)
  + FacilityID: INT (Foreign Key)
  + Timestamp: DATETIME
  + MachineID: VARCHAR(50)
  + Status: ENUM(‘Active’, ‘Inactive’)
* Weather Data Table
  + WeatherID: INT(Primary Key)
  + FacilityID: INT (Foreign Key)
  + Timestamp: DATETIME
  + Temperature: INT
  + Condition: VARCHAR(50)
* Model Table
  + ModelID: INT (Primary key)
  + Model Type: VARCHAR(50)
  + Training Date: DATE
  + Accuracy: INT
  + Additional Note: VARCHAR(50)
* Prediction Table
  + PredictionID: INT (Primary Key)
  + ModelID: INT (Foreign key)
  + Timestamp: DATETIME
  + Predicted Energy: INT
  + Predicted duration: VARCHAR(50)
* Optimized Recommendation Table
  + RecommendationID: INT (Primary Key)
  + PredictionID: INT (Foreign key)
  + Timestamp: DATETIME
  + Recommendation Note: VARCHAR(255)
  + Predicted savings: INT
  + Priority Level: INT

#### Tables, Fields, and Data Types

* Facility Table:
  + Fields: Facility Type, Facility Location, Date of Establishment
  + Data Types:
    - FacilityID: INT (Primary Key, Auto increment)
    - Facility Type: VARCHAR(50)
    - Facility Location: VARCHAR(255)
    - Date of Establishment: DATETIME

| Facility | Data Type |
| --- | --- |
| Facility ID | INT (Primary Key, Auto increment) |
| Facility Type | VARCHAR(50) |
| Facility location | VARCHAR(255) |
| Date of Establishment | DATETIME |

* Sensor Table:
  + Fields: SensorID, FacilityID, Installation Date, Unit, Status
  + Data Types:
    - SensorID: INT (Primary key, Auto Increment)
    - FacilityID: INT (Foreign Key)
    - Installation Date: DATETIME
    - Unit: VARCHAR(5)
    - Status: ENUM(‘Active’, ‘Inactive’)

| Sensor | Data Type |
| --- | --- |
| SensorID | INT (Primary key, Auto increment) |
| Facility ID | INT (Foreign Key) |
| Installation Date | DATETIME |
| Unit | VARCHAR(5) |
| Status | ENUM(‘Active’, ‘Inactive’) |

* Sensor Data Table:
  + Fields: Sensor Data ID, Timestamp, Value, Processed
  + Data Types:
    - Sensor Data ID: INT (Primary Key, Auto increment)
    - SensorID: INT (Foreign Key)
    - Timestamp: DATETIME
    - Value: VARCHAR(20)
    - Processed: BOOLEAN

| Sensor Data | Data Type |
| --- | --- |
| Sensor Data ID | INT (Primary Key, Auto increment) |
| SensorID | INT (Foreign Key) |
| Timestamp | DATETIME |
| Value | VARCHAR(20) |
| Processed | BOOLEAN |

* Energy Consumption history Table:
  + Fields: ecID, FacilityID, Timestamp, Energy Consumed, Source Type, Session
  + Data Types:
    - ecID: INT (Primary Key)
    - FacilityID: INT (Foreign Key)
    - Timestamp: DATETIME
    - Energy Consumed: VARCHAR(20)
    - Source Type: VARCHAR(50)
    - Session: VARCHAR(255)

| Energy Consumption | Data Type |
| --- | --- |
| ecID | INT (Primary Key) |
| FacilityID | INT (Foreign Key) |
| Timestamp | DATETIME |
| Energy Consumed | VARCHAR(20) |
| SourceType | VARCHAR(50) |
| Session | VARCHAR(255) |

* Production Log Table:
  + Fields: Production Log ID, FacilityID, Timestamp, MachineID, Status
  + Data Types:
    - Production Log ID: INT(Primary Key)
    - FacilityID: INT (Foreign Key)
    - Timestamp: DATETIME
    - MachineID: VARCHAR(50)
    - Status: ENUM(‘Active’, ‘Inactive’)

| Production Log | Data Type |
| --- | --- |
| Production Log ID | INT (Primary Key) |
| FacilityID | INT (Foreign Key) |
| Timestamp | DATETIME |
| MachineID | VARCHAR(50) |
| Status | ENUM(‘Active’,’Inactive’) |

* Weather Data Table:
  + Fields: WeatherID, FacilityID, Timestamp, Temperature, Condition
  + Data Types:
    - WeatherID: INT(Primary Key)
    - FacilityID: INT (Foreign Key)
    - Timestamp: DATETIME
    - Temperature: INT
    - Condition: VARCHAR(50)

| Weather Data | Data Type |
| --- | --- |
| WeatherID | INT (Primary Key) |
| FacilityID | INT (Foreign Key) |
| Timestamp | DATETIME |
| Temperature | INT |
| Condition | VARCHAR(50) |

* Model Table:
  + Fields: Model ID, Model Type, Training Date, Accuracy, Additional Notes.
  + Data Types:
    - ModelID: INT (Primary key)
    - Model Type: VARCHAR(50)
    - Training Date: DATE
    - Accuracy: INT
    - Additional Note: VARCHAR(50)

| Model | Data Type |
| --- | --- |
| ModelID | INT (Primary Key) |
| Model Type | VARCHAR(50) |
| Training Date | DATE |
| Accuracy | INT |
| Additional Note | VARCHAR(50) |

* Prediction Table:
  + Fields: PredictionID, ModelID, Timestamp, Predicted Energy, Predicted Duration.
  + Data Types:
    - PredictionID: INT (Primary Key)
    - ModelID: INT (Foreign key)
    - Timestamp: DATETIME
    - Predicted Energy: INT
    - Predicted duration: VARCHAR(50)

| Prediction | Data Type |
| --- | --- |
| PredictionID | INT (Primary Key) |
| ModelID | INT (Foreign Key) |
| Timestamp | DATETIME |
| Predicted Energy | INT |
| Predicted duration | VARCHAR(50) |

* Optimized Recommendation Table:
  + Fields: RecommendationID, PredictionID, Timestamp, Recommendation Note, Predicted savings, Priority Level
  + Data Type:
    - RecommendationID: INT (Primary Key)
    - PredictionID: INT (Foreign key)
    - Timestamp: DATETIME
    - Recommendation Note: VARCHAR(255)
    - Predicted savings: INT
    - Priority Level: INT

| Optimized Recommendation | Data Type |
| --- | --- |
| RecommendationID | INT (Primary Key) |
| PredictionID | INT (Foreign Key) |
| Timestamp | DATETIME |
| Recommendation Note | VARCHAR |
| Predicted Savings | INT |
| Priority Level | INT |