# Machine learning to predict clinical outcomes from RNS background ECoG

**Byron Sun**
**Xiaojing Wu**
**Anli Liu**

# Limitations of current RNS practice

1. Long intervals between seizures and between programming visits (2-3 months)

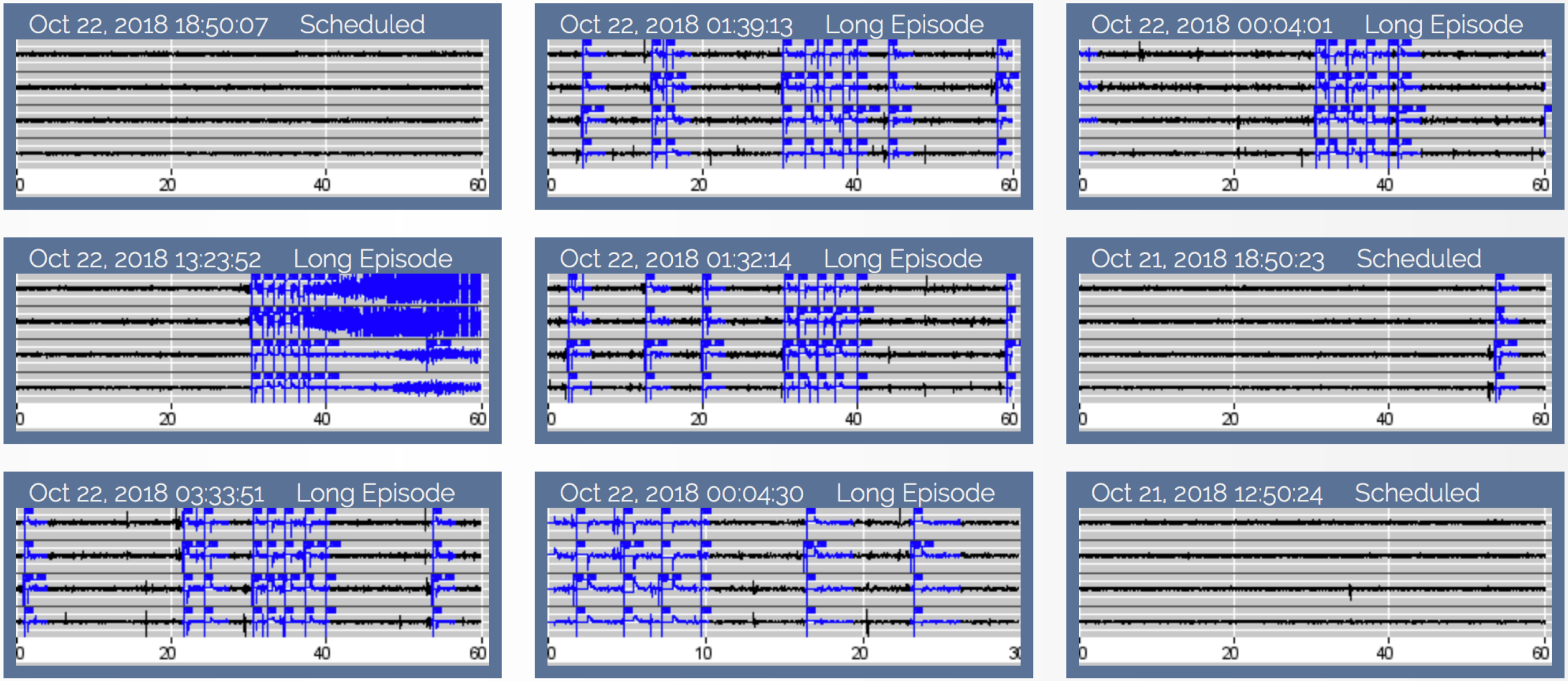
2. Very limited understanding of how stimulation parameters should be adjusted

Can we build a reliable classifier for an individual patient which predicts clinical outcome, based on retrospective review of that patient's EEG?

## RNS parameter adjustment for a given patient
## - How to quickly and accurately determine effectiveness?

- Goal: build reliable classifier for *individual* patient → generate multiple predictions per day with newly recorded "scheduled" ECoGs → help physicians better understand the patients' clinical conditions and make better parameter adjustment based on the prediction

# Patient 231



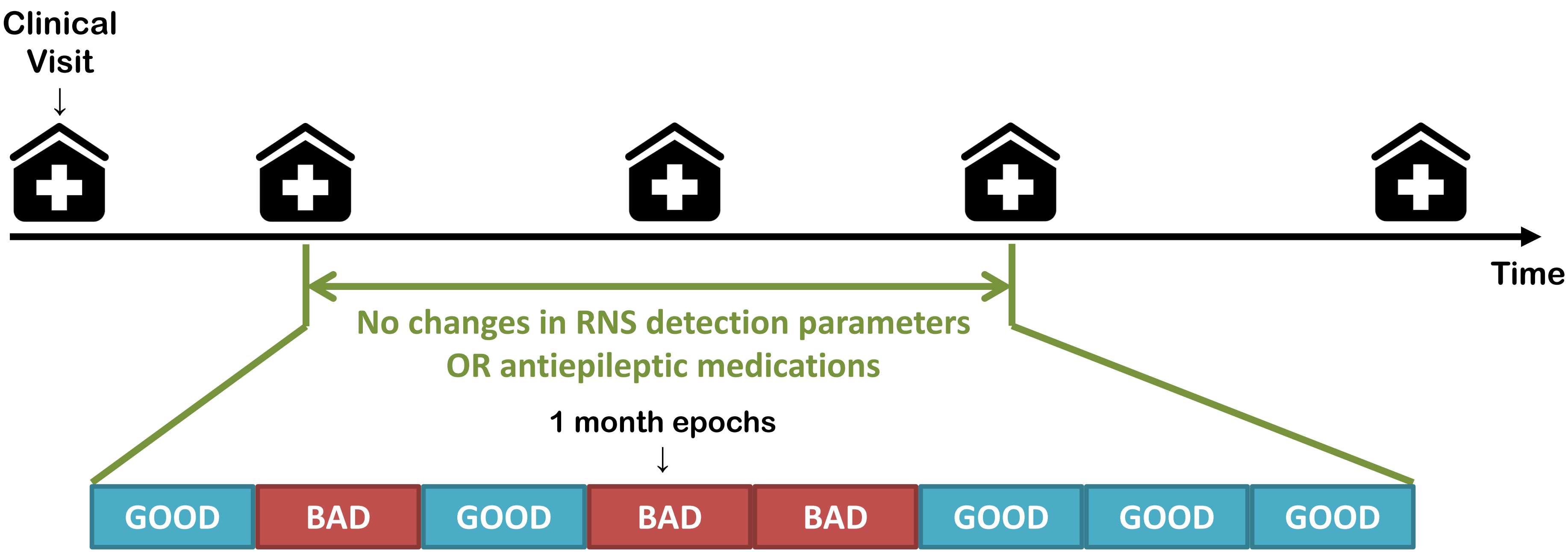

Representative patient, demonstrating difference in EEG characteristics between (1) Scheduled ECoGs and (2) Long Episode Detection (Triggering Stimulation). Some LE become electroclinical seizures.

## RNS parameter adjustment for a given patient
## - How to quickly and accurately determine effectiveness?

- What data should we use to predict patients' clinical conditions?

- Make use of "scheduled" ECoG segments
  - Recorded multiple times per day, data abundance
  - Indicative of long-term neuromodulatory effects of chronic electrical stimulation

# EEG Selection Criteria




**Find long periods of time:**

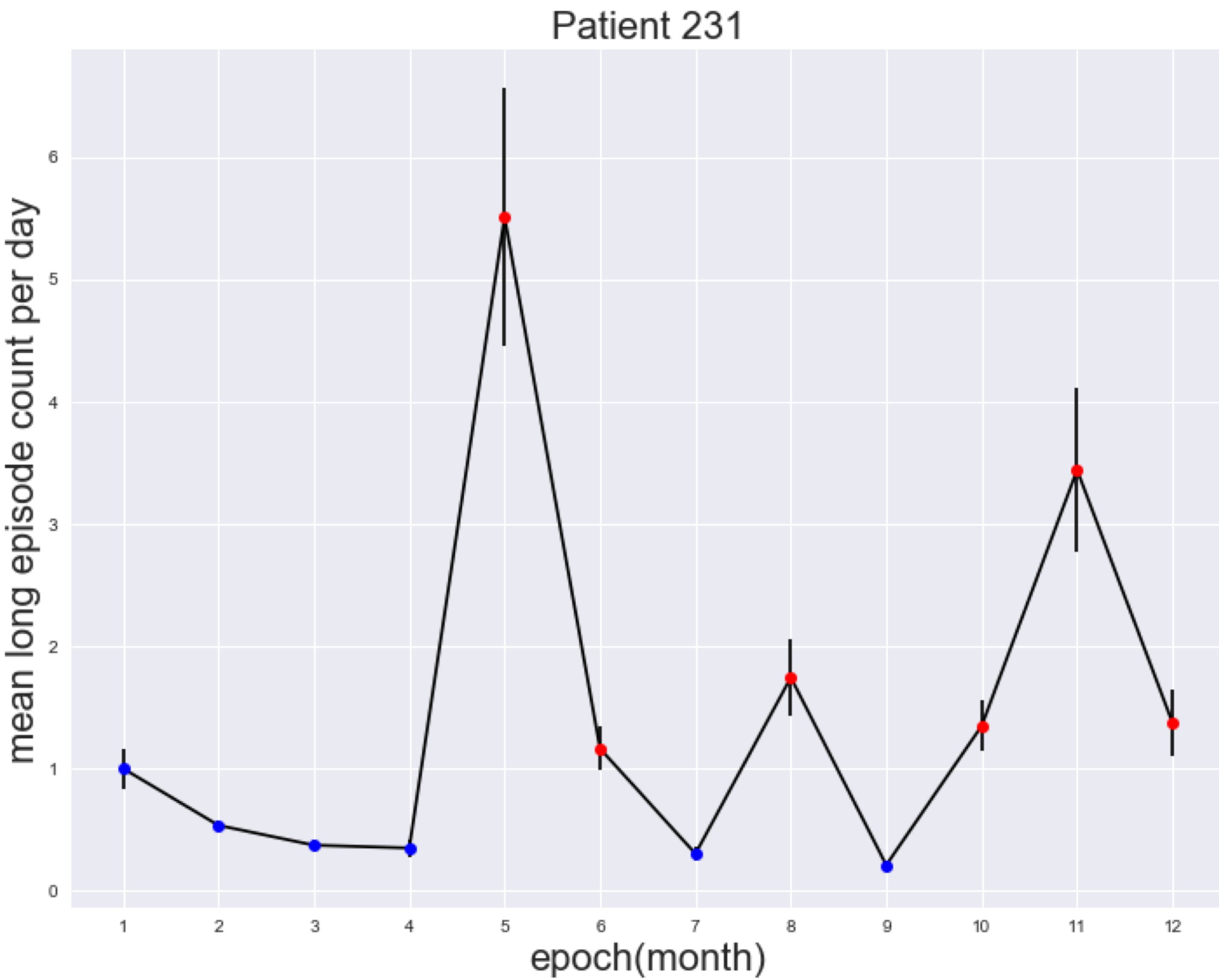
- **Long episodes detected by same parameters**
- **Any EEG pattern change cannot be caused by AED changes**

**GOOD: relatively <u>low</u> average daily long-episode count**

**BAD: relatively <u>high</u> average daily long-episode count**

**Methods**

Average long-episode number per day, over one-month epochs



**Methods**

## Feature calculation

EEG segment

↓

| Channel 1 Delta | Channel 2 Delta | Channel 3 Delta | Channel 4 Delta | Channel 1 Theta | Channel 2 Theta | Channel 3 Theta | Channel 4 Theta | ••• |
|---|---|---|---|---|---|---|---|---|

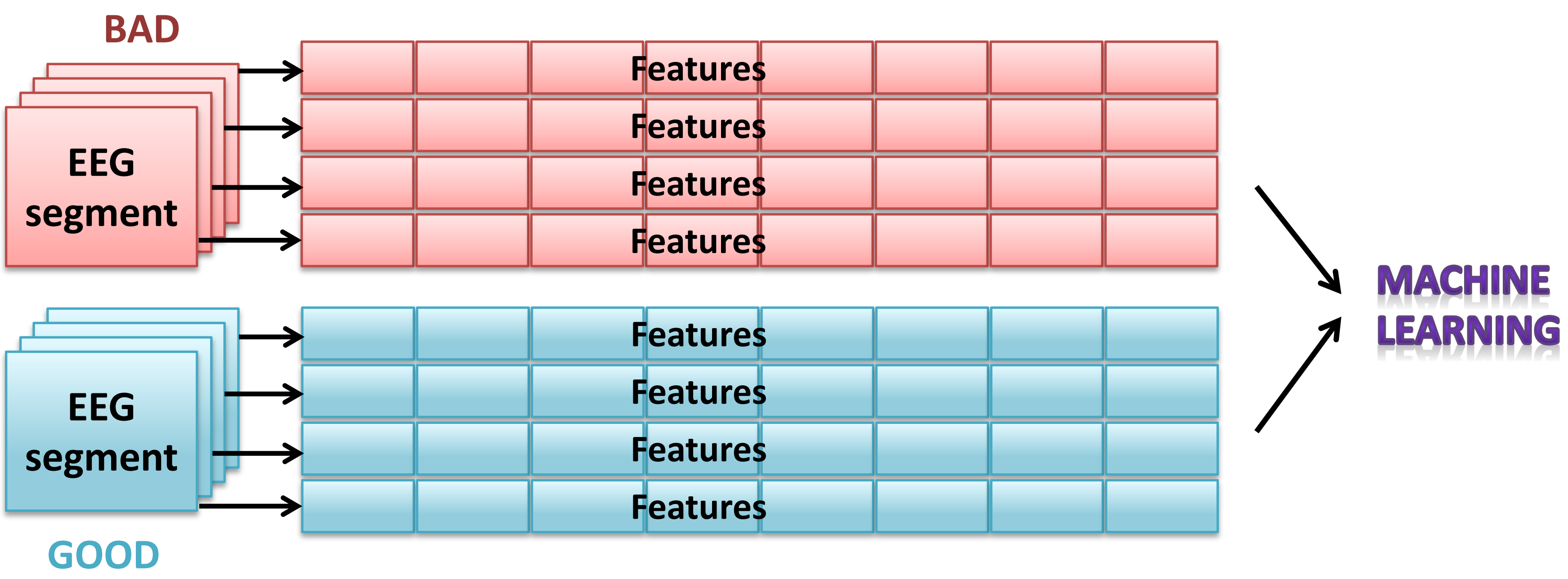
(band power)

Classic frequency bands:

Delta = 0.5 - 4 Hz
Theta = 4 - 8 Hz
Alpha = 8 - 12 Hz
Beta = 12 - 25 Hz
Low gamma = 25 - 50 Hz
High gamma = 50 - 124.9 Hz
Entire band = 0.01 - 124.9 Hz

→ 4 channels X 7 frequency bands
= 28 features per EEG segment

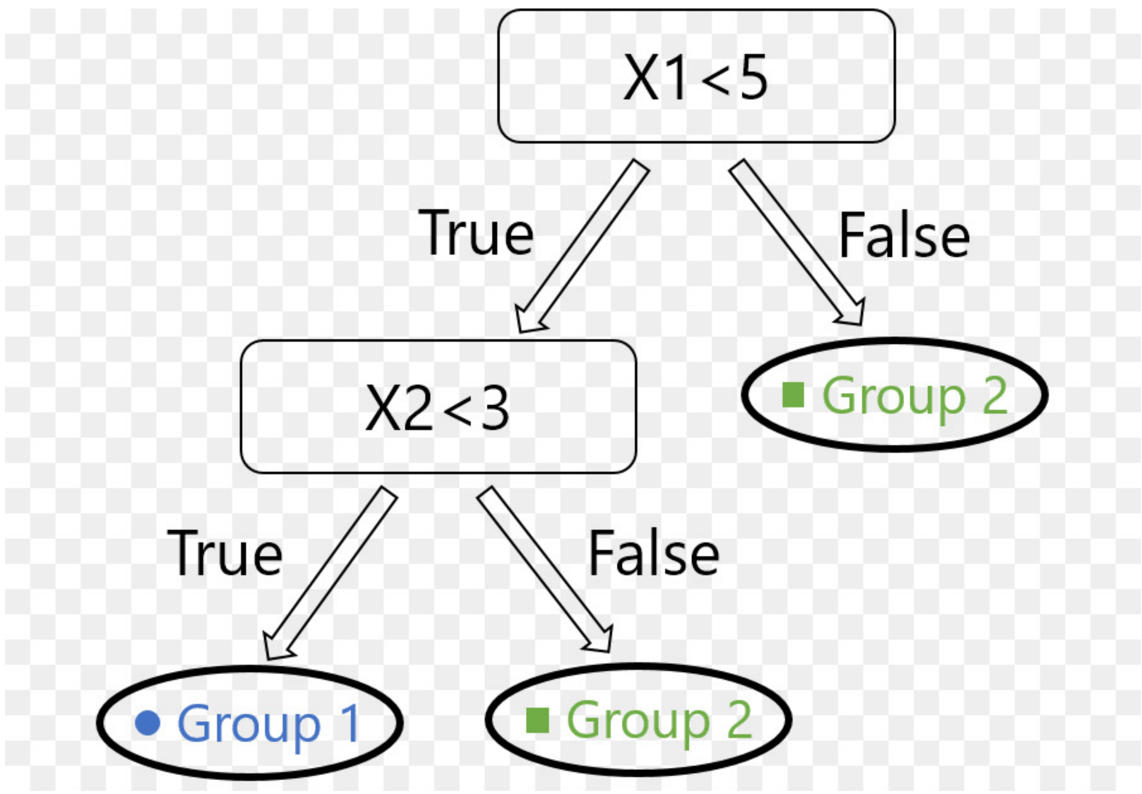
# Methods


BAD
EEG segment
Features
Features
Features
Features
EEG segment
Features
Features
Features
Features
GOOD
MACHINE LEARNING

**Methods**

# Machine learning methods

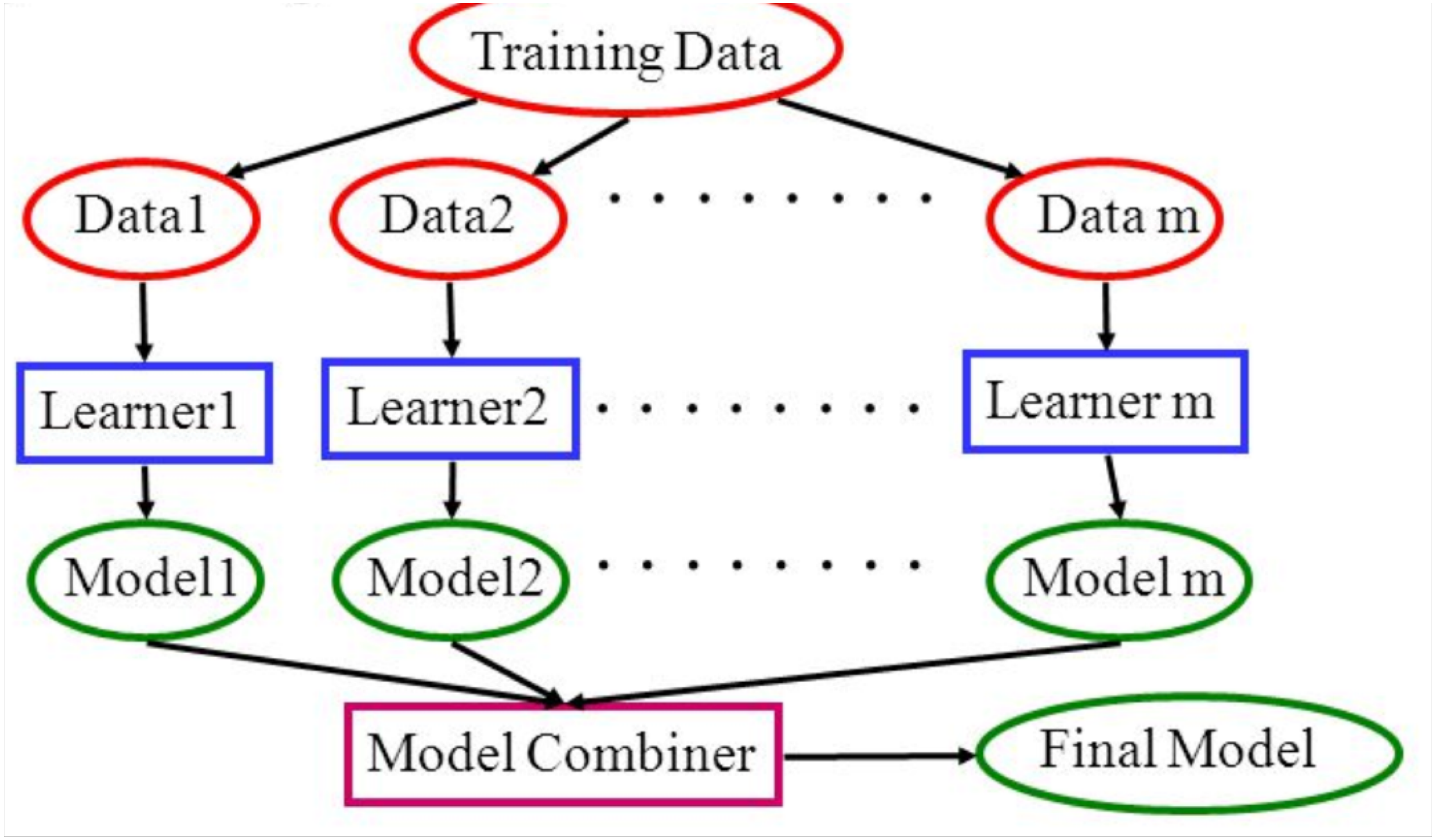
## Decision tree



**Methods**

# Machine learning methods

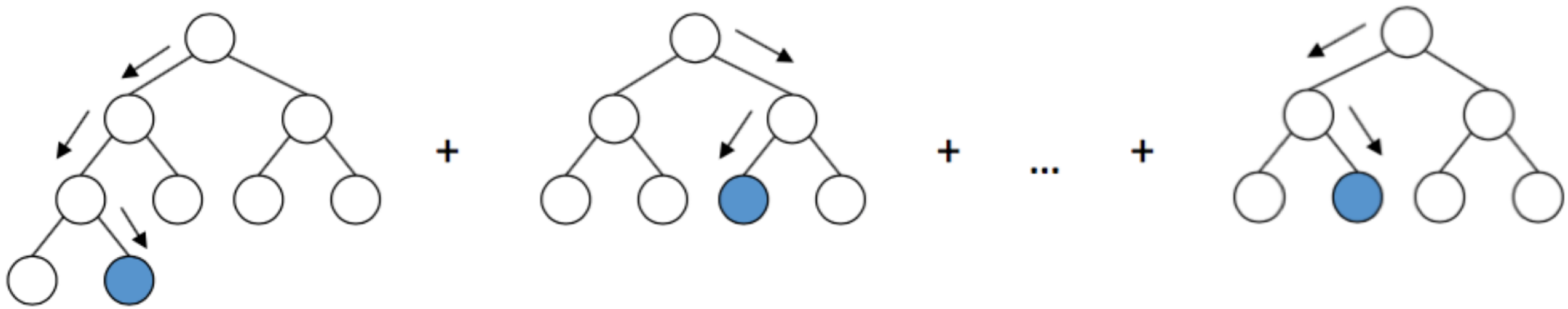
Ensemble Model: Random Forest, Gradient Boosting

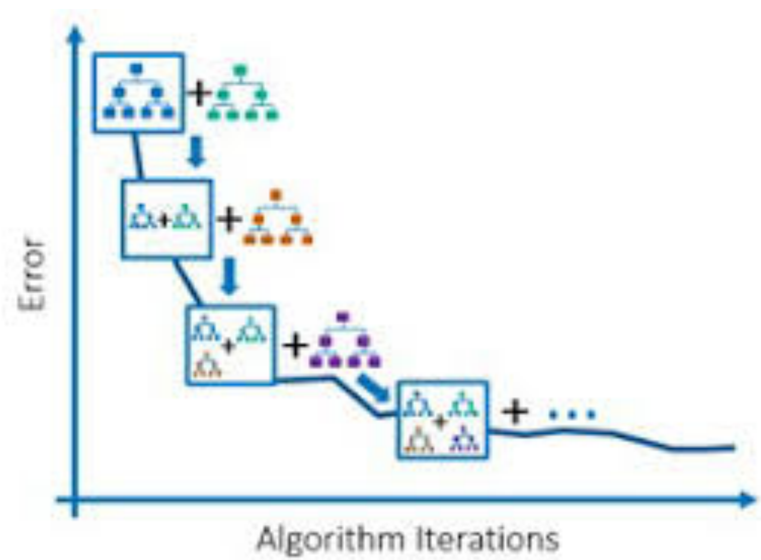


**Methods**

# Machine learning methods

## Gradient Boosting



**Methods**

# Patient Selection Criteria

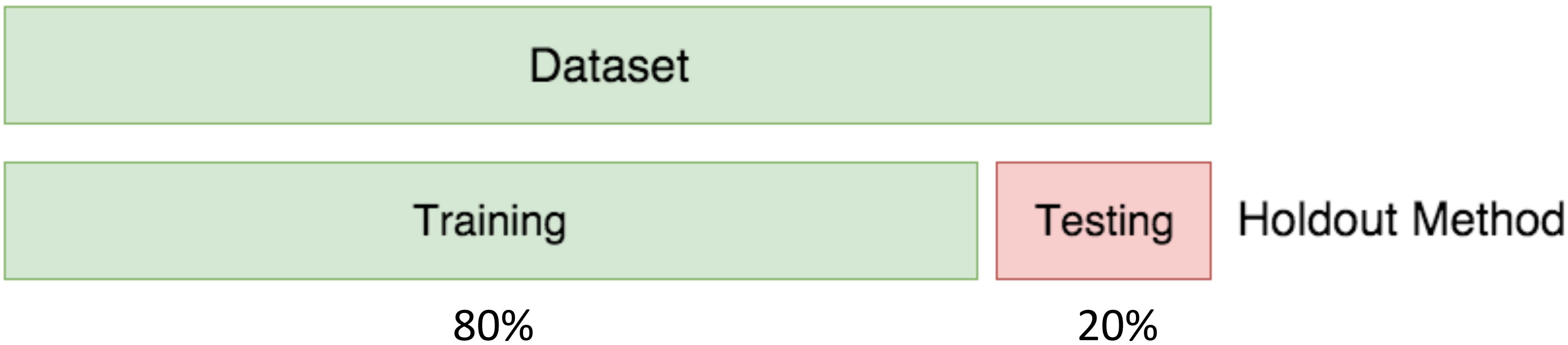
1. Good RNS upload compliance
2. Have scheduled ECoGs multiple times per day
3. Have a relatively clean EEG background
4. Good clinical correlation between long episodes and clinical seizures
5. Infrequent medication changes
6. Past the 6 month window of implant effect, and past the detection adjustment stage.

## Analyzed patients

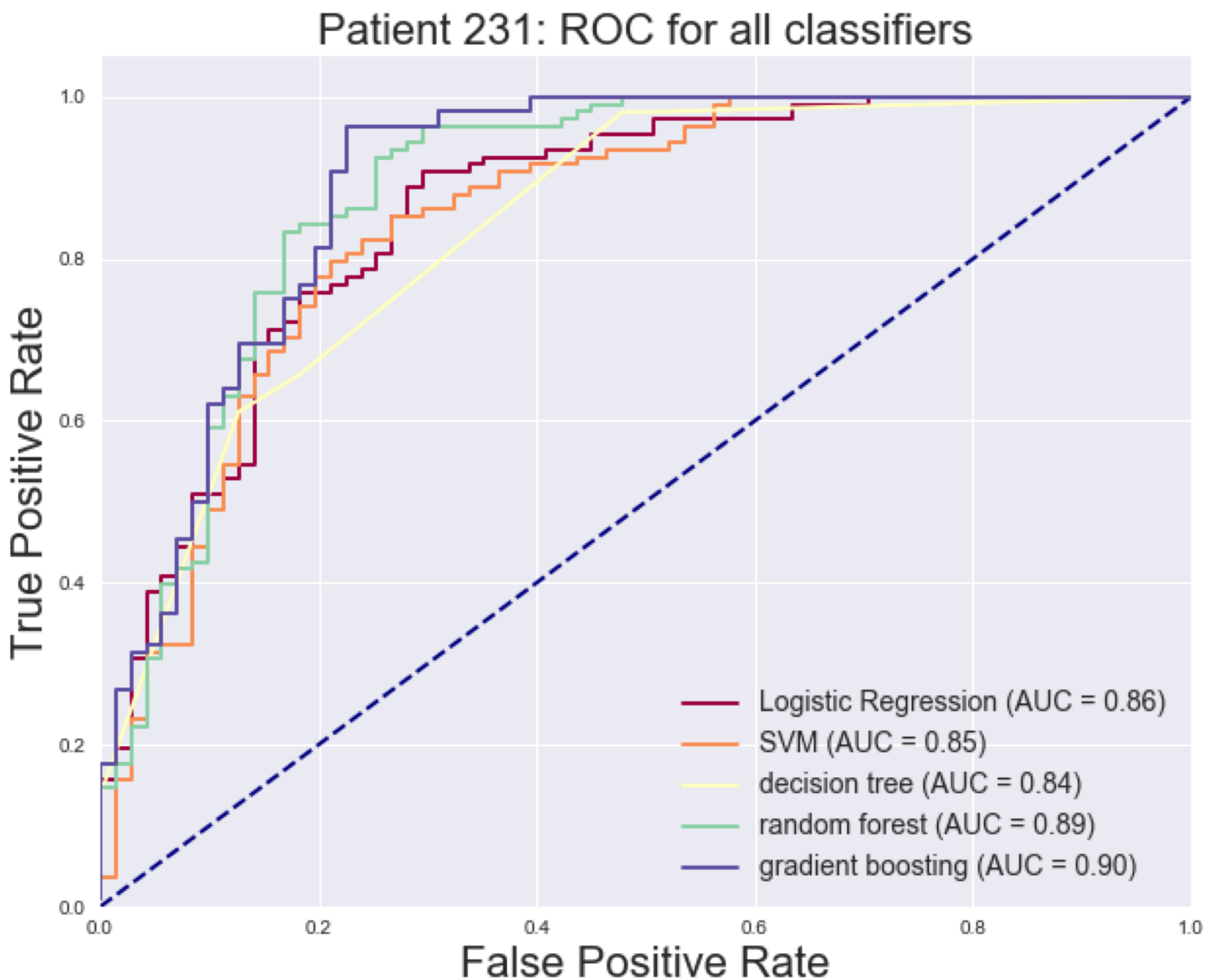
| Patient ID | Patient Initials | Gender | Age | Clean periods (Days) | Leads location |
|---|---|---|---|---|---|
| NY231 | DB | F | 33 | 379 | L hippocampal depth<br>R hippocampal depth |
| NY222 | AJ | M | 25 | 255, 215 | L insular depth<br>L superior temporal strip |
| NY229 | TR | M | 26 | 318 | Insular depth<br>L mid-central strip |

**Methods**

# Training & Testing



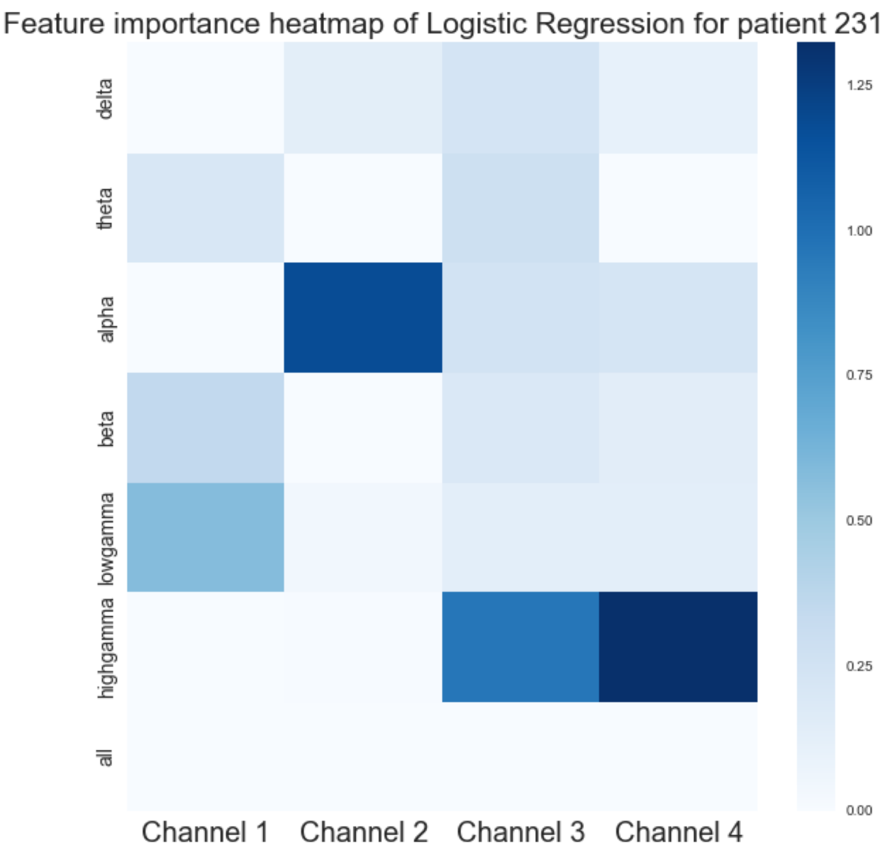
# Results



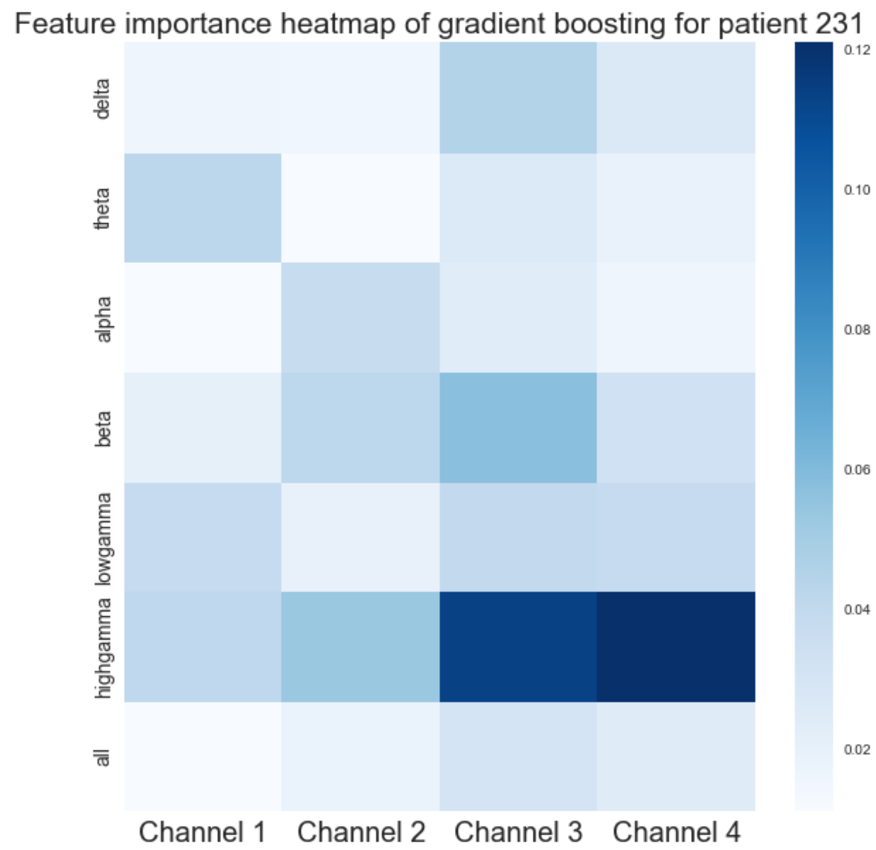

Best classifier for 231 is gradient boosting with AUC = 0.90, accuracy = 87.15%

# Results

## Logistic Regression



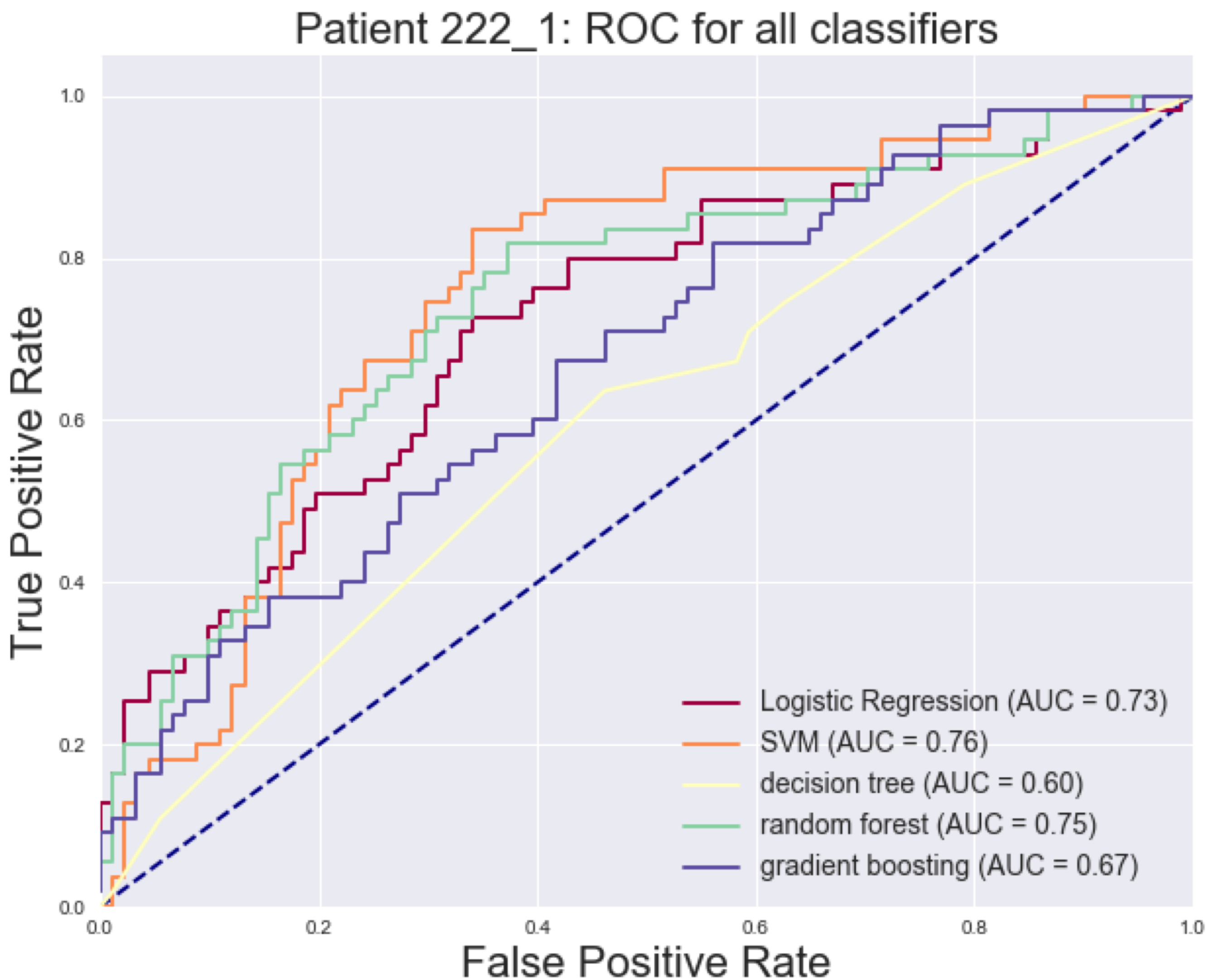

## Gradient boosting




Important features:
- high gamma Channel 4
- alpha Channel 2
- high gamma Channel 3

Important features:
- high gamma Channel 4
- high gamma Channel 3
- beta Channel 3

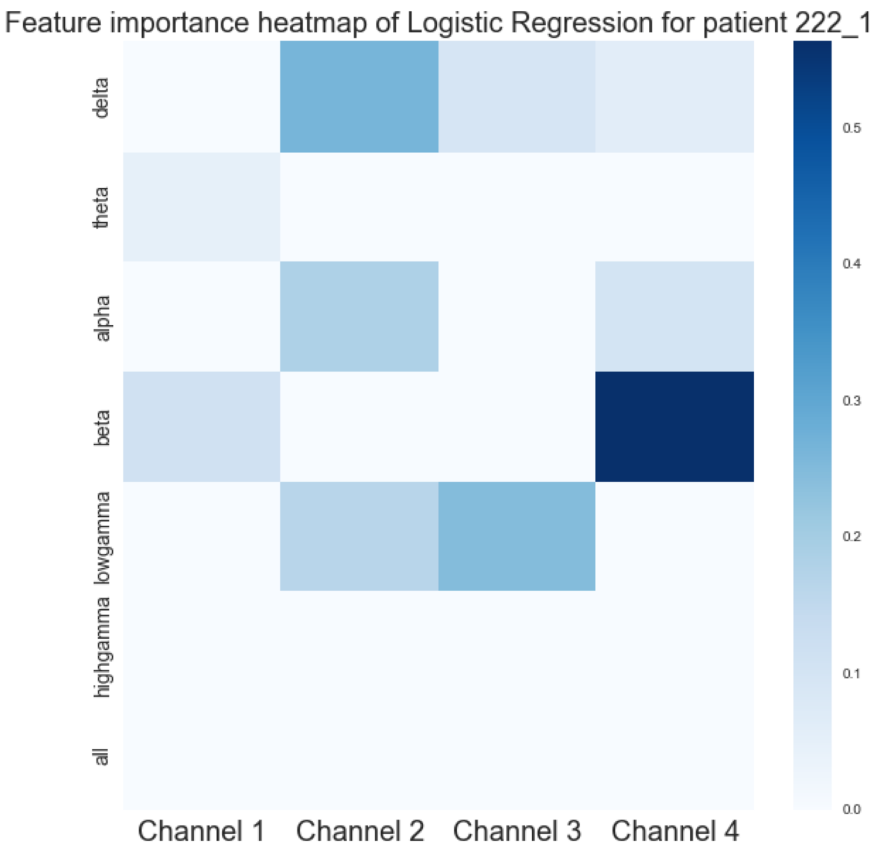
## Results



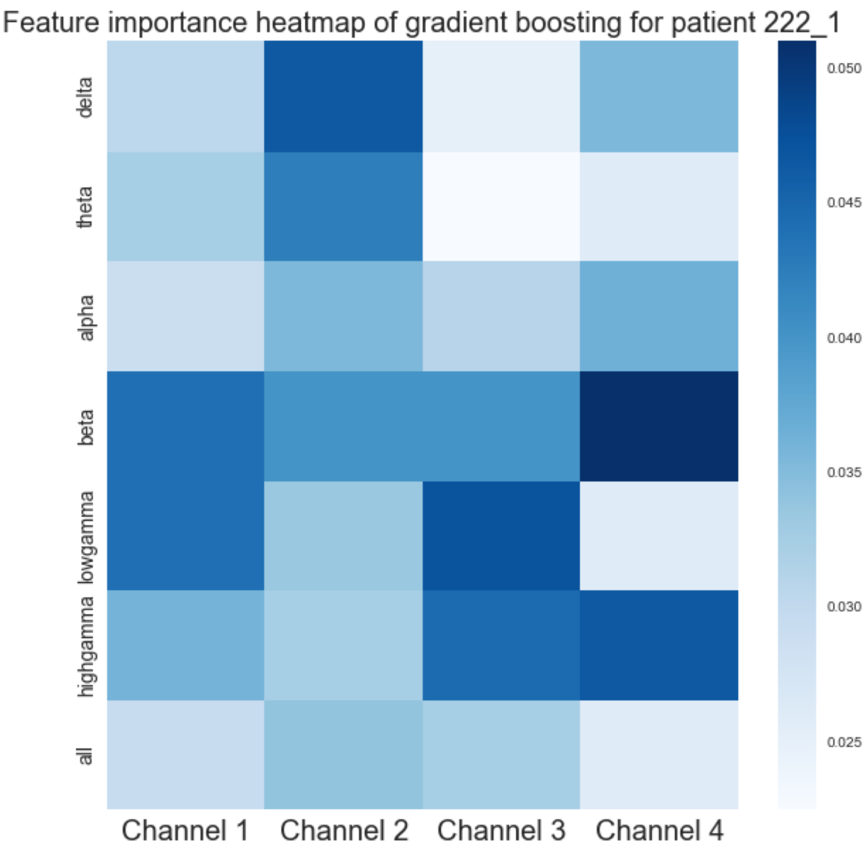

Best classifier for 222_1 is SVM with AUC = 0.76, accuracy = 71.92%

# Results

## Logistic Regression




## Gradient boosting



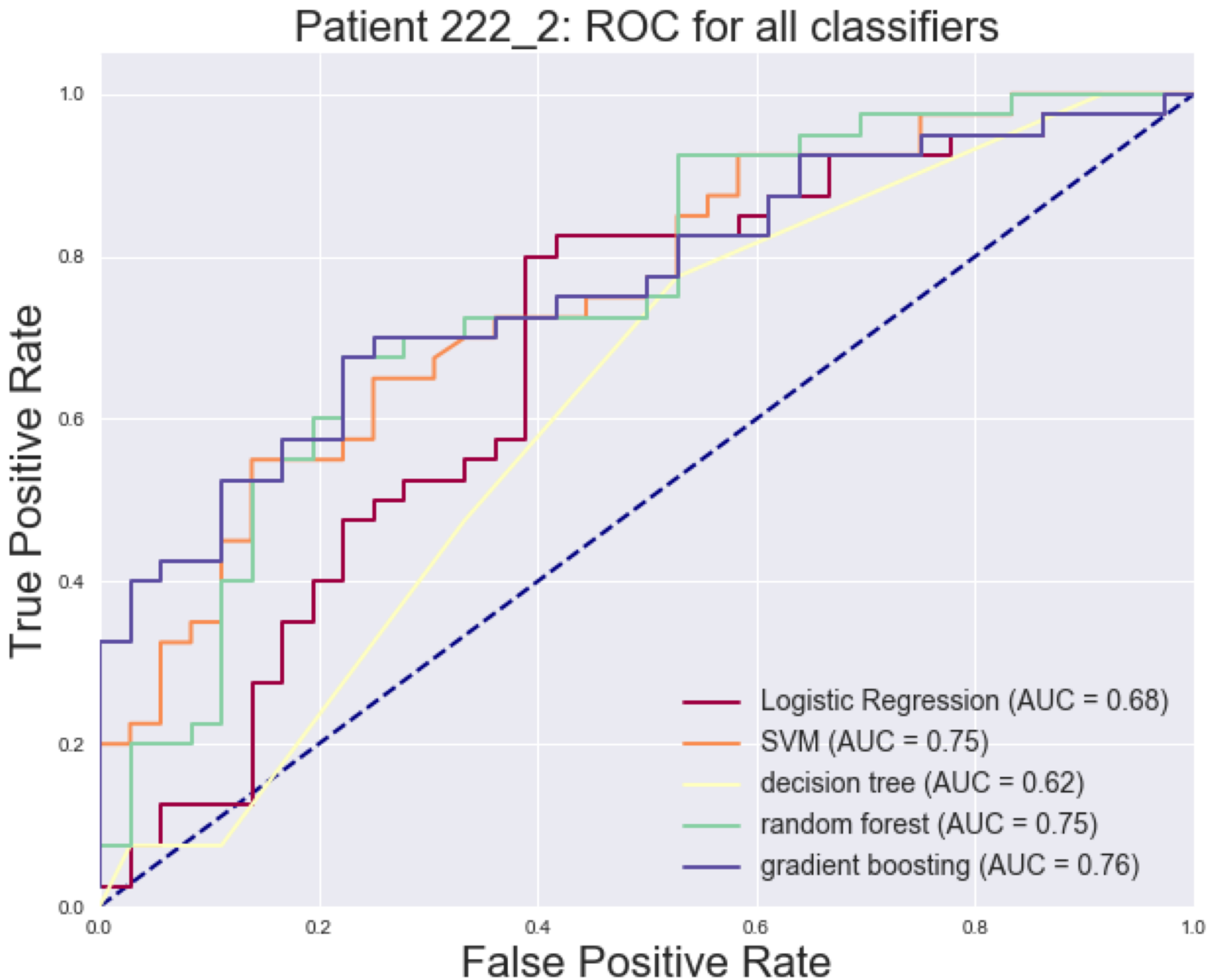

Important features:

- beta Channel 4
- low gamma Channel 3
- delta Channel 2

Important features:

- beta Channel 4
- low gamma Channel 3
- delta Channel 2

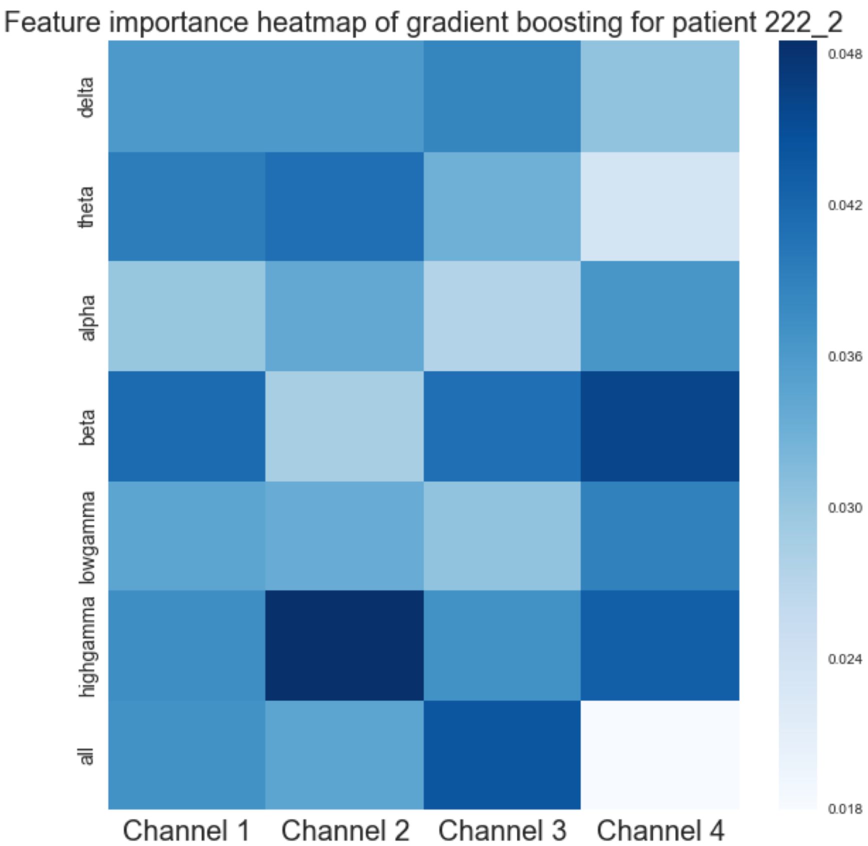
# Results




Best classifier for 222_2 is gradient boosting with AUC = 0.76 , accuracy = 72.37%

# Results

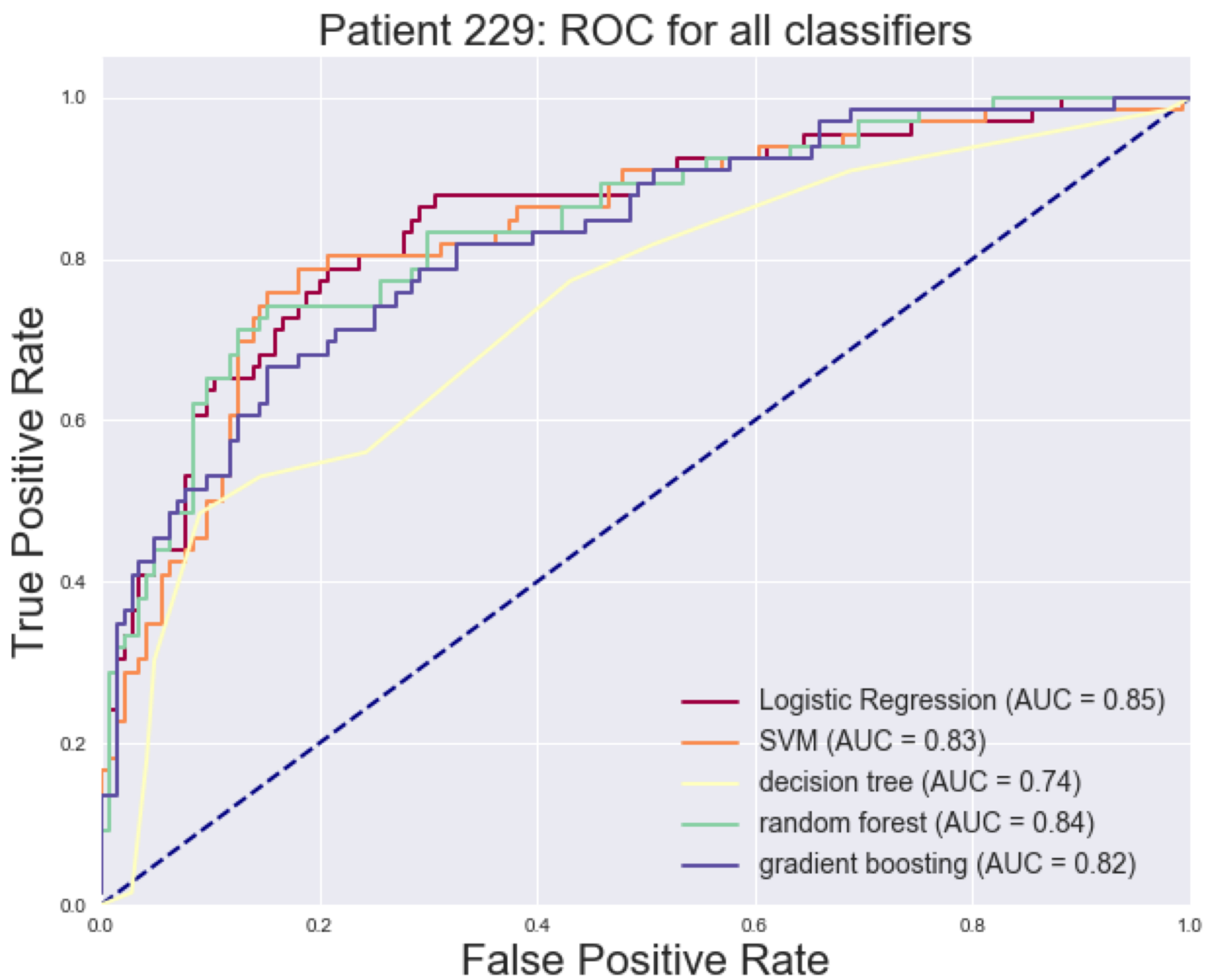
## Gradient boosting




Important features:

- high gamma Channel 2
- beta Channel 4
- all Channel 3

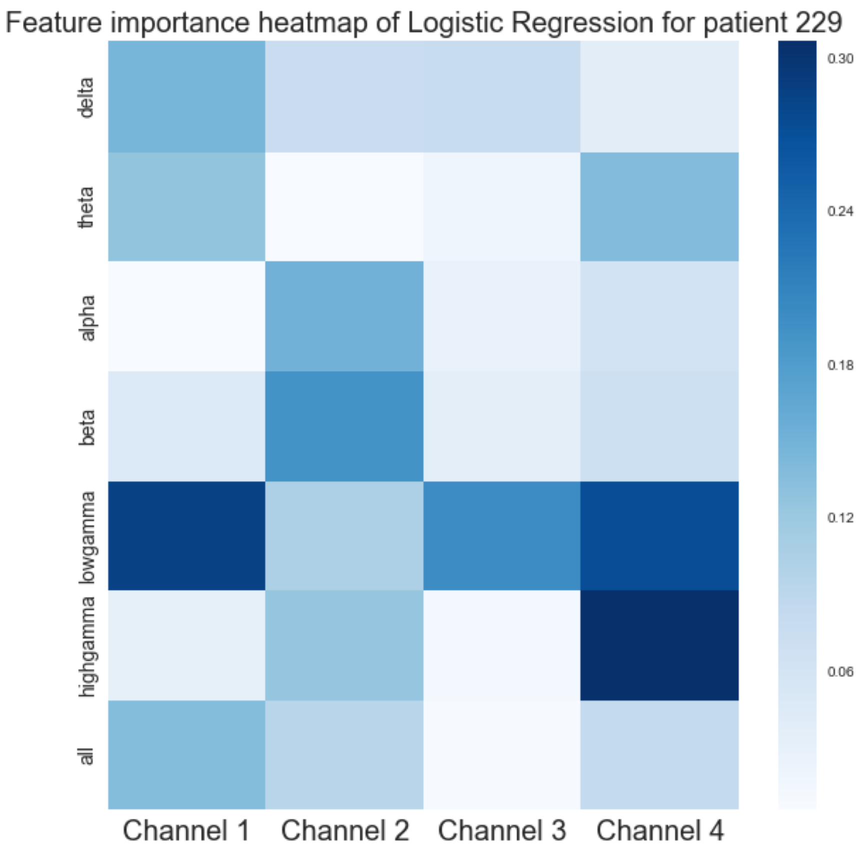
# Results




Best classifier for 229 is Logistic Regression with AUC = 0.85 , accuracy = 81.43%

# Results

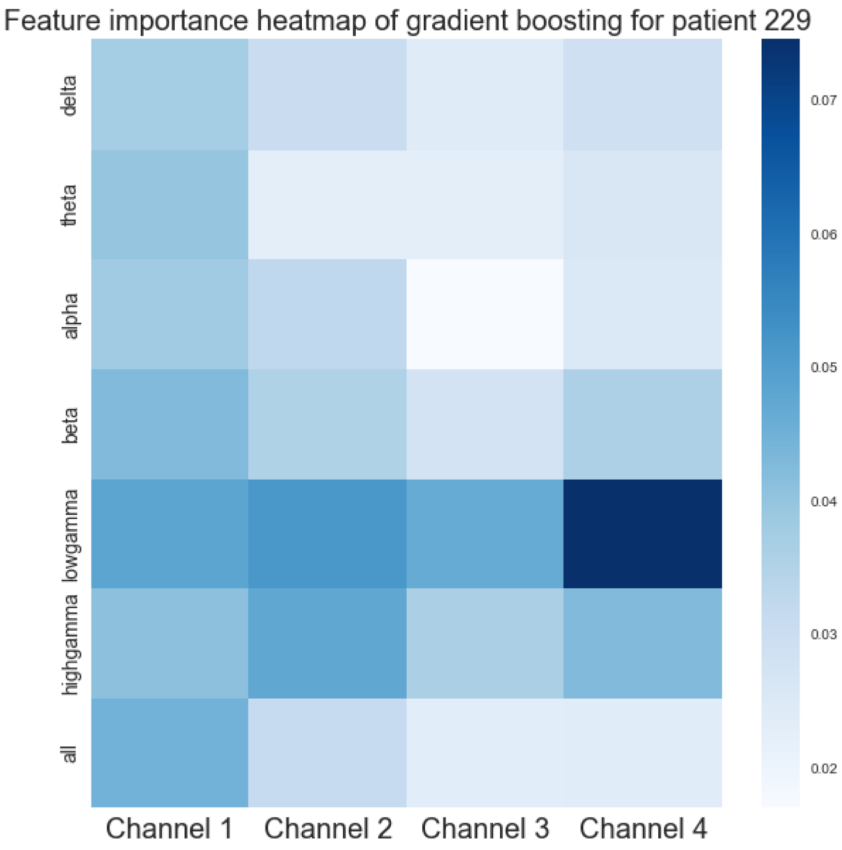
## Logistic Regression




Important features:

- high gamma Channel 4
- low gamma Channel 1
- low gamma Channel 4

## Gradient boosting

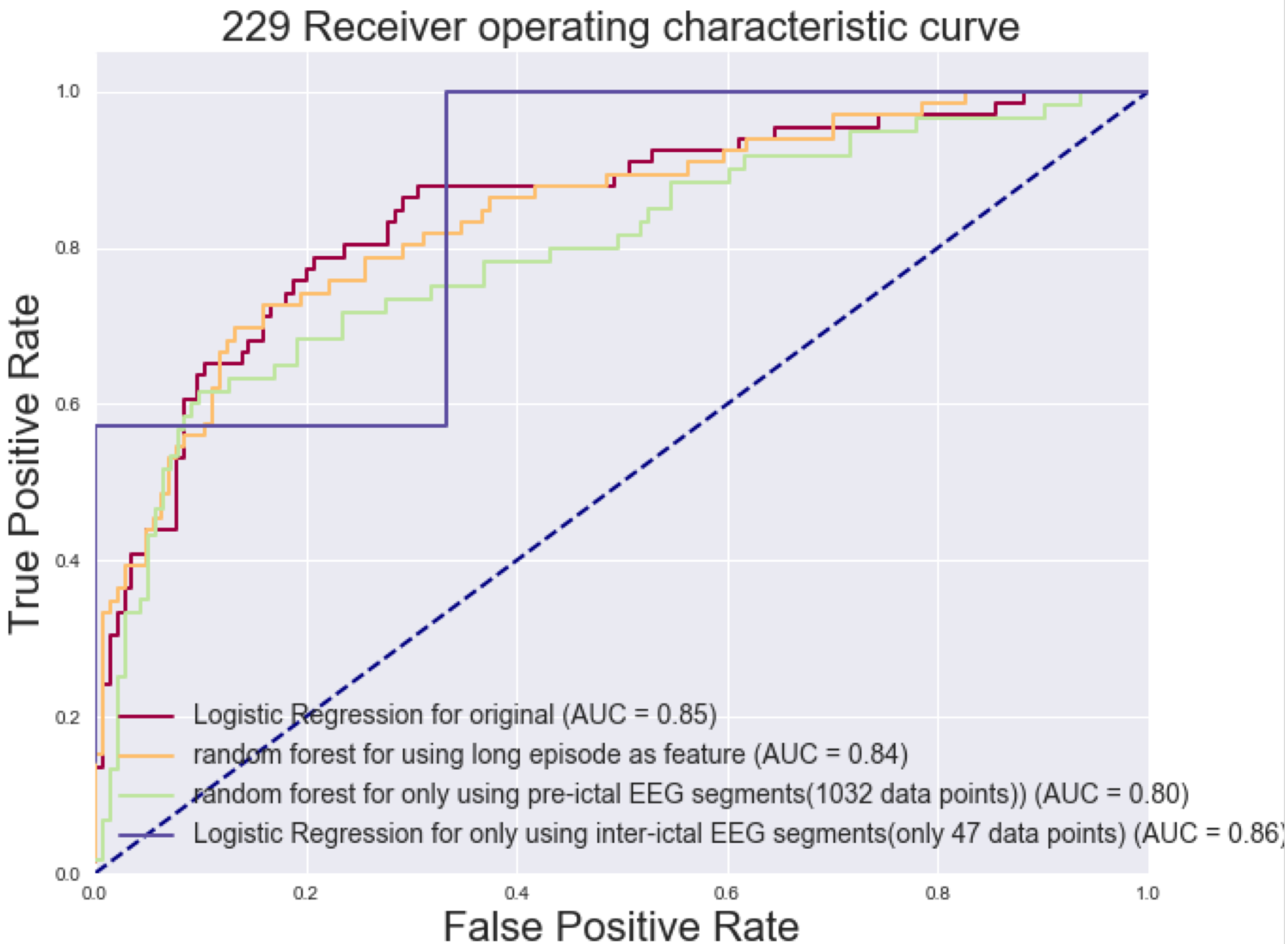



Important features:

- low gamma Channel 4
- low gamma Channel 2
- low gamma Channel 1

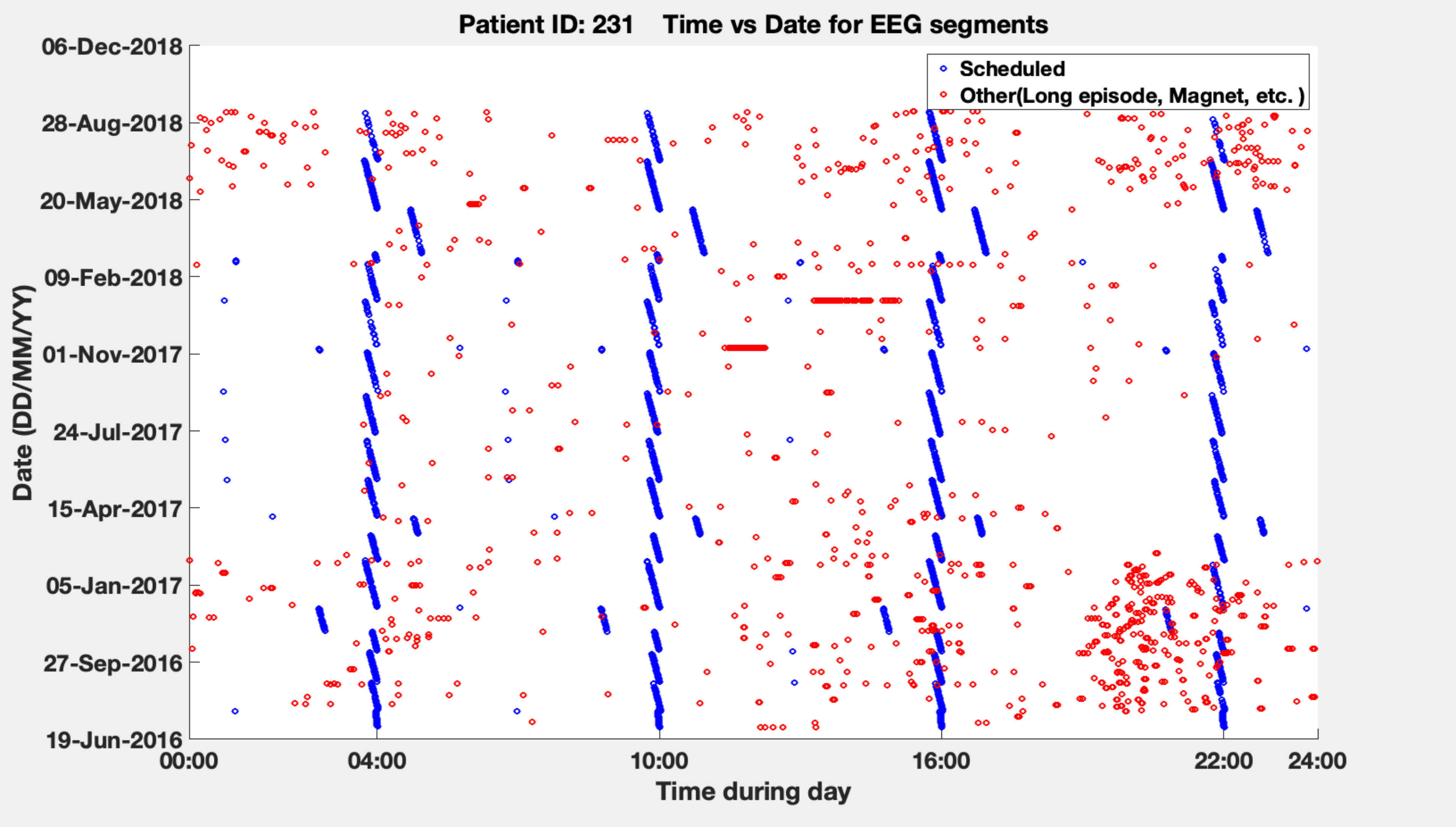
## Results – discussion 1

Discover the role of pre-ictal EEG segments in predicting clinical condition



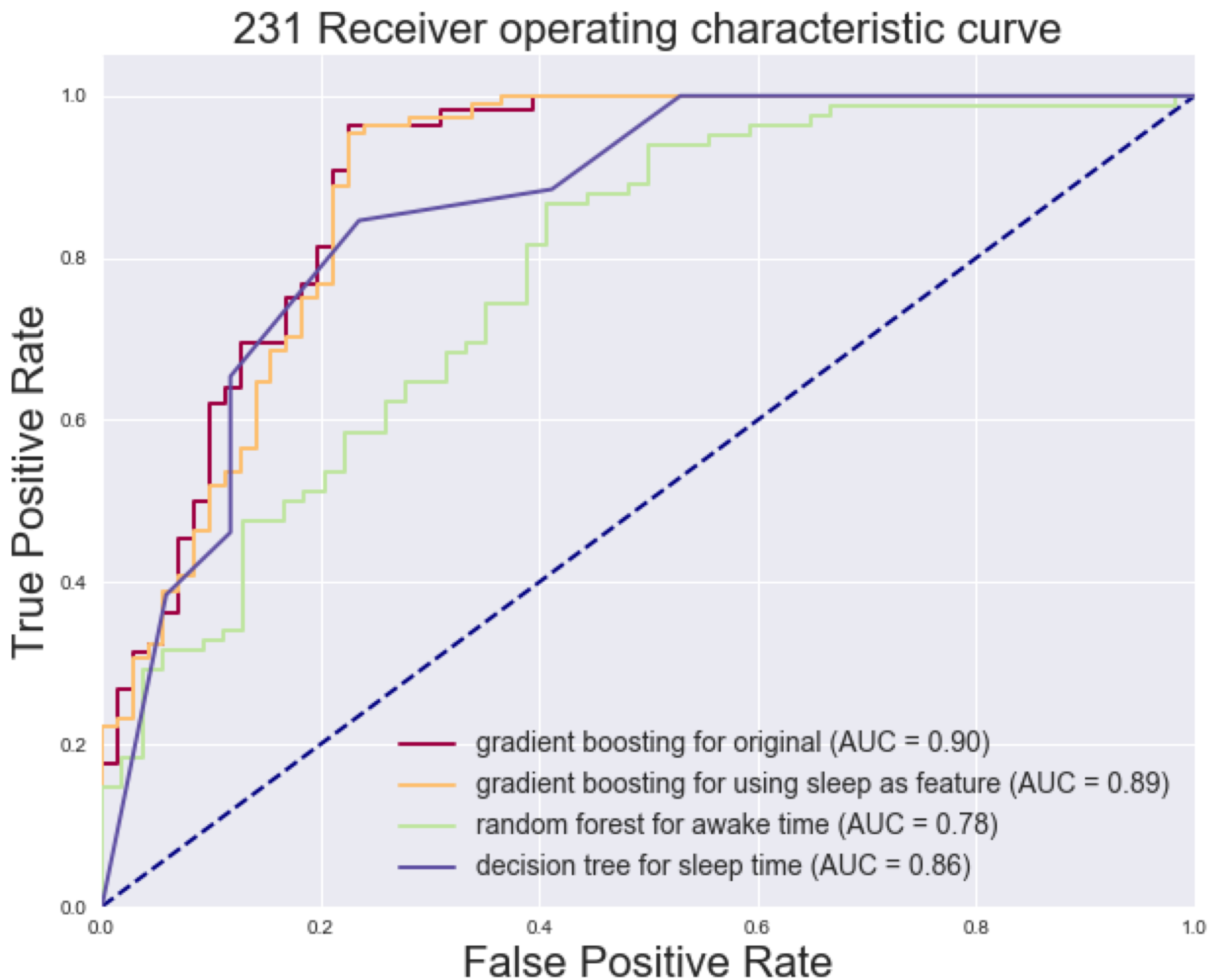
# Results – discussion 2

Discover the role of sleep in predicting clinical conditions



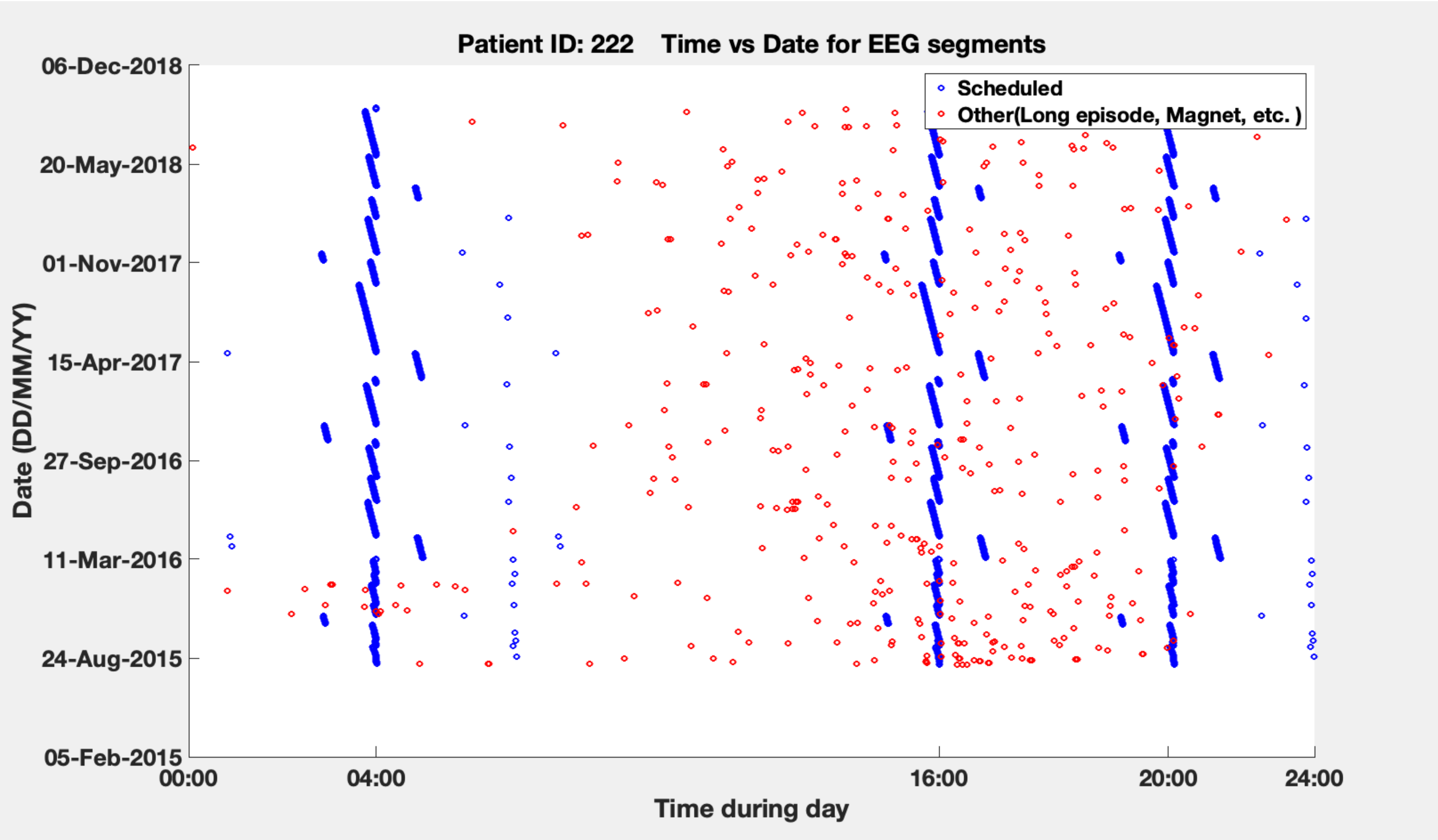
# Results – discussion 2

Discover the role of sleep in predicting clinical conditions



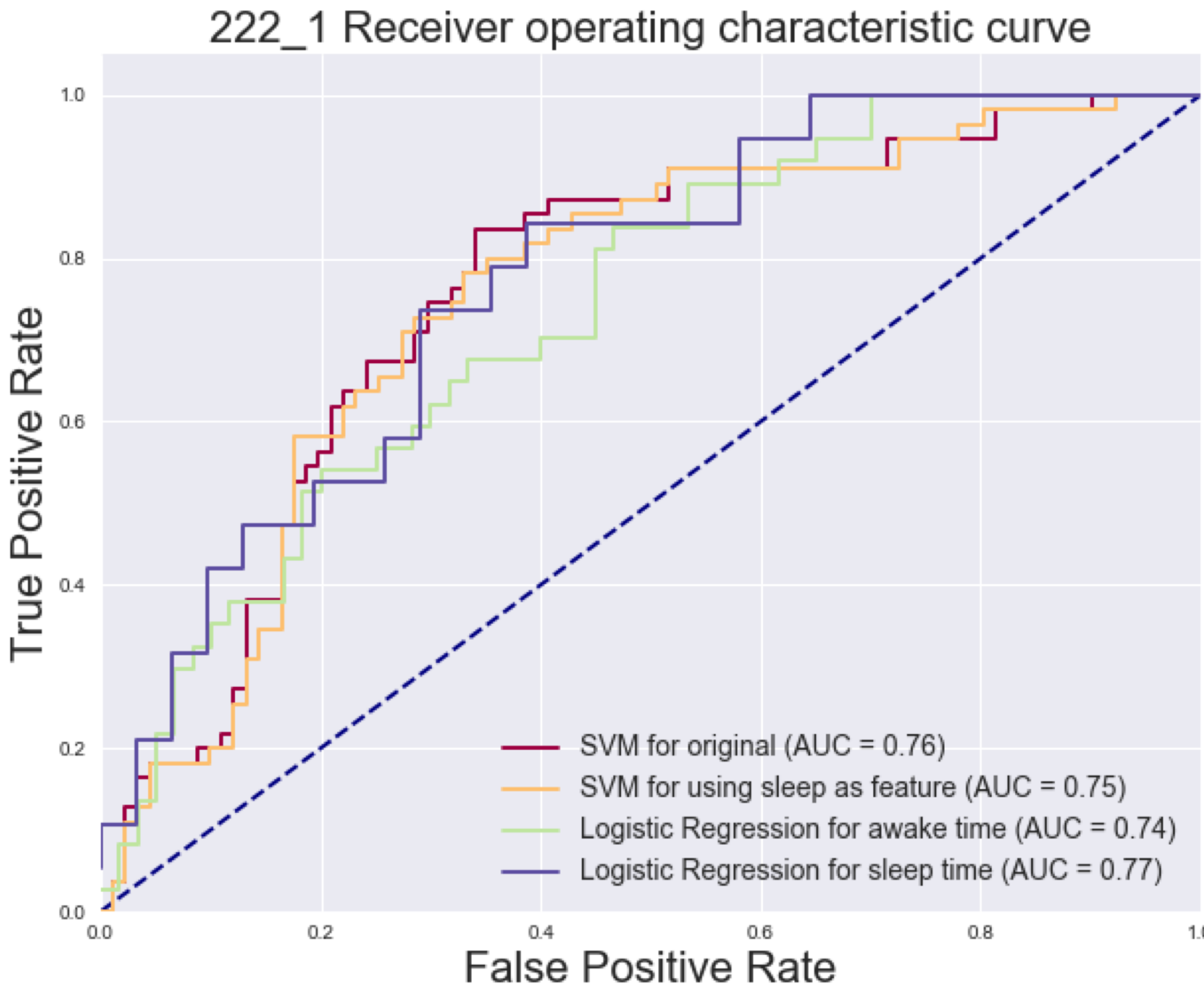
# Results – discussion 2

Discover the role of sleep in predicting clinical conditions



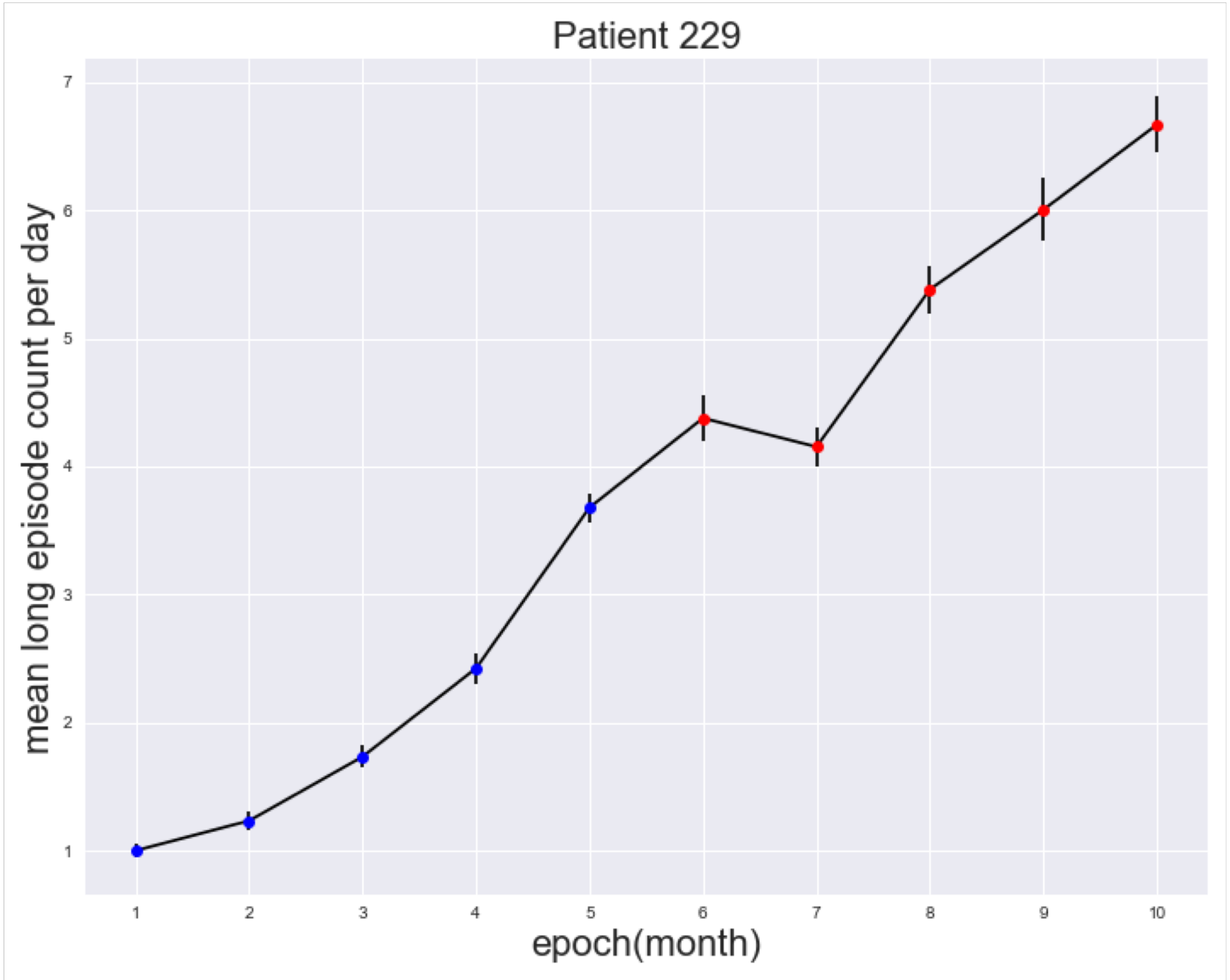
# Results – discussion 2

Discover the role of sleep in predicting clinical conditions



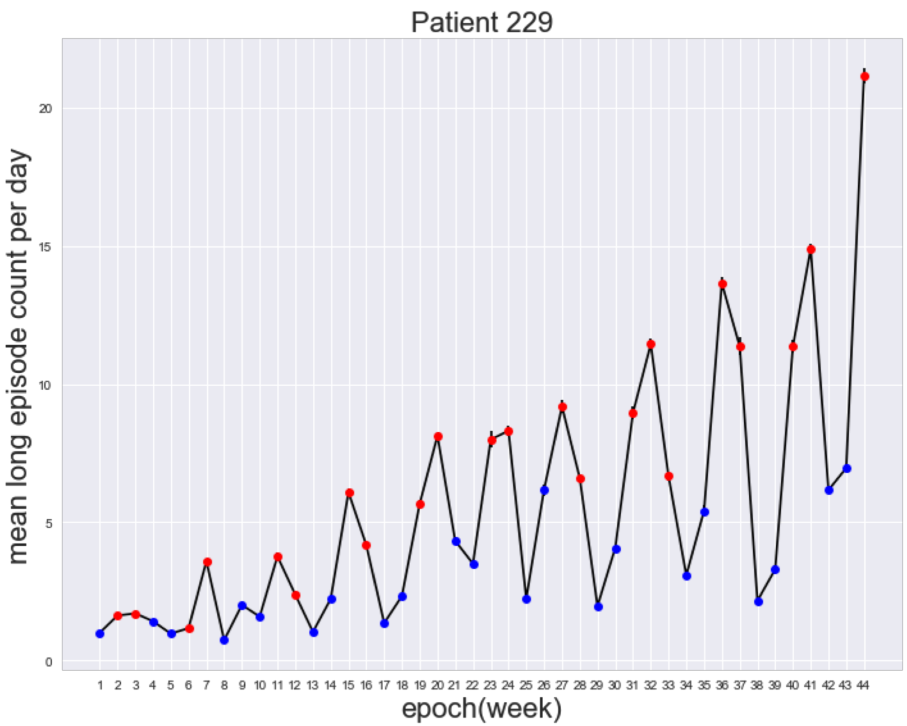
# Results – discussion 3

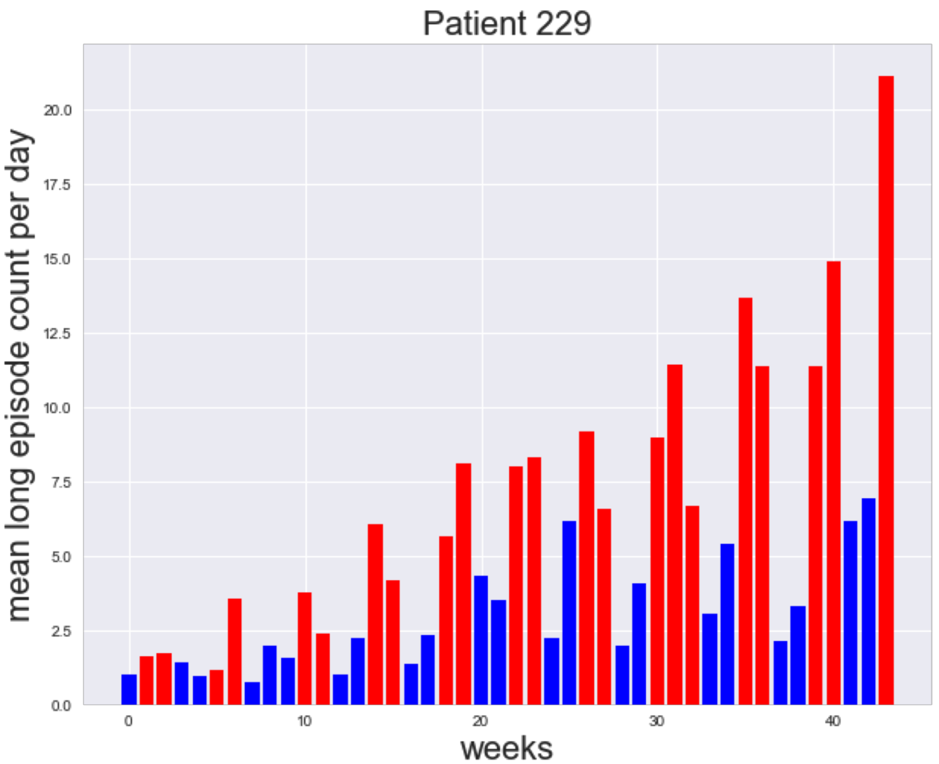
Average long-episode number per day, over one-month epochs



# Results – discussion 3

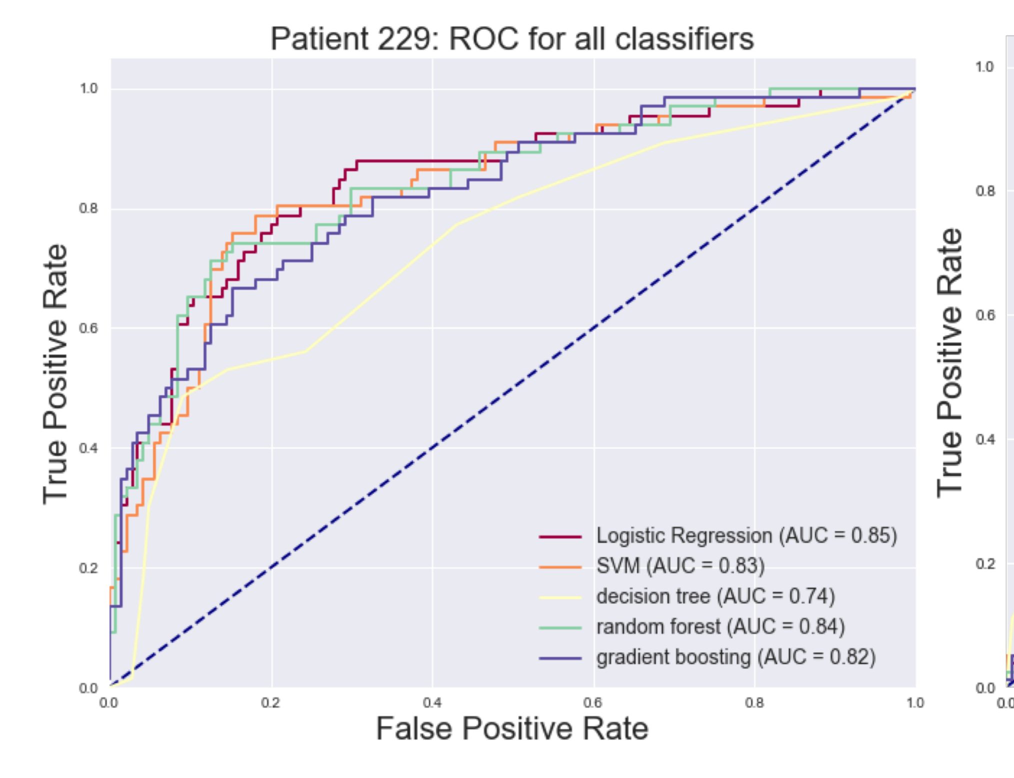
Average long-episode number per day, over one-week epochs



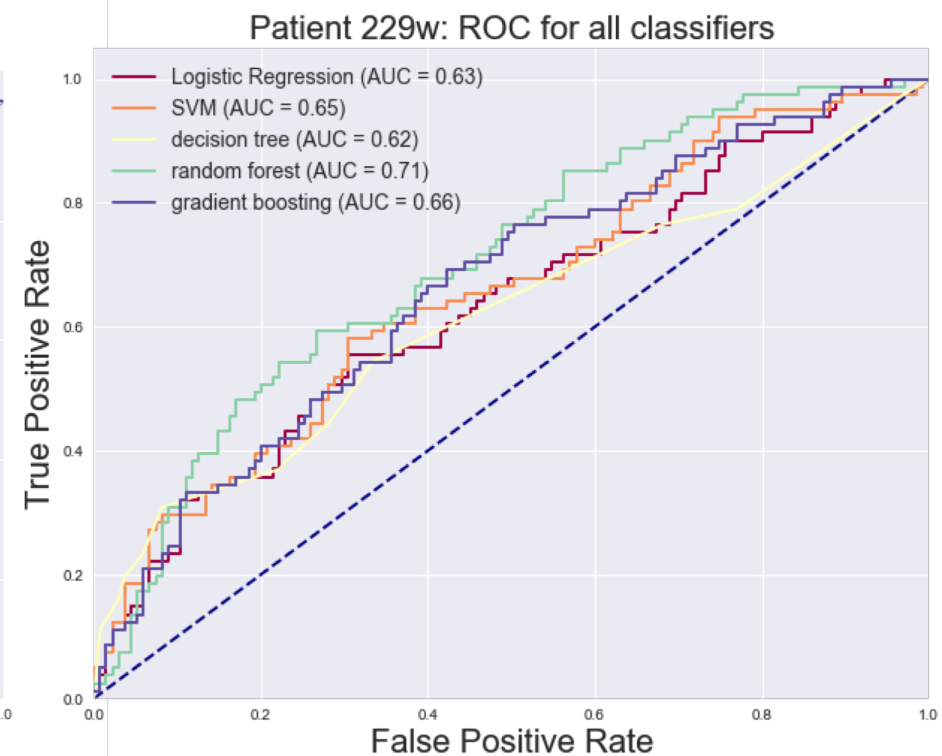



# Results – discussion 3

## 229 monthly ROC plot




## 229 weekly ROC plot



# Summary and Discussion

- Background ECoG can be used to predict clinical outcomes for an individual patient.
- The implication of best classifier. Gradient boosting and Random forest combine the prediction of large number of weak classifiers to achieve better results than using single classifier and prevent overfitting problem.
- Feature importance analysis implicates greater contributions from higher frequency bands.
- Sleep ECoG achieve better classification performance when compared to awake ECoG. The reason may be that sleep EcoG shares stronger correlations with ictal activity.
- Background EEG appears to be equally valuable as pre-ictal EEG in predicting clinical outcome.

# Future steps

- Train the classifier on more patients to see if the machine learning method can generalize to more clinical cases

- Incorporate multidien rhythm information directly in classifier to improve performance

# Appendix

**Experiments**

# Training, Testing and Cross-validation