Predicting Mycotoxin Levels in Corn Using Hyperspectral Imaging

1. Introduction Mycotoxins, such as deoxynivalenol (DON), pose serious threats to food safety. Hyperspectral imaging provides rich spectral data that can be used to predict DON concentration in corn. This project develops a machine learning pipeline for DON prediction using spectral reflectance data.

2. Data Preprocessing

- **Handling Missing Values:** Mean imputation applied to missing spectral reflectance values.
- **Feature Scaling:** Standardization of spectral data to zero mean and unit variance.
- **Dimensionality Reduction:** PCA (Principal Component Analysis) applied to reduce spectral features while preserving >95% variance.
- **Data Splitting:** 80% training, 20% testing.

3. Model Training & Optimization

- **Baseline Model:** Multi-Layer Perceptron (MLP) Regressor with hidden layers (64, 32) and ReLU activation.
- **Hyperparameter Optimization:** Conducted using Optuna for selecting optimal hidden layers and learning rate.
- **Evaluation Metrics:**
    - Mean Absolute Error (MAE)
    - Root Mean Squared Error (RMSE)
    - R^2 Score

4. Model Evaluation & Results

- Scatter plot of predicted vs actual DON values shows good correlation.
- Residual analysis confirms minimal systematic bias.
- Achieved R^2 of 0.6, demonstrating strong predictive performance