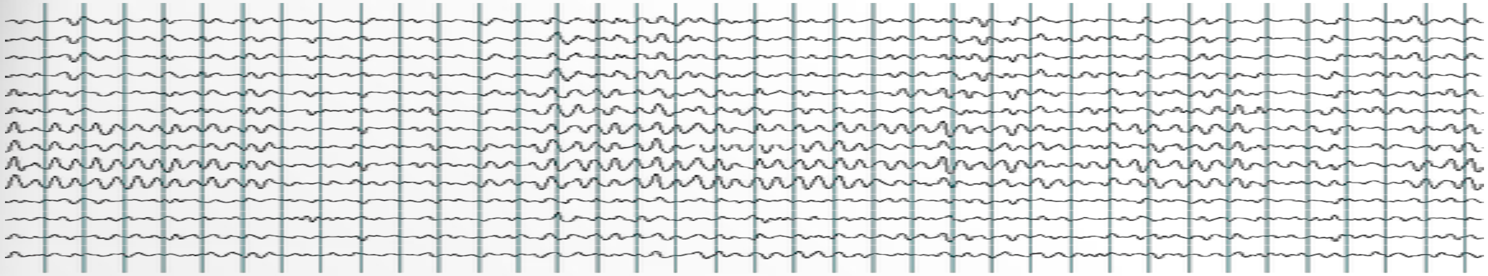
Enhancing EEG-Based User Verification with a Normalized Neural Network Ensemble Approach

Roberto Saia, Riccardo Balia, Alessandro Sebastian Podda, Livio Pompianu, Salvatore Carta, and Alessia Pisu

Department of Mathematics and Computer Science
University of Cagliari, Italy

Outline

- **Overview**
  Introduction ▪ State of the Art ▪ Open Problems
- **Intuition**
  Idea ▪ Architecture
- **Formalization**
  Data Input ▪ Preprocessing ▪ Optimization ▪ Verification
- **Results**
  Experiments ▪ Conclusions and Future Directions



OVERVIEW

Introduction

- The exponential increase in systems/services requiring a reliable **user authentication** process needs robust biometric systems to perform such a task, leading many researchers toward developing effective systems

- Within this research context, researchers consider **EEG-based approaches** a great opportunity due to their reliance on unique patterns of user brain activity

Introduction

- The main strength of **EEG-based biometric systems** is their capability to capture/analyze complex brain activity patterns, finding **unique patterns** that characterize each user

- However, the **inherent complexity** and variability of this data make designing reliable solutions challenging

- This work proposes a ***Normalized Neural Network Ensemble*** approach to address these problems and **improve** the EEG-based systems performance

State of the Art

- In this research domain, the state of the art is characterized by an ever increasing number of works
- However, it should be noted that these works have a high level of heterogeneity given by a wide variability in the choices of components, approaches, and strategies
- This makes comparisons between their performances difficult and sometimes impossible

State of the Art ■■

- To overcome this issue, the proposed work was validated by exploiting the **Biometric EEG Dataset** (BED)
- Because it provides **benchmark** values that allow us the comparisons with the widely adopted approaches in literature

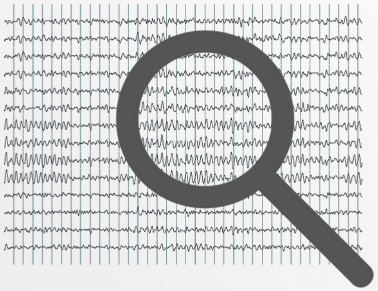
Open Problems

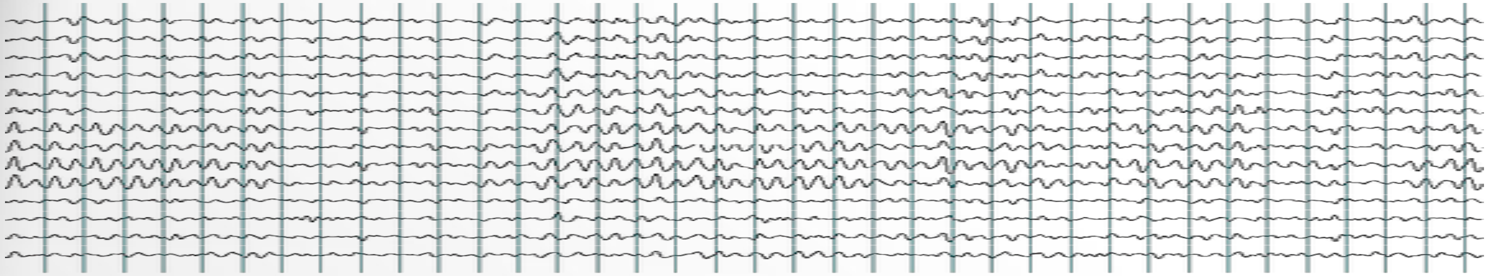
Two **main challenges** in EEG-based biometric systems are:

- The **instability** and **complexity** of EEG patterns over time, making data extraction difficult

- The **variability** in user responses, which complicates calibration and system setup

These issues highlight the **need** for methods that improve data **reliability** and **usability** without requiring complex hardware or extensive preparation





INTUITION

Idea

- We leverage an architecture based on an ensemble of **Multi-Layer Perceptron** (MLP) Artificial Neural Networks (ANNs) governed by a **soft voting** criterion that considers the output class's predicted probability

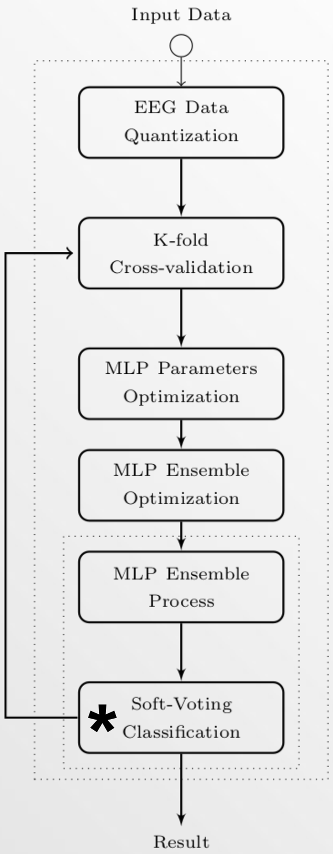
- This approach allows us to **better capture** spatial and temporal patterns in EEG data

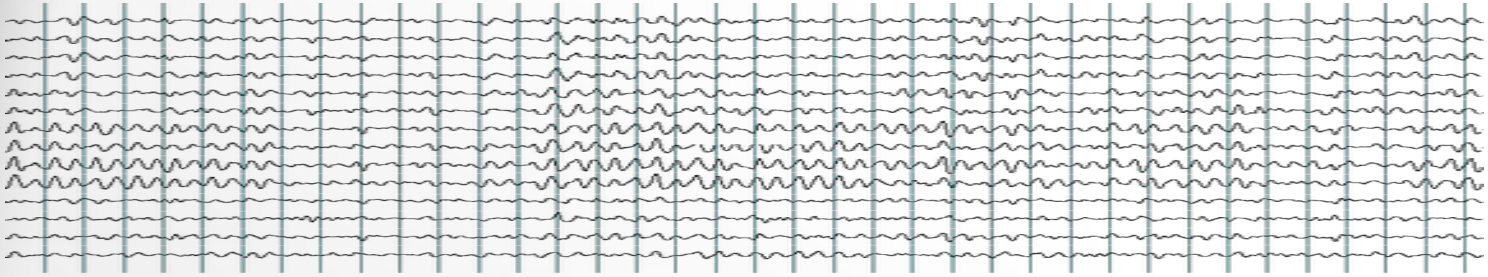
- A preliminary step normalizes the input data according to a **quantile strategy**, and an **automatic optimization** of the MLP hyperparameters is also performed

Architecture

High-level **architecture** of the proposed approach:

* While the approach categorizes input into two classes, its primary function is to verify if the **EEG** data matches the claimed user's stored data





FORMALIZATION

Input Data

The **Biometric EEG Dataset** (BED) is used to verify biometric approaches with the **Emotiv Epoc+** low-cost EEG acquisition device

EEG Data Collection:
- **Channels**: 14-channel signals
- **Sampling frequency**: 256 Hz
- **Device**: Emotiv wireless EEG headset

Data Features:
- **Feature Extraction Techniques**: MFCC, ARRC, SPEC
- **Stimuli Types**: 12 different stimuli
- **Sessions**: 3 disjointed sessions

Dataset Configuration:
- **Number of different datasets**: 3 × 12 × 3 = **108**

EEG Data Preprocessing

Quantile Normalization (QN):
- Ensures uniform distribution across samples, facilitating comparison

Preprocessing Steps:
- Sort values in each column (feature) of the dataset
- Assign ranks to sorted values (ties get average rank)
- Calculate the average value for each rank across columns (quantile values)
- Replace original values with corresponding quantile values.

Formal Representation:
- Given dataset **X** with **n** samples (rows) and **m** features (columns), the normalized value $\mathbf{Y_{ij}}$ for column **j** calculated by:

$$Y_{ij} = \text{Quantile\_Value}(X_{ij}) = \frac{1}{n}\sum_{k=1}^{n} X_{kj}$$

MLP Optimization

Objective:
Optimize key hyperparameters of the *Multi-Layer Perceptron* (MLP) neural network for ensemble classification

Hyperparameters:
- **activation**: Activation function for hidden layers
- **hidden_layer_sizes**: Number of neurons in each hidden layer
- **learning_rate_init**: Weight update step size
- **alpha**: L2 regularization term to prevent overfitting
- **max_iter**: Maximum number of training iterations

Optimization Method:
- **Grid Search**: Evaluates all combinations of hyperparameters
- **Validation**: k-fold cross-validation (k=5)
- **Selection Criterion**: Based on Equal Error Rate (EER)

Result:
Optimal hyperparameters dynamically chosen based on training data characteristics

MLP Ensemble Optimization

Objective:
Determine the optimal number of MLP classifiers for the ensemble

Approach:
- **Uniform Hyperparameters**: All MLP classifiers in the ensemble use the same optimized hyperparameters
- **DIfferent Initialization**: Each classifier is initialized differently via the random_state parameter to enhance ensemble performance
- **Optimization Objective**: Minimize Equal Error Rate (EER) by selecting the optimal number of MLP classifiers (**N_o**)

$$N_{\mathrm{o}} = \arg\min_{N} \mathrm{EER}\left(\mathrm{Ensemble}\left(N\right)\right)$$

Outcome:
- **Trade-off**: Optimal ensemble size balances model complexity with performance
- **EER Metric**: Used to evaluate and minimize ensemble performance in binary classification

MLP Ensemble Classification

Purpose:
Perform user verification using an ensemble of optimized MLP classifiers

Classification Method:
- **Soft Voting Criterion**: Combines probabilities from each MLP classifier in the ensemble to predict the user class, aggregating decisions probabilistically, considering the confidence of each classifier

$$\hat{y} = \mathrm{argmax}_j \left(\frac{1}{N} \sum_{i=1}^{N} P_{ij} \right)$$

Performance Metrics:
- **AUC** (Area Under the Curve)
- **EER** (Equal Error Rate)

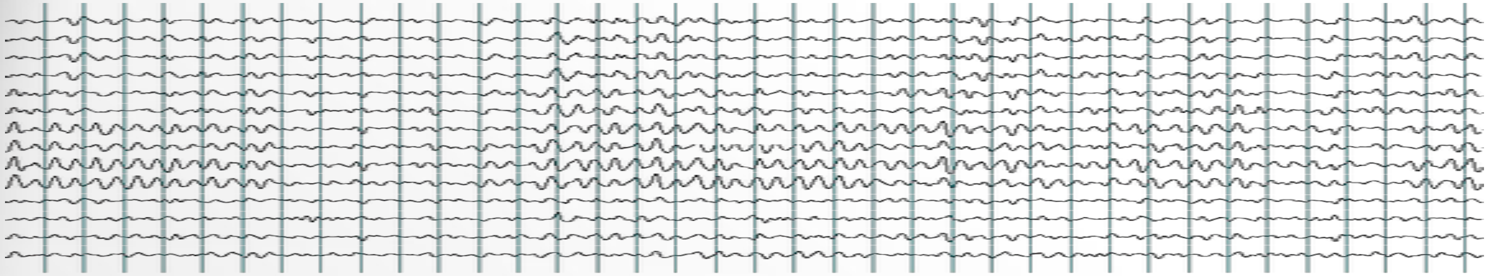
Advantages:
- **Handles EEG Variability**: Soft Voting accommodates the inherent complexity of EEG signals by considering the uncertainty in predictions
- **Improved Decision-Making**: Enhances accuracy by leveraging the strengths of multiple classifiers

User Verification

Objective:
Classify each user's input into a specific identity class

Process Overview:
Class Definition: Represents the identity of the user being verified
Verification vs. Identification: While the approach categorizes input into classes, its primary function is to verify if the EEG data matches the claimed user's stored data
Practical Application: Ensures that the EEG data from a user matches the stored data, confirming the user's identity.

Algorithm Implementation:
- **Inputs**:
  - Training set
  - Unclassified EEG data
  - Random seed initialization value
- **Output**:
  - Classification of unclassified EEG data



RESULTS

Experiments

Training and Testing:

- **Training Set**: Session 1 used for optimizing MLP hyperparameters and determining the optimal number of networks in the ensemble

- **Test Sets**: Sessions 2 and 3 used for validation using AUC and EER metrics

Results Overview:

- **Comparison**: Results compared to baseline values from the BED dataset reference paper

- **Outcome Indicators**:
  - **+** NNNE approach outperforms baseline
  - **−** NNNE approach underperforms baseline
  - **=** Performance is equal

Experiments

NNNE Approach Trained With Session 1 and Tested With Session 2

Dataset	Stimulation	HMM AUC	HMM EER	NNNE AUC	NNNE EER	Comparison AUC	Comparison EER
ARRC	AS	0.7320	0.2911	0.7265	0.3105	-	-
ARRC	MC	0.7034	0.3180	0.7335	0.3254	+	-
ARRC	RC	0.7285	0.2961	0.6579	0.3749	-	-
ARRC	RO	0.6765	0.3355	0.5792	0.4313	-	-
ARRC	VC3	0.5420	0.4594	0.5105	0.4921	-	-
ARRC	VC5	0.6611	0.3282	0.5146	0.4870	-	-
ARRC	VC7	0.6714	0.3409	0.6284	0.3890	-	-
ARRC	VC10	0.5725	0.4234	0.5260	0.4894	-	-
ARRC	VF3	0.6425	0.3604	0.5633	0.4645	-	-
ARRC	VF5	0.6813	0.3314	0.5667	0.4509	-	-
ARRC	VF7	0.7478	0.2865	0.6104	0.4158	-	-
ARRC	VF10	0.7479	0.2733	0.6217	0.4208	-	-
MFCC	AS	0.7699	0.2633	0.8629	0.1921	+	+
MFCC	MC	0.6917	0.3085	0.8279	0.2364	+	+
MFCC	RC	0.7567	0.2792	0.7713	0.2592	+	+
MFCC	RO	0.6617	0.3601	0.7354	0.3035	+	+
MFCC	VC3	0.6099	0.4092	0.6225	0.4087	+	+
MFCC	VC5	0.6846	0.3315	0.7215	0.3160	+	+
MFCC	VC7	0.6476	0.3643	0.7185	0.3132	+	+
MFCC	VC10	0.6250	0.3886	0.7451	0.2895	+	+
MFCC	VF3	0.6602	0.3503	0.7622	0.2742	+	+
MFCC	VF5	0.6632	0.3573	0.6808	0.3403	+	+
MFCC	VF7	0.7017	0.3183	0.7215	0.3171	+	+
MFCC	VF10	0.7431	0.2826	0.7619	0.2758	+	+
SPEC	AS	0.7465	0.2880	0.7993	0.2558	+	+
SPEC	MC	0.7035	0.3135	0.7790	0.2734	+	+
SPEC	RC	0.7192	0.3164	0.7210	0.3015	+	+
SPEC	RO	0.6528	0.3660	0.6148	0.3950	-	-
SPEC	VC3	0.5934	0.4209	0.5214	0.5012	-	-
SPEC	VC5	0.7567	0.2669	0.6576	0.3720	-	-
SPEC	VC7	0.6771	0.3392	0.6210	0.4048	-	-
SPEC	VC10	0.6424	0.3729	0.5729	0.4328	-	-
SPEC	VF3	0.6687	0.3593	0.6783	0.3593	+	=
SPEC	VF5	0.6606	0.3667	0.6020	0.4212	-	-
SPEC	VF7	0.7380	0.2837	0.7051	0.3409	-	-
SPEC	VF10	0.7551	0.2622	0.6973	0.3424	-	-

Experiments

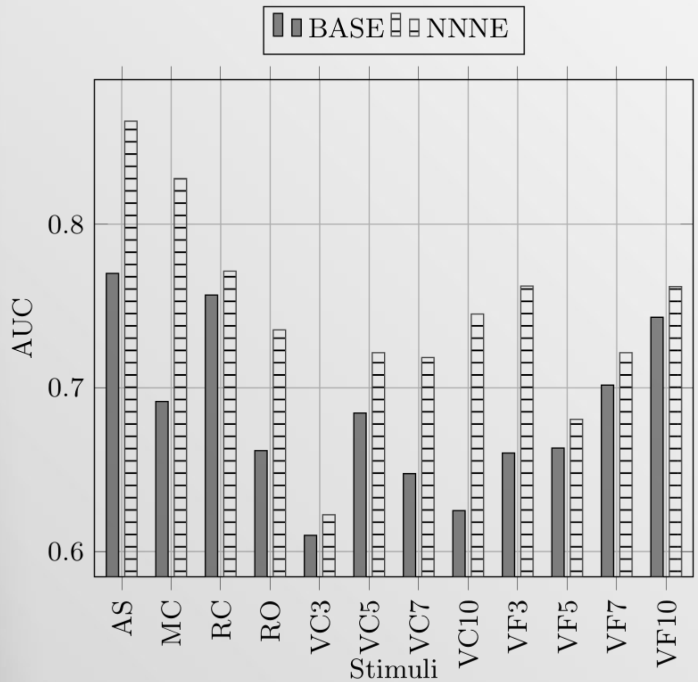
NNNE Approach Trained With Session 1 and Tested With Session 3

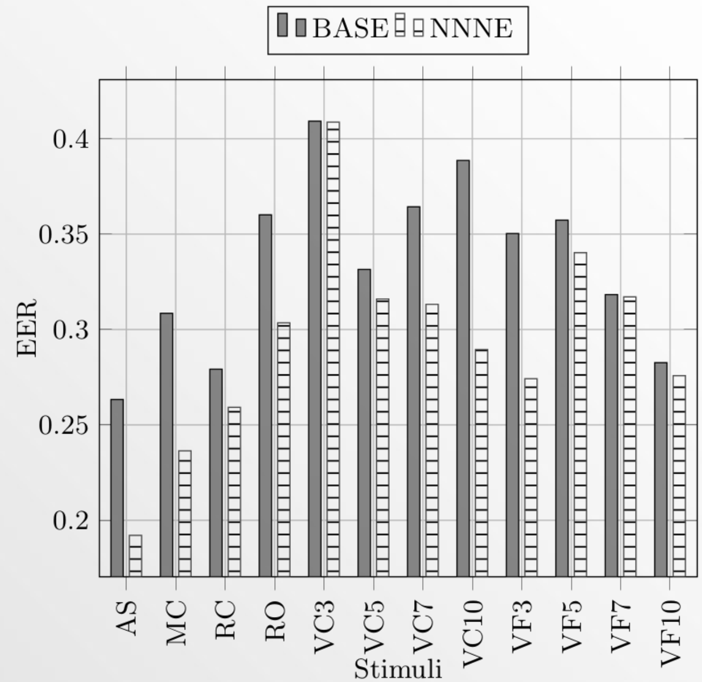
Dataset	Stimulation	HMM AUC	HMM EER	NNNE AUC	NNNE EER	Comparison AUC	Comparison EER
ARRC	AS	0.6487	0.3587	0.7173	0.3258	+	+
ARRC	MC	0.6668	0.3348	0.6934	0.3504	+	-
ARRC	RC	0.7643	0.2567	0.6642	0.3655	-	-
ARRC	RO	0.6797	0.3354	0.6604	0.3635	-	-
ARRC	VC3	0.5932	0.3959	0.5711	0.4315	-	-
ARRC	VC5	0.7251	0.2922	0.6464	0.3942	-	-
ARRC	VC7	0.6069	0.3860	0.6255	0.3996	+	-
ARRC	VC10	0.6059	0.3874	0.5285	0.4732	-	-
ARRC	VF3	0.5905	0.4008	0.5487	0.4599	-	-
ARRC	VF5	0.7061	0.2997	0.6920	0.3356	-	-
ARRC	VF7	0.6514	0.3582	0.5667	0.4521	-	-
ARRC	VF10	0.6636	0.3454	0.6002	0.4232	-	-
MFCC	AS	0.7109	0.3113	0.7808	0.2691	+	+
MFCC	MC	0.6706	0.3389	0.7899	0.2664	+	+
MFCC	RC	0.7030	0.3190	0.8081	0.2245	+	+
MFCC	RO	0.7321	0.2963	0.7407	0.2982	+	-
MFCC	VC3	0.6505	0.3574	0.7109	0.3164	+	+
MFCC	VC5	0.7070	0.3172	0.7281	0.2871	+	+
MFCC	VC7	0.6514	0.3616	0.7103	0.3102	+	+
MFCC	VC10	0.5539	0.4386	0.6237	0.4025	+	+
MFCC	VF3	0.6044	0.4010	0.6743	0.3394	+	+
MFCC	VF5	0.7575	0.2642	0.7928	0.2397	+	+
MFCC	VF7	0.6870	0.3209	0.6800	0.3439	-	-
MFCC	VF10	0.6493	0.3728	0.7590	0.2729	+	+
SPEC	AS	0.6385	0.3821	0.7406	0.3146	+	+
SPEC	MC	0.6204	0.3847	0.7617	0.2878	+	+
SPEC	RC	0.6922	0.3357	0.7437	0.2853	+	+
SPEC	RO	0.7122	0.3152	0.6702	0.3465	-	-
SPEC	VC3	0.6322	0.3776	0.6144	0.3965	-	-
SPEC	VC5	0.6958	0.3322	0.7120	0.3160	+	+
SPEC	VC7	0.6579	0.3637	0.6049	0.4023	-	-
SPEC	VC10	0.5569	0.4445	0.5511	0.4606	-	-
SPEC	VF3	0.6356	0.3495	0.5415	0.4696	-	-
SPEC	VF5	0.6756	0.3278	0.7238	0.3068	+	+
SPEC	VF7	0.6803	0.3337	0.6251	0.3986	-	-
SPEC	VF10	0.6255	0.3842	0.6503	0.3749	+	+

Conclusions

MFCC Comparison by AUC and EER

The **MFCC** results in terms of **AUC** and **EER** highlight the effectiveness of employing a conventional ensemble strategy based on MLP neural networks combined with the quantile normalization of input data and the real-time optimization of the involved hyperparameters, which proves to be effective in modeling spatial and temporal patterns in EEG data





Conclusions

- It is important to highlight that an important contribution of the proposed approach lies in the use of a **canonical ensemble of neural classifiers**, which proves capable of **handling the complexity** of the involved data, improving the performance of biometric user verification systems without the use of more sophisticated and complex approaches

- As **future work**, we will extend the experiments to include additional EEG datasets, investigating the factors contributing to the **varying performance** of the **NNNE** approach across different conditions and datasets

Enhancing EEG-Based User Verification with a Normalized Neural Network Ensemble Approach

THANK YOU FOR YOUR ATTENTION