Multi-modal Unsupervised Variational Autoencoder framework for artifact detection

Student: Indrajeet Ghosh, Juan Arizpe Vega
Faculty: Dr. Kasthuri Jayarajah and Dr. Nirmalya Roy
ARL Collaborator: Dr. Nicholas Waytowich

Introduction - Artifacts

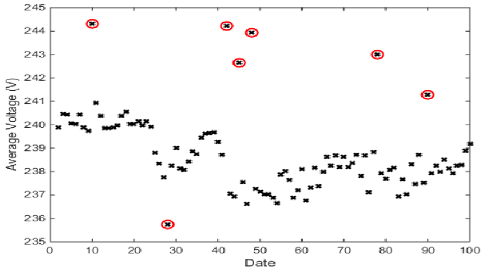
Artifact – A normal unpredictable distortion within data caused by cognitive and/or physical characteristics.

- Eye Artifacts
- Face Movements
- Muscle Artifacts
- Motion Artifacts
- Connectivity Noise
- Malfunction Channel Noise

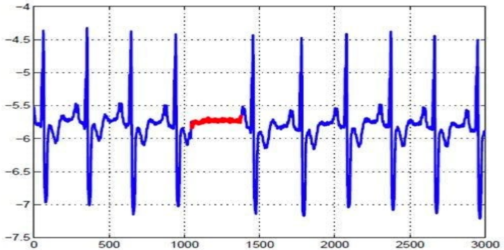
Sensors for collecting data:
- EEG - Electroencephalogram
  - Brain wave recordings
  - Neural Activity

Anomalies

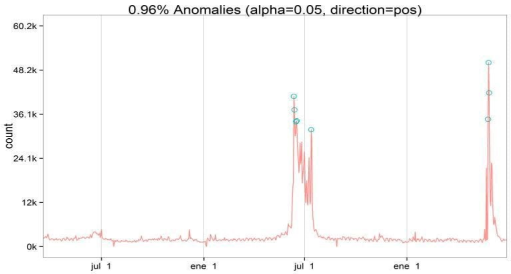
Point Anomalies



Collective Anomalies



Contextual Anomalies



Challenges / Motivation

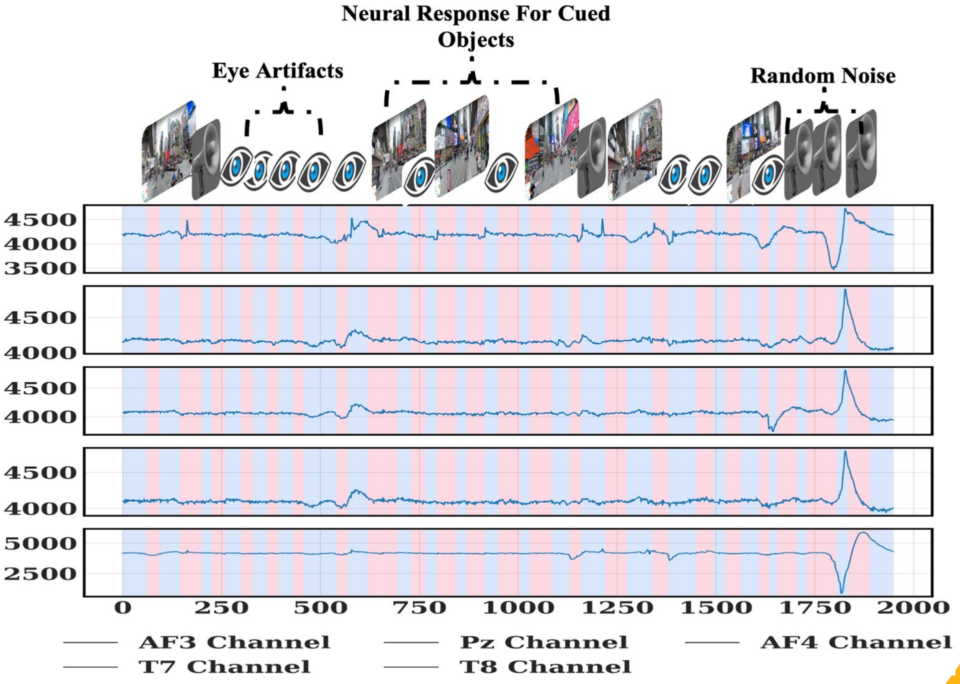
Being able to properly identify artifacts from data with high accuracy

Challenges:

- Parameters of Signal to Noise Ratio (SNR)
  - EEG = 0.526
- Artifacts Noise
- Data Requires Annotation
- Preprocessing Step
- Scalability

Solution:

- End-to-end multimodal Unsupervised framework for artifact detection
- Reduce artifacts without the need of human annotation



Related Works

Robust PPG-Based Mental Workload Assessment System Using Wearable Devices
(Win-Ken Beh et al.)

High Working Memory (WM) Problems:

- Mental health
- Driving
- Learning

Proposal: Wearable Smart Watches are able to collect PPG information to determine WM.

Problem: Corrupted PPG data due to picking up external artifacts.

Utilized Pre & Post Processing of data to acquire accurate data

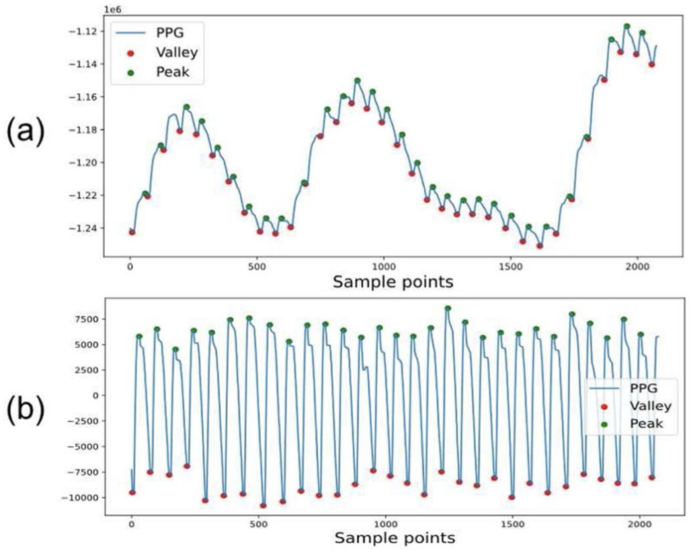


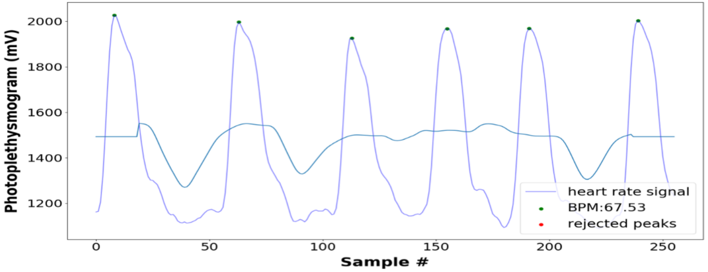
Fig. 3. Peak detection on: (a) raw signal and (b) processed signal by EMD.

Related Works

Deep Recurrent neural network-based autoencoder for Photoplethysmorgram (PPG) Artifact Filtering (Joseph Azar et al.)

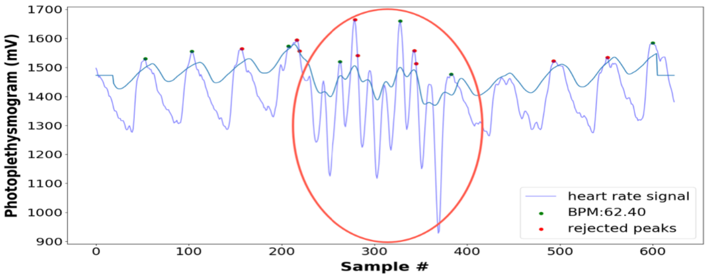
Motivation:

Proposes a neural network-based filtering method to remove corrupted zones from the collected PPG data in an unsupervised manner. It also proposed PPG data summarization and augmentation strategies.



(a) clean PPG signal

Discrete Wavelet Transform (DWT) - sequence summarization approach towards data



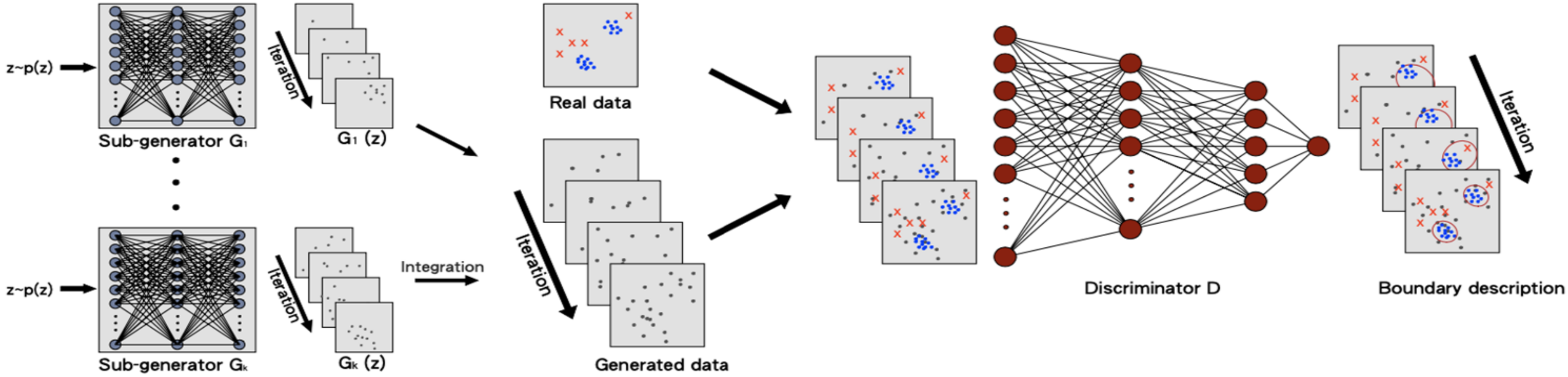
(b) PPG signal with artifacts

Collective Anomalies

Related Works

Generative Adversarial Active Learning for Unsupervised Outlier Detection
(Yezheng Liu et al.)

Multiple-Objective Generative Adversarial Active Learning (MO-GAAL) Architecture



10 Generators with different objectives of producing artificial outlier data points.
Discriminator Learns from Generators and is able to learn to identify Outlier data.

Real world Datasets: Pima, Shuttle, Stamps, PageBlacks, PenDigits, Annthyroid, waveform, WDBC, Ionosphere, spamBase, APS, Arrhthmai, HAR, p53 Mutant

Related Works

Unsupervised EEG Artifact Detection and Correction
(Sari Saba-Sadiya et al.)

Motivation:

End-to-end pre-processing pipeline for the automated identification, rejection and removal/correction of EEG artifacts using a combination of feature-based and deep-learning models.

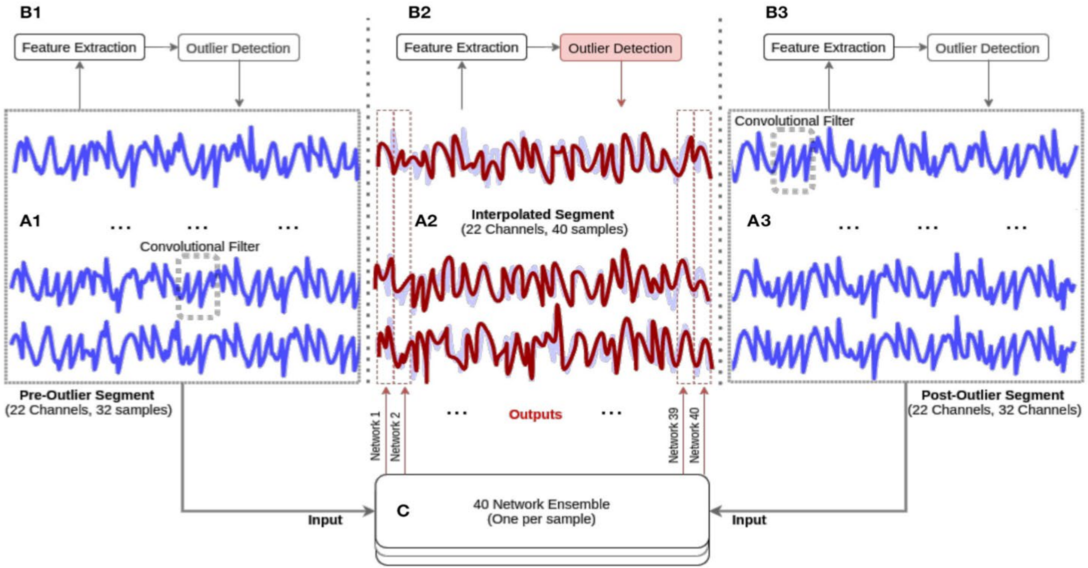


FIGURE 1 | Our methodological approach. The EEG data is first segmented into epochs (see A_1, A_2, A_3). Next, 58 features are extracted and an ensemble of unsupervised outlier detection methods are used (see B_1, B_2, B_3) to identify EEG epochs that are artifact-ridden and require interpolation (see A_2 and B_2). The artifact-ridden epochs are then interpolated by an ensemble of deep encoder-decoder networks (see red line in **C**).

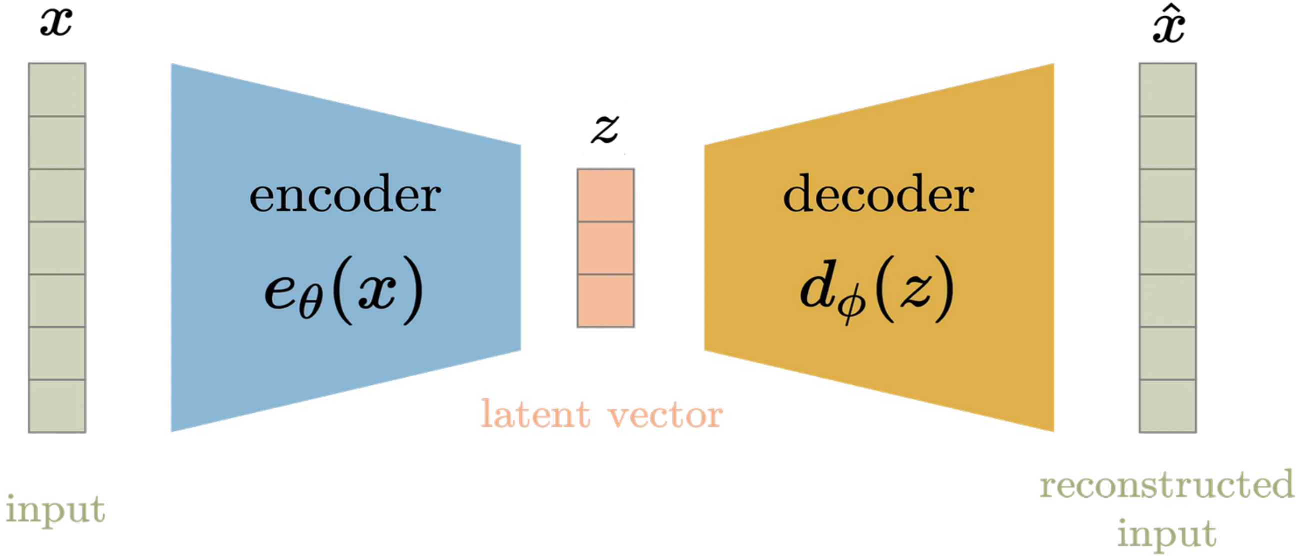
AutoEncoder (AE)

Unsupervised Technique to reduce the dimensionality of data into fewer values than decode the data back to its original dimensionality

Purpose:

- Extract important features from data.
- Reduction of noise

Example:
Sari Saba-Sadiya et al.



$$loss = \|x - \hat{x}\|_2 = \|x - d_\phi(z)\|_2 = \|x - d_\phi(e_\theta(x))\|_2$$

Variational Autoencoder

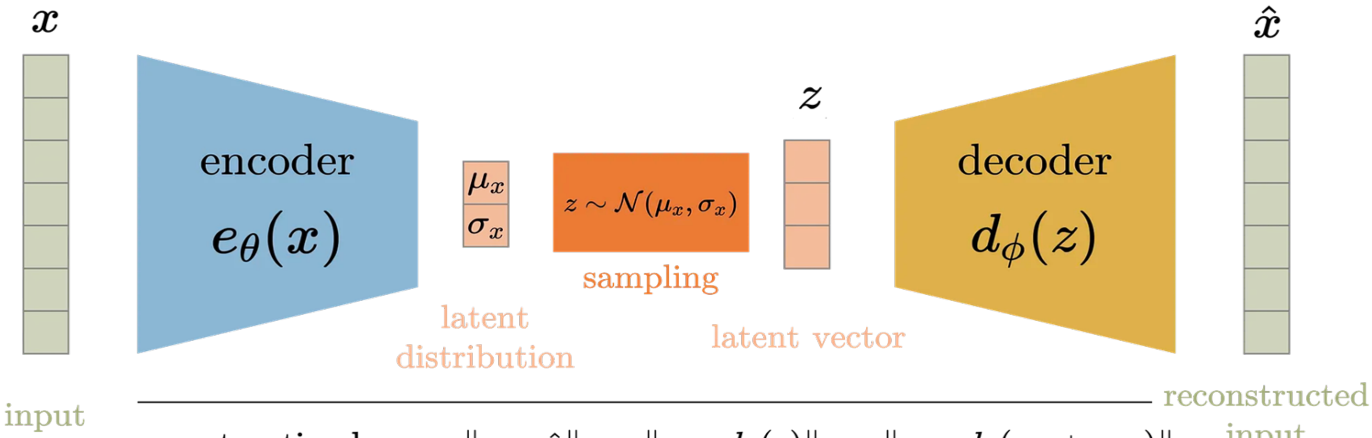
Unsupervised Technique to reduce the dimensionality of data into fewer values and regularizing the encoded information than feeding it to the decoder to output the data to its normal form

Purpose:

- Extract important features from data.
- Reduction of noise
- Regularized latent space

Regularized Latent space by using constraints - Normal Distribution

Latent distribution - mean & variance



$$\text{reconstruction loss} = \|x - \hat{x}\|_2 = \|x - d_\phi(z)\|_2 = \|x - d_\phi(\mu_x + \sigma_x \epsilon)\|_2$$

$$\mu_x, \sigma_x = e_\theta(x), \quad \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

$$similarity\ loss = KL\ Divergence = D_{KL}(\mathcal{N}(\mu_x, \sigma_x) \parallel \mathcal{N}(\mathbf{0}, \mathbf{I}))$$

$$loss = reconstruction\ loss + similarity\ loss$$

Variational AutoEncoder (AE) VS AE

Autoencoder (AE)

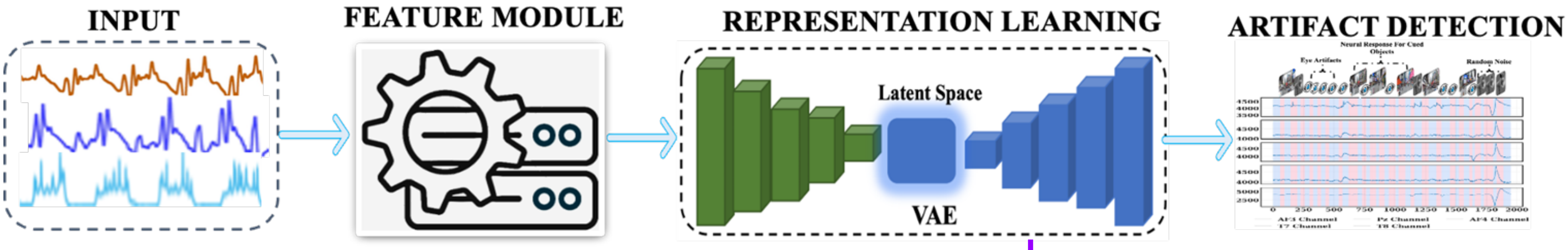
- Used to generate a compressed transformation of input in a latent space.

Variational Autoencoder (VAE)

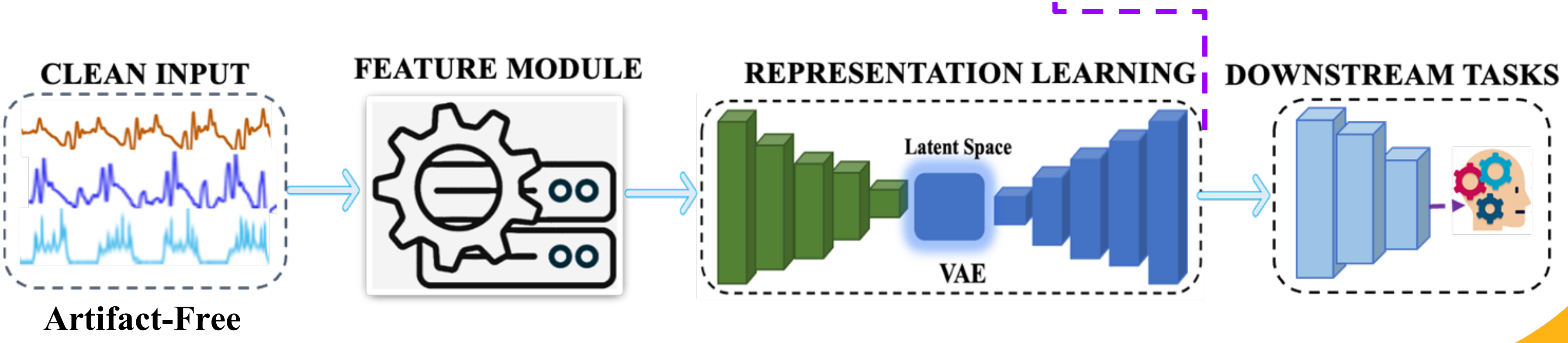
- Enforces conditions on the latent variable to be the unit norm.
- The latent variable in the compressed form is mean and variance.
- Regularized latent space.

Approach

TASK 1 - Artifact Detection



TASK 2 - Artifact Correction and Downstream tasks



Approach

Our proposed framework three module:-

- **Feature Module :-**
  - EEG Signals - Total 58 handcrafted features
    - **Continuity features** - Bursts, spikes, and unusual changes in the mean and standard deviation in the frequency and power domains (median frequency, alpha, beta, gamma, etc.).
    - **Connectivity features** - statistical dependence of EEG signal activity across two or more channels (mutual information, coherence, etc.).
    - **Complexity features** - information-theoretic perspective and are known to correlate with impaired cognitive functions and the presence of degenerative illnesses (Shannon entropy, information quality, false neighbour, etc.).
  - Physiological Signals -
    - **PPG Signals**- beats per minute, interbeat interval, standard deviation of RR intervals, median absolute deviation of RR intervals, etc. [Win-Ken, et,al IEEE Journal of Biomedical and Health Informatics 2023]
    - **GSR Signals**- Phasic and Tonic Signals.

Approach

Shannon entropy-a way to quantify, in a statistical sense, the amount of uncertainty or randomness in the pattern

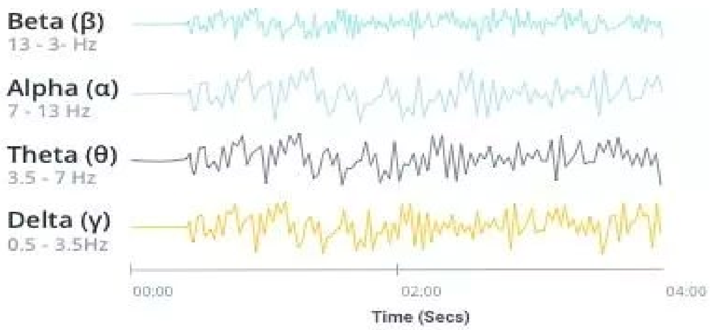


TABLE 1 | EEG Features.

Signal Descriptor	References	Brief description
Complexity features		Degree of randomness or irregularity
Shannon entropy	(22)	Additive measure of signal stochasticity
Tsalis entropy (n = 10)	(23)	Non-additive measure of signal stochasticity
Information quantity ($\delta, \alpha, \theta, \beta, \gamma$)	(24)	Entropy of a wavelet decomposed signal
Cepstrum coefficients (n = 2)	(25)	Rate of change in signal spectral band power
Lyapunov exponent	(26)	Separation between signals with similar trajectories
Fractal embedding dimension	(27)	How signal properties change with scale
Hjorth mobility	(28)	Mean signal frequency
Hjorth complexity	(28)	Rate of change in mean signal frequency
False nearest neighbor	(29)	Signal continuity and smoothness
ARMA coefficients (n = 2)	(30)	Autoregressive coefficient of signal at (t-1) and (t-2)
Continuity features		Clinically grounded signal characteristics
Median frequency		The median spectral power
δ band power		Spectral power in the 0–3 Hz range
θ band power		Spectral power in the 4–7 Hz range
α band power		Spectral power in the 8–15 Hz range
β band power		Spectral power in the 16–31 Hz range
γ band power		Spectral power above 32 Hz
Standard deviation	(31)	Average difference between signal value and it's mean
α/δ ratio	(14)	Ratio of the power spectral density in α and δ bands
Regularity (burst-suppression)	(14)	Measure of signal stationarity/spectral consistency
Voltage < (5, 10, 20 μ)		Low signal amplitude
Diffuse slowing	(32)	Indicator of peak power spectral density <8 Hz
Spikes	(32)	Signal amplitude exceeds μ by 3σ for 70 ms or less
Delta burst after spike	(32)	Increased δ after spike, relative to δ before spike
Sharp spike	(32)	Spikes lasting <70 ms
Number of bursts		Number of amplitude bursts
Burst length μ and σ		Statistical properties of bursts
Burst band powers ($\delta, \alpha, \theta, \beta, \gamma$)		Spectral power of bursts
Number of suppressions		Segments with contiguous amplitude suppression
Suppression length μ and σ		Statistical properties of suppressions
Connectivity features		Interactions between EEG electrode pairs
Coherence – δ	(14)	Correlation in 0–4 Hz power between signals
Mutual information	(18)	Measure of dependence
Granger causality – All	(33)	measure of causality
Phase lag index	(34)	Association between the instantaneous phase of signals
Cross-correlation magnitude	(35)	Maximum correlation between two signals
Cross-correlation – lag	(35)	Time-delay that maximizes correlation between signals

The 58 EEG features fell into three EEG signal property domains: Complexity features (25 in total), Category features (27 in total), Connectivity features (six in total).

Approach

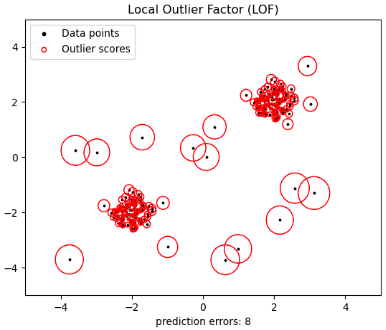
Our proposed framework three module

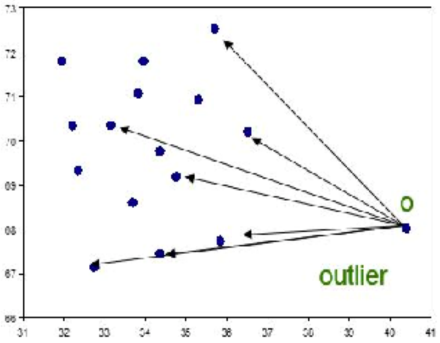
- **Representation Learning Module: -**
  - We employed Variational AutoEncoder to learn the underlying data distribution
    - operates with a probability distribution for each latent variable instead outputs a single value to describe each latent variable
    - Sampling technique - mean and variance (reparameterization trick)
- **Artifact Detection and Correction Module: -**
  - Propagate and reconstructed signals using the trained model to quantify the divergence from ground-truth (noise-free and with-noise)
  - Downstream Task - Cognitive monitoring
    - Quantify the cognitive overhead in performing working memory tasks.

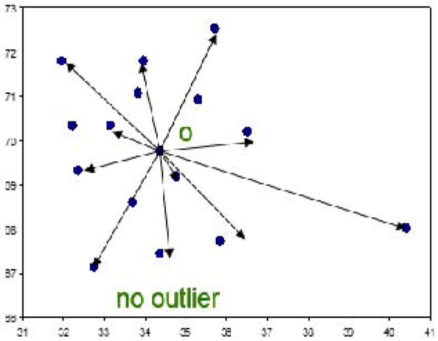
Baseline

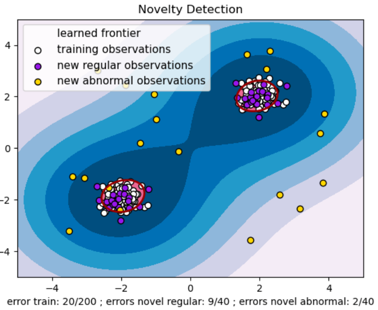
Baseline Algorithms -

- **Statistical -based Algorithms : -**
  - Angle Based Outlier Detection -
    - Detects anomalies based on data points angles.
  - One Class Support Vector Machine -
    - Classifies datasets into groups by drawing vectors.
  - Local Outlier Factor -
    - Determines Outliers by calculating distances between variables.









Baseline

Baseline Algorithms -

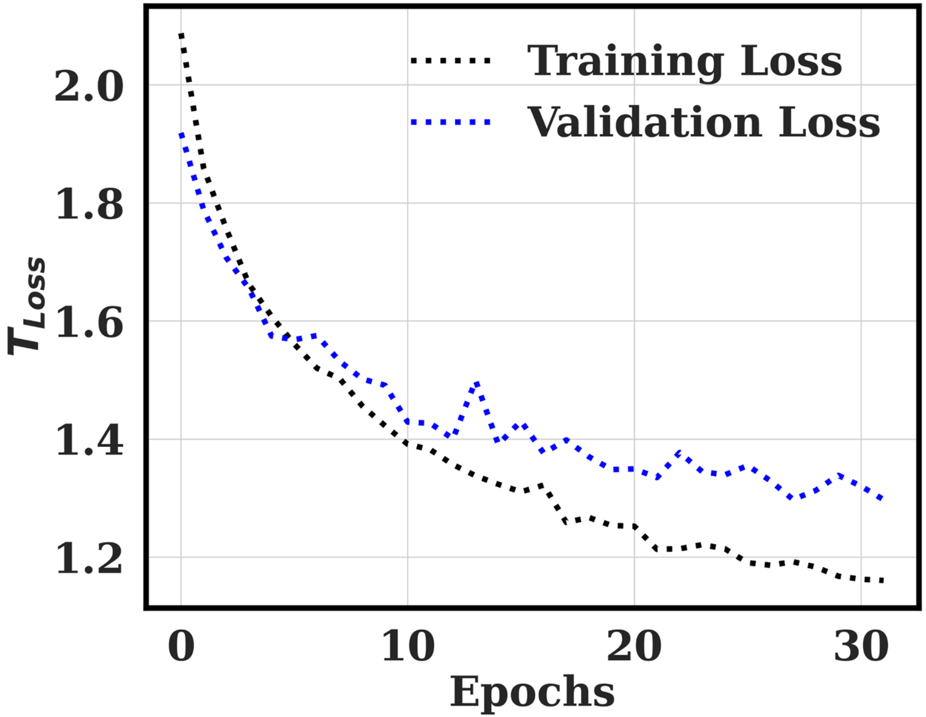
- **Deep Learning Representation -based Algorithms : -**
  - AutoEncoder [Sari Saba-Sadiya et al.]
    - Determines outlier data by interpolation of data.
  - GAAL
    - Single Anomaly Generative Model with discriminator.
  - MO-GAAL
    - Multiple Anomaly Generative Model with discriminator.

Our Proposed Framework

- VAE
  - Representation Learning
    - Encoder -
      - Downsampling Technique
      - Reduces dimensionality of data - mean & variance
      - Regularized latent space from previous data
    - Decoder
      - Upsampling Technique
      - Regenerates data but without outlier data
  - Model Learning Module
    - KL-Divergence Loss + MSE loss (Similarity loss)
    - Loss functions assures data is reconstructed properly.

Training phase

- Trained for 30 Epochs
- Every ten epochs
- Overfitting Indicator



Results

F1-score

-Accuracy measure of Precision & Recall

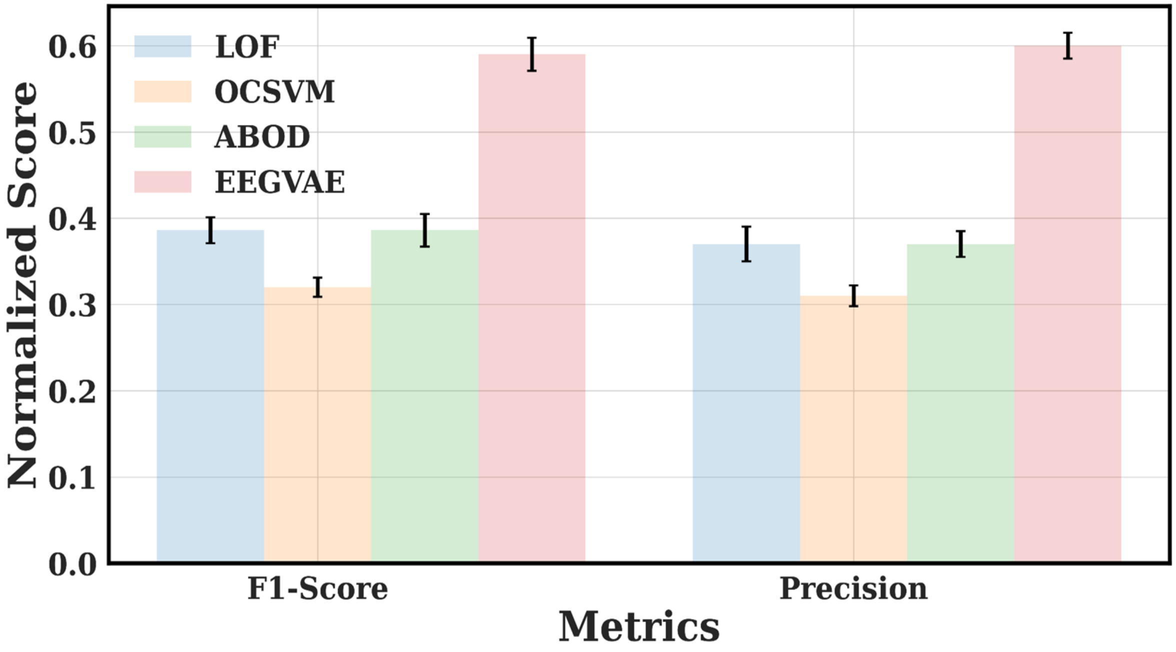
Precision

-Overall the times it got true positive.

Worst: OCSVM

Alright: LOF, ABOD

Best: EEG-VAE

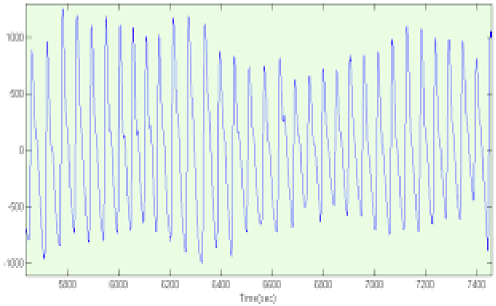


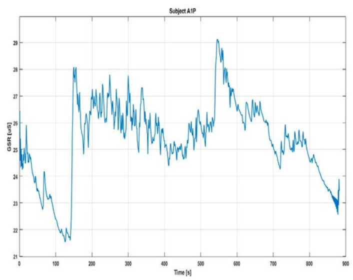
Future work

Include: EEG + GSR + PPG = 0.567

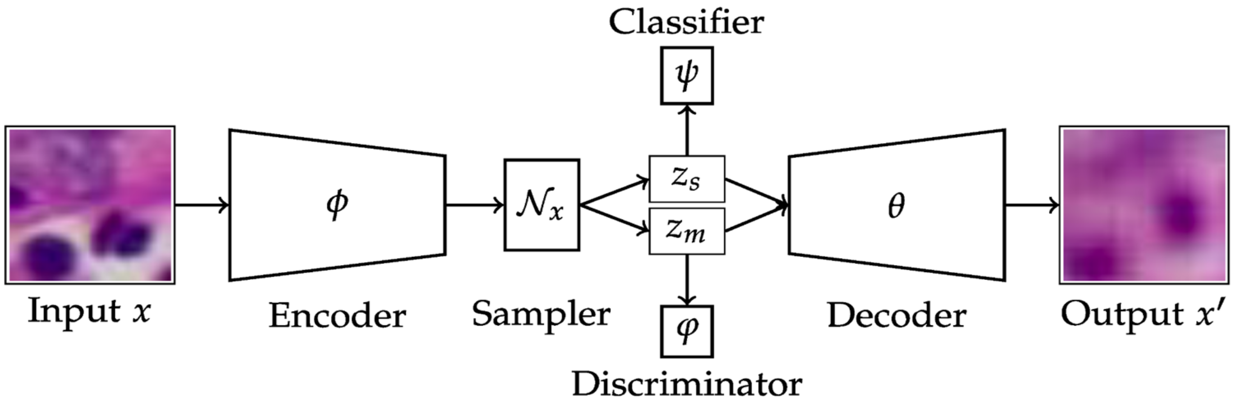
photoplethysmography (PPG)

Galvanin Skin Response (GSR)





Implement disantagle Variational Autoencoder

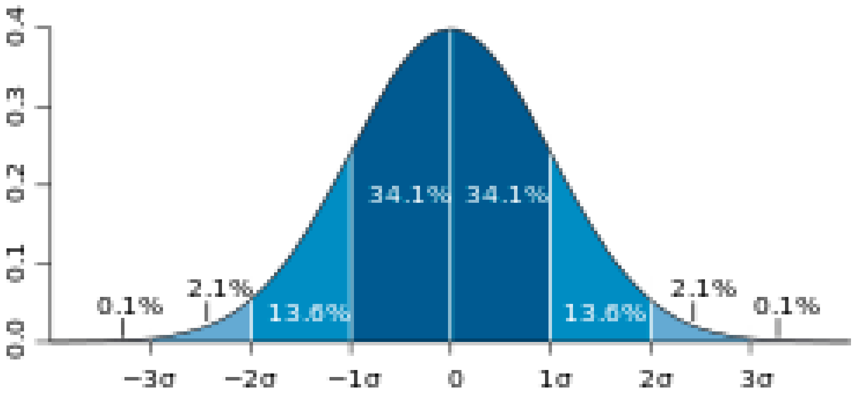


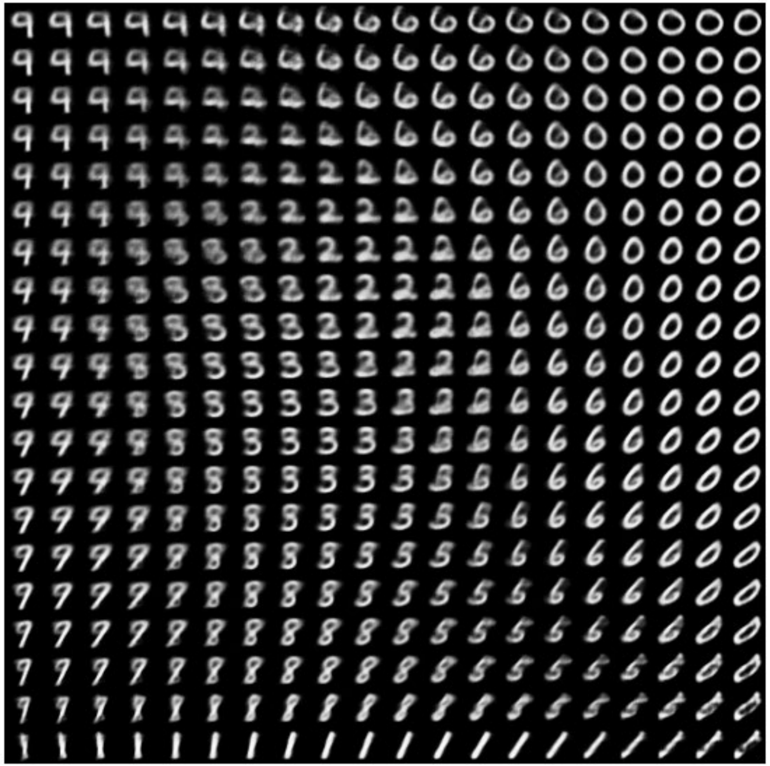
Acknowledgementc

This research could not have been possible for the mentorship and hard work of Indrajeet Ghosh. Also along side the support of Dr. Nirmalya Roy and Dr. Kasthuri Jayarajah.

This research is supported by the NSF Research Experience for Undergraduates (REU) grant \# CNS-2050999, NSF CAREER Award \# 1750936 and U.S. Army Grant \# W911NF2120076.

Supplemental Material





Implementation

Indrajeet Ghosh,
Juan Arizpe Vega

Environments

Jarvis: Remote Server

- Use Docker - start jupyter environment
- Install necessary libraries
- Store scripts, files, and data

Jupyter Notebook

- GUI - for file management and coding

Github

- Version Control - Debugging
- History Log
- Cloning existing projects

Python 3 - Coding Language implemented

Methodology - File Structure

Artifact Detection - Extracts the features from data

Outlier detection - Models

- Depv-Decoding EEG for passive viewing
- EEG-acvc - EEG artifact correction via completion
- GAAL - GAAL Based outlier detection

Outlier detection - Stats

- Angle Based Outlier Detection
- Local Outlier Factor
- One class support vector machine

```
.
└── REU_2023_EEG/
    ├── EEG22_results.ipynb
    ├── EEG32_results.ipynb
    ├── GSR_results.ipynb
    ├── PPG_results.ipynb
    ├── EEGExtract.py
    ├── artifact_detection/
    │   ├── EEG_32ch_Features.py
    │   ├── EEG_Features.py
    │   ├── PPG_Features.py
    │   └── GSR_Features.py
    ├── depv_Base_Models/
    │   └── DEPV.py
    ├── eeg-acvc_Base_Models/
    │   └── EEG_ACVC.py
    ├── gaal_Base_Models/
    │   └── GAAL.py
    ├── stat_Base_Models/
    │   ├── ABOD.py
    │   ├── LOF.py
    │   └── OCSVM.py
    └── data/
        ├── EEG/
        │   ├── bcidatasetIV2a
        │   └── OSF
        ├── PPG
        └── GSR/
            ├── UTDallas
            └── UTDallas_raw
```

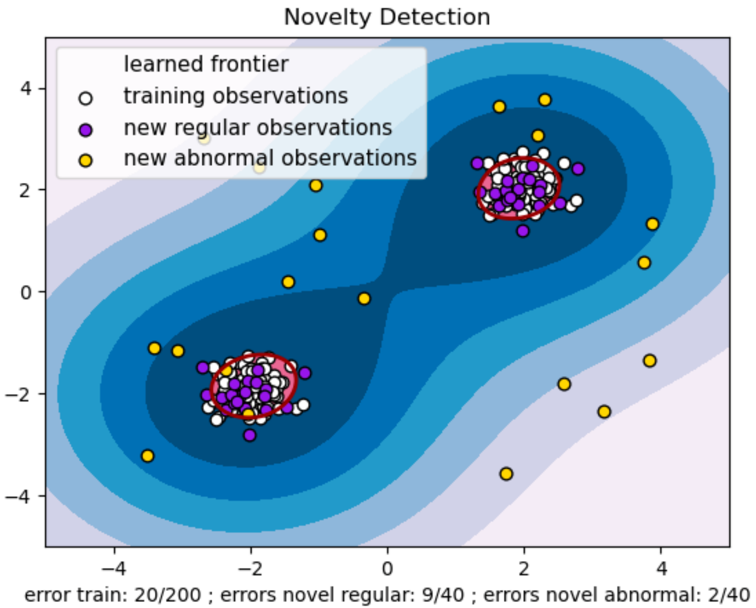
One Class Support Vector Machine

Unsupervised model for anomaly detection

-Sklearn Function-

Initializing - Data is feed into function
- Creates data variable
- Creates SVM Object/Model

Forward -
- Returns a vector of scores of likely how data is incorrect.



Local Outlier Factor

Utilizes distances to determine groups and points that are not within clusters

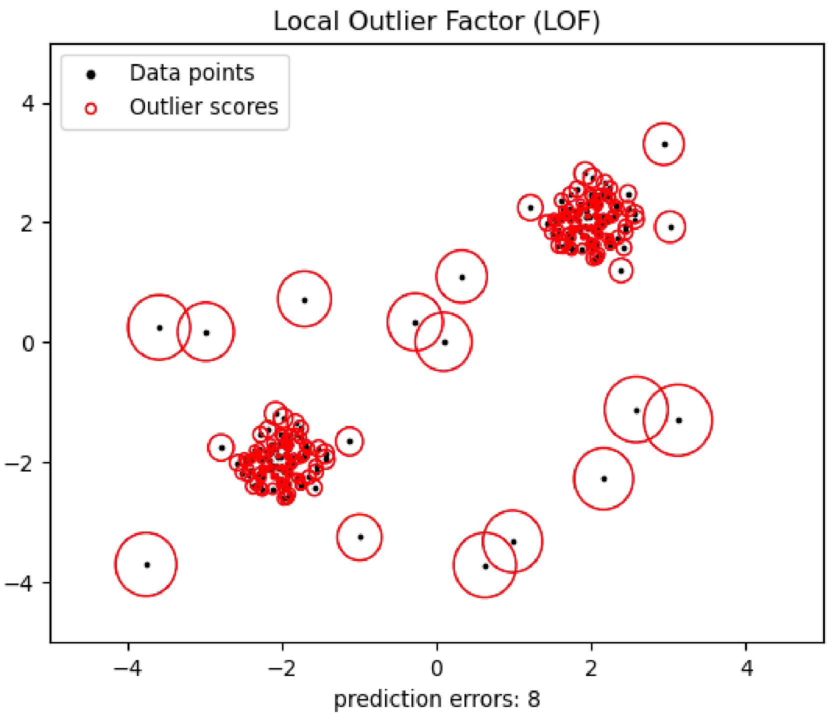
-Sklearn Function-

Initializing - Ground truth and outlier data is feed into the function.

- Creates data variable w/ Ground truth
- Creates boolean vector indexing the combined information of both data sets

Forward -

- Returns a vector of LOF scores of likely how data is incorrect.
- 1 indicates inlear higher means outlier



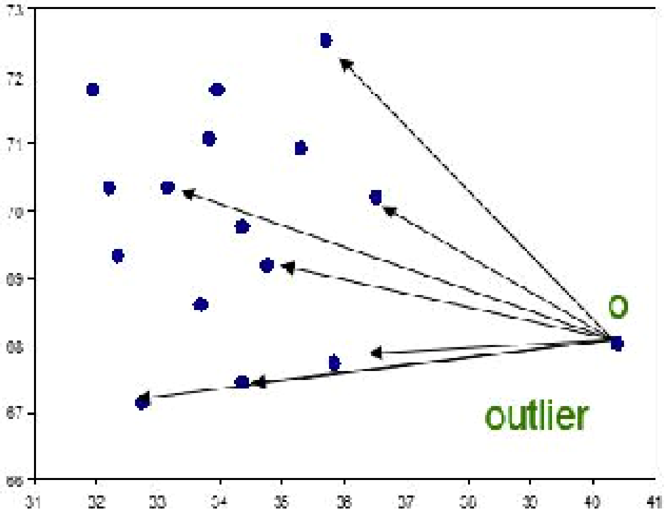
Angle Based Outlier Detection

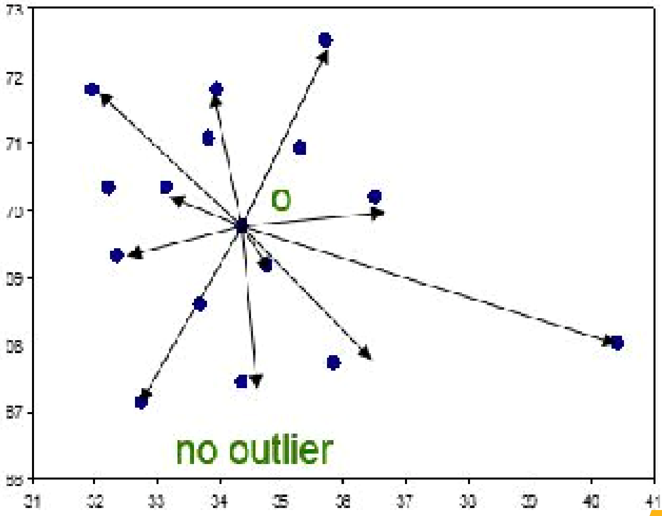
Determines outliers based on angles of points in comparison to other points

-Pyod package-

Initializing - Data is feed
- 60/40 Split
- Xd and Yd Train
- Unzip Xd by column

Forward -
- Returns a vector of scores
- Overall performance of ABOD





Implementation

Indrajeet Ghosh,
Juan Arizpe Vega

UMBC

Implementation

Summary -

- Wrapped up EEG ABOD
  - Error, F1 - Score, and Cohen Kappa Score
  - Still have some bugs to work out
- Started to work on PPG and GSR feature extraction.
  - Came across preprocessing issues

Tools I've used

- Python Libraries-
  - Pickle
  - Numpy
  - Pandas
- Files Format
  - CSV
  - Pickle
  - Npy, and npz

Baseline

- Dataset -
  - In-house dataset - Working Memory (WoM) Dataset
    - Comprised of multi-modal data - GSR, PPG, EEG
    - Collection 35 subjects performing four 4 different visual-spatial tasks
  - BCI Dataset -
    - EEG data from 9 subjects where they were tasked with four motor imagery tasks
    - 22 channels
  - MAUS Dataset - Mental Workload
    - PixArt PPG watch & Procomp Infiniti PPG
    - 22 subjects
  - Sari Saba-Sadiya et al. Dataset - Anomaly detection through interpolation
    - EEG data 32 channels, Visual-event potentials (VEP)

GSR Features

The 4 statistical features we construct for each 5 second time window are as follows:

1. mean, 2. standard deviation, 3. maximum, and 4. Minimum.

24 EDA features, from feature 1 to 24:

Feature 1 to 4 are the 4 statistical features of the raw EDA data.

Feature 5 to 8 are the 4 statistical features of the first derivative of the raw EDA data.

Feature 9 to 12 are the 4 statistical features of the second derivative of the raw EDA data.

Feature 13 to 16 are the 4 statistical features of the 1Hz wavelet coefficients of the raw EDA data.

Feature 17 to 20 are the 4 statistical features of the 2Hz wavelet coefficients of the raw EDA data.

Feature 21 to 24 are the 4 statistical features of the 4Hz wavelet coefficients of the raw EDA data.

GSR Features cont.

96 Acceleration features, from feature 25 to 120:

25 - 28: are the 4 statistical features of the 3-axis acceleration magnitude data.

Feature 29 to 32 are the 4 statistical features of the first derivative of the 3-axis acceleration magnitude data.

Feature 33 to 36 are the 4 statistical features of the second derivative of the 3-axis acceleration magnitude data.

Feature 37 to 40 are the 4 statistical features of the x axis acceleration data.

Feature 41 to 44 are the 4 statistical features of the first derivative of the x axis acceleration data.

Feature 45 to 48 are the 4 statistical features of the second derivative of the x axis acceleration data.

Feature 49 to 52 are the 4 statistical features of the y axis acceleration data.

Feature 53 to 56 are the 4 statistical features of the first derivative of the y axis acceleration data.

Feature 57 to 60 are the 4 statistical features of the second derivative of the y axis acceleration data.

Feature 61 to 64 are the 4 statistical features of the z axis acceleration data.

Feature 65 to 68 are the 4 statistical features of the first derivative of the z axis acceleration data.

Feature 69 to 72 are the 4 statistical features of the second derivative of the z axis acceleration data.

Feature 73 to 76 are the 4 statistical features of the 1Hz wavelet coefficients of the 3-axis acceleration magnitude data.

Feature 77 to 80 are the 4 statistical features of the 2Hz wavelet coefficients of the 3-axis acceleration magnitude data.

Feature 81 to 84 are the 4 statistical features of the 4Hz wavelet coefficients of the 3-axis acceleration magnitude data.

Feature 85 to 88 are the 4 statistical features of the 1Hz wavelet coefficients of the x axis acceleration data.

Feature 89 to 92 are the 4 statistical features of the 2Hz wavelet coefficients of the x axis acceleration data.

Feature 93 to 96 are the 4 statistical features of the 4Hz wavelet coefficients of the x axis acceleration data.

Feature 97 to 100 are the 4 statistical features of the 1Hz wavelet coefficients of the y axis acceleration data.

Feature 101 to 104 are the 4 statistical features of the 2Hz wavelet coefficients of the y axis acceleration data.

Feature 105 to 108 are the 4 statistical features of the 4Hz wavelet coefficients of the y axis acceleration data.

Feature 109 to 112 are the 4 statistical features of the 1Hz wavelet coefficients of the z axis acceleration data.

Feature 113 to 116 are the 4 statistical features of the 2Hz wavelet coefficients of the z axis acceleration data.

Feature 117 to 120 are the 4 statistical features of the 4Hz wavelet coefficients of the z axis acceleration data.

PPG Features

	Feature
Time	SDNN
	NN50
	PNN50
	RMSSD
	SDSD
	TINN
	TRI Index
Frequency	TF
	LF
	HF
	LFn
	HFn
	LF/HF

EEG 22ch Results ABOD

Results

```
feature: bandPwr_alpha       Error % = 23.15340909090909       F1-Score = 0.5084978070175439       Cohen Kappa = 0.02
8250118555653514
feature: bandPwr_beta       Error % = 20.170454545454543       F1-Score = 0.5689803661605954       Cohen Kappa = 0.14
84693095282713
feature: bandPwr_gamma       Error % = 18.25284090909091       F1-Score = 0.6125274122807017       Cohen Kappa = 0.23
392724070184956
feature: std_res       Error % = 19.03409090909091       F1-Score = 0.5977057080108318       Cohen Kappa = 0.20423794
581097954
feature: ratio_res       Error % = 23.970170454545457       F1-Score = 0.48721697859977975       Cohen Kappa = -0.012
934525890367699
feature: regularity_res       Error % = 21.022727272727273       F1-Score = 0.5556749610865904       Cohen Kappa = 0.
12109862671660432
feature: volt05_res       Error % = 21.697443181818183       F1-Score = 0.5388994373508458       Cohen Kappa = 0.0884
6939553965738
feature: volt10_res       Error % = 21.661931818181817       F1-Score = 0.5401589912280701       Cohen Kappa = 0.0908
475035566697
feature: volt20_res       Error % = 20.845170454545457       F1-Score = 0.5618119811734641       Cohen Kappa = 0.1327
2835640472447
feature: df_res       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa = 0.0
feature: spikeNum       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa = 0.0
feature: deltaBurst       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa = 0.0
feature: sharpSpike_res       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa = 0.
0
feature: numBursts_res       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa = 0.0
feature: burstLenMean_res       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa =
0.0
feature: burstLenStd_res       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa =
0.0
feature: numSupps_res       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa = 0.0
feature: supplenMean_res       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa =
0.0
feature: suppLenStd_res       Error % = 17.329545454545457       F1-Score = 0.4525660964230171       Cohen Kappa = 0.
0
feature: coherence_res       Error % = 23.082386363636363       F1-Score = 0.5078447990294561       Cohen Kappa = 0.0
2744244567619958
--------------------------------------------------------------------------------
Mean: Error Percent = 19.99067826704546    F1-Score = 0.5141297023617234    Cohen Kappa = 0.06764291726825149
```

Implementation

Indrajeet Ghosh,
Juan Arizpe Vega

UMBC

F1 and cohen kappa score

Error Percent: Overall percentage of correct

F1 score: Harmonic mean. Calculates mean of precision vs recall.

$$F1\,score = 2 * \frac{Precision \ * Recall}{Precision + Recall}$$

Cohen Kappa: A measure of consistency.

$$kappa = \frac{totalAccuracy - randomAccuracy}{1 - randomAccuracy}$$

EEG Results ABOD

- BCI Dataset -

EEG data from 9 subjects where they were tasked with four motor imagery tasks.

EEG Results ABOD
22 Channels:
Error Percent: 20.0%
F1_Score: 0.51
Cohen Kappa: 0.07

TABLE 1 | EEG Features.

Signal Descriptor	References	Brief description
Complexity features		Degree of randomness or irregularity
Shannon entropy	(22)	Additive measure of signal stochasticity
Tsalis entropy (n = 10)	(23)	Non-additive measure of signal stochasticity
Information quantity ($\delta, \alpha, \theta, \beta, \gamma$)	(24)	Entropy of a wavelet decomposed signal
Cepstrum coefficients (n = 2)	(25)	Rate of change in signal spectral band power
Lyapunov exponent	(26)	Separation between signals with similar trajectories
Fractal embedding dimension	(27)	How signal properties change with scale
Hjorth mobility	(28)	Mean signal frequency
Hjorth complexity	(28)	Rate of change in mean signal frequency
False nearest neighbor	(29)	Signal continuity and smoothness
ARMA coefficients (n = 2)	(30)	Autoregressive coefficient of signal at (t-1) and (t-2)
Continuity features		Clinically grounded signal characteristics
Median frequency		The median spectral power
δ band power		Spectral power in the 0–3 Hz range
θ band power		Spectral power in the 4–7 Hz range
α band power		Spectral power in the 8–15 Hz range
β band power		Spectral power in the 16–31 Hz range
γ band power		Spectral power above 32 Hz
Standard deviation	(31)	Average difference between signal value and it's mean
α/δ ratio	(14)	Ratio of the power spectral density in α and δ bands
Regularity (burst-suppression)	(14)	Measure of signal stationarity/spectral consistency
Voltage < (5, 10, 20 μ)		Low signal amplitude
Diffuse slowing	(32)	Indicator of peak power spectral density <8 Hz
Spikes	(32)	Signal amplitude exceeds μ by 3σ for 70 ms or less
Delta burst after spike	(32)	Increased δ after spike, relative to δ before spike
Sharp spike	(32)	Spikes lasting <70 ms
Number of bursts		Number of amplitude bursts
Burst length μ and σ		Statistical properties of bursts
Burst band powers ($\delta, \alpha, \theta, \beta, \gamma$)		Spectral power of bursts
Number of suppressions		Segments with contiguous amplitude suppression
Suppression length μ and σ		Statistical properties of suppressions
Connectivity features		Interactions between EEG electrode pairs
Coherence – δ	(14)	Correlation in 0–4 Hz power between signals
Mutual information	(18)	Measure of dependence
Granger causality – All	(33)	measure of causality
Phase lag index	(34)	Association between the instantaneous phase of signals
Cross-correlation magnitude	(35)	Maximum correlation between two signals
Cross-correlation – lag	(35)	Time-delay that maximizes correlation between signals

The 58 EEG features fell into three EEG signal property domains: Complexity features (25 in total), Category features (27 in total), Connectivity features (six in total).

Update

Researching Other Data Sets:

Requirements for new dataset:

- Contains artifact annotations
- Is GSR or PPG

Implemented a save file for dataframe system with numpy library.

Next Steps

- Try to run tmux
- Keep working on EEG data