Water Quality Prediction using RNN:

machine learning model using a Recurrent Neural Network (RNN) that can predict water quality parameters for the next 5 days based on historical data.

# Dataset:

- File: trend\_water\_quality.csv

- Features: Multiple water quality indicators (e.g., pH, Dissolved Oxygen, BOD, etc.)

- Note: The Date column is dropped for model training since we only use numerical values.

# Preprocessing Steps:

1. Drop Date Column: If present, we drop it.

2. Drop Missing Values: Ensures clean data for training.

3. Normalization:

- Scales all feature values between 0 and 1 using MinMaxScaler.

- Necessary for RNNs to ensure stable and efficient training.

# Data Preparation for Time Series:

We prepare the dataset using sliding windows:

- look\_back = 5: Use the past 5 days' data as input.

- look\_forward = 5: Predict the next 5 days.

This results in:

- X.shape: (samples, 5, number\_of\_features)

- y.shape: (samples, 5 \* number\_of\_features)

# Model Architecture:

We use a Sequential RNN model:

- SimpleRNN(64): A simple recurrent layer with 64 units.

- Dense layer: Output layer that flattens predictions for 5 future time steps.

# Model Compilation:

- Loss: Mean Squared Error (MSE)

- Metric: Mean Absolute Error (MAE)

# Model Training:

- Epochs: 30 passes through the dataset.

- Batch size: 16 samples per update.

- Validation split: 10% of the training data used for validation.

# Model Evaluation:

- Reports MSE and MAE for the test dataset.

# Model Saving:

- Saves the trained model and the scaler for future use.

# Prediction Script (Using Sample Input):

1. Load the last 5 days of data from sample\_5\_days\_input.csv

2. Normalize using the same scaler.

3. Predict the next 5 days.

4. Inverse transform the results to get actual values.

5. Save predictions to predicted\_next\_5\_days.csv

# Output:

- Predictions for next 5 days saved in predicted\_next\_5\_days.csv

- Each row represents predicted water quality values for a future day.