Theory of Machine Learning (ML) and Deep Learning (DL)

Machine Learning: Random Forest Classifier

Random Forest Classifier:
- A Random Forest is an ensemble learning method used for classification and regression.
- It operates by constructing multiple decision trees during training and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.
- Key characteristics:
  - Ensemble of Trees: Combines multiple decision trees to improve model accuracy and robustness.
  - Randomness: Each tree is trained on a random subset of the data, and within each tree, splits are chosen from a random subset of features.
  - Overfitting Reduction: By averaging the results, it reduces the risk of overfitting compared to individual decision trees.

Advantages:
- Robust to overfitting due to averaging of multiple trees.
- Handles large datasets with higher dimensionality.
- Provides feature importance, which helps in feature selection.

Disadvantages:
- Can be computationally intensive due to the large number of trees.
- Model interpretability is lower compared to single decision trees.

Deep Learning: Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN):
- A CNN is a class of deep neural networks, most commonly applied to analyze visual imagery.
- It is composed of layers that apply convolution operations to the input, using filters (also known as kernels) to detect patterns and features.
- Key layers in a CNN:
  - Convolutional Layer: Applies convolution operations to extract features from the input data.
  - Pooling Layer: Performs down-sampling to reduce the dimensionality and retain important features.
  - Flatten Layer: Converts the 2D matrix data to a 1D vector to be fed into a fully connected layer.
  - Fully Connected Layer: Performs classification based on the extracted features.
  - Dropout Layer: Randomly sets a fraction of input units to 0 during training to prevent overfitting.

Advantages:
- Excellent performance in image classification tasks.
- Automatically detects important features without manual feature extraction.
- Can handle large-scale datasets.

Disadvantages:
- Requires a large amount of data for training.
- Computationally intensive and requires significant processing power (GPUs).

Confusion Matrix

A Confusion Matrix is a performance measurement for classification problems. It is a table with four different combinations of predicted and actual values. It allows the visualization of the performance of an algorithm.

Components:
- True Positives (TP): Correctly predicted positive observations.
- True Negatives (TN): Correctly predicted negative observations.
- False Positives (FP): Incorrectly predicted positive observations (Type I error).
- False Negatives (FN): Incorrectly predicted negative observations (Type II error).

Metrics Derived from the Confusion Matrix:

- Accuracy: (TP + TN) / (TP + TN + FP + FN)
  - The overall correctness of the model.
- Precision: TP / (TP + FP)
  - The accuracy of the positive predictions.
- Recall (Sensitivity): TP / (TP + FN)
  - The ability of the model to find all the relevant cases within a dataset.
- F1 Score: 2 * (Precision * Recall) / (Precision + Recall)
  - The harmonic mean of precision and recall, providing a balance between the two.

Visualizing Confusion Matrix
Visualizing the confusion matrix helps to understand the performance of a classifier by showing where the model gets confused and makes incorrect predictions. It is particularly useful for multi-class classification problems to analyze which classes are being misclassified and why.

Application in EEG Analysis
In the provided script, both Random Forest and CNN classifiers are applied to classify EEG data.

1. Random Forest:
   - Used to classify the features extracted from the power spectral density (PSD) of EEG epochs.
   - Performance is evaluated using classification reports and confusion matrices.

2. Convolutional Neural Network (CNN):
   - Used to classify EEG data directly from the time-series information of the epochs.
   - Performance is evaluated similarly with classification reports and confusion matrices.

The confusion matrices for both models are visualized to understand their performance in distinguishing between different classes of EEG data. This approach helps in identifying the strengths and weaknesses of each classifier in the context of EEG signal classification.