Salvatore Giugliano

Machine Learning and XAI methods for improving EEG-based BCI classification systems

Tutor: Roberto Prevete

co-Tutors: Francesco Isgrò
Andrea Apicella

Cycle: XXXV

Year: Third

Background

- Master's degree in Computer Science at Università degli Studi di Napoli "Federico II"

- Research group / laboratory
  - Artificial Intelligence, Privacy & Applications (AIPA) Lab
  - Augmented Reality for Health Monitoring Laboratory (ARHeMLab)

- PhD
  - start date 1/11/2019
  - end date 31/10/2022

- Scholarship type
  - No scholarship

- Periods abroad
  - 30/06/2021 - 30/09/2021, Istituto Superior Tècnico, University of Lisbon

Summary of study activities 1/2

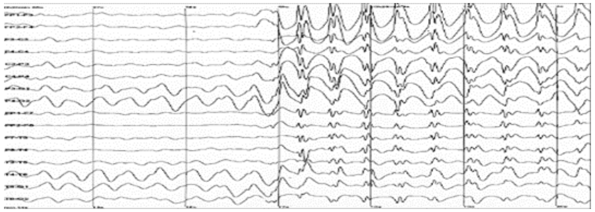
- Attended Courses (relevant for the research field)
  - Accelerated Computing With Cuda C/C++
    (Luigi Troiano)
  - Intelligenza Artificiale ed Etica
    (Roberto Prevete, Guglielmo Tamburrini)
  - Deep Learning for Computer Vision
    (Luigi Troiano)
  - Scientific Programming and Visualization with Python
    (Alessio Botta)
  - Machine Learning
    (Marco Aiello, Anna Corazza, Diego Gragnaniello, Francesco Isgrò, Roberto Prevete, Francesco Raimondi, Carlo Sansone)
  - Data science for patient records analysis
    (Marcello Cinque)

Summary of study activities 2/2

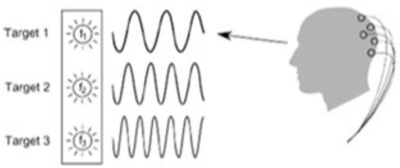
- Attended Seminars (relevant for the research field)
  - Large Scale Training of Deep Neural Networks
    (Giuseppe Fiameni)
  - Bias from the wild
    (Nello Cristianini)
  - Adversarial attacks on image classifiers
    (Andrea Cavallaro)
  - Webinar series on Deep Learning for CINI AIIS Labs
    (Christian Hundt, Gunter Roeth, Niki Loppi, Giuseppe Fiameni
  - Wearable Brain-Computer Interface for Augmented Reality-based Inspection in Industry 4.0
    (Pasquale Arpaia)
  - Machine learning: Causality lost in translation
    (Edwin A. Valentijn)
  - Visual Interaction and Communication in Data Science
    (Marco Quartulli)
  - Picariello Lectures on Data Science
    (Flora Amato, V.K. Gubani, Simon Pietro Romano, Roberto Maranca)
- Study activities for the research areas
  - Machine Learning techniques
  - Analysis and interpretation of EEG signals
  - eXplainable Artificial Intelligence methods

Research area(s) 1/2

Analysis and interpretation of Electroencephalographic (EEG) signals for Brain-Computer Interface (BCI) systems



- EEG types
  - Spontaneous (for passive BCI)
    - emotional and cognitive engagement classification
  - Steady-State Visually Evoked Potential (SSVEP) (for active BCI)
    - different fixed-frequency flickering visual stimuli for specific commands



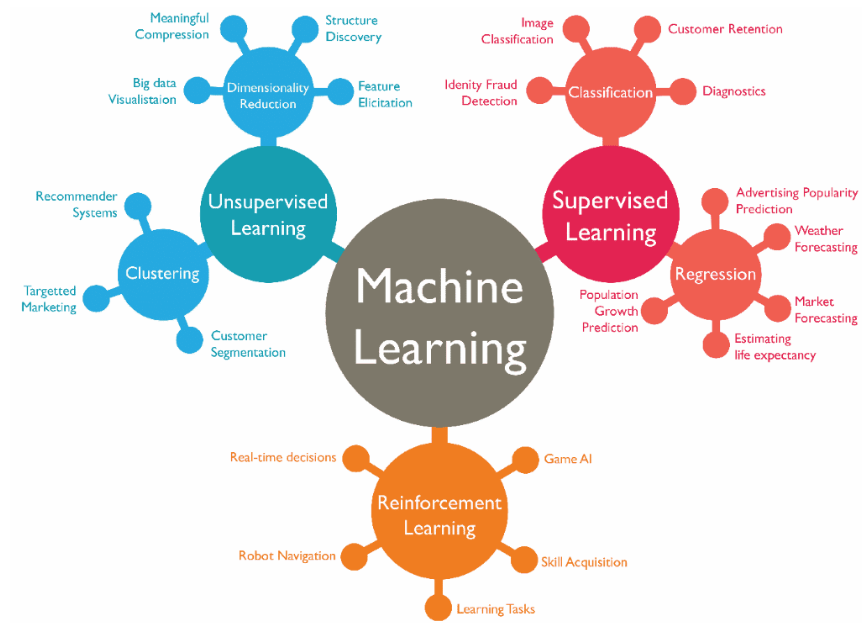
- Non-stationarity
  - significant signal's differences both across different times and different subjects

Research area(s) 2/2

Machine Learning techniques

"Programming computers to optimise a performance criterion using example data or past experience"

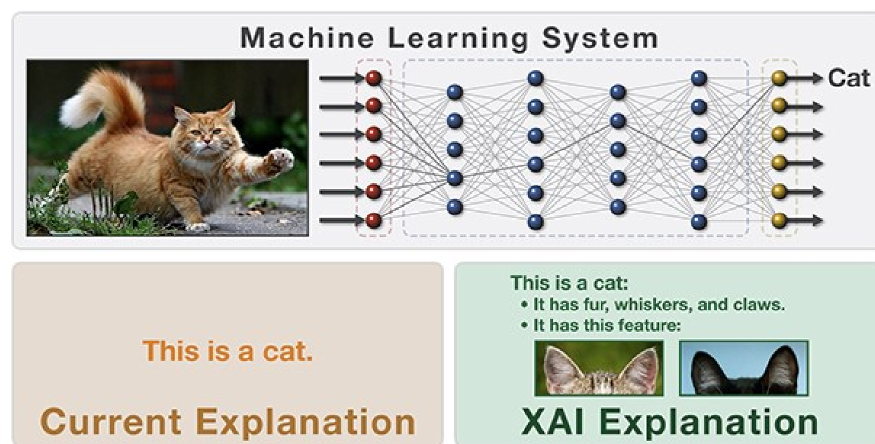
- Deep (and shallow) Neural Networks
- Approaches to handle unbalanced data
- Methods to handle the 'Dataset Shift problem'



eXplainable Artificial Intelligence (XAI)

Explanations of the model's output that can be easily interpreted by the human beings

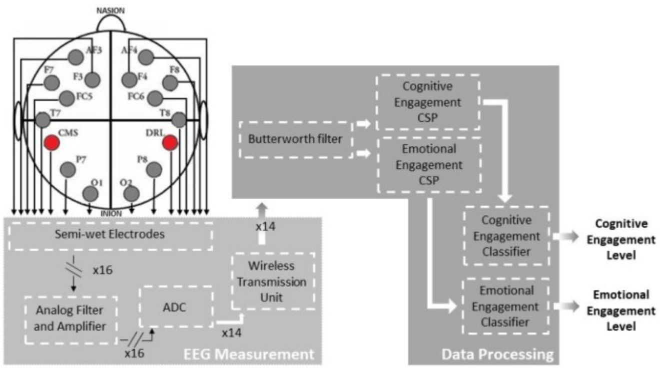
- White-box or Black-box methods for Image Recognition
- Multiple explanations in terms of Middle-Level input Features (MLF)
- Methods to handle the 'Dataset Shift problem'



Research results

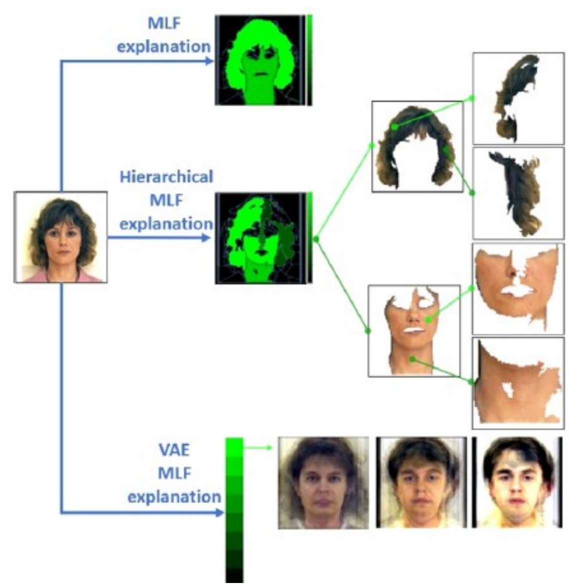
EEG-based BCI classification systems

- Engagement recognition
  - Emotional and cognitive engagement classification system on highly non-stationarity and unbalaced EEG dataset
- SSVEP classification
  - in different feature spaces



Modern XAI

- a method for multiple Explanations in terms of Middle-Level input Features (MLF) (image classification task)
- preliminary experiments for Explanations on EEG data (emotion recognition task)



itee PhD
information technology
electrical engineering

Research products 1/2

Journal papers

[P1]	Apicella A., Giugliano S., Isgrò F., & Prevete R., Exploiting auto-encoders and segmentation methods for middle-level explanations of image classification systems, **Knowledge-Based Systems**, vol. 255 (109725), 2022, DOI: 10.1016/j.knosys.2022.109725
[P2]	Apicella A., Arpaia P., Giugliano S., Mastrati G., & Moccaldi N, High-wearable EEG-based transducer for engagement detection in pediatric rehabilitation, **Brain-Computer Interfaces**, vol. 9 (3), pp. 129-139, 2022, DOI: 10.1080/2326263X.2021.2015149
[P3]	Apicella A., Arpaia P., De Benedetto E., Donato N., Duraccio L., Giugliano S., & Prevete R., Enhancement of SSVEPs classification in BCI-based wearable instrumentation through machine Learning Techniques, **IEEE Sensors Journal**, vol. 22 (9), pp. 9087-9094, 2022, DOI: 10.1109/JSEN.2022.3161743

Research products 2/2

Conference papers

[P4]	Angrisani L., Apicella A., Arpaia P., De Benedetto E., Donato N., Duraccio L., Giugliano S, & Prevete R., A ML-based Approach to Enhance Metrological Performance of Wearable Brain-Computer Interfaces, **IEEE International Instrumentation and Measurement Technology Conference (I2MTC)**, Ottawa, Canada, May. 2022, pp. 1-5, IEEE, DOI: 10.1109/I2MTC48687.2022.9806518
[P5]	Apicella A., Arpaia P., Cataldo A., De Benedetto E., Donato N., Duraccio L., Giugliano S, & Prevete, R., Adoption of Machine Learning Techniques to Enhance Classification Performance in Reactive Brain-Computer Interfaces, **IEEE International Symposium on Medical Measurements and Applications (MeMeA),** Messina, Italy, Jun. 2022, pp. 1-5, IEEE, DOI: 10.1109/MeMeA54994.2022.9856441
[P6]	Apicella A., Giugliano S., Isgrò F., & Prevete R., Explanations in terms of Hierarchically organised Middle Level Features, **Italian Workshop on Explainable Artificial Intelligence**, Milano, Italy, Dec. 2021, vol. 3014 (4), CEUR Workshop, URL: http://ceur-ws.org/Vol-3014/paper4.pdf
[P7]	Apicella A., Giugliano S., Isgrò F., & Prevete R., A general approach to compute the relevance of middle-level input features, **International Conference on Pattern Recognition - International Workshops and Challenges**, Virtual-Milano, Italy, Jan. 2021, pp. 189-20, Springer, DOI: 10.1007/978-3-030-68796-0_14

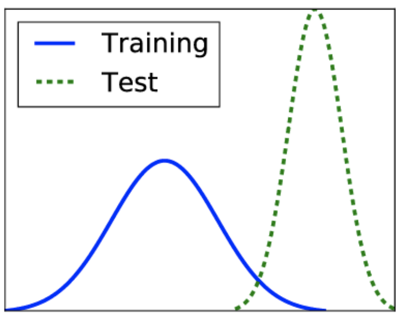
PhD thesis overview 1/2

Problem statement

Machine Learning point of view

Domain or Dataset Shift Problem

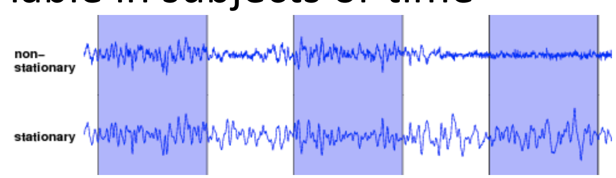
Training and test data can belong to different probability distributions



Bio-signal point of view

Non-stationarity of EEG signals

EEG signals highly variable in subjects or time acquisitions



Poor generalisation performances

PhD thesis overview 2/2

Objective

Improve generalisation across different subjects and different times

Methodology

Part 1

EEG classification problem

- unbalanced data
- non-stationarity

Part 2

XAI on Image Recognition task

- using XAI methods presente in literature
- new XAI methods for explanations in terms of Middle-Level input Features (MLF)

Part 3

Apply methodologies and knowledge learnt in parts 1 and 2 on EEG classification problem

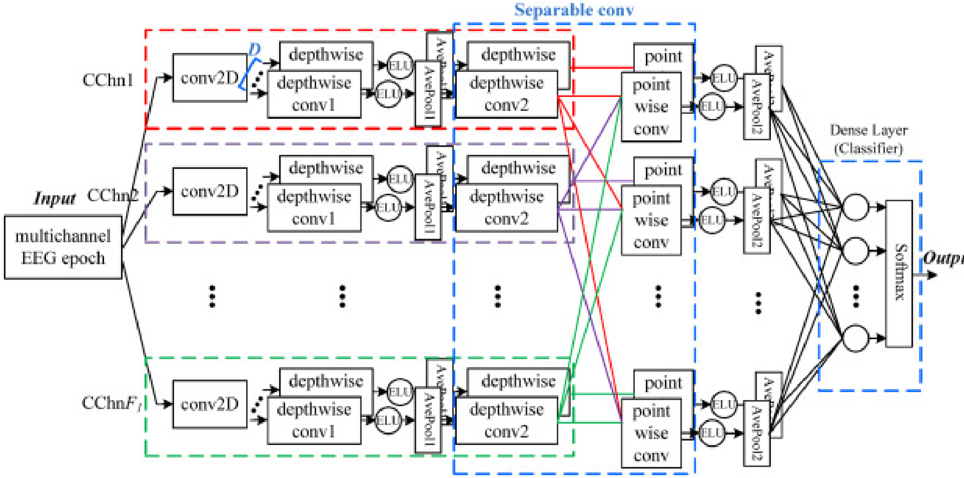
- XAI methods for Image Recognition on Emotion classification task (EEG dataset)

PhD thesis - Proposal

Machine Learning and XAI methods for improving EEG-based BCI classification systems

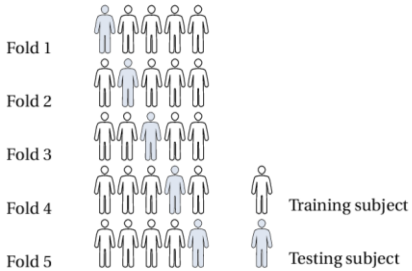
To handle the 'Dataset Shift' / non-stationarity Problem

- Deep Neural Networks architectures
  - Convolutional Neural Networks
  - Depthwise and Separable Convolutions
  - specific for Raw EEG datasets
- Domain Adaptation methods
  - different data normalisation schemes
  - considering the similarity between subjects
- XAI methods
  - usually used on image recognition datasets
  - applied on emotion recognition dataset (EEG)
  - locate the relevant characteristics of the input
  - exploit the properties of relevant characteristics



Inter-subjective validation strategy

- Leave-One-Subject-Out Cross-Validation (LOSO CV)



PhD thesis – Results, Part 1a

Machine Learning and XAI methods for improving EEG-based BCI classification systems

Table 5.3. Overall mean of the intra-individual performances on cognitive engagement using three different classifiers: the balanced accuracy (BA) and the Matthews correlation coefficient (MCC) at varying the oversampling methods.

Oversampling	Metric	k-NN	SVM	ANN	Mean
none	BA	67.1	67.4	73.7	69.4 ± 3.0
	MCC	0.31	0.34	0.45	0.36 ± 0.06
SMOTE	BA	68.6	69.8	72.0	70.1 ± 1.4
	MCC	0.33	0.36	0.40	0.36 ± 0.03
BorderlineSMOTE	BA	70.3	70.9	73.6	71.6 ± 1.4
	MCC	0.36	0.38	0.43	0.39 ± 0.03
ADASYN	BA	68.1	68.3	72.5	69.6 ± 2.0
	MCC	0.33	0.33	0.42	0.36 ± 0.04
SVMSMOTE	BA	69.0	69.4	72.9	70.4 ± 1.7
	MCC	0.34	0.36	0.42	0.37 ± 0.03
KMeansSMOTE	BA	69.8	71.1	74.5	$\mathbf{71.8 \pm 1.98}$
	MCC	0.35	0.39	0.46	$\mathbf{0.39 \pm 0.04}$

Results on Spontaneous EEG classification problem

Table 6.3. Classification Accuracy and Corresponding 1-σ Reproducibility on the Four Data Sets

Data set #1 (Moverio BT-200)

T (s)	CCA [30] (%)	Deep SCU (%)	FR* (%)
0.5	70.8 ± 10.0	74.4 ± 9.5	**75.0 ± 9.5**
1.0	74.8 ± 18.1	81.6 ± 9.6	**82.1 ± 9.8**
2.0	84.9 ± 12.1	87.5 ± 8.0	**89.2 ± 7.8**
3.0	91.0 ± 9.4	91.9 ± 7.3	**93.7 ± 5.6**
5.0	95.4 ± 5.6	95.7 ± 4.9	**96.7 ± 3.9**
10.0	-	97.7 ± 4.5	**99.4 ± 2.7**

Data set #2 (Moverio BT-350)

T (s)	CCA [30] (%)	Deep SCU (%)	FR* (%)
0.5	-	30.9 ± 7.1	**39.2 ± 13.5**
1.0	-	35.8 ± 10.4	**46.3 ± 19.2**
2.0	51.9 ± 27.0	42.8 ± 13.2	**53.9 ± 23.5**
3.0	53.3 ± 25.6	43.5 ± 21.1	**56.7 ± 24.9**
5.0	56.7 ± 23.9	41.4 ± 17.9	**57.5 ± 23.7**
10.0	-	47.2 ± 23.0	**62.2 ± 24.5**

Data set #3 (Hololens)

T (s)	CCA [30] (%)	Deep SCU (%)	FR* (%)
0.5	-	**48.4 ± 11.3**	44.9 ± 10.0
1.0	-	56.9 ± 13.9	**66.8 ± 16.7**
2.0	58.9 ± 20.6	72.3 ± 14.4	**76.4 ± 16.9**
3.0	70.5 ± 18.5	77.0 ± 15.8	**82.6 ± 13.1**
5.0	72.9 ± 28.3	80.0 ± 13.8	**88.9 ± 8.6**
10.0	-	75.0 ± 19.3	**94.4 ± 8.3**

Data set #4 (Oculus Rift S)

T (s)	CCA [30] (%)	Deep SCU (%)	FR* (%)
0.5	-	36.7 ± 10.5	**42.7 ± 16.8**
1.0	-	40.6 ± 16.2	**54.0 ± 21.5**
2.0	56.1 ± 24.2	46.4 ± 18.6	**62.3 ± 23.5**
3.0	64.8 ± 20.9	56.3 ± 20.6	**65.7 ± 25.3**
5.0	68.5 ± 23.2	55.3 ± 18.6	**70.6 ± 23.8**
10.0	-	48.9 ± 21.2	**72.2 ± 23.3**

*Only the best result is reported for brevity.

Results on SSVEP EEG classification problem

PhD thesis – Results, Part 1b

Machine Learning and XAI methods for improving EEG-based BCI classification systems

Results on SSVEP classification problem (Benchmark dataset)

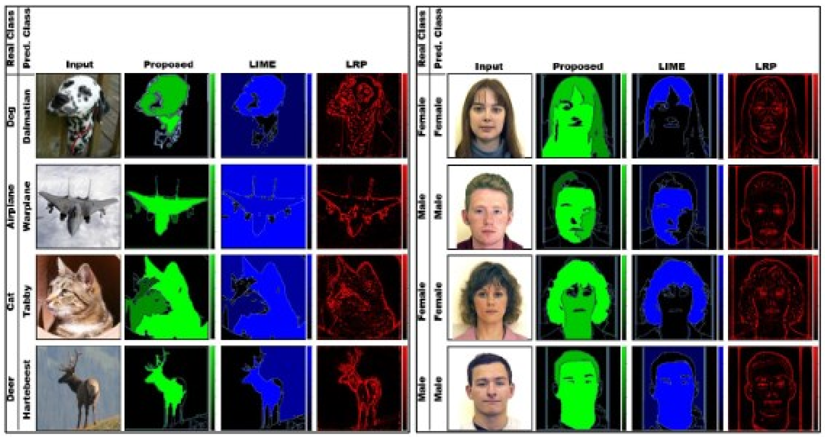
TABLE II
CLASSIFICATION ACCURACY WITH A 5-S ACQUISITION TIME

Method	Accuracy (%)
Proposal (FE + VAF ANN)	**95.5 ± 3.8**
1D SSVEP Convolutional Unit [25]	68.6
PodNet [25]	86.2
Filter Bank CCA [25]	97.9

DATASET BENCHMARK RAW STRATEGY 25-5-5 (10 RUN)				
BASELINE			Normalization (Domain Adaptation)	
epoch	NORM. ONLINE	NORM. PER SET	NORM. PER SUBJECT	NORM. PER BLOCK AND PER SUBJECT
1.5	72.75±6.92	74.17±6.15	**76.42±6.1**	76.24±6.1
5	93.51±4.53	95.77±1.95	**96.59±1.71**	**96.68±1.49**

PhD thesis – Results, Part 2

Machine Learning and XAI methods for improving EEG-based BCI classification systems



(a) STL10 (b) Aberdeen

Figure 8.5. Explanations obtained by GMLF using the flat strategy (second columns), LIME (third columns) and LRP (fourth columns) for VGG16 network responses using images from STL10 (a) and Aberdeen datasets (b). In both (a) and (b), for each input (first columns) the explanation in terms of most relevant segments are reported for the proposed flat approach (second columns) and LIME (third columns). For better clarity, we report a colormap where only the first two most relevant segments are highlighted both for MLRF and LIME.

(a) Proposed approach (b) LIME

Figure 8.7. Examples of a two-layer hierarchical explanation on images classified as *Female* and *Male* by VGG16. (a) First column: segment heat map. Left to right: segments sorted in descending relevance order. Top-down: the coarsest (second row) and the finest (third row) hierarchical level. (b) LIME explanation: same input, same segmentation used in (a).

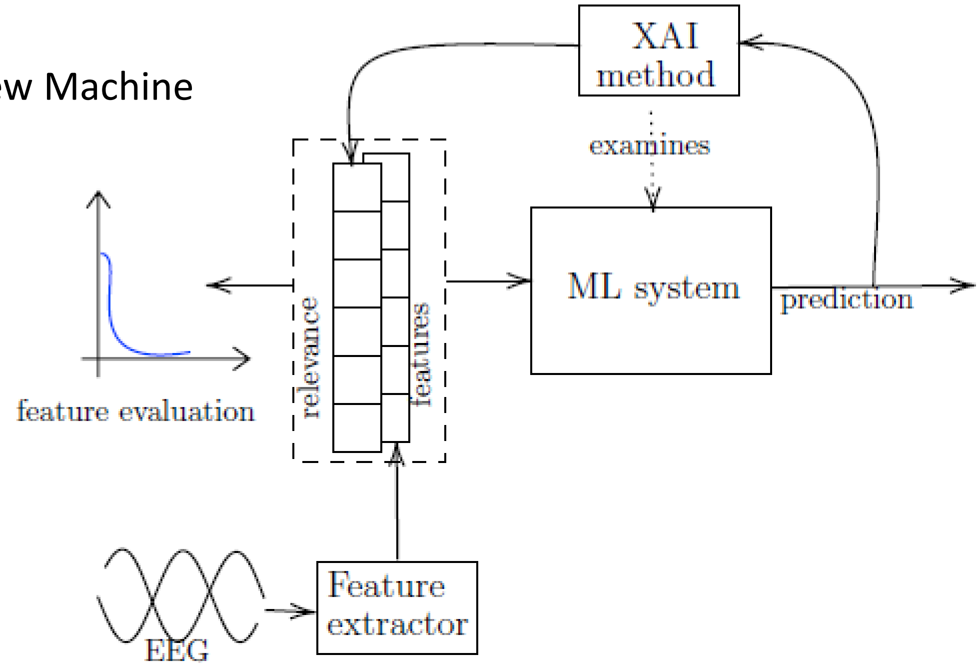
Explanations in terms of MLF

PhD thesis – Results, Part 3

Machine Learning and XAI methods for improving EEG-based BCI classification systems

Experimental assessments on ***preliminary*** results:

- Better generalisation across different subjects / sessions
- Many components found by XAI methods are shared across the sessions / subjects
- Relevant components can be used to build a new Machine Learning system



Thanks for your attention