**1. Data Loading and Initial Processing**

You began by loading a CSV file named TASK-ML-INTERN.csv. This dataset seems to include various features and a target variable called vomitoxin\_ppb. Here are the initial steps taken:

* **Reading the CSV file**: The dataset is loaded into a pandas DataFrame.
* **Displaying initial data**: The first few rows (data.head()) and summary statistics (data.describe()) are displayed to get an overview of the data.

**2. Data Standardization**

* **Dropping non-feature columns**: Columns hsi\_id and vomitoxin\_ppb are dropped from the features set.
* **Standardization**: Feature values are standardized using StandardScaler, which normalizes the feature data to have a mean of zero and a standard deviation of one.

**3. Visualization of Spectral Reflectance**

* **Calculating average reflectance**: You calculated the mean of standardized features across each wavelength band.
* **Plotting**: A line plot shows how average spectral reflectance varies across different bands.

**4. Principal Component Analysis (PCA)**

* **Applying PCA**: PCA is performed to reduce dimensionality while retaining 95% of the variance in the dataset.
* **Explained Variance Plot**: A plot displaying the cumulative variance explained by the PCA components was created, showing the number of components needed to retain 95% of the variance.

**5. 3D Visualization of PCA Results**

* **3D Scatter Plot**: The first three principal components are used to create a 3D scatter plot, color-coded by the vomitoxin\_ppb values, providing insights into the data’s structure and potential clustering.

**6. Normalization of Target Variable**

* **MinMax Scaling**: The target variable vomitoxin\_ppb is normalized using MinMaxScaler.

**7. Preparation for LSTM Model**

* **Data Splitting**: The dataset is split into training (80%) and testing (20%) sets.
* **Reshaping**: The feature matrices are reshaped to fit LSTM input requirements.

**8. LSTM Model Training**

* **Model Definition**: An LSTM model is defined and trained to predict vomitoxin\_ppb.
* **Model Training**: The model is trained, and the training/validation loss over epochs is plotted to assess model performance.

**9. Prediction and Evaluation**

* **Prediction**: The trained LSTM model is used to predict the vomitoxin\_ppb values on the test set.
* **Performance Metrics**: Mean Absolute Error, Root Mean Squared Error, and R² Score are computed to evaluate the model's accuracy.

**10. Detailed Reporting**

* **Interpretation**: Each step's output needs to be carefully interpreted to assess the model's performance, feature importance, and the potential need for further tuning or data preprocessing.
* **Final Report**: Summarize the entire workflow, findings, and recommendations for future work or improvements.