**Background**

- **Time Series data** is a collection of observed measurements that maintain a sequential order.
- A major problem regarding time series data is time series classification (TSC). TSC seeks to predict a certain class of a new time series based on the training data.
- With regards to this TSC problem, researchers have sought to classify EEG data which can be used to identify certain microsleep events.
- In the current literature and online repositories, there are no all encompassing pipelines that contain newly developed methods of TSC analysis which is why we decided to develop one.
  - **Specific attributes of time series data**
    - Seasonality: Variations that repeats itself over time
    - Trend: Gradual upward or downward movement
    - Level: Average value or moving average value of the series
    - Stationarity: Time series whose mean and variance are constant over time

**Current Research/Pipelines**

- **Sktime**: Sktime is a unified framework for time series data analysis and classification. This was a major framework utilized in the pipeline since other advanced models were based on it.
- **TapNet**: TapNet is a multivariate time series classification tool that can be used for time series data analysis.
- **AutoMLPipe**: AutoMLPipe is an auto-ML pipeline that can be used for time series data analysis.

**GOAL:** Develop an encompassing pipeline (AutoMLPipe-TSC) that processes univariate and multivariate time series data in order to apply recent contributions in computer vision technology such as YOLO and SSD to classify said data.

**Analysis & Preliminary Results**

- **AutoMLPipe-TSC** is an enhancement of the already built AutoMLPipe-PNC which handles general binary classification.
- To build the basis of the pipeline, I utilized the Sktime unified framework since it already has prebuilt classification techniques that can be applied to time series sets usually been too computationally complex. However, this research shows how EDN are more accurate than traditional manual feature learning methods.
- HIVE-COTE was used as the baseline algorithm in all future algorithms would be competitive since it is the oldest/most popular TSC algorithm.
- Besides YOLO, I have started the implementation of InceptionTime into AutoMLPipe-TSC because of its general breadth when handling different series sets.
- This pipeline is a combination of all the current cutting-edge method in time series classification.
- The UCR data set was used for testing since it is a collection of multiple time series data sets.
- The data sets used for training were all preprocessed which allowed for an easier analysis process.

**Strategies for Analyzing Time Series Data**

- **Segmentation:** By segmenting data, we can make an acceptable understanding of the data. A well-known example is this sound data, where the time series can be segmented by different classes or sounds. Through segmentation, our training time drastically decreases as the model only searches for key features in each discrete window instead of both local and global features.
  - Types of segmentation:
    - Rolling Window: A window of a set value moves one increment along the time series (more time required than sliding window)
    - Sliding Window: Set of default events are created and are a specific discrete length
  - Rolling Window vs Sliding Window:
    - Rolling Window: A window of set value moves one increment along the time series (more time required than sliding window)
    - Sliding Window: Like the normal sliding window, the adjusted sliding window finds the center of the default events and adjusts the bounds of the window
- **Image Classification Techniques:** Since computer vision and image classification techniques have been around longer and studied more than regular time series techniques, it is useful to be able to use the same time data to a two-dimensional image that can be analyzed.
  - Gramian Angular Fields (GAF): GAF essentially scales the data, converts it to polar, and then constructs a gram matrix using the inner product. The density of the gram matrix is plotted which can be analyzed using image classification techniques.
  - Markov Transition Fields: More complicated than a GAF, a MTF transforms 1D-time series data into a Markov transition matrix by assigning a probability along the temporal order
- **Convolutional Neural Networks (CNNs):** Extract features in a time series data and extract features automatically in order to classify the data set, unlike regular machine learning where feature extraction is manually completed. Recently, convolutional neural networks are being utilized in order to analyze TS data.

**Conclusions**

- Based on the preliminary results data sets, the pipeline is able to handle unsegmented data and apply image classification techniques.
- Future work would include a method to manually segment the data using methods such as rolling window or adjusted sliding window.

**References**


**Acknowledgements**

This work was supported by NIH grants AI131794, DK122127, ES015058, EY023200, HL134015, LM000101, LM010909, LM013560, and TR01263.

**Code Availability**

GitHub: [https://github.com/Urbanowicz](https://github.com/Urbanowicz)