Detecting Seizures From Randomized One Second Single Channel EEG Epochs

comparing classical and algorithmic approaches

Oliver Shetler

Last Revised April 2022

Introduction

In a hospital setting, interpreting data from instruments that measure the physiological state of patients can be a matter of life and death. Epileptic patients are often monitored by a single or multi-channel electroencephalogram (EEG). One important function of EEG monitoring is to alert doctors to when a patient is having a seizure. This requires that machines automatically classify short epochs of EEG signals into the category of normal activity or seizure activity, accounting for the relative costs of false positives and false negatives. Moreover, distinguishing between EEG signals originating from different brain states and locations is important to many non-invasive machine-brain interfacing applications. In this report, I compare and contrast classical and algorithmic approaches to classifying EEG signals, and discuss their respective merits.

Data

Andrzejak et. al (2001) did sample selection and artifact filtering on EEG recordings and posted the standardized data set on the University of California Irvine's Machine Learning Repository.[1] They divided the recordings into 4,097 randomized epochs of single channel EEG feeds from patients under five conditions:

1. Seizure: EEG recording of seizure activity
2. Tumor: recording of a known epileptic tumor while the patient is not having a seizure.[2]
3. Non-tumor: recording from a healthy brain area in a person who's tumor has been identified.
4. Eyes closed: data from a healthy person while eyes are closed.
5. Eyes open: data from a healthy person while eyes are open.

Each epoch in the data set is 1 second long and consists of 178 data points. I included some plots of a random sample of epochs in Appendix A, so you can get a feel for what they are. The goal of this investigation is to distinguish between seizure vs normal brain activity (1 vs 3).

External validity

One reason to believe the data are of good quality is that they were collected in the normal course of clinical work. Therefore, the sample is representative of the cohort doctors would be interested in monitoring. Moreover, the standards of data collection at a European hospital are likely comparable to other hospitals in the first world.

---

1 The data set is accessible here: https://archive.ics.uci.edu/ml/datasets/Epileptic+Seizure+Recognition

2 Epileptic patients often have a tumor that causes the seizures. These recordings were taken from patients with such a tumor.

Data Frame

The data set is organized so that each row includes a unique identifier and a response variable `y` that encodes the five-level recording condition (numerically coded as above). Additionally, each row contains 178 variables `t1` through `t178` that horizontally encode the time series epochs mentioned above.

Feature Extraction: Classical VS Algorithmic

The classical approach to feature extraction is to think conceptually about which features ought to be relevant to a research question and then to conduct an exploratory analysis of those features. By contrast, the algorithmic, or machine learning oriented, approach is to automate the process of extracting every possible feature, and then algorithmically filter features. Here, I demonstrate the two approaches, starting with the more arduous classical approach.

Classical Feature Extraction via Exploratory Analysis

Before doing any exploratory analysis, 15% of the data (n=690) were set aside as a test set. All the exploratory analysis that follows (up to the results) was conducted only on the training set (85% of the original set, n=3910).

Response Conditions. The epochs from each group were aggregated into a single vector and their summary statistics are presented here (I included only groups 1 and 3 because the others are not relevant to the question being investigated):

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Var.
1 seizure	-1885	-175	-4	-4.20	179	2047	117072.61
3 off tumor	-412	-43	-7	-8.73	26	623	3519.65

The means differ by two fold while the variances differ by one or two orders of magnitude. Therefore, it is likely that variance will play an important role in distinguishing between EEG recordings of seizures vs normal recordings.

The time series data could not be treated as a list of factors in a linear model. In stead, several features were extracted from the time series data, and these were written to a secondary data set, which was subsequently analyzed.

Classical Data Transformations. I extracted several features from each EEG epoch. First, for each epoch, I extracted the mean and variance.

Next, I computed autocorrelations at displacement lengths 1 through 15 for each epoch. In order to select three autocorrelation displacement values for use in modeling, the means of each autocorrelation, divided up by group were computed.

	ac1	ac2	ac3	ac4	ac5	ac6	ac7	ac8	ac9	ac10	ac11	ac12	ac13	ac14	ac15
all	0.94	0.80	0.63	0.46	0.33	0.24	0.17	0.13	0.11	0.10	0.10	0.10	0.11	0.11	0.10
seizure	0.94	0.79	0.59	0.39	0.23	0.11	0.04	-0.01	-0.04	-0.06	-0.08	-0.11	-0.14	-0.17	-0.19
off.tumor	0.97	0.91	0.82	0.74	0.66	0.58	0.51	0.44	0.38	0.33	0.28	0.24	0.20	0.17	0.14

I computed the variance of each column to capture how well each column differentiates the groups.

ac1	ac2	ac3	ac4	ac5	ac6	ac7	ac8	ac9	ac10	ac11	ac12	ac13	ac14	ac15
0.00	0.01	0.03	0.05	0.08	0.08	0.08	0.07	0.05	0.03	0.02	0.01	0.02	0.02	0.03

Additionally, I computed the correlations among each of the autocorrelation values, which are included in Appendix B.

Displacements of 3, 8 and 15 were selected based on the trade-off between correlation and variance.

	ac3	ac8	ac15
ac3	1.00	0.72	0.02
ac8	0.72	1.00	0.18
ac15	0.02	0.18	1.00

I considered the moderate to high correlation between ac3 and ac8 to be a fair trade-off for the high variance of these two autocorrelation factors.

Additionally, I computed the Hurst exponent, which estimates the rate at which a signal's autocorrelaton decays with displacement. This factor can be seen as a way of measuring how causally contingent future activity is on past activity. In the case of brain signals, H should quantify neural feedback.

Next, I computed the frequency band profile for each epoch. In EEG literature, it is conventional to break the data, in the frequency domain (via fourier transform), into five distinct bands: Delta (0-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (12-30 Hz) and Gamma (30-45 Hz). In very broad strokes, delta waves have been associated with the catecholaminergic reward system and the unconstrained pursuit of biological drives (Knyazev, 2007). Theta waves have been associated with emotional decision making (Aftanas et al., 2001). Alpha waves have been associated with quieting mental processes, improving signal to noise ratios (Klimesch et al., 2007). Beta waves have been associated with voluntary action (Nam et al., 2011) and attention / alertness (Gola et al., 2013). Gamma

waves have been associated with the intense cognition involved with sensory binding and mental imagery (Tallon-Baudry and Bertrand, 1999). In the context of a seizure

Lastly, I computed the fractal dimension using the box count method. The fractal dimension helps to quantify the self-similarity and “roughness” of a signal. I observed qualitatively rough and self-simiar signals in the tumor and seizure group when I explored the data set, so I though the fractal dimension might add some useful information to the models.

Algorithmic Feature Extraction

In contrast to the classical approach to feature extraction, the algorithmic, machine learning oriented, method is both more complex and much simpler to implement. The `tsfresh` package was used to automatically extract relevant features. The package identified and returned 411 initial factors (see documentation on features for more information), and then filtered them for relevance. A procedure for multiple significance tests, developed by Benjamini and Yekutieli (2001), was used to find the most significant factors at an alpha level of 0.95. In contrast to the many subjective decisions made in the classical approach, the algorithmic approach to feature extraction entailed deploying a broad net using a one size fits all algorithm, with subjective decisions being implicitly made by the designers of the tsfresh package. This way of doing things reduces the chances of non-experts making conceptual mistakes, but it also increases the risk that experts, who understand the methods, will mindlessly apply inappropriate algorithms.

Modeling

The classical approach to modeling usually involves assessing the merits of different models, using subsets of the engineered features, in one model family (*maybe* two). Moreover, the kinds of models considered are usually restricted to those with well established relationships to probability distributions (parametric models). This enables detailed comparisons and discussion later on. Unfortunately, this can also limit the predictive accuracy of available models. So much so, in fact, that in this report, I was unable to target all five outcome variables because multinomial regressions performed very poorly on the classically engineered features. As a consequence, I had to use logistic regression and pool the non-seizure data into one group. By contrast, in the algorithmic approach, models from several different families are often tested, and nonparametric models are quite common. This limits the possible comparisons and discussion but opens the door to higher performing models. To compensate for the dearth of statistical logic engendered by the algorithmic approach, machine

learning models are trained and assessed on different partitions of the data. This helps to empirically justify the induction of results to the target population without classical statistical arguments.

Classical Modeling Approach

I used logistic regression to compute two models. I first looked at the covariance matrix (see Appendix E) of the feature variables. There were two highly correlated clusters of variables. As a result, I decided to take two different approaches to modeling the data. First, I selected one variable from each cluster so that the cumulative correlation with the rest of the cluster was maximized. Second, I ran a step algorithm on a full model with no interactions in order to find a high performing model.

I evaluated the success of each model by looking a plot of the response rates vs predicted probabilities. Additionally, I evaluated metrics from the confusion matrix.

Algorithmic Modeling Approach

Five machine learning models were tested: (1) a support vector machine, (2) decision trees with bagging (500 estimators), (3) a random forest (500 estimators), (4) decision trees with boosting (500 estimators) and (5) a voting model using (2)-(4). The support vector machine was chosen to get a baseline for what a single linear discriminator could do while the ensemble models were selected because they tend to perform well on data sets with large numbers of features.

The data were divided into a test set and a training set (80% train, 20% test) before the models were implemented. Each model was validated on the test set.

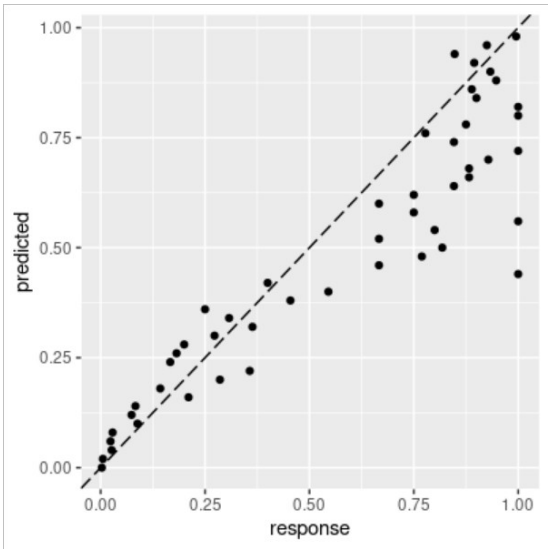
Analysis

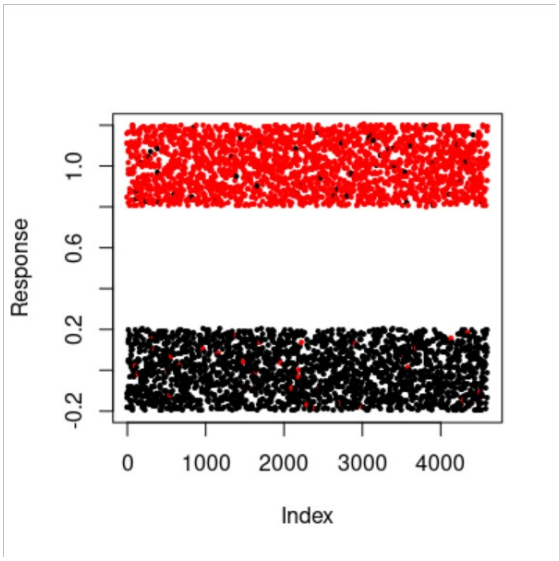
As discussed previously, classical and algorithmic ways of reporting and discussing results differ significantly. The former relies implicitly on the soundness of parametric models; in this case, that of the logistic regression. Details of how every variable relates to the model are attended to as shown below. By contrast, the

Classical Analysis of Results

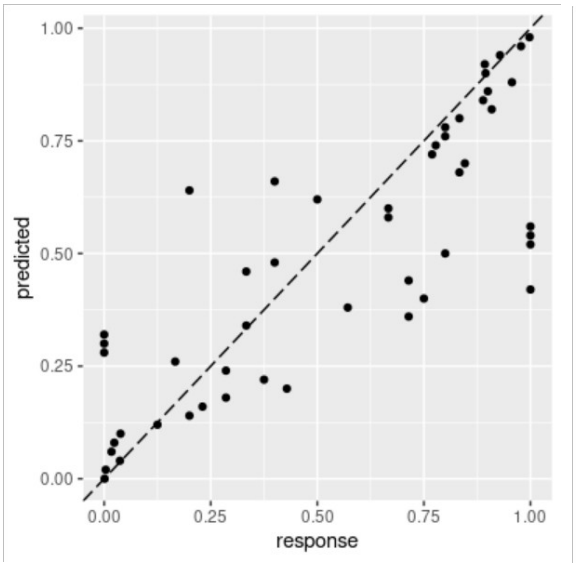
Model 1. I computed the first model using only variance and H (these variables maximized their respective cumulative correlations within each cluster). The summary is available in Appendix C. Both variables were strongly predictive ($p < 2e\text{-}16$). The variance coefficient was very small. A 1,000 unit increase in variance predicted a 1.76 fold increase in the epoch being from a seizure (this will be discussed further in the discussion section). However the orientation of the H coefficient flipped from positive to negative. A 0.1 unit increase in H predicted a 3.96 fold increase in the chances of the epoch

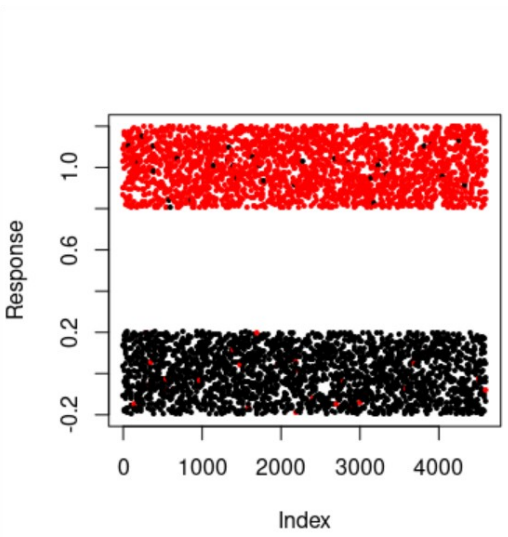
not being from a seizure. The expected vs actual plot was more reasonable (less stratified) than the full model. It is important to note that this plot was constructed using buckets of width 0.01 because the response was Bernoulli. As a consequence, the points in the middle of the distribution may appear to have more weight than they should, because the dots are not scaled according to sample size. Lastly, the model remained a very good predictor of seizures (accuracy = 0.981, sensitivity = 0.983, specificity = 0.986).





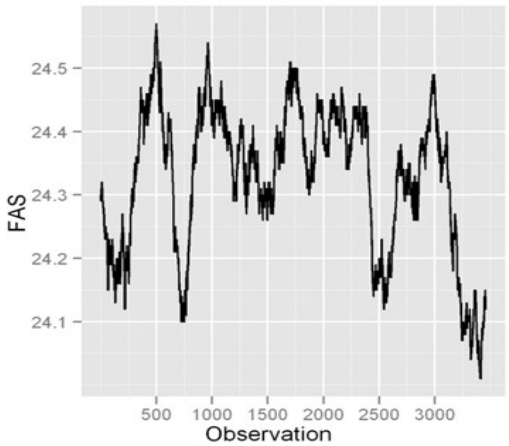
Model 2. I computed the second model using the step function in R with direction set to "both ways." The algorithm selected a model with mean, fD, H, ac15, Delta, Theta, Alpha, Beta, Gamma and variance as coefficients. All coefficients were significantly related to the response (see Appendix D for p values). Given the multicolinearity discussed earlier, it does not make sense to interpret the coefficients individually. This model serves as a predictive performance benchmark, rather than an explanatory model. The expected vs actual plot showed a more or less symmetric distribution about the optimal line, with notable variability towards the middle regions. However, as noted before, the plot is not scaled in proportion to the sample sizes. There were far more predictions near 0 and 1 than in between. As a result, the plot indicates acceptable performance. The model's performance was strong (accuracy = 0.987, sensitivity = 0.983, specificity = 0.986)

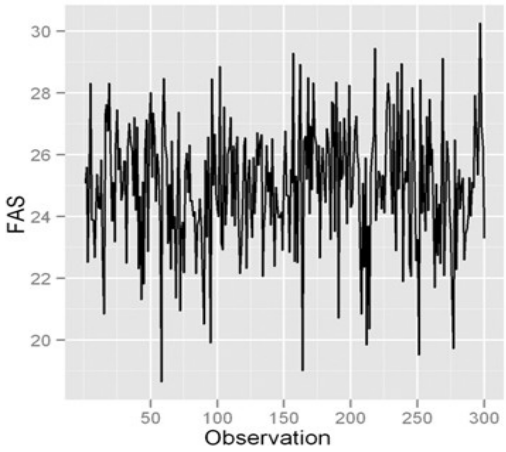




The first model was designed to capture the two main factors that appeared in the covariance matrix. After computing the models, I made another covariance matrix which included the response variate (see Appendix F). H and variance were indeed the most highly (anti)correlated with the response. They were fair representations of the two main factors characterizing whether a signal was from a seizure or not.

The Hurst exponent quantifies the rate at which a signal's autocorrelation decays with displacement. Higher values indicate smoothness and fewer long term dependencies while lower values indicate the opposite. A persistent time series from the stock market (left) has a high Hurst exponent (0.95) while a Brownian time series (right) has a moderate Hurst exponent (0.53)[3]





3 Images were taken from an online article on the Hurst exponent which can be found here: http://analytics-magazine.org/the-hurst-exponent-predictability-of-time-series/

Seizures involve the wide-spread over-synchronization of neurons. During a seizure, electrical signals in the brain form runaway feedback loops that cause convulsions in the afflicted person. Therefore, it might be natural to think that EEG readings from seizures should have high, rather than low Hurst exponents. However, it is important to remember that a single EEG signal is only listening to activity within one local region of the brain. Information from distant, independent regions of the brain can more easily enter a single EEG feed during a seizure than during normal brain activity. On the other hand, during normal brain activity, it is natural to expect local feedback loops---due to activity visible to the EEG feed---to dominate, driving up H. After evaluating the models, I checked to make sure that this intuition was correct by plotting a correlation matrix (see Appendix E) that included the response variable. As one would expect, H was highly negatively correlated (cor = -0.67) with the response variable. Moreover, it had the highest absolute value of any correlation with the response variable.

The association of high variance with seizures is much easier to understand. Seizures lead to very high overall elevated brain activity. A a consequence, it is quite natural to expect that EEG activity during a seizure will vary over a wider range of values. One thing that could appear puzzling is why the variance coefficient from the first model would predict such a low change in probability per change in unit variance. The reason is that the variability of variance values is very high. One standard deviation along the variance dimension is 94,446 units. Thus, it makes sense to think of the predicted change in terms of 1,000 unit changes rather than single unit changes.

The second model demonstrated that, while performance improvement was possible, the simple linear approach was not capable of improving over the two factor model by more than 0.04 percent accuracy. This is encouraging from an explanatory perspective. Over 98 percent of variation between EEG feeds from seizures vs normal tissue comes from differences in variability and autocorrelation related factors. The high performance of the optimized linear model also suggests that further improvements would likely need to come from improved feature extraction, rather than improved model structure.

Algorithmic Analysis of Results

The support vector machine did poorly, with an accuracy of 0.41. A random model would have an accuracy of 0.20. However, the model did successfully predict seizures quite well. This is likely due to the fact that the variances of seizure recordings are often orders of magnitude higher than normal recordings.

	precision	recall	f1-score	support
seizure	0.91	0.90	0.90	460
tumor	0.80	0.03	0.05	457
non-tumor	0.23	0.07	0.10	457
eyes closed	0.39	0.11	0.17	469
eyes open	0.28	0.97	0.44	457
accuracy			0.41	2300
macro avg	0.52	0.41	0.33	2300
weighted avg	0.52	0.41	0.33	2300

The ensemble models all did better.

The bagged tree model did significantly better, with an accuracy score of 0.71. It also performed most strongly on the seizure group. It did more poorly on the tumor and non-tumor groups. This may be because they have similar properties due to the measurements being taken during the same procedures.

	precision	recall	f1-score	support
seizure	0.90	0.95	0.92	460
tumor	0.53	0.46	0.49	457
non-tumor	0.57	0.60	0.58	457
eyes closed	0.84	0.81	0.82	469
eyes open	0.70	0.75	0.73	457
accuracy			0.71	2300
macro avg	0.71	0.71	0.71	2300
weighted avg	0.71	0.71	0.71	2300

The random forest model performed the best over-all with an accuracy score of 0.84. The main improvements came from the tumor and non-tumor groups, although there was substantial improvement in the eyes closed and eyes open groups as well.

	precision	recall	f1-score	support
seizure	0.97	0.98	0.97	460
tumor	0.77	0.70	0.73	457
non-tumor	0.73	0.79	0.76	457
eyes closed	0.92	0.90	0.91	469
eyes open	0.84	0.87	0.85	457
accuracy			0.84	2300
macro avg	0.84	0.84	0.84	2300
weighted avg	0.85	0.84	0.84	2300

The boosted tree model did a little bit worse than the random forest model, with an accuracy score of 0.81. However, this difference is likely not significant for outside validity (if the models were tested on a truly external group of readings from another data set, they would likely perform comparably).

	precision	recall	f1-score	support
seizure	0.94	0.97	0.95	460
tumor	0.69	0.61	0.65	457
non-tumor	0.70	0.74	0.72	457
eyes closed	0.92	0.88	0.90	469
eyes open	0.81	0.86	0.83	457
accuracy			0.81	2300
macro avg	0.81	0.81	0.81	2300
weighted avg	0.81	0.81	0.81	2300

Lastly, the voting model did not improve on the three ensemble models, with an accuracy score of 0.83. This is likely due to the similarity of the models tried. It would have been interesting to test a neural network. Unfortunately, as the fourth out of five final projects, due each on one day of the week, this project had to be simple.

	precision	recall	f1-score	support
seizure	0.95	0.97	0.96	460
tumor	0.72	0.71	0.71	457
non-tumor	0.77	0.71	0.74	457
eyes closed	0.93	0.87	0.90	469
eyes open	0.80	0.89	0.84	457
accuracy			0.83	2300
macro avg	0.83	0.83	0.83	2300
weighted avg	0.83	0.83	0.83	2300

In the algorithmic sections, it was demonstrated how ensemble methods can drastically improve the performance of EEG prediction models when compared with a linear discriminator such as the support vector machine. One of the primary applications of these kinds of models is to detect seizures and alert doctors. The precision and recall for seizures for all the ensemble methods (95-98%) is acceptable considering that each epoch is only one second long (the chances of a string of false

predictions lasting more than a few seconds is likely minuscule, even if the epochs are highly correlated). However, there is clearly still room for improvement.

One drawback of these ensemble methods is that they are hard to interpret scientifically. Moreover, the best performing EEG models use deep neural networks to extract additional information.

Discussion

We have seen throughout this paper that the classical and algorithmic approach to predictive modeling differs significantly. Both approaches have major benefits and drawbacks. The greatest strength of the classical approach is that it relies on explanatory logic, which makes it easier to use for persuasion. In science and business, persuasion plays an important role in building credibility, so classical models are often preferred due to their persuasive powers. Unfortunately, this great strength is attached to a major weakness. By relying on interpretable models, the classical approach entails many subjective decisions and limits the algorithms available. In practice, this often severely hampers the quantifiable performance of classical models. By contrast, the algorithmic approach has as its greatest strength its ability to produce very high performing models. Unfortunately, many machine learning models are not supported by explanatory logic, so, despite often performing well, they can be uninformative and unconvincing. However, many machine learning models can be justified using cross validation, which is a naive but effective inductive argument. In practice, the algorithmic approach can often compensate for its lack of explanatory rigor by showing a model to be stable across time and sample spaces. As a result, in business, the high performing algorithmic models are often accepted because the performance trade-off is worth the risk of not having a clear explanation for why the model works.

In this report, we saw the strengths and weaknesses of the two approaches to modeling in the context of seizure detection and, more generally, EEG classification. We found that the classical feature extraction approach produced fewer features to work with; and moreover the feature quality was not particularly high. By contrast, the algorithmic approach extracted hundreds of features and relied on statistical testing theory to safely find a subset of relevant ones. Ironically, in this instance, the algorithmic approach was more statistically sound than the classical approach, because the correct statistical theory was automatically applied to a pre-defined set of features. Nevertheless, the algorithmic approach was less interpretable because the features were only shown to be relevant post hoc, rather than chosen for their popularity in the EEG literature.

The classical approach produced models that were much weaker than the machine learning approach. Multinomial models were not even included in this report as they did worse than chance at classifying the EEG epochs. Even so, the accuracy, specificity and sensitivity of logistic regressions for seizure detection was quite strong (98-99%). Therefore, despite the performance deficits, logistic regressions might be preferred for emergency medical applications where stakeholders would want to know exactly how and why the models work.

In contrast to the relative weakness of the classical models, most of the machine learning models were able to classify all five conditions and still detect seizures at rates rivaling the logistic regressions, with recall scores of 96-98%. These models are more versatile yet harder to interpret, so they might be preferred in a commercial setting where fine grained distinctions are more important and risk is lower.

References

Aftanas, L., Varlamov, A., Pavlov, S., Makhnev, V., & Reva, N. (2001). Event-related synchronization and desynchronization during affective processing: emergence of valence-related time-dependent hemispheric asymmetries in theta and upper alpha band. *International journal of Neuroscience*, *110*(3-4), 197-219.

Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., & Elger, C. E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, *64*(6), 061907.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. Annals of statistics, 1165–1188

Gola, M., Magnuski, M., Szumska, I., & Wróbel, A. (2013). EEG beta band activity is related to attention and attentional deficits in the visual performance of elderly subjects. *International Journal of Psychophysiology*, *89*(3), 334-341.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain research reviews*, *53*(1), 63-88.

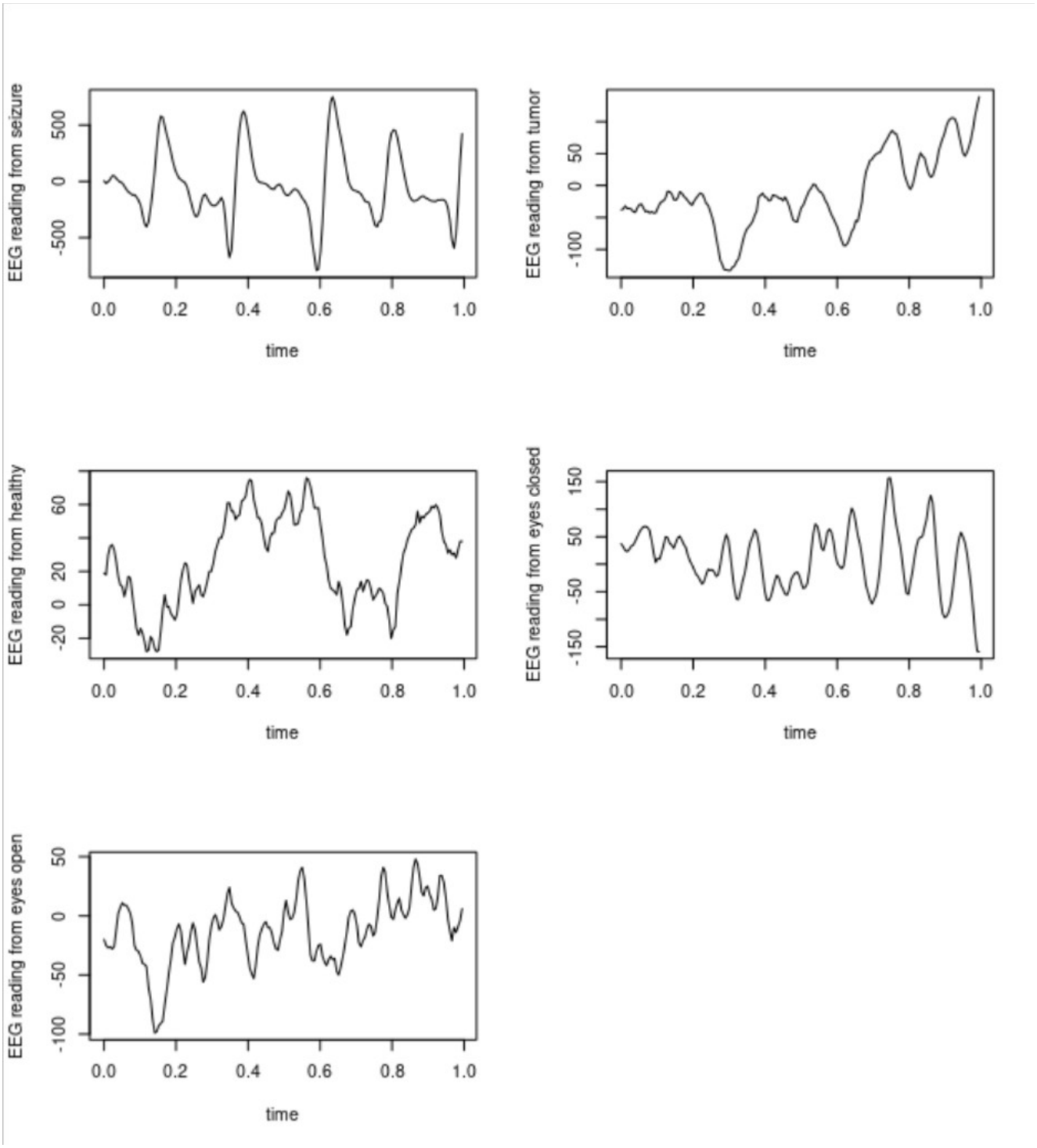
Knyazev, G. G. (2007). Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neuroscience & Biobehavioral Reviews*, *31*(3), 377-395.

Nam, C. S., Jeon, Y., Kim, Y. J., Lee, I., & Park, K. (2011). Movement imagery-related lateralization of event-related (de) synchronization (ERD/ERS): motor-imagery duration effects. *Clinical Neurophysiology*, *122*(3), 567-577.

Patel, A. A., Jannati, A., Dhamne, S. C., Sapuwa, M., Kalanga, E., Mazumdar, M., ... & Rotenberg, A. (2020). EEG markers predictive of epilepsy risk in pediatric cerebral malaria–A feasibility study. *Epilepsy & Behavior*, *113*, 107536.

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences*, *3*(4), 151-162.

Appendix A


EEG reading from seizure
-500
0
500
0.0
0.2
0.4
0.6
0.8
1.0
time
EEG reading from tumor
-100
0
50
0.0
0.2
0.4
0.6
0.8
1.0
time
EEG reading from healthy
-20
0
20
60
0.0
0.2
0.4
0.6
0.8
1.0
time
EEG reading from eyes closed
-150
-50
50
150
0.0
0.2
0.4
0.6
0.8
1.0
time
EEG reading from eyes open
-100
-50
0
50
0.0
0.2
0.4
0.6
0.8
1.0
time

Appendix B

	ac1	ac2	ac3	ac4	ac5	ac6	ac7	ac8	ac9	ac10	ac11	ac12	ac13	ac14	ac15
ac1	1.00	0.98	0.96	0.91	0.84	0.75	0.65	0.56	0.47	0.37	0.27	0.17	0.07	0.02	0.01
ac2	0.98	1.00	0.99	0.96	0.91	0.83	0.74	0.65	0.55	0.44	0.33	0.21	0.10	0.03	0.01
ac3	0.96	0.99	1.00	0.99	0.95	0.89	0.81	0.72	0.62	0.50	0.38	0.25	0.13	0.05	0.02
ac4	0.91	0.96	0.99	1.00	0.99	0.94	0.88	0.80	0.70	0.57	0.44	0.30	0.17	0.08	0.03
ac5	0.84	0.91	0.95	0.99	1.00	0.98	0.94	0.87	0.78	0.65	0.51	0.36	0.22	0.11	0.05
ac6	0.75	0.83	0.89	0.94	0.98	1.00	0.98	0.94	0.86	0.74	0.59	0.44	0.28	0.15	0.08
ac7	0.65	0.74	0.81	0.88	0.94	0.98	1.00	0.98	0.93	0.82	0.69	0.53	0.36	0.22	0.12
ac8	0.56	0.65	0.72	0.80	0.87	0.94	0.98	1.00	0.98	0.91	0.80	0.64	0.46	0.30	0.18
ac9	0.47	0.55	0.62	0.70	0.78	0.86	0.93	0.98	1.00	0.97	0.90	0.76	0.58	0.40	0.27
ac10	0.37	0.44	0.50	0.57	0.65	0.74	0.82	0.91	0.97	1.00	0.97	0.87	0.71	0.54	0.39
ac11	0.27	0.33	0.38	0.44	0.51	0.59	0.69	0.80	0.90	0.97	1.00	0.96	0.85	0.70	0.56
ac12	0.17	0.21	0.25	0.30	0.36	0.44	0.53	0.64	0.76	0.87	0.96	1.00	0.96	0.86	0.74
ac13	0.07	0.10	0.13	0.17	0.22	0.28	0.36	0.46	0.58	0.71	0.85	0.96	1.00	0.97	0.89
ac14	0.02	0.03	0.05	0.08	0.11	0.15	0.22	0.30	0.40	0.54	0.70	0.86	0.97	1.00	0.98
ac15	0.01	0.01	0.02	0.03	0.05	0.08	0.12	0.18	0.27	0.39	0.56	0.74	0.89	0.98	1.00

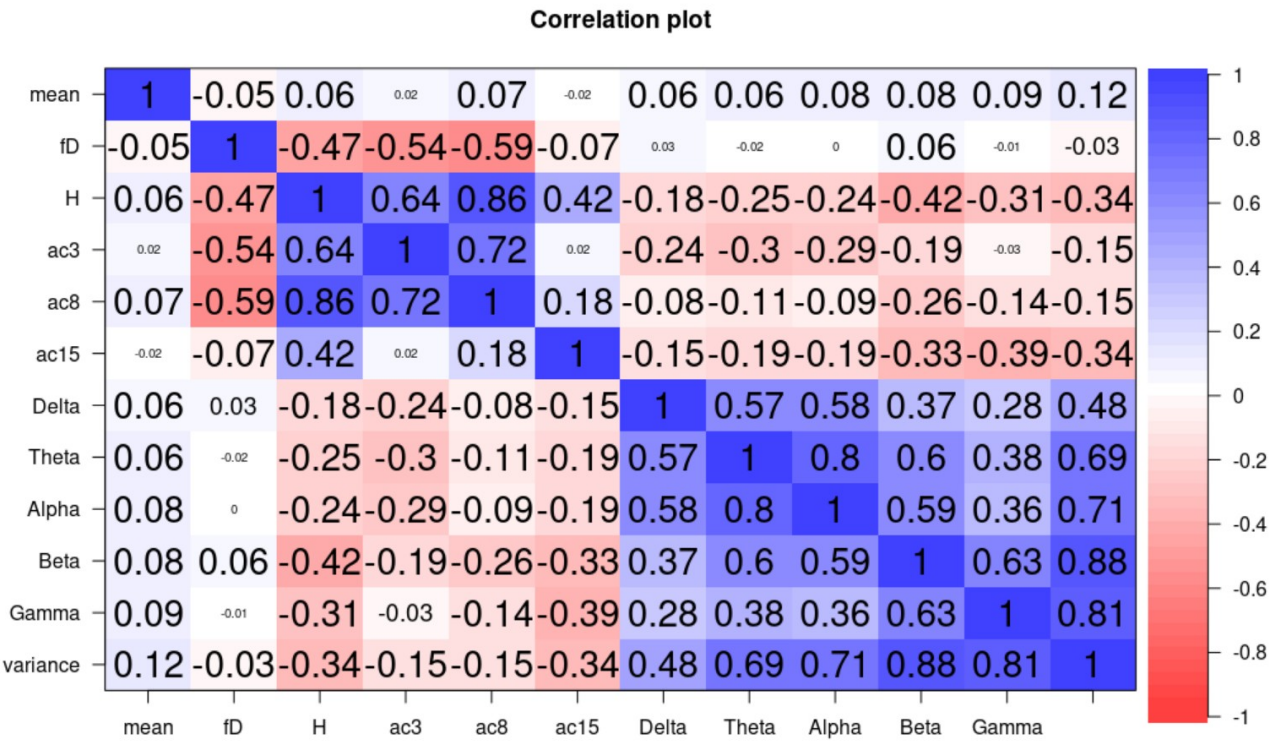
Appendix C

```
Call:
glm(formula = response ~ variance + H, family = binomial(link = "logit"),
    data = fd[2:14], control = list(maxit = 500))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-4.1874  -0.1035  -0.0113   0.0000   3.5307

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.639e+00  9.175e-01   6.146 7.93e-10 ***
variance     5.648e-04  2.777e-05  20.335  < 2e-16 ***
H           -1.375e+01  1.170e+00 -11.751  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6376.95  on 4599  degrees of freedom
Residual deviance:  569.71  on 4597  degrees of freedom
AIC: 575.71
Number of Fisher Scoring iterations: 12
```

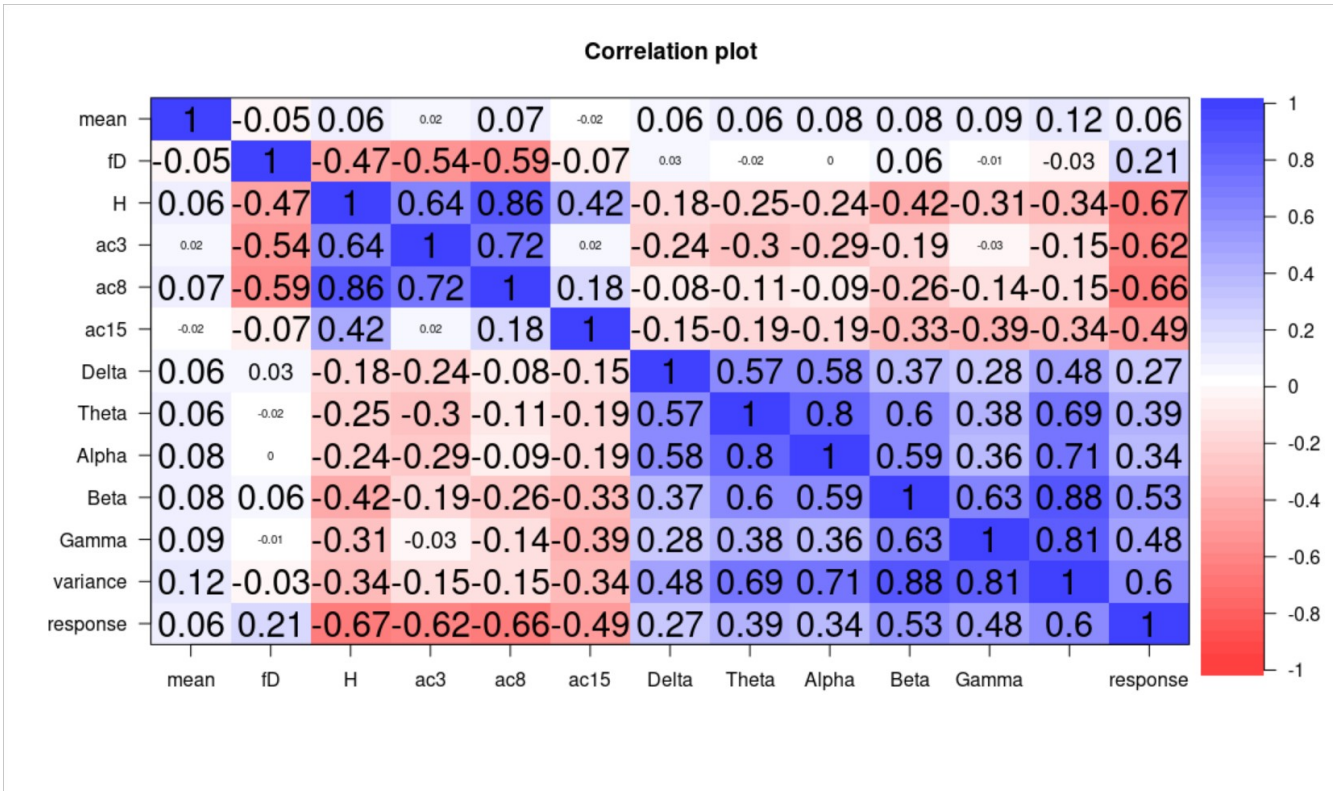
Appendix D

```
glm(formula = response ~ mean + fD + H + ac15 + Delta + Theta +
    Alpha + Beta + Gamma + variance, family = binomial(link = "logit"),
    data = fd[2:14], control = list(maxit = 500))

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.5852  -0.0741  -0.0034   0.0000   3.4988

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.976e+00  2.917e+00   2.049 0.040494 *
mean        -2.174e-02  4.977e-03  -4.369 1.25e-05 ***
fD          -4.581e+00  1.086e+00  -4.218 2.46e-05 ***
H           -7.993e+00  2.828e+00  -2.827 0.004704 **
ac15         2.380e+00  1.133e+00   2.101 0.035661 *
Delta       -2.200e-02  2.028e-03 -10.851  < 2e-16 ***
Theta        8.980e-04  2.466e-04   3.642 0.000270 ***
Alpha        1.441e-04  3.820e-05   3.773 0.000161 ***
Beta         1.033e-04  1.572e-05   6.570 5.03e-11 ***
Gamma        1.352e-05  4.664e-06   2.899 0.003739 **
variance     2.428e-04  4.285e-05   5.665 1.47e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6376.95  on 4599  degrees of freedom
Residual deviance:  368.13  on 4589  degrees of freedom
AIC: 390.13

Number of Fisher Scoring iterations: 15
```

Appendix E


Correlation plot
mean 1 -0.05 0.06 0.02 0.07 -0.02 0.06 0.06 0.08 0.08 0.09 0.12
fD -0.05 1 -0.47 -0.54 -0.59 -0.07 0.03 -0.02 0 0.06 -0.01 -0.03
H 0.06 -0.47 1 0.64 0.86 0.42 -0.18 -0.25 -0.24 -0.42 -0.31 -0.34
ac3 0.02 -0.54 0.64 1 0.72 0.02 -0.24 -0.3 -0.29 -0.19 -0.03 -0.15
ac8 0.07 -0.59 0.86 0.72 1 0.18 -0.08 -0.11 -0.09 -0.26 -0.14 -0.15
ac15 -0.02 -0.07 0.42 0.02 0.18 1 -0.15 -0.19 -0.19 -0.33 -0.39 -0.34
Delta 0.06 0.03 -0.18 -0.24 -0.08 -0.15 1 0.57 0.58 0.37 0.28 0.48
Theta 0.06 -0.02 -0.25 -0.3 -0.11 -0.19 0.57 1 0.8 0.6 0.38 0.69
Alpha 0.08 0 -0.24 -0.29 -0.09 -0.19 0.58 0.8 1 0.59 0.36 0.71
Beta 0.08 0.06 -0.42 -0.19 -0.26 -0.33 0.37 0.6 0.59 1 0.63 0.88
Gamma 0.09 -0.01 -0.31 -0.03 -0.14 -0.39 0.28 0.38 0.36 0.63 1 0.81
variance 0.12 -0.03 -0.34 -0.15 -0.15 -0.34 0.48 0.69 0.71 0.88 0.81 1
mean fD H ac3 ac8 ac15 Delta Theta Alpha Beta Gamma
1 0.8 0.6 0.4 0.2 0 -0.2 -0.4 -0.6 -0.8 -1

Appendix F


Correlation plot
mean 1 -0.05 0.06 0.02 0.07 -0.02 0.06 0.06 0.08 0.08 0.09 0.12 0.06
fD -0.05 1 -0.47 -0.54 -0.59 -0.07 0.03 -0.02 0 0.06 -0.01 -0.03 0.21
H 0.06 -0.47 1 0.64 0.86 0.42 -0.18 -0.25 -0.24 -0.42 -0.31 -0.34 -0.67
ac3 0.02 -0.54 0.64 1 0.72 0.02 -0.24 -0.3 -0.29 -0.19 -0.03 -0.15 -0.62
ac8 0.07 -0.59 0.86 0.72 1 0.18 -0.08 -0.11 -0.09 -0.26 -0.14 -0.15 -0.66
ac15 -0.02 -0.07 0.42 0.02 0.18 1 -0.15 -0.19 -0.19 -0.33 -0.39 -0.34 -0.49
Delta 0.06 0.03 -0.18 -0.24 -0.08 -0.15 1 0.57 0.58 0.37 0.28 0.48 0.27
Theta 0.06 -0.02 -0.25 -0.3 -0.11 -0.19 0.57 1 0.8 0.6 0.38 0.69 0.39
Alpha 0.08 0 -0.24 -0.29 -0.09 -0.19 0.58 0.8 1 0.59 0.36 0.71 0.34
Beta 0.08 0.06 -0.42 -0.19 -0.26 -0.33 0.37 0.6 0.59 1 0.63 0.88 0.53
Gamma 0.09 -0.01 -0.31 -0.03 -0.14 -0.39 0.28 0.38 0.36 0.63 1 0.81 0.48
variance 0.12 -0.03 -0.34 -0.15 -0.15 -0.34 0.48 0.69 0.71 0.88 0.81 1 0.6
response 0.06 0.21 -0.67 -0.62 -0.66 -0.49 0.27 0.39 0.34 0.53 0.48 0.6 1
mean fD H ac3 ac8 ac15 Delta Theta Alpha Beta Gamma response
1 0.8 0.6 0.4 0.2 0 -0.2 -0.4 -0.6 -0.8 -1