Predicting seizures using artificial neural networks

Jeremy Angel [1] Matthew Wootten [1]

[1] Loudoun Academy of Science

2018-10-18

What are seizures? (Background)

What if you could go unconscious at any time, with minimal warning?

What are seizures? (Background)

What if you could go unconscious at any time, with minimal warning?

- You can't drive — you might crash

What are seizures? (Background)

What if you could go unconscious at any time, with minimal warning?

- You can't drive — you might crash
- You must avoid sharp objects — you could hurt yourself

What are seizures? (Background)

What if you could go unconscious at any time, with minimal warning?

- You can't drive — you might crash
- You must avoid sharp objects — you could hurt yourself

This is the everyday reality for about 1% of people: people with epilepsy.

Figure: Image: Damnsoft 09

Objectively, seizures are bad (Background)

Camfield and Camfield, 2010 studied a group of 15-30 year olds with idiopathic generalized epilepsy (no brain damage outside of the epilepsy) tonic-clonic seizures (seizures that cause whole-body convulsions). This is considered *mild*:

Objectively, seizures are bad (Background)

Camfield and Camfield, 2010 studied a group of 15-30 year olds with idiopathic generalized epilepsy (no brain damage outside of the epilepsy) tonic-clonic seizures (seizures that cause whole-body convulsions). This is considered *mild*:

- 40% couldn't graduate high school

Objectively, seizures are bad (Background)

Camfield and Camfield, 2010 studied a group of 15-30 year olds with idiopathic generalized epilepsy (no brain damage outside of the epilepsy) tonic-clonic seizures (seizures that cause whole-body convulsions). This is considered *mild*:

- 40% couldn't graduate high school
- 33% were unemployed

Objectively, seizures are bad (Background)

Camfield and Camfield, 2010 studied a group of 15-30 year olds with idiopathic generalized epilepsy (no brain damage outside of the epilepsy) tonic-clonic seizures (seizures that cause whole-body convulsions). This is considered *mild*:

- 40% couldn't graduate high school
- 33% were unemployed
- 27% had a psychiatric diagnosis

Seizure biology (Background)

- Seizures occur when a defective cluster of neurons sends an unusually strong signal, causing a positive feedback loop
- This signal then spreads across the brain, affecting a large area resulting in a seizure
- Thus, if we look into brain activity for signs of future abnormality, we may be able to predict an upcoming seizure

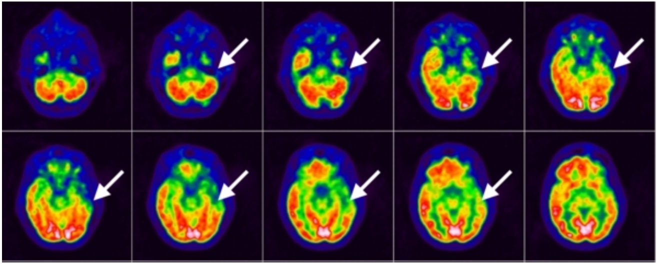


Figure: Seizure spreading through a brain. Credit: University of Iowa Medical Center.

What is an EEG? (Background)

An electroencephalogram (EEG) measures brain activity by recording the voltage difference between pairs of electrodes. EEGs can be either taken from the scalp (extracranial) or inside the skull (intracranial).

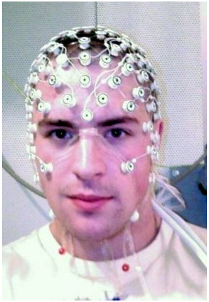
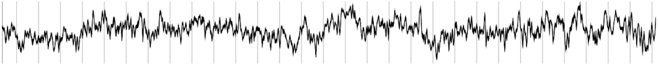


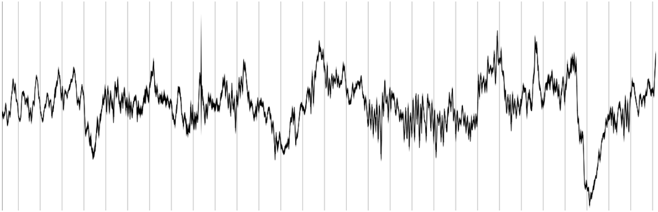
Figure: Man with an extracranial EEG

Examples of EEGs (Background)

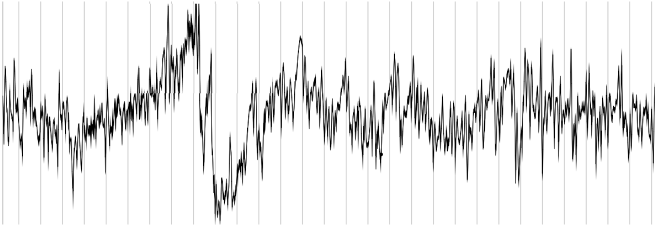
- Outside of a seizure



- Immediately before a seizure



- During a seizure

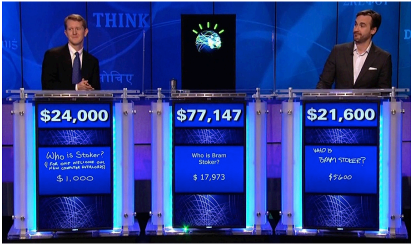


What is machine learning? (Background)

Machine learning is a class of statistical methods that let computers solve a wide variety of problems using data, rather than a hand designed algorithm.

Machine learning can do all sorts of things:

- YouTube can recommend videos the viewer might like
- AlphaGo can beat world champions at Go
- IBM Watson can dominate Jeopardy games



Why machine learning? (Background)

- In the past, differential equations were the standard for math modeling projects.
- Differential equations are limited though in that they can practically only be used in two or three dimensions.
- Machine learning algorithms, such as neural networks and support vector machines, can be used for problems with hundreds of dimensions.
- Machine learning algorithms are also able to solve a wide variety of given problems while differential equations must be designed for a specific problem

What are our inputs?

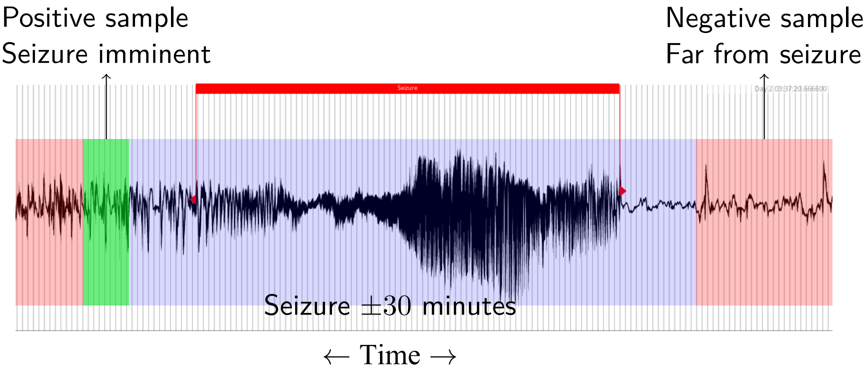


Figure: Credit for EEG: University of Pennsylvania and Mayo Clinic, 2014

What are our inputs?

- We take each pair of inputs and feed it in seperately, so that the network can learn cross-channel features.
- For example, if the EEG had two channels (C1 and C2), we would feed in the following combinations:
  - C1, C1
  - C1, C2
  - C2, C1
  - C2, C2

Fourier Transform (Preprocessing)

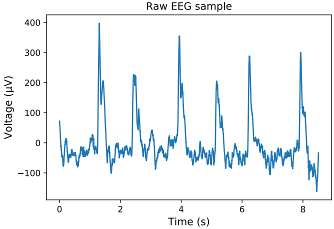


Figure: Complex multi-component signal

→

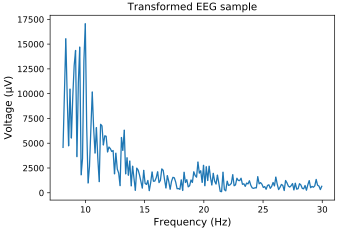


Figure: Frequency decomposition

The EEG channels are transformed into their individual frequencies. We then take out alpha and beta brain waves (the 8-30 Hz range) for input to the neural networks, because they carry much of the relevant information (Darcey & Williamson, 1985).

Traditional neural network (Model)

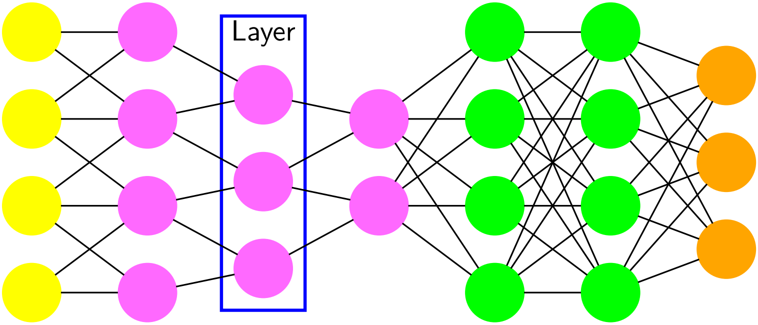


Figure: A neural network uses “neurons” and “synapses” to “learn”.

Analyzing with confusion matrices (Analysis)

Collect the data into something that looks like this:

Confusion Matrix		Actual	
		Seizure	Not seizure
Predicted	Seizure	True pos.	False pos.
	Not seizure	False neg.	True neg.

While our classifier can be tuned, its output at one setting looks like this:

Confusion Matrix		Actual	
		Seizure	Not seizure
Predicted	Seizure	17	12
	Not seizure	14	19

Summary metrics (Analysis)

- True positive rate (sensitivity): how often a seizure **was predicted** and **actually occurred**
- False positive rate: how often a seizure **was predicted** but **never occurred**
- There's a tradeoff between the two, so we need to pick a threshold.
- What if we try all the thresholds at once? What does that look like?

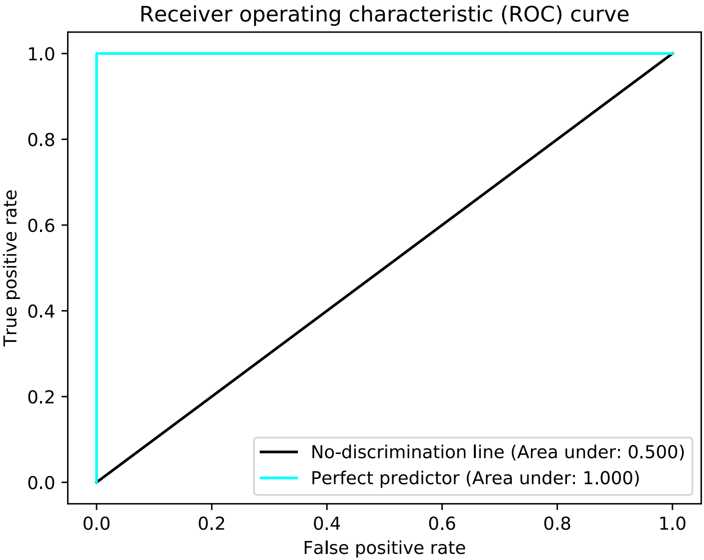


Figure: ROC graph of detection methods, with the area under them (AUROC)

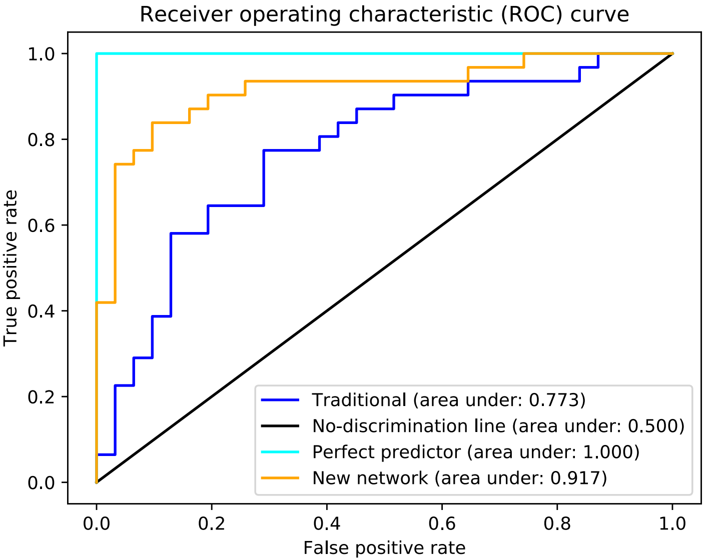


Figure: ROC graph of detection methods, with the area under them (AUROC)

Future work

- Determine how to choose optimal channel pairs
- Tune hyperparameters (settings)
- Combine data from multiple patients, using transfer learning to keep multiple patients partially separate
- Predict in real-time
- Test our algorithm on implantable hardware
- Transition to a handpicked feature-based model

Acknowledgments

- Ms. Na-eun Roh at DSHS
- Mr. Writer at AOS
- Dr. Scheffer at HHMI
- Dr. Turaga at HHMI
- Dr. Oguz at NIH

Cited in presentation I

Camfield, P. & Camfield, C. (2010). Idiopathic generalized epilepsy with generalized tonic-clonic seizures (IGE-GTC): a population-based cohort with > 20 year follow up for medical and social outcome. *Epilepsy & Behavior*, *18*(1-2), 61–63. doi:10.1016/j.yebeh.2010.02.014

D'Alessandro, M., Vachtsevanos, G., Esteller, R., Echauz, J., Cranstoun, S., Worrell, G., … Litt, B. (2005). A multi-feature and multi-channel univariate selection process for seizure prediction. *Clinical Neurophysiology*, *116*(3), 506–516. doi:10.1016/j.clinph.2004.11.014

Netoff, T., Park, Y., & Parhi, K. (2009). Seizure prediction using cost-sensitive support vector machine. In *2009 annual international conference of the IEEE engineering in medicine and biology society*. IEEE. doi:10.1109/iembs.2009.5333711

Major influences I

Ghosh-Dastidar, S. & Adeli, H. (2009). A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection. *Neural Networks*, *22*, 1419–1431. doi:10.1016/j.neunet.2009.04.003

University of Pennsylvania and Mayo Clinic. (2014). *IEEG-Portal documentation: Collaborative research in the cloud*. Retrieved from https://main.ieeg.org/sites/default/files/IEEGDocumentation.pdf

Predicting with Electroencephalograms (EEGs)

Figure: Non-exhaustive list of previous EEG-based prediction algorithms

D'Alessandro et al., 2005	Netoff, Park, and Parhi, 2009
▶ 10 minutes ▶ 100% sensitivity ▶ 1.1 false positives/hour	▶ 5 minutes ▶ 79% sensitivity ▶ 0.0 false positives/hour

SVM (Support Vector Machine)

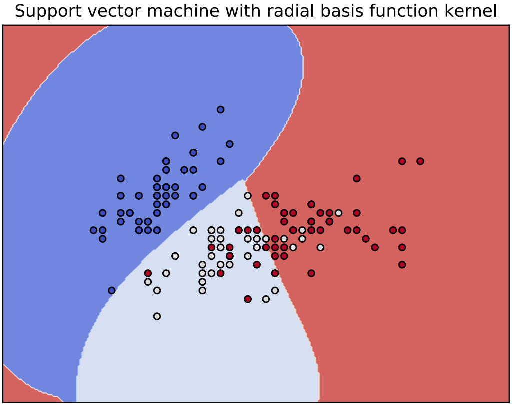


Figure: A sample of a support vector machine with an radial basis function kernel (Pedregosa et al., 2011)

Converting into spikes

The procedure for converting real outputs into spikes borrows from nature. Similar to how the eye's neurons detect light, this model spikes with different intensity depending on how close the value is to the target.

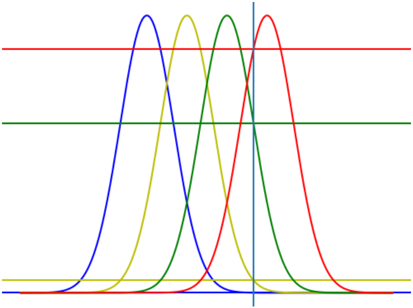


Figure: Mimicking receptive fields. Adapted from Ghosh-Dastidar and Adeli (2009).

Multi-spiking internal state function

$$x_j(t) = \sum_{i=1}^{N_{l+1}} \sum_{k=1}^{K} \sum_{g=1}^{G_i} w_{ij}^k \varepsilon(t - t_i^{(g)} - d^k) + \rho(t - t_j^{(f)})$$

- Sum across all neuron indices in the previous layer
- Sum across all synapses
- Sum across all spike times

- Typical neuron connections
- Refractoriness (carryover based on previous neuron state)

Backpropagation, conceptually

- The idea behind backpropagation is to treat the network as an optimization problem
  - The function being optimized—the cost function—is the error of the network on its training examples
  - The variables on that function are the weights of the network
- Since this optimization is far too complex to manage analytically, we use gradient descent to approximate it.

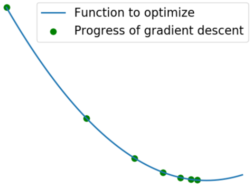


Figure: Gradient descent in action

Gradient descent

$$x' = x - \eta \frac{\partial E}{\partial x}$$

Network parameters

- Traditional network: $135 \to 45 \to 15 \to 5 \to 1$
- Multi-spiking (1): $21 \to 5 \to 1$
- Multi-spiking (2): $61 \to 15 \to 4 \to 1$
- Single-spiking (1): $5 \to 3 \to 2$
- Single-spiking (2): $15 \to 6 \to 2$

The end

This is the end. The end is here.

The Math

Figure: Credit: Randall Munroe

Bonus credits

- Saunak, for lending us a charger
- Alexandra Elbakyan, for creating Sci-Hub
- Rohan and Jessica and Jubin, for real (giving feedback)
- Our brains
- Douglas Myers (aka EEG man)
- Adrian Joseph, for some reason
- Samantha Wootten, for inspiring the project