## Project Title:

Efficient and Generalizable Seizure Detection Algorithm

## Related Keywords:

Seizure Detection, Time Series, EEG, Detach-ROCKET, Epilepsy

## Introduction:

Epilepsy is a neurological disorder affecting millions worldwide. Developing reliable and efficient seizure detection methods is crucial for diagnosing and managing epilepsy. Current methodologies are often tailored for one particular experimental setting, often struggling with generalization to new datasets, leading to performance degradation. Addressing this challenge is essential for developing a truly robust detection algorithm that can be used in real-world applications, improving patient outcomes and reducing the time-intensive task of manual EEG annotation [1].

Accurate scalp EEG-based seizure detection algorithms serve two main purposes:

- **Improve Diagnosis**: Detection algorithms can optimize and facilitate the diagnostic workup performed on people with epilepsy to improve patients' care and quality of life [1]. The automatic annotation reduces the need for expert manual annotation and can potentially improve its accuracy.
- **Live Monitoring**: Detection algorithms enable real-time monitoring of epileptic patients, which is essential for immediate intervention. Efficiency is critical in this context as the algorithm must operate in real time without significant computational overhead.

Recent advancements in machine learning for time series analysis provide new opportunities for EEG analysis. One particularly interesting algorithm is Detach-ROCKET (Random Convolutional Kernel Transform), an extension of the ROCKET framework refined through Sequential Feature Detachment (SFD) that has demonstrated exceptional performance in terms of efficiency and accuracy on time-series classification [2]. By leveraging a large ensemble of random convolutional kernels, Detach-ROCKET extracts discriminative features from time-series data without extensive preprocessing while significantly reducing the number of features and enhancing model interpretability. These advancements position Detach-ROCKET as a promising tool for overcoming the challenges of seizure detection.

## Research Questions:

What is the maximum achievable accuracy for seizure detection algorithms while maintaining efficiency for practical use? To what extent does the model's accuracy decrease when applied to novel subjects and EEG recordings collected under different experimental conditions (i.e., datasets not used during training), and how does this impact its robustness and reliability in diverse clinical environments?

## Hypothesis:

Due to their remarkable trade-off between accuracy and efficiency, convolutional time series algorithms, such as those based on ROCKET, are expected to deliver state-of-the-art performance in seizure detection. Their highly parallelizable architecture enables significantly faster processing compared to traditional feature extraction methods, making them ideal for real-time and large-scale clinical applications. The core hypothesis of this thesis is that their speed advantage, combined with their high accuracy, can effectively bridge the gap between machine learning advancements and clinical adoption, thereby enhancing the reliability and scalability of EEG-based seizure detection in practical settings.

## Project Description:

The objective of this thesis is to assess whether new time series algorithms, such as Detach-ROCKET, can provide meaningful insights into seizure detection by achieving efficiency and generalizability.

The datasets we work with are:

1. **CHB-MIT Scalp EEG Database:** Composed of 24 different patients, 982 hours of recordings, and 198 seizures episodes [3].
2. **TUH EEG Sz Corpus:** Composed of 675 different patients, 1476 hours of recordings, and 4029 seizures episodes [4].
3. **Siena Scalp EEG Database:** Composed of 14 different patients, 128 hours of recordings, and 47 seizures episodes [5].

We will first implement a traditional seizure detection algorithm such as discrete wavelet transforms (DWT) combined with support vector machines (SVM). This model will serve as a benchmark to compare against and test how difficult the segmentation task is. Next, we will implement a convolutional approach, such as ROCKET, to explore these datasets and evaluate whether the extracted features can effectively distinguish seizure events. Next, we will further implement Detach-ROCKET, to enhance prediction performance and select and study the features that are really important for prediction.

After implementing these algorithms, we will evaluate their generalizability to new patients within the same dataset (with leave-one-subject out cross-validation) and, more importantly, their generalizability across different datasets (with leave-one-dataset out cross-validation).

The dataset formats (for both input and output), evaluation methodology, and performance metrics we will use for this thesis will obey the recently proposed SzCORE standards [6], which are also used in the Seizure detection challenge (2025) [7]. This will enable a fair comparison of our results to other methodologies.

Additional tasks, if time permits, will include exploring alternative approaches to improve the accuracy and the output of additional information from the model, such as event probability or event location (primary seizure channels) [8].

## Proposed Plan:

1. **Literature Review:** Review the relevant literature on time-series classification, EEG seizure detection, and convolutional algorithms like ROCKET.
2. **Dataset Preparation:** Download the EEG datasets and implement a pipeline to convert them into the standard SzCORE format.
3. **Benchmark Implementation:** Implement a traditional seizure detection method using DWT for feature extraction and SVM for classification. Assess baseline performance.
4. **ROCKET Implementation:** Implement a standard ROCKET model for feature extraction and classification of the EEG data.
5. **Detach-ROCKET Implementation:** Implement the Detach-ROCKET model for improved performance and feature selection. Evaluate the impact on performance of different hyperparameters, mainly the ones related to the data (such as window length and normalization).
6. **Segmentation model heuristics:** Design the meta-algorithm to segment the data given the output of the seizure detection algorithm. Evaluate the impact on model performance of strategies as: changing the event threshold, ignoring short events or merging consecutive events that are close in time.
7. **Comprehensive Performance Evaluation:**
    - Metrics: Sensitivity, specificity, accuracy, ROC-AUC, and confusion matrix.
    - Generalization: Evaluate intra-dataset and inter-dataset performance.
8. **Comprehensive Efficiency Evaluation:** Careful evaluation of the efficiency of the algorithms (including profiling and scaling).
9. **Comparison:** Compare classification performance of the Detach-Rocket model against the benchmark method both in terms of accuracy and efficiency.
10. **Optional Tasks:** If time allows:
    - Modify or try alternative algorithms to enhance performance.
    - Modify algorithms to get event probability and/or primary channels involved in the seizure.
11. **Thesis Report:** Write a comprehensive thesis report summarizing methodologies, results, and providing a discussion on potential applications of results in future epilepsy research.

## Bibliography:


1. Baumgartner, Christoph, and Johannes P. Koren. "Seizure detection using scalp-EEG." Epilepsia 59 (2018): 14-22.
2. Uribarri, G., Barone, F., Ansuini, A. et al. Detach-ROCKET: sequential feature selection for time series classification with random convolutional kernels. Data Min Knowl Disc 38, 3922–3947 (2024). https://doi.org/10.1007/s10618-024-01062-7.
3. Guttag, John. "CHB-MIT Scalp EEG Database" (version 1.0.0). PhysioNet (2010), https://doi.org/10.13026/C2K01R.
4. Shah, Vinit, et al. "The temple university hospital seizure detection corpus." Frontiers in neuroinformatics 12 (2018): 83.
5. Detti, P. (2020). Siena Scalp EEG Database (version 1.0.0). PhysioNet. https://doi.org/10.13026/5d4a-j060.
6. Dan, Jonathan, et al. "SzCORE: Seizure Community Open-Source Research Evaluation framework for the validation of electroencephalography-based automated seizure detection algorithms." Epilepsia (2024).
7. https://epilepsybenchmarks.com/challenge/
8. Solana, Adrià, Erik Fransén, and Gonzalo Uribarri. "Classification of Raw MEG/EEG Data with Detach-Rocket Ensemble: An Improved ROCKET Algorithm for Multivariate Time Series Analysis." arXiv preprint arXiv:2408.02760 (2024).