Decoding Semantics During an Artificial Language Learning Task Using EEG

Chris Foster, Chad C. Williams,
Olave E. Krigolson, and Alona Fyshe

University of Victoria

Can machine learning algorithms be used to determine what a person is thinking about, based solely on the electrical signals from their brain?

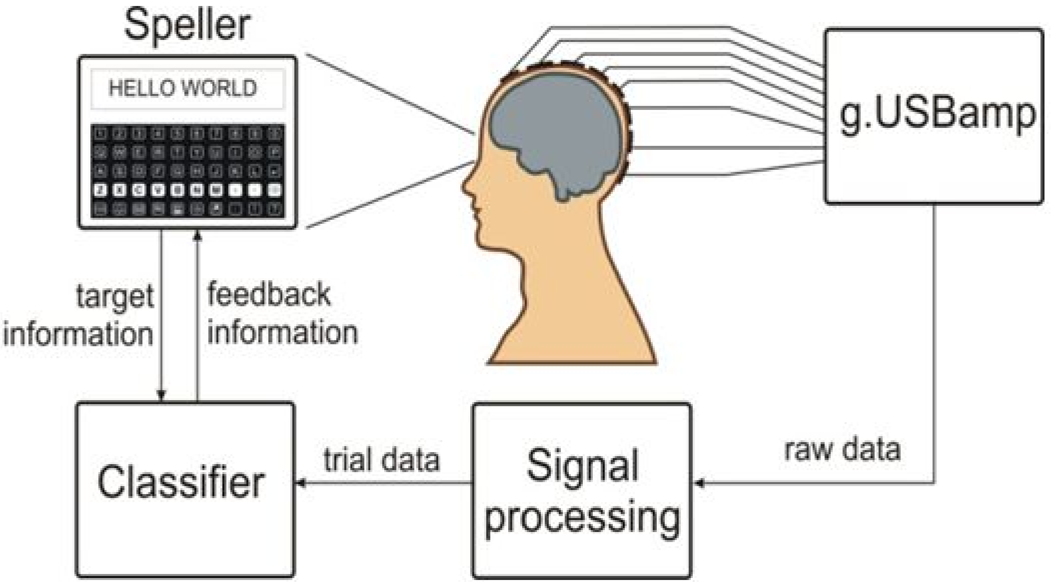
Current applications

Neuroscience

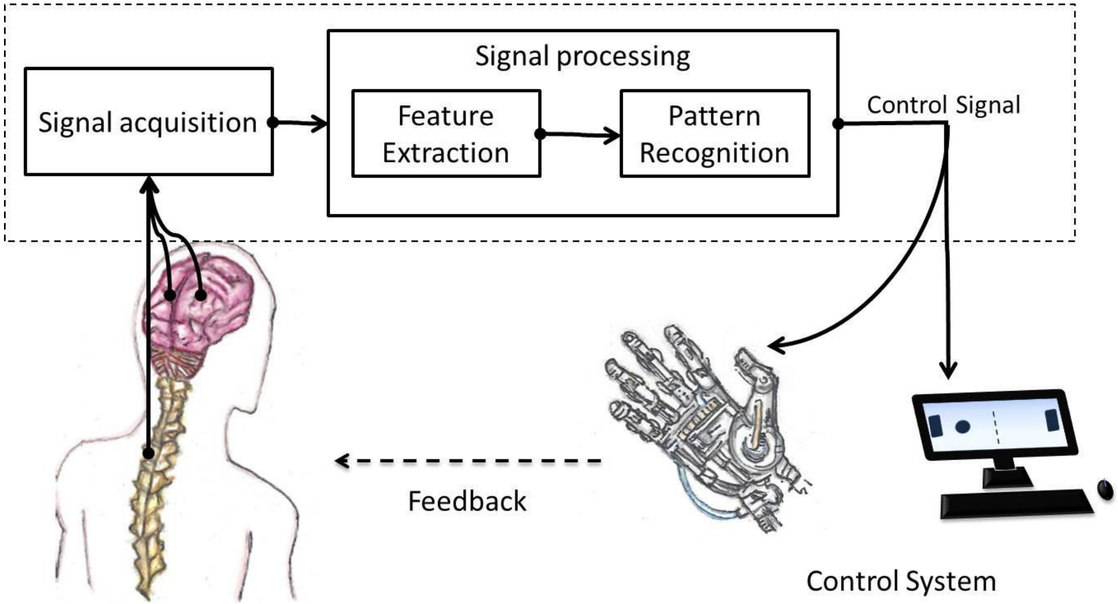
Accessibility

Linguistics

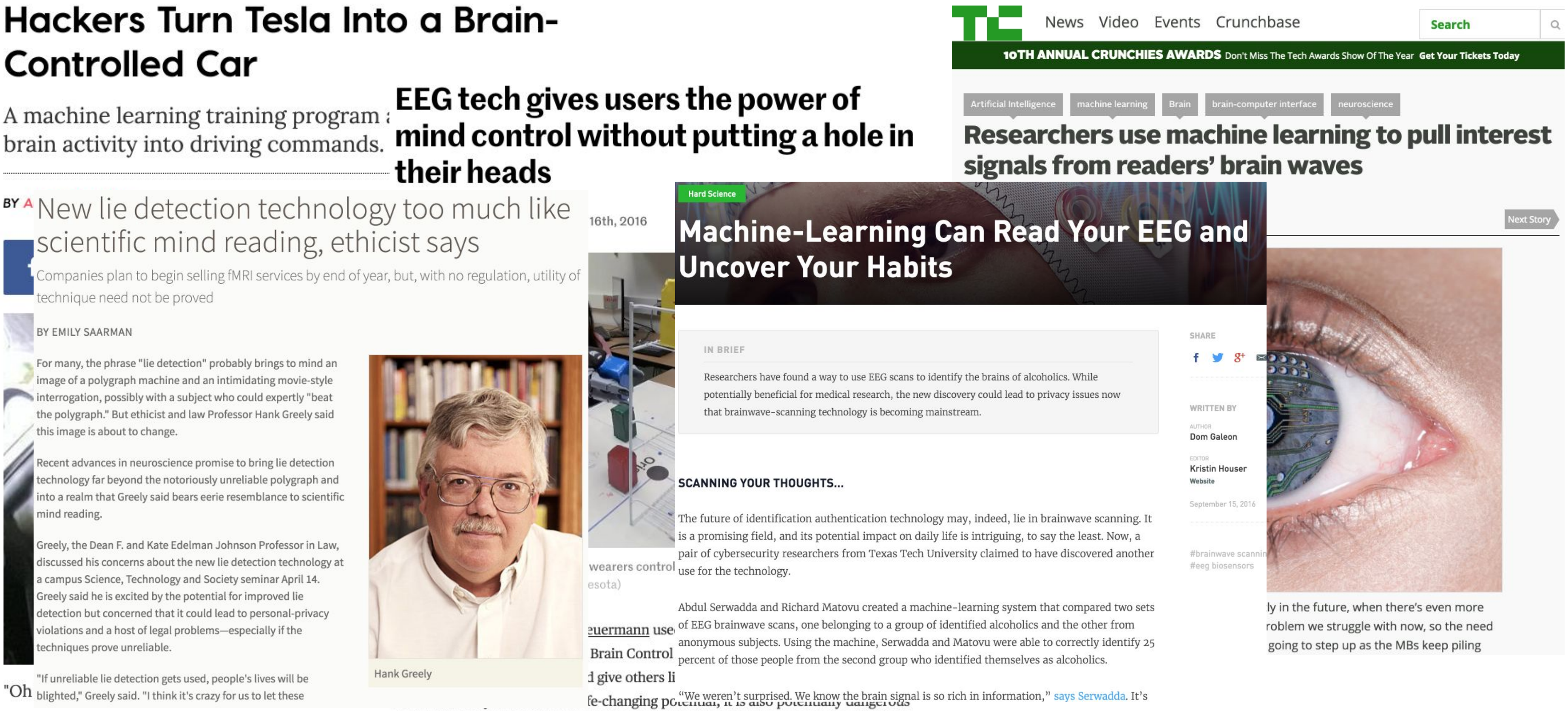
P300 Speller



Brain-computer interfaces



‘In the news’ research



Hackers Turn Tesla Into a Brain-Controlled Car

A machine learning training program a… brain activity into driving commands.

BY A…

EEG tech gives users the power of mind control without putting a hole in their heads

…16th, 2016

…wearers control …esota)

…euermann use… Brain Control …d give others li… …fe-changing potential, it is also potentially dangerous

New lie detection technology too much like scientific mind reading, ethicist says

Companies plan to begin selling fMRI services by end of year, but, with no regulation, utility of technique need not be proved

BY EMILY SAARMAN

For many, the phrase "lie detection" probably brings to mind an image of a polygraph machine and an intimidating movie-style interrogation, possibly with a subject who could expertly "beat the polygraph." But ethicist and law Professor Hank Greely said this image is about to change.

Recent advances in neuroscience promise to bring lie detection technology far beyond the notoriously unreliable polygraph and into a realm that Greely said bears eerie resemblance to scientific mind reading.

Greely, the Dean F. and Kate Edelman Johnson Professor in Law, discussed his concerns about the new lie detection technology at a campus Science, Technology and Society seminar April 14. Greely said he is excited by the potential for improved lie detection but concerned that it could lead to personal-privacy violations and a host of legal problems—especially if the techniques prove unreliable.

"If unreliable lie detection gets used, people's lives will be blighted," Greely said. "I think it's crazy for us to let these

Hank Greely

"Oh…

News Video Events Crunchbase Search

10TH ANNUAL CRUNCHIES AWARDS Don't Miss The Tech Awards Show Of The Year Get Your Tickets Today

Artificial Intelligence machine learning Brain brain-computer interface neuroscience

Researchers use machine learning to pull interest signals from readers’ brain waves

Next Story

…ly in the future, when there’s even more …roblem we struggle with now, so the need … going to step up as the MBs keep piling

Hard Science

Machine-Learning Can Read Your EEG and Uncover Your Habits

IN BRIEF

Researchers have found a way to use EEG scans to identify the brains of alcoholics. While potentially beneficial for medical research, the new discovery could lead to privacy issues now that brainwave-scanning technology is becoming mainstream.

SHARE

WRITTEN BY

AUTHOR Dom Galeon

EDITOR Kristin Houser Website

September 15, 2016

#brainwave scannin… #eeg biosensors

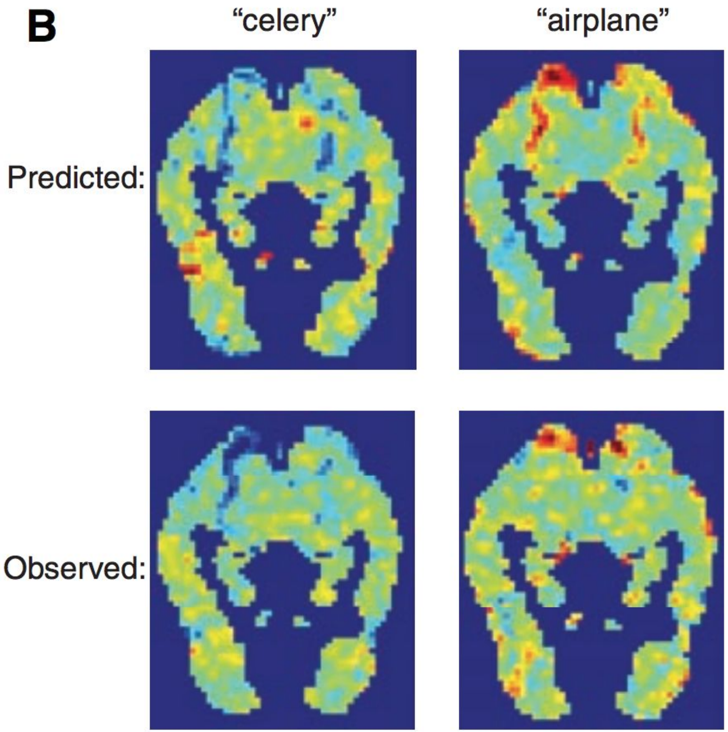
SCANNING YOUR THOUGHTS...

The future of identification authentication technology may, indeed, lie in brainwave scanning. It is a promising field, and its potential impact on daily life is intriguing, to say the least. Now, a pair of cybersecurity researchers from Texas Tech University claimed to have discovered another use for the technology.

Abdul Serwadda and Richard Matovu created a machine-learning system that compared two sets of EEG brainwave scans, one belonging to a group of identified alcoholics and the other from anonymous subjects. Using the machine, Serwadda and Matovu were able to correctly identify 25 percent of those people from the second group who identified themselves as alcoholics.

“We weren’t surprised. We know the brain signal is so rich in information,” says Serwadda. It’s

Predicting fMRI Activity



Mitchell, et al (2008)

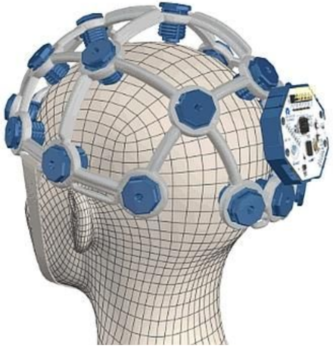
Can we do this with EEG?

EEG is great due to its low cost and the ability to collect data anywhere

EEG also has the advantage of introducing time resolution, which we can use to explore brain data in different ways than fMRI

But, using EEG can be challenging compared to techniques like fMRI or MEG due to the lower signal-to-noise ratio and poor spatial resolution

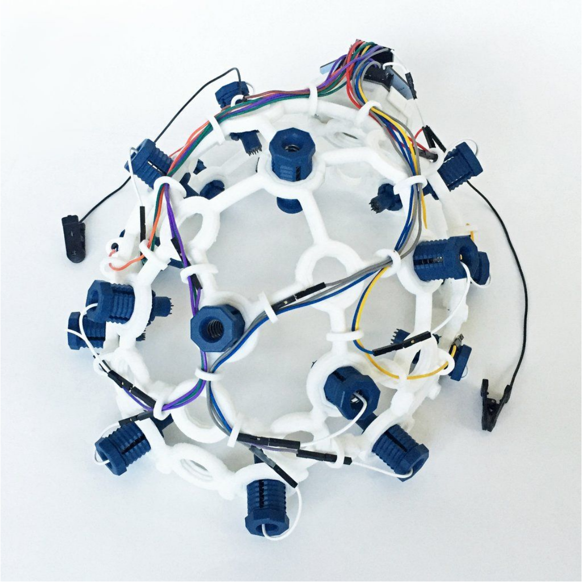
EEG is a tradeoff!



Accurate
Detailed
High Quality

Portable
Affordable
Simple

OpenBCI Headsets



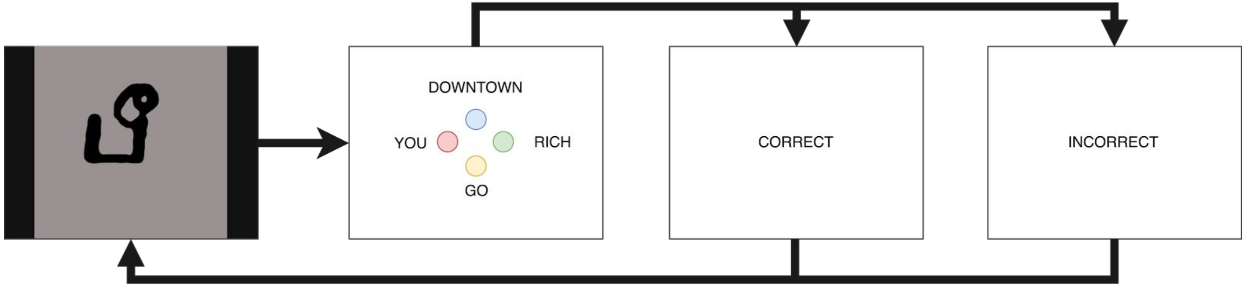
- Cost effective
- Self contained
- Wireless
- Simple to setup
- Easy to collect

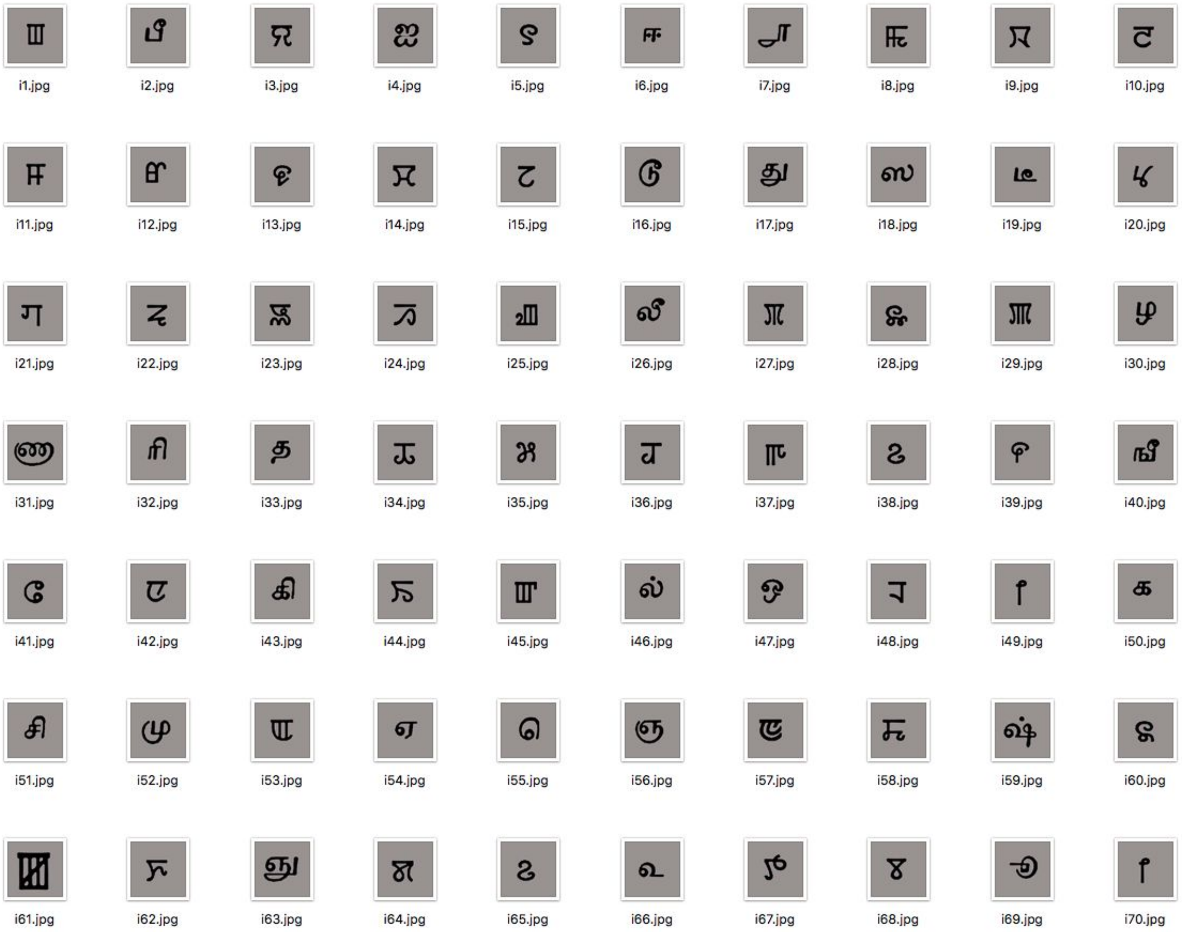
Experiment Paradigm

In this experiment 24 subjects learn to translate 60 words

We discard words that were presented less than six times to a subject, and also discard the first two exposures of each word for each subject

Finally we average the remaining exposures together across all subjects for each word, which gives us a single noise-reduced exposure for every word





i1.jpg
i2.jpg
i3.jpg
i4.jpg
i5.jpg
i6.jpg
i7.jpg
i8.jpg
i9.jpg
i10.jpg
i11.jpg
i12.jpg
i13.jpg
i14.jpg
i15.jpg
i16.jpg
i17.jpg
i18.jpg
i19.jpg
i20.jpg
i21.jpg
i22.jpg
i23.jpg
i24.jpg
i25.jpg
i26.jpg
i27.jpg
i28.jpg
i29.jpg
i30.jpg
i31.jpg
i32.jpg
i33.jpg
i34.jpg
i35.jpg
i36.jpg
i37.jpg
i38.jpg
i39.jpg
i40.jpg
i41.jpg
i42.jpg
i43.jpg
i44.jpg
i45.jpg
i46.jpg
i47.jpg
i48.jpg
i49.jpg
i50.jpg
i51.jpg
i52.jpg
i53.jpg
i54.jpg
i55.jpg
i56.jpg
i57.jpg
i58.jpg
i59.jpg
i60.jpg
i61.jpg
i62.jpg
i63.jpg
i64.jpg
i65.jpg
i66.jpg
i67.jpg
i68.jpg
i69.jpg
i70.jpg

Ridge Regression

Simple linear regression model with regularization via weight decay

Basic linear regression: $$y = w_0x_0 + w_1x_1 + ... + w_mx_m = \sum_{j=0}^{m} = \mathbf{w}^T\mathbf{x}$$

Trained via gradient descent using a loss function: $$J(\mathbf{w}) = \frac{1}{2}\sum_i(\text{target}^{(i)} - \text{output}^{(i)})^2$$

Regression with EEG

The features extracted from the EEG data are a simple concatenation of the raw values in microvolts at each timestamp

With a sampling rate of 250Hz, exposure length of 700ms, and 61 usable sensors we have a total of 10,675 features for each regressor

What are we going to predict?

Word Vectors

The word “cat” may mean something different to different people

The concept of “fuzziness” is very universal and generalizable

We have a dictionary of many words augmented with many different word features on a scale of 1 to 5 and we call these “word vectors”

These are generated by Amazon’s Mechanical Turk service, in which human operators evaluate the attributes of the word vector

Is it a kitchen item?

Can you hold it?

Is it large?

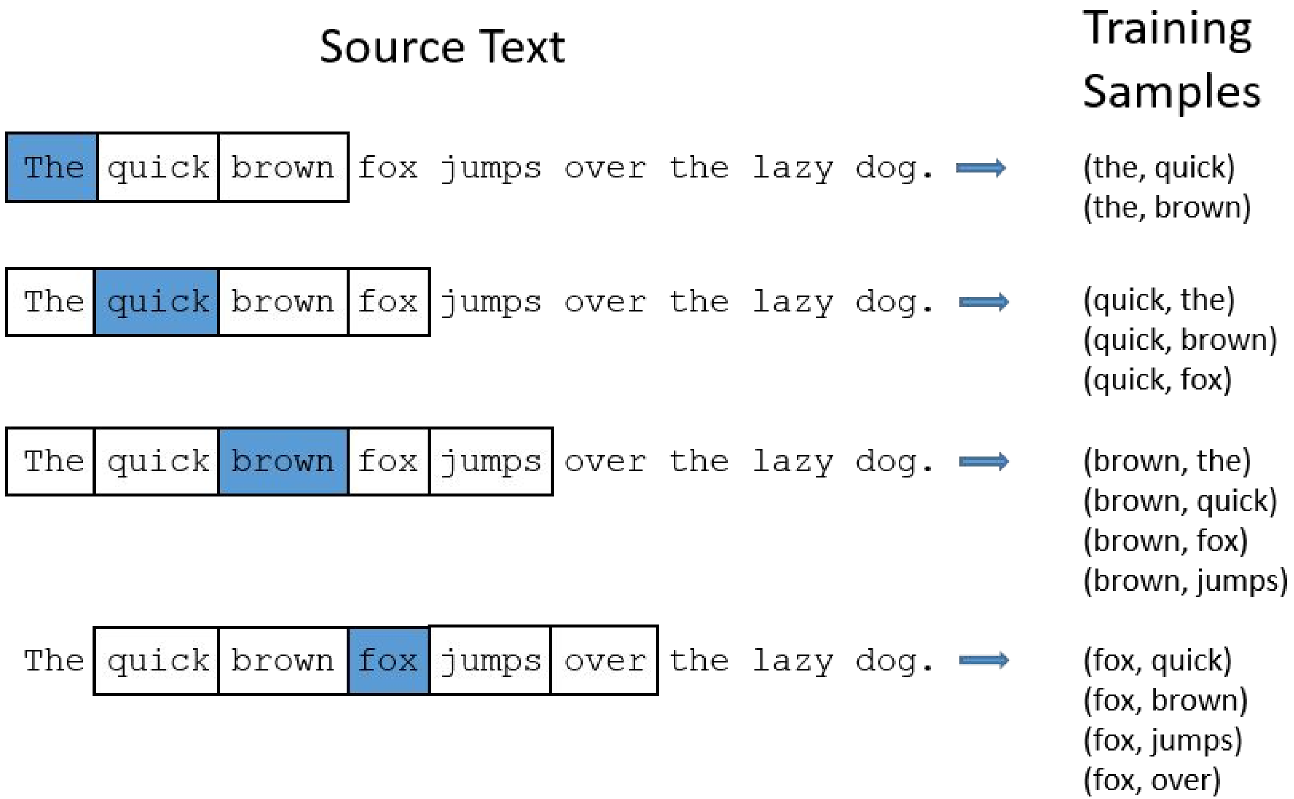
Would you eat it?

Corpus-based Word Vectors

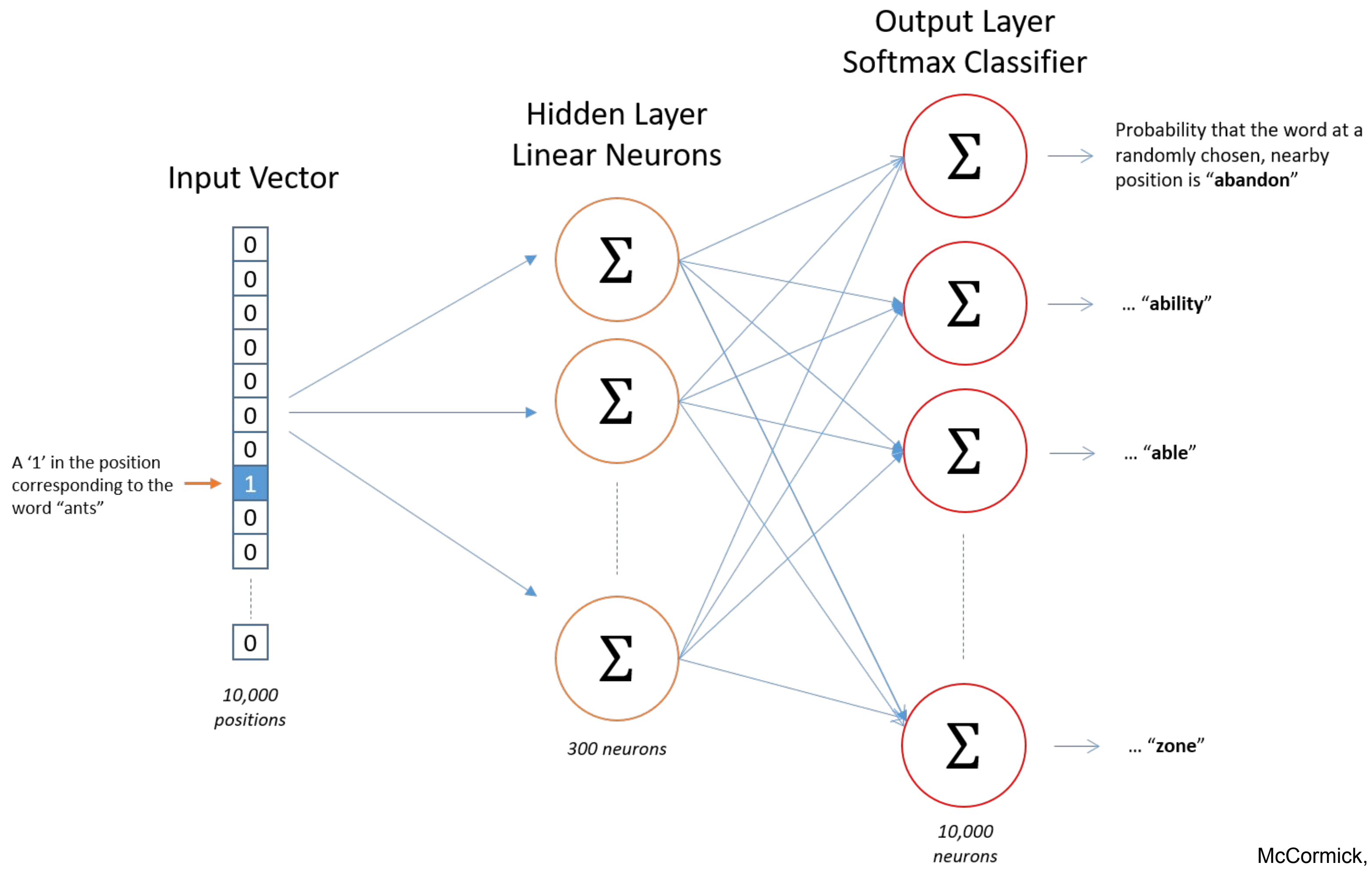
We use the Skipgram word vector set in this research

This 300-element word vector set is generated by training a neural network on a large text corpus to predict nearby words

The location of these words in high dimensional vector space is a representation of their use and semantics in language, which makes them a good target for learning semantics in the brain



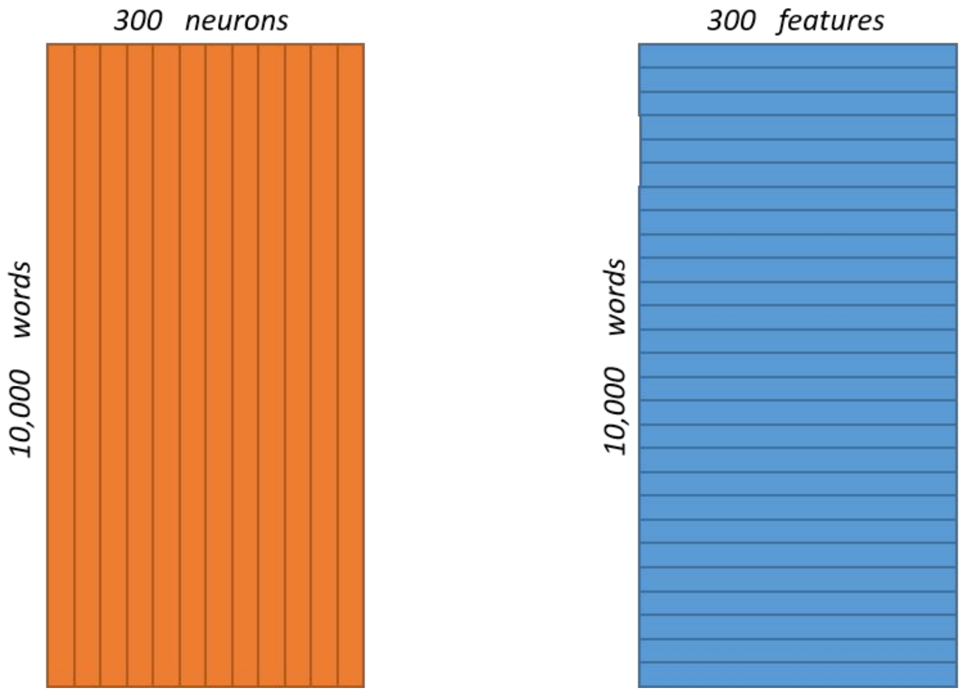
McCormick, C. (2016)



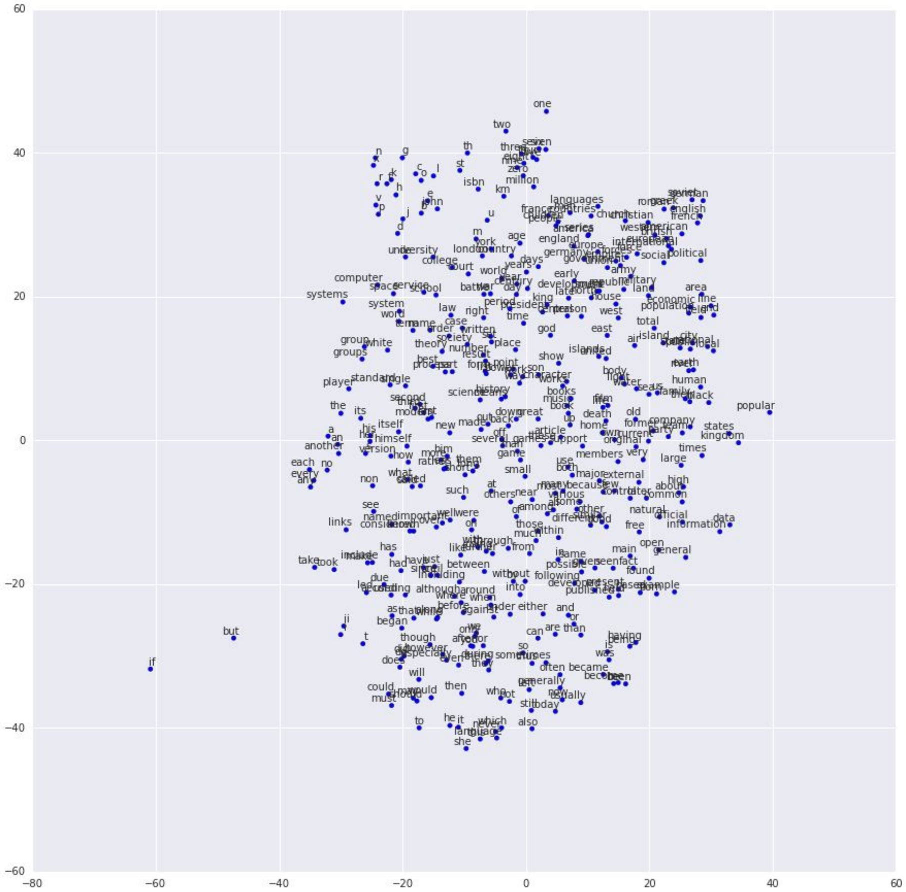
McCormick, C. (2016)

Hidden Layer Weight Matrix

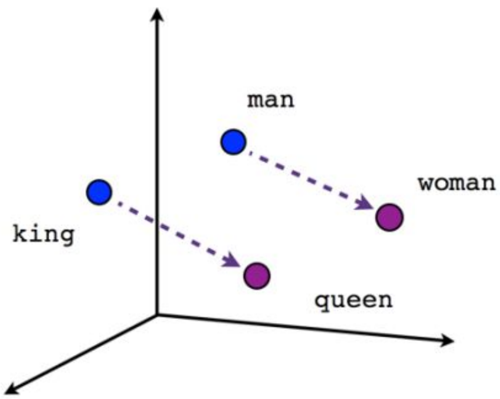
Word Vector Lookup Table!



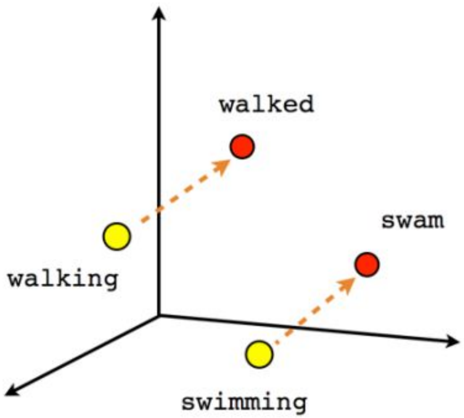
McCormick, C. (2016)



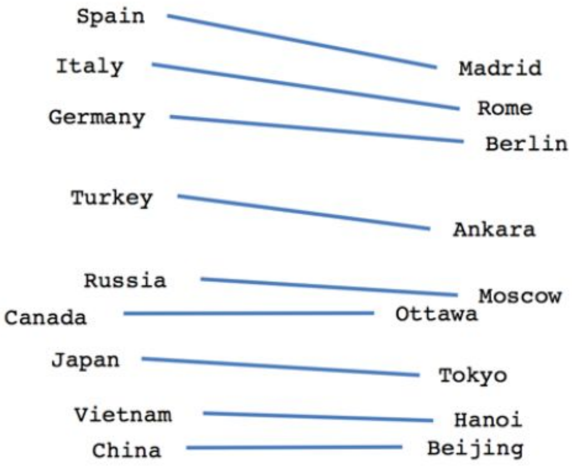
tensorflow.org



Male-Female



Verb tense



Country-Capital

tensorflow.org

Multi-regressor model

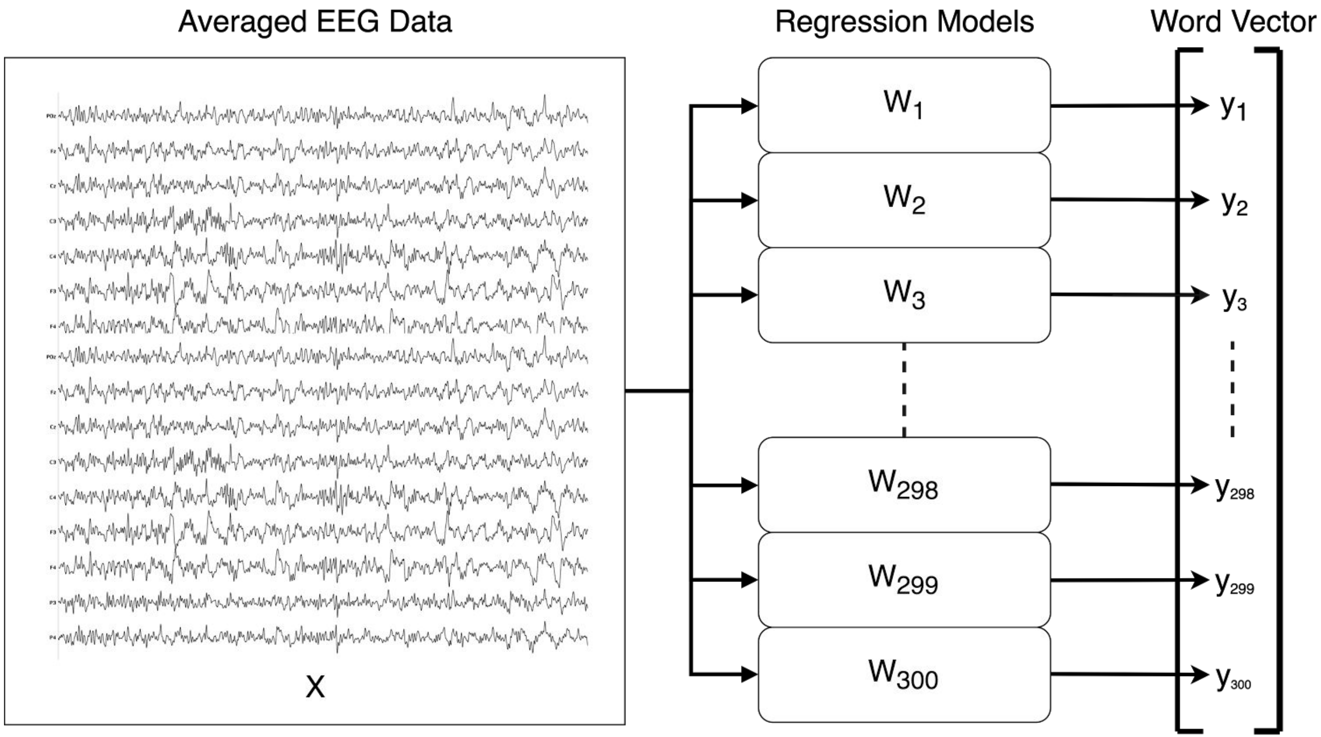
We train a series of simple machine learning ridge regressors such that there exists one for each element of the word vector

Each regressor receives all timesteps of the full 61 channels of averaged EEG data, but only predicts a single index of the vector

For 300 regressors with 10,675 features, we have 3.2 million weights per model

Collectively these regressors form a model that allow us to predict a word vector from any given EEG activity

Multi-regressor model



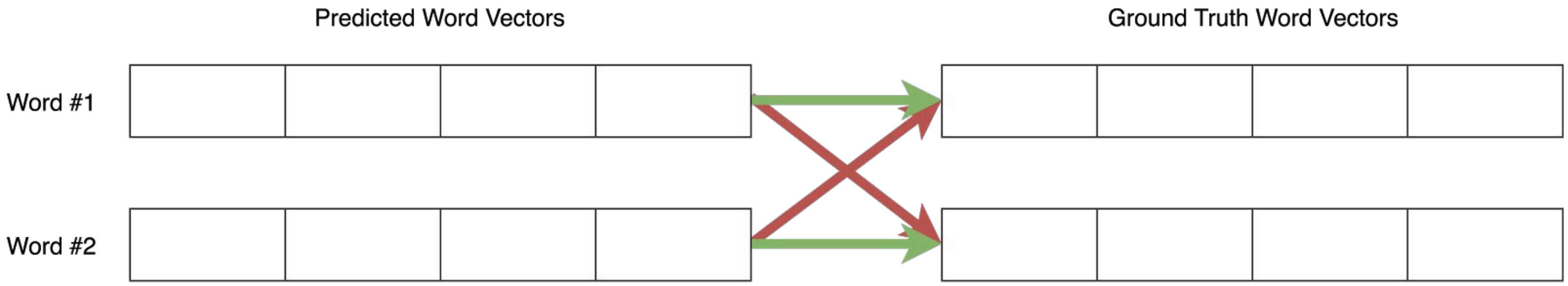
2 vs 2 accuracy

Null hypothesis: the EEG word data and word vectors are not correlated with each other in any way

To disprove the null hypothesis we use the 2 vs 2 test, which evaluates the model in a “leave two out” fashion

The trained models predict the two target word vectors and are successful if the distance is smaller between the matched pairs of predicted / ground truth vectors than the mismatched pairs

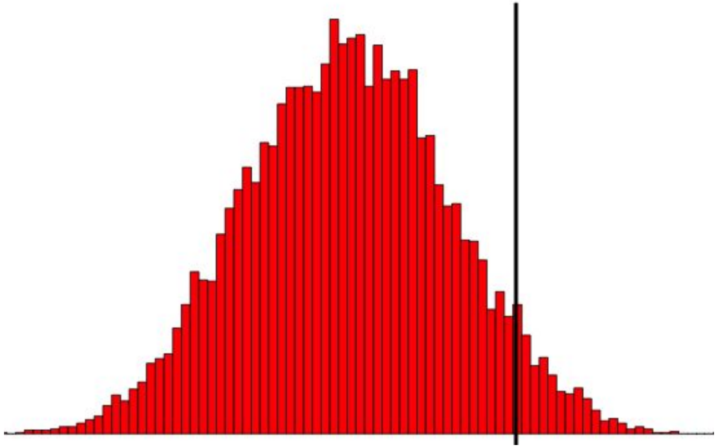
If the null hypothesis is correct, our 2 vs 2 accuracy would be in a range near the chance value of 50%


Predicted Word Vectors
Ground Truth Word Vectors
Word #1
Word #2

Measuring significance

Significant is measured with a permutation test

1. Randomly permute the word vectors so that they no longer pair to the correct EEG
2. Run the entire model on the permuted data
3. Repeat 1,000 times



This creates a null distribution we can use

Our machine learning model shows an accuracy of **75.51%** in the 2 vs 2 test when trained on this data which is significant with $p < 0.01$

This provides evidence that EEG activity is correlated with the representation of word semantics in the brain

Measuring learning

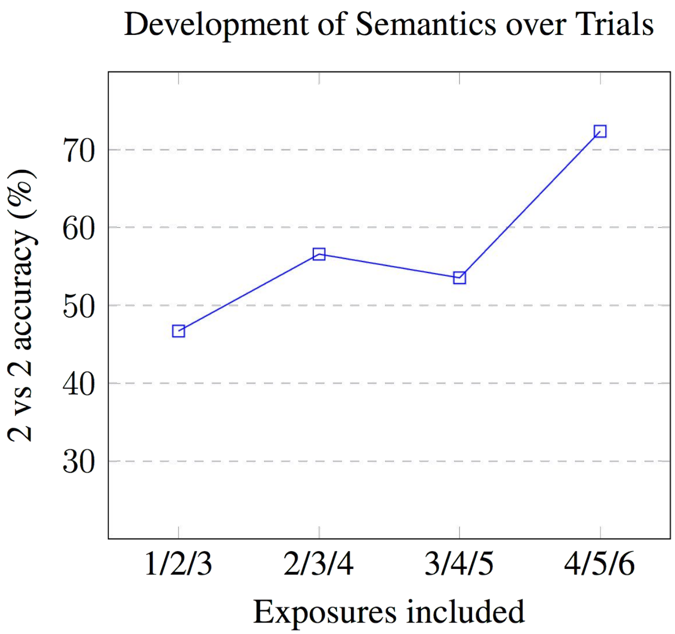
As before, we only consider the subject-word pairs that have six or more exposures (subjects saw between 0–20 exposures of each word)

We demonstrate learning by comparing the 2 vs 2 accuracy of four averaged overlapping subsets of three exposures each

The first averaged three have a non-statistically significant accuracy while the last block is a successful detection

We would expect the last to be slightly lower than the original 2 vs 2 accuracy as it does not include as many exposures

This provides evidence of subjects learning, which is supported by their improvements in accuracy on the task



Conclusion

We show that semantics can be detected via EEG, and further that we can detect learning of semantic concepts as they develop a language mapping in the brain

This opens new avenues for studying language semantics and learning via EEG

Thanks!

CIFAR University of Victoria NSERC CRSNG