EEG Data Analysis Case Study

# Introduction

This case study focuses on the analysis and preprocessing of an Electroencephalogram (EEG) dataset. EEG datasets are complex and contain valuable information about the brain's electrical activity. The primary goal is to prepare the dataset for further analysis or machine learning modeling.

# Objective

The main objectives of this case study are to perform exploratory data analysis (EDA) to understand the dataset's characteristics and to apply necessary preprocessing steps to clean and normalize the data. These steps are essential for ensuring that the dataset is suitable for building reliable and accurate models.

# Dataset Description

The EEG dataset contains several features that represent different aspects of brain activity, measured through EEG. The features include measures of attention, meditation, and various brain wave types (delta, theta, lowAlpha, highAlpha, lowBeta, highBeta, lowGamma, highGamma). The dataset also includes a 'classification' column, which might represent a binary target for classification tasks.

# Tasks

1. Exploratory Data Analysis (EDA):

* a. Check for missing values across all columns.
* b. Visualize the distribution of key features including EEG measures and the target variable.
* c. Analyze the balance of the classification target variable.

2. Preprocessing:

* a. Normalize or standardize the EEG feature columns to ensure they are on a similar scale.
* b. Identify and handle any outliers in the dataset, if necessary.

# SQL Preprocessing Replication

Replicate the preprocessing steps performed in Python in an SQL environment. Focus on applying the normalization or standardization process for each feature column in SQL. Provide SQL queries that demonstrate how to calculate and apply these preprocessing steps directly in an SQL database.