Summary of Feature Correlation Analysis

Background:
The exercise aimed to identify the most correlating features for a given dataset with multiple classes. By examining correlations on a per-class basis, the analysis sought to determine if the most correlating features for individual classes differed from those correlating with all classes collectively.

Methodology:
1. **Data Overview:** The dataset consisted of features like 'RI', 'Na', 'Mg', 'Al', 'Si', 'K', 'Ca', 'Ba', 'Fe', with 'Type' being the target variable.
2. **Feature Correlation:**
   - A binary representation was used for each class in the target variable 'Type' to find the most correlated features.
   - The top 2 correlating features were identified for each class.

Key Findings:
1. Correlating Features per Class:
   a. Type 1: 'Al' and 'Ba'
   b. Type 2: 'Mg' and 'RI'
   c. Type 3: 'Fe' and 'Mg'
   d. Type 5: 'Mg' and 'Na'
   e. Type 6: 'Mg' and 'Na'
2. The feature 'Mg' emerged as a consistent top correlator across multiple classes.
3. When correlations were examined on a per-class basis, it became evident that the most correlating features for individual classes differed from those correlating with all classes collectively.

Conclusion:
Relying solely on overall feature correlation might not yield the most relevant features for specific classes. This observation underscores the importance of a granular, class-specific analysis. By examining correlations on a per-class basis, more nuanced and potentially impactful insights can be derived, allowing for better feature selection and, potentially, improved model performance.

Link to the google colab:
https://colab.research.google.com/drive/1hp1MXhVYhXZ_rPHplhNx3WVzdCiSoakX?usp=sharing