# Project Plan: Epileptic Seizure Detection Based On EEG Signals and Machine Learning

Martin Nguyen, s194378

September 2022

## Introduction

Today, 50 million people worldwide suffer from epilepsy and for some, epileptic seizures (ES) are crippling and interfere with day-to-day activities [1]. Epilepsy is a neurological disorder in which abnormal brain activity results in seizures or periods of unusual behavior, sensations, and loss of awareness [2]. Electroencephalography (EEG) is an efficient diagnostic tool for examining the functional anatomy of the brain and analysis of the EEG signals is the primary clinical method to identify ES activities in the brain. Experienced specialists can detect disorders simply by visually observing the EEG signals. However, due to the high variability across patients caused by the temporal and spatial aspects of the dynamic non-linear EEG data, this procedure is time-consuming and prone to error. Therefore, analysis and feature extraction of EEG signals is an essential part of automated epileptic seizure detection. A review of the state-of-the-art techniques for the classification of ES based on ictal and inter-ictal states of the EEG signals was published in 2020 [3]. The scientific papers included in this review used various feature extraction methods, e.g. Independent Component Analysis (ICA), Fourier Transformation (FT) or Entropy, and machine- (ML) and deep learning (DL) models such as Support Vector Machine (SVM) or Convolutional Neural Network (CNN).

## Project Purpose

The purpose of the project will be to investigate and deploy different advanced machine learning models on EEG signals from real epileptic patients. This project will primarily focus on a binary classifier that can detect epileptic seizures based on whether a given EEG signal is in its ictal or inter-ictal state. Another important aspect of such a task is of course the feature extraction of the EEG signals. However, due to time constraints a dimensionality reduction method will probably only be applied. The goal of this project is to have a better-performing machine learning model with the chosen feature extraction technique. Hopefully, this project will become a stepping stone for practical use.

## Methodology

First, the EEG signals obtained from the CHB-MIT database will be preprocessed using the Python library MNE. The CHB-MIT database consists of recordings from 22 subjects grouped into 23 cases; 5 males, ages 3-22; and 17 females, ages 1.5-19. The sample rate of all signals is 256 Hz and each case have recordings ranging from 1-4 hours and contain between 9-42 continuous .edf files from a single subject. The feature extraction used for this project will be a dimensional reduction method such as PCA, LDA, or t-SNE. The advanced machine learning models will probably be some sort of supervised learning algorithms such as Logistic Regression, Naïve-Bayes, or K-NN. The models will be evaluated with cross-validation and the performance metrics used are F1 score, confusion matrix, and specificity/sensitivity.

If time allows, this project will also look at different feature extraction methods such as Fourier Transformation or non-linear methods e.g. Kolmogorov entropy.

## Applied Techniques

- Detailed literature study on epilepsy and EEG signals and the state-of-the-art preprocessing and feature extraction techniques thereof.
- Preprocessing, feature extraction and machine learning models will be done on the CHB-MIT dataset
- Python library like MNE to extract and preprocess EEG signals and SciKit-Learn and TensorFlow for machine- and deep learning purposes.

Since training various machine- and deep learning models will be computationally expensive, this project will make use of the available HPC clusters offered by DTU for this purpose.

## Project Risks

Since this project will make use of an already existing and well-tested database, there is minimal risk associated with the project in itself. However, the value of the project outcome would be greatly diminished if there are issues with data and the chosen feature extraction methods. If the performance is terrible too this could also affect the project outcome.

## Learning Objective

The learning objective that follows from working on this project is of course knowing that the signals are retrieved from actual epileptic patients. Furthermore, advanced machine learning methods that have been discussed theoretically will be applied practically. The outcome of this project is hoping that it will become applicable as an aid for health experts.

## Time Plan

- **September:** Become familiar with MNE-library for EEG preprocessing (Python) along with understanding the underlying theory. Preprocessing includes loading, understanding the data structure, and filtering for noise and artifacts if needed.
- **October:** Apply feature extraction on preprocessed EEG signals and prepare train-test data for various ML/DL models.
- **November:** Train and evaluate models and report writing.

# References


[1] - WHO 2019, Epilepsy: A Public Health Imperative - Summary

[2] - Mayo Clinic 2021, Epilepsy

[3] - Rasheed, K. et al 2020, Machine Learning For Predicting Epileptic Seizures Using EEG Signals: A Review

[4] Shoeb, A. 2009, Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment

[5] Shoeb, A. and Guttag, J. 2010, Application of Machine Learning To Epileptic Seizure Detection

[5] Gramfort, A. et al. 2013, MEG and EEG Data Analysis with MNE-Python

[7] Pedregosa, F. et al. 2011, Scikit-learn: Machine Learning in Python

[8] Abadi, M. et al. 2015, TensorFlow: Large-scale Machine Learning On Heterogenous Systems