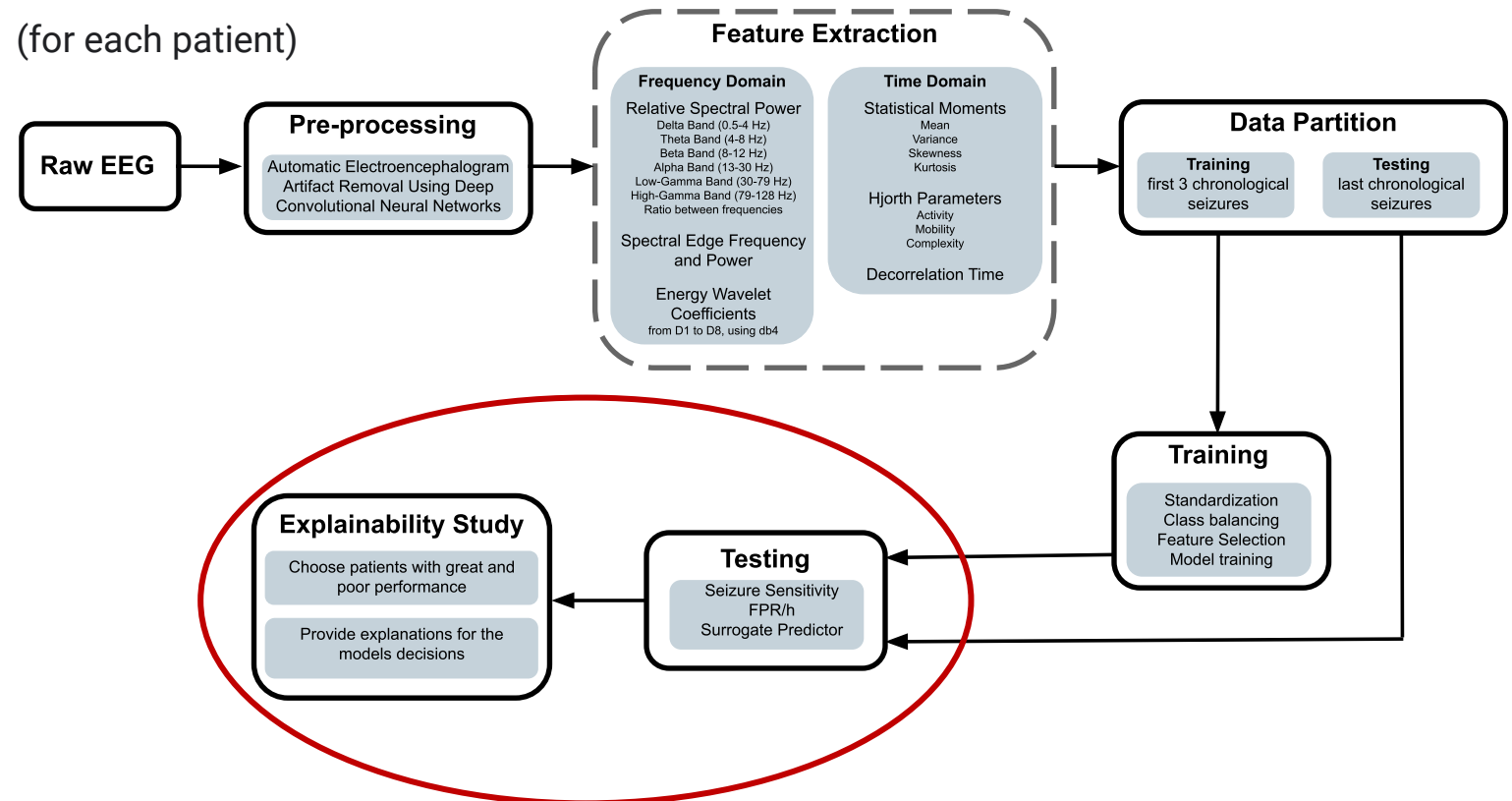# Towards the explanation of Machine Learning methods in EEG seizure prediction

Motivation

- Current state-of-the-art methods for electroencephalogram seizure prediction use Machine Learning (ML). However, as they deal with multidimensionality, they may create black boxes where clinicians may not have sufficient trust to use them for high-risk decisions.

- This work studies which explanations may be provided about the models to increase trust in their decisions.

- First stage: data scientists. **Second stage: clinicians**

General Methodology

- 40 Patients from EPILEPSIAE in pre-surgical monitoring
  - 104 tested seizures and 2055 hours of data

(for each patient)

Raw EEG →

Pre-processing
Automatic Electroencephalogram Artifact Removal Using Deep Convolutional Neural Networks

→

Feature Extraction

Frequency Domain
Relative Spectral Power
Delta Band (0.5-4 Hz)
Theta Band (4-8 Hz)
Beta Band (8-12 Hz)
Alpha Band (13-30 Hz)
Low-Gamma Band (30-79 Hz)
High-Gamma Band (79-128 Hz)
Ratio between frequencies

Spectral Edge Frequency and Power

Energy Wavelet Coefficients
from D1 to D8, using db4

Time Domain
Statistical Moments
Mean
Variance
Skewness
Kurtosis

Hjorth Parameters
Activity
Mobility
Complexity

Decorrelation Time

→

Data Partition

Training
first 3 chronological seizures

Testing
last chronological seizures

↓

Training
Standardization
Class balancing
Feature Selection
Model training

↓

Testing
Seizure Sensitivity
FPR/h
Surrogate Predictor

→

Explainability Study
Choose patients with great and poor performance

Provide explanations for the models decisions

Logistic Regression

$$\phi(z) = \frac{1}{1 + e^{-z}}$$

1

0

- 1) With the first 3 seizures:
  - Find the optimal pre-ictal period within a range of values: 30-60 minutes
  - Find the optimal number of features

- 2) Label windows of signal of 5 seconds as either interictal or preictal

- 3) Feed this examples to a Machine Learning Algorithm, and obtain a model:
  - Logistic Regression

- 4) Apply the model to new data and smooth the output over time:
  - Evaluate Performance

Seizure onset

Firing Power

Alarm

Threshold

SPH

SOP

Time
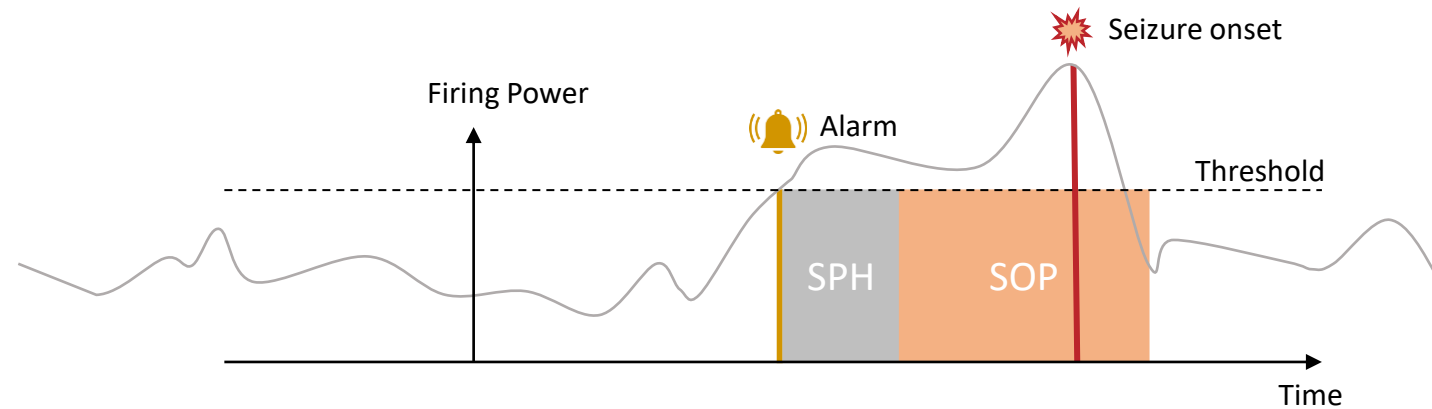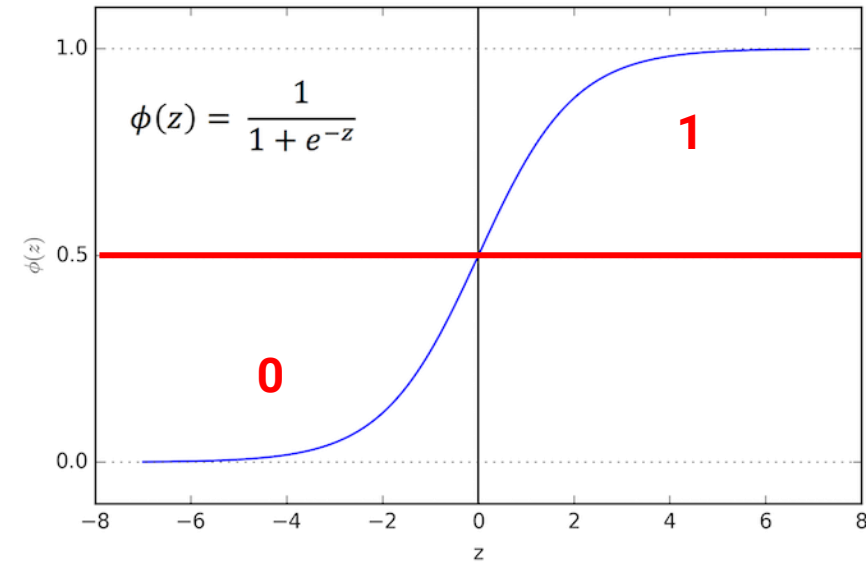
Patient 8902 – Performance

- SS:100%, FPR/h=0.11

- Logistic Regression with 7 features

- All seizures during the night. Sleep related?

- Surgery decision: not offered

- Focus localization=tpl
  - Temporal polar left

	Onset	Type	Pattern	Vigilance state at onset
Training Seizures	Day 0 00:51:14	Unclassified	Rhythmic alpha waves	Awake
	Day 1 00:03:23	FOIA	Rhythmic beta waves	Awake
	Day 1 06:37:05	FOIA	Rhythmic alpha waves	Awake
Testing Seizures	Day 2 01:35:56	FOIA	Amplitude depression	Awake
	Day 2 06:10:26	FOIA	Rhythmic alpha waves	Awake

	Onset	Predicted	#False Alarms	Hours of analysed data
Testing Seizures	Seizure 4 Day 2 01:35:56	Yes	1	≈17
	Seizure 5 Day 2 06:10:26	Yes	1	≈4

Patient 8902 –Regression coefficients

- What can we say about these features?

- Ratios between gamma bands

- Strong influence of the gamma bands

- Electrodes: F8 and T7

- Surgery decision: not offered

- Focus localization= tpl
    Temporal polar left
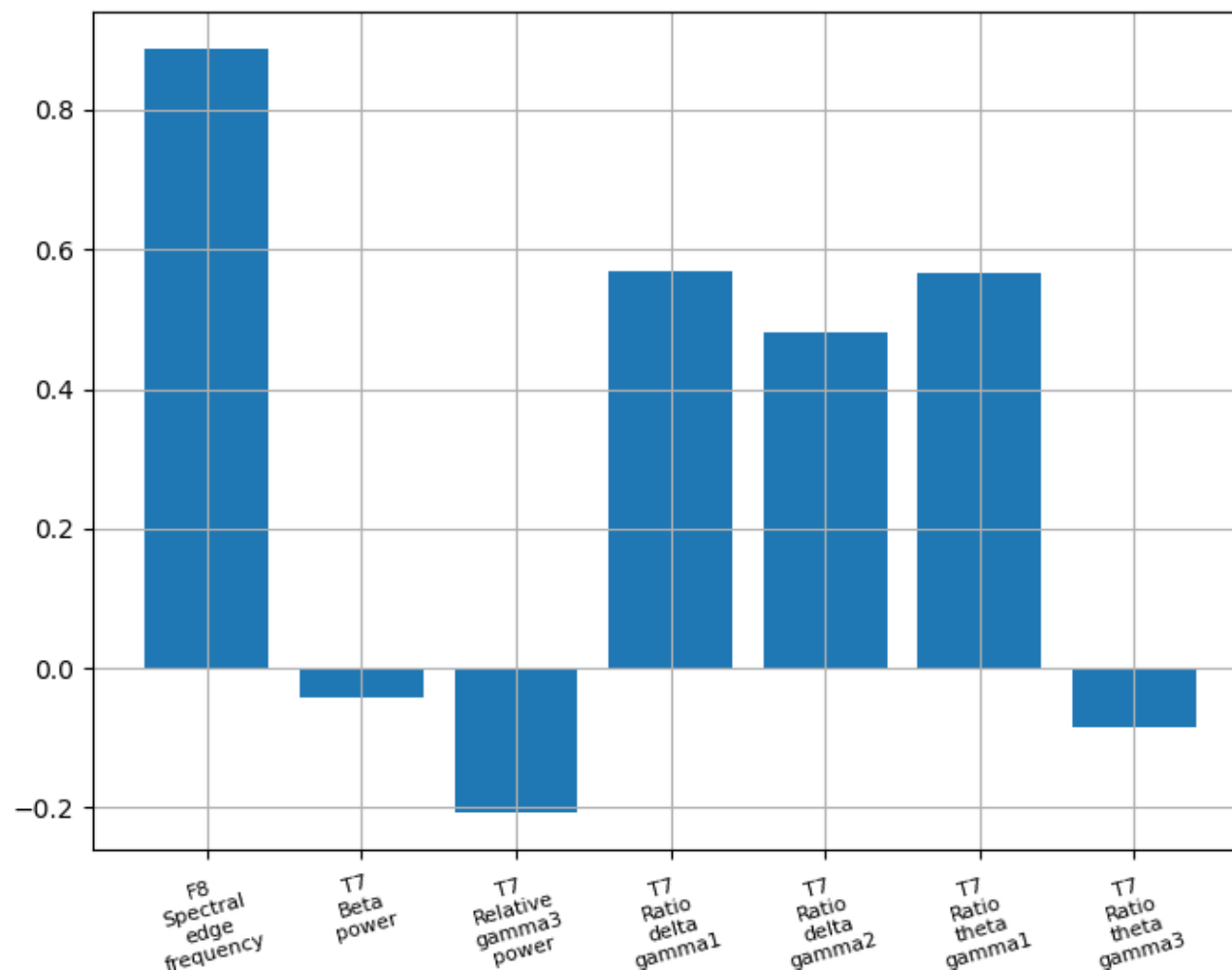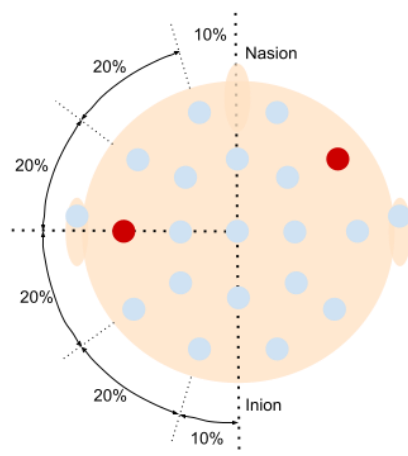
Patient 8902 – Plot over time

- Point A (10h00-10h30):
  - FP: 0
  - Morning control

- Point B (13h50-14h15):
  - Small peak
  - Related to lunch, sleep after lunch?

- Point C (21h30-22h00):
  - FP: 0
  - But interestingly, the Firing Power curve is already going down

- Point D (23h30-01h00):
  - FP started to rise at 00h00.
  - Capture the before 00h00 and the ungoing
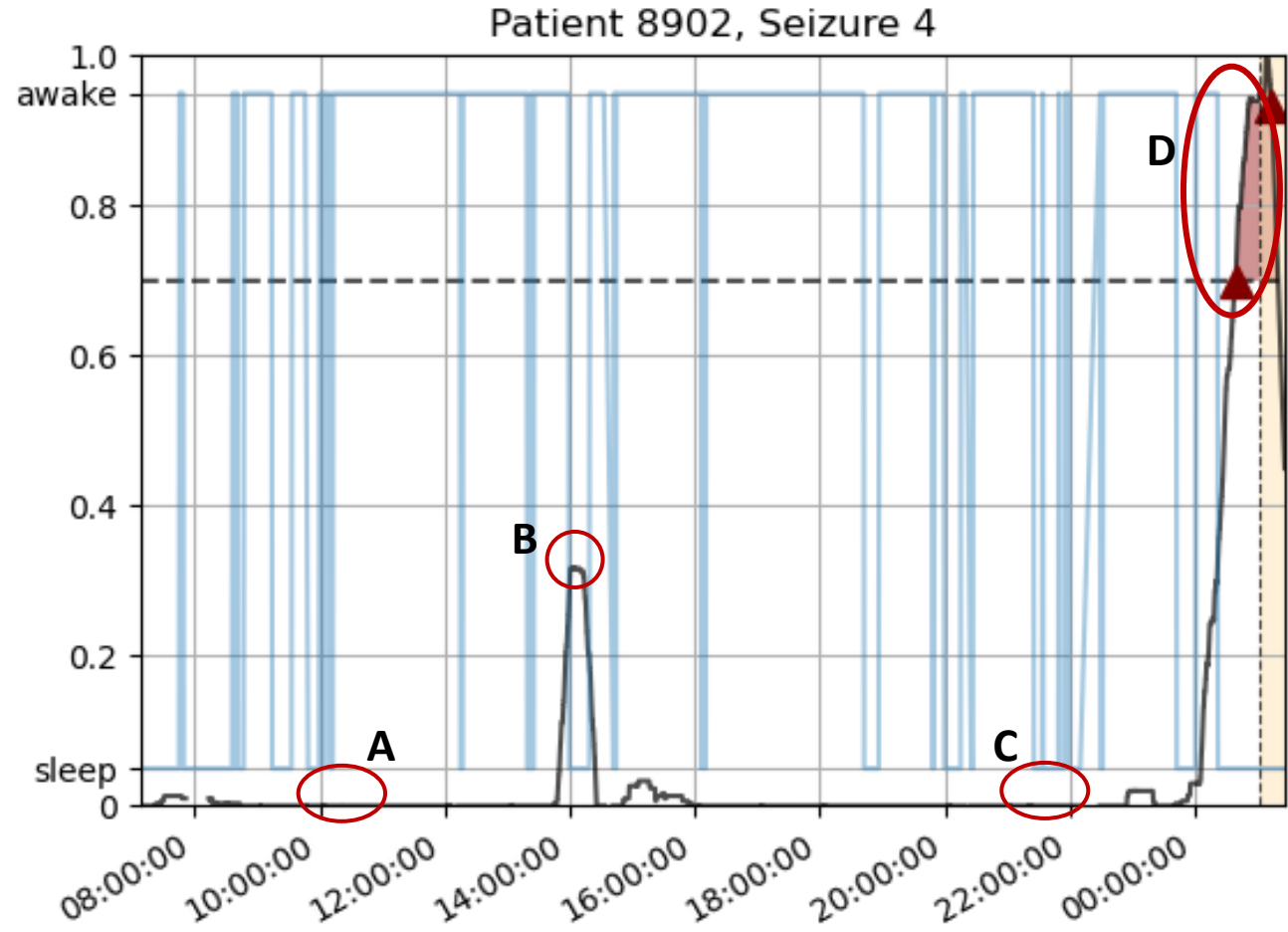
Patient 8902, Seizure 4

Patient 8902 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to have the following features (top 3 by order of importance):

  - T7_Beta_power: ↑

  - T7_Ratio_theta_gamma3: ↓

  - T7_Relative_gamma3_power: ↓

- Or, there was a peak because:

  - T7_Beta_power was low

  - T7_Ratio_theta_gamma3 were high

  - T7_Relative_gamma3_power were high
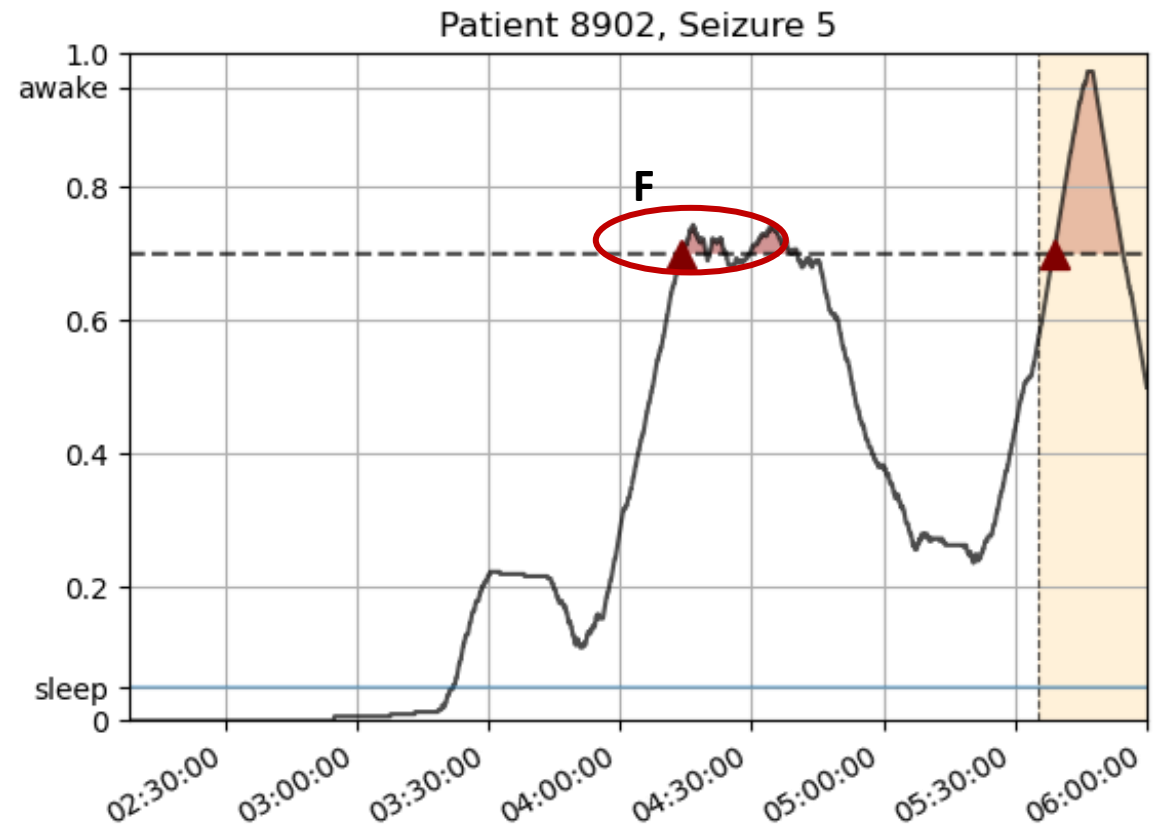

Patient 8902, Seizure 4
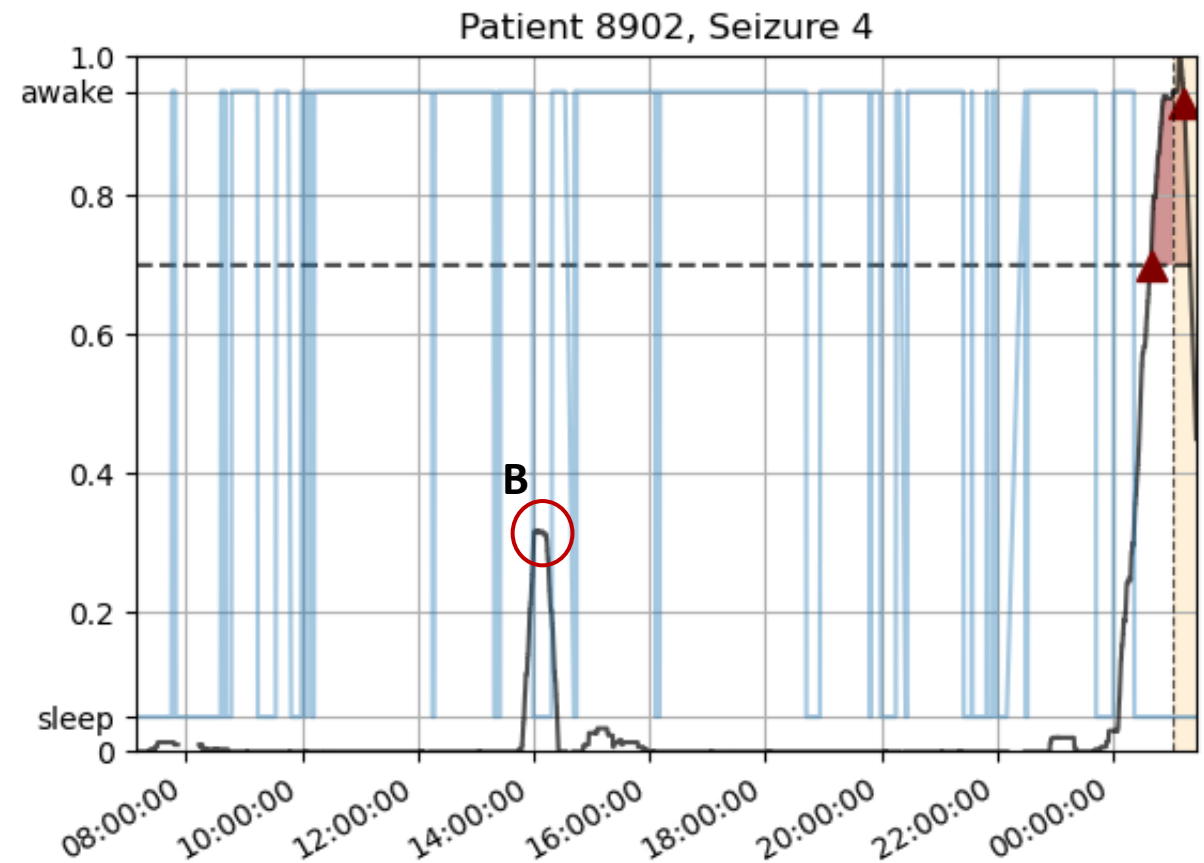
Patient 8902 – Counterfactual explanations

- For the prediction to be different (no alarms), we would have to have the following features (top 3 by order of importance):

    - T7_Relative_gamma3_power: ↑

    - T7_Beta_power: ↑

    - T7_Ratio_theta_gamma3: ↑

- Or, there was a peak because:

    - T7_Relative_gamma3_power were low

    - T7_Beta_power were low

    - T7_Ratio_theta_gamma3 were low

Patient 8902, Seizure 4
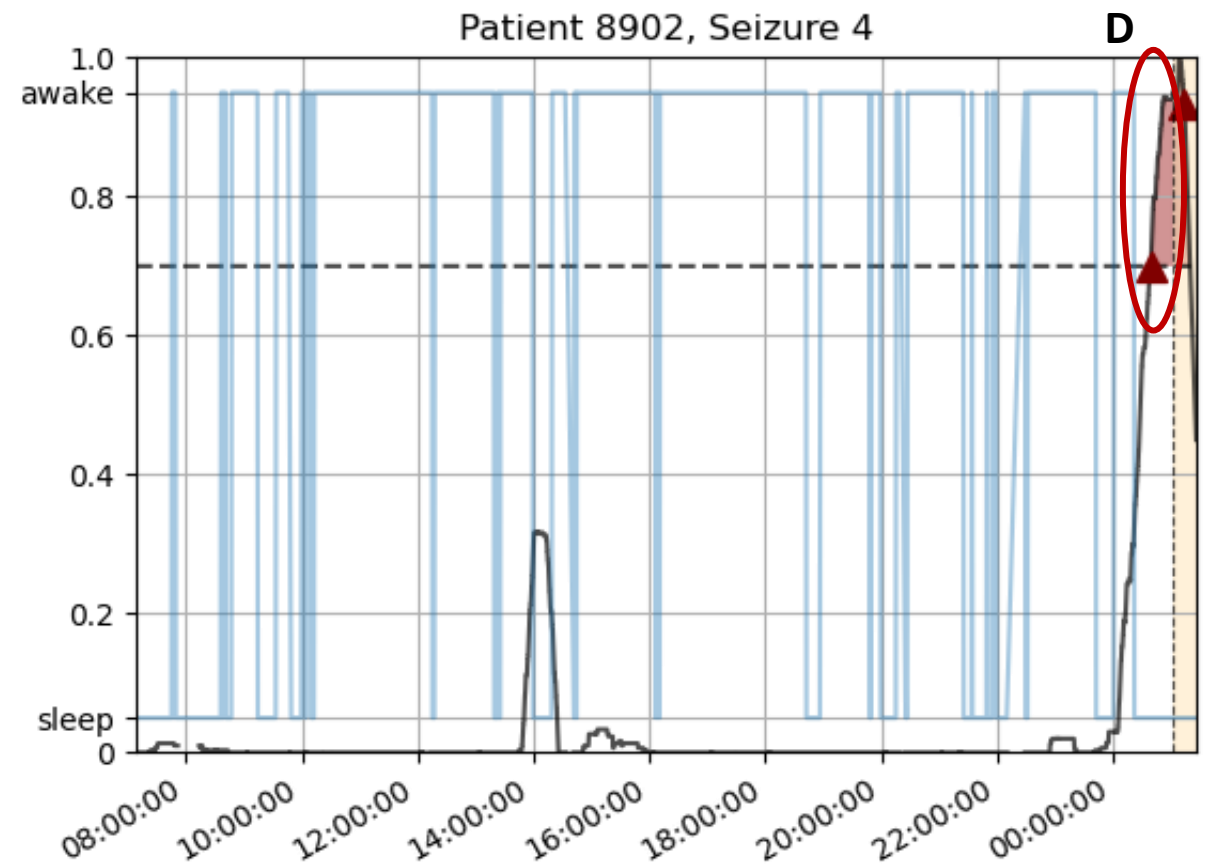
Patient 8902 – Over time

- Point E (02h30-03h00):
  - FP=0
  - Control

- Point F (04h00-04h40):
  - A false alarm

- Point G (05h00-05h45):
  - A control part
  - And then FP rises and an alarm is correctly triggered
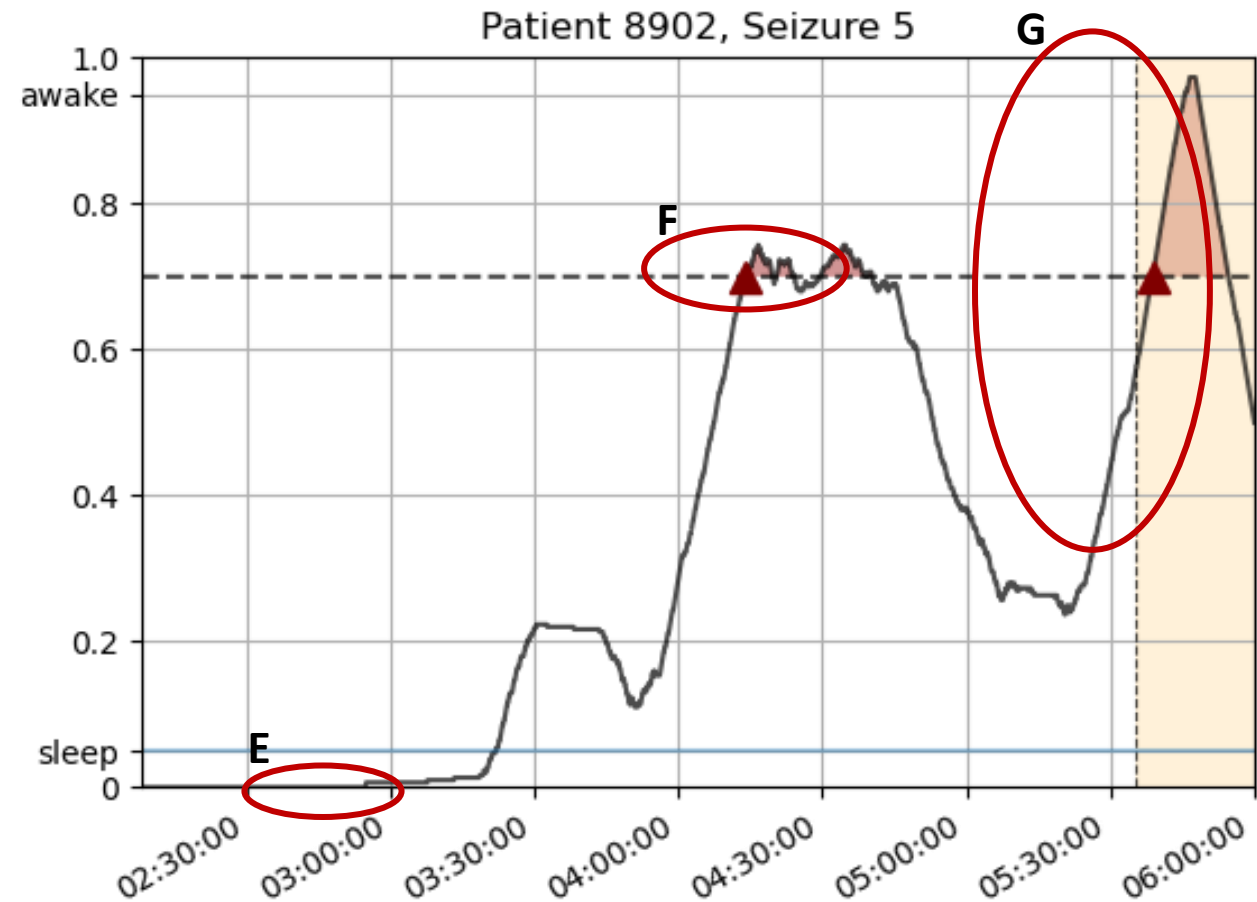
Patient 8902, Seizure 5

Patient 8902 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to have the following features (top 3 by order of importance):

  - F8_Spectral_edge_frequency: ↓

  - T7_Ratio_delta_gamma1: ↓

  - T7_Ratio_theta_gamma1: ↓

- Or, there was a peak because:

  - F8_Spectral_edge_frequency were high

  - T7_Ratio_delta_gamma1 were high

  - T7_Ratio_theta_gamma1 were high

Patient 8902, Seizure 5

Patient 8902 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to change the following features (top 3 by order of importance):

    - T7_Ratio_theta_gamma3: ↓

    - T7_Relative_gamma3_power: ↑

    - F8_Spectral_edge_frequency: ↓

- Or, there was a peak because:

    - T7_Ratio_theta_gamma3 were high

    - T7_Relative_gamma3_power were low

    - F8_Spectral_edge_frequency were high
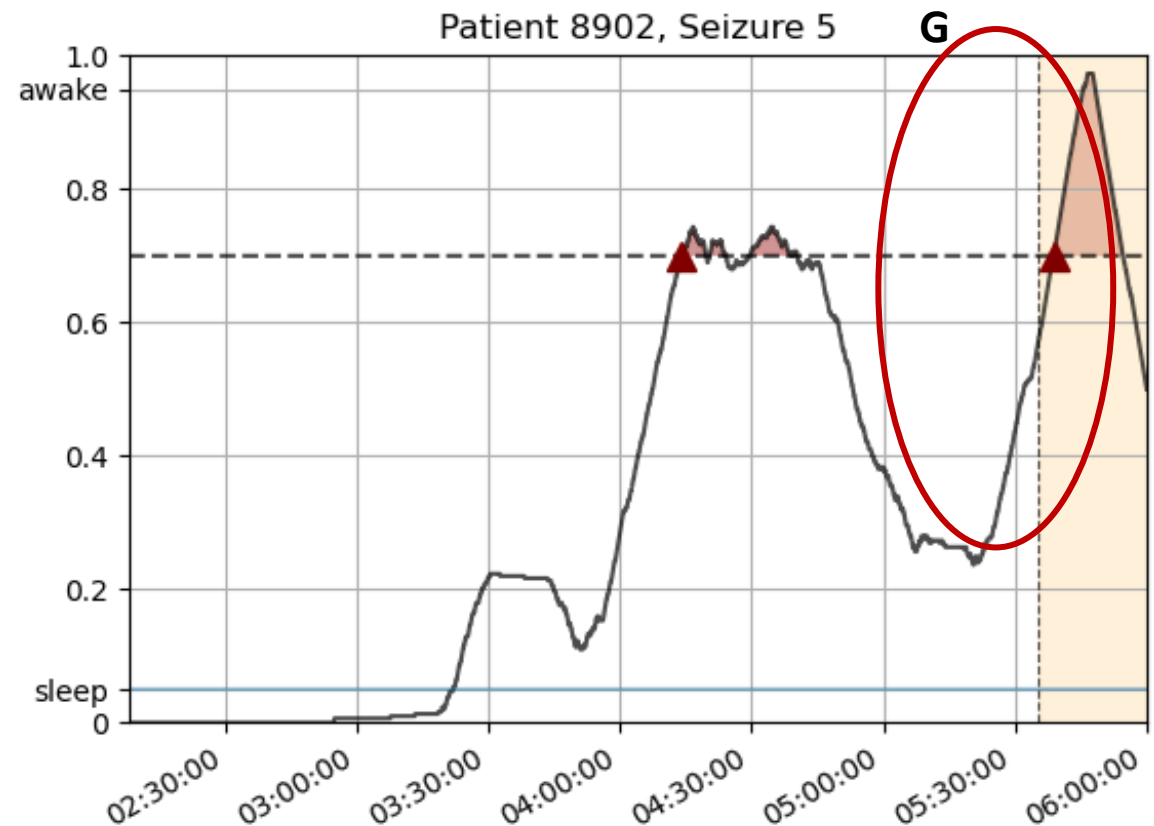
Patient 93402 – Overall performance

- SS:100%, FPR/h=0.50

- Above chance but FPR/h too high

- Logistic Regression with 3 features

- Surgery decision: not offered

- Focus localization= tpl
  Temporal polar left

	Onset	Type	Pattern	Vigilance state at onset
Training Seizures	Day 0 22:17:50	FBTC	Rhythmic theta waves	Non-REM II
	Day 1 10:21:34	FOIA	Rhythmic theta waves	Non-REM II
	Day 1 23:20:24	FOIA	Rhythmic theta waves	Non-REM II
Testing Seizures	Day 4 00:59:09	Unclassified	Rhythmic theta waves	Non-REM II
	Day 4 06:26:26	Unclassified	Rhythmic theta waves	Non-REM II

	Onset	Predicted	#False Alarms	Hours of analysed data
Testing Seizures	Seizure 4 Day 4 00:59:09	Yes	15	≈49
	Seizure 5 Day 4 06:26:26	Yes	6	≈5

Patient 93402 – Analysing regression coefficients

- What can we say about these features?

- Theta band power, which makes a lot of sense

- Influence of the gamma bands

- Electrodes: C3 , P3, and O2.
 - C3 and P3 make sense. Near focus
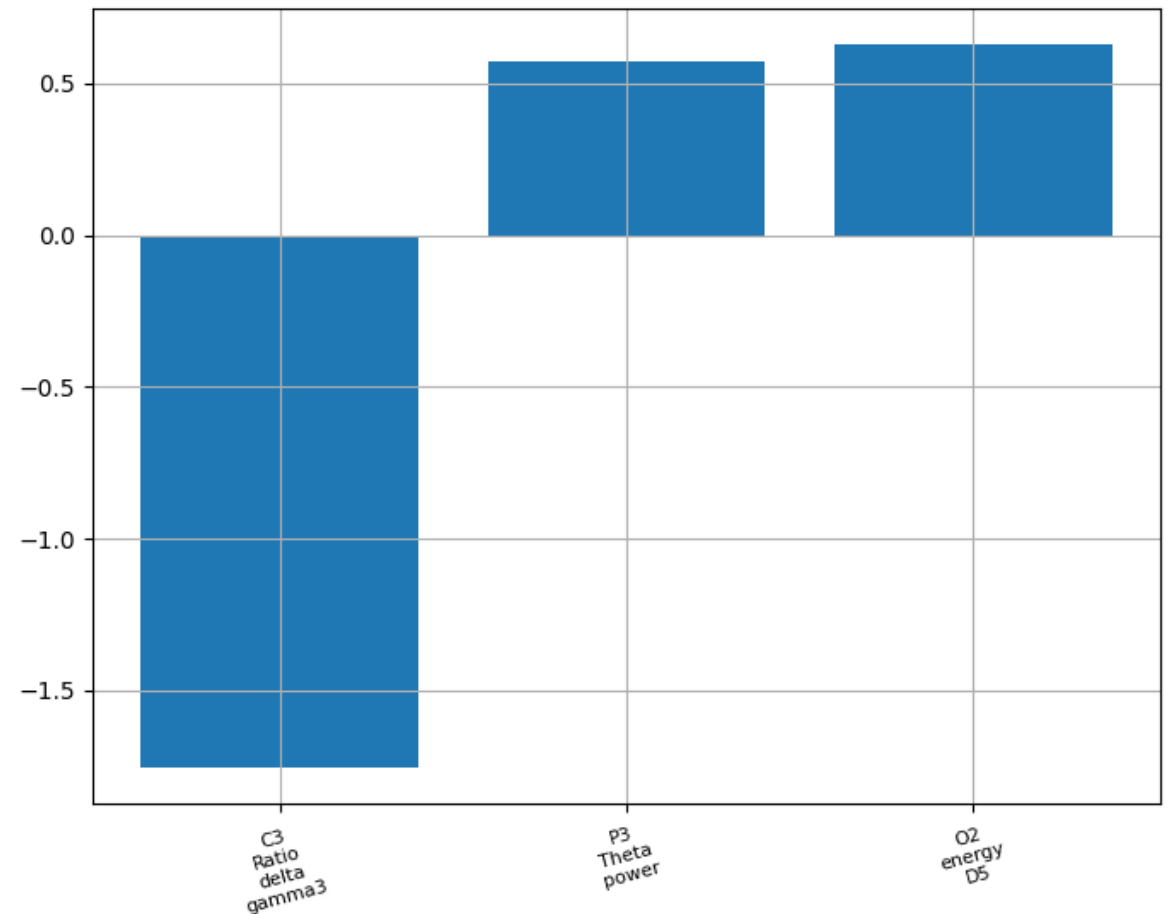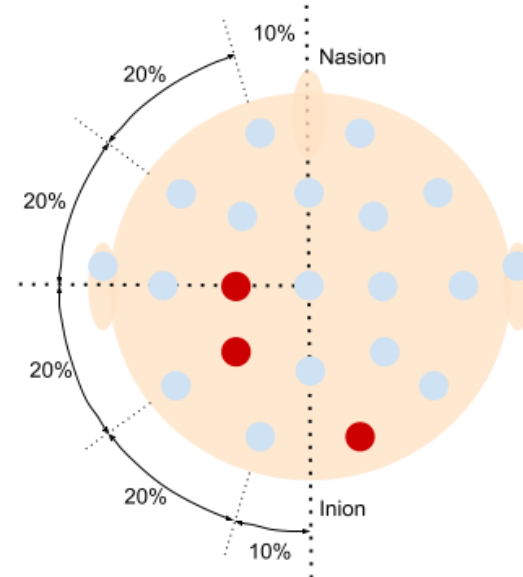
Patient 93402 – Over time

- Cluster A (00h00-04h00, Day 0):
  - A series of false alarms

- Cluster B (00h00-06h00, Day 1):
  - A series of false alarms

- Point D (23h20-00h35, Day 2):
  - False alarms but seizure predicted

- Green and Yellow circles

- Point C (14h00-18h00, Day 2):
  - FP=0
  - Control method

- Circadian cycles influence? Sleep stages?

- Seizure susceptibility?

Patient 93402, Seizure 4
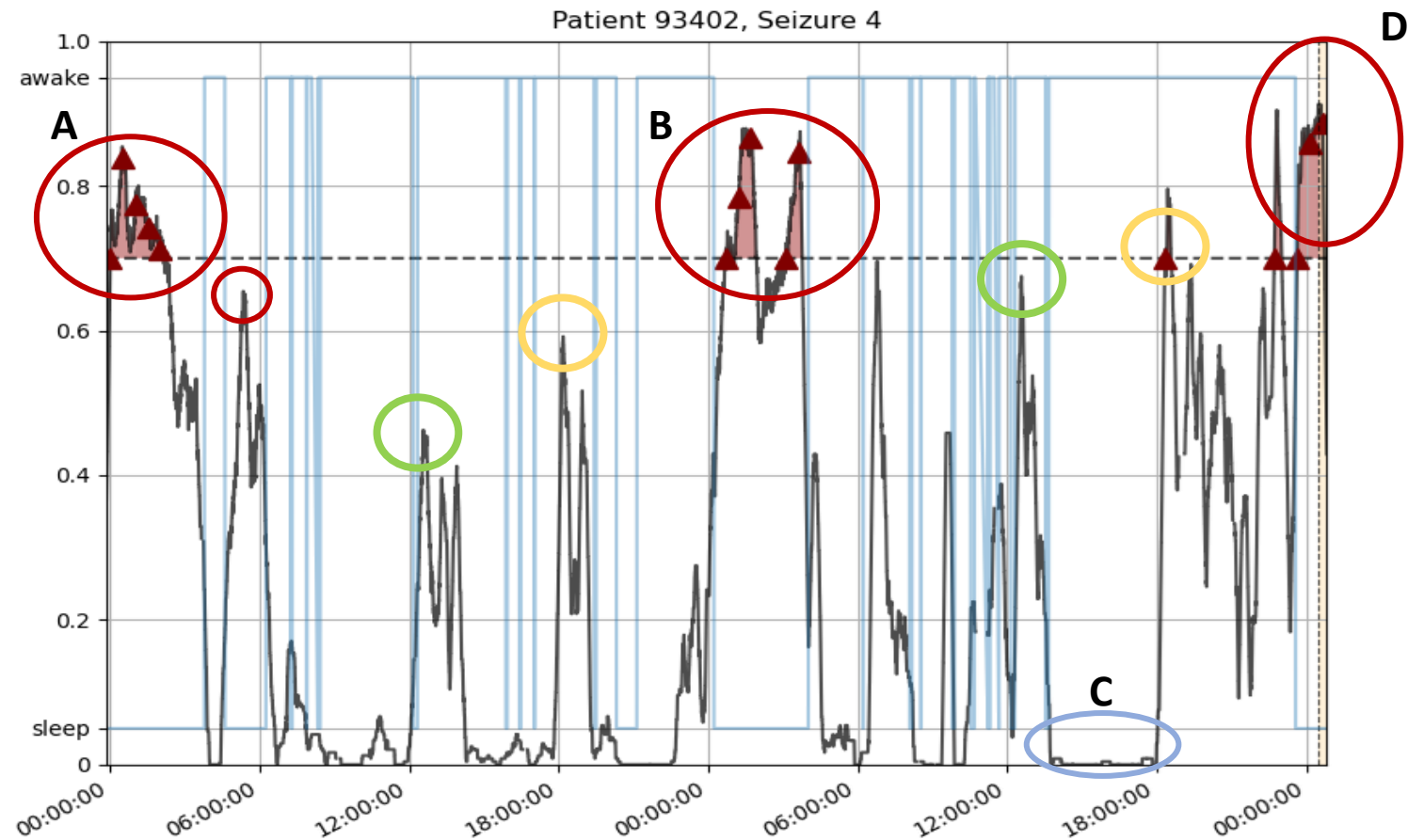
Patient 93402 – Over time

- Point E (01h30-02h00):
  - Two false alarms

- Point F (02h30-03h30):
  - Two false alarms

- Point G (4h00-06h00):
  - From a decreasing and low FP
  - To a rising FP, false alarms, and predicted Seizure
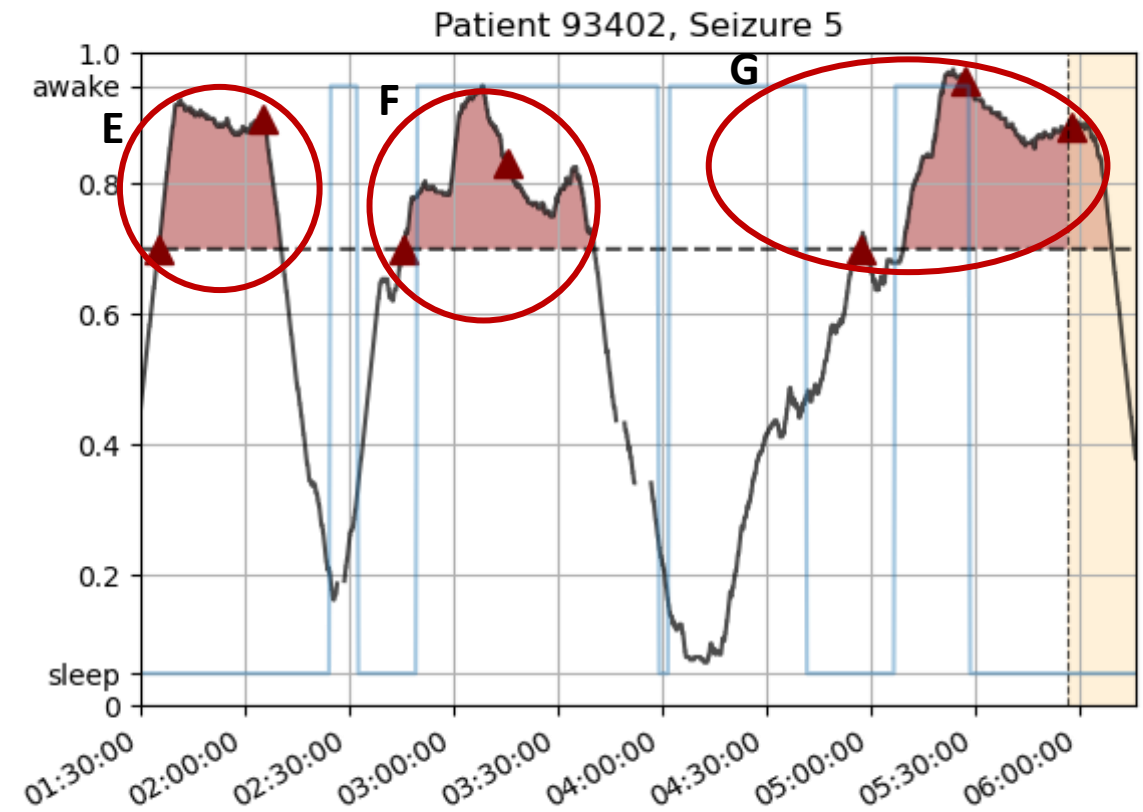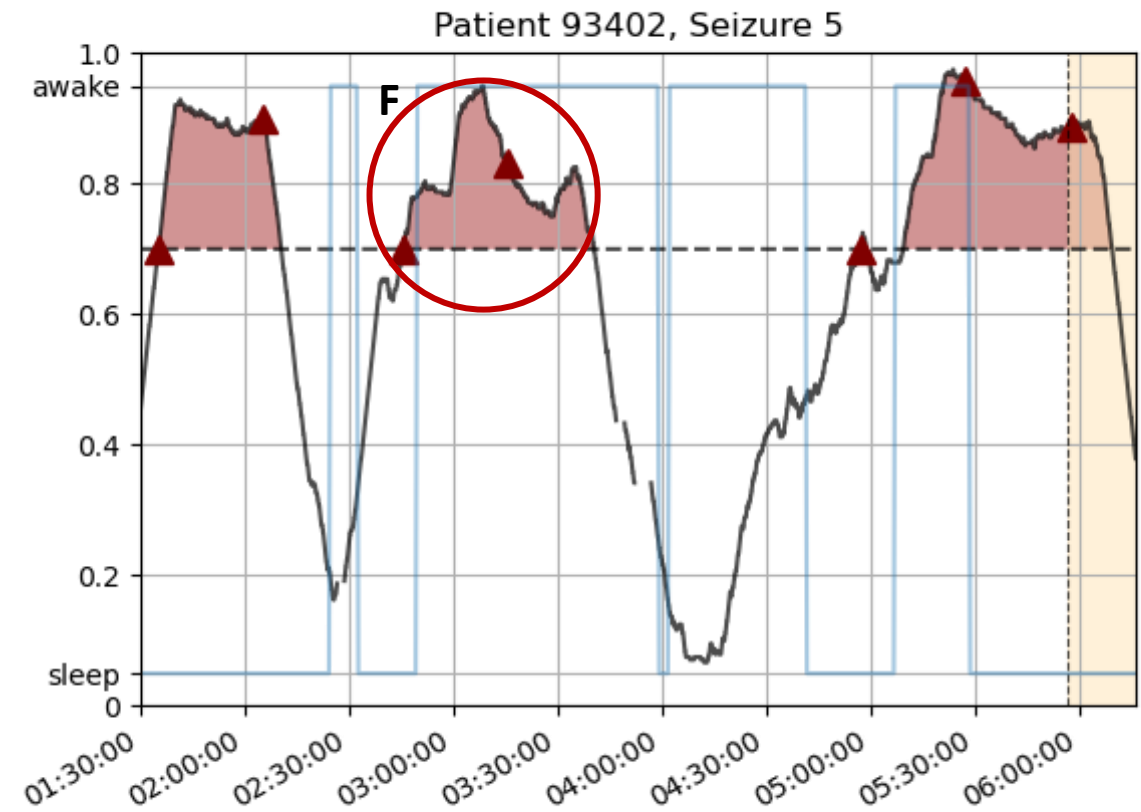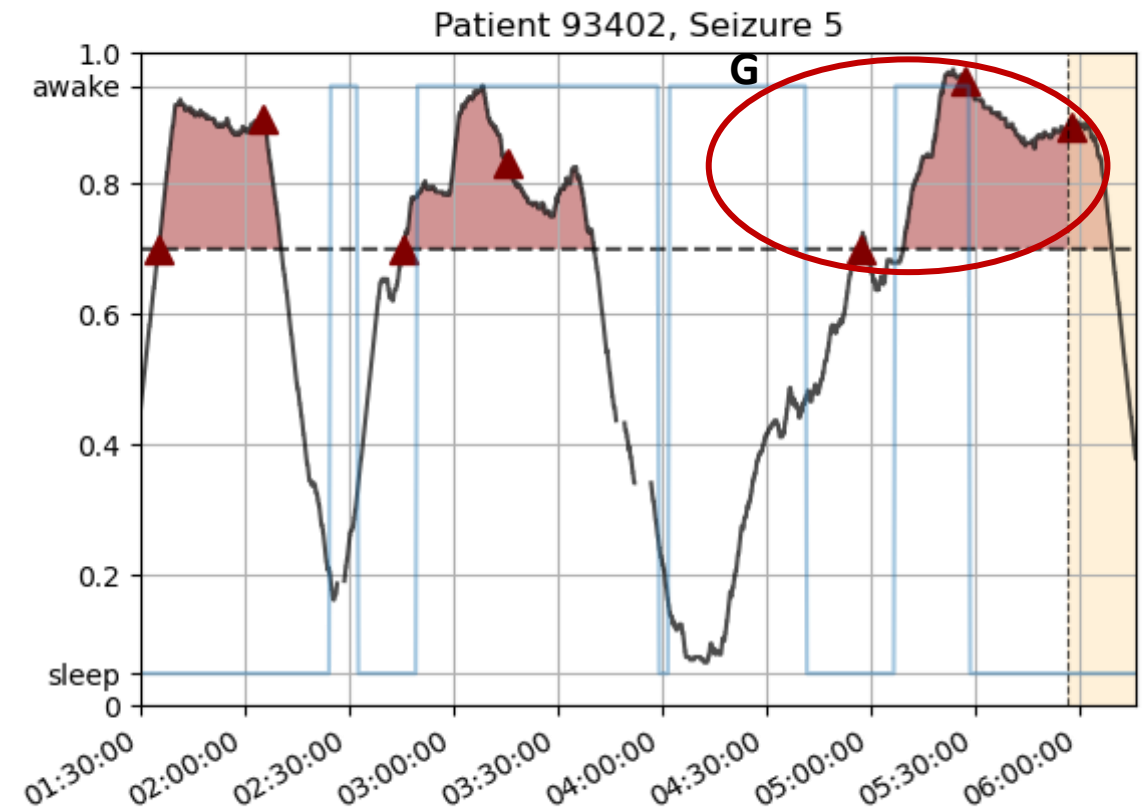

Patient 93402, Seizure 5

Patient 93402 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to have the following features (top 3 by order of importance):

  - C3_Ratio_delta_gamma3: ↑

  - O2_energy_D5: ↓

  - P3_Theta_power↓

- Or, there was a peak because:

  - C3_Ratio_delta_gamma3 were low

  - O2_energy_D5 were high

  - P3_Theta_power were high

Patient 93402, Seizure 5

Patient 93402 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to have the following features (top 3 by order of importance):

  - C3_Ratio_delta_gamma3: ↑

  - O2_energy_D5: ↓

  - P3_Theta_power: ↓

- Or, there was a peak because:

  - C3_Ratio_delta_gamma3 was low

  - O2_energy_D5: was high
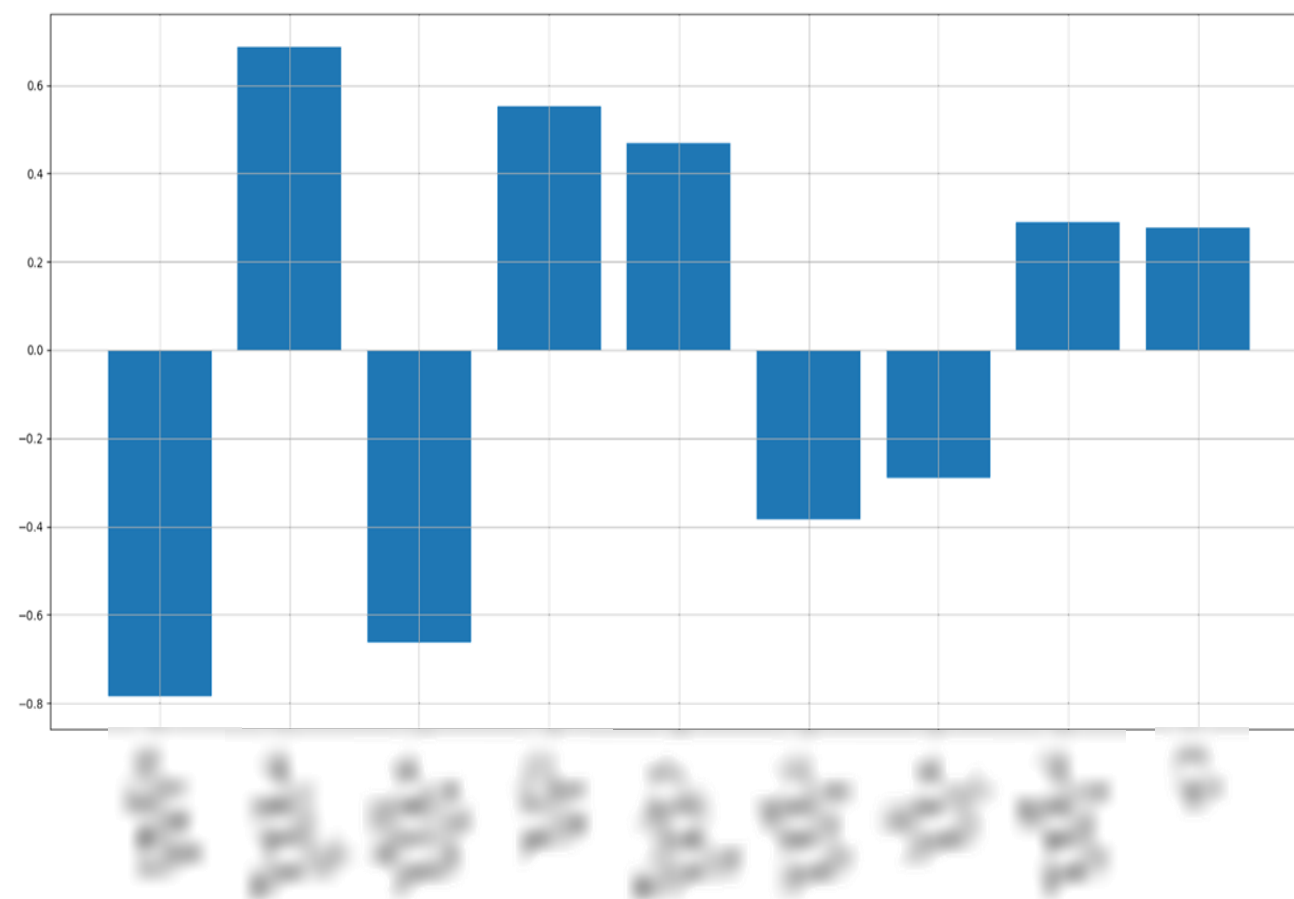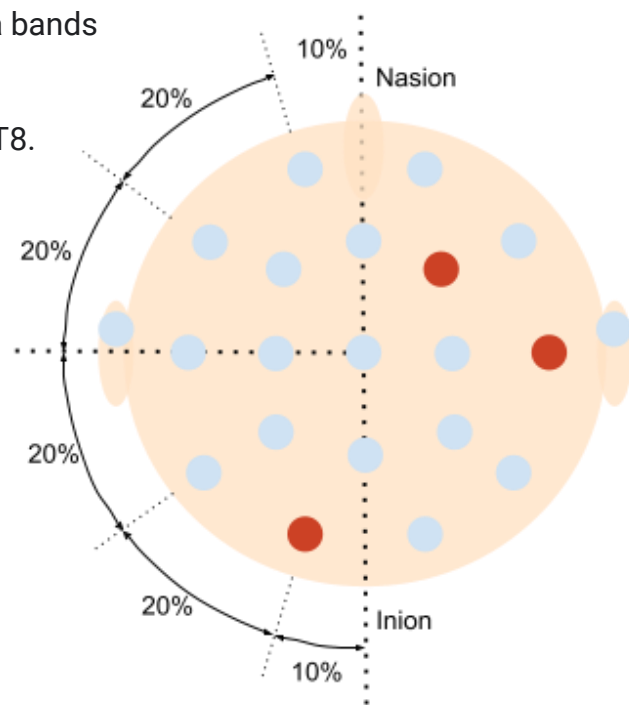
  - P3_Theta_power was high

Patient 93402, Seizure 5

Patient 93402 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to have the following features (top 3 by order of importance):

  - C3_Ratio_delta_gamma3: ↑

  - O2_energy_D5: ↓

  - P3_Theta_power: ↓

- Or, there was a peak because:

  - C3_Ratio_delta_gamma3 was low

  - O2_energy_D5 was high

  - P3_Theta_power was high

Patient 93402, Seizure 5

Patient 402 – Overall performance

- SS:0%, FPR/h=0.00

- No single alarm was raised

- Logistic Regression with 20 features

- Surgery decision: invasive

- Focus localization= t-r, t-l
    Temporal right, temporal left

	Onset	Type	Pattern	Vigilance state at onset
Training Seizures	Day 0 23:45:26	FOIA	Rhythmic theta waves	Awake
	Day 2 22:27:34	FBTC	Rhythmic theta waves	Awake
	Day 3 03:13:30	FOIA	Rhythmic theta waves	Awake
Testing Seizures	Day 3 09:53:21	FBTC	Rhythmic theta waves	Awake
	Day 4 09:57:27	FOIA	Rhythmic theta waves	Awake

|	Onset	Predicted	#False Alarms	Hours of analysed data
Testing Seizures	Day 3 09:53:21	No	0	≈6
	Day 4 09:57:27	No	0	≈24

Patient 402 – Analysing regression coefficients

- What can we say about these features?

- Different gama bands, kurtosis, activity, energy

- Influence of the gamma bands

- Electrodes: O1, F4 and T8.
  - Do these make sense?

Patient 402 – Over time

- Cluster A (04h30-05h10):
  - A peak of 0.47, sleep?

- Cluster B (08h55-09h35):
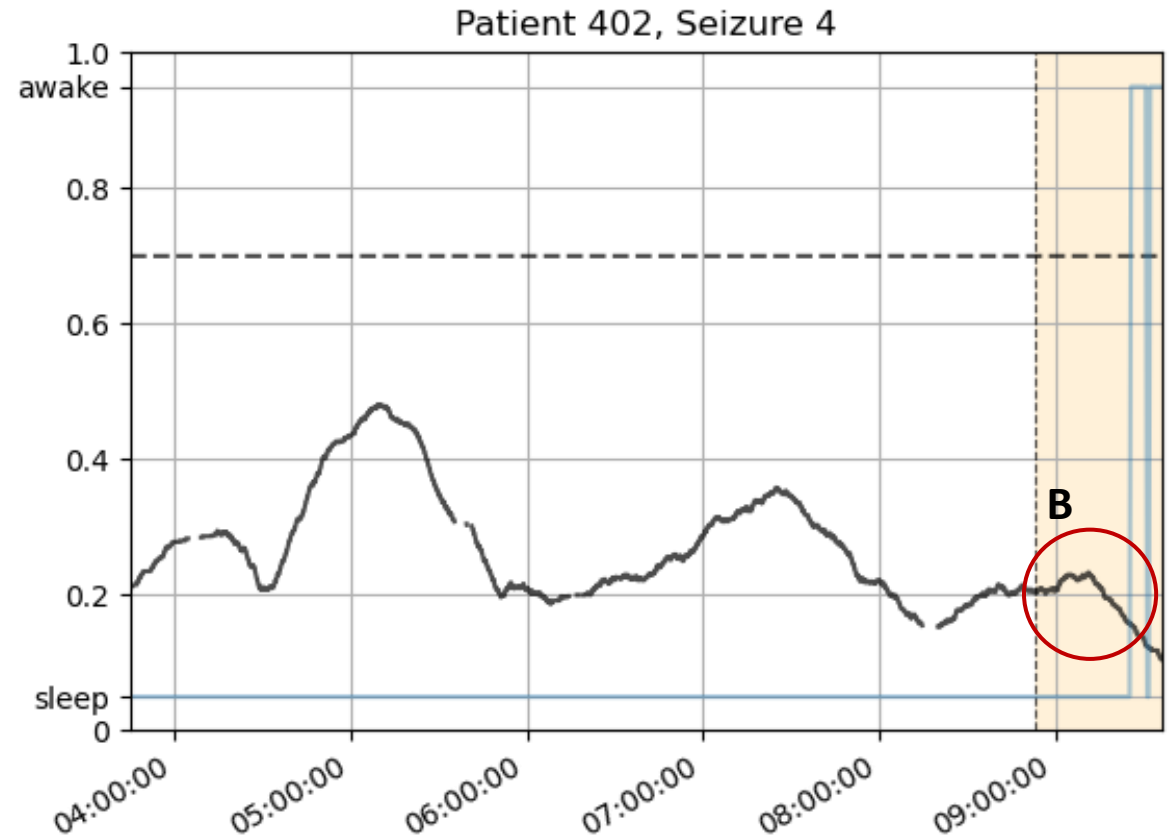 - Did not predict the seizure
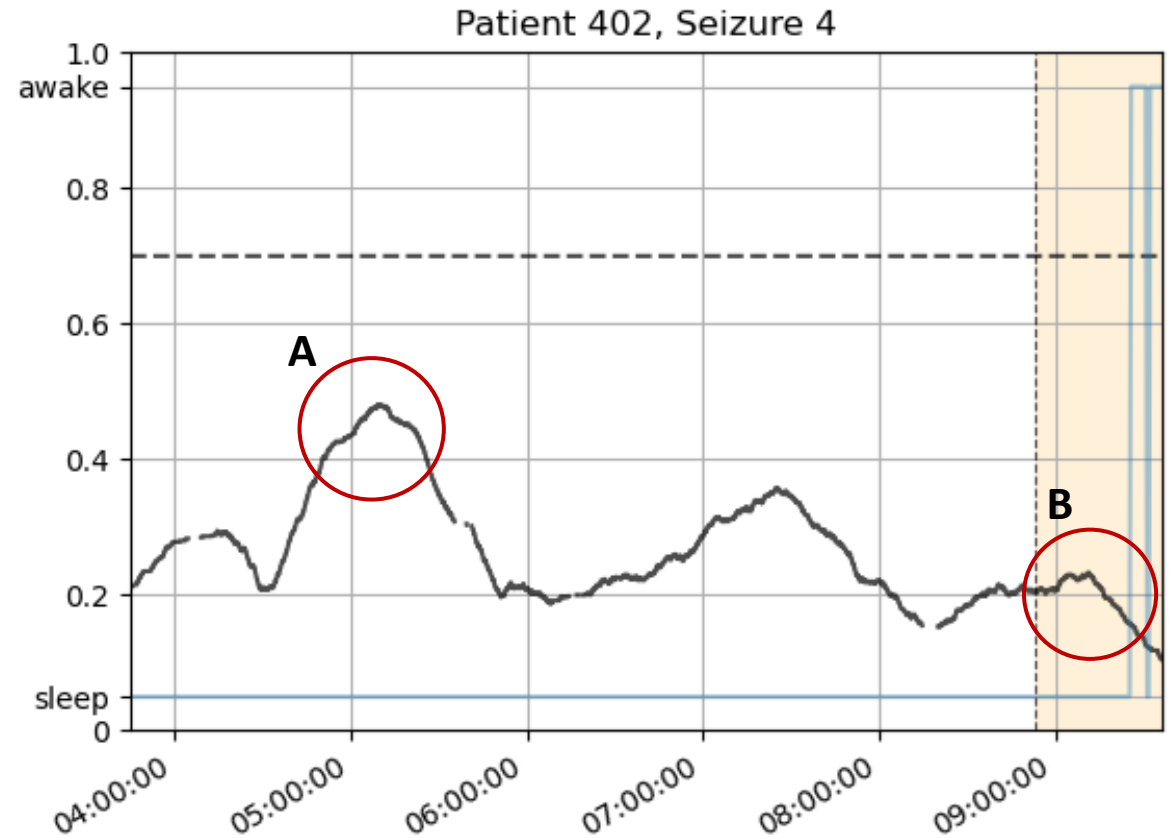
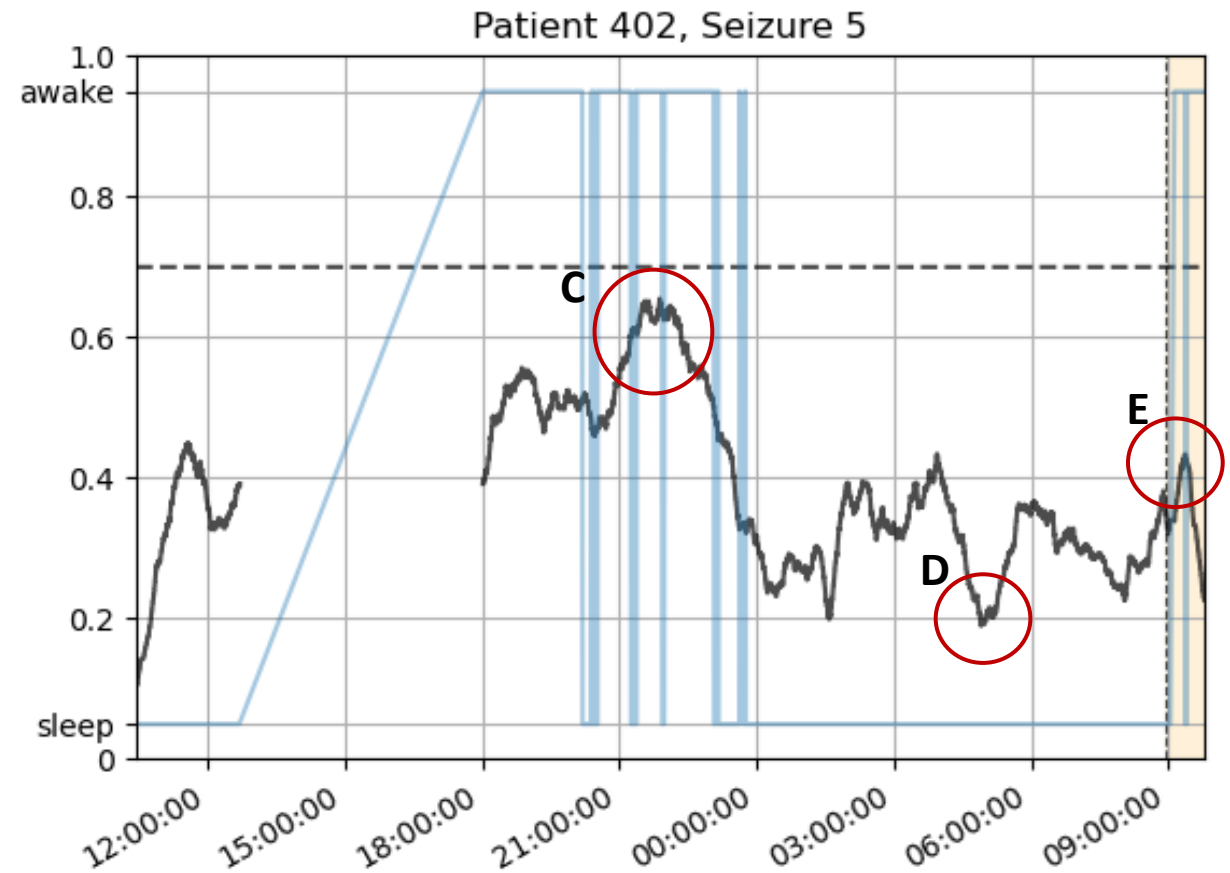
Patient 402, Seizure 4

Patient 402 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to change the following features (top 3 by order of importance):

  - O1_Ratio_delta_theta :↓

  - T8_Ratio_theta_gamma2: ↑

  - F4_Relative_gamma1_power: ↓

- Or, there was a peak because:

  - O1_Ratio_delta_theta were high

  - T8_Ratio_theta_gamma2 were low

  - F4_Relative_gamma1_power were high
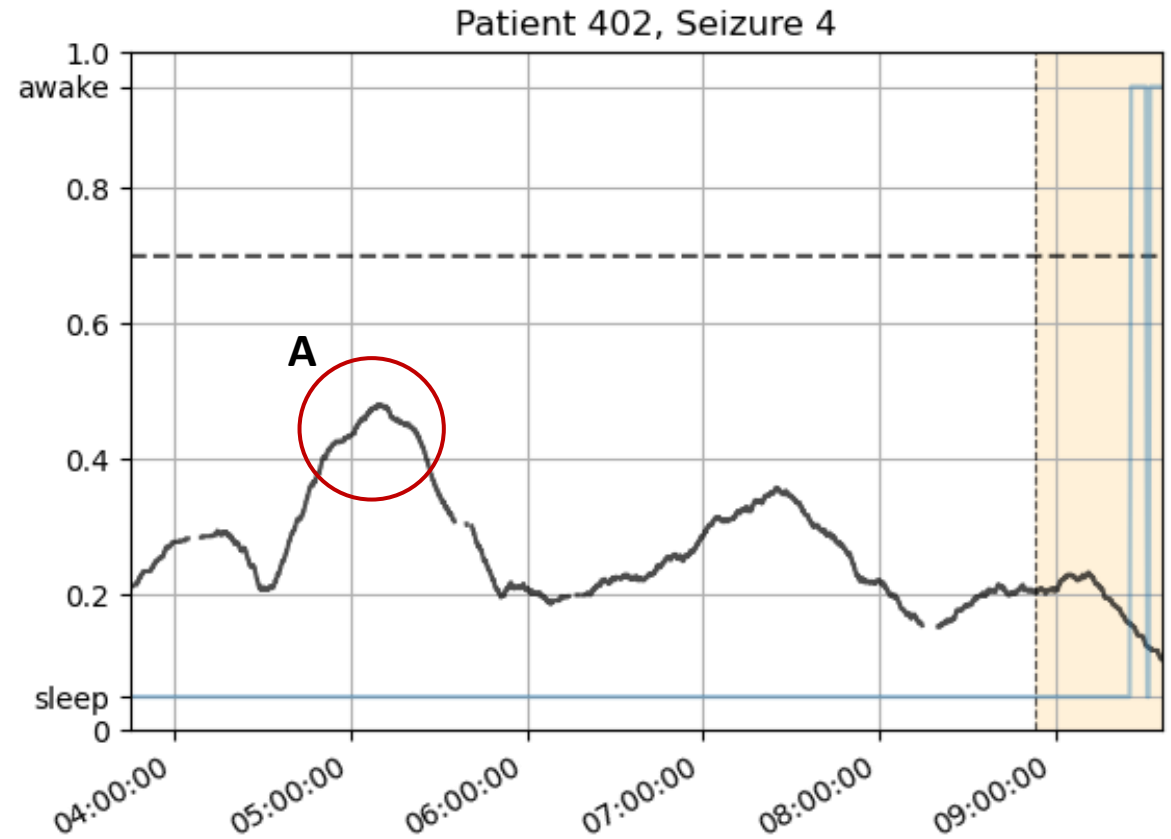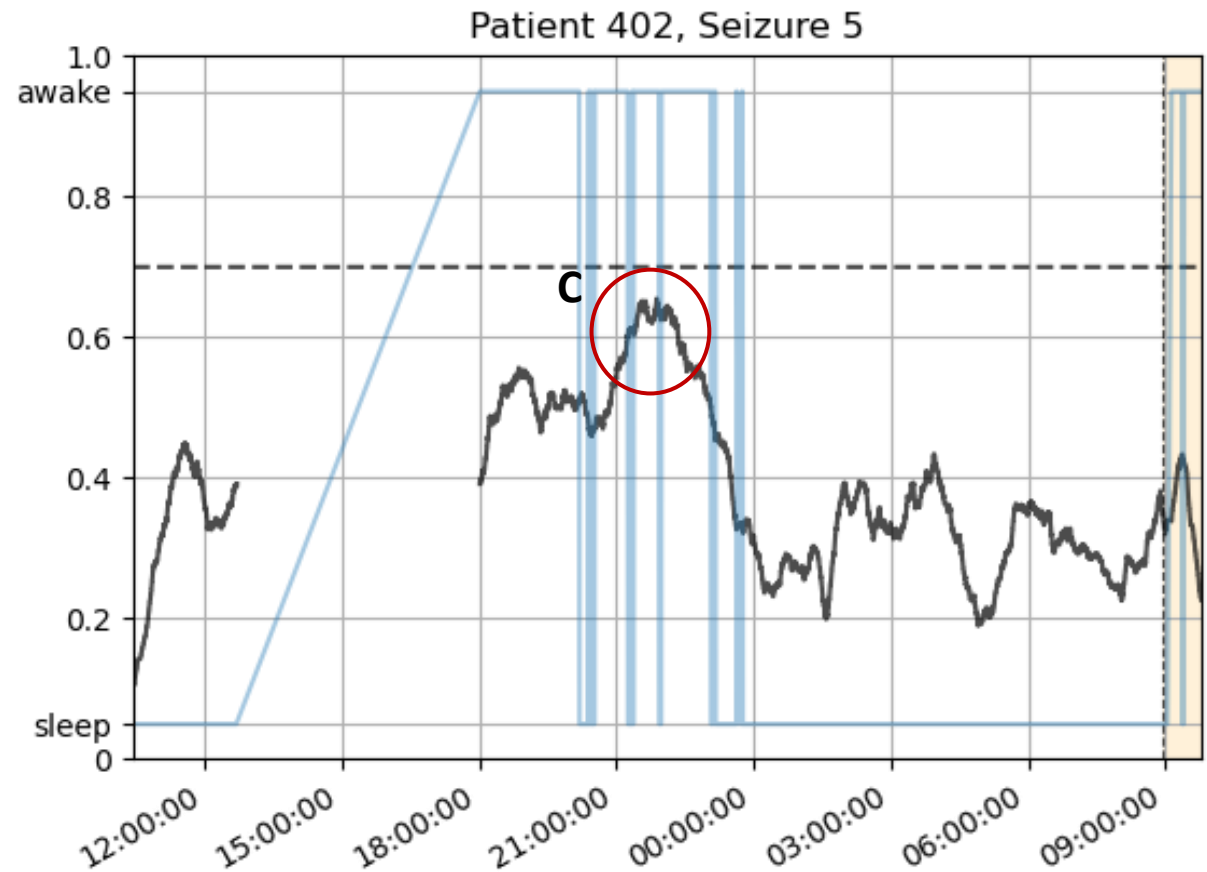

Patient 402, Seizure 4

Patient 402 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to change the following features (top 3 by order of importance):

  - F4_Kurtosis: ↑

  - O1_Ratio_delta_theta: ↓

  - T8_Ratio_theta_gamma2: ↑

- Or, there was a peak because:

  - F4_Kurtosis: were low

  - O1_Ratio_delta_theta: were high

  - T8_Ratio_theta_gamma2: were low

Patient 402, Seizure 4

Patient 402 – Over time

- Cluster C (21h00-00h00):
  - A peak of 0.66, sleep?

- Cluster E (08h10-09h40):
  - Again, at the same time, a seizure occurred but it was not predicted


Patient 402, Seizure 5

Patient 402 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to change the following features (top 3 by order of importance):

    - F4_Kurtosis: ↓

    - O1_Ratio_delta_theta: ↑

    - T8_Ratio_theta_gamma2: ↓

- Or, there was a peak because:

    - F4_Kurtosis was high

    - O1_Ratio_delta_theta was low

    - T8_Ratio_theta_gamma2 was high

Patient 402 – Counterfactual explanations

- For the prediction to be different (no peak), we would have to change the following features (top 3 by order of importance):

  - F4_Kurtosis: ↑
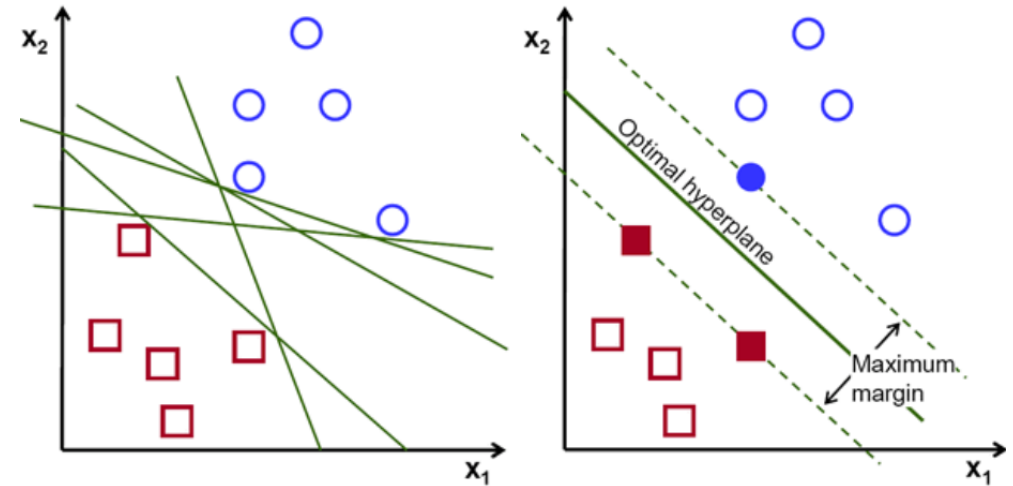
  - O1_Ratio_delta_theta: ↓

  - T8_Ratio_theta_gamma2: ↑

- Or, there was a peak because:

  - O1_Ratio_delta_theta was low

  - T8_Ratio_theta_gamma2 was high

  - F4_Relative_gamma1_power was low
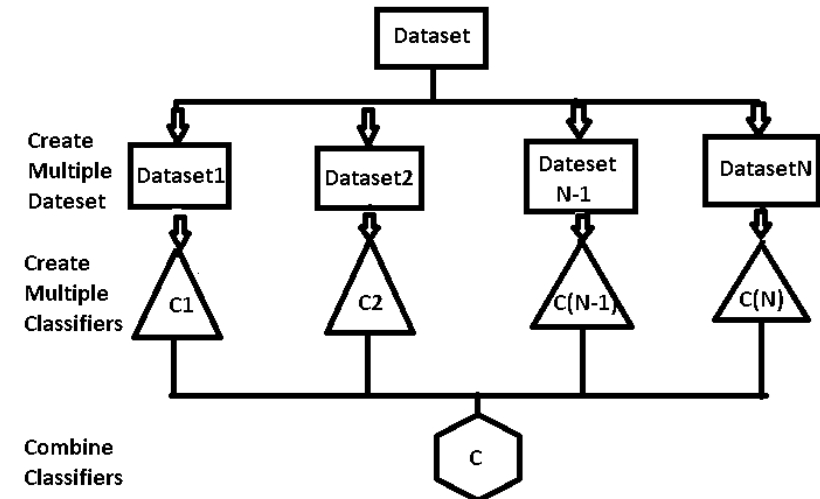

Patient 402, Seizure 5
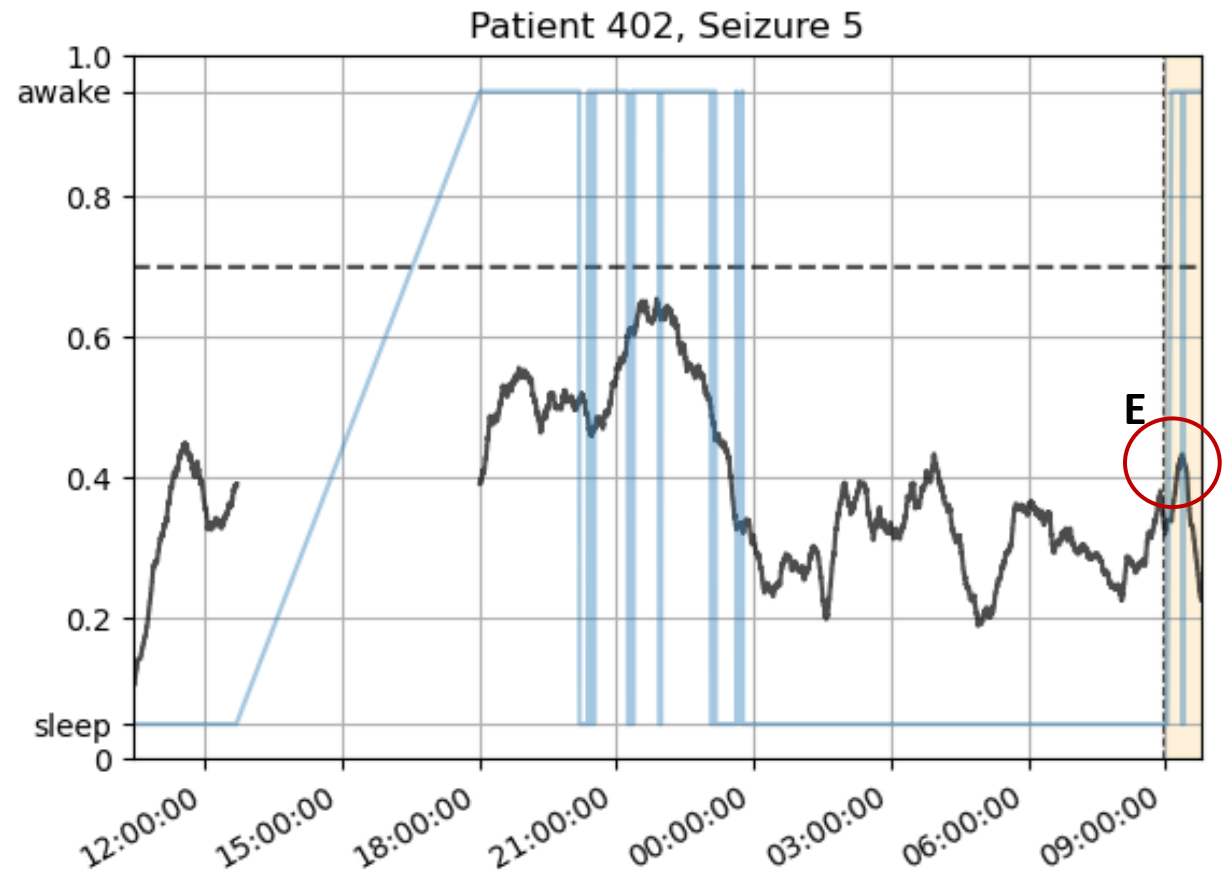
Support Vector Machines

- 1) With the first 3 seizures:
  - Find the optimal pre-ictal period within a range of values: 30-60 minutes
  - Find the optimal number of features

- 2) Do the following 15 times:
  - Balance the data: choose random samples from inter-ictal
  - Train an SVM classifier

- 3) Each classifier then makes a prediction. The prediction is obtained by a voting system
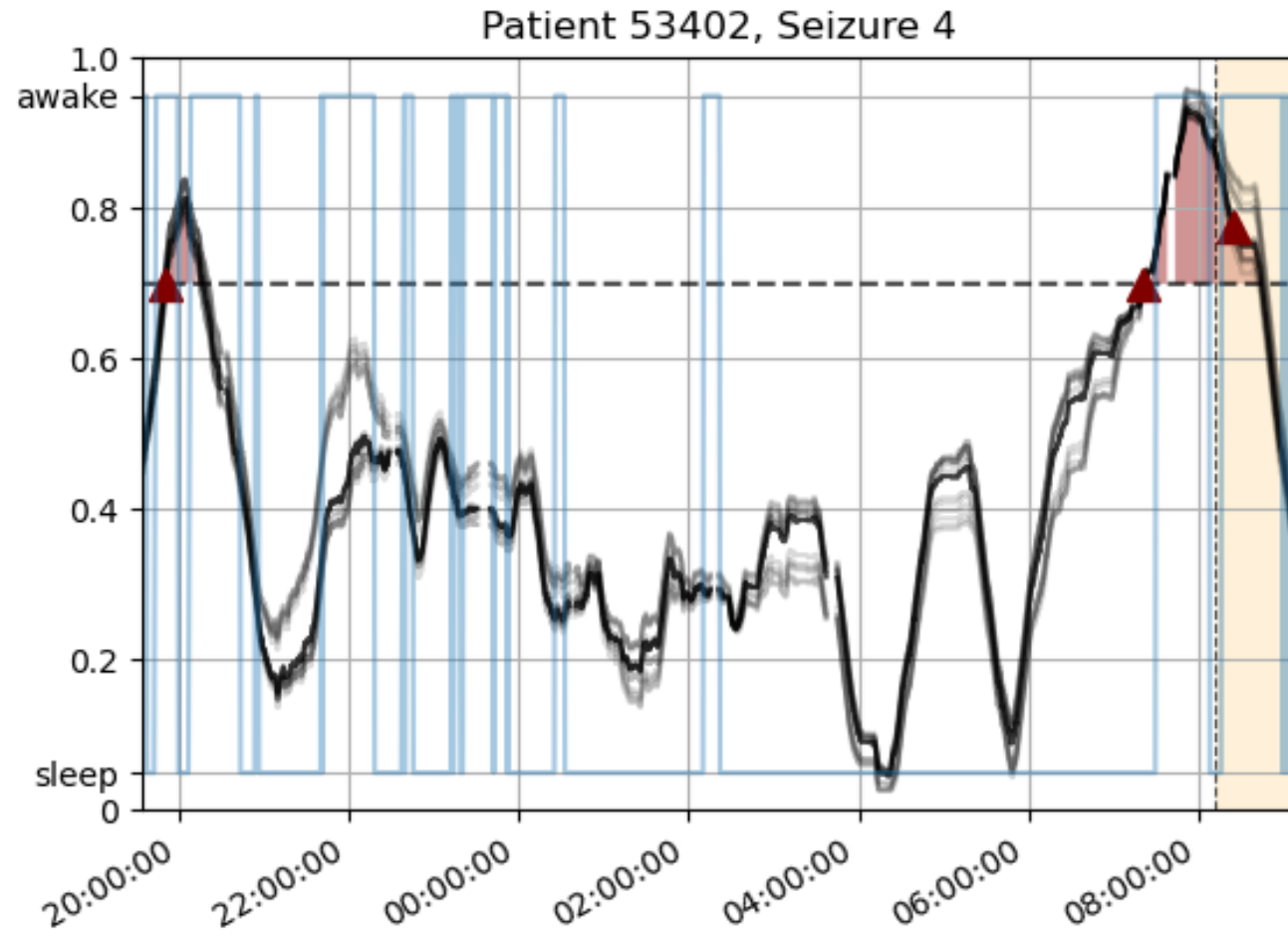

Source: https://medium.com/machine-learning-intuition/document-classification-part-3-detection-algorithm-support-vector-machines-gradient-descent-282316b0838e
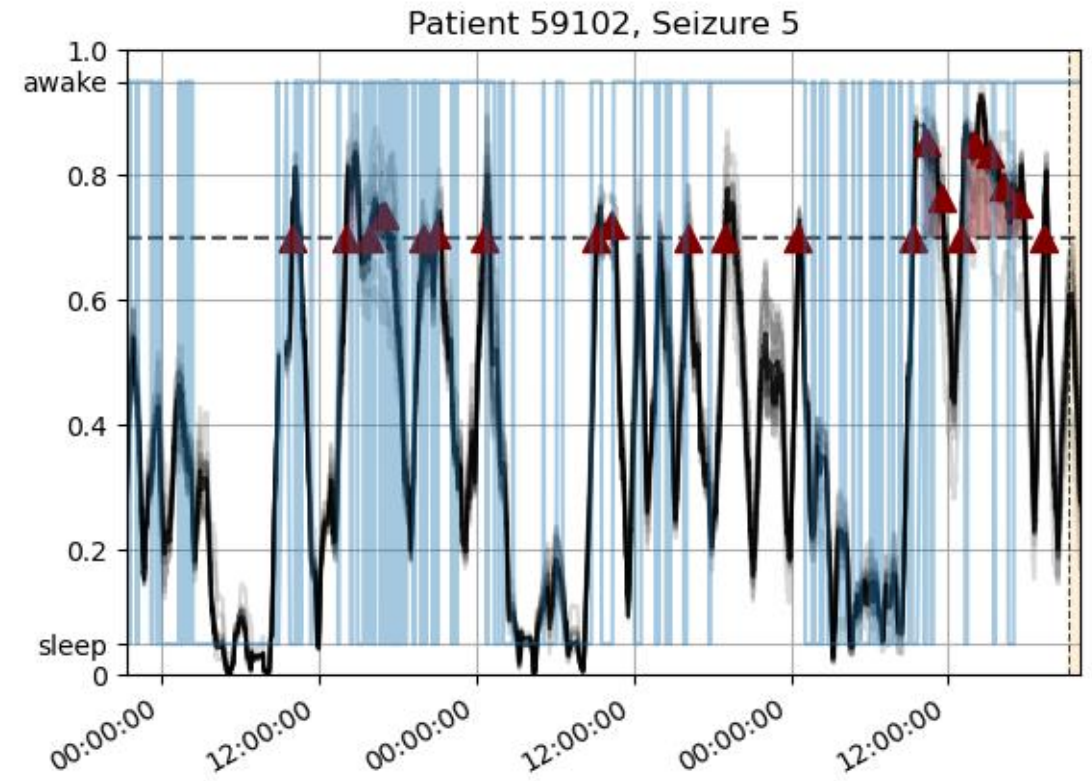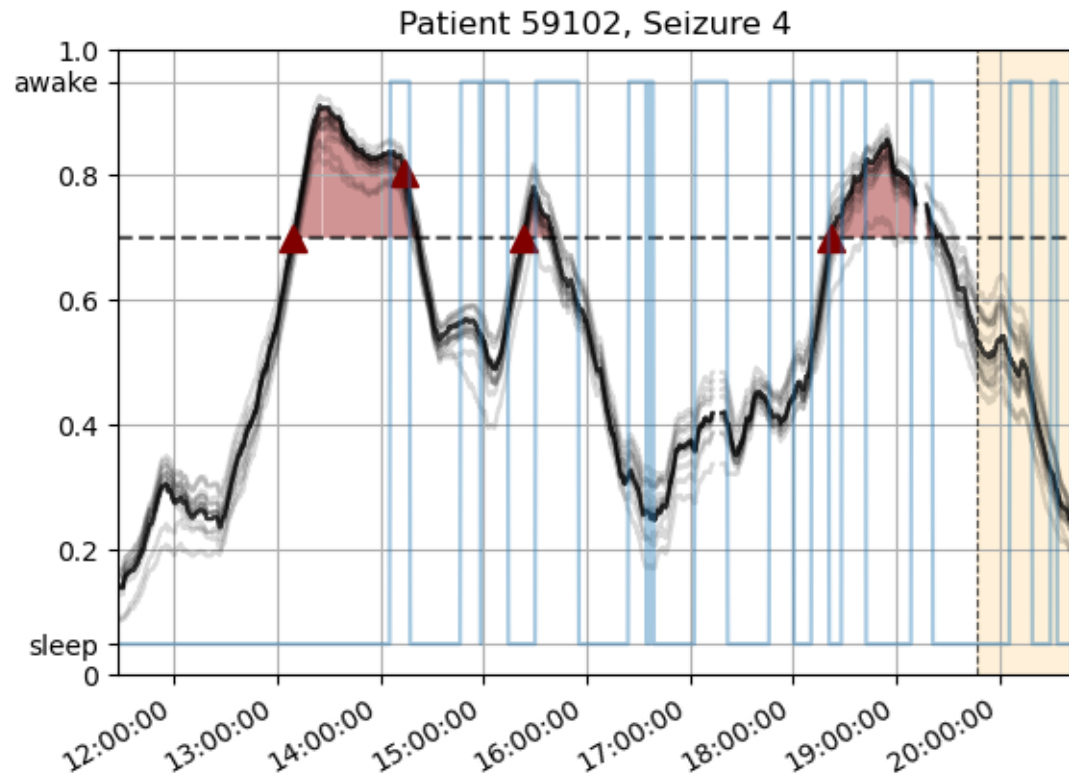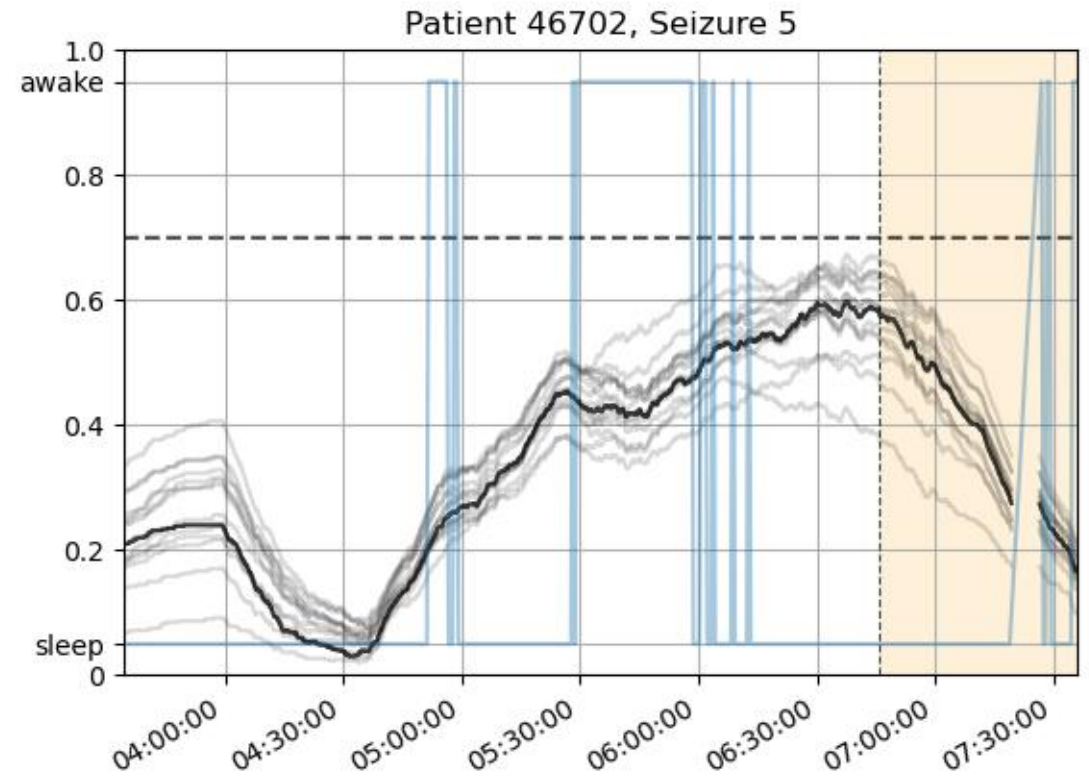

Source: https://www.geeksforgeeks.org/ensemble-classifier-data-mining/
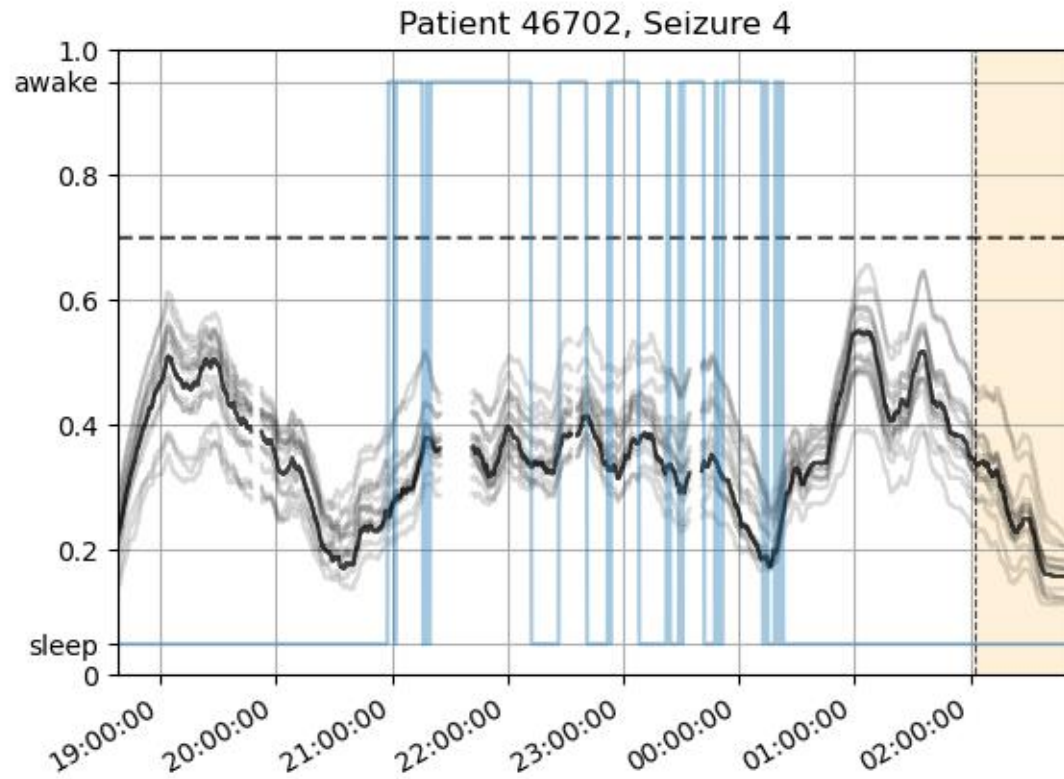
Patient 53402 – Over time

Patient 53402, Seizure 4
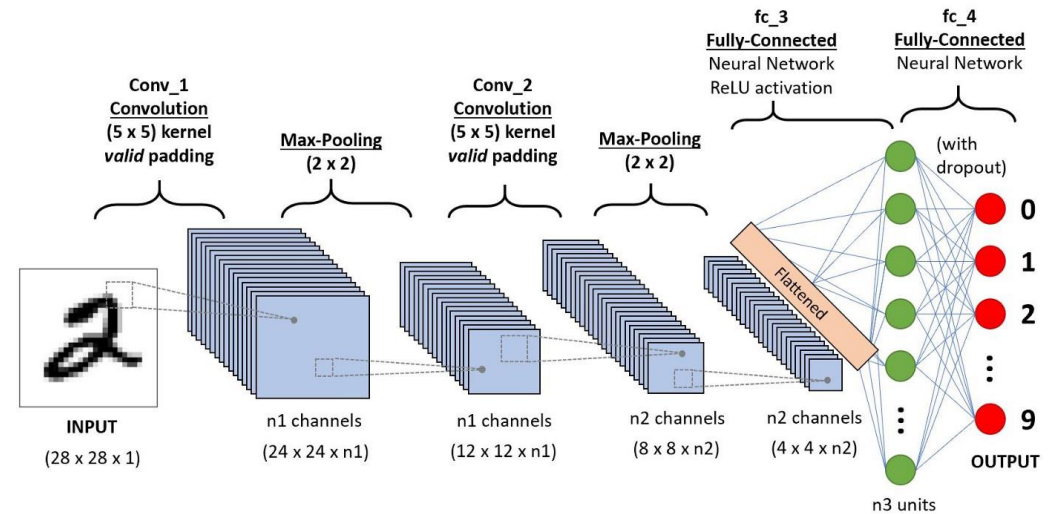
Patient 59102 – Over time
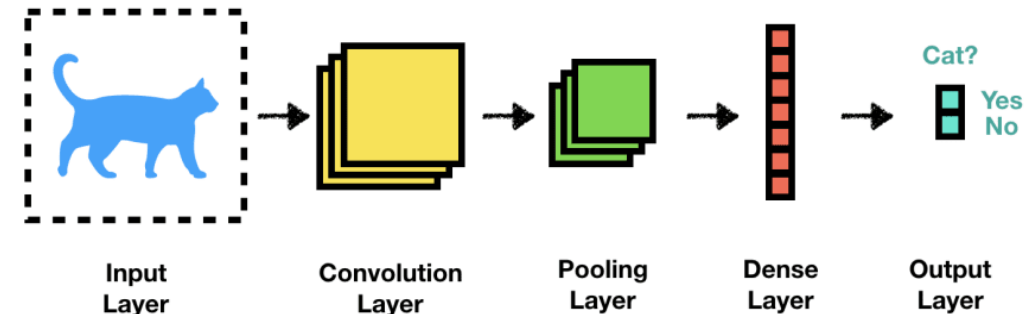
Patient 46702 – Over time

Convolutional Neural Networks

- 1) With the first 3 seizures:
  - Find the optimal pre-ictal period within a range of values: 30-60 minutes
  - Find the optimal number of features

- 2) Do the following 3 times:
  - Balance the data: choose random samples from inter-ictal
  - Train an CNN classifier

- 3) Each classifier then makes a prediction. The prediction is obtained by a voting system

Source: https://www.analyticsvidhya.com/blog/2022/03/basic-introduction-to-convolutional-neural-network-in-deep-learning/

Source: https://towardsdatascience.com/convolutional-neural-network-a-step-by-step-guide-a8b4c88d6943
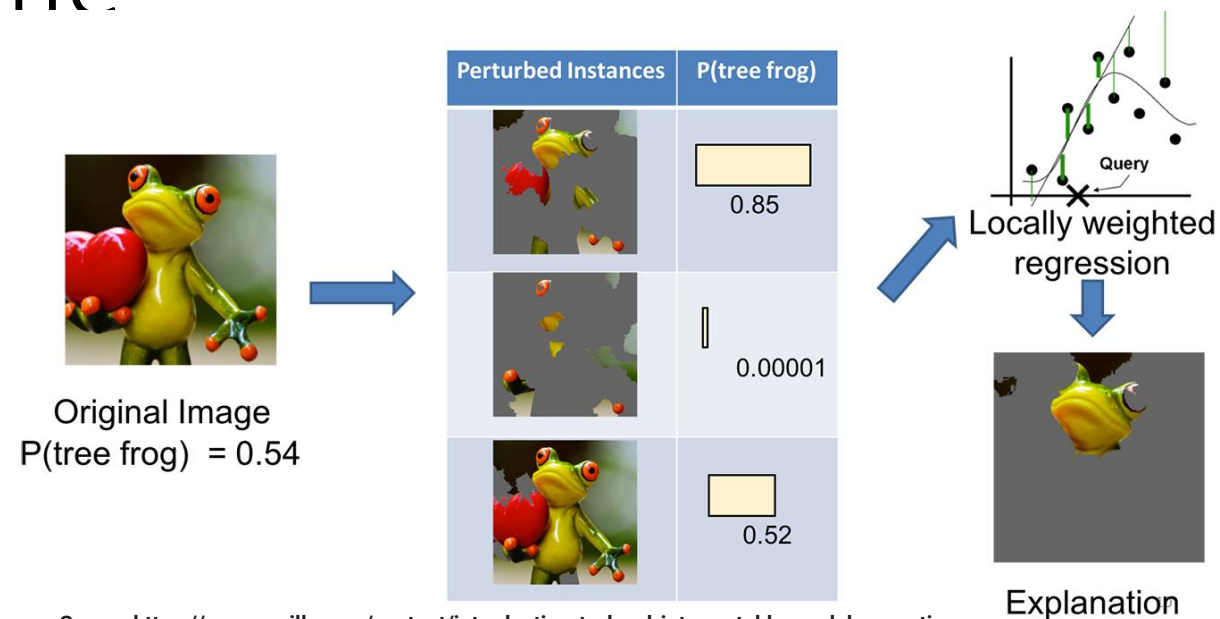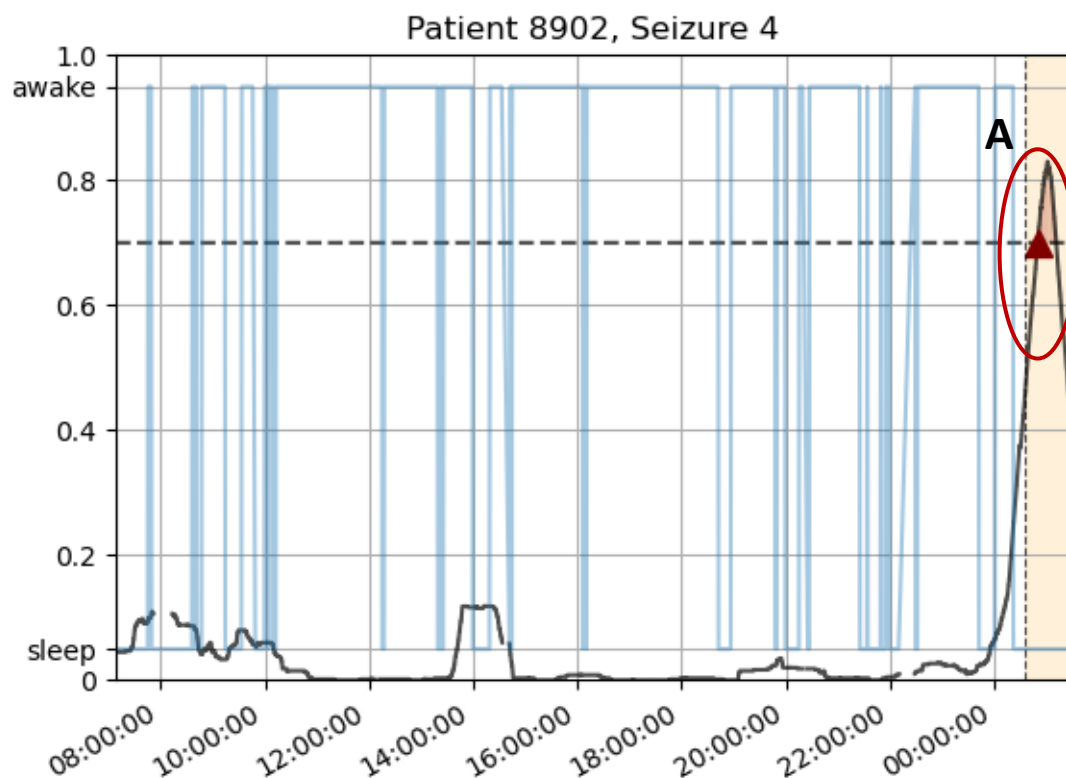
Patient 8902 – Overall performance

- SS:50%, FPR/h=0.00

- Above chance

- All seizures during the night. Sleep related?

- Surgery decision: not offered

- Focus localization=tpl
  Temporal polar left

	Onset	Type	Pattern	Vigilance state at onset
Training Seizures	Day 0 00:51:14	Unclassified	Rhythmic alpha waves	Awake
	Day 1 00:03:23	FOIA	Rhythmic beta waves	Awake
	Day 1 06:37:05	FOIA	Rhythmic alpha waves	Awake
Testing Seizures	Day 2 01:35:56	FOIA	Amplitude depression	Awake
	Day 2 06:10:26	FOIA	Rhythmic alpha waves	Awake

	Onset	Predicted	#False Alarms	Hours of analysed data
Testing Seizures	Seizure 4 Day 2 01:35:56	Yes	1	≈17
	Seizure 5 Day 2 06:10:26	Yes	1	≈4

Patient 8902 – Over time

- Ver ficheiro 8902_seizure_4_DL_00h30_01h00 para ver o que a rede usou para dizer que as janelas correspondiam a um período pre-ictal

Patient 8902, Seizure 4

Original Image
P(tree frog) = 0.54

Perturbed Instances	P(tree frog)
	0.85
	0.00001
	0.52

Locally weighted regression

Explanation

Source:https://www.oreilly.com/content/introduction-to-local-interpretable-model-agnostic-explanations-lime/

(a) Original Image (b) Explaining *Electric guitar* (c) Explaining *Acoustic guitar* (d) Explaining *Labrador*
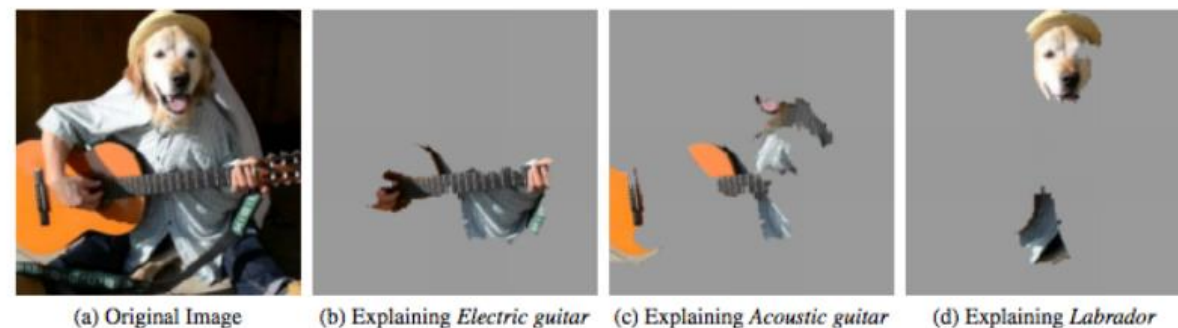
Figure 4: Explaining an image classification prediction made by Google's Inception network, highlighting positive pixels. The top 3 classes predicted are "Electric Guitar" ($p = 0.32$), **"Acoustic guitar"** ($p = 0.24$) **and "Labrador"** ($p = 0.21$)

Source: https://homes.cs.washington.edu/~marcotcr/blog/lime/