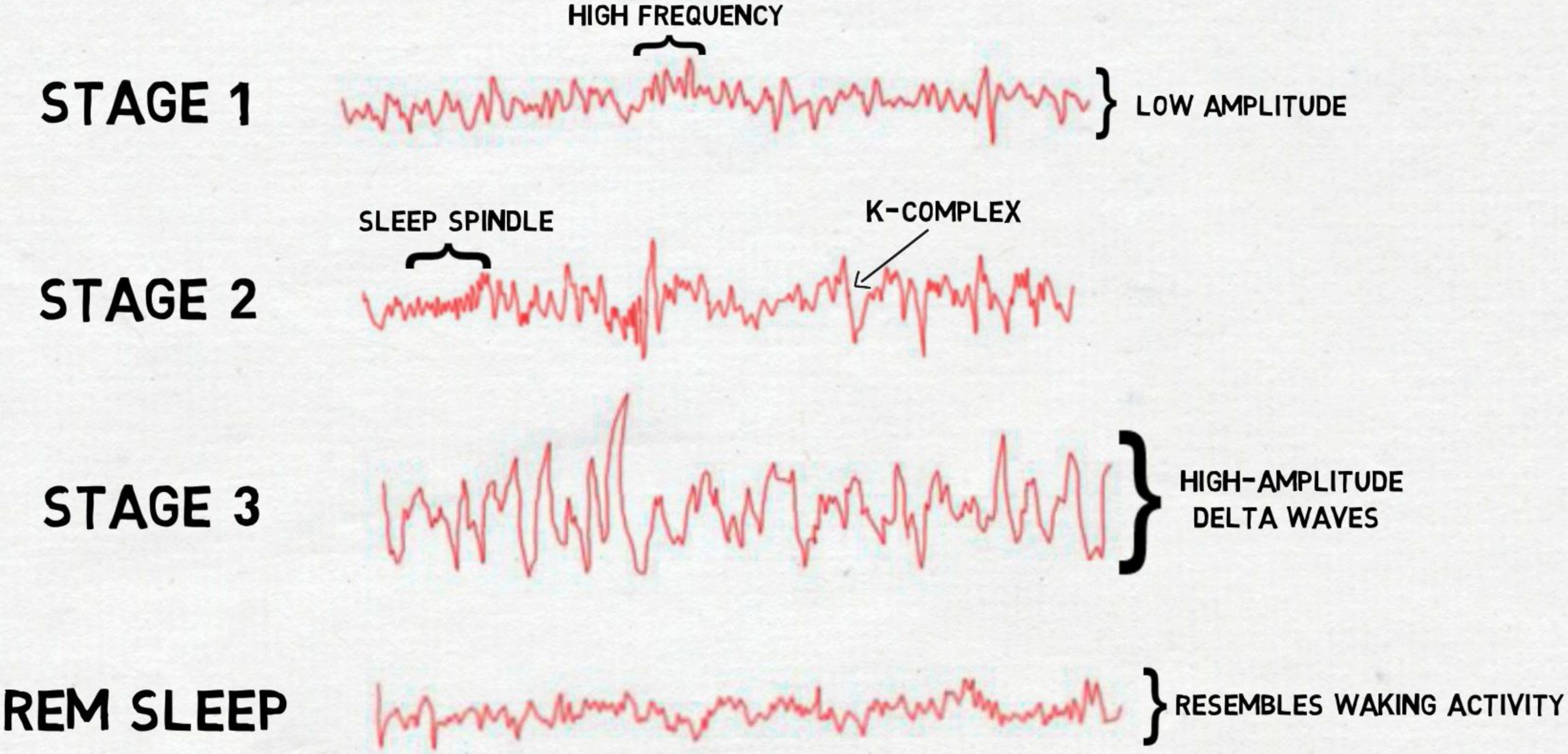
Brain Deep Sleep Slow Oscillation

Jack Dong, Zhentao Hou, Xu Liu, Jingwen Yu

SLEEP STAGES



HIGH FREQUENCY

STAGE 1

LOW AMPLITUDE

SLEEP SPINDLE

K-COMPLEX

STAGE 2

STAGE 3

HIGH-AMPLITUDE DELTA WAVES

REM SLEEP

RESEMBLES WAKING ACTIVITY

Dataset - by dreem

Dataset - by dreem

- Size: 2.7G (261,634 x 1,261)
- URL: https://challengedata.ens.fr/participants/challenges/10/
- Features (1261 columns)
    - Time spent in different sleep stages
    - Current sleep stage
    - Amplitude/duration of previous slow oscillations
    - EEG signals (column 12-1261)
        - represents 10 seconds of recording starting 10 seconds before the end of a slow oscillation.
        - sampling frequency: 125Hz -> 1250 data points

Dataset - by dreem

- Label: if there is any slow oscillation in the following second
    - 0: no slow oscillation
    - 1: a slow oscillation of low amplitude
    - 2: a slow oscillation of high amplitude

Analytical Goals dreem

Predict whether or not a slow oscillation will be followed by another one in sham conditions.

Contributions to this field:

- Predict normal brain activity
- Know when to stimulate
- Compare the impact of individual stimulation

Related Works

- Sleep stage prediction

  Logistic regression

  Random forest

  Deep learning

- Slow oscillations vs heart rate

Our approach uses previous slow oscillation data directly.

Preprocessing Algorithms and Time Efficiency

- No missing data
- Imbalance: not severe (11 : 8 : 7)
- 1250 EEG data points
  min, max, mean, median
- Preprocessing runtime: 87.17 s
  Specs: r5.2xlarge, 5 instances

Model Outcome and Runtime Comparison

r5.2xlarge, 5 instances, 64 G memory, default configure

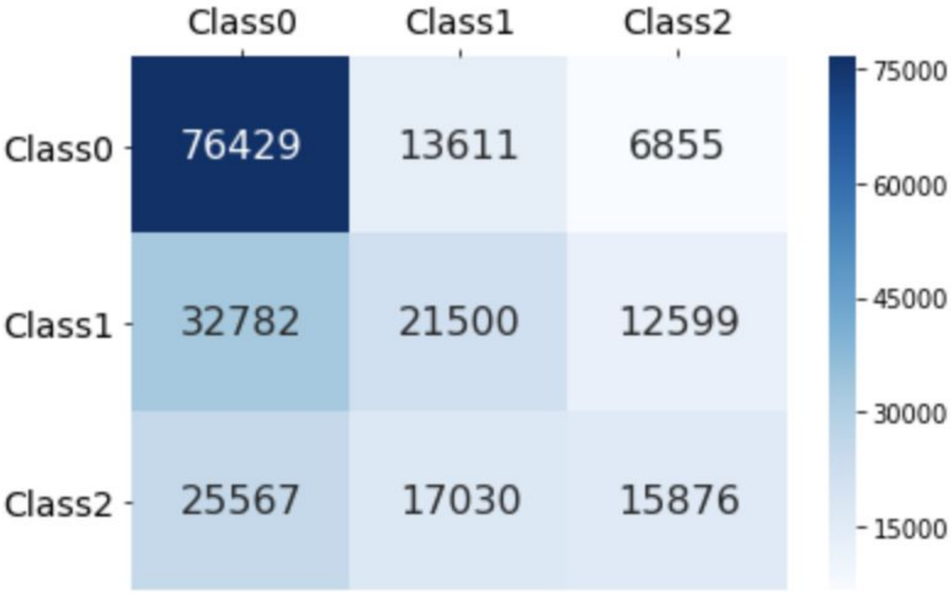
Model	Platform	Accuracy	Runtime
Decision Tree	Spark ML	0.4943	299 s
Random Forest	Spark ML	0.5036	222 s
Gradient Boosting	H2O	0.5121	148 s
XGBoost	H2O	0.5094	301 s
Deep Learning	H2O	0.5055	292 s

Runtime Comparison on the Best Model

Instance Type	Memory Size	Number of Instances	Executor Memory	Runtime
r5.2xlarge	64 G	5	default	147 s
r5.2xlarge	64 G	4	default	141 s
r5.2xlarge	64 G	3	default	151 s
m5.4xlarge	64 G	3	default	139 s

Performance of the Best Model

- Accuracy: **0.5121**
- Confusion Matrix:



	Class0	Class1	Class2
Class0	76429	13611	6855
Class1	32782	21500	12599
Class2	25567	17030	15876

Conclusion & Lesson Learned

- Winner: **Gradient Boosting Machine**
  Highest accuracy and most efficient

- Class0 (no slow oscillation) is relatively easier to predict
- Accuracy generally around 50%
  Missing important features? Better feature engineering?

- Many weird errors occur 😂

Q&A