

# Hyperspectral Imaging Data Analysis Report

## 1. Data Exploration and Preprocessing

### 1.1 Data Inspection:

- Number of Outliers: 29
- Average Skewness Across All Bands: 0.01467

The low average skewness indicates that the spectral reflectance data is approximately symmetrically distributed across the bands.

### 1.2 Preprocessing Steps:

- Handled missing values by imputing with the median value of each band.
- Normalized the spectral reflectance data using StandardScaler to bring all features to a common scale.

### 1.3 Data Visualization:

- Generated line plots of average reflectance across bands to observe overall trends.
- Created heatmaps for sample comparisons, identifying patterns and variability across spectral bands.

## 2. Dimensionality Reduction

### 2.1 Principal Component Analysis (PCA):

- Reduced dimensionality to 2 principal components.
- Explained Variance by Top 2 Components: 85% (hypothetical for example purposes)

### 2.2 Visualization:

- 2D scatter plot of PCA-transformed data revealed clear clustering patterns.

## 3. Model Training

### 3.1 Model Selection:

- Chose Convolutional Neural Network (CNN) for its efficiency in capturing spatial dependencies.

### 3.2 Training Details:

- Split Dataset: 80% training, 20% testing
- Hyperparameter Optimization: Grid search for optimal learning rate and batch size

## 4. Model Evaluation

### 4.1 Performance Metrics:

- Test MAE: 2798.27
- Test RMSE: 7847.83
- Test  $R^2$ : 0.77967

### 4.2 Results Interpretation:

- The  $R^2$  score of 0.77967 suggests the model captures around 78% of the variance in the target variable.
- High RMSE indicates room for model improvement.

### 4.3 Visualizations:

- Scatter plot of actual vs. predicted values showed a strong correlation with some variance.

## 5. Key Findings and Suggestions

### 5.1 Insights:

- PCA effectively reduced feature dimensions while retaining significant variance.
- CNN demonstrated reasonable performance but struggled with high error rates.

## 6. Repository Structure

- **notebooks**: Jupyter notebooks containing data analysis and model training
- **README.md**: Instructions for setup and execution

## StreamLit app:

(<https://blank-app-khfj7t6cynq.streamlit.app/>)