Functional connectivity magnetic resonance imaging classification of autism spectrum disorder using the multisite ABIDE dataset

Xin Yang

Middle Tennessee State University

Southern Arkansas University

Jan 31st, 2020

Agenda

1. Introduction
2. ABIDE Datasets
3. Experiments
4. Results
5. Discussion

Agenda

1. Introduction
2. ABIDE Datasets
3. Experiments
4. Results
5. Discussion

Autism Spectrum Disorder (ASD)

- Autism spectrum disorder (ASD) is a brain disorder that is characterized by social and communication impairments as well as restricted interests and repetitive behaviors.
- According to the Centers for Disease Control, autism affects an estimated 1 in 59 children in the United States today.

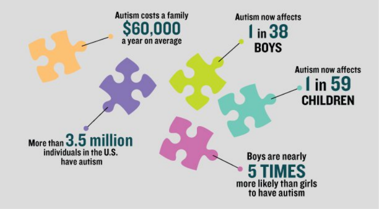


Figure: Early intervention can change a life.

Autism Spectrum Disorder (ASD)

- Indicators of autism usually appear by age 2 or 3. Some associated development delays can appear even earlier, and often, it can be diagnosed as early as 18 months.
- Research shows that early intervention leads to positive outcomes later in life for people with autism.

Figure: World Autism Awareness Day is on 2 April every year

What cause Autism?

- Research suggests that autism develops from a combination of genetic and nongenetic, or environmental, influences.

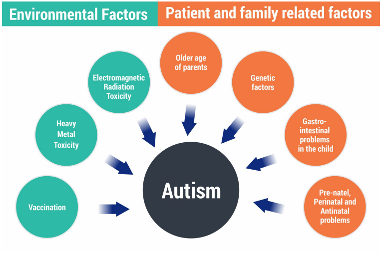


Figure: These mentioned are a few probable causes of autism

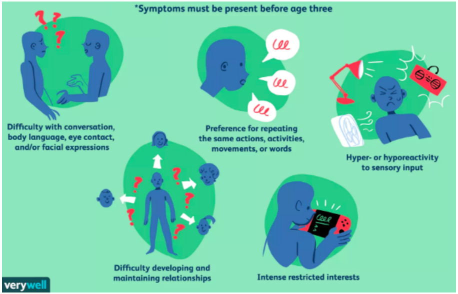
- Although ASD has been identified since the early 1960s, its exact cause is still unknown.

Main goals of Autism Research

- Seek the causes and types of autism.
- Lower the average age of diagnosis to under 2 years.
- Enhance medical treatment of autism's health conditions.

Symptom-based Diagnosis of Autism

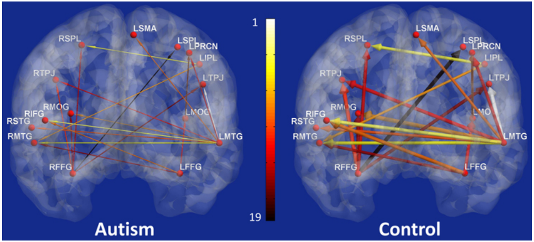
- Generally, the symptom-based diagnosis of ASD requires a very significant amount of time behavioral assessments under the guidance of a highly experienced multidisciplinary team.



- However, symptom-based diagnosis often results in poor treatment due to lack of knowledge of neuropathology.

Data-driven Diagnosis of Autism

- In the past few years, an increasing number of neuroscience research studies have used machine learning to implement neuroimaging data-driven diagnosis of ASD, which would lead to more effective treatment outcomes.



Neuroimaging Techniques

- Anatomical Technique: track the normal and abnormal development of the brain of the patients.

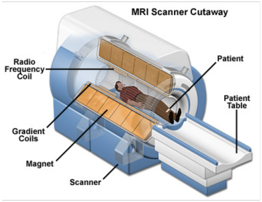
- Functional Technique: track the brain activity.

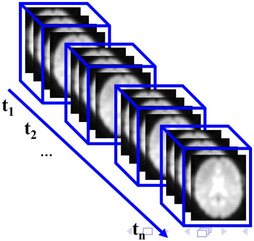
Magnetic resonance imaging (MRI)

- MRI provides good contrast between the different soft tissues of the body based on the facts that the human body is largely composed of water molecules.

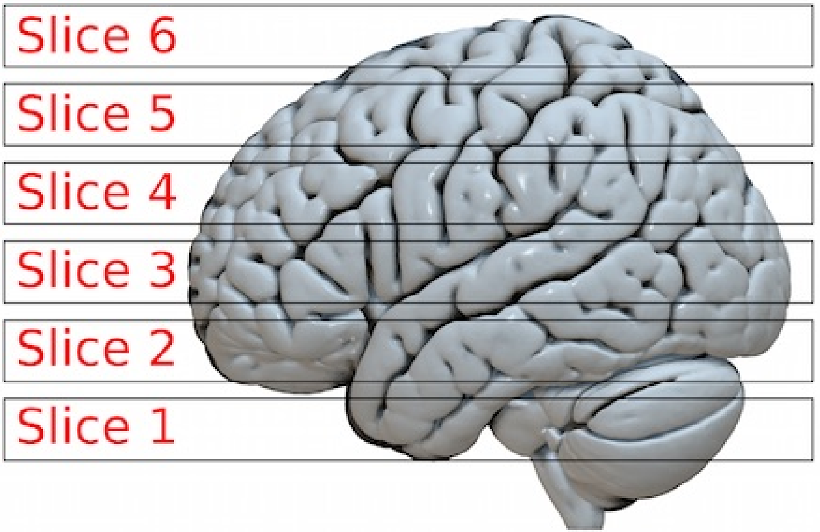
- Most of the disease tissue are tend to have higher water molecules than normal tissue.

- MRI is a 3D data cube, which is usually composed of a lot of MRI slices. Each MRI slice is a 2D image.

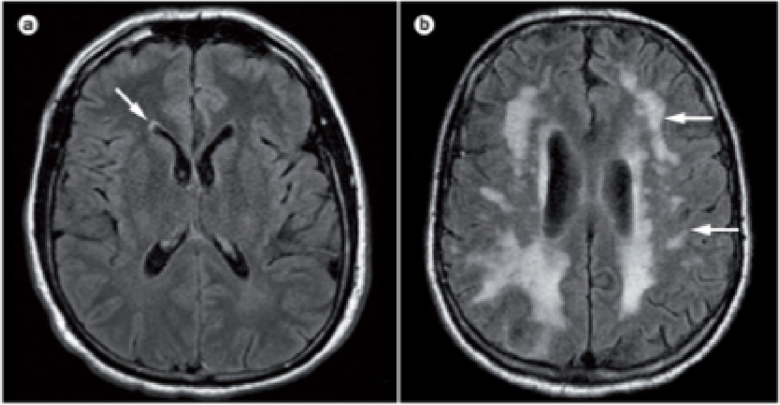




Magnetic resonance imaging (MRI)



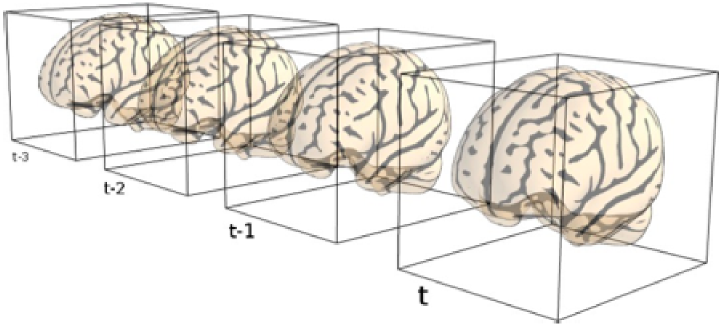
Magnetic resonance imaging (MRI)



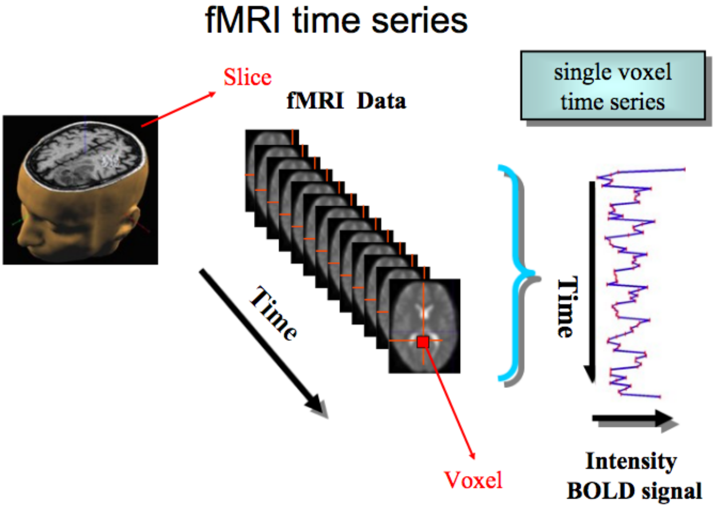
Nature Reviews | Neurology

functional Magnetic resonance imaging (fMRI)

- During the course of an fMRI experiment, a series of brain images are acquired while the subject performs a set of tasks.



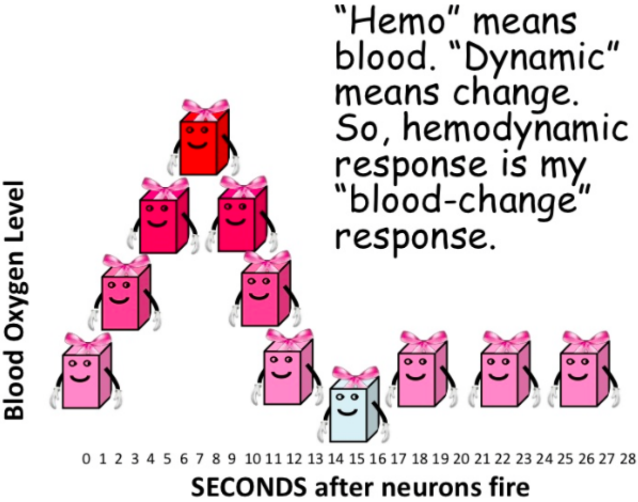
functional Magnetic resonance imaging (fMRI)



functional Magnetic resonance imaging (fMRI)

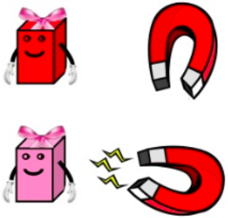
- The general purpose of fMRI data studies is to analyze the brain voxel's time series data to detect whether the Blood Oxygen Level Dependent (BOLD) signal changes in response to a particular stimulus and hence to infer neuronal activity of the human brain.

functional Magnetic resonance imaging (fMRI)

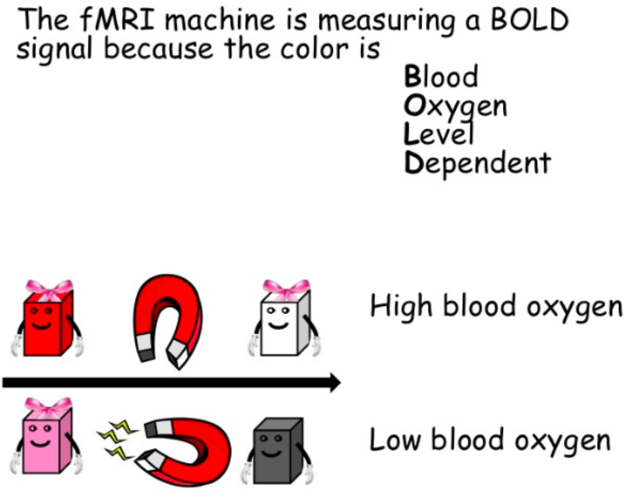


functional Magnetic resonance imaging (fMRI)

The fMRI machine can see my color change because blood with a lot of oxygen (red) is less attracted to magnets than blood without much oxygen (pink or blue).



functional Magnetic resonance imaging (fMRI)



Resting state fMRI

- Resting state fMRI (rs-fMRI) is a method of functional magnetic resonance imaging (fMRI) that is used in brain mapping to evaluate regional interactions that occur in a resting state.
- Because brain activity is intrinsic, present even in the absence of an externally prompted task, any brain region will have spontaneous fluctuations in BOLD signal.
- The resting state approach is useful to explore the brain's functional organization and to examine if it is altered in neurological or mental disorders.

The goal of this study

- The goal of this study is to apply machine learning algorithms to classify autism spectrum disorder (ASD) patients and typically developing (TD) participants using resting-state functional MRI (rs-fMRI) data from a large multisite data repository ABIDE (Autism Brain Imaging Data Exchange)

Agenda

Datasets

ABIDE Datasets

- The preprocessed rs-fMRI data with ASD and TD are downloaded from a large multisite data repository ABIDE (Autism Brain Imaging Data Exchange).
- ABIDE is a multisite platform that has aggregated functional and structural brain imaging data collected from 17 different laboratories around the world.

ABIDE

Autism Brain Imaging
Data Exchange

ABIDE data Phenotypical Information

ABIDE Datasets

- In total, there are 1112 subjects consist of structural and preprocessed resting state functional MRI data along with phenotypic information.
- From these 1112 subjects, 1035 subjects are screened for our study since only 1035 subjects have been given the corresponding completed phenotypic information.
- In these 1035 subjects, there are 505 ASD and 530 TD, 157 females and 878 males.

Data Summary

TABLE I
ABIDE DATA PHENOTYPICAL INFORMATION

site	*ASD*	*TD*	*M*	*F*	*Age Range*
CALTECH	19	18	29	8	17~56
CMU	14	13	21	6	19~40
KKI	20	28	36	12	8~13
LEUVEN	29	34	55	8	12~32
MAX_MUN	24	28	48	4	7~58
NYU	75	100	139	36	6~39
OHSU	12	14	26	0	8~15
OLIN	19	15	29	5	10~24
PITT	29	27	48	8	9~35
SBL	15	15	30	0	20~64
SDSU	14	22	29	7	9~17
STANFORD	19	20	31	8	8~13
TRINITY	22	25	47	0	12~26
UCLA	54	44	86	12	8~18
UM	66	74	113	27	8~29
USM	46	25	71	0	9~50
YALE	28	28	40	16	7~18
TOTAL	505	530	157	878	6~64

Agenda

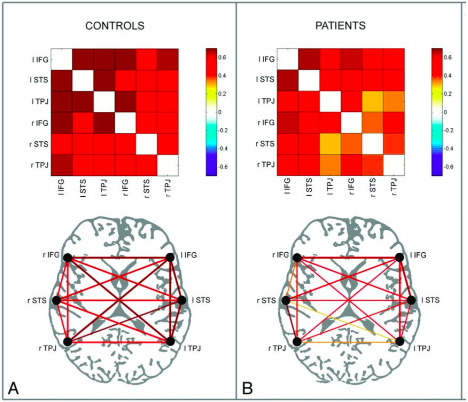
Feature Extraction

Keywords

- resting state functional magnetic resonance imaging (rs-fMRI)
- function connectivity (rsfc-fMRI)
- region of interests (ROIs)
- craddock 400 (cc400, 400 ROIs)

Feature Extraction

- Since the functional connectivity is a manifestation of the co-activation level of the brain regions, in this study, we use functional connectivity of ROIs to classify ASD patients and TD participants.



Brain Atlas

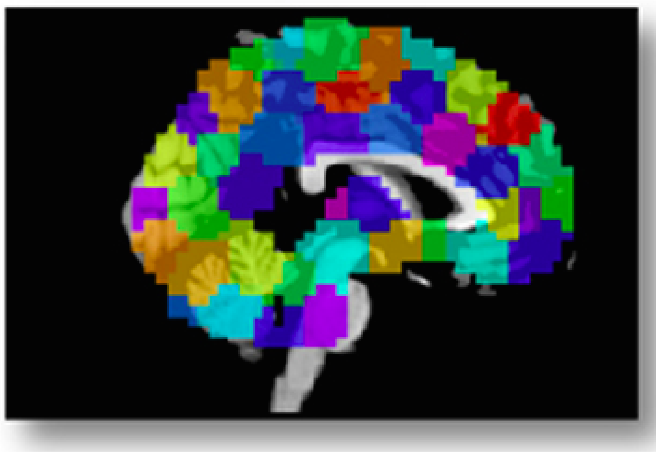
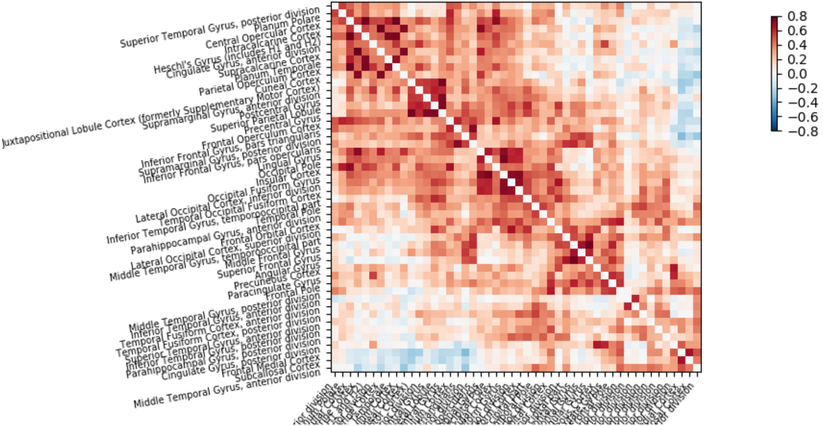
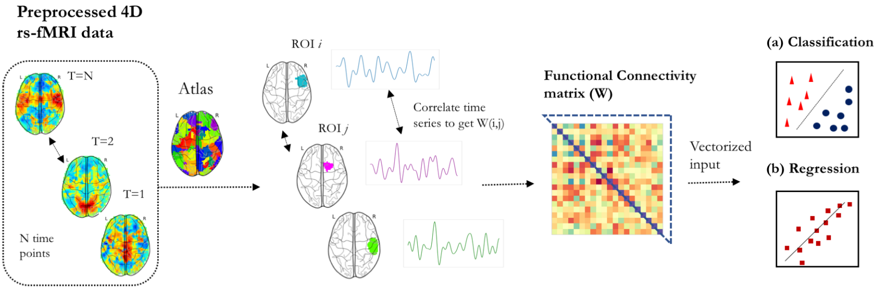


Figure: craddock 400 (cc400, 400 ROIs)

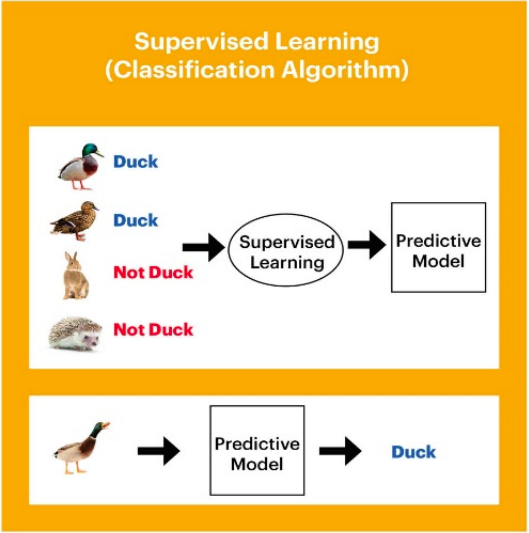
Connectivity Matrix



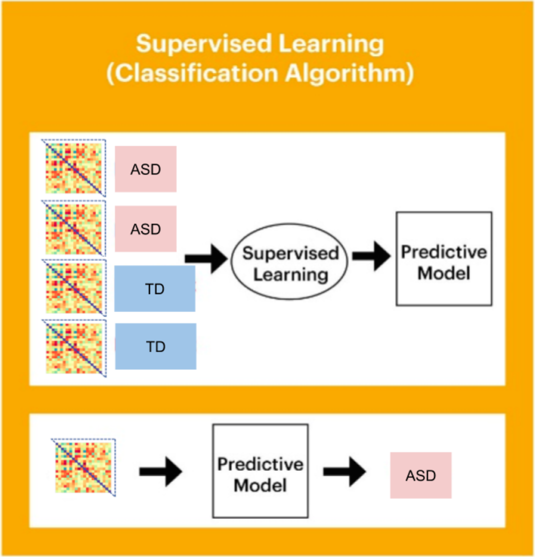
Preprocess and feature extraction



Classification



Classification

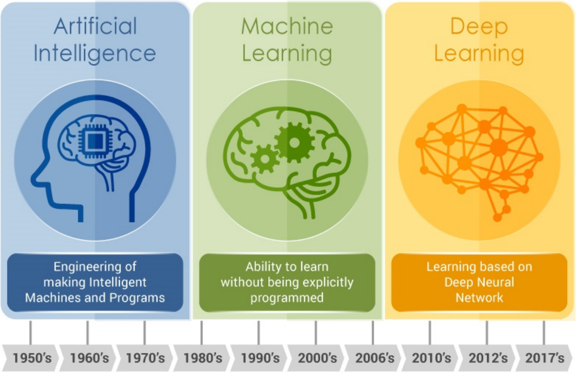


Machine Learning Models

Classifiers

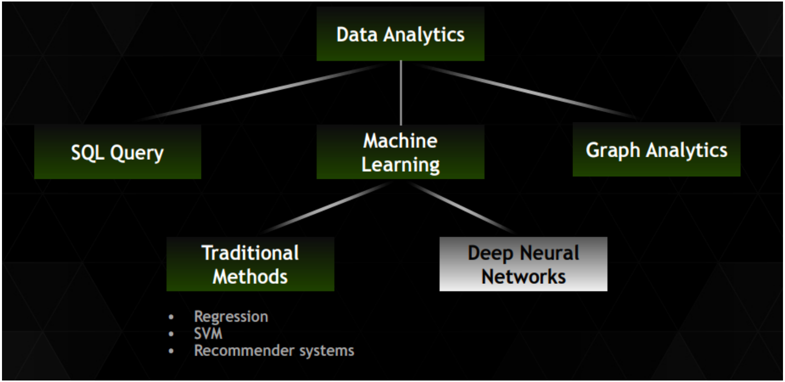
- Ridge
- Logistic Regression
- linear Support Vector Machine
- Kernel Support Vector Machine
- Deep Neural Network

AI v.s. Machine Learning v.s. Deep Learning



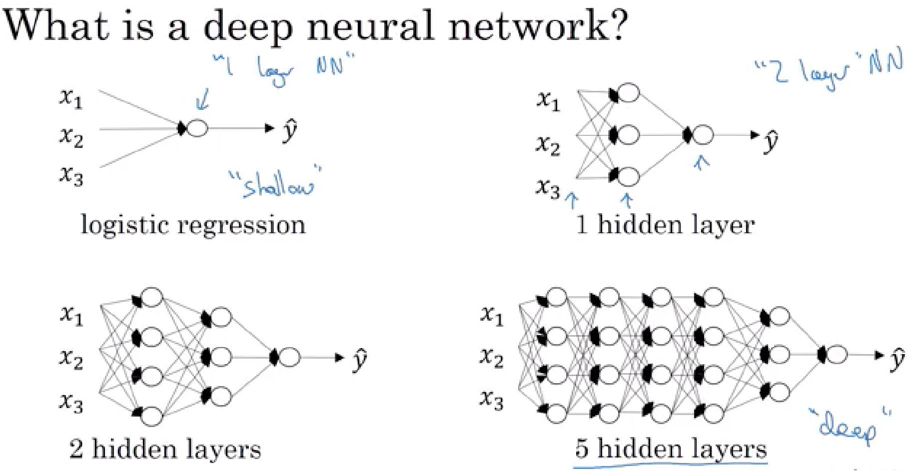
- Figure courtesy of www.embedded-vision.com

Deep Neural Networks



- Slide courtesy of Larry Brown, NVIDIA.com

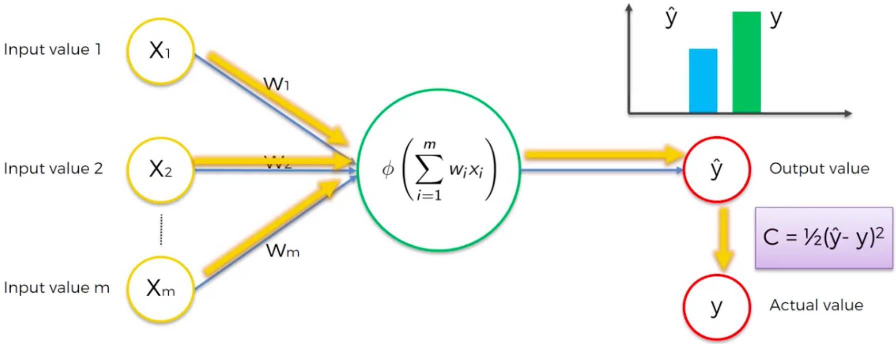
What make Deep Neural Networks deep?



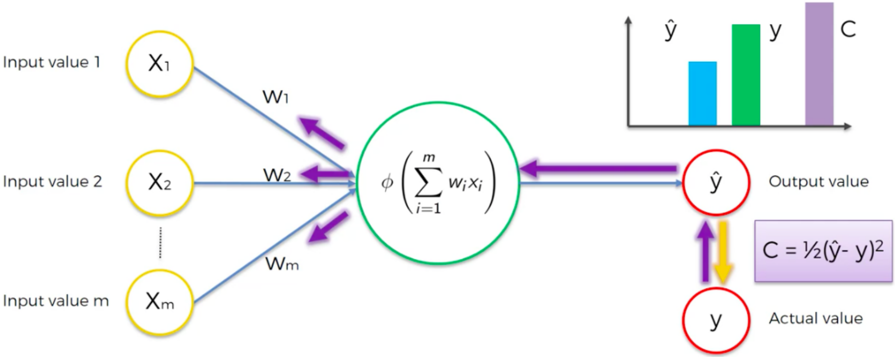
Naming Conventions

- The overall length of the chain gives the depth of the model. The name “deep learning” arose from this terminology.
- You may also hear these networks interchangeably referred to as “Artificial Neural Networks” (ANN).
- Many people do not like the analogies between Neural Networks and real brains and prefer to refer to neurons as units.

How do Neural Networks work ?

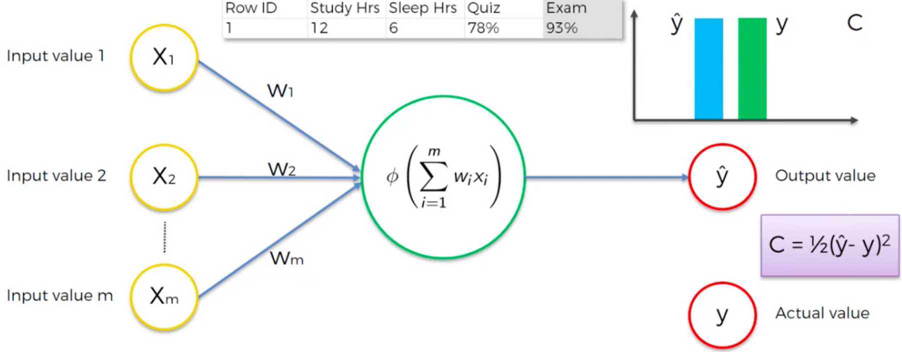


How do Neural Networks learn?

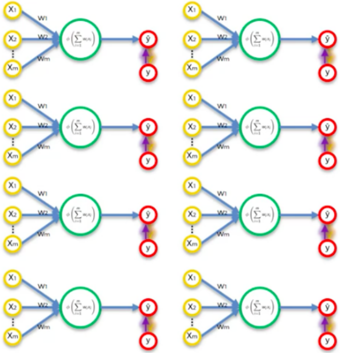


How do Neural Networks work ?

Row ID	Study Hrs	Sleep Hrs	Quiz	Exam
1	12	6	78%	93%



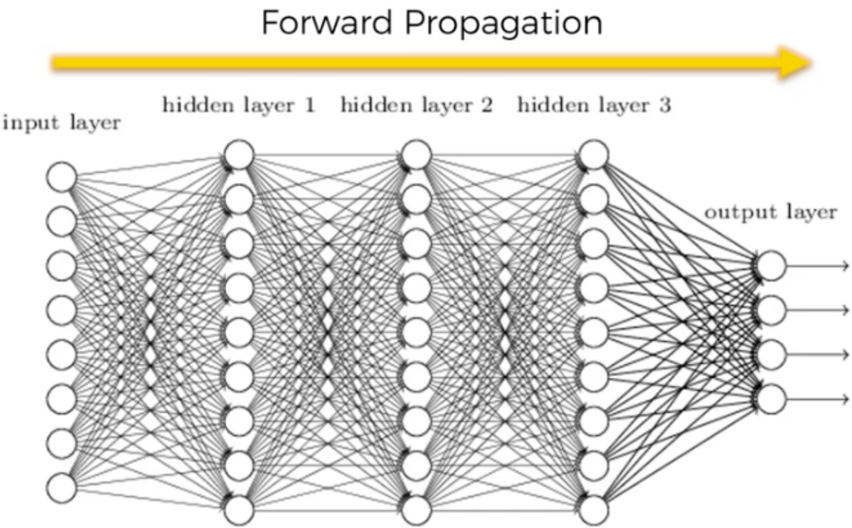
How do Neural Networks learn?



Row ID	Study Hrs	Sleep Hrs	Quiz	Exam
1	12	6	78%	93%
2	22	6.5	24%	68%
3	115	4	100%	95%
4	31	9	67%	75%
5	0	10	58%	51%
6	5	8	78%	60%
7	92	6	82%	89%
8	57	8	91%	97%

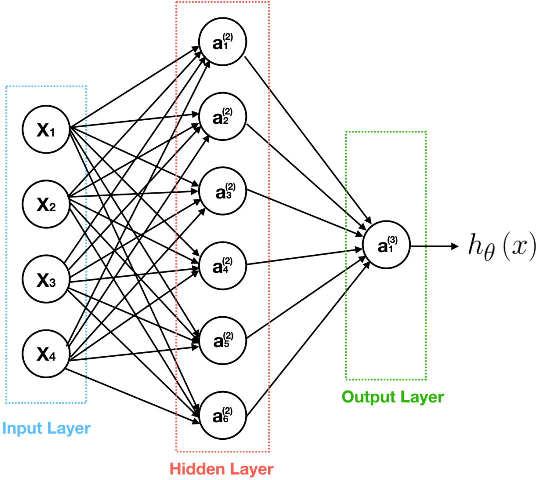
$$C = \sum \tfrac{1}{2}(\hat{y} - y)^2$$

Forward Propagation



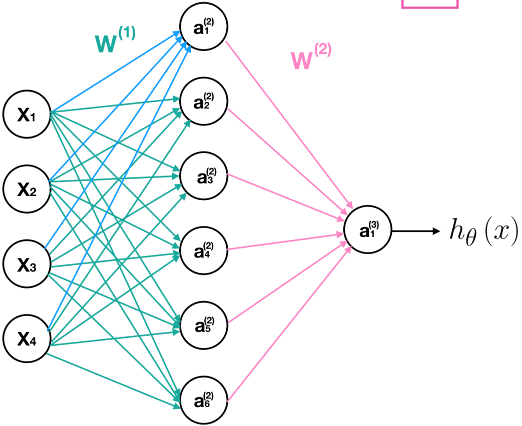
How do Neural Networks work ?

- Input Layer: 4 units, Output Layer: 1 units, Hidden Layer : 6 units



How do Neural Networks work ?

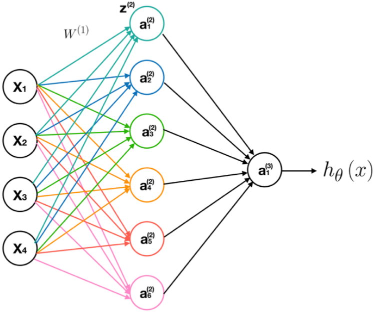
$$W^{(1)} = \begin{bmatrix} \theta_{11}{}^{(1)} & \theta_{21}{}^{(1)} & \theta_{31}{}^{(1)} & \theta_{41}{}^{(1)} & \theta_{51}{}^{(1)} & \theta_{61}{}^{(1)} \\ \theta_{12}{}^{(1)} & \theta_{22}{}^{(1)} & \theta_{32}{}^{(1)} & \theta_{42}{}^{(1)} & \theta_{52}{}^{(1)} & \theta_{62}{}^{(1)} \\ \theta_{13}{}^{(1)} & \theta_{23}{}^{(1)} & \theta_{33}{}^{(1)} & \theta_{43}{}^{(1)} & \theta_{53}{}^{(1)} & \theta_{63}{}^{(1)} \\ \theta_{14}{}^{(1)} & \theta_{24}{}^{(1)} & \theta_{34}{}^{(1)} & \theta_{44}{}^{(1)} & \theta_{54}{}^{(1)} & \theta_{64}{}^{(1)} \end{bmatrix} \qquad W^{(2)} = \boxed{\begin{bmatrix} \theta_{11}{}^{(2)} \\ \theta_{12}{}^{(2)} \\ \theta_{13}{}^{(2)} \\ \theta_{14}{}^{(2)} \\ \theta_{15}{}^{(2)} \\ \theta_{16}{}^{(2)} \end{bmatrix}}$$



- From Input Layer to Hidden Layer

$$W^TX = \begin{bmatrix} \theta_{11}{}^{(1)} & \theta_{12}{}^{(1)} & \theta_{13}{}^{(1)} & \theta_{14}{}^{(1)} \\ \theta_{21}{}^{(1)} & \theta_{22}{}^{(1)} & \theta_{23}{}^{(1)} & \theta_{24}{}^{(1)} \\ \theta_{31}{}^{(1)} & \theta_{32}{}^{(1)} & \theta_{33}{}^{(1)} & \theta_{34}{}^{(1)} \\ \theta_{41}{}^{(1)} & \theta_{42}{}^{(1)} & \theta_{43}{}^{(1)} & \theta_{44}{}^{(1)} \\ \theta_{51}{}^{(1)} & \theta_{52}{}^{(1)} & \theta_{53}{}^{(1)} & \theta_{54}{}^{(1)} \\ \theta_{61}{}^{(1)} & \theta_{62}{}^{(1)} & \theta_{63}{}^{(1)} & \theta_{64}{}^{(1)} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} z_1{}^{(1)} \\ z_2{}^{(1)} \\ z_3{}^{(1)} \\ z_4{}^{(1)} \\ z_5{}^{(1)} \\ z_6{}^{(1)} \end{bmatrix} = Z^{(2)}$$

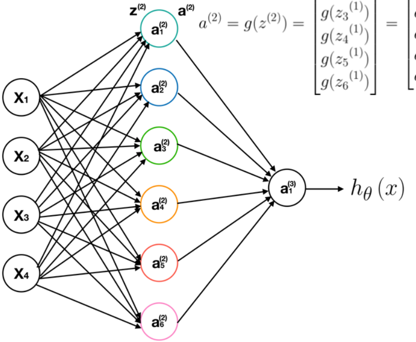
$$z_1{}^{(2)} = \theta_{11}{}^{(1)}x_1 + \theta_{12}{}^{(1)}x_2 + \theta_{13}{}^{(1)}x_3 + \theta_{14}{}^{(1)}x_4$$
$$z_2{}^{(2)} = \theta_{21}{}^{(1)}x_1 + \theta_{22}{}^{(1)}x_2 + \theta_{23}{}^{(1)}x_3 + \theta_{24}{}^{(1)}x_4$$
$$z_3{}^{(2)} = \theta_{31}{}^{(1)}x_1 + \theta_{32}{}^{(1)}x_2 + \theta_{33}{}^{(1)}x_3 + \theta_{34}{}^{(1)}x_4$$
$$z_4{}^{(2)} = \theta_{41}{}^{(1)}x_1 + \theta_{42}{}^{(1)}x_2 + \theta_{43}{}^{(1)}x_3 + \theta_{44}{}^{(1)}x_4$$
$$z_5{}^{(2)} = \theta_{51}{}^{(1)}x_1 + \theta_{52}{}^{(1)}x_2 + \theta_{53}{}^{(1)}x_3 + \theta_{54}{}^{(1)}x_4$$
$$z_6{}^{(2)} = \theta_{61}{}^{(1)}x_1 + \theta_{62}{}^{(1)}x_2 + \theta_{63}{}^{(1)}x_3 + \theta_{64}{}^{(1)}x_4$$



How do Neural Networks work ?

- From Input Layer to Hidden Layer

only calculate value by sigmoid function and pass it

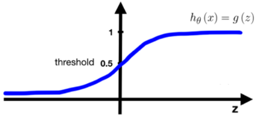


$h_\theta(x)$

$$a^{(2)} = g(z^{(2)}) = \begin{bmatrix} g(z_1{}^{(1)}) \\ g(z_2{}^{(1)}) \\ g(z_3{}^{(1)}) \\ g(z_4{}^{(1)}) \\ g(z_5{}^{(1)}) \\ g(z_6{}^{(1)}) \end{bmatrix} = \begin{bmatrix} a_1{}^{(2)} \\ a_2{}^{(2)} \\ a_3{}^{(2)} \\ a_4{}^{(2)} \\ a_5{}^{(2)} \\ a_6{}^{(2)} \end{bmatrix}$$

Sigmoid (logistic) Activation Function

$$h_\theta(x) = g(-\theta^T x) = \frac{1}{1 + e^{-\theta^T x}}$$

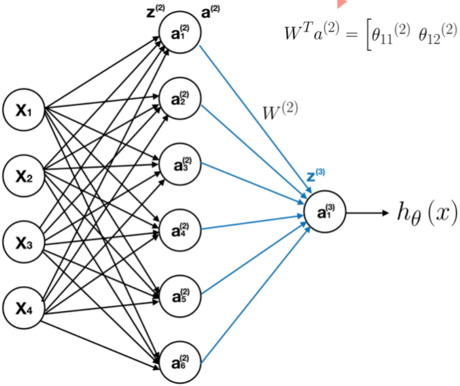


How do Neural Networks work ?

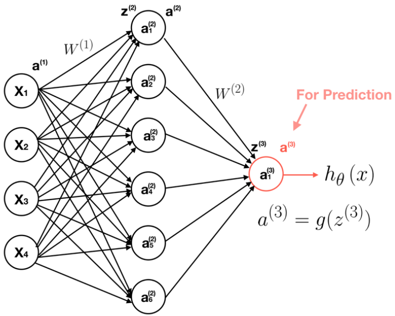
- From Hidden Layer to Output Layer

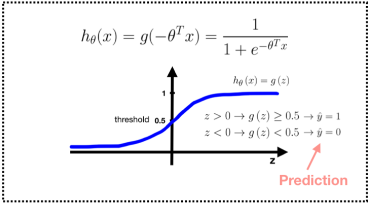
Passing Value

$$W^T a^{(2)} = \begin{bmatrix} \theta_{11}{}^{(2)} & \theta_{12}{}^{(2)} & \theta_{13}{}^{(2)} & \theta_{14}{}^{(2)} & \theta_{15}{}^{(2)} & \theta_{16}{}^{(2)} \end{bmatrix} \begin{bmatrix} a_1{}^{(2)} \\ a_2{}^{(2)} \\ a_3{}^{(2)} \\ a_4{}^{(2)} \\ a_5{}^{(2)} \\ a_6{}^{(2)} \end{bmatrix} = z^{(3)}$$



- From Hidden Layer to Output Layer





- **Loss Function** is also called **Cost Function**, but this name is used more commonly in NN.
- The goal of Loss Function is to measure the error that the model made.

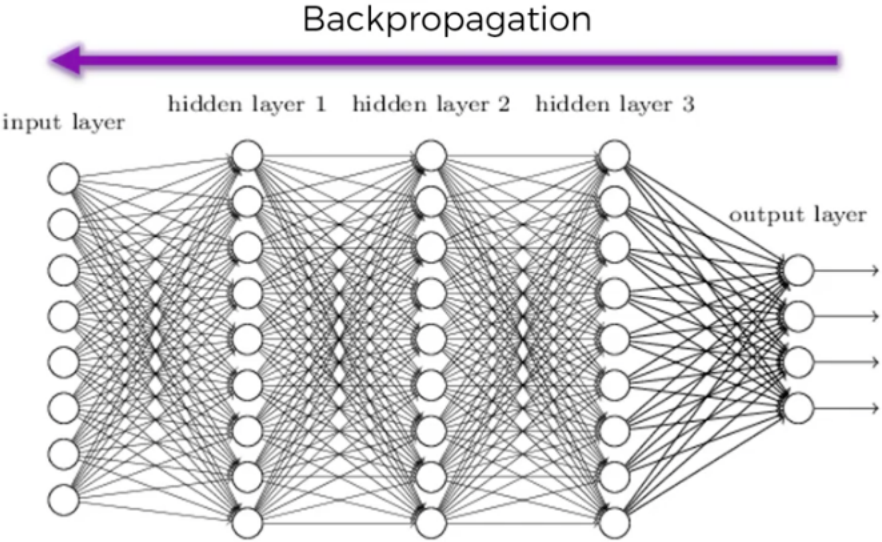
data number

Regularization Term

$$J(\Theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{k=1}^{K} y_k^{(i)} \log(h_\Theta(x^{(i)}))_k + (1-y_k^{(i)})\log(1-(h_\Theta(x^{(i)}))_k)\right] + \boxed{\frac{\lambda}{2m}\sum_{l=1}^{L-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(\Theta_{ji}^{(l)})^2}$$

label number

Back Propagation



- This procedure is for **minimizing Loss Function**, just like the technique – gradient descent.
- There are two steps:
  - Compute the partial derivative of $J(\Theta)$
  - Update each element of weighting matrix Θ

Repeat until converge {

$$\Theta_{ij}{}^{(l)} = \Theta_{ij}{}^{(l)} - \boxed{\alpha} \frac{\partial J(\Theta)}{\partial \Theta_{ij}{}^{(l)}}$$

learning rate

}

Back Propagation

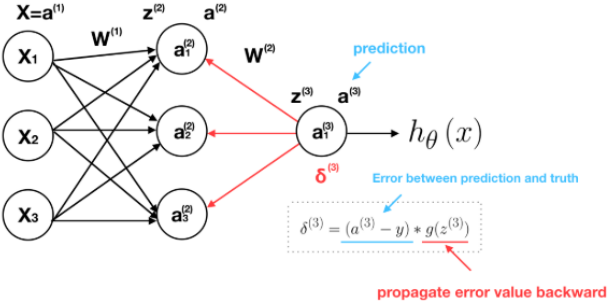
$$\frac{\partial J(\Theta)}{\Theta^{(2)}} = \delta^{(3)} * a^{(2)}$$

$$\frac{\partial J(\Theta)}{\Theta^{(1)}} = \delta^{(2)} * a^{(1)}$$

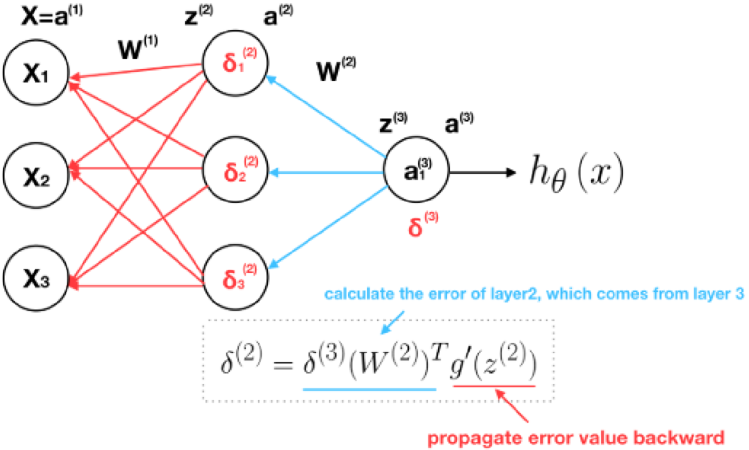
$$\delta^{(3)} = (a^{(3)} - y) * g'(z^{(3)})$$

$$\delta^{(2)} = \delta^{(3)} (W^{(2)})^T g'(z^{(2)})$$

- δ^3 is the error of layer3, the Output Layer in this example.
- δ^2 is the error of layer2, the Hidden Layer in this example.
- δ^1 is not exist since layer1 is the Input Layer.



Back Propagation



Warm Up

STEP 1: Randomly initialise the weights to small numbers close to 0 (but not 0).

⇩

STEP 2: Input the first observation of your dataset in the input layer, each feature in one input node.

⇩

STEP 3: Forward-Propagation: from left to right, the neurons are activated in a way that the impact of each neuron's activation is limited by the weights. Propagate the activations until getting the predicted result y.

⇩

STEP 4: Compare the predicted result to the actual result. Measure the generated error.

⇩

STEP 5: Back-Propagation: from right to left, the error is back-propagated. Update the weights according to how much they are responsible for the error. The learning rate decides by how much we update the weights

⇩

STEP 6: Repeat Steps 1 to 5 and update the weights after each observation (Reinforcement Learning). Or:
Repeat Steps 1 to 5 but update the weights only after a batch of observations (Batch Learning).

⇩

STEP 7: When the whole training set passed through the ANN, that makes an epoch. Redo more epochs.

Deep Learning Frameworks

Deep Learning Frameworks

- Caffe/Caffe2
- Keras
- TensorFlow
- Theano
- Torch

Agenda

Results

Atlas	Classifier	Accuracy	Recall	Precision
CC400	DNN	72.95%	74.76%	**71.96%**
CC400	Ridge	71.98%	70.89%	71.53%
CC400	LR	71.79%	70.69%	71.29%
CC400	linearSVM	71.40%	70.10%	70.93%
CC400	kernelSVM	71.40%	69.90%	71.12%

Table: 5 fold Cross-Validation Results

Deep Neural Network Results

Layer	Accuracy	Recall	Precision
128-64-2	**73.42%**	70.30 %	75.00%
256-128-2	72.95%	74.76%	71.96%
512-256-2	71.98 %	69.39%	70.83%
1024-512-2	72.46%	68.37%	72.53%

Table: 5 fold Cross-Validation Results

Results

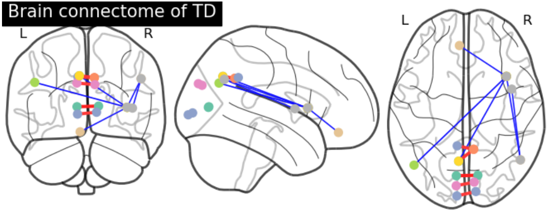


Figure: Top 5 correlated and anti-correlated ROIs of TD

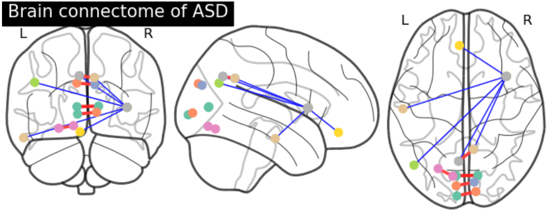


Figure: Top 5 correlated and anti-correlated ROIs of ASD

Results

Table: ROIs Analysis for CC400

Top 5 Correlation Connectivity ROIs	
ASD	Left Intracalcarine Cortex → Right Intracalcarine Cortex
	Left Occipital Pole → Right Occipital Pole
	Left Cuneal Cortex → Right Cuneal Cortex
	Left Precuneous Cortex → Right Precuneous Cortex
	Left Occipital Fusiform Gyrus → Left Lingual Gyrus
TD	Left Intracalcarine Cortex → Right Intracalcarine Cortex
	Left Precuneous Cortex → Right Precuneous Cortex
	Left Cingulate Gyrus → Right Precuneous Cortex
	Left Occipital Pole → Right Occipital Pole
	Left Cuneal Cortex → Right Cuneal Cortex

Results

Table: ROIs Analysis for CC400

Top 5 Anti-Correlation Connectivity ROIs	
ASD	Left Lateral Occipital Cortex → Right Frontal Operculum Cortex
	Right Frontal Medial Cortex → Right Frontal Operculum Cortex
	Left Middle Temporal Gyrus → Right Frontal Operculum Cortex
	Right Precuneous Cortex → Right Frontal Operculum Cortex
	Left Precuneous Corte → Right Frontal Operculum Cortex
TD	Left Lateral Occipital Cortex → Right Frontal Operculum Cortex
	Left Precuneous Cortex → Right Frontal Operculum Cortex
	Right Frontal Medial Cortex → Right Frontal Operculum Cortex
	Right Lateral Occipital Cortex → Right Frontal Operculum Cortex
	Right Lateral Occipital Cortex → Right Central Opercular Cortex

Agenda

- hyperparameter

How to find the best hyper-parameters?

Future Work

- more data samples

Can we get more data samples?

Future Work

- clinical application

Can fMRI research apply to the clinic?

Thank you! :)

Questions?

Functional connectivity magnetic resonance imaging classification of autism spectrum disorder using the multisite ABIDE dataset

Xin Yang

Middle Tennessee State University

Southern Arkansas University

Jan 31st, 2020