SQL Queries and Machine Learning Models

Today, I explored key **SQL commands** used in databases for data retrieval and manipulation, along with the use of **machine learning models** to compare actual versus predicted values.

SQL Commands:

1. **SELECT**:
    - The **SELECT** statement is used to **fetch data** from a database. The data returned is stored in a result table, known as the result set.
    - Example:

      ```
      sql
      Copy code
      SELECT * FROM users;
      ```

2. **FROM**:
    - The **FROM** clause specifies **the table** from which the data will be retrieved.
    - Example:

      ```
      sql
      Copy code
      SELECT name, age FROM users;
      ```

3. **WHERE**:
    - The **WHERE** clause is used to **filter records** based on a condition.
    - Example:

      ```
      sql
      Copy code
      SELECT * FROM users WHERE age > 30;
      ```

4. **OFFSET**:
    - The **OFFSET** keyword is used to **skip rows** before starting to return the rows.
    - Example:

      ```
      sql
      Copy code
      SELECT * FROM users ORDER BY name OFFSET 5;
      ```

5. **LIMIT**:
    - The **LIMIT** clause is used to **set a maximum number of rows** that will be returned.
    - Example:

      ```
      sql
      Copy code
      SELECT * FROM users LIMIT 10;
      ```

6. **GROUP BY**:
    - The **GROUP BY** clause is used to group rows that have the same values into summary rows, often used with aggregate functions like `COUNT()`, `SUM()`, `AVG()`.
    - Example:

```
sql
Copy code
SELECT department, COUNT(*) FROM employees GROUP BY department;
```

These SQL commands form the basis of interacting with **tables** (which are collections of data in rows and columns) and **attributes** (which are the columns representing the characteristics of the dataset).

Using Machine Learning Models to Compare Actual vs. Predicted Values:

In **machine learning**, models are built to make predictions based on patterns in data. Once a model is trained, it's crucial to evaluate its performance by comparing **actual** versus **predicted values**. This can be done using:

1. **Confusion Matrix**: For classification tasks, a confusion matrix helps visualize the number of correct and incorrect predictions.
2. **Mean Squared Error (MSE)**: For regression tasks, MSE calculates the average squared difference between actual and predicted values, indicating the model’s accuracy.