CLASSIFICATION OF AUTISM SPECTRUM DISORDER USING MACHINE LEARNING

Andrew Bender

Mentors: Dr. Mark Wallace and David Tovar

Thesis committee: Dr. Mark Wallace, Dr. Ramnarayan Ramachandran, Dr. Randolph Blake, and Dr. Eli Zimmerman

April 18, 2019

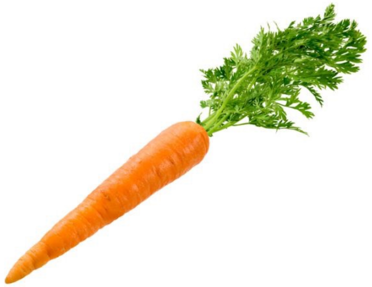
OUTLINE

- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

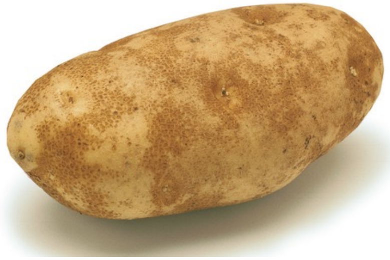
OUTLINE

- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

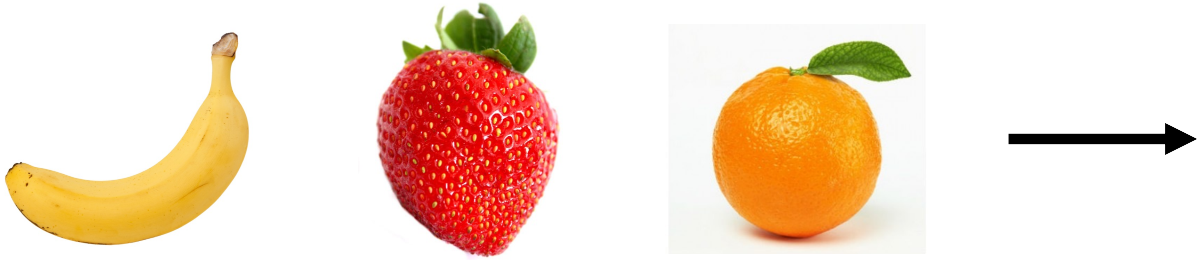
CLASSIFICATION





CLASSIFICATION



CLASSIFICATION

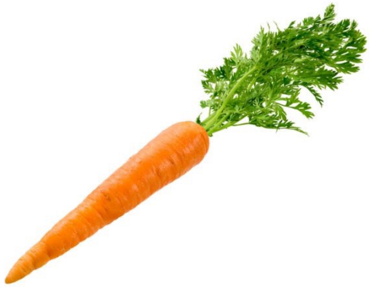
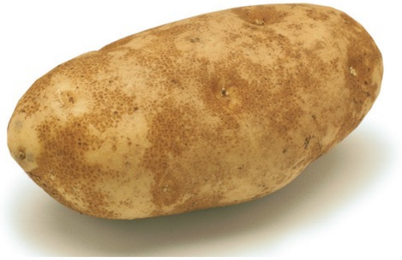
  

Fruits

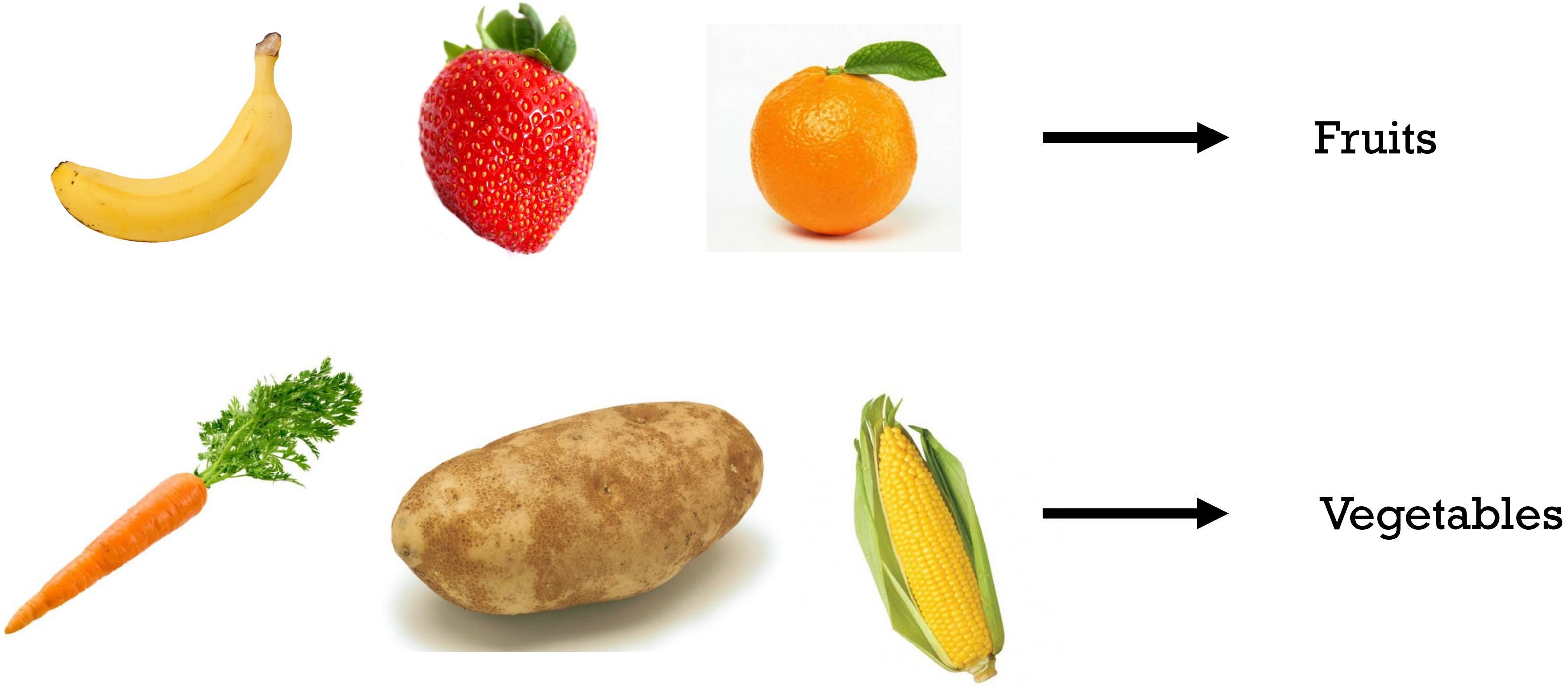
CLASSIFICATION

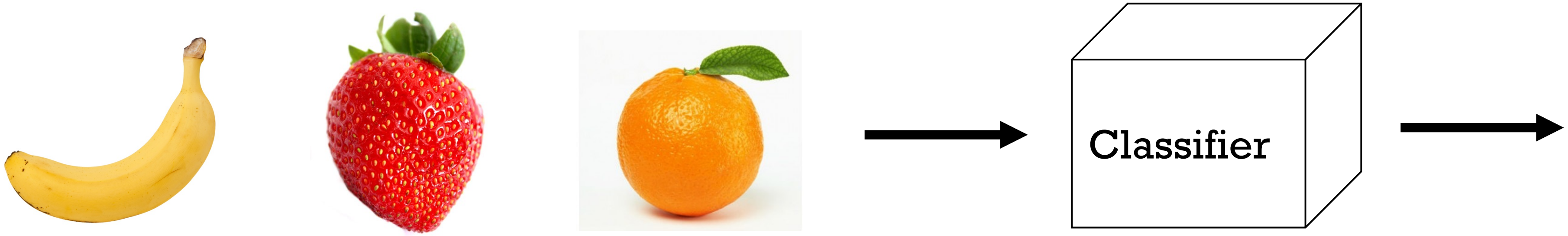
Fruits

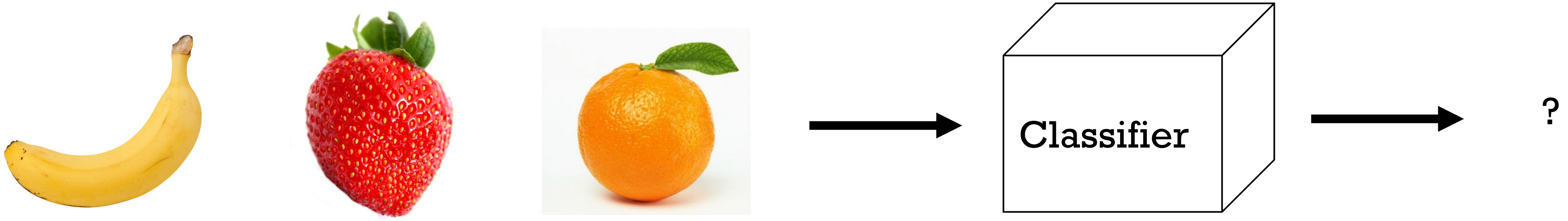
CLASSIFICATION



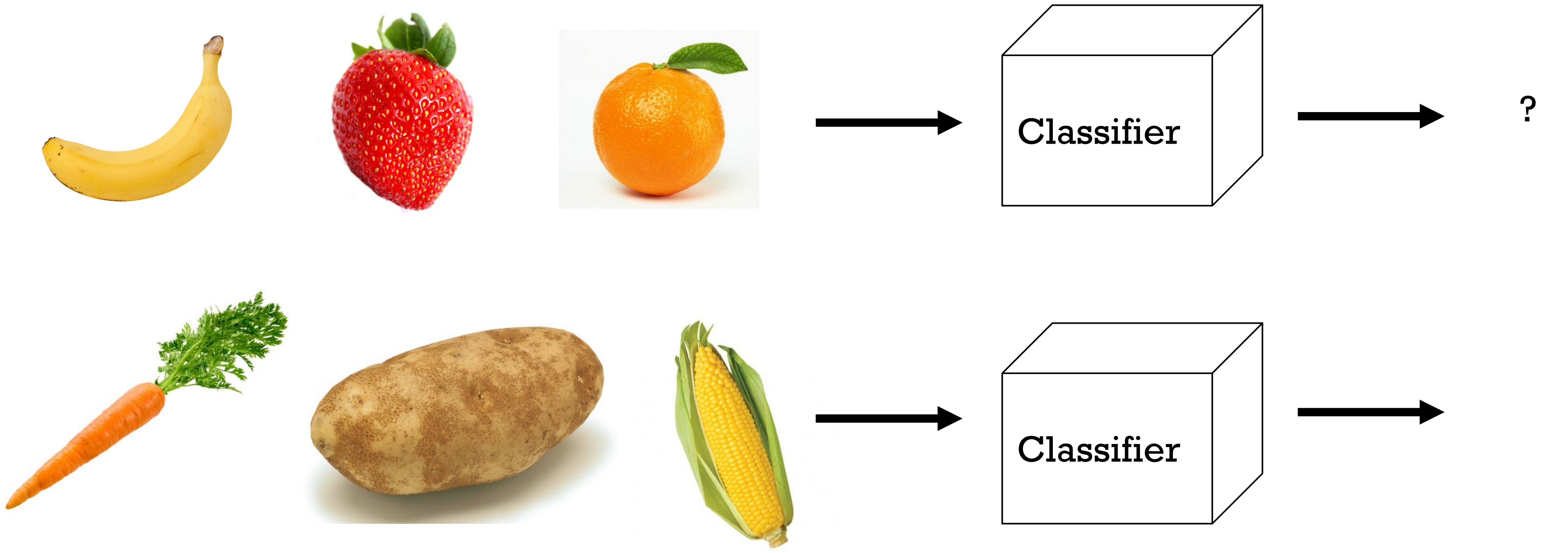
CLASSIFICATION



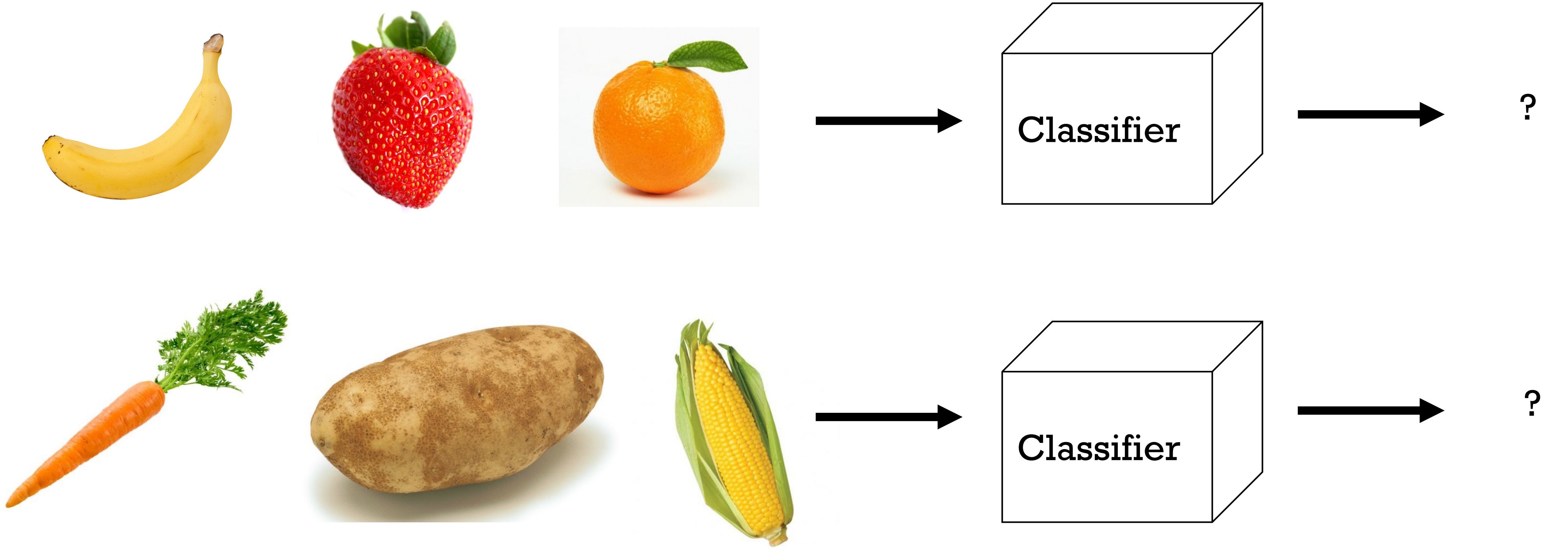
CLASSIFICATION



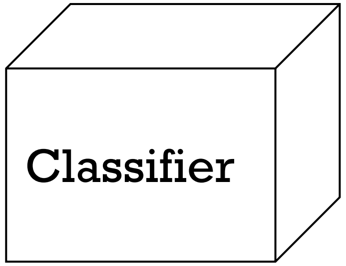
CLASSIFICATION



CLASSIFICATION

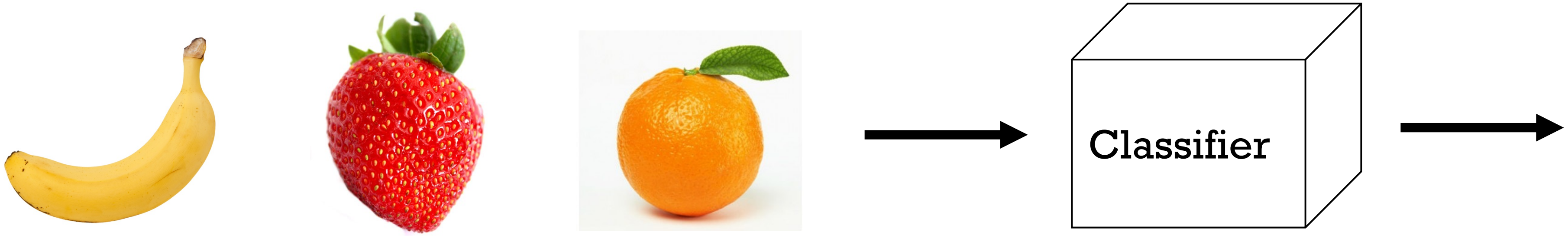


CLASSIFIER TRAINING

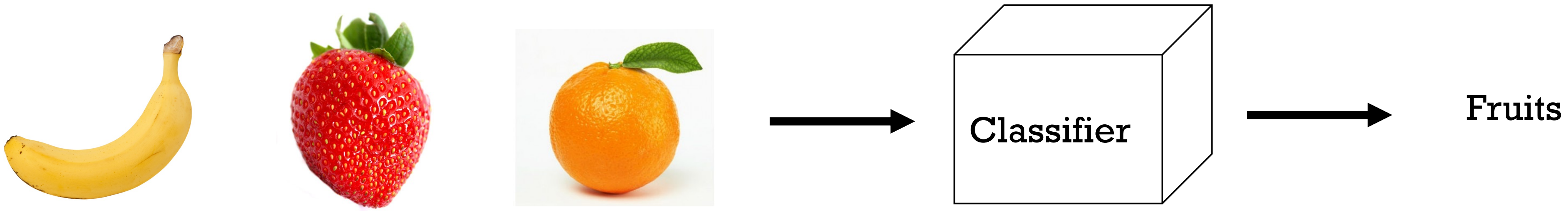


Picture from muscleandfitness.com

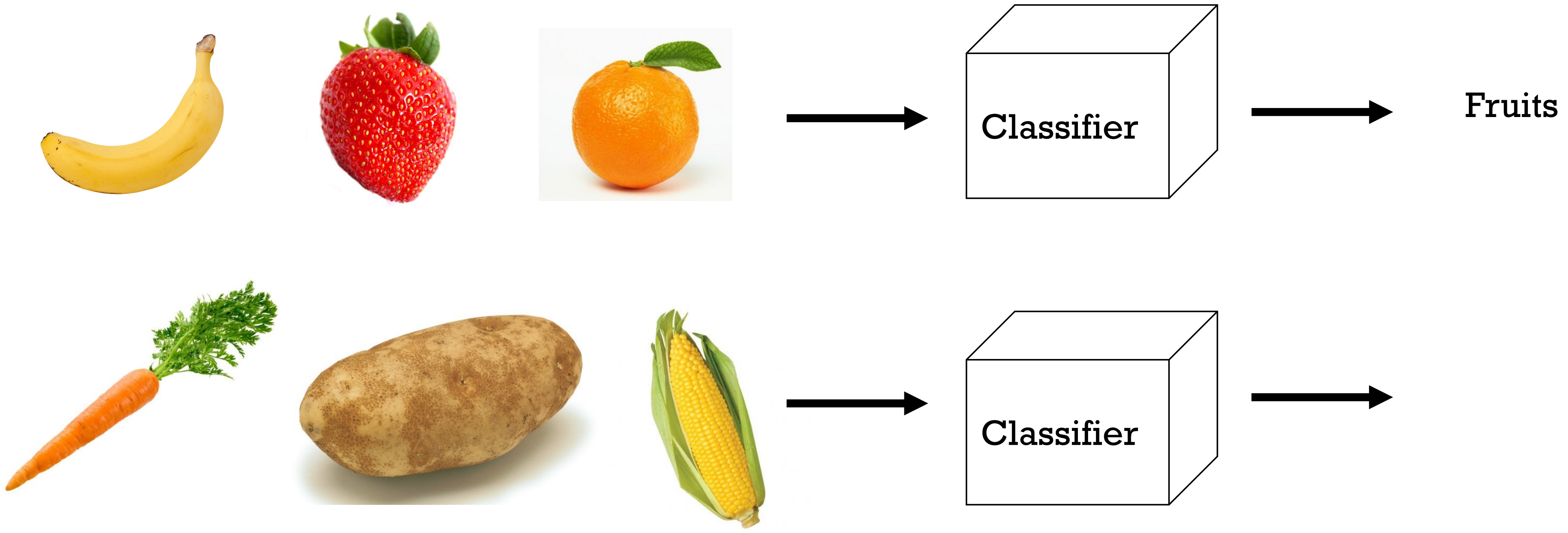
CLASSIFICATION AFTER TRAINING



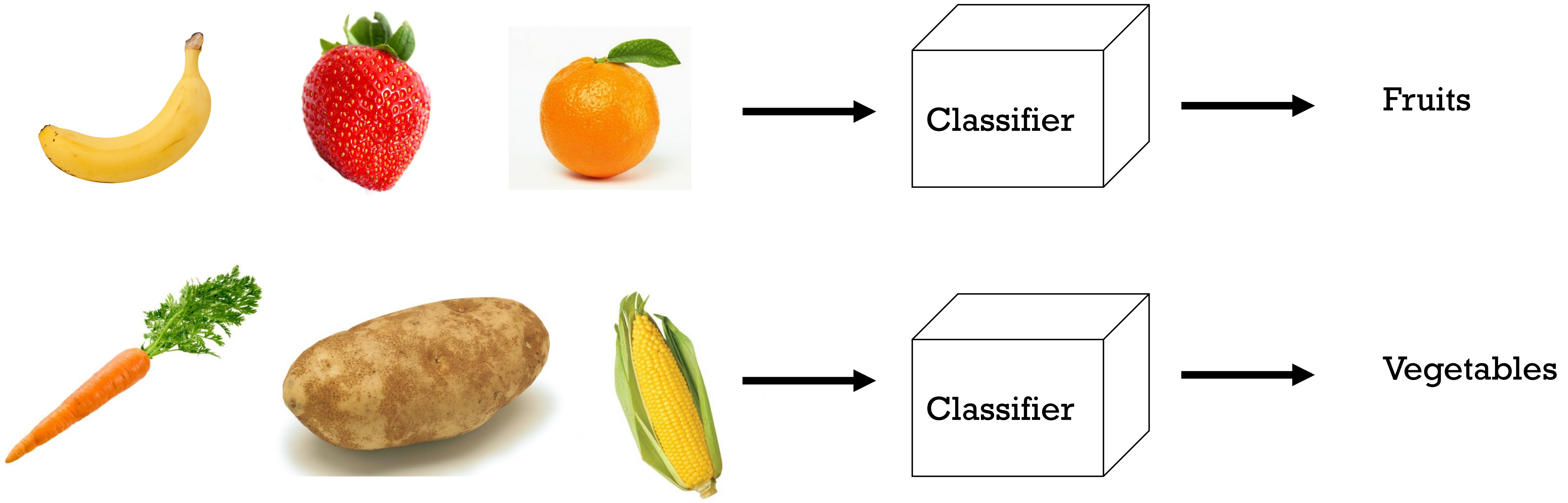
CLASSIFICATION AFTER TRAINING



CLASSIFICATION AFTER TRAINING

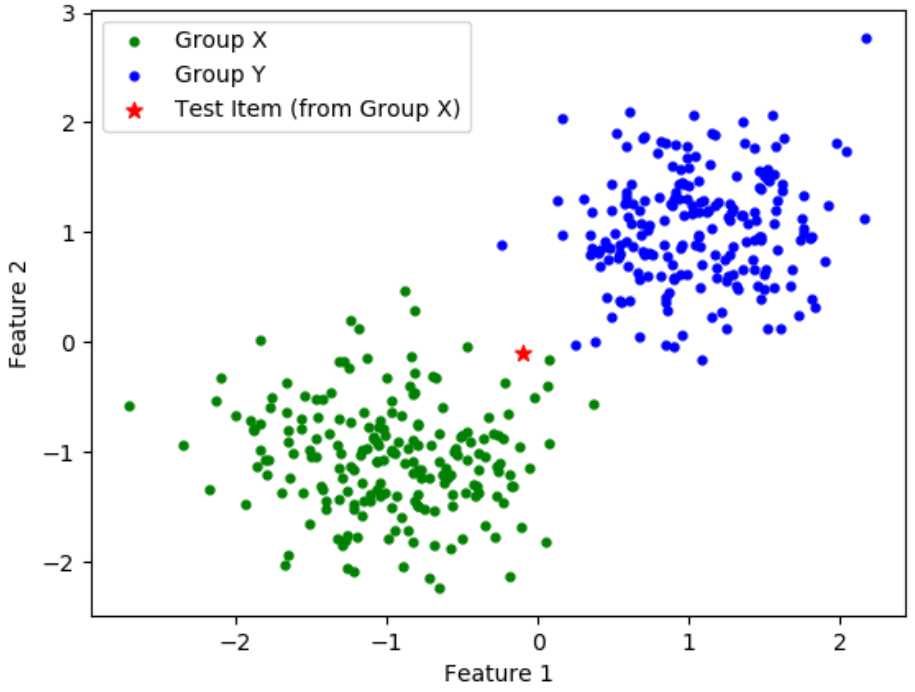


CLASSIFICATION AFTER TRAINING



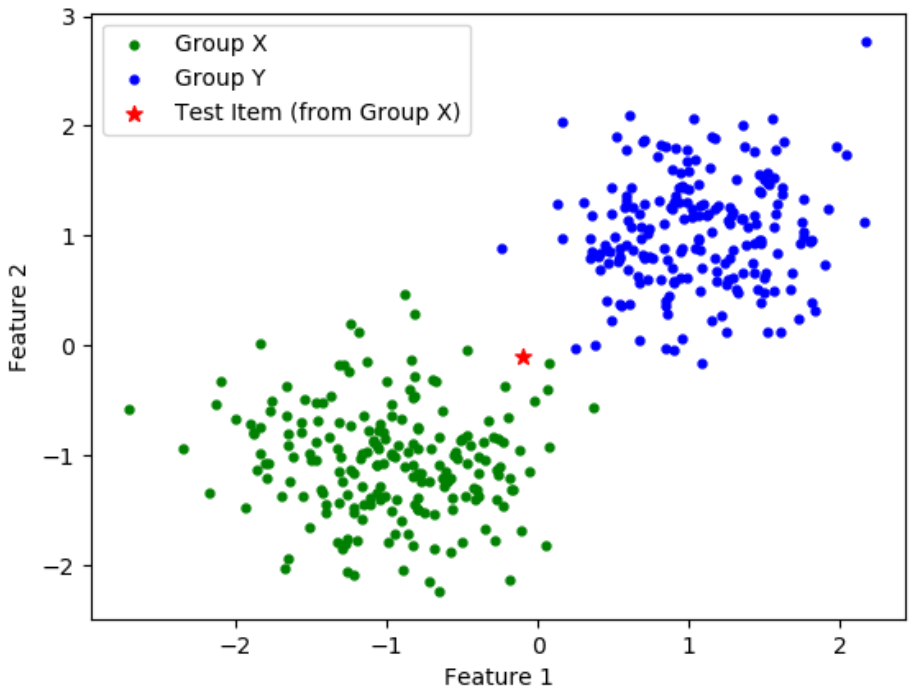
CLASSIFICATION TERMINOLOGY

- Observations
- Classes
- Features
- Feature space



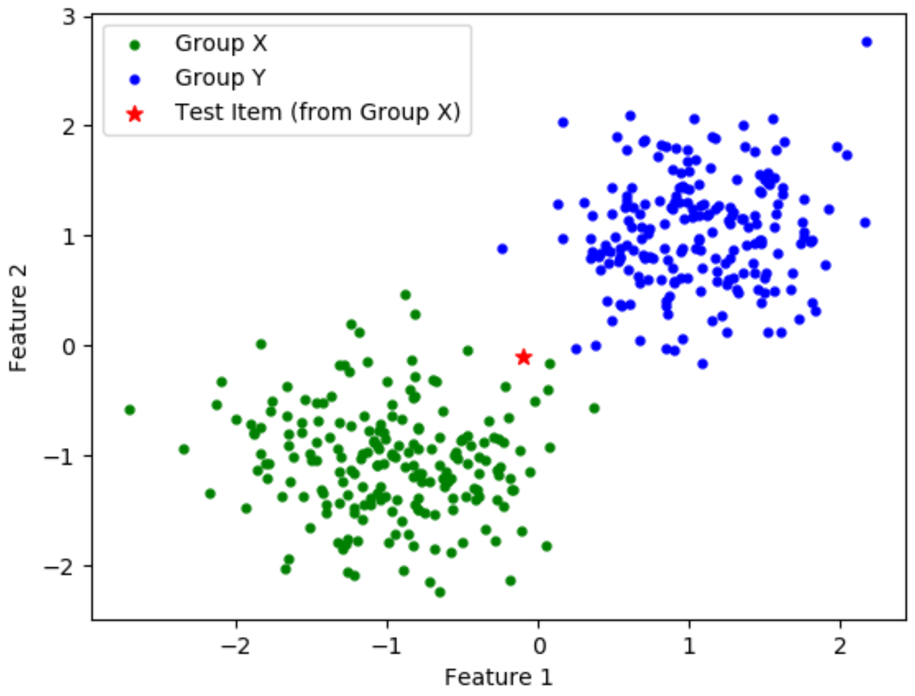
CLASSIFICATION TERMINOLOGY

- Observations
- Classes
- Features
- Feature space



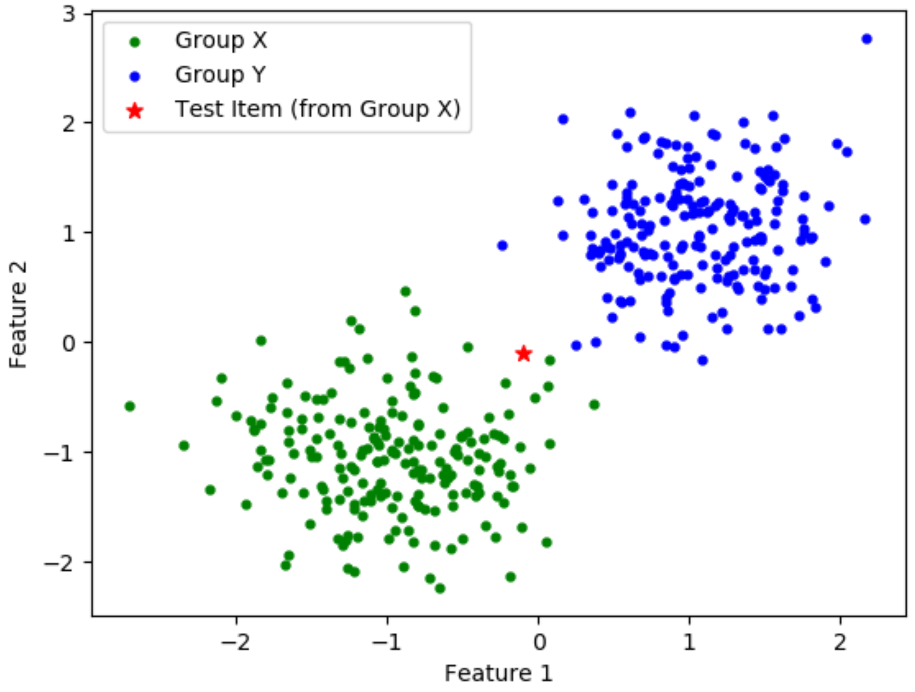
CLASSIFICATION TERMINOLOGY

- Observations
- Classes
- Features
- Feature space



CLASSIFICATION TERMINOLOGY

- Observations
- Classes
- Features
- Feature space

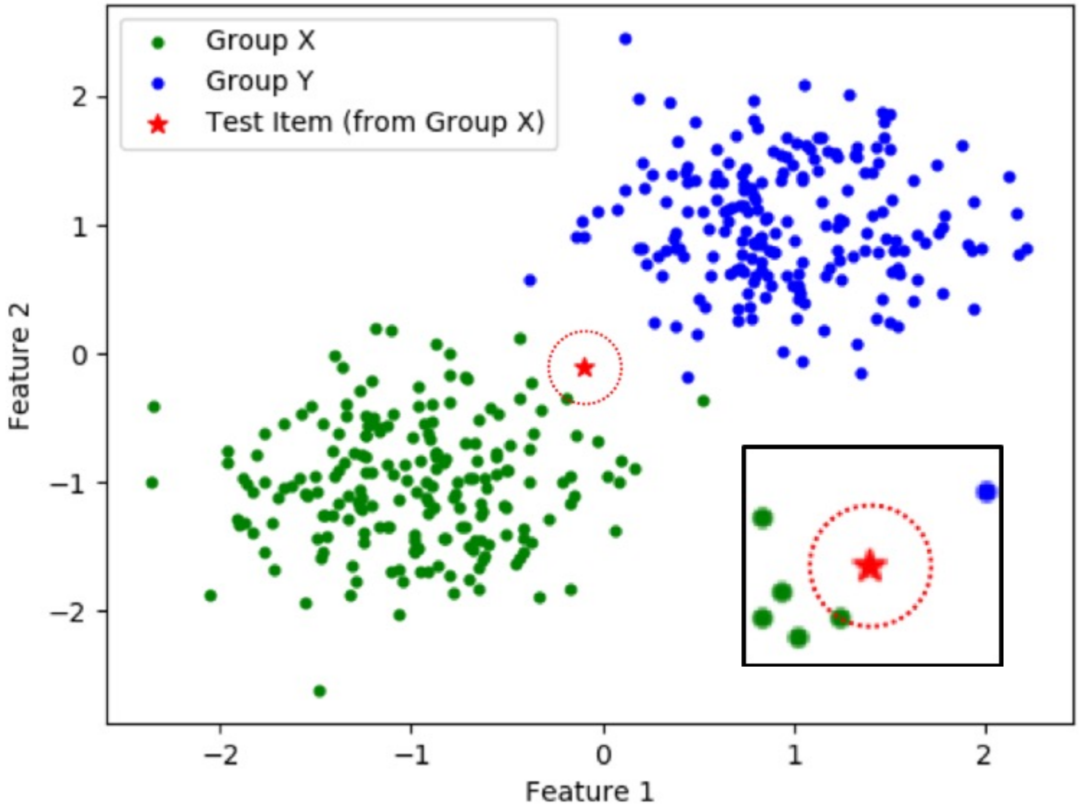


OUTLINE

- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

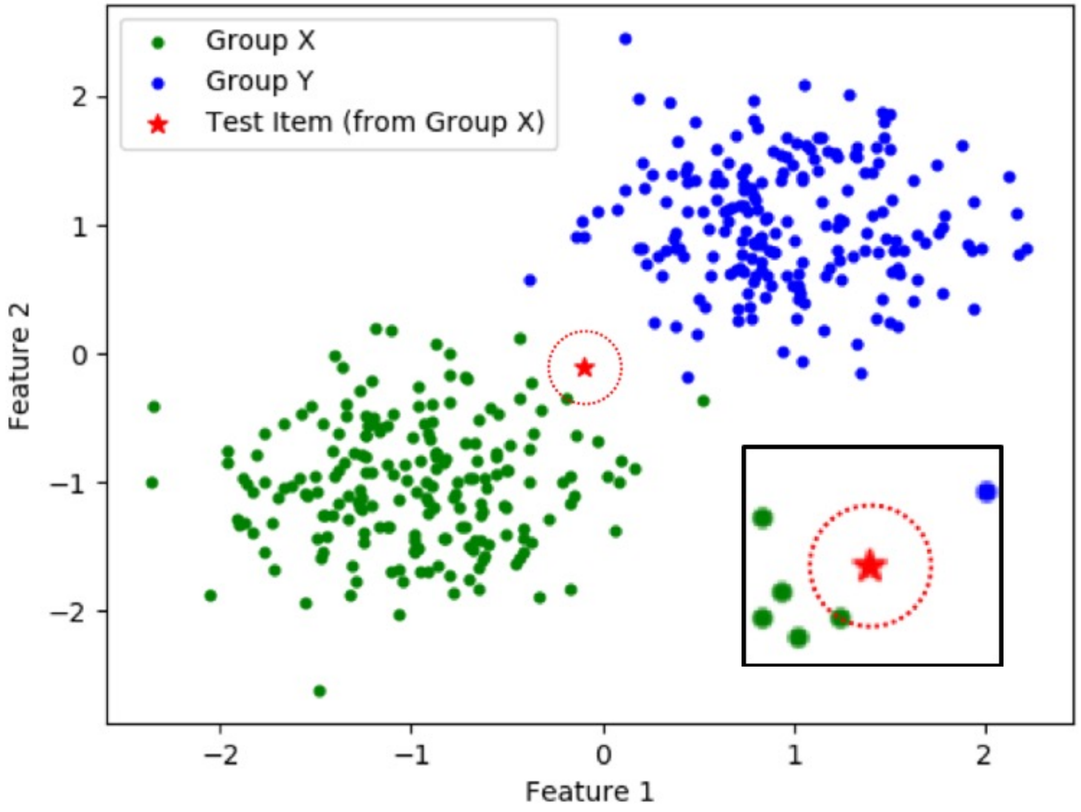
K-NEAREST NEIGHBORS

- No training required
- Euclidean distance between test item and every item in training set
- K = 1: nearest neighbor determines class label
- K > 1: k-nearest neighbors determine class label
  - Majority
  - Weighting



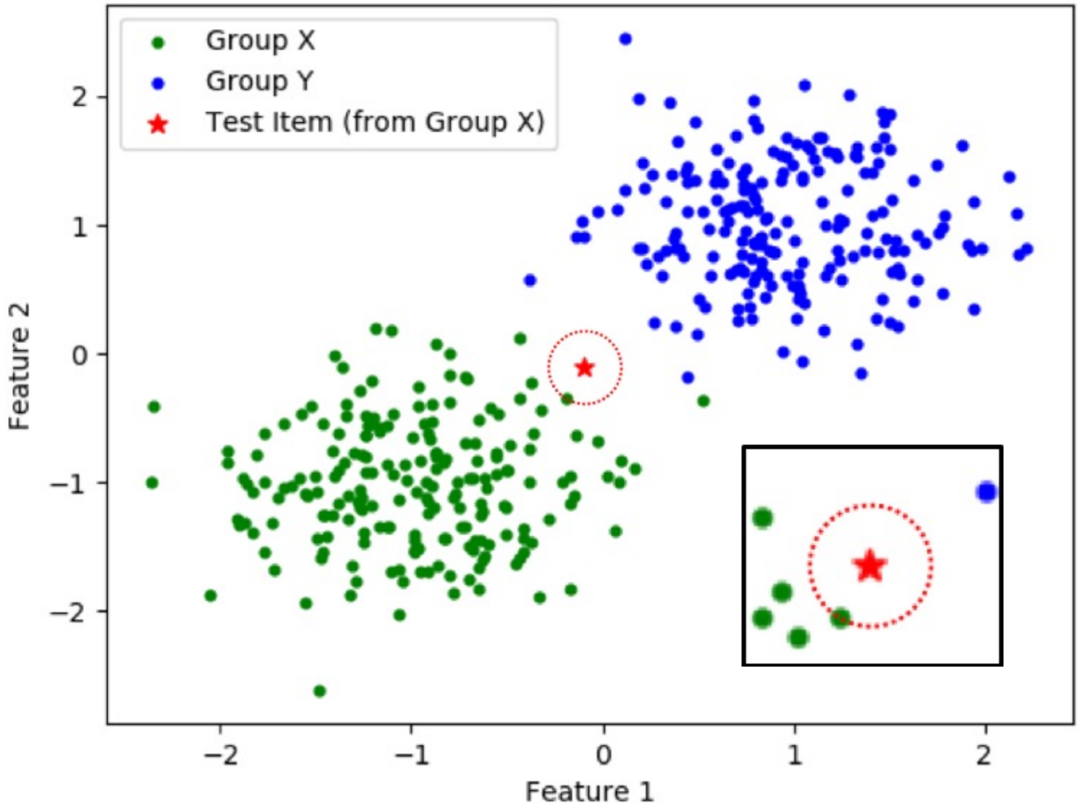
K-NEAREST NEIGHBORS

- No training required
- Euclidean distance between test item and every item in training set
- K = 1: nearest neighbor determines class label
- K > 1: k-nearest neighbors determine class label
  - Majority
  - Weighting



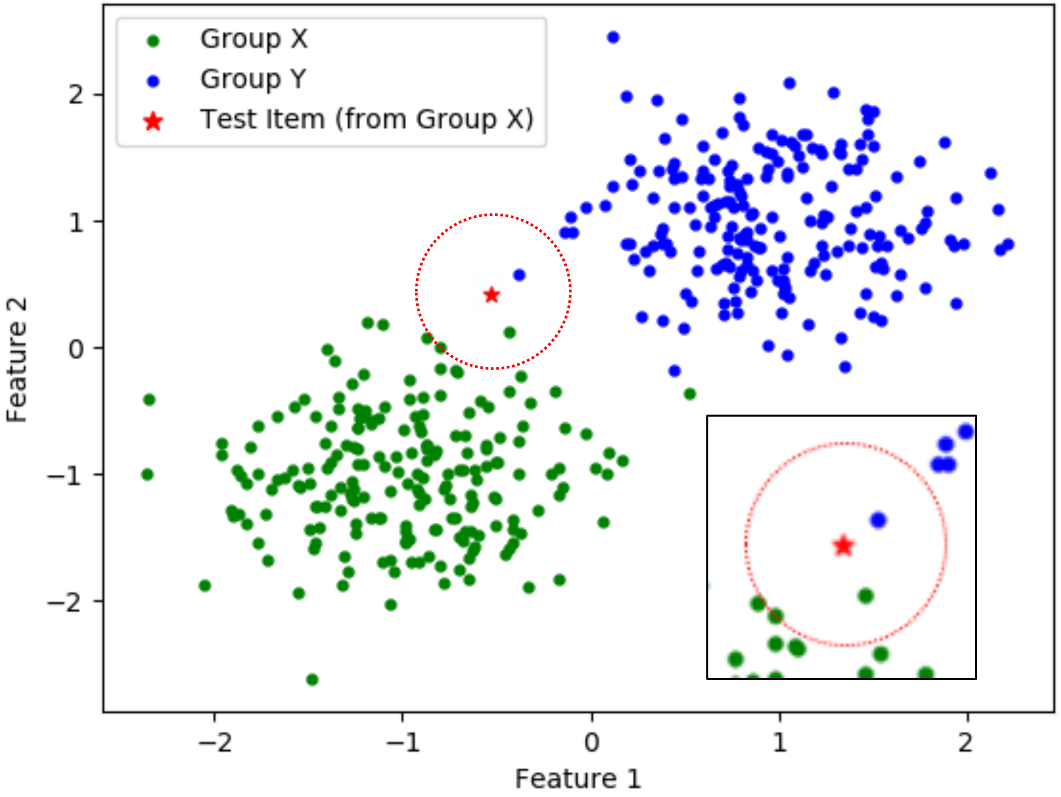
K-NEAREST NEIGHBORS

- No training required
- Euclidean distance between test item and every item in training set
- K = 1: nearest neighbor determines class label
- K > 1: k-nearest neighbors determine class label
  - Majority
  - Weighting



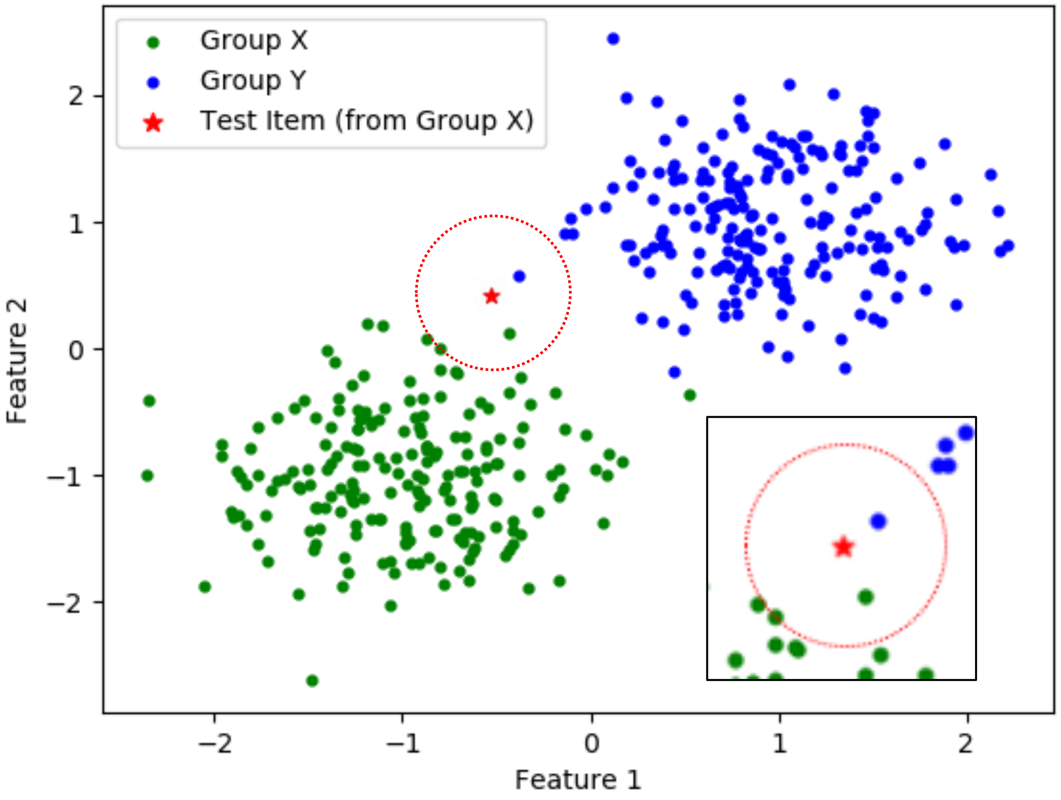
K-NEAREST NEIGHBORS

- No training required
- Euclidean distance between test item and every item in training set
- K = 1: nearest neighbor determines class label
- K > 1: k-nearest neighbors determine class label
  - Majority
  - Weighting



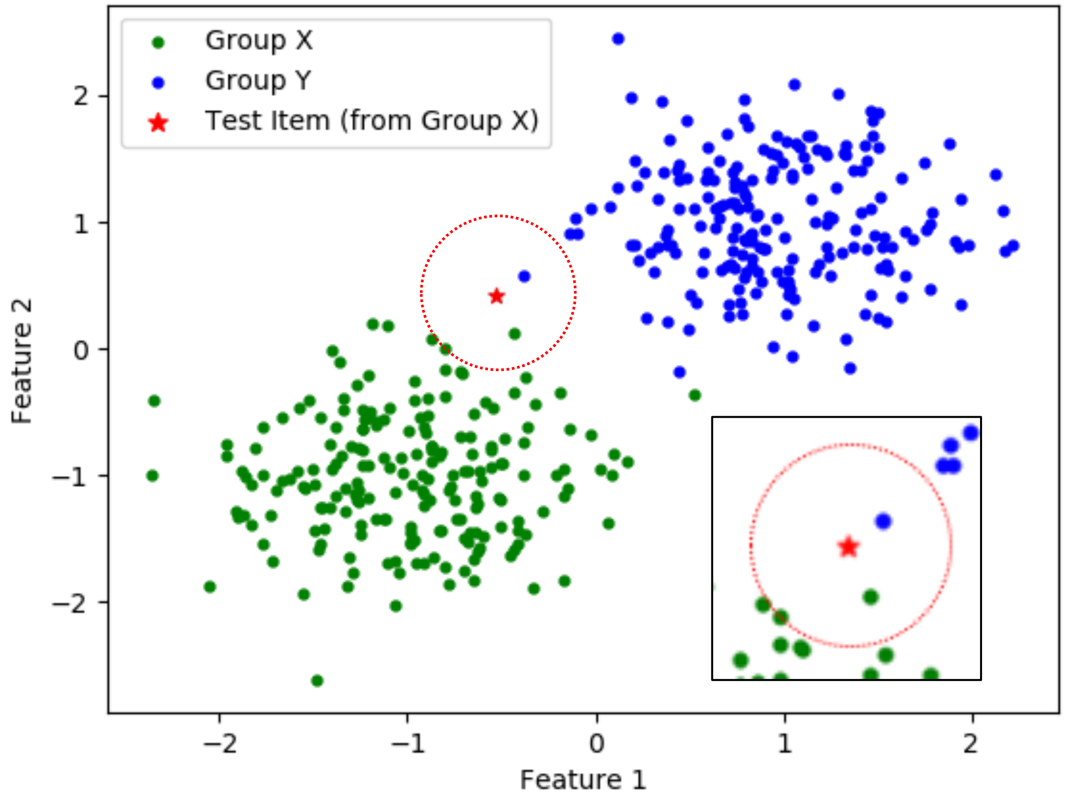
K-NEAREST NEIGHBORS

- No training required
- Euclidean distance between test item and every item in training set
- K = 1: nearest neighbor determines class label
- K > 1: k-nearest neighbors determine class label
  - Majority
  - Weighting



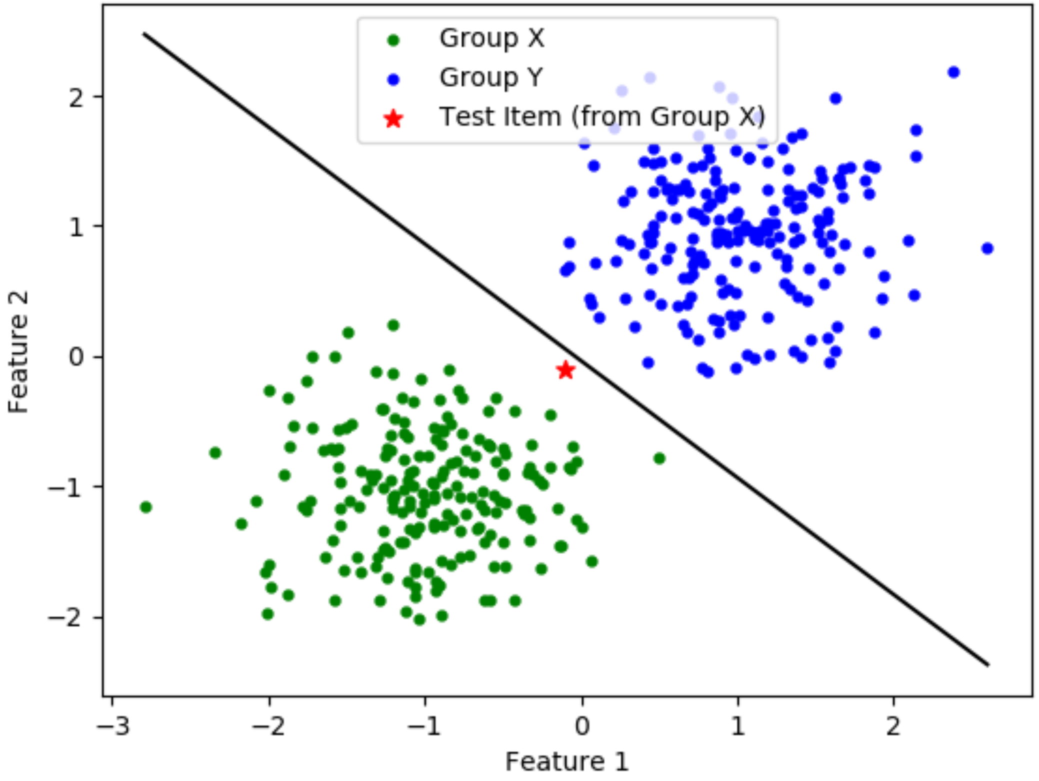
K-NEAREST NEIGHBORS

- No training required
- Euclidean distance between test item and every item in training set
- K = 1: nearest neighbor determines class label
- K > 1: k-nearest neighbors determine class label
  - Majority
  - Weighting



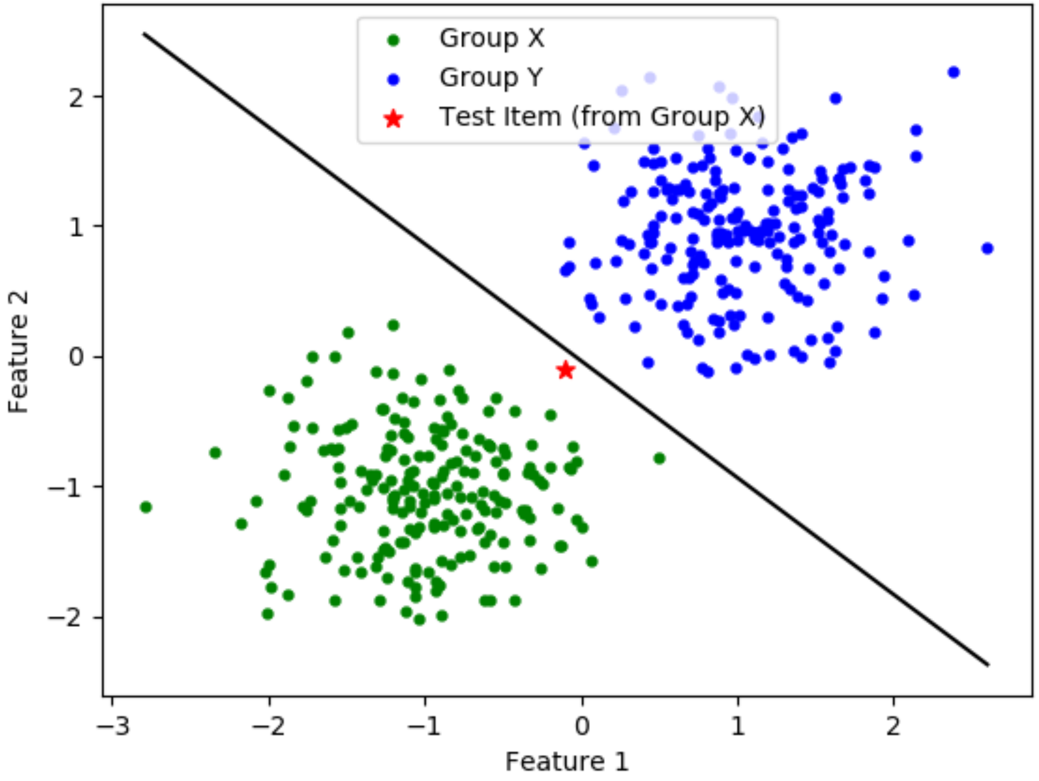
LINEAR DISCRIMINANT ANALYSIS

- Provides a line that best separates the two features
- Linear decision boundary
- Importance of linear separability



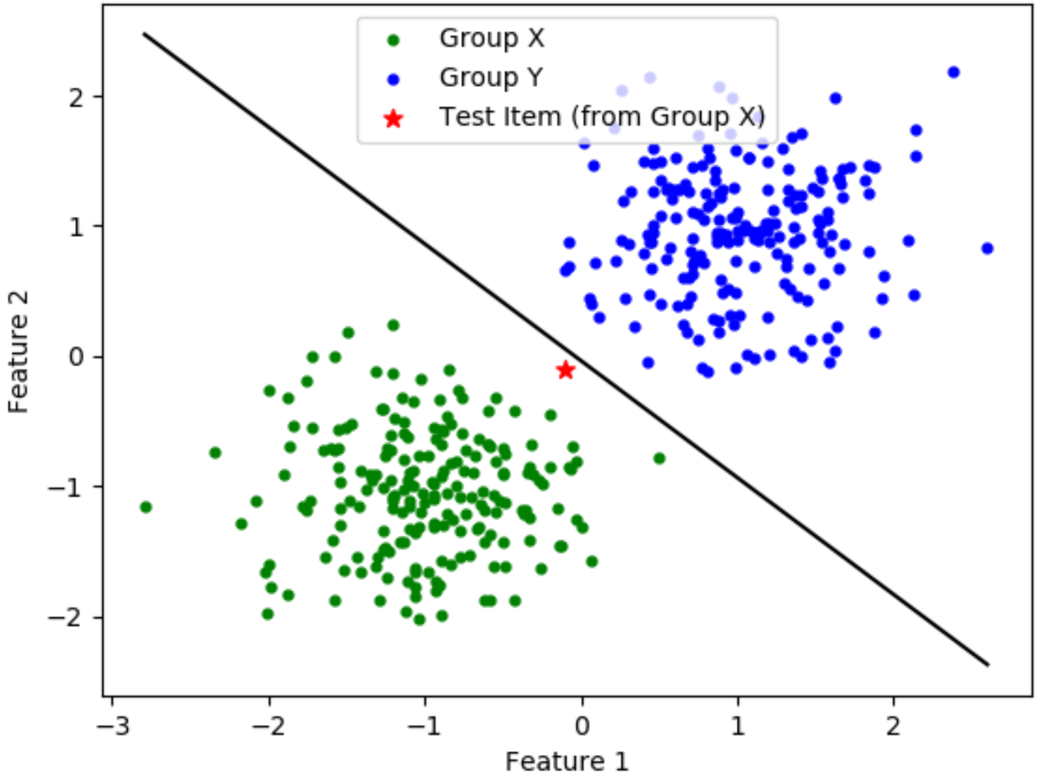
LINEAR DISCRIMINANT ANALYSIS

- Provides a line that best separates the two features
- Linear decision boundary
- Importance of linear separability



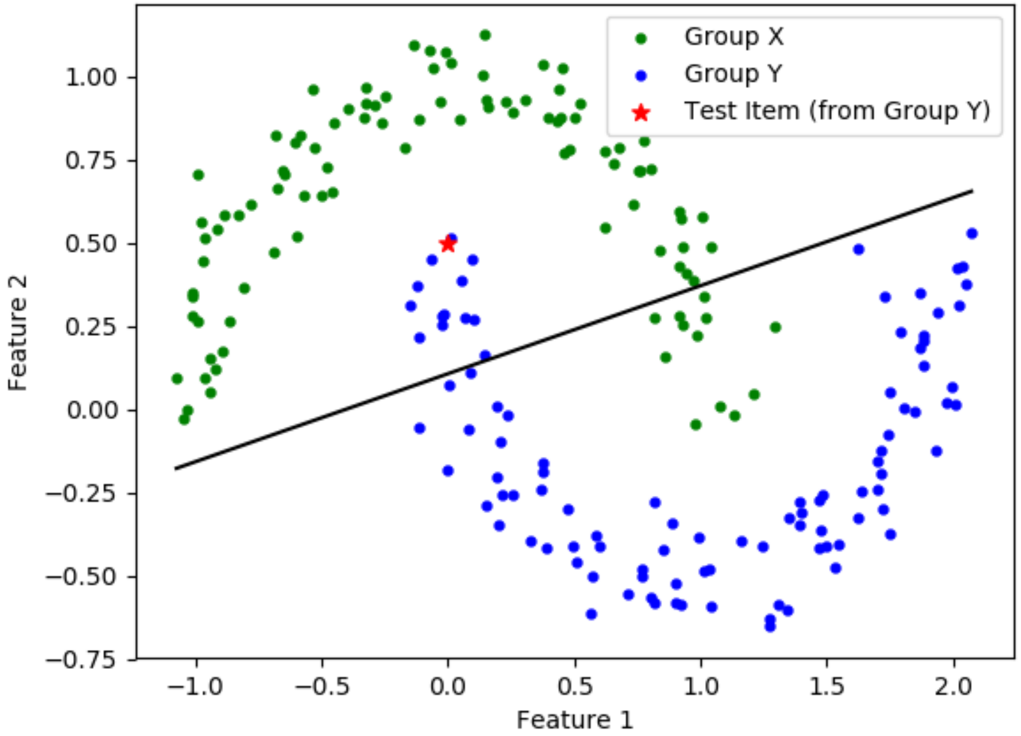
LINEAR DISCRIMINANT ANALYSIS

- Provides a line that best separates the two features
- Linear decision boundary
- Importance of linear separability



LINEAR DISCRIMINANT ANALYSIS

- Provides a line that best separates the two features
- Linear decision boundary
- Importance of linear separability

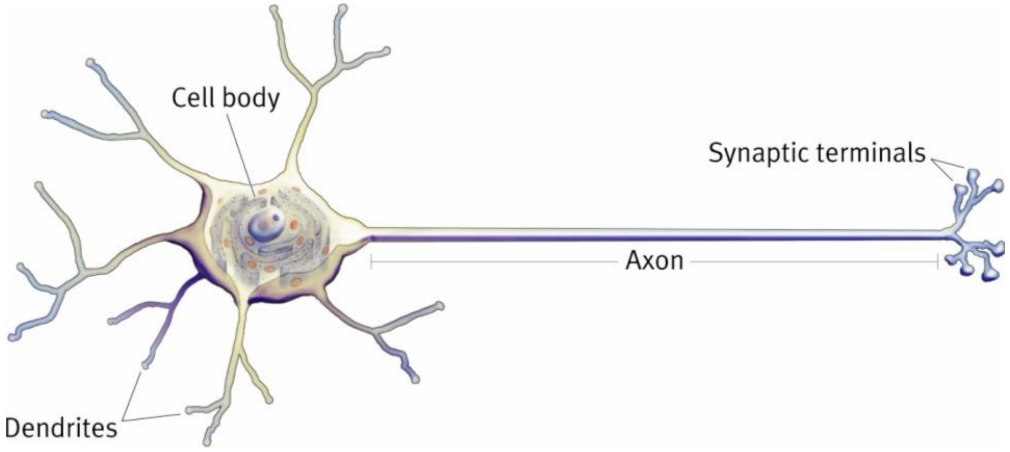


OUTLINE

- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

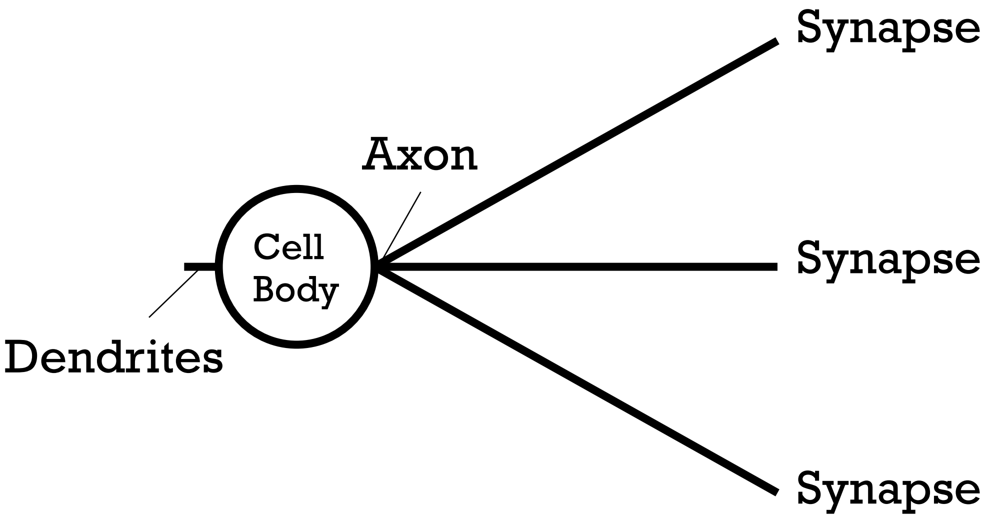
NEURON → NODE IN NEURAL NETWORK

Neuron

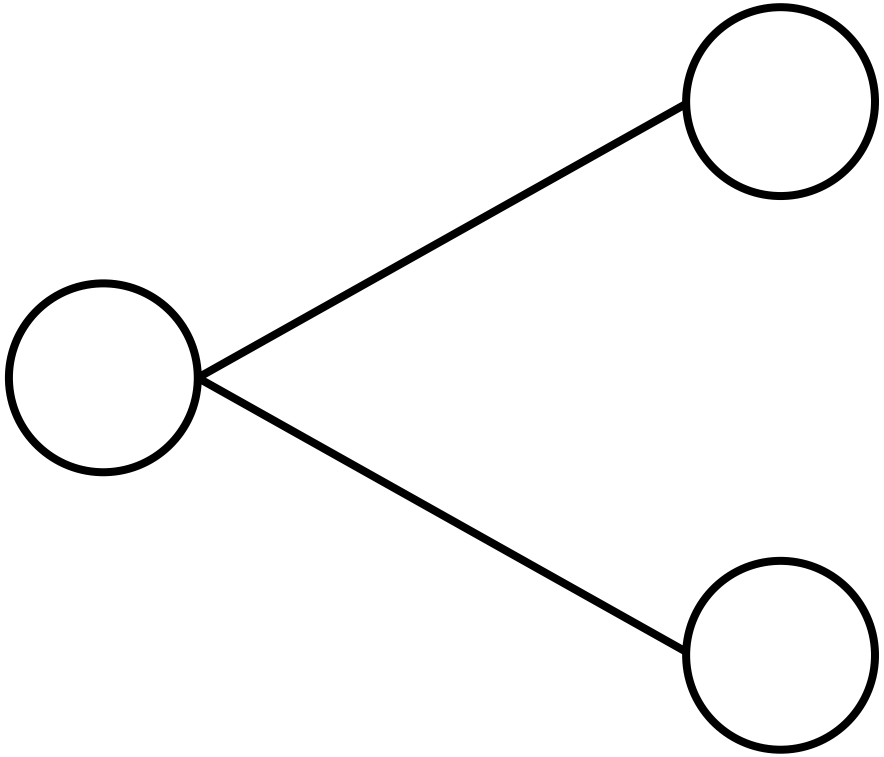


From Memorang

Node in NN

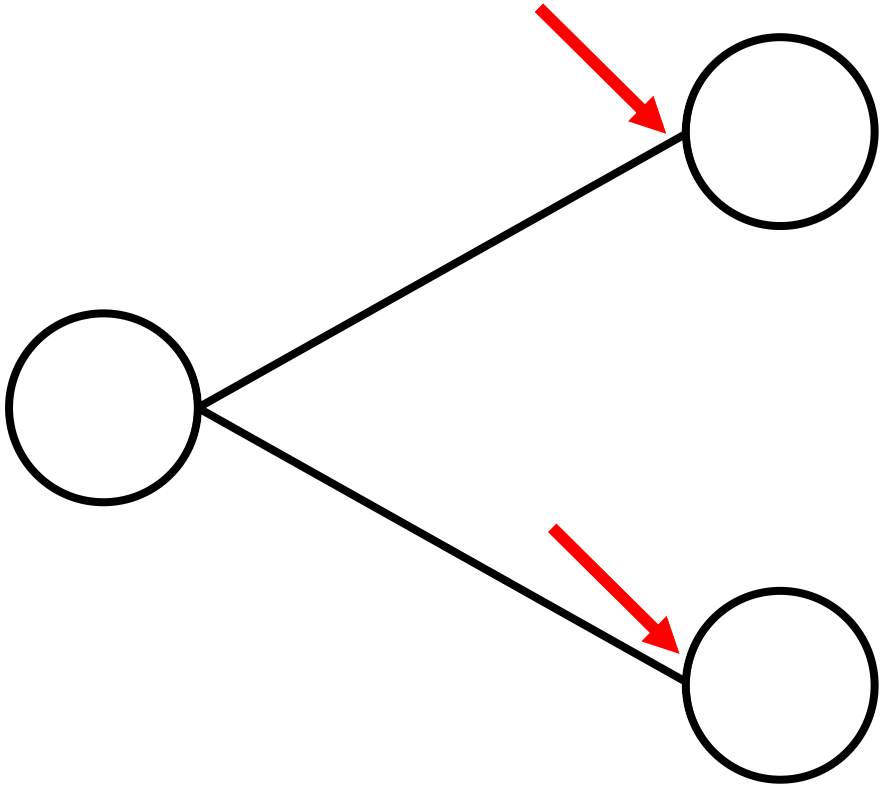


ARTIFICIAL NEURAL NETWORK – 3 NEURONS



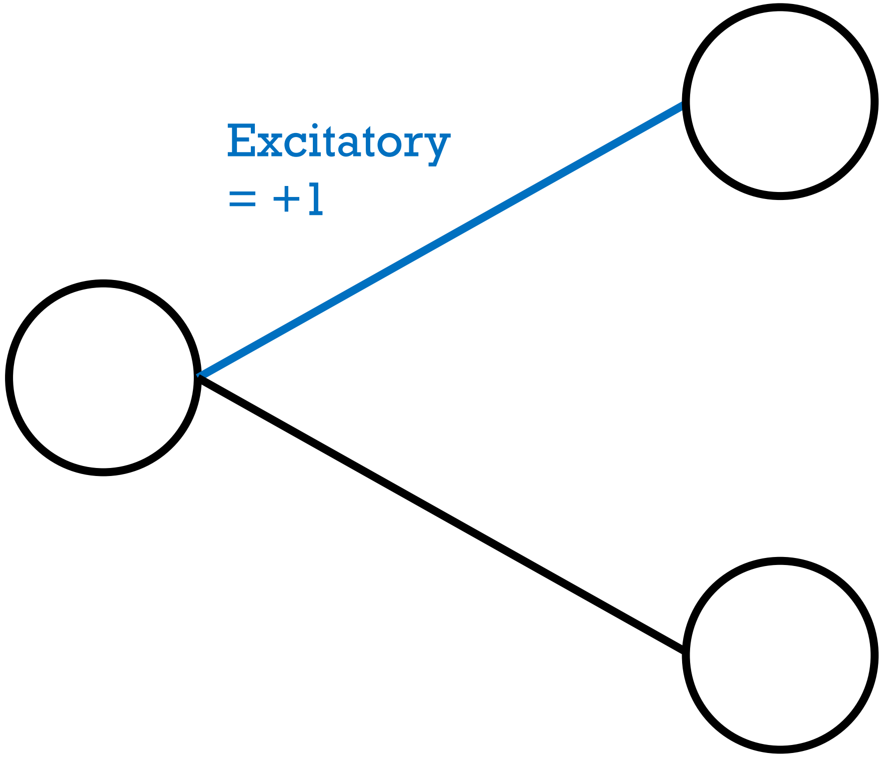
NEURAL NETWORK TERMINOLOGY

- Inputs
- Weights
- Activation
- Outputs
- Layers



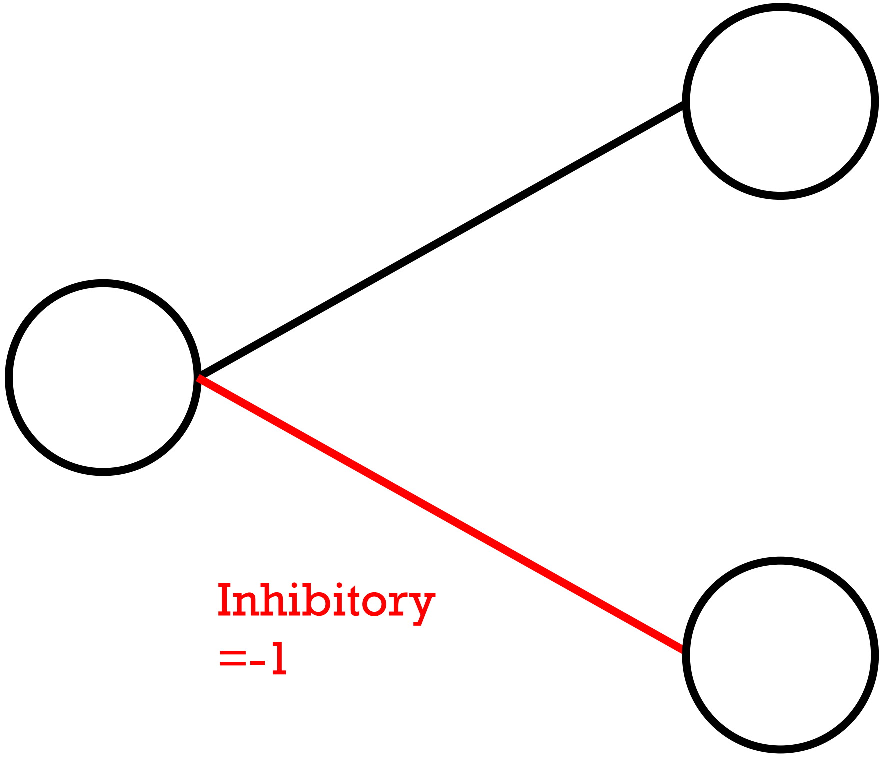
NEURAL NETWORK TERMINOLOGY

- Inputs
- Weights
- Activation
- Outputs
- Layers



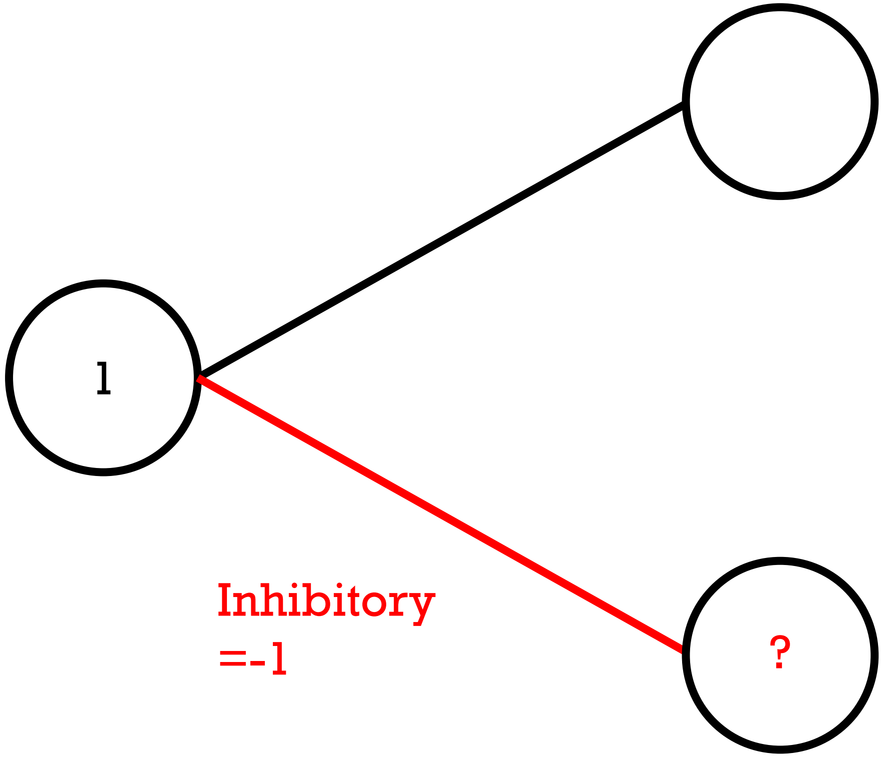
NEURAL NETWORK TERMINOLOGY

- Inputs
- Weights
- Activation
- Outputs
- Layers



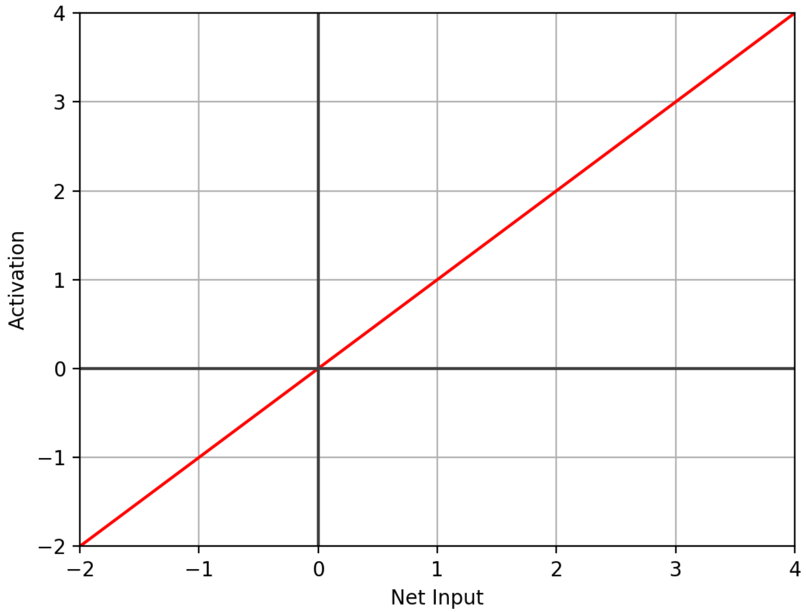
NEURAL NETWORK TERMINOLOGY

- Inputs
- Weights
- Activation
- Outputs
- Layers

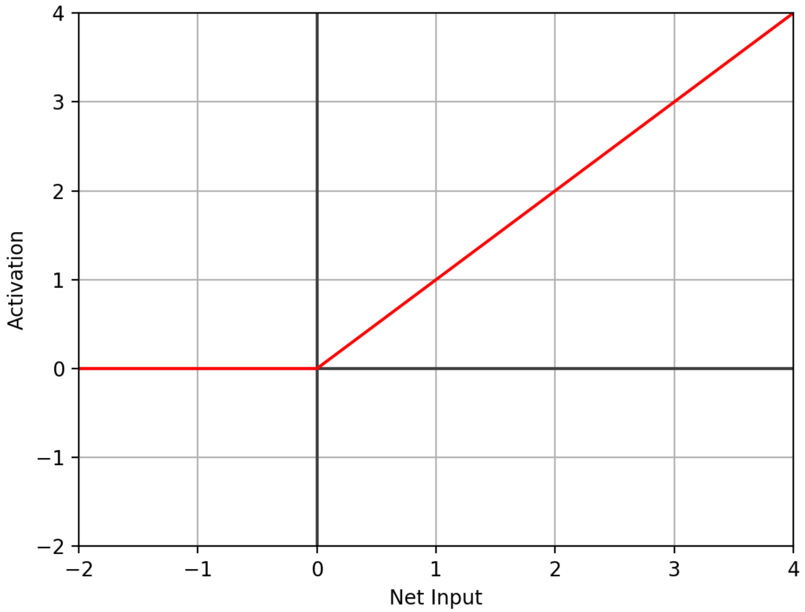


ACTIVATION FUNCTIONS

Linear

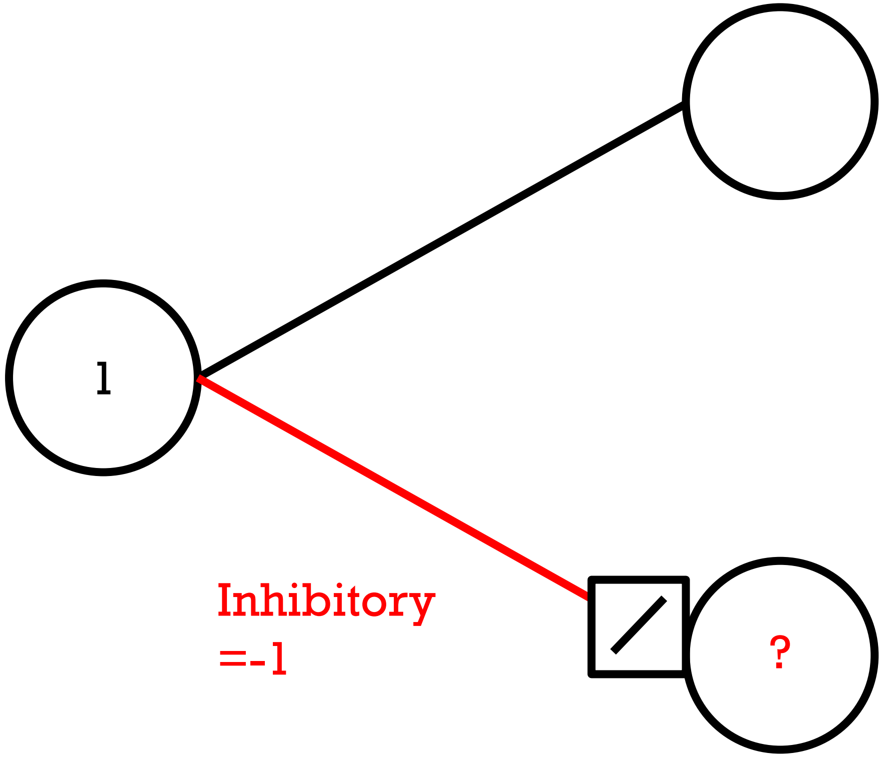


Rectified Linear Unit (ReLU)



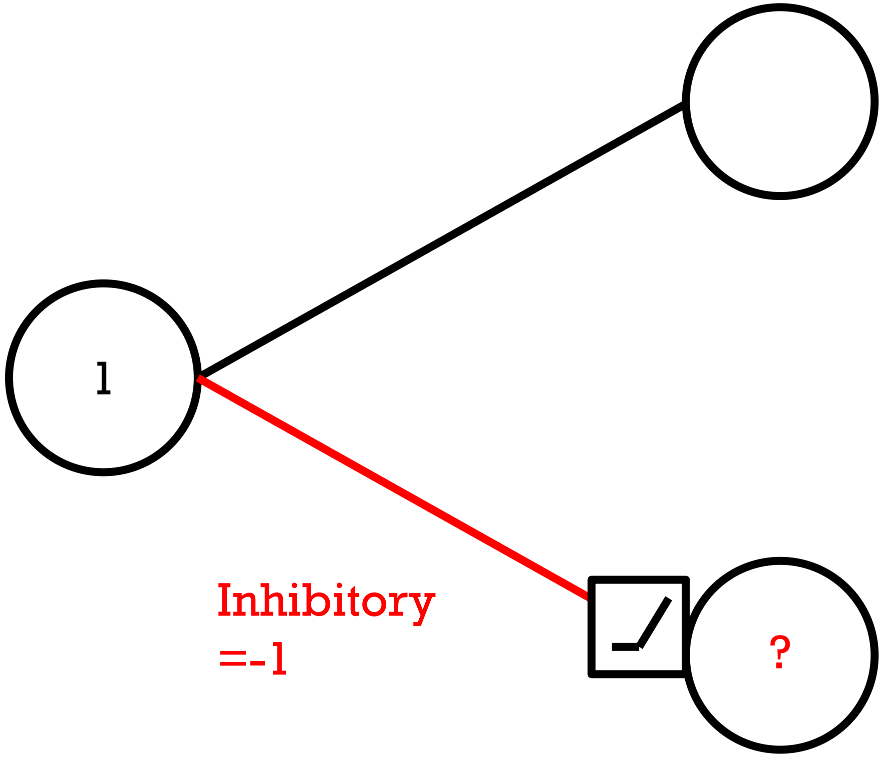
NEURAL NETWORK TERMINOLOGY

- Inputs
- Weights
- Activation
- Outputs
- Layers



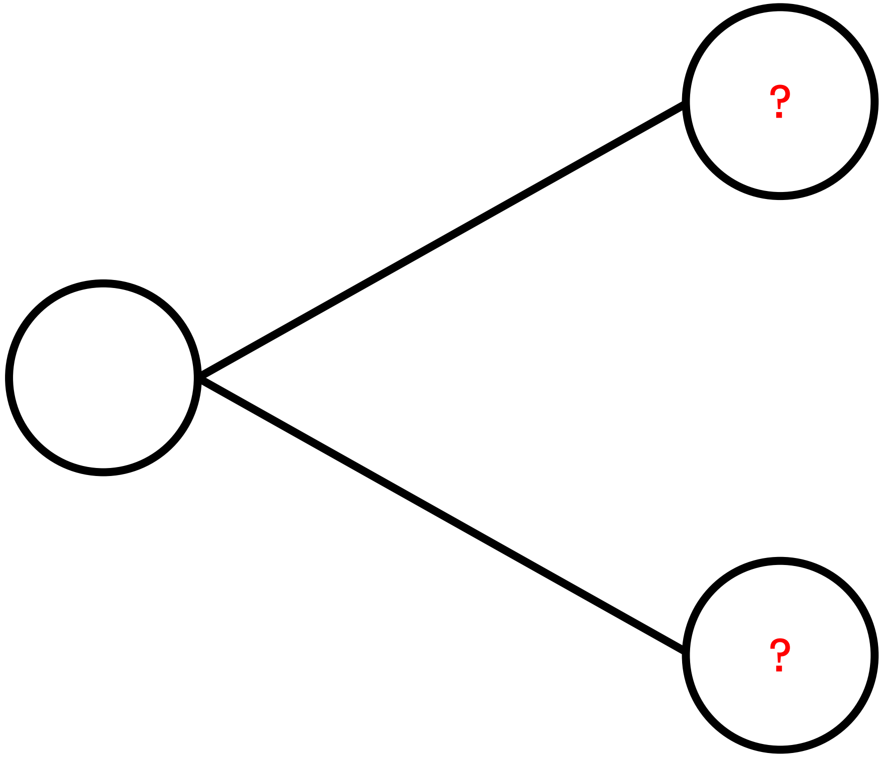
NEURAL NETWORK TERMINOLOGY

- Inputs
- Weights
- Activation
- Outputs
- Layers



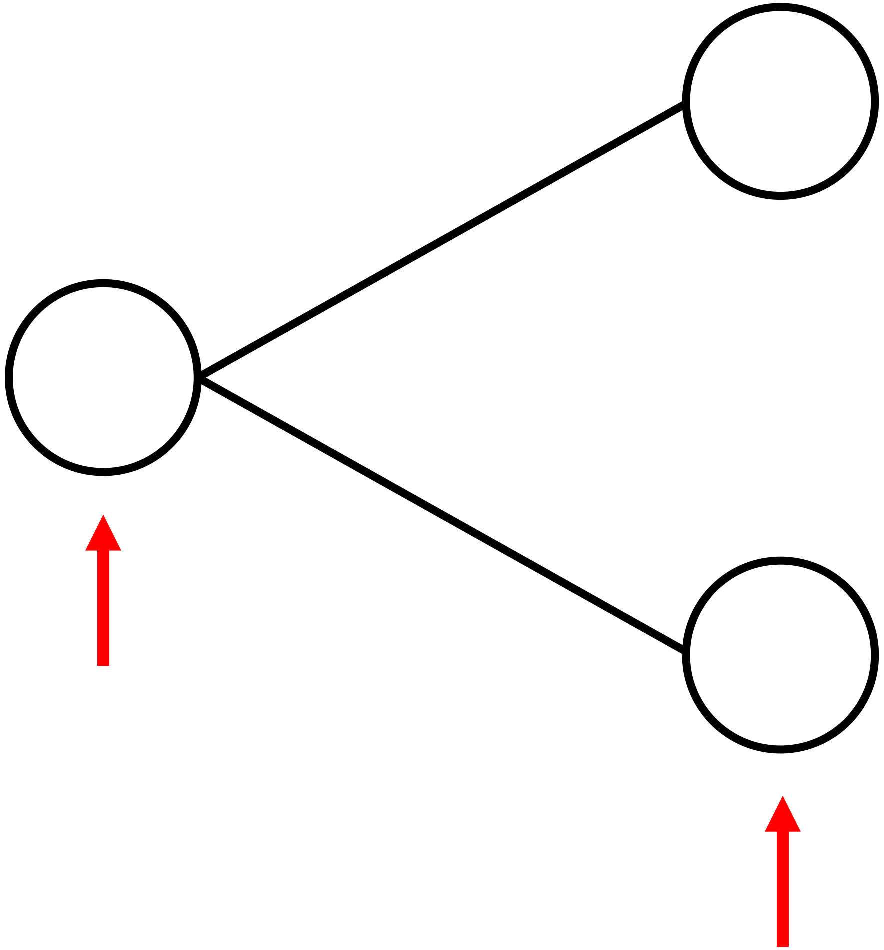
NEURAL NETWORK TERMINOLOGY

- Inputs
- Weights
- Activation
- Outputs
- Layers



NEURAL NETWORK TERMINOLOGY

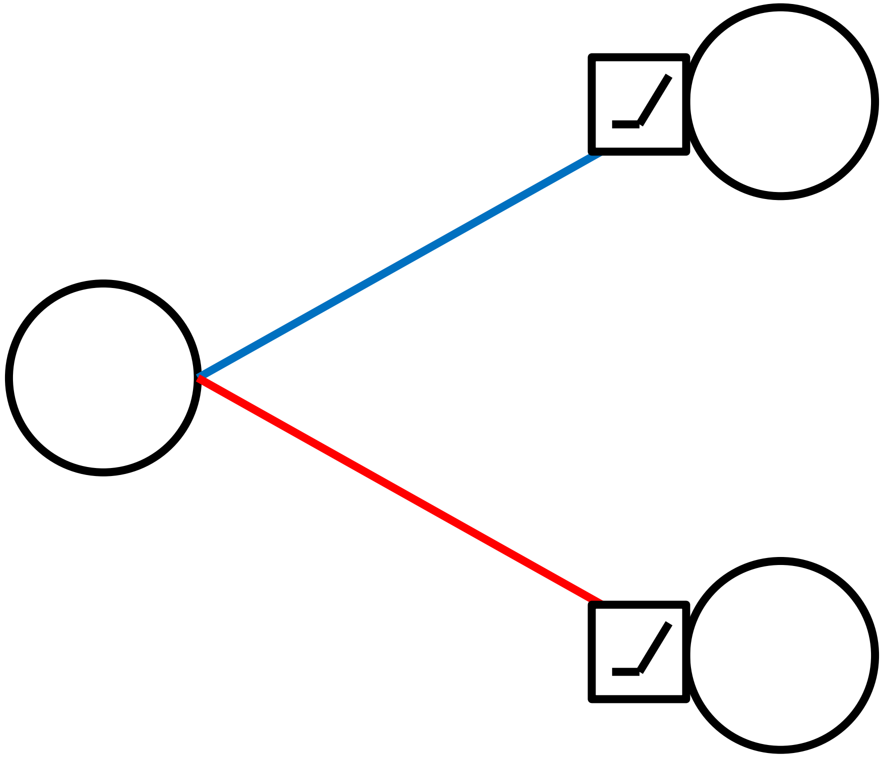
- Inputs
- Weights
- Activation
- Outputs
- Layers



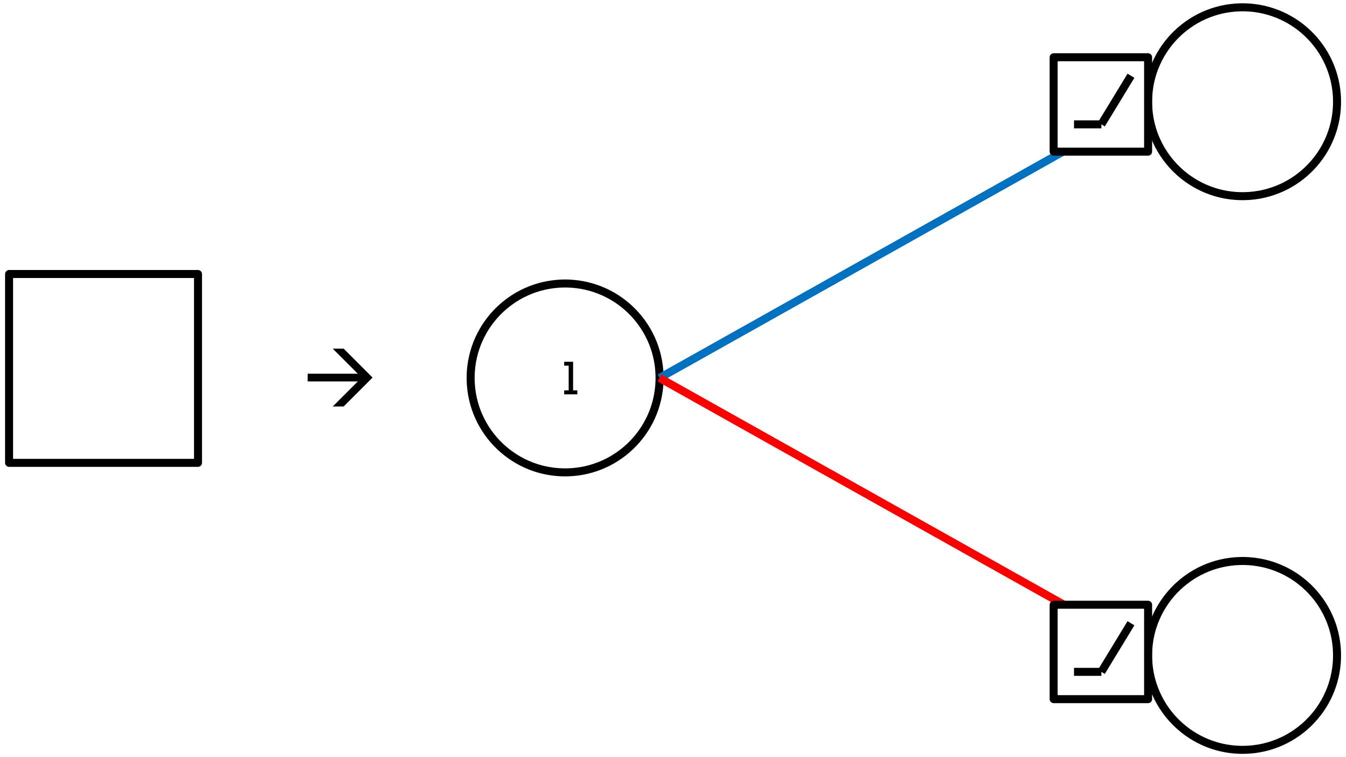
OVERVIEW

- Why ASDNet?
- Neurons → Neural Network
- Basic Neural Network Examples
- Weight Tuning/Optimization
- ASDNet
- Visualization
- Results

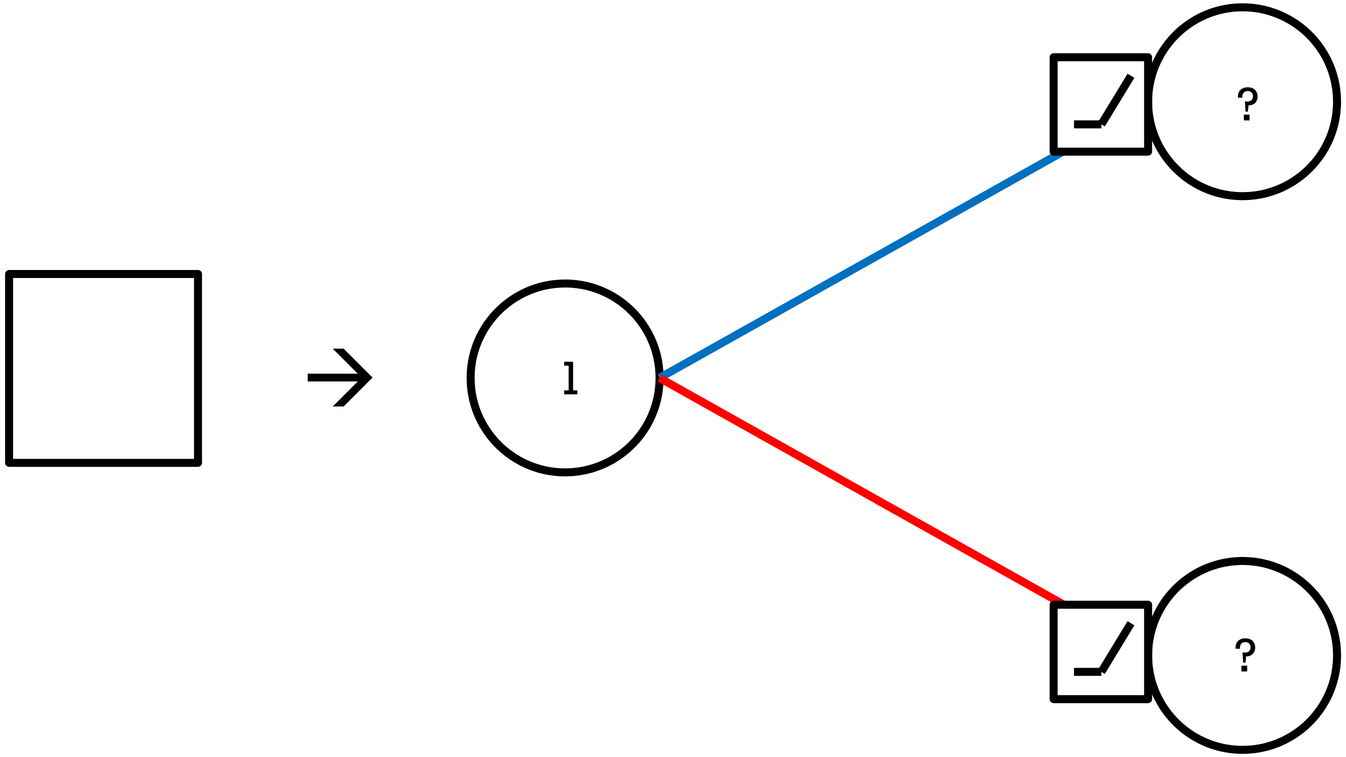
BASIC NEURAL NETWORK



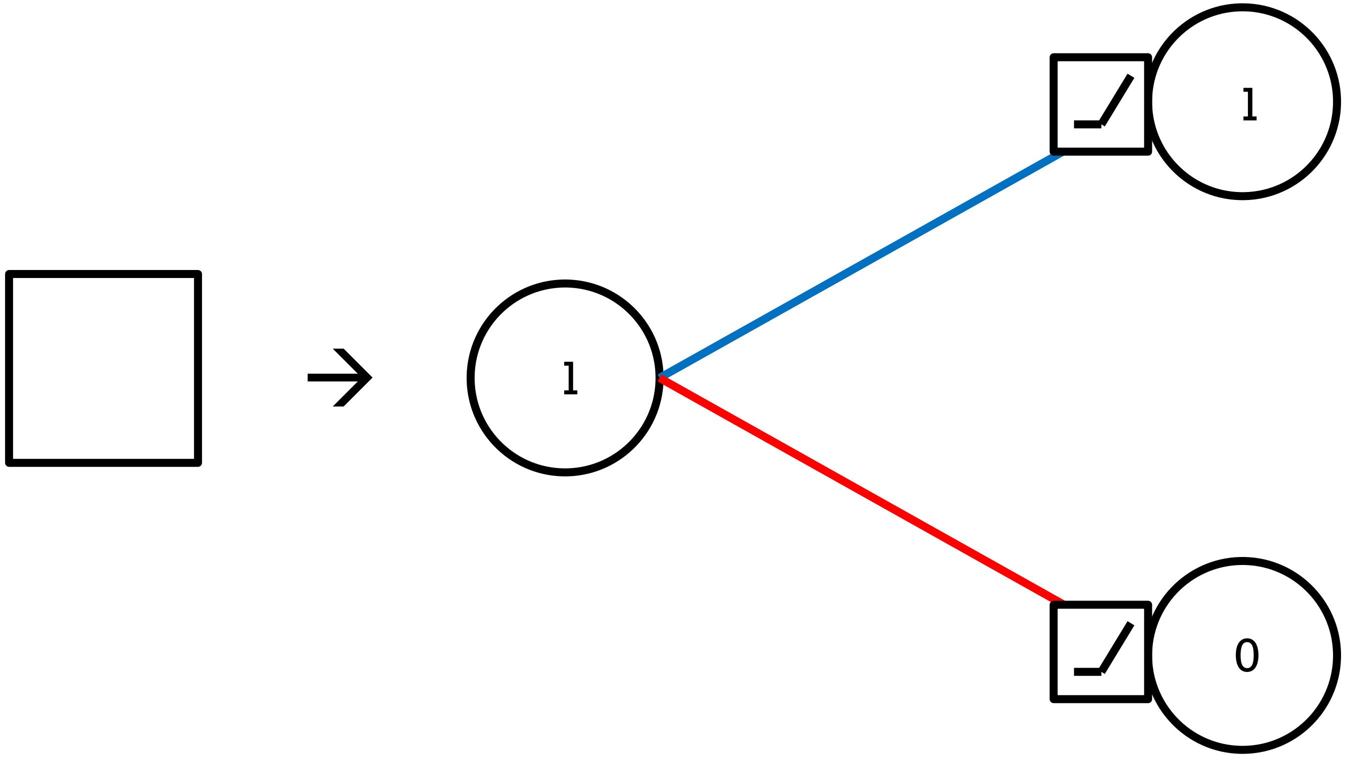
BASIC NEURAL NETWORK – 1 PIXEL



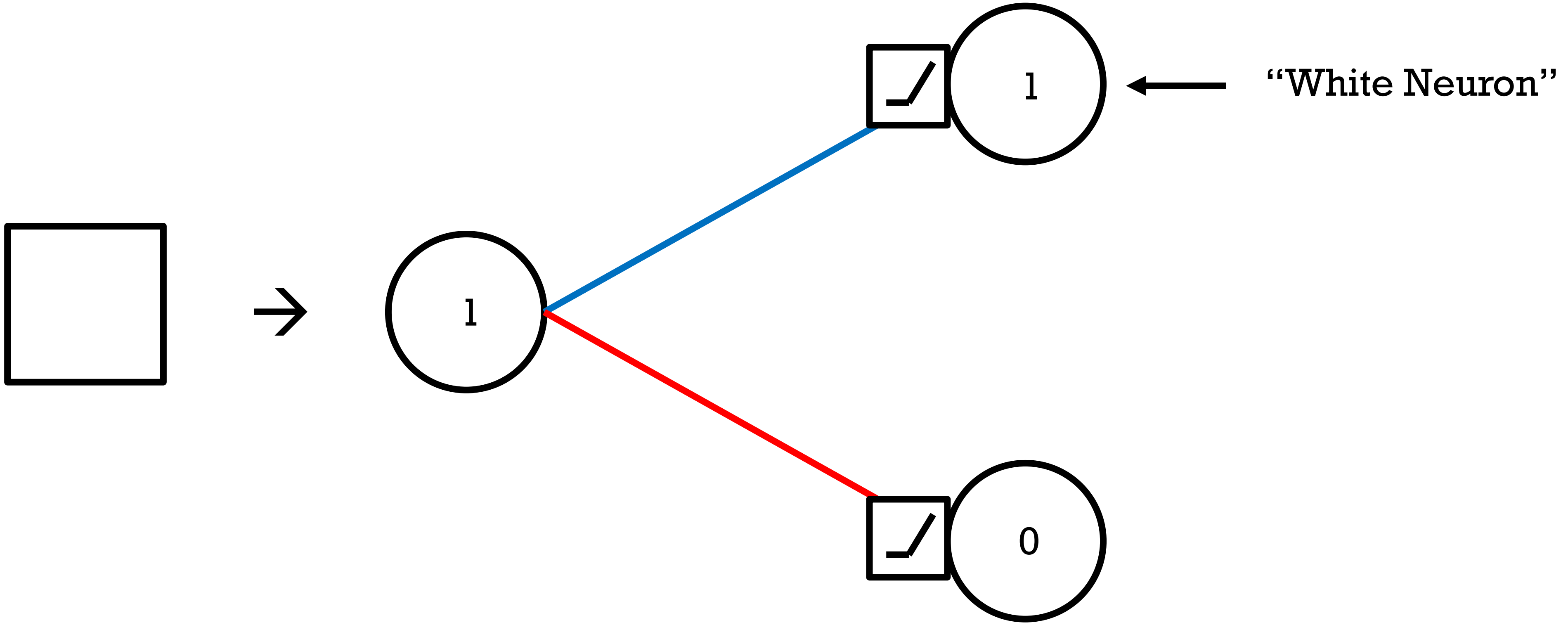
BASIC NEURAL NETWORK – 1 PIXEL



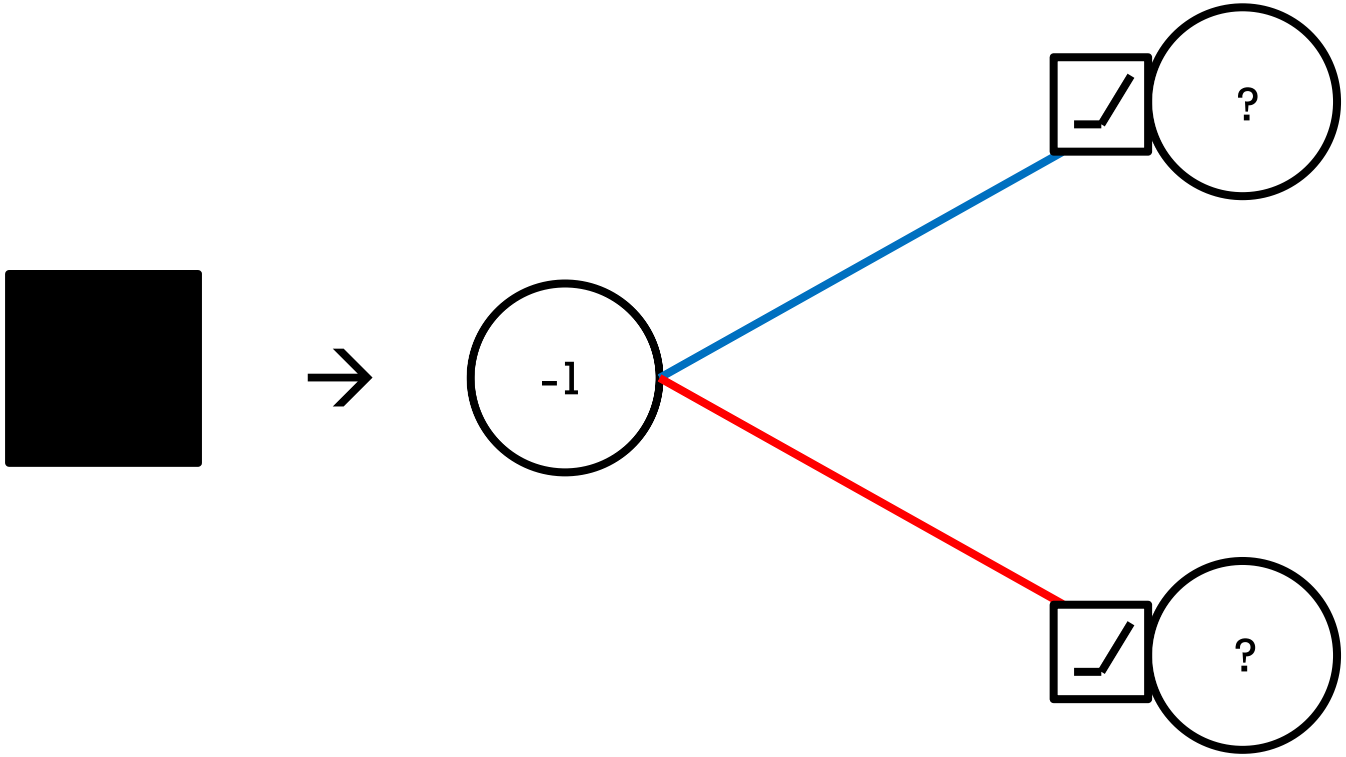
BASIC NEURAL NETWORK – 1 PIXEL



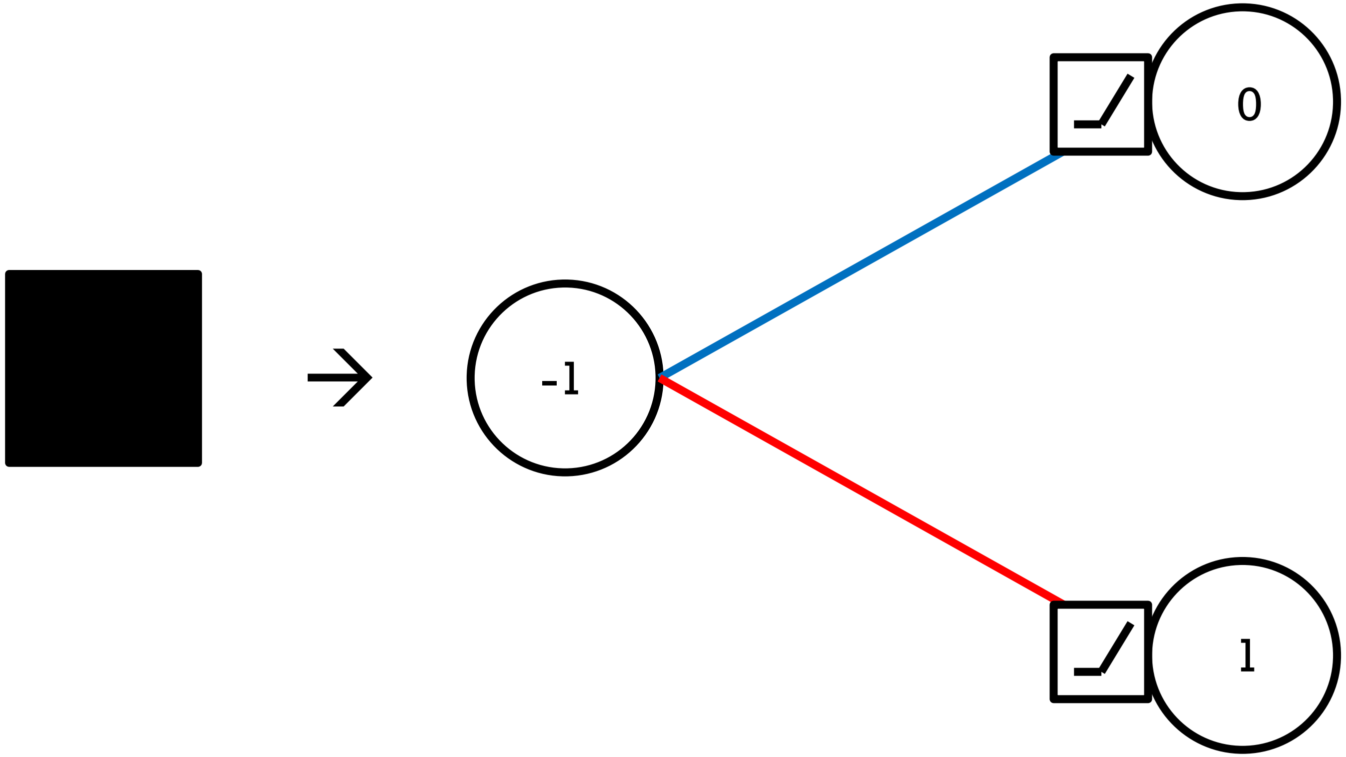
BASIC NEURAL NETWORK – 1 PIXEL



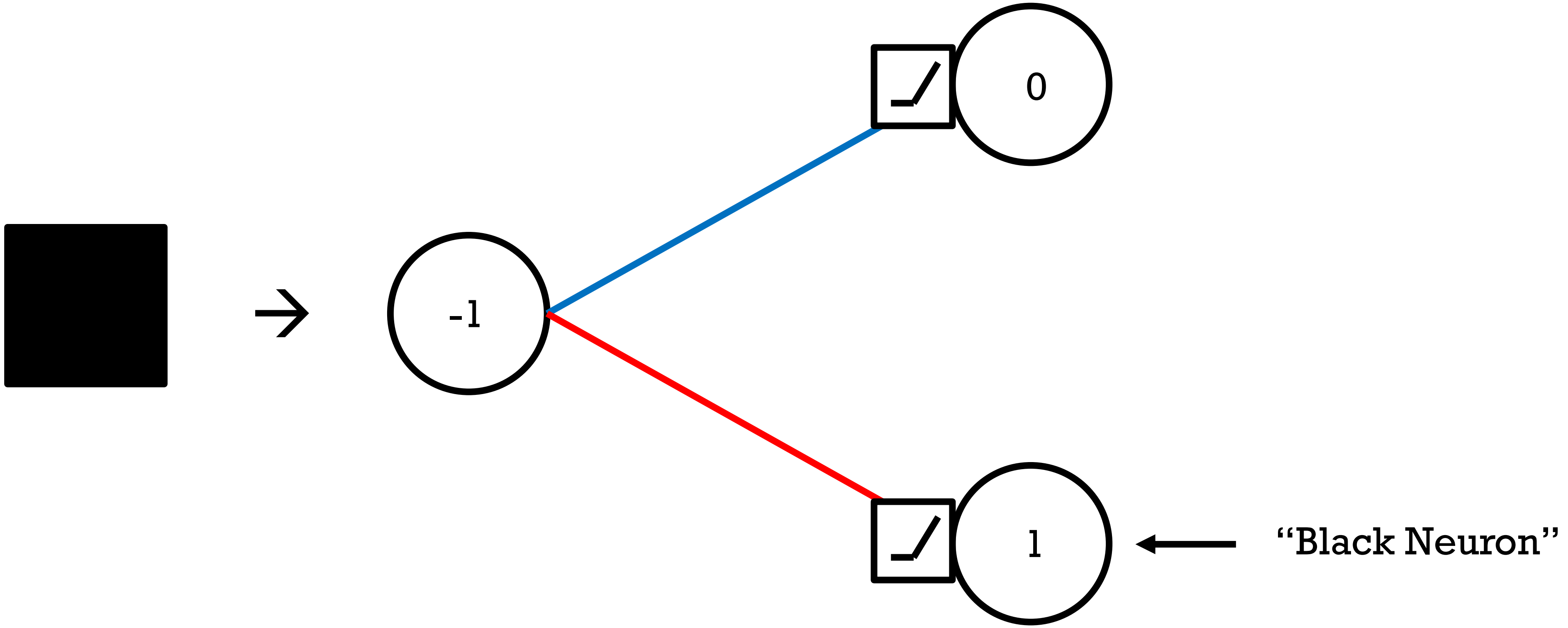
BASIC NEURAL NETWORK – 1 PIXEL



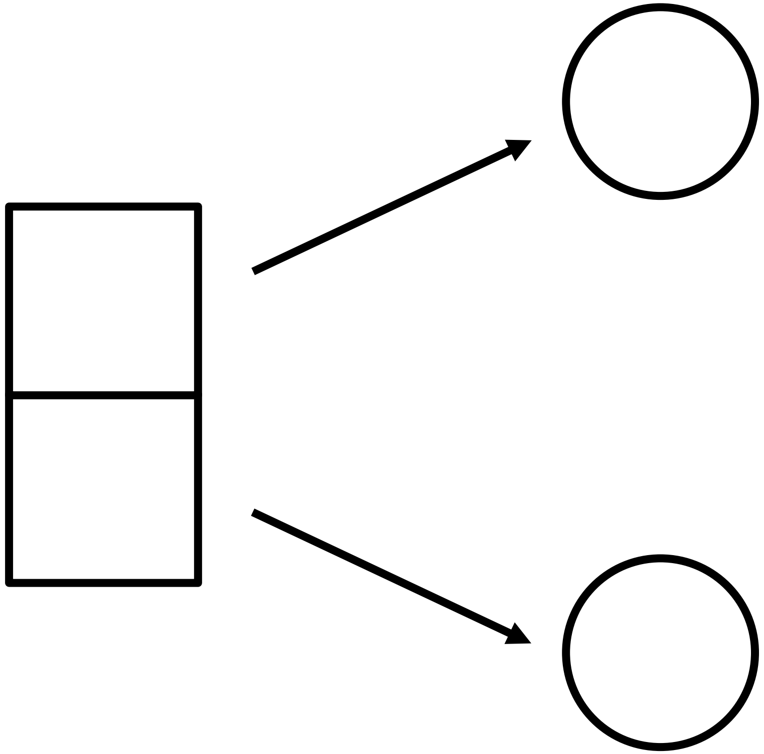
BASIC NEURAL NETWORK – 1 PIXEL



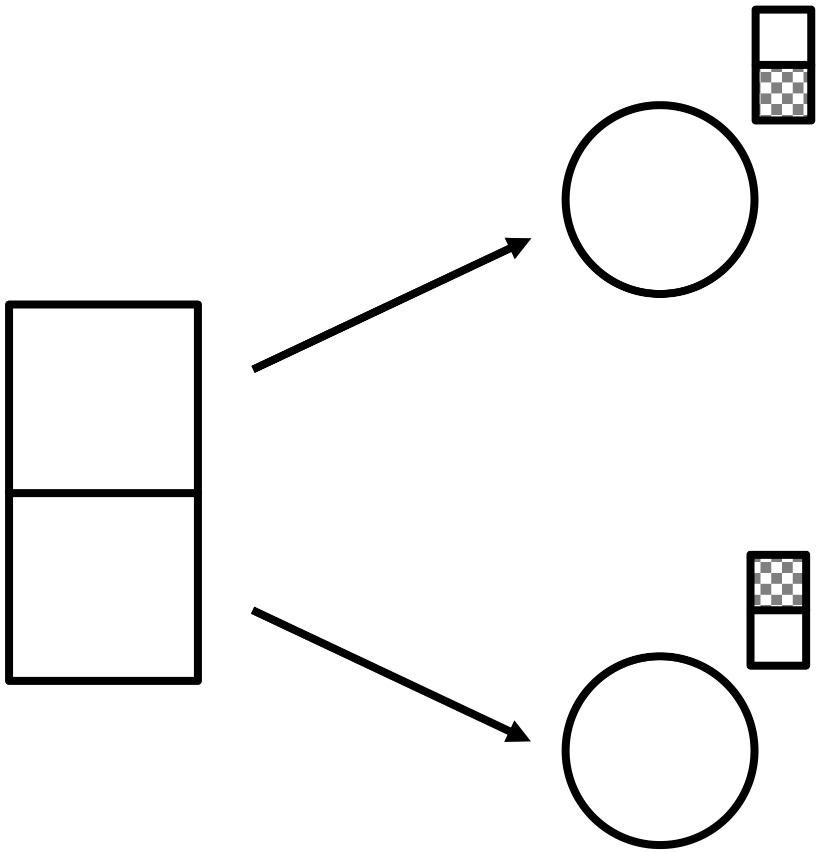
BASIC NEURAL NETWORK – 1 PIXEL



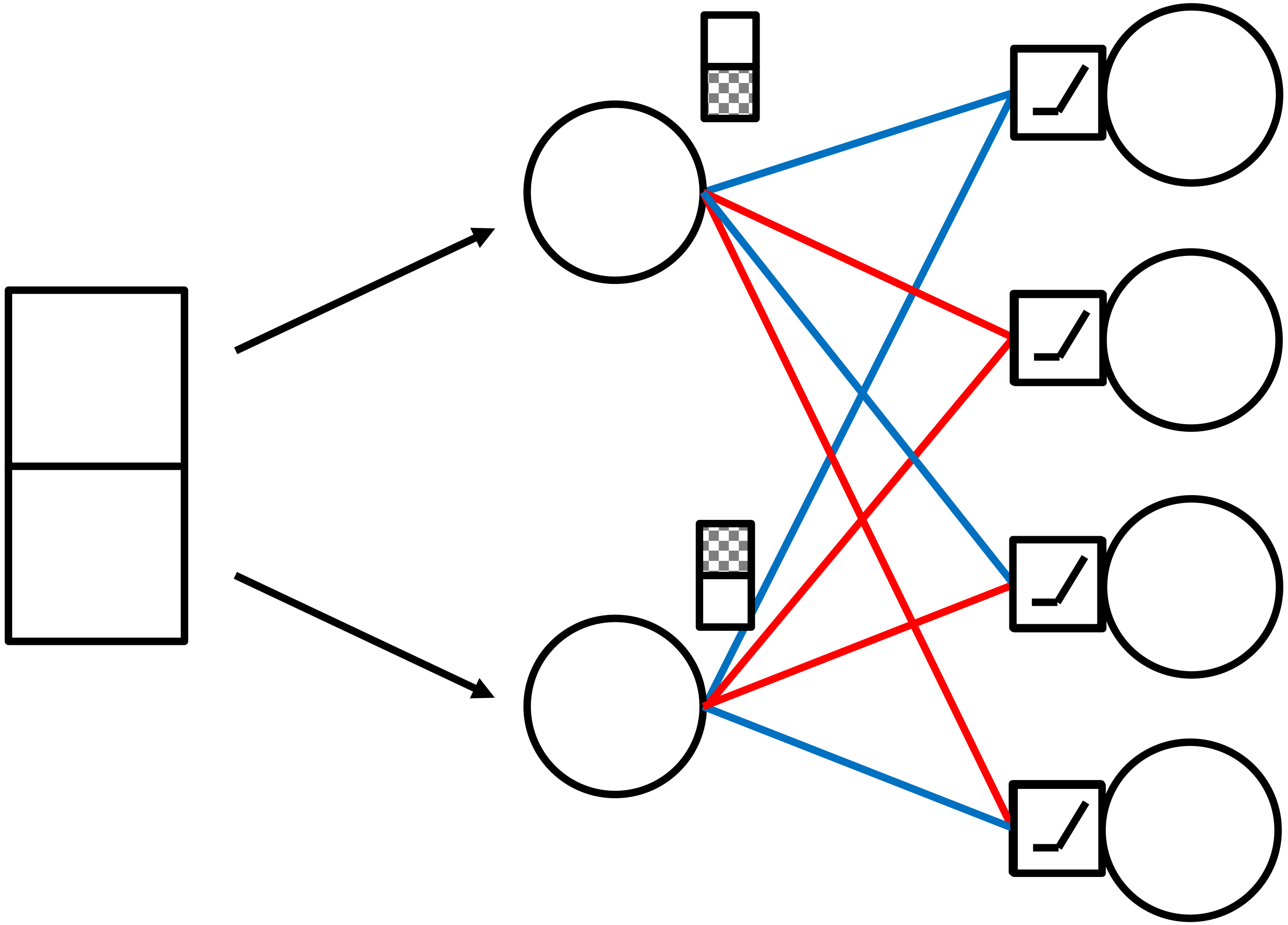
BASIC NEURAL NETWORK – 2 PIXELS



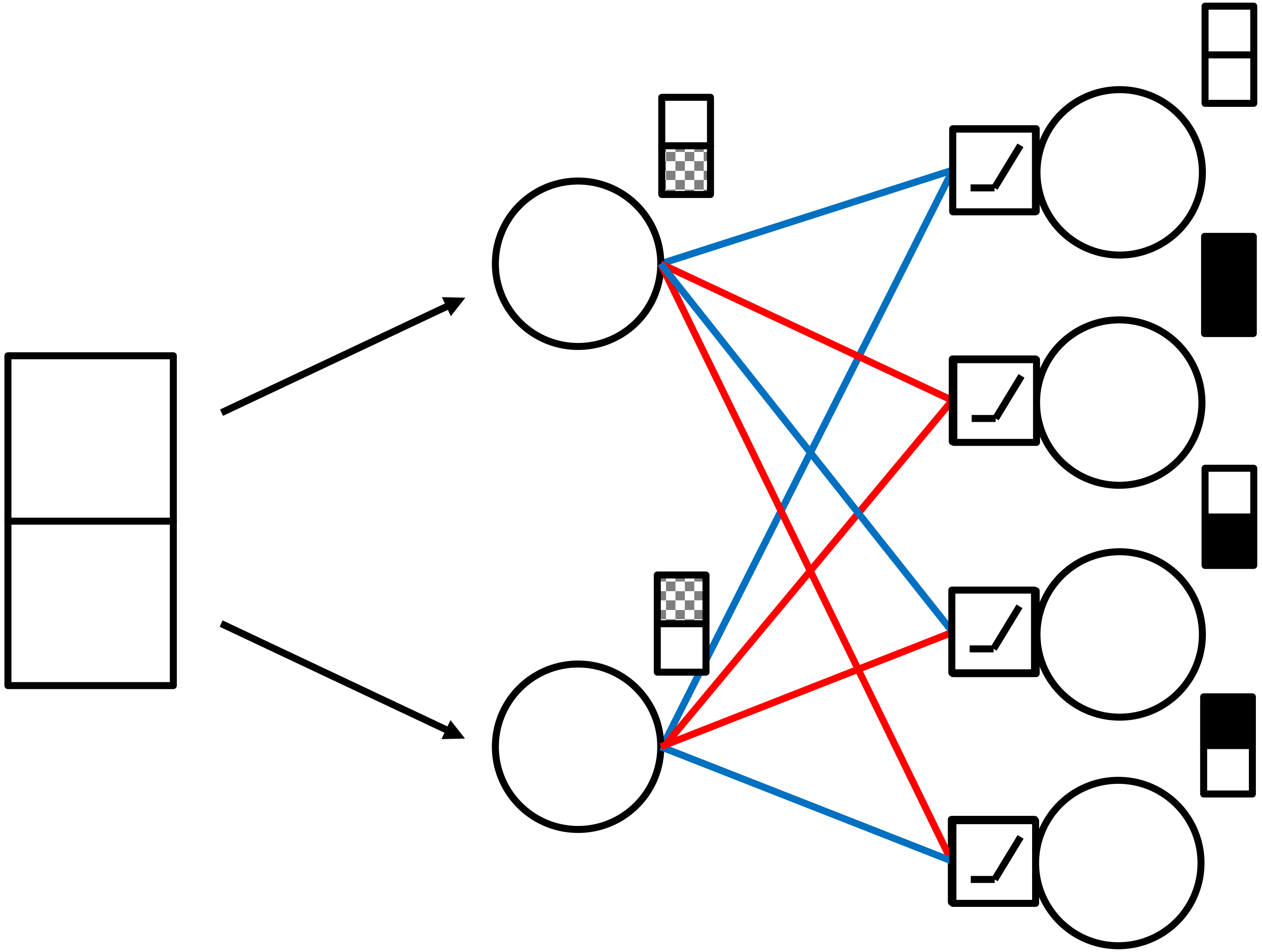
BASIC NEURAL NETWORK – 2 PIXELS



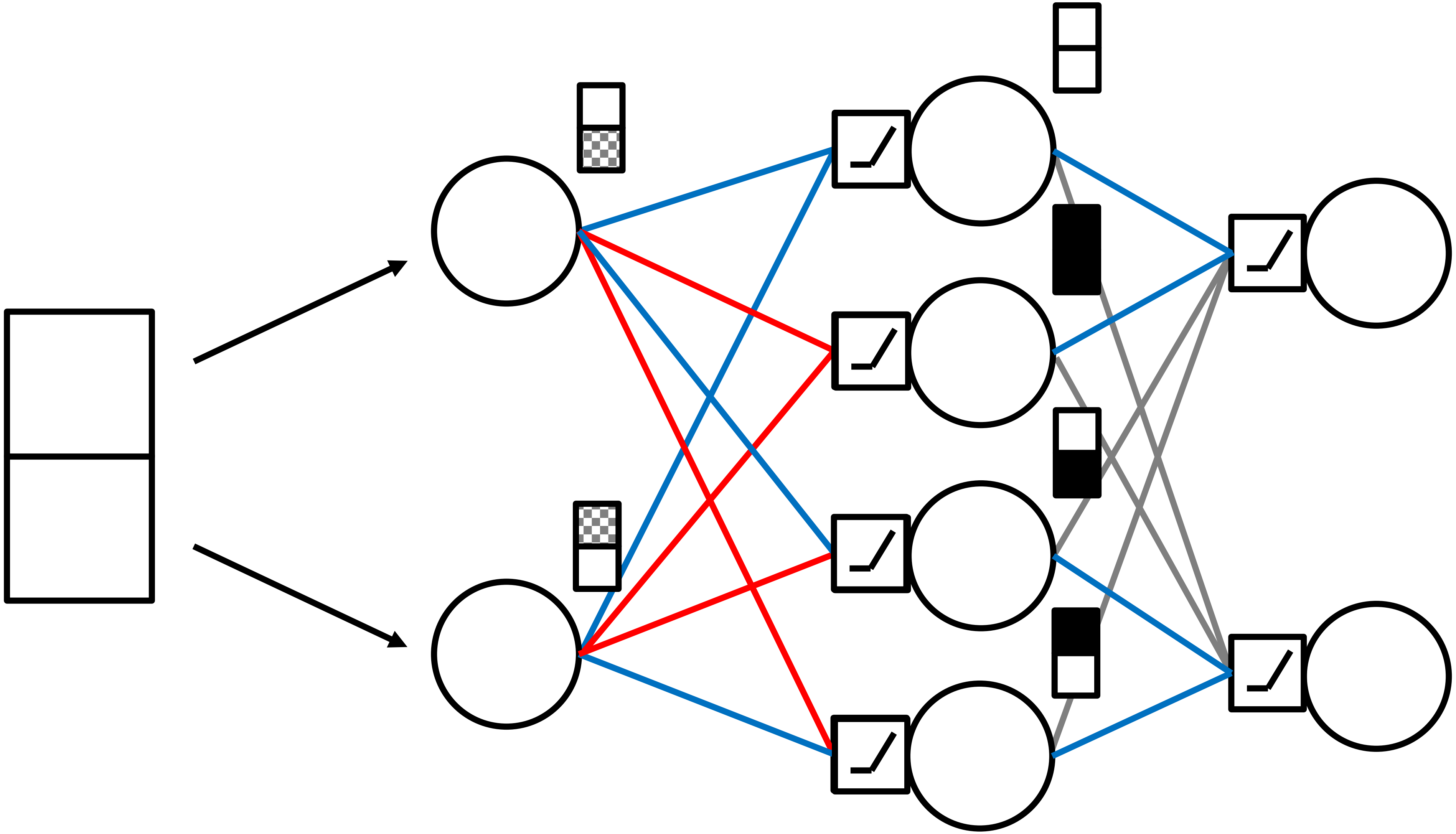
BASIC NEURAL NETWORK – 2 PIXELS



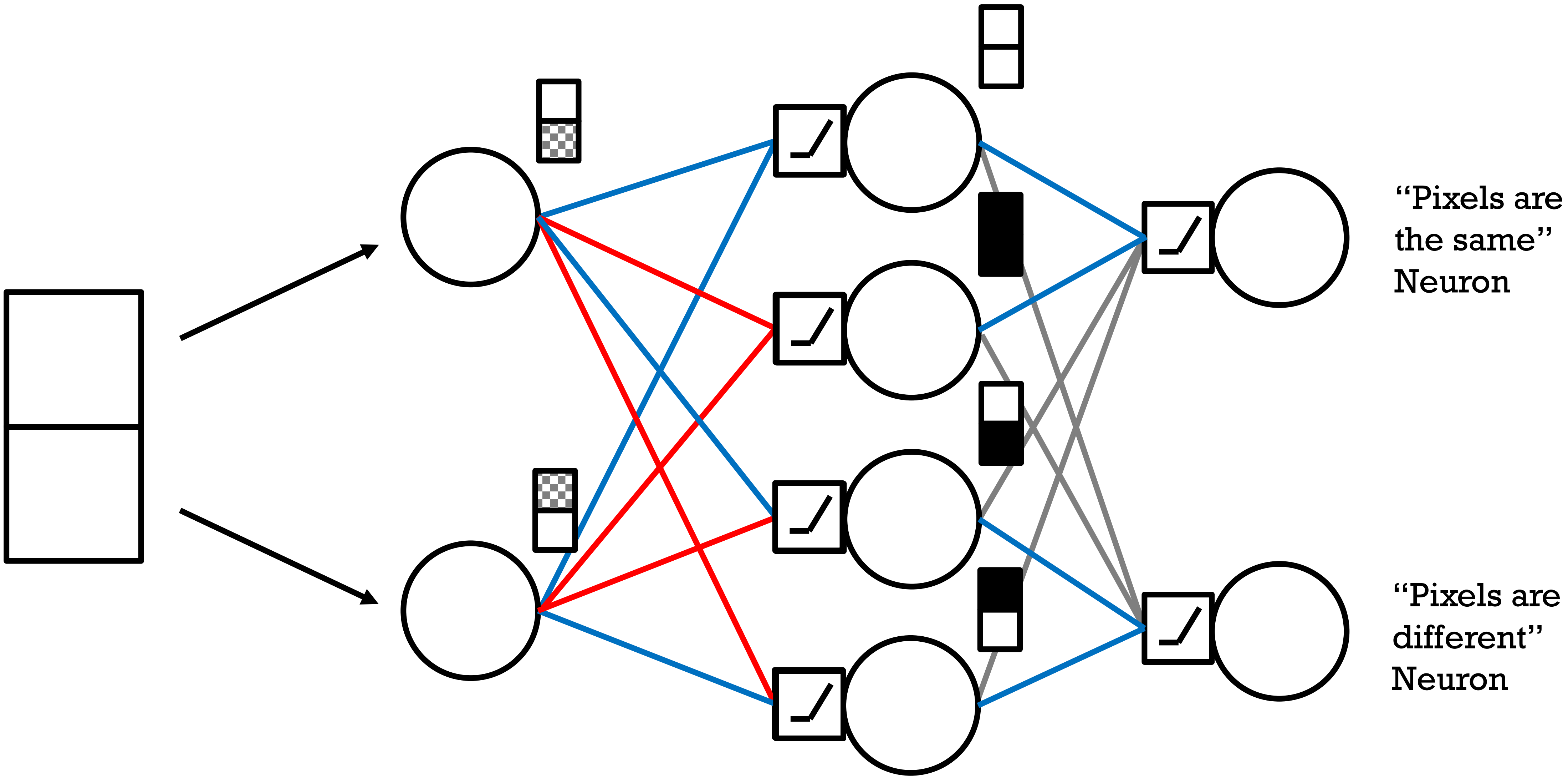
BASIC NEURAL NETWORK – 2 PIXELS



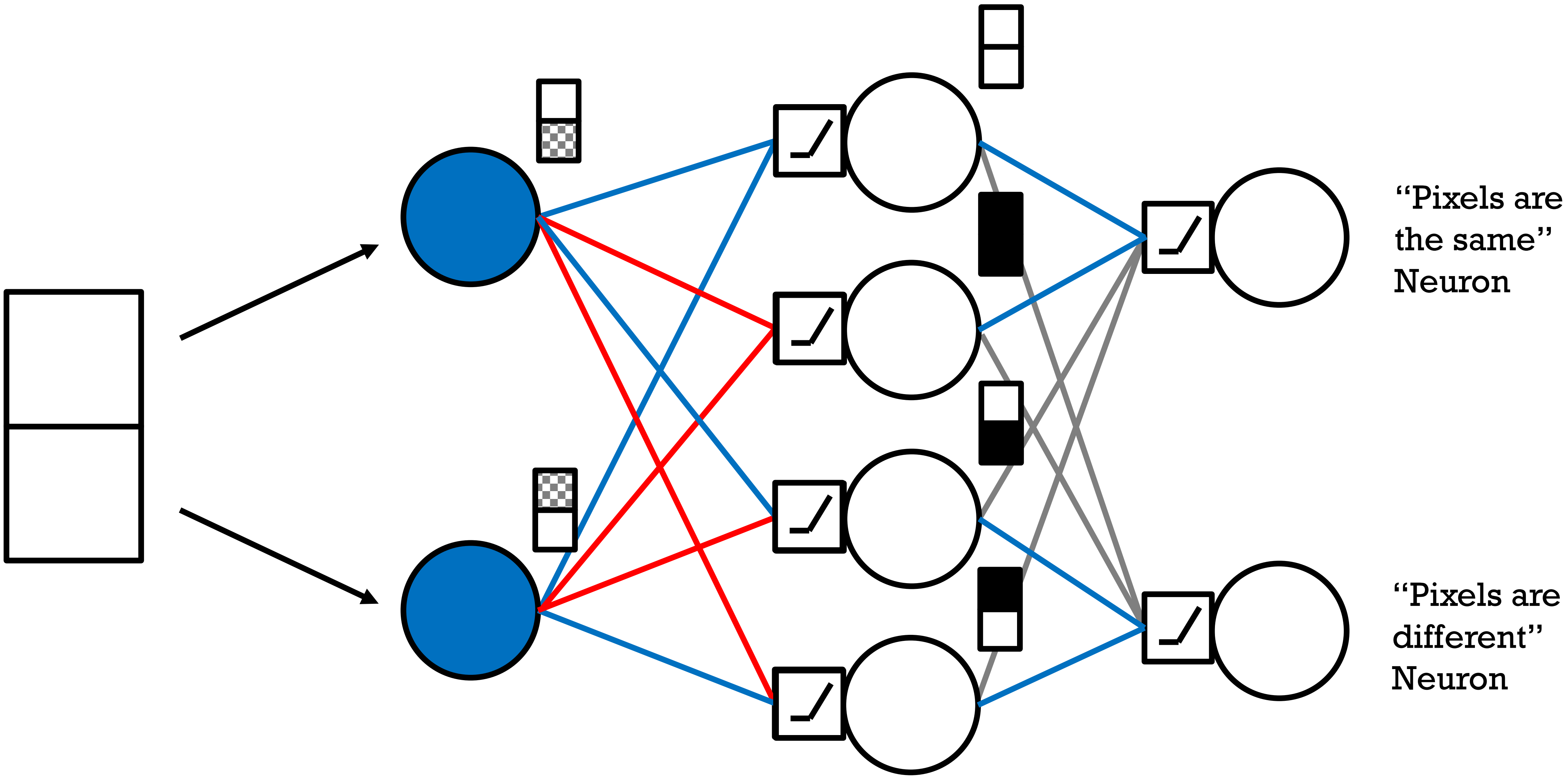
BASIC NEURAL NETWORK – 2 PIXELS



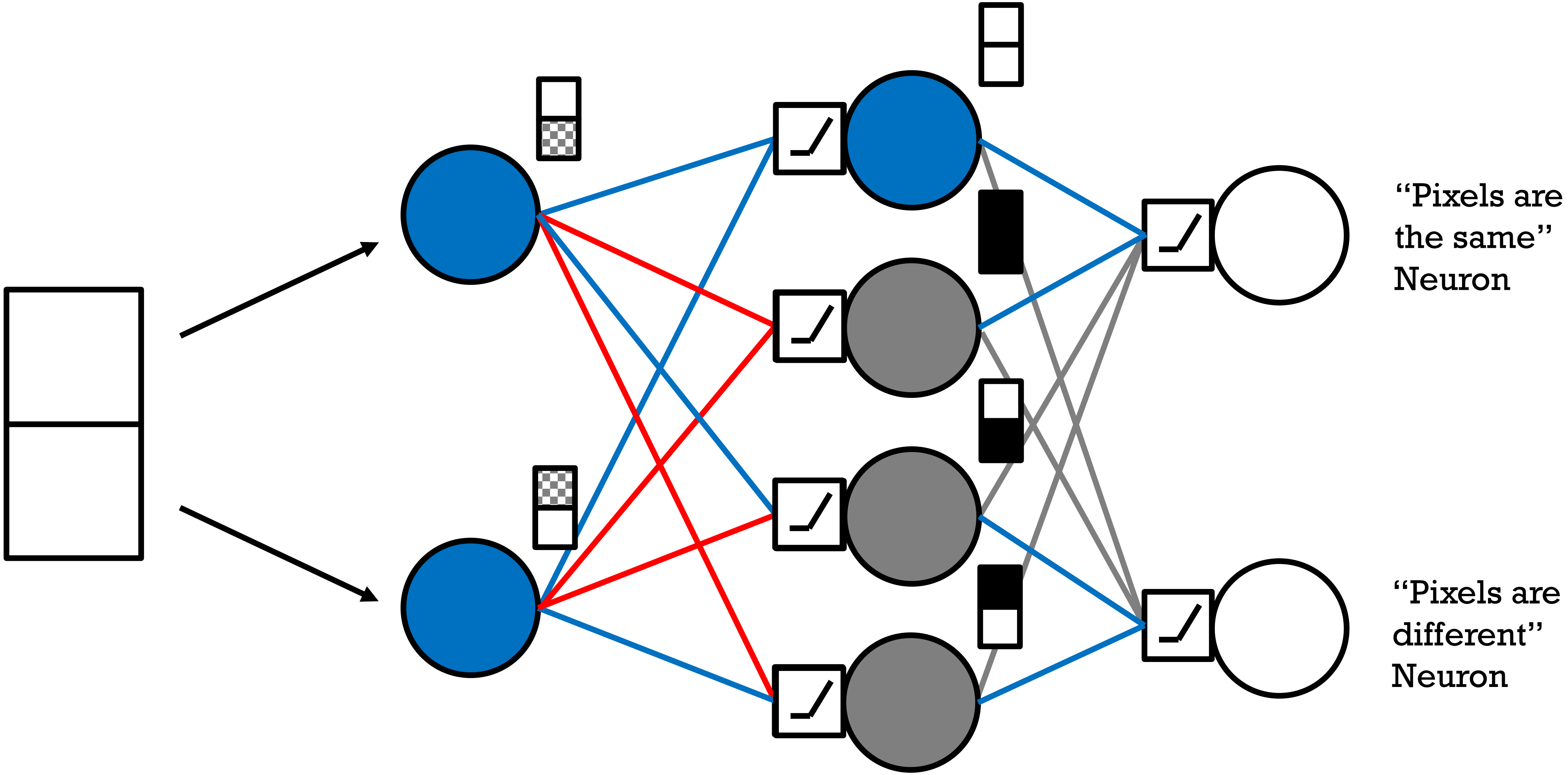
BASIC NEURAL NETWORK – 2 PIXELS



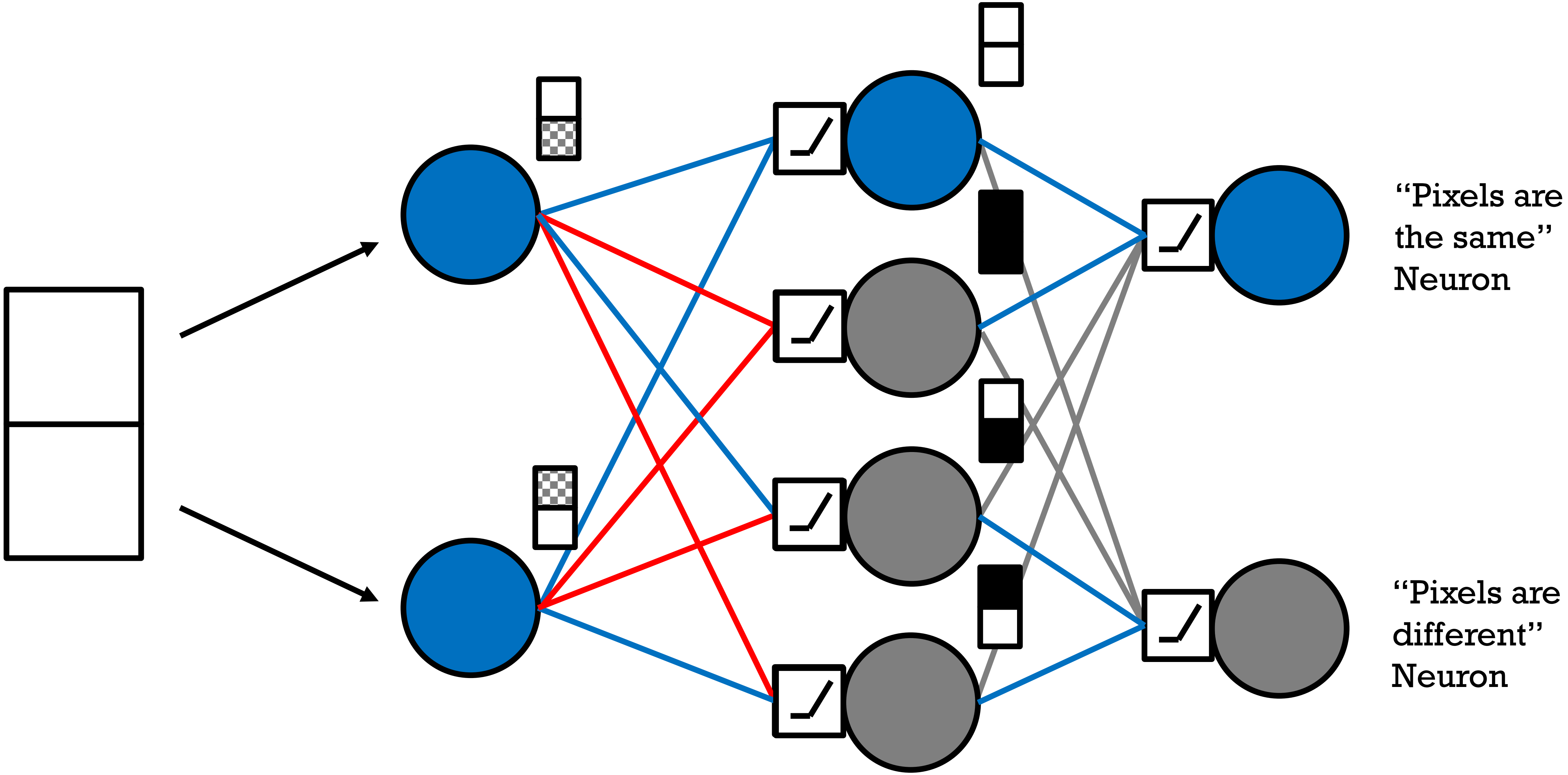
BASIC NEURAL NETWORK – 2 PIXELS



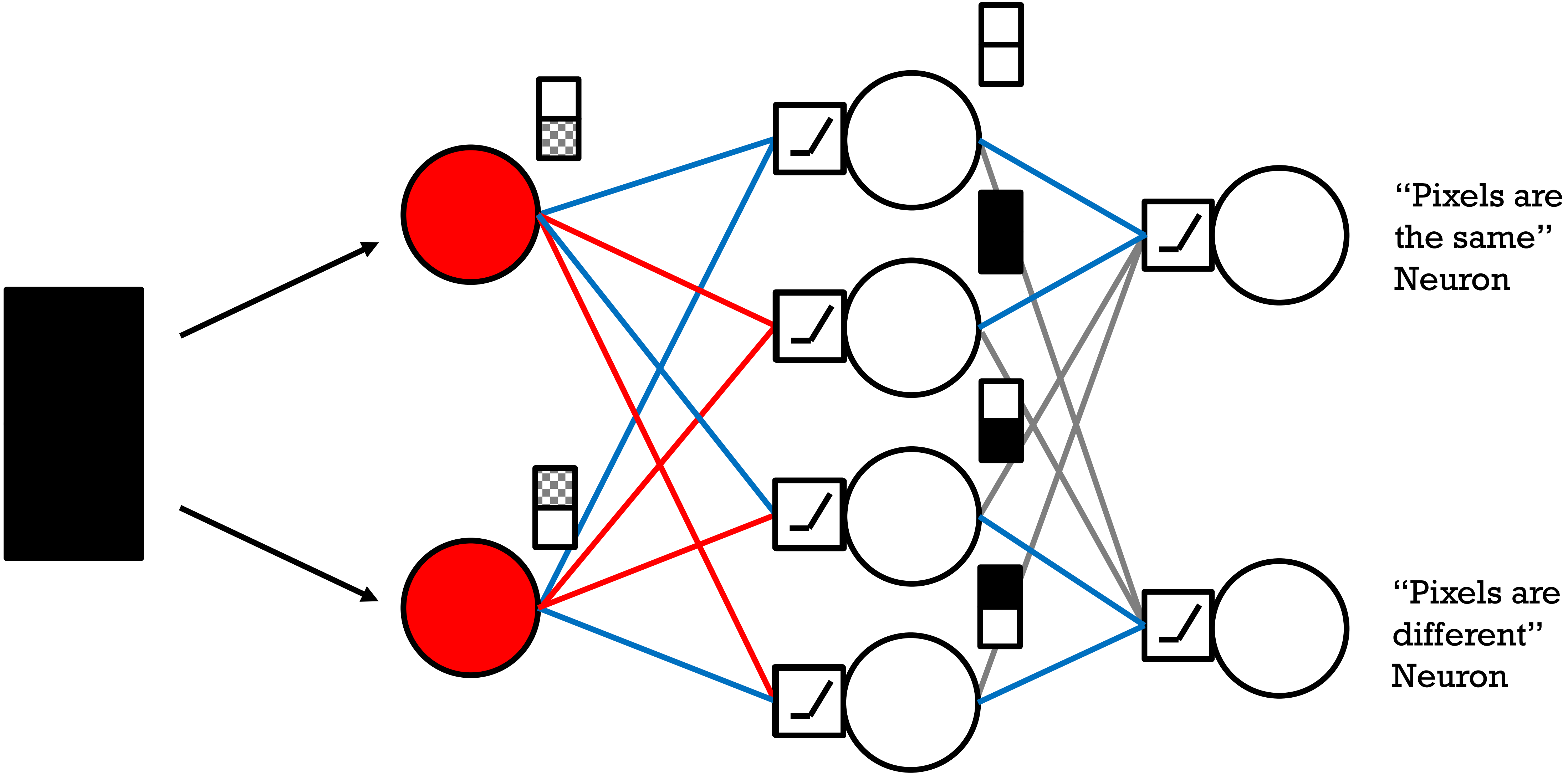
BASIC NEURAL NETWORK – 2 PIXELS



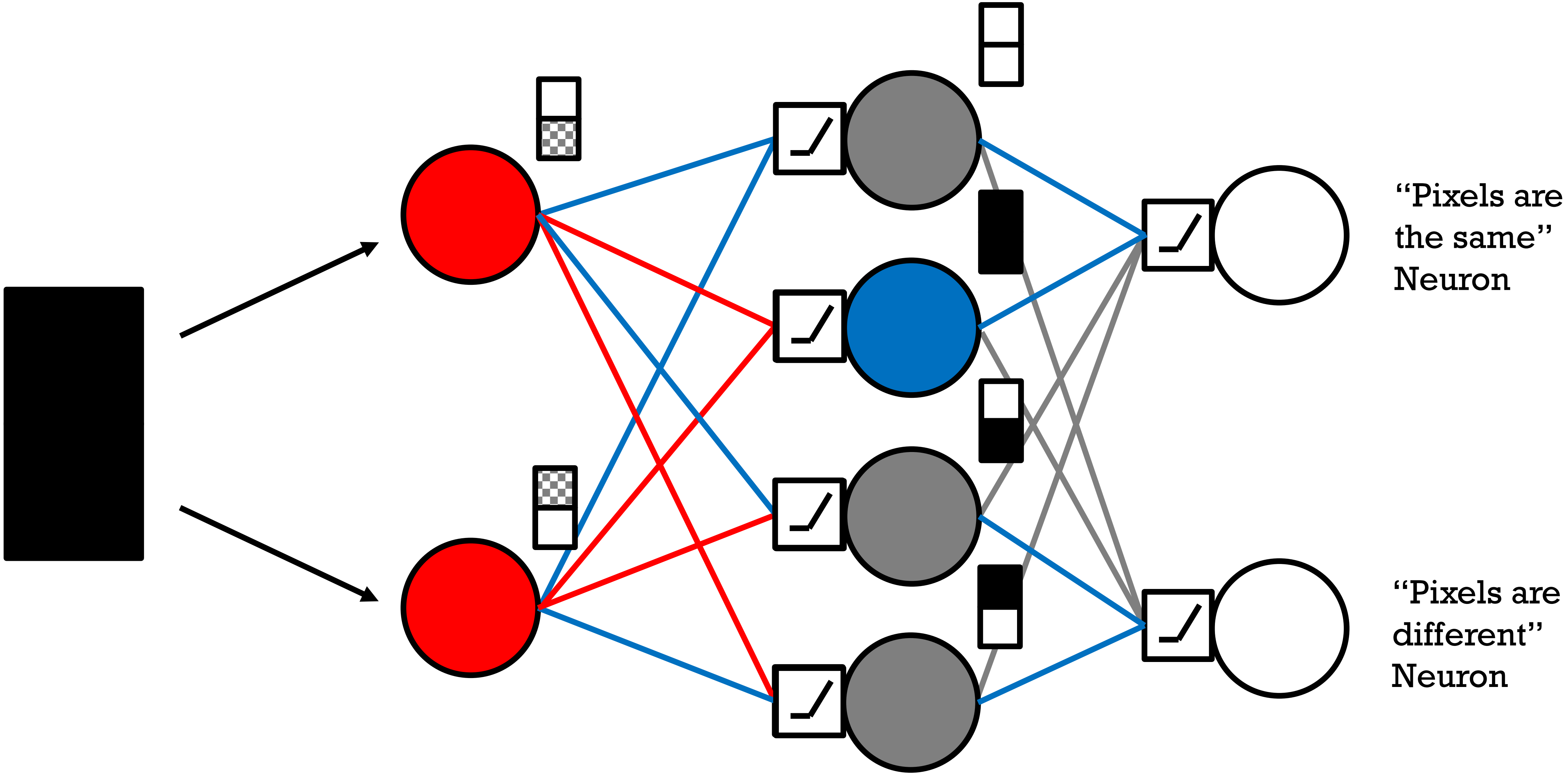
BASIC NEURAL NETWORK – 2 PIXELS



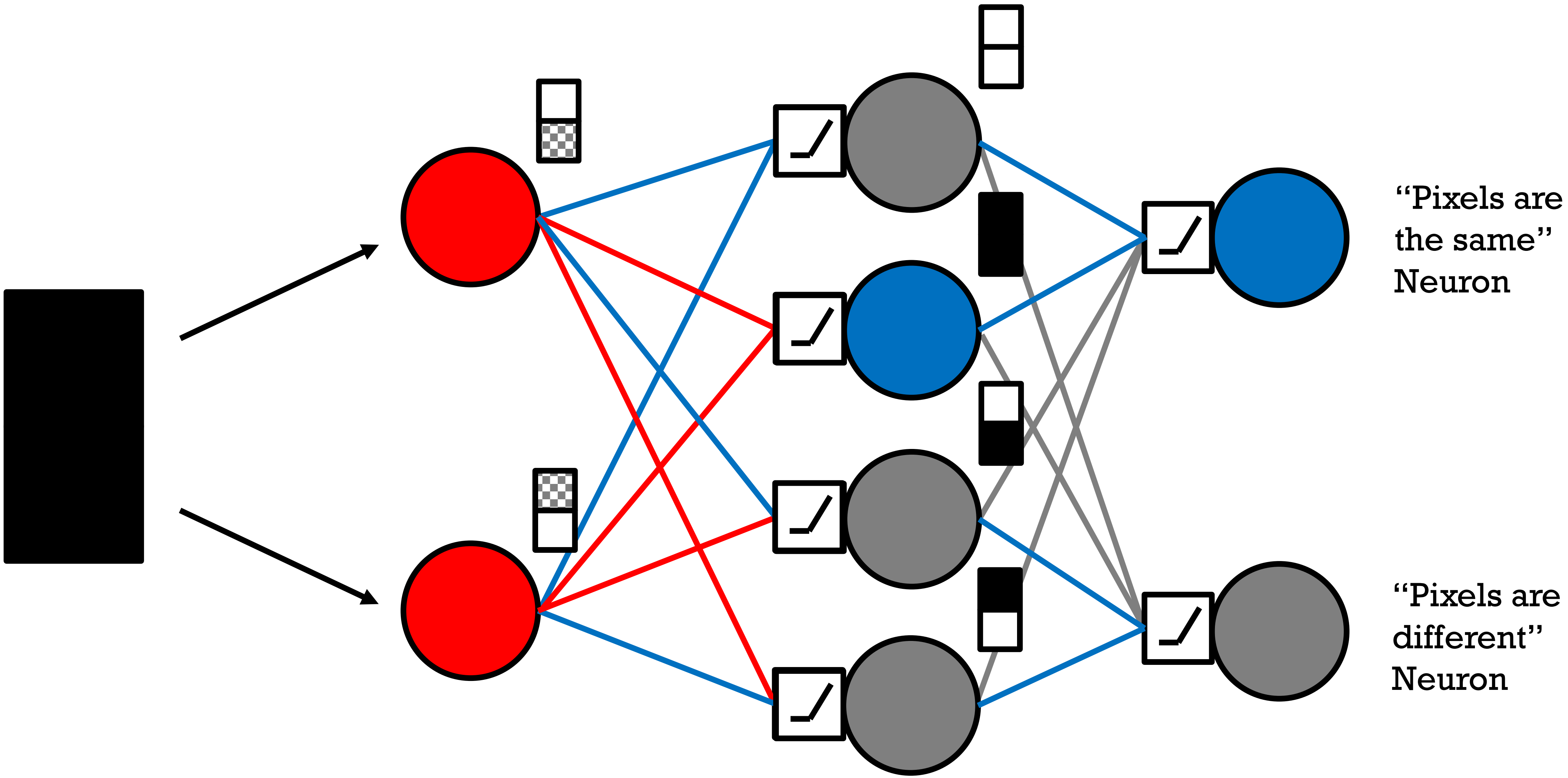
BASIC NEURAL NETWORK – 2 PIXELS



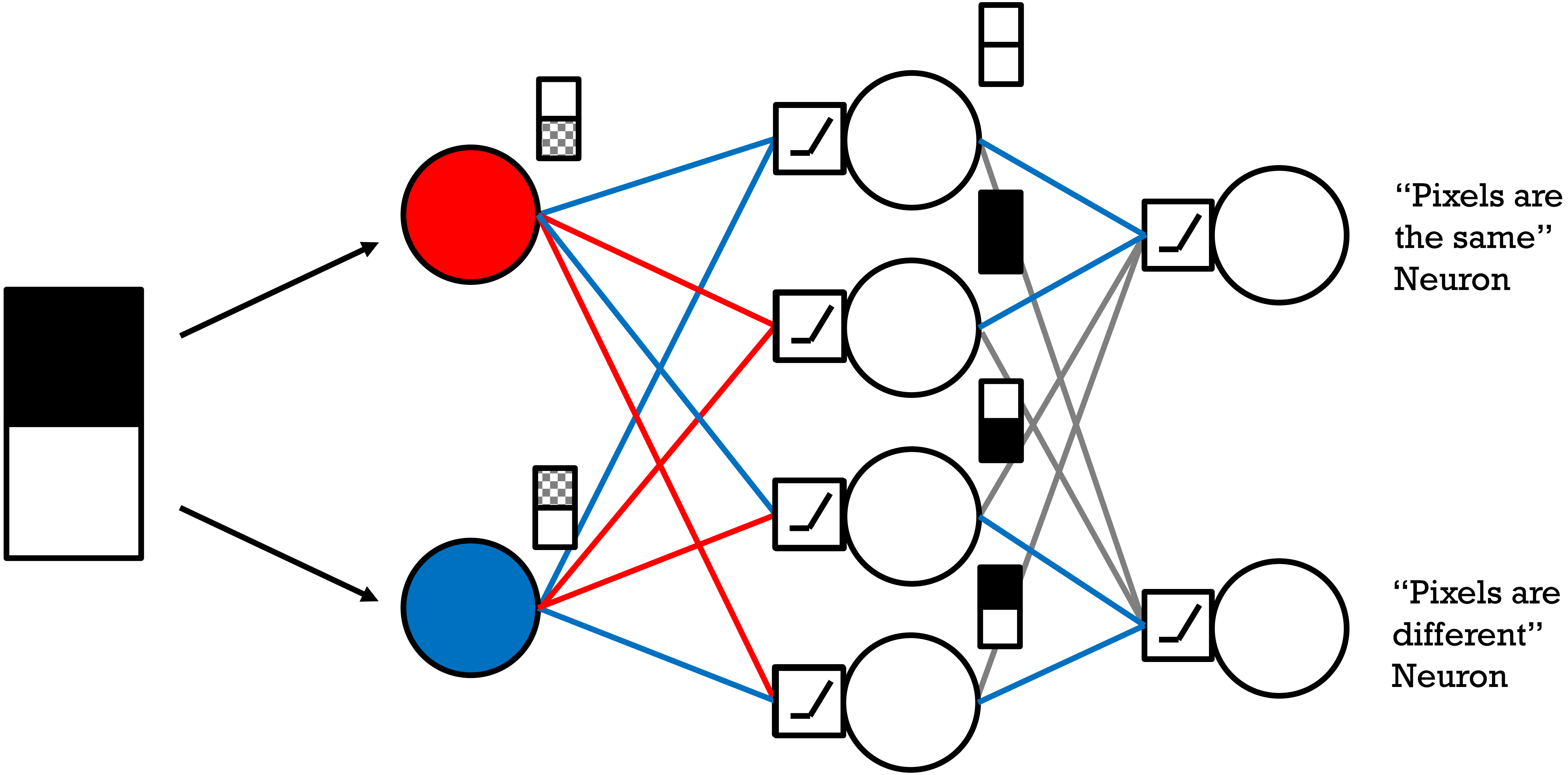
BASIC NEURAL NETWORK – 2 PIXELS



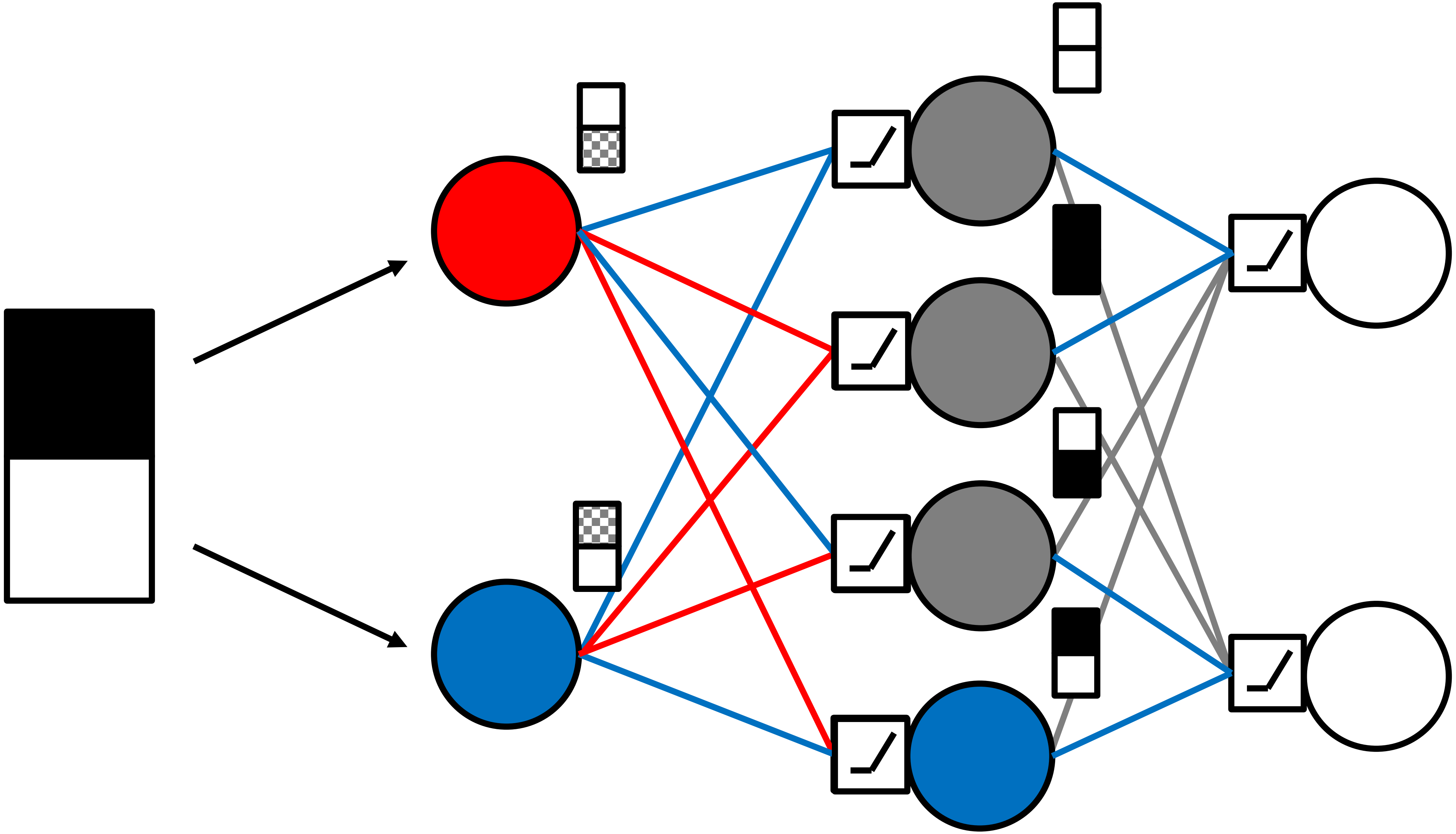
BASIC NEURAL NETWORK – 2 PIXELS



BASIC NEURAL NETWORK – 2 PIXELS



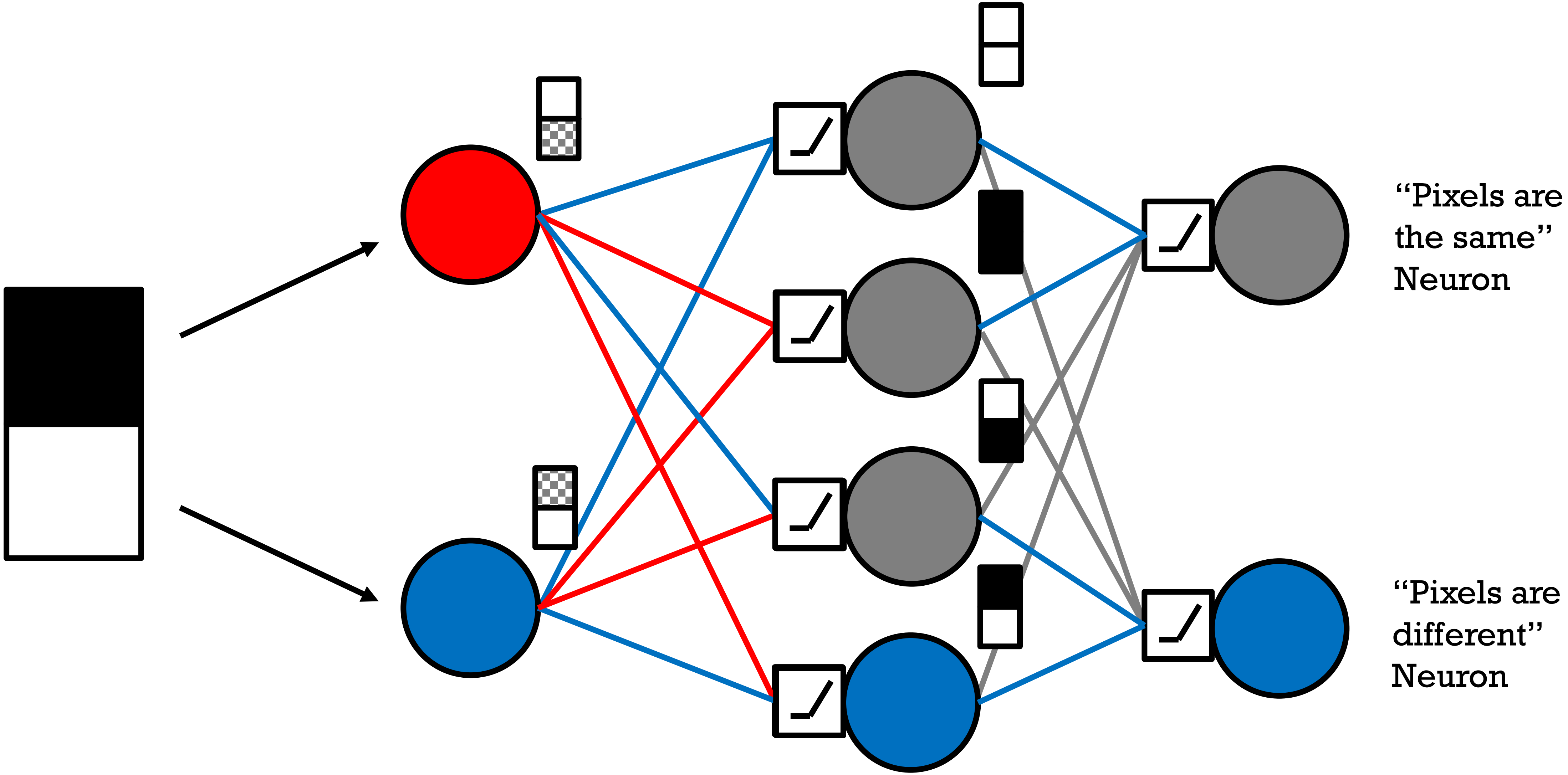
BASIC NEURAL NETWORK – 2 PIXELS



"Pixels are the same" Neuron

"Pixels are different" Neuron

BASIC NEURAL NETWORK – 2 PIXELS



OVERVIEW

- Why ASDNet?
- Neurons → Neural Network
- Basic Neural Network Examples
- Weight Tuning/Optimization
- ASDNet
- Visualization
- Results
- Future Directions

IN REALITY

- Best answer is not given
- Weights are not just -1, 0, or 1
- Many more pixels/inputs
- Many more layers
- Many more nodes within each layer

IN REALITY

- Best answer is not given
- Weights are not just -1, 0, or 1
- Many more pixels/inputs
- Many more layers
- Many more nodes within each layer

IN REALITY

- Best answer is not given
- Weights are not just -1, 0, or 1
- Many more pixels/inputs
- Many more layers
- Many more nodes within each layer

IN REALITY

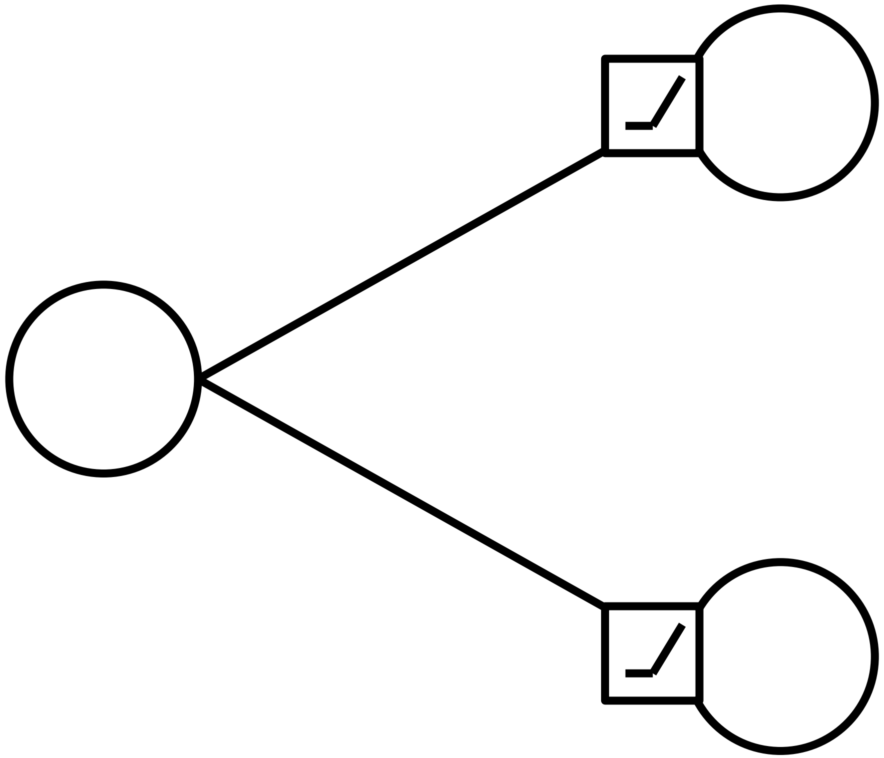
- Best answer is not given
- Weights are not just -1, 0, or 1
- Many more pixels/inputs
- Many more layers
- Many more nodes within each layer

IN REALITY

- Best answer is not given
- Weights are not just -1, 0, or 1
- Many more pixels/inputs
- Many more layers
- Many more nodes within each layer

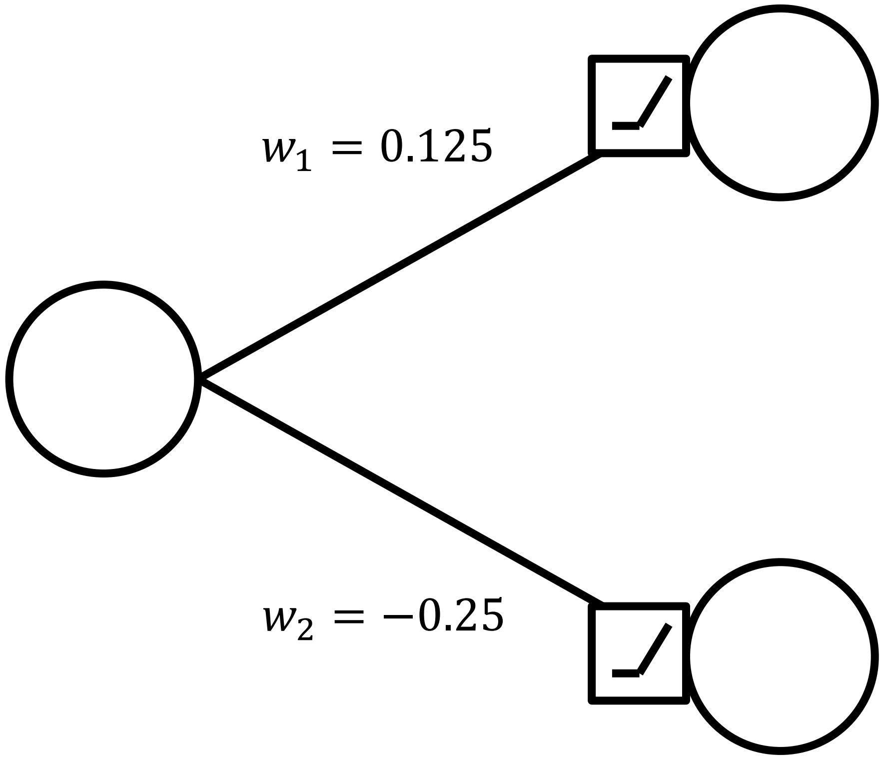
HOW IT REALLY WORKS

1. Set up neural network architecture
2. Initialize weights to small values with mean 0
3. TRAINING



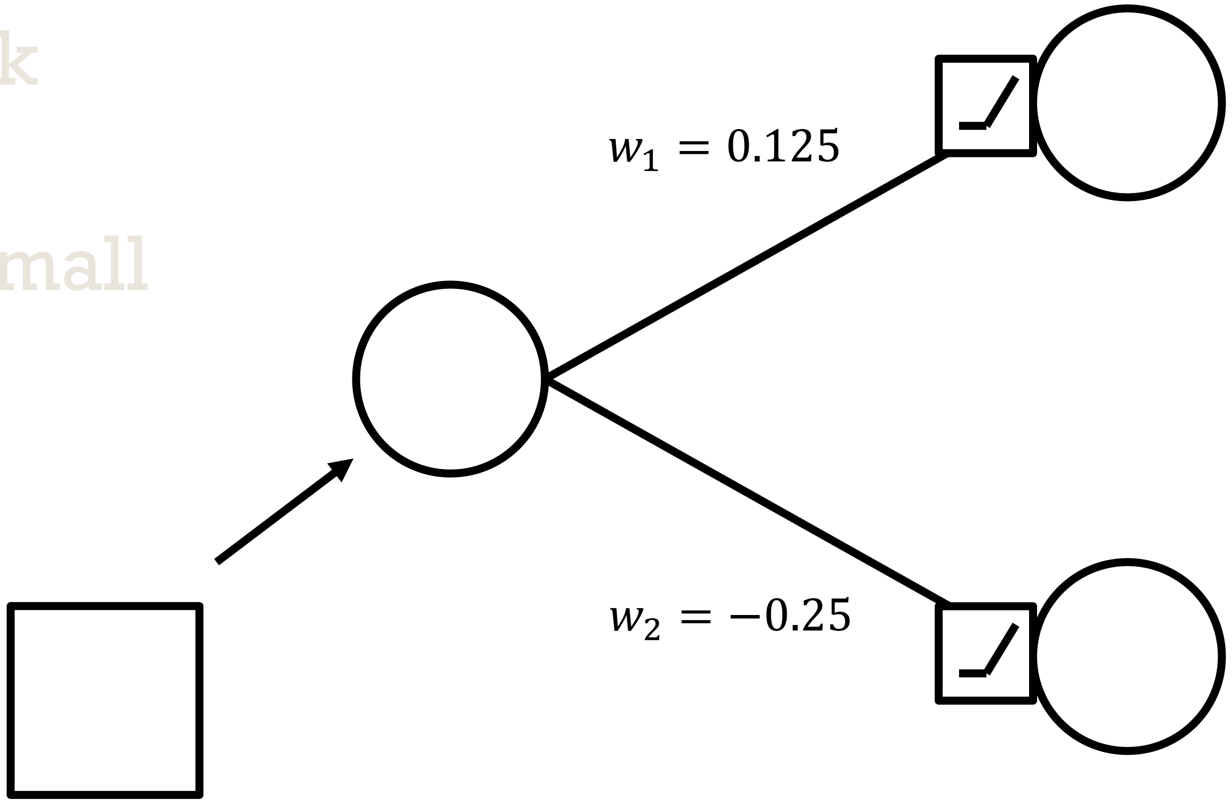
HOW IT REALLY WORKS

1. Set up neural network architecture
2. Initialize weights to small values with mean 0
3. TRAINING



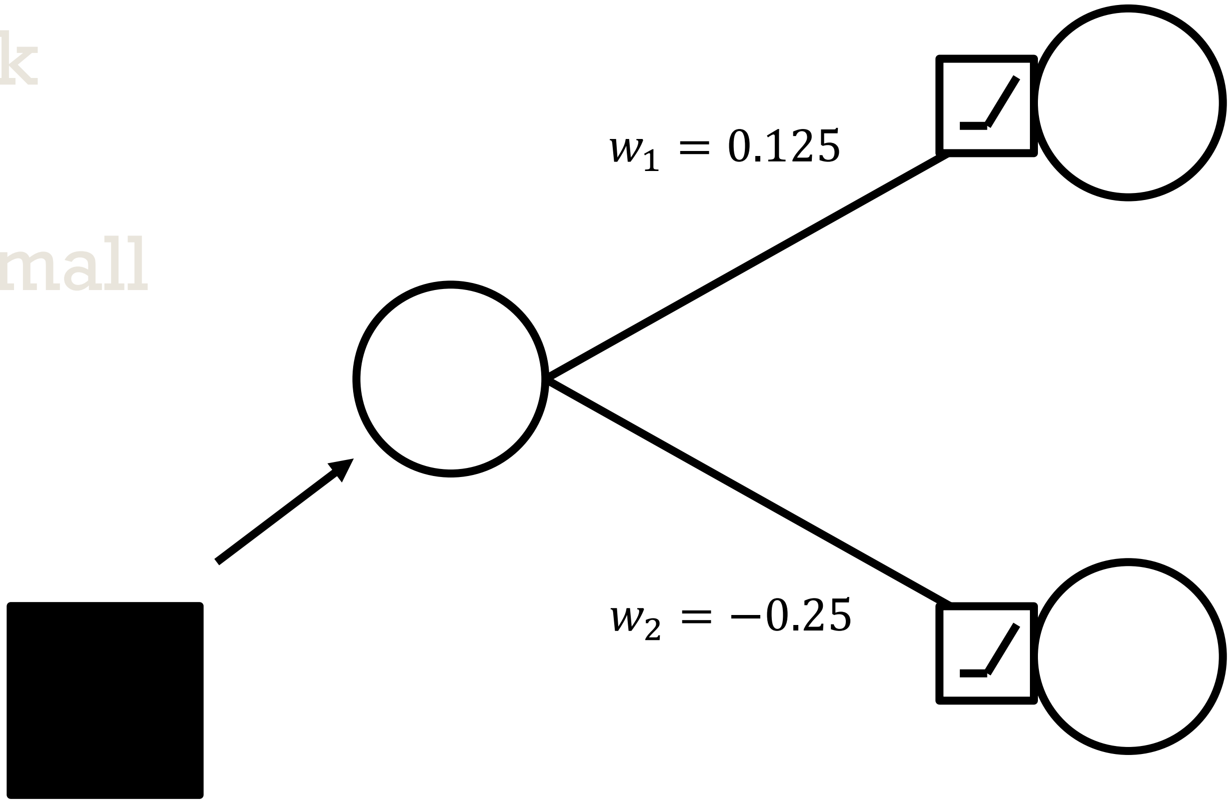
HOW IT REALLY WORKS

1. Set up neural network architecture
2. Initialize weights to small values with mean 0
3. TRAINING

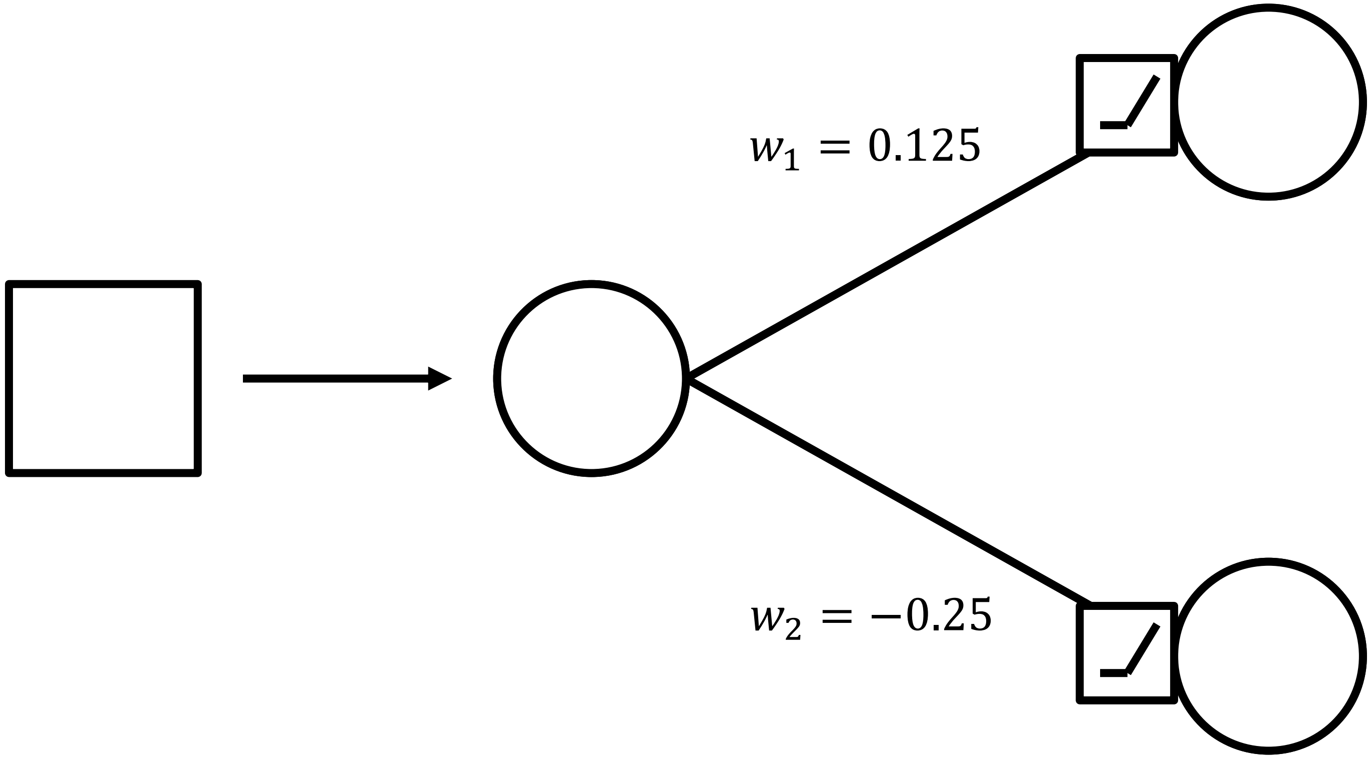


HOW IT REALLY WORKS

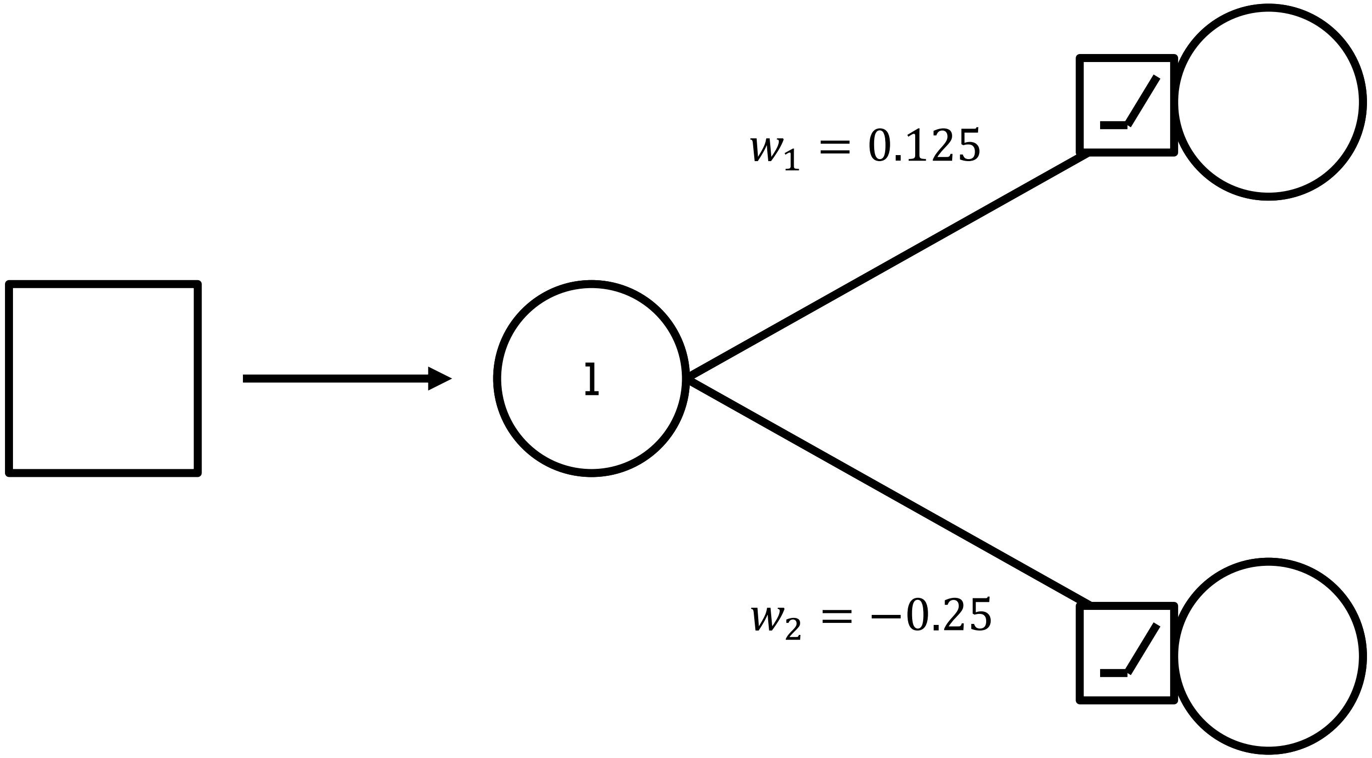
1. Set up neural network architecture
2. Initialize weights to small values with mean 0
3. TRAINING



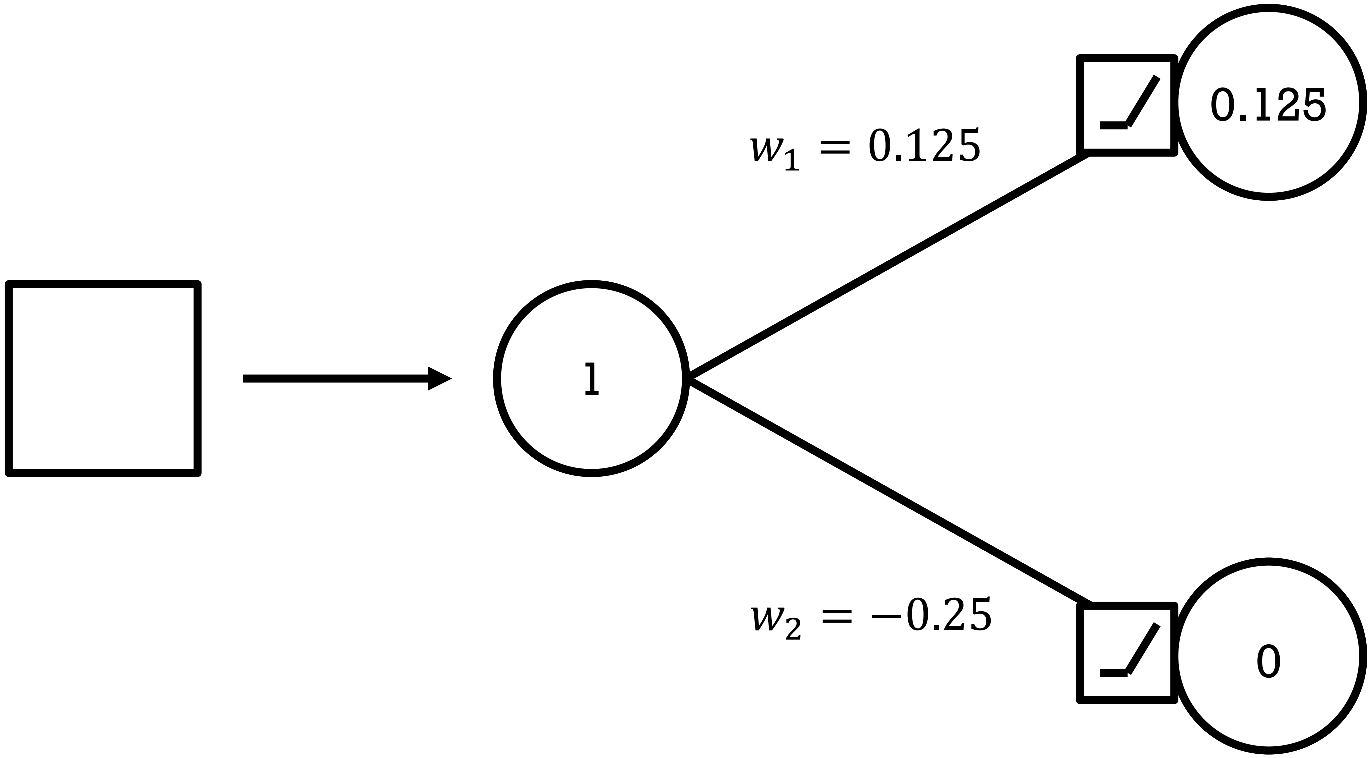
TRAINING



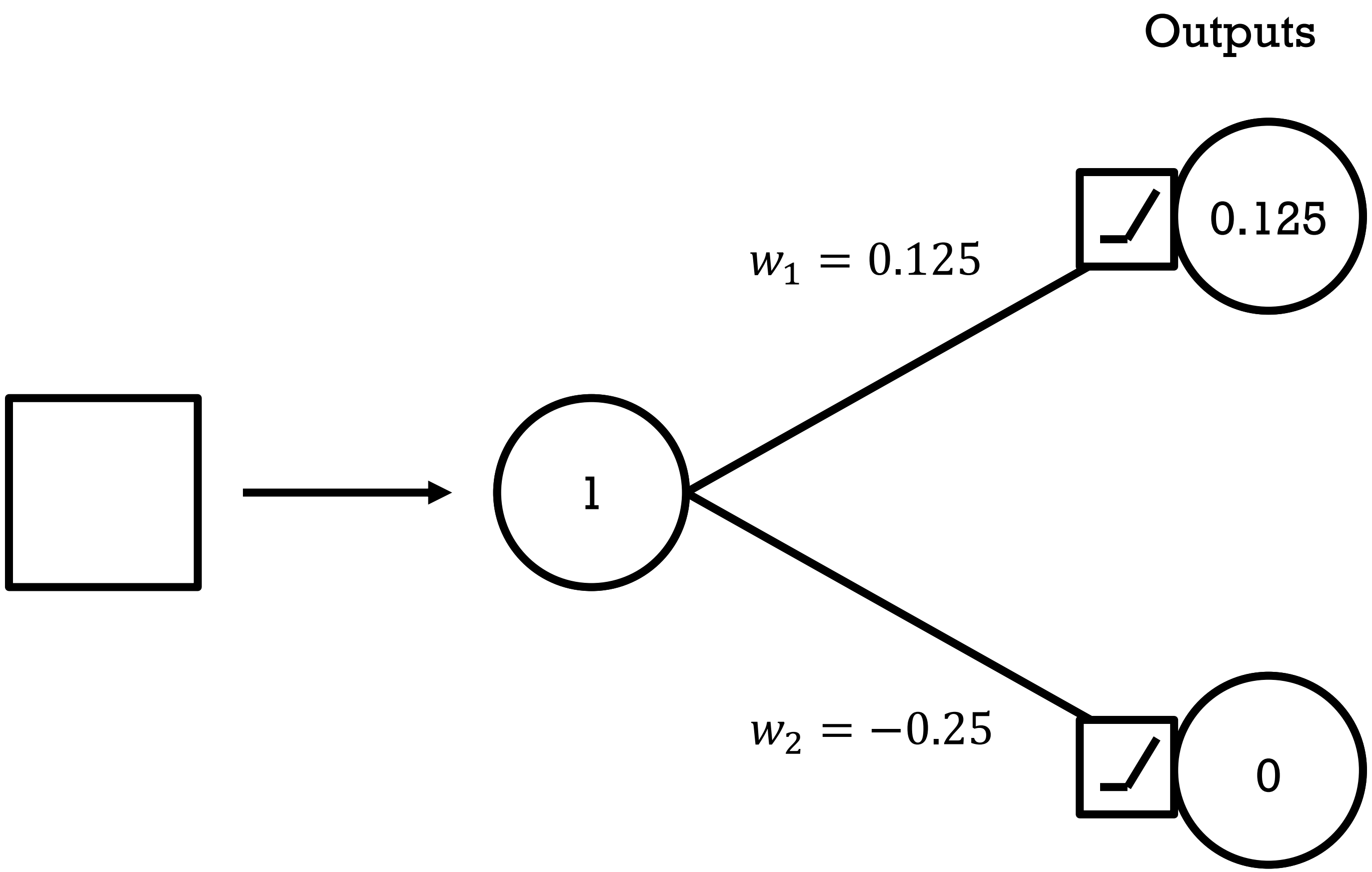
TRAINING



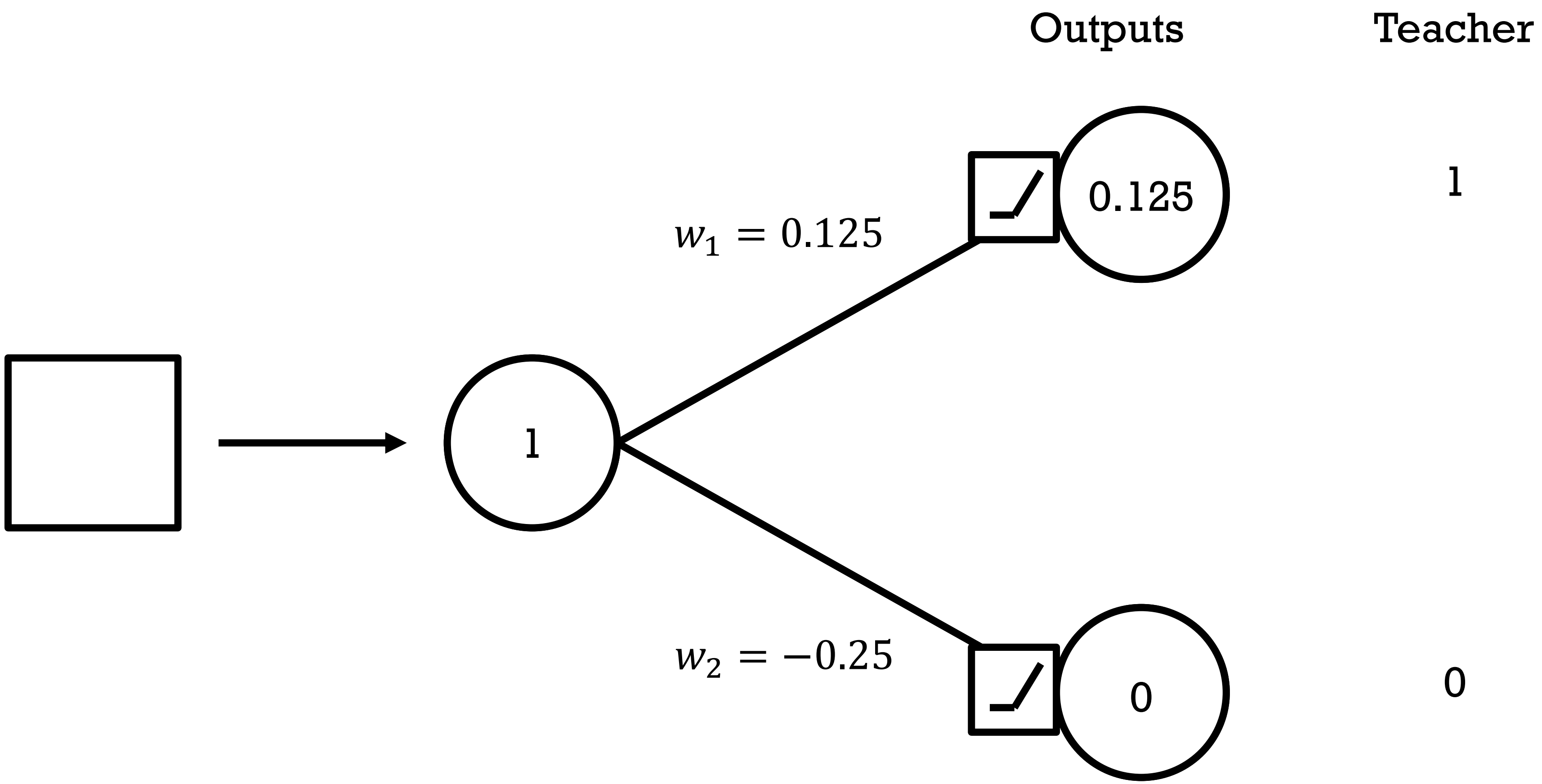
TRAINING



TRAINING



TRAINING



ERROR

- Sum of Squared Error

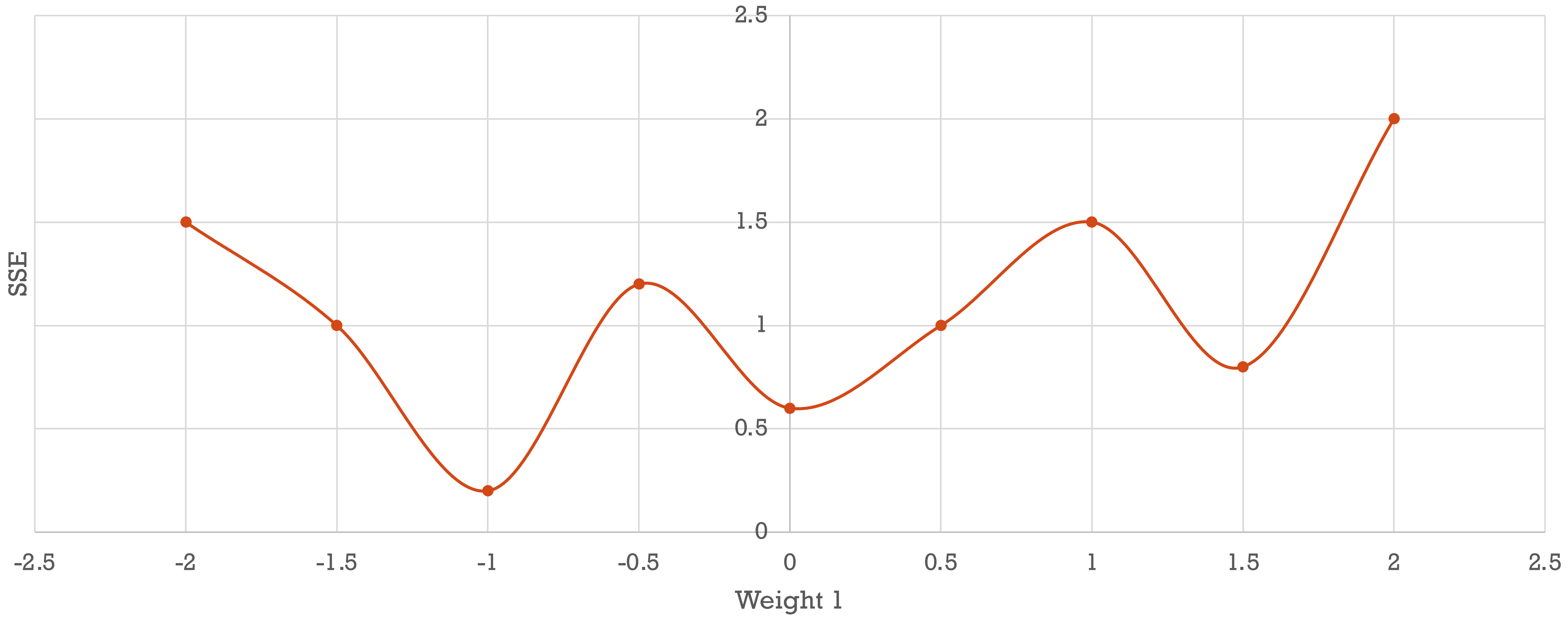
$$SSE = \sum (teacher - output)^2$$

ERROR CALCULATION

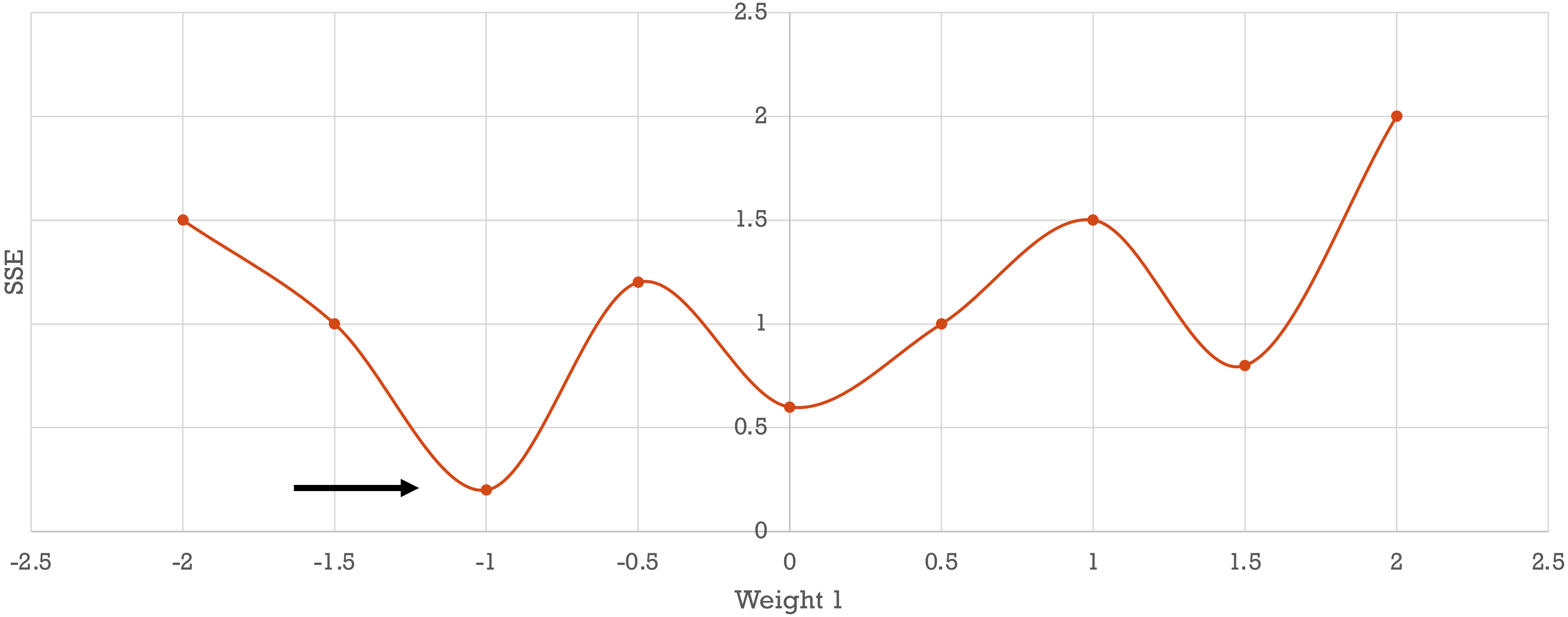
Output	Teacher	SE
0.125	1	(1 – 0.125) ^ 2 = 0.875 ^ 2 = 0.765625
0	0	(0 – 0)^2 = 0

SSE = 0.765625 + 0 = 0.765625

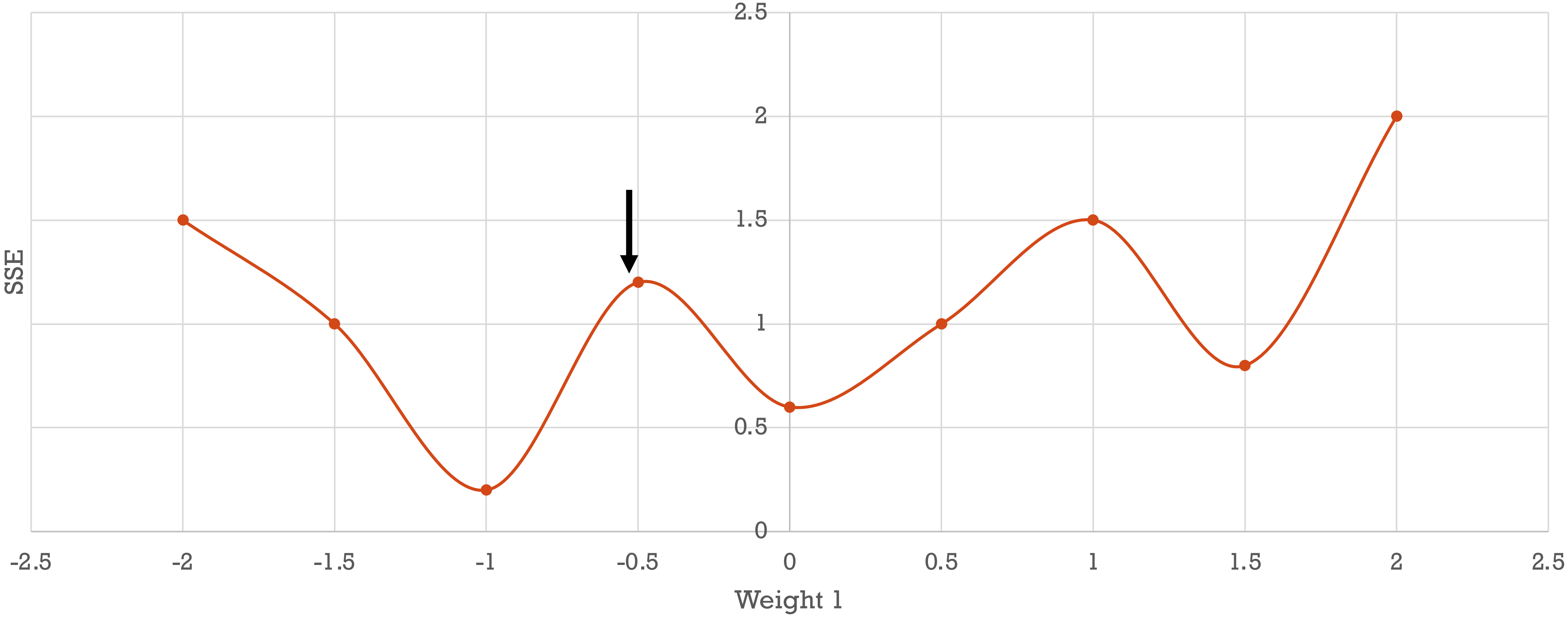
ERROR PLOT



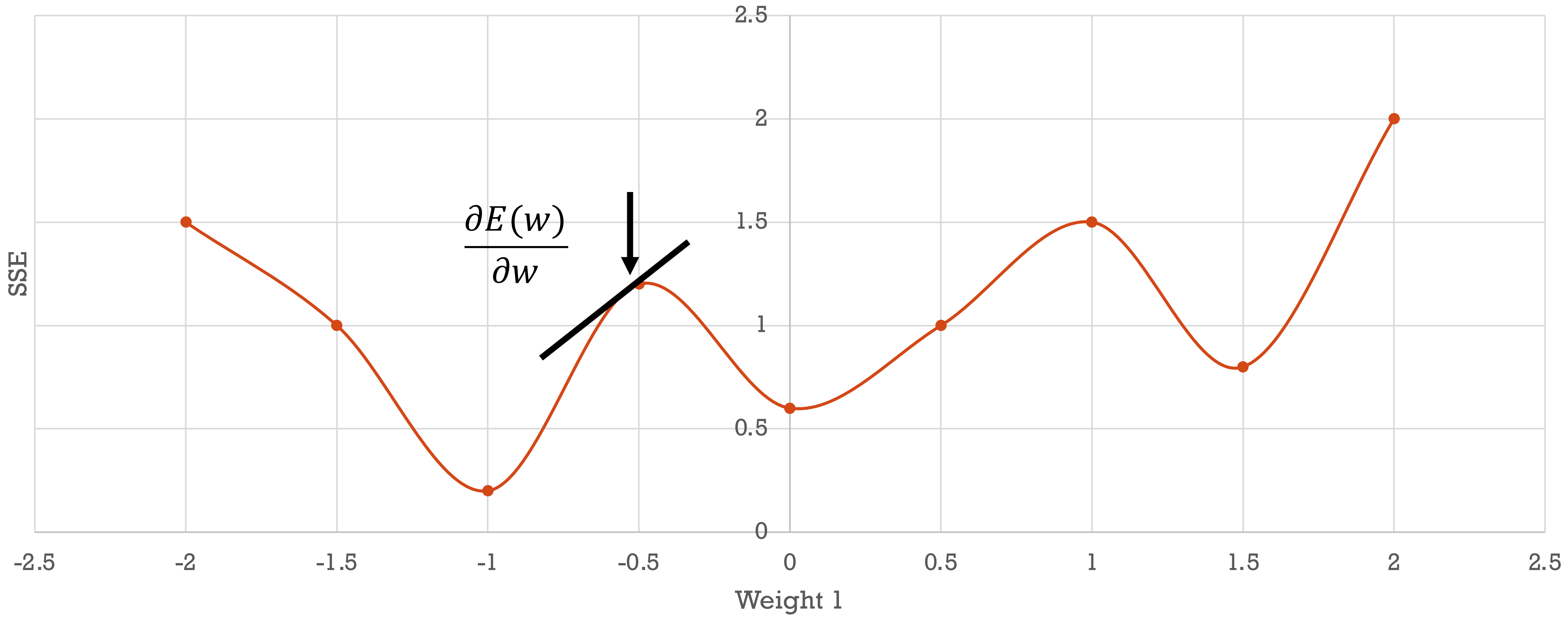
WEIGHT TUNING AS OPTIMIZATION



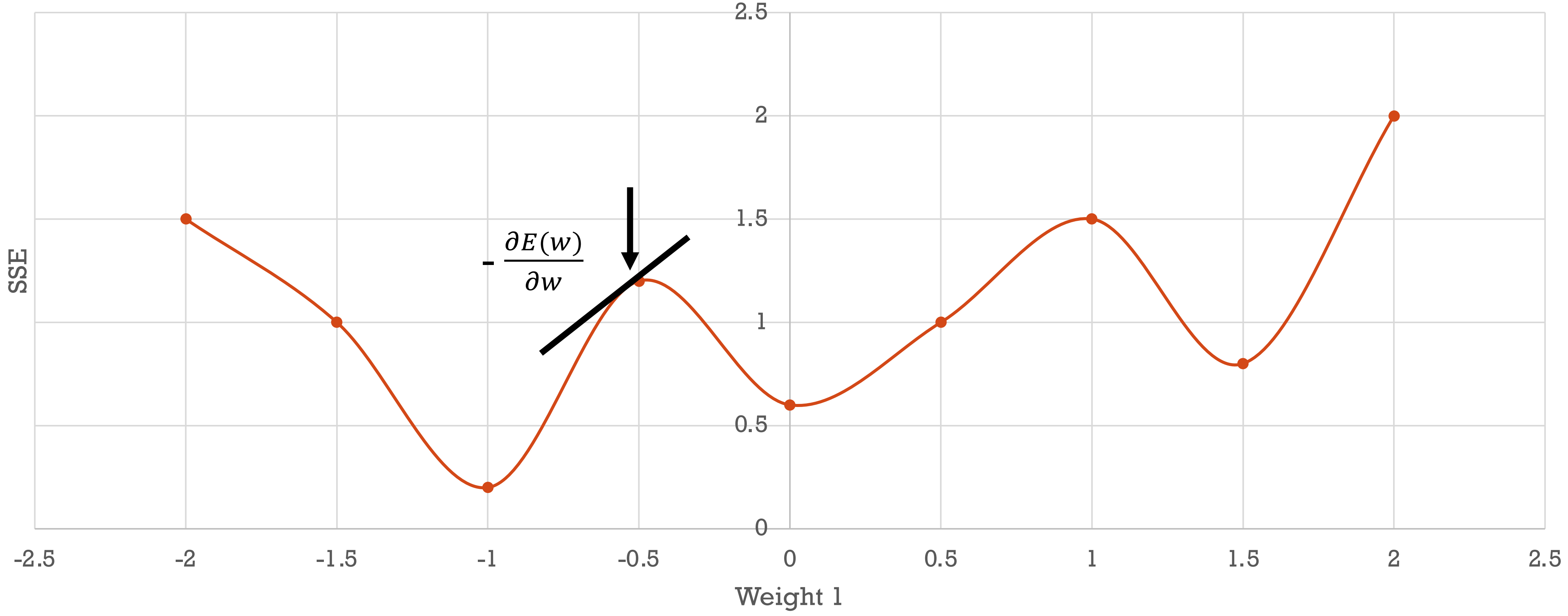
WEIGHT TUNING AS OPTIMIZATION



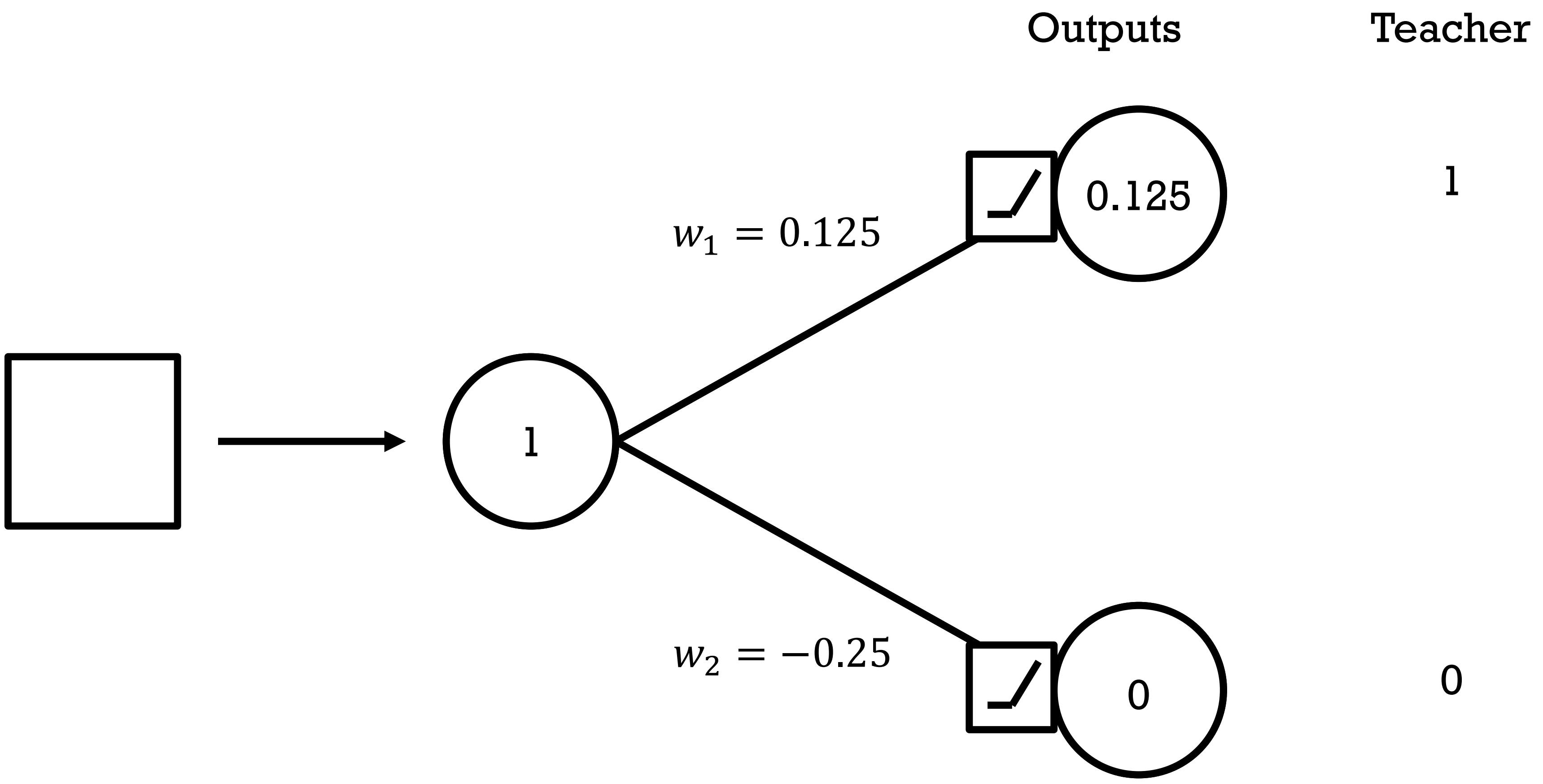
GRADIENT DESCENT



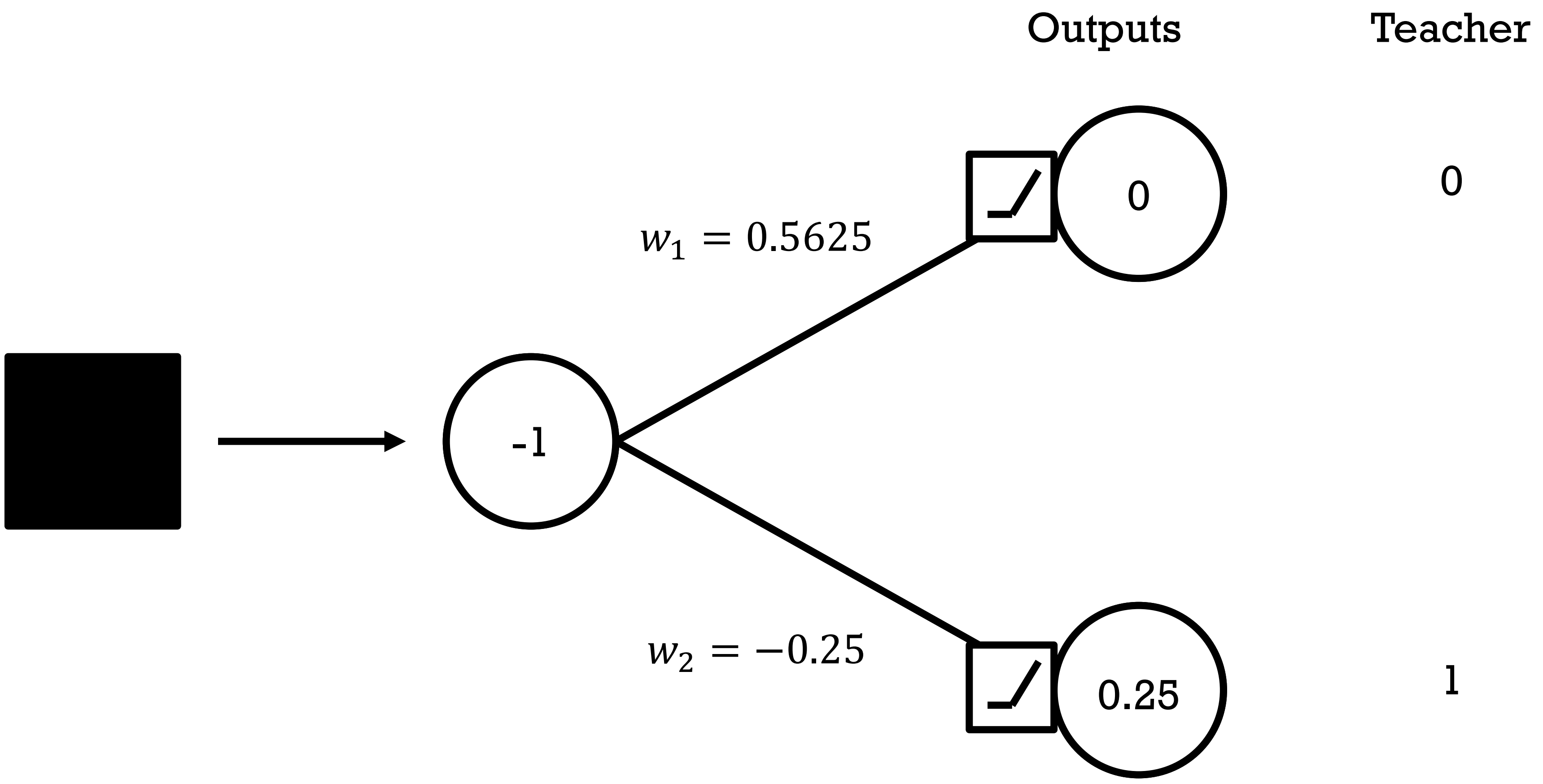
GRADIENT DESCENT



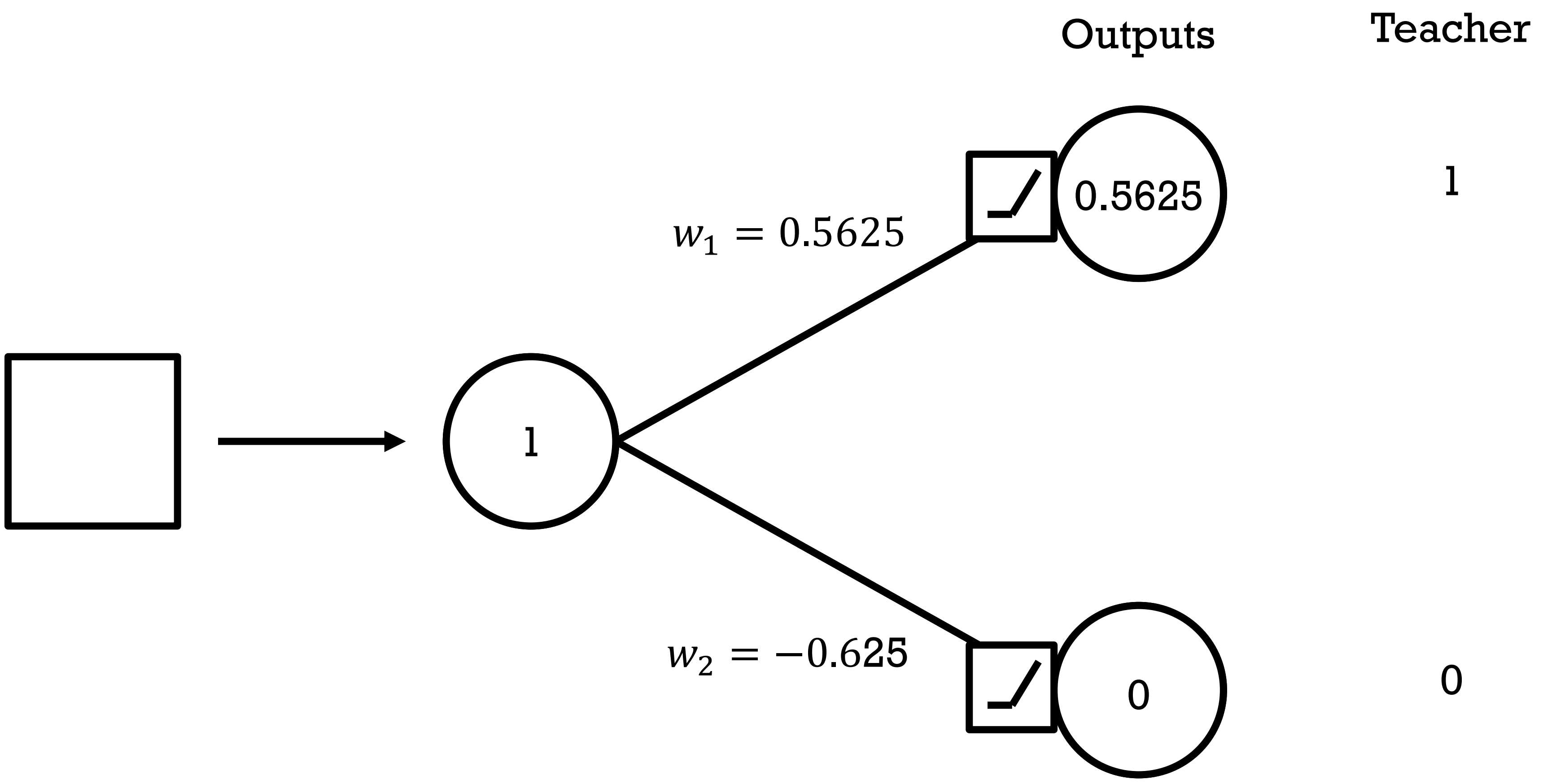
TRAINING – 1ST STEP



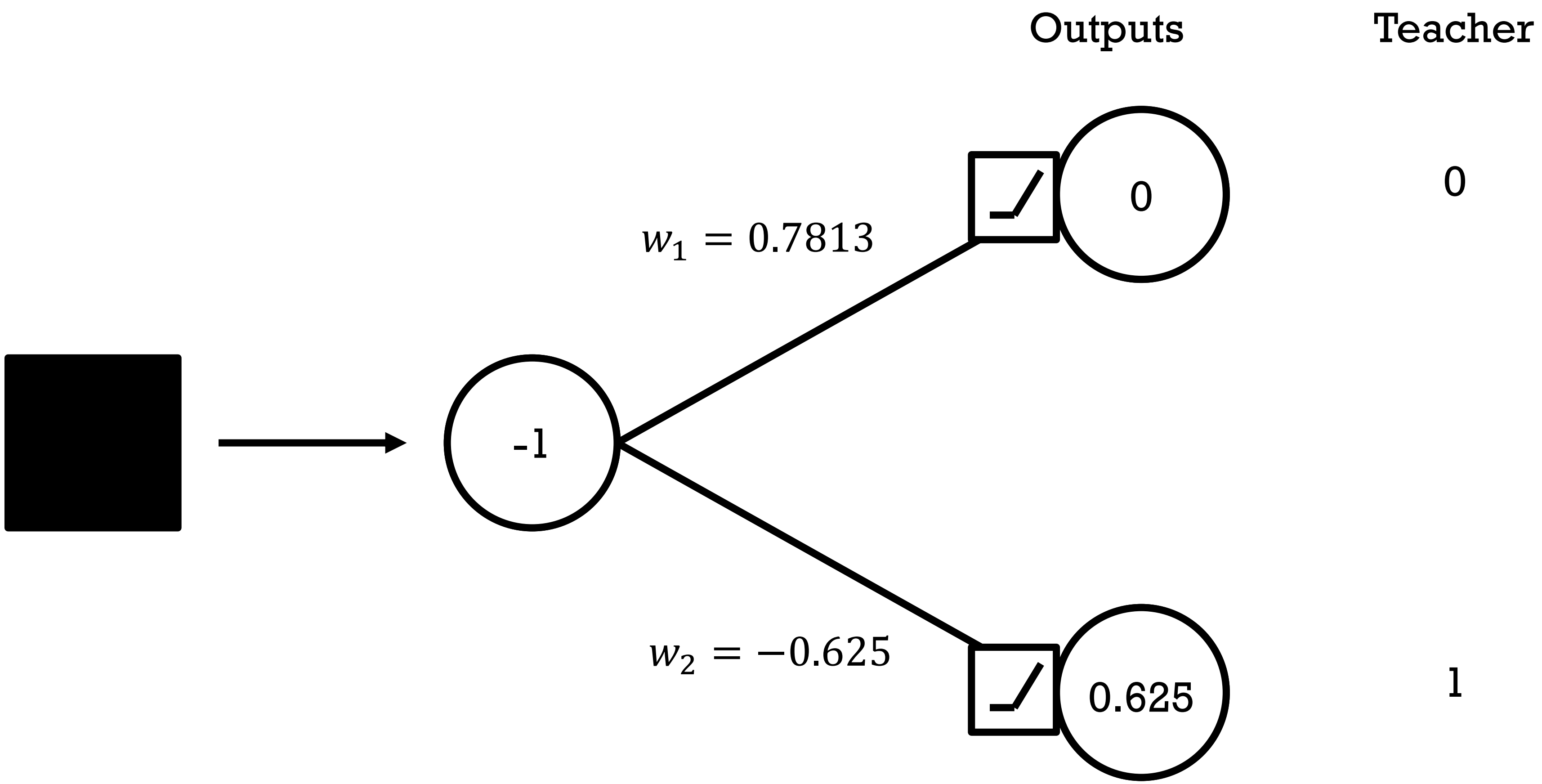
TRAINING – 2ND STEP



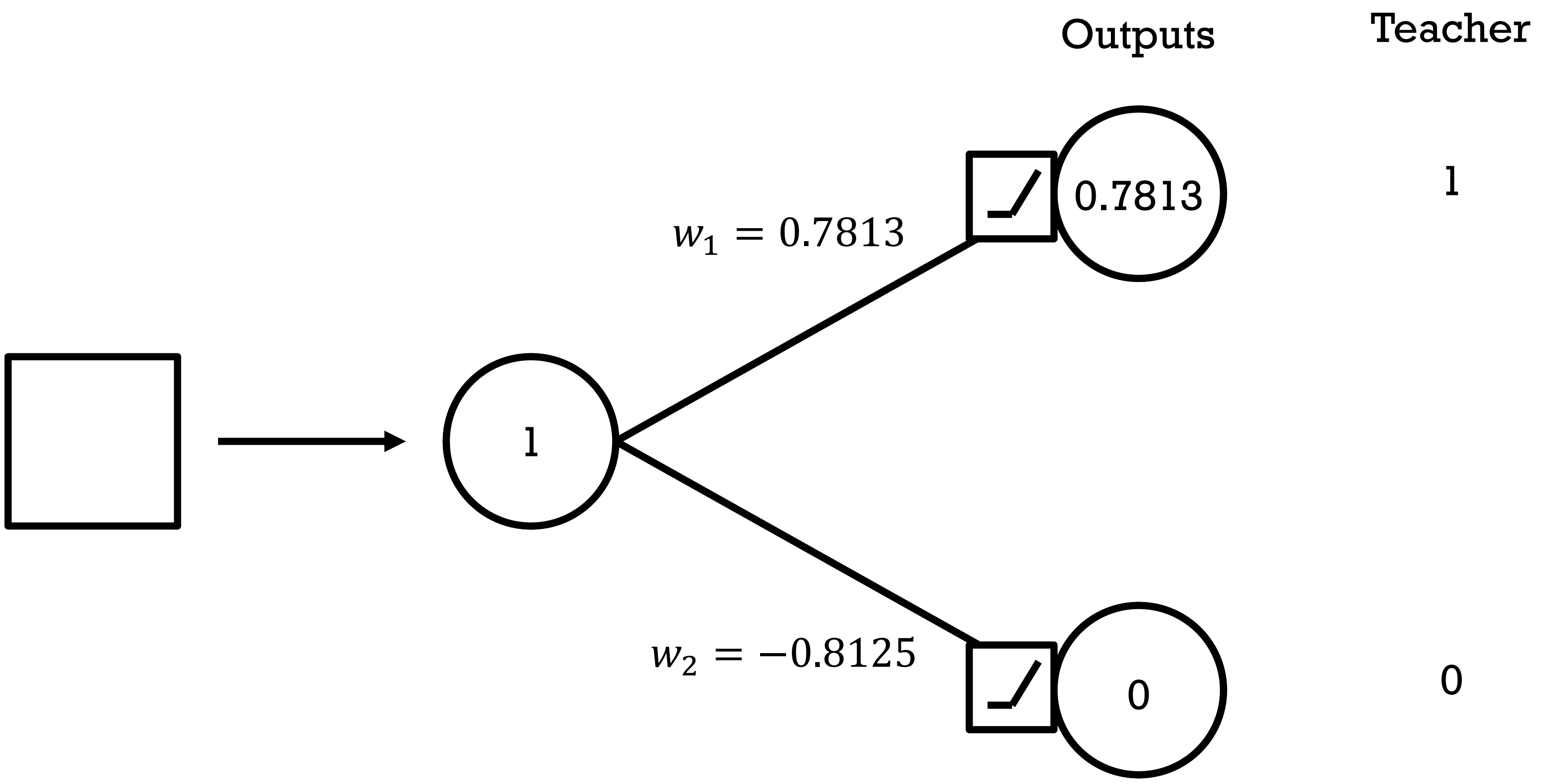
TRAINING – 1ST STEP OF 2ND EPOCH



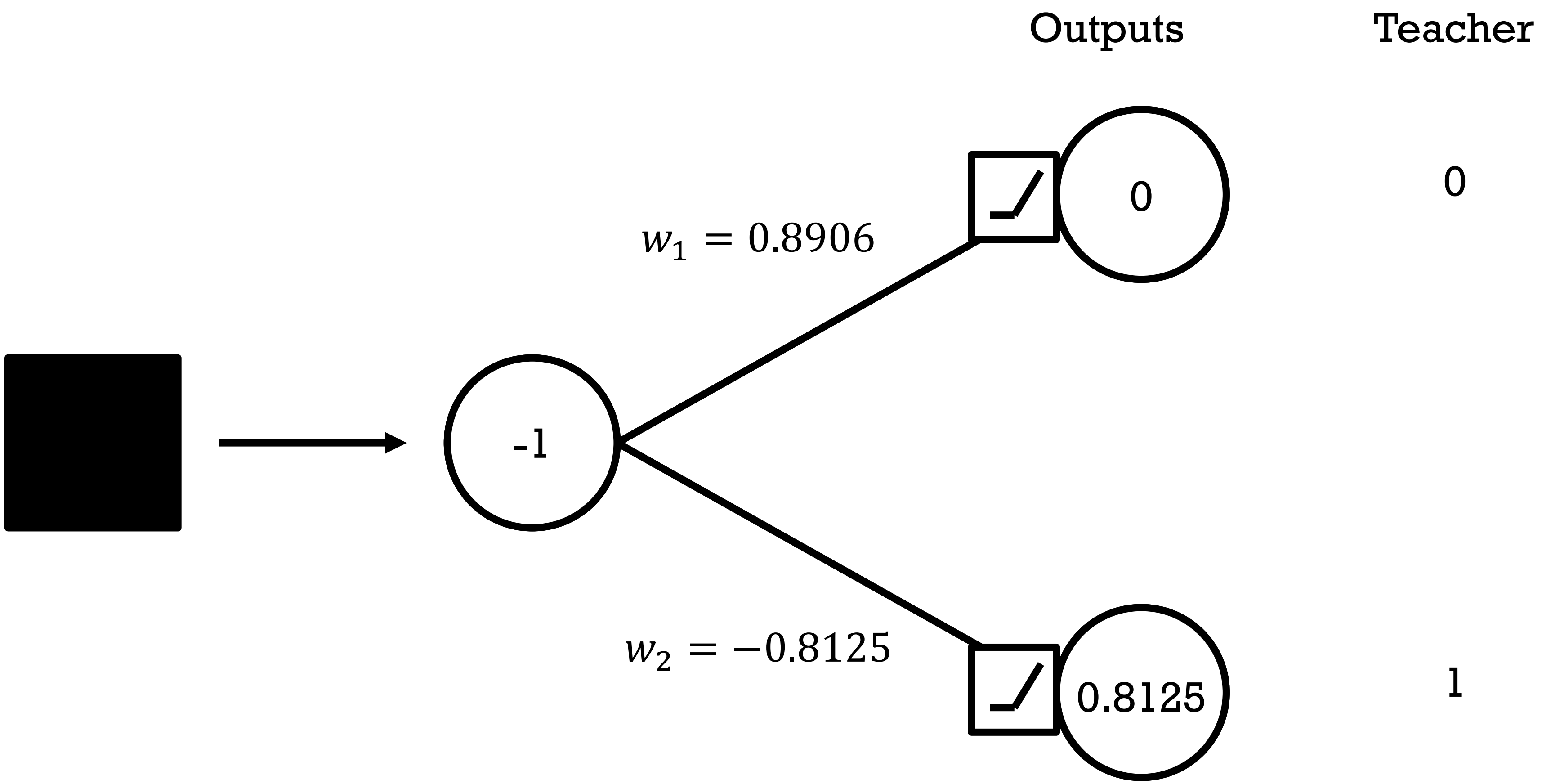
TRAINING – 2ND STEP OF 2ND EPOCH



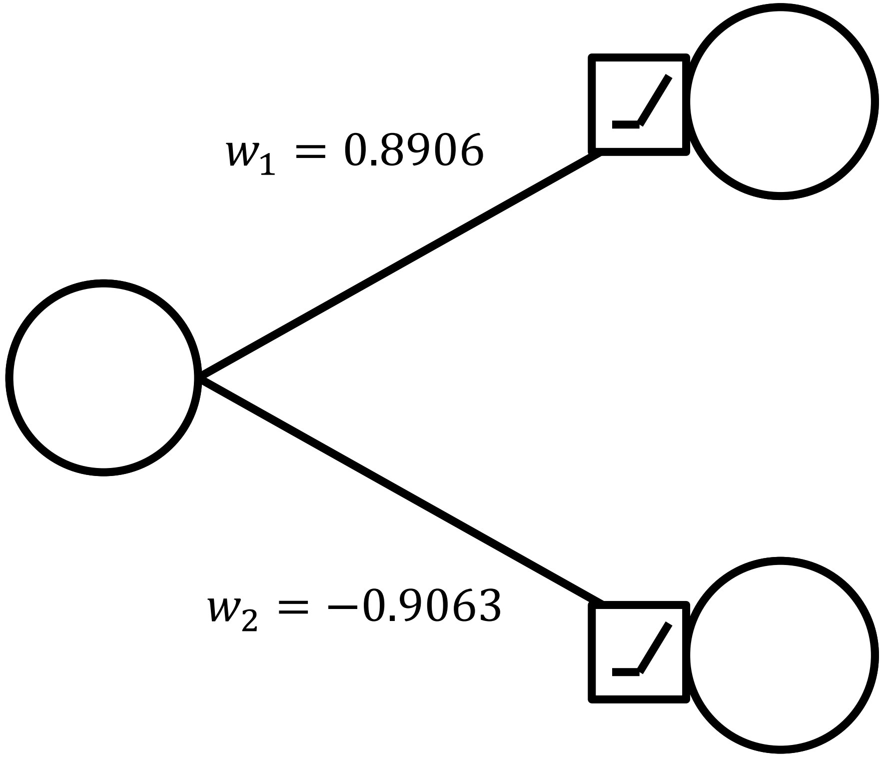
TRAINING – 1ST STEP OF 3RD EPOCH



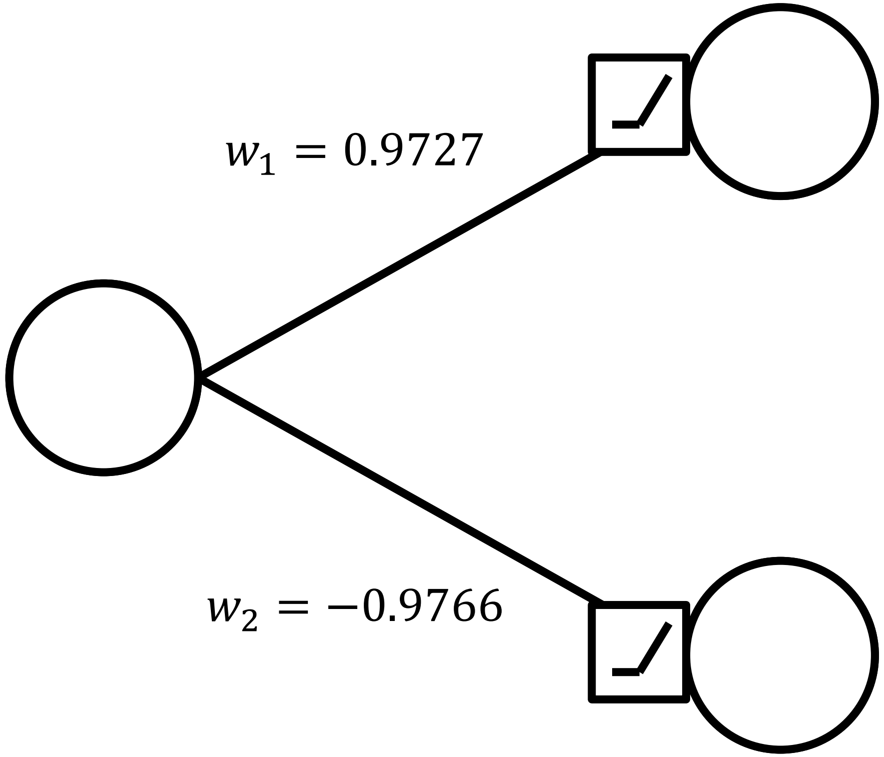
TRAINING – 2ND STEP OF 3RD EPOCH



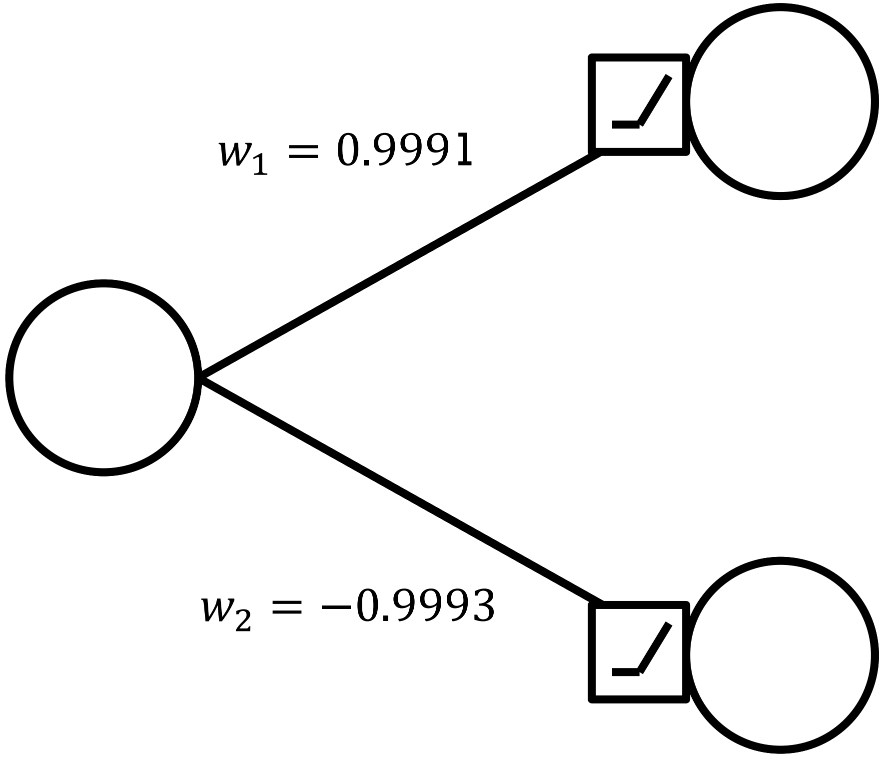
AFTER 3 EPOCHS



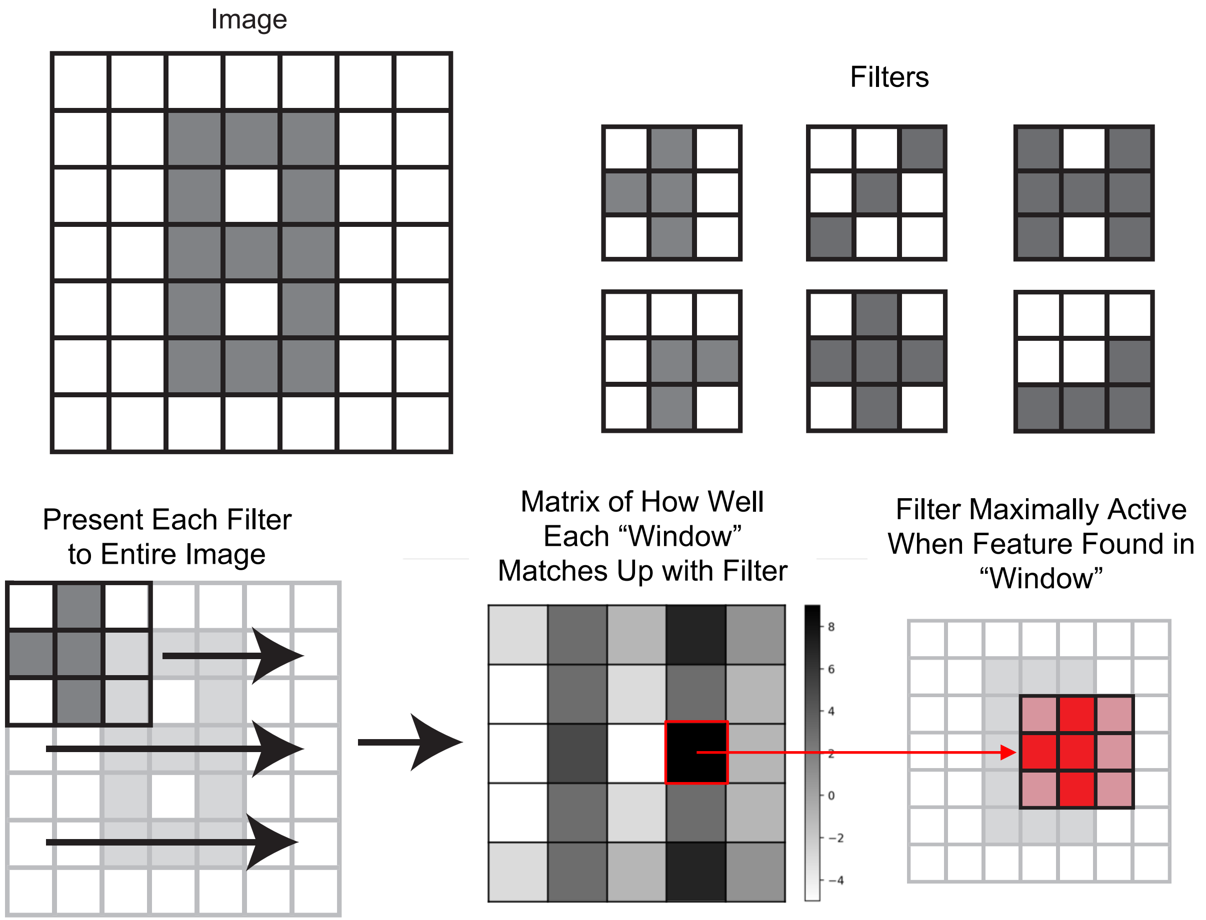
AFTER 5 EPOCHS



AFTER 10 EPOCHS

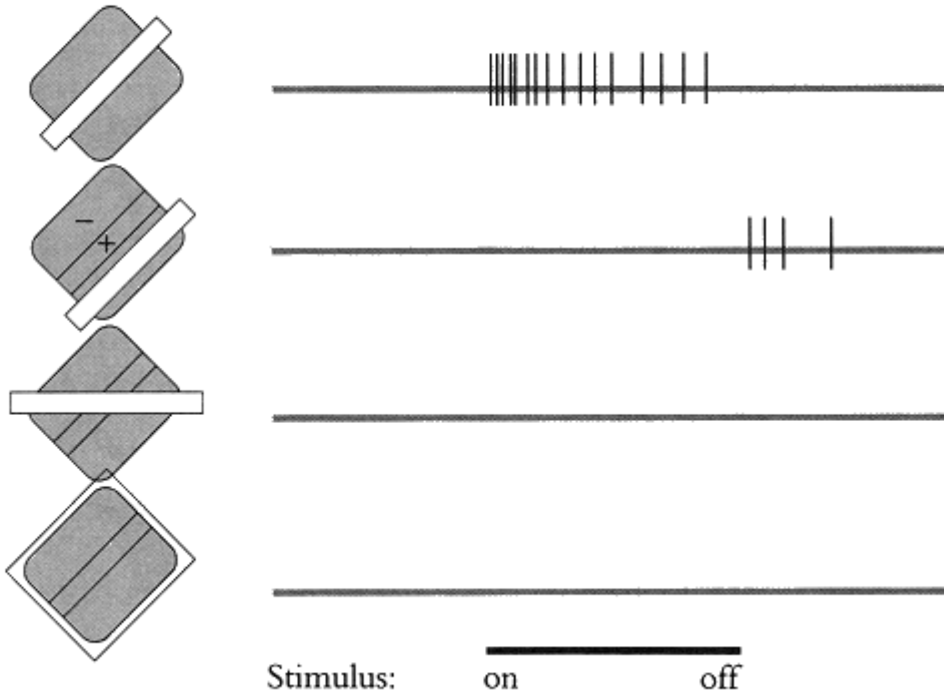


CONVOLUTION



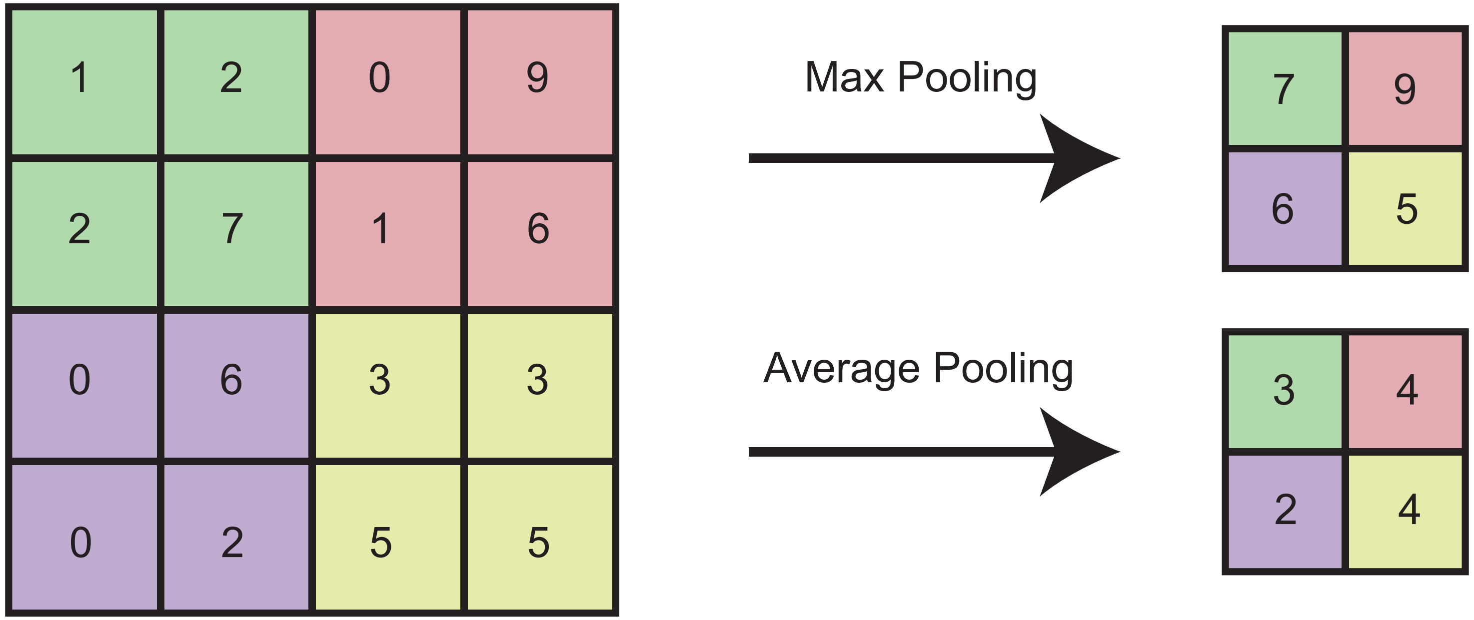
CONVOLUTION LAYERS

- Modeled after V1 simple cells



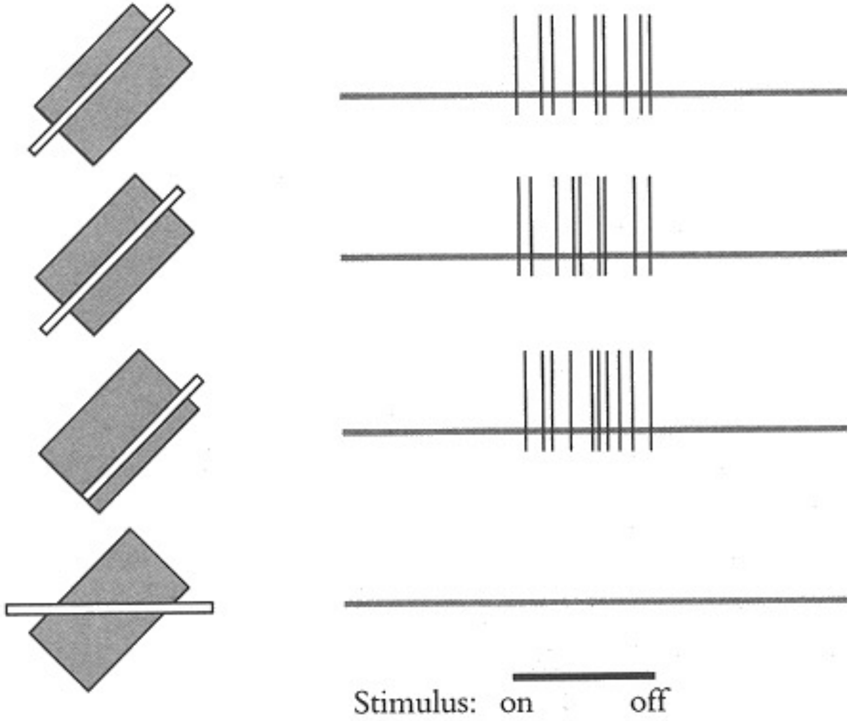
Picture from fourier.eng.hmu.edu

POOLING



POOLING AFTER CONVOLUTION

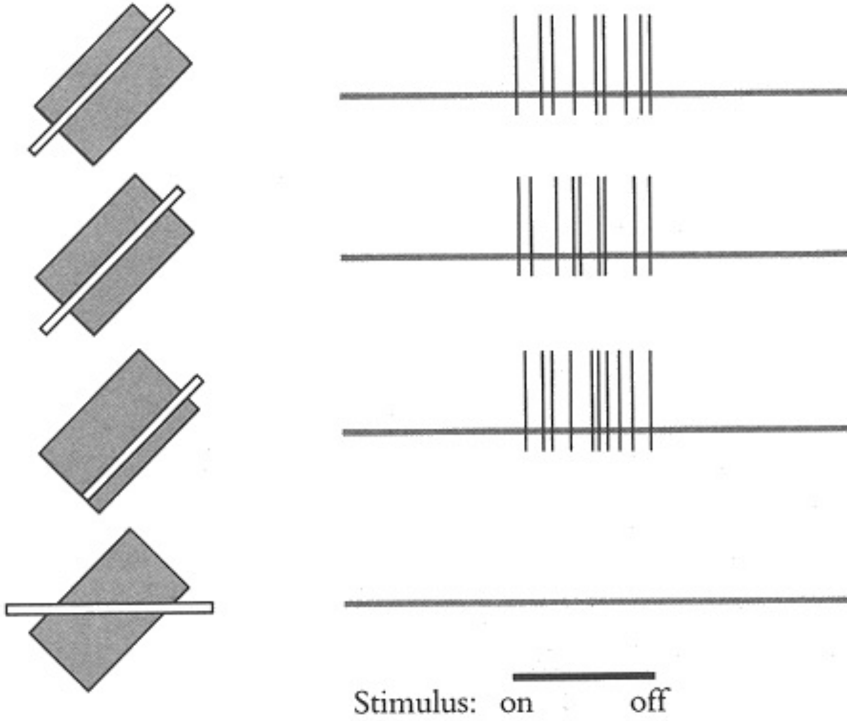
- Functions like V1 complex cells
- Shift invariance



Picture from fourier.eng.hmu.edu

POOLING AFTER CONVOLUTION

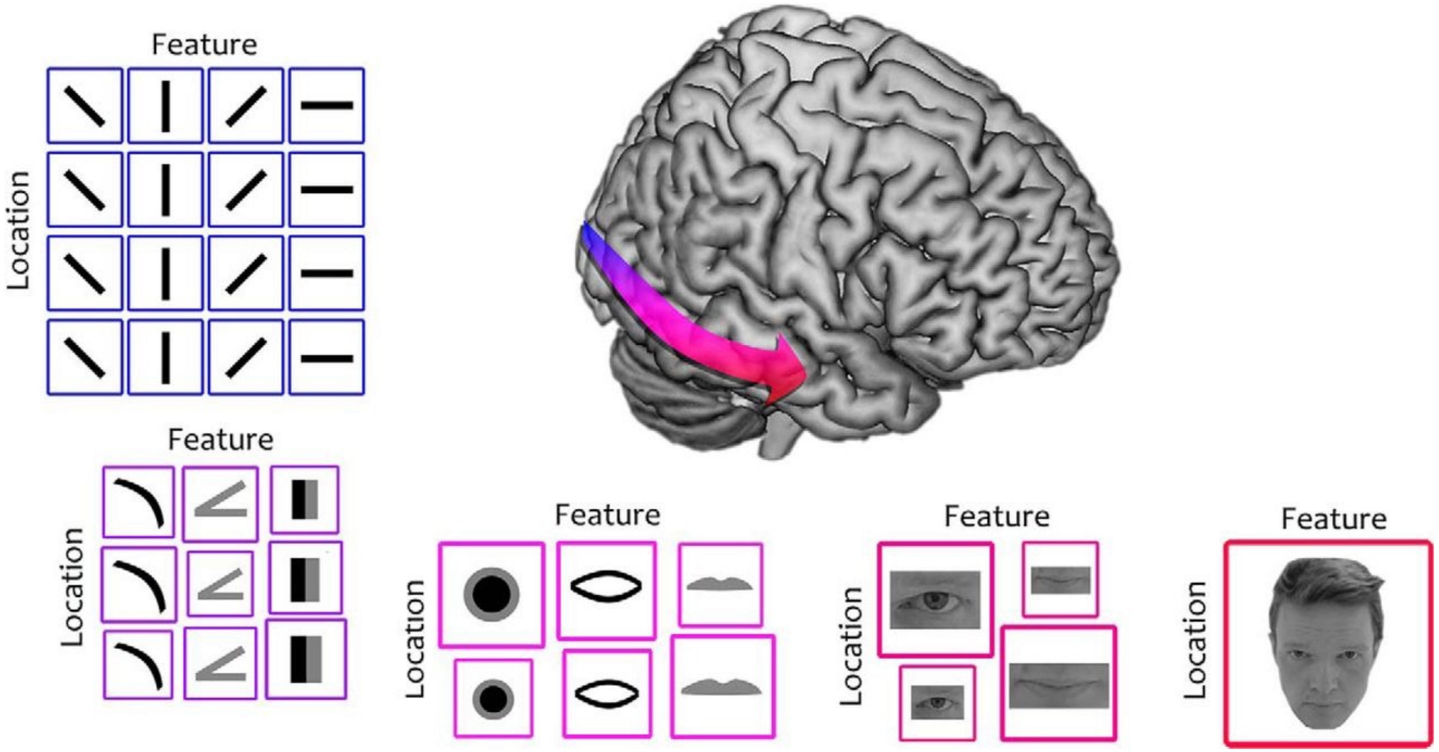
- Functions like V1 complex cells
- Shift invariance



Picture from fourier.eng.hmu.edu

STACKING CONVOLUTION LAYERS

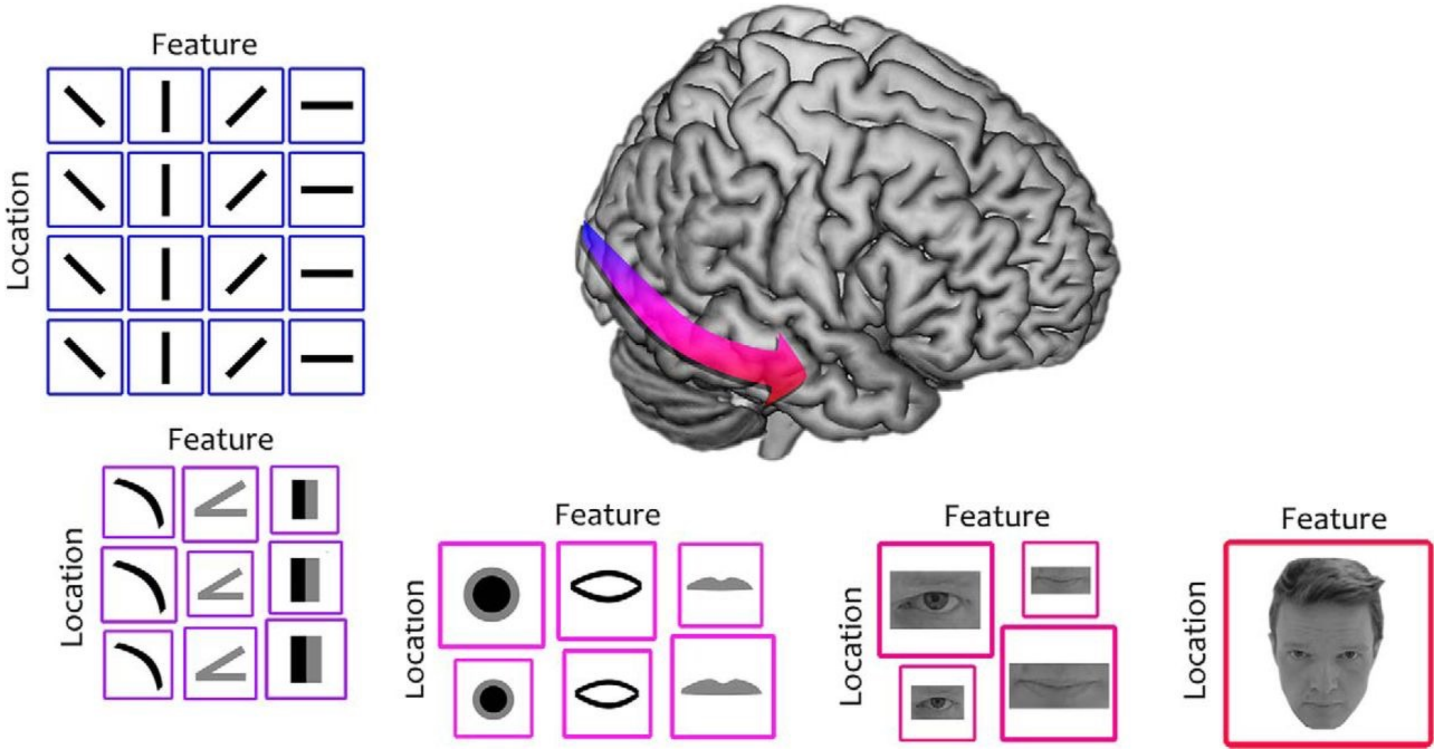
- Similar to ventral visual pathway
- Successive layers extract more complex features



Picture from Kaiser and Haselhahn (2017)

STACKING CONVOLUTION LAYERS

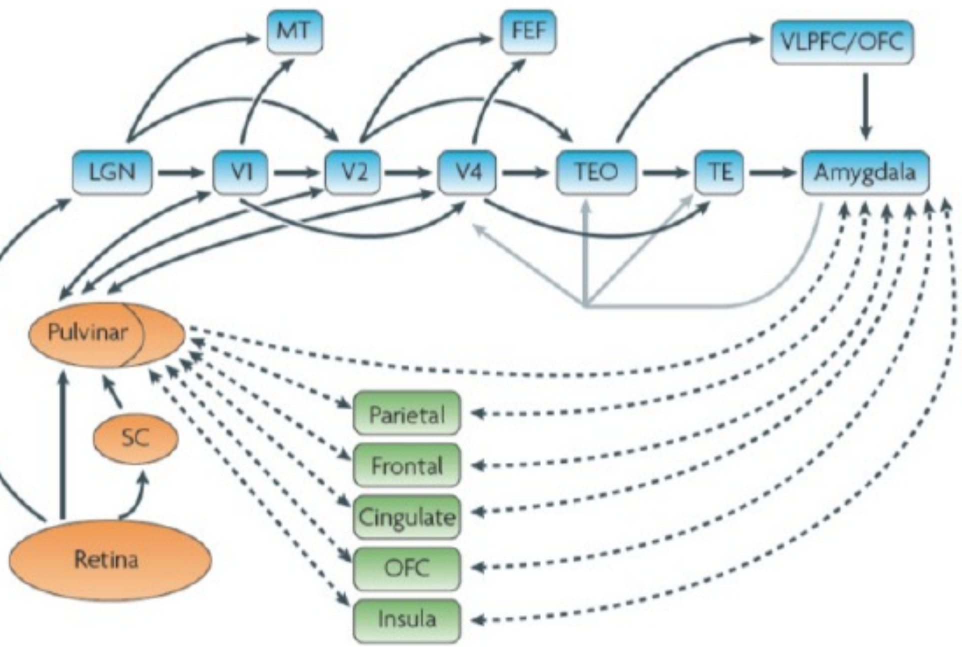
- Similar to ventral visual pathway
- Successive layers extract more complex features



Picture from Kaiser and Haselhahn (2017)

SKIP CONNECTIONS

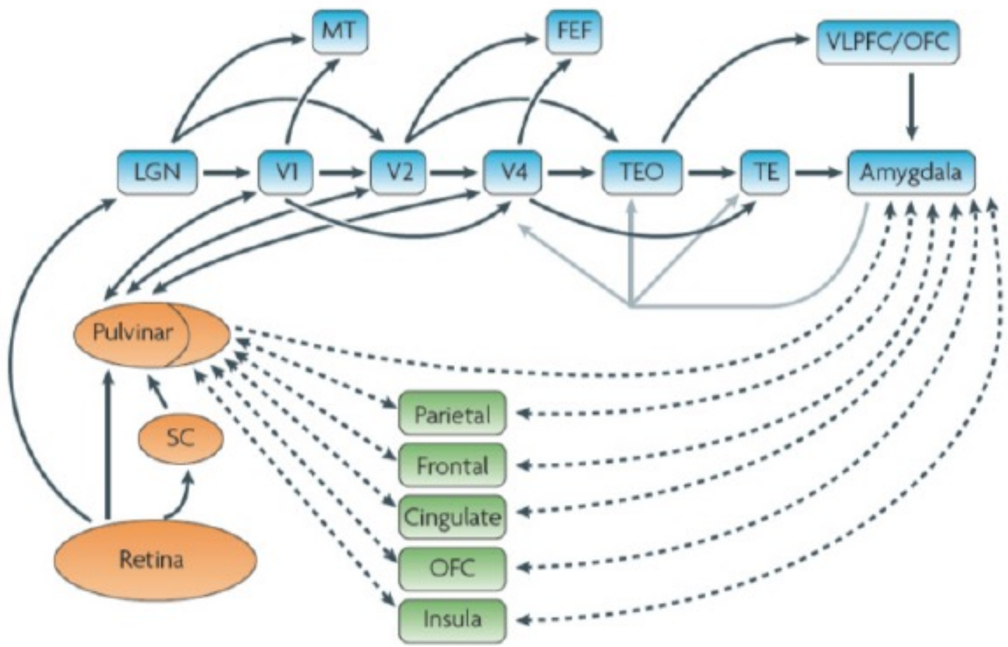
- Similar to skip connections in brain
- Make training of networks easier



Picture from Peter Walla

SKIP CONNECTIONS

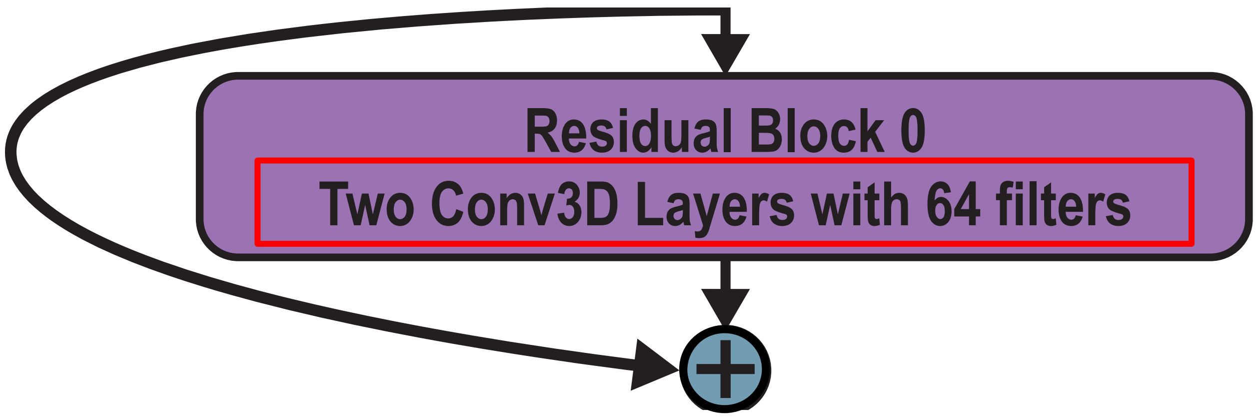
- Similar to skip connections in brain
- Make training of networks easier



Picture from Peter Walla

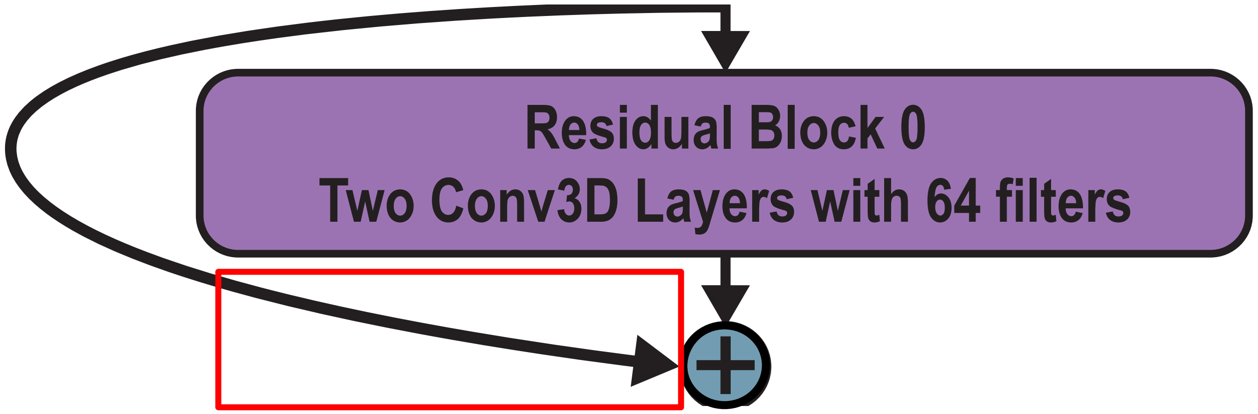
RESIDUAL BLOCK

- Stack of 2 convolution layers
- Skip connection from beginning to end



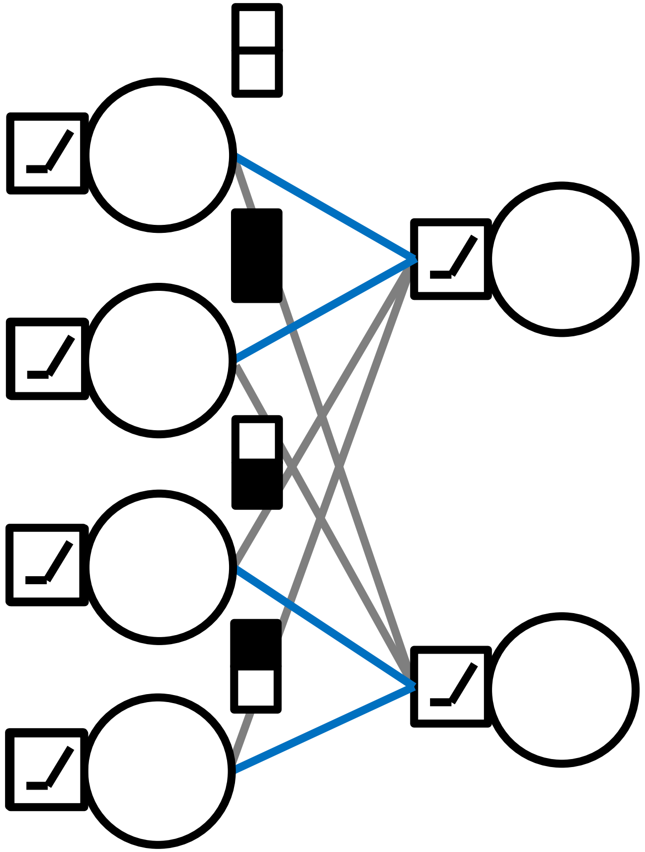
RESIDUAL BLOCK

- Stack of 2 convolution layers
- Skip connection from beginning to end



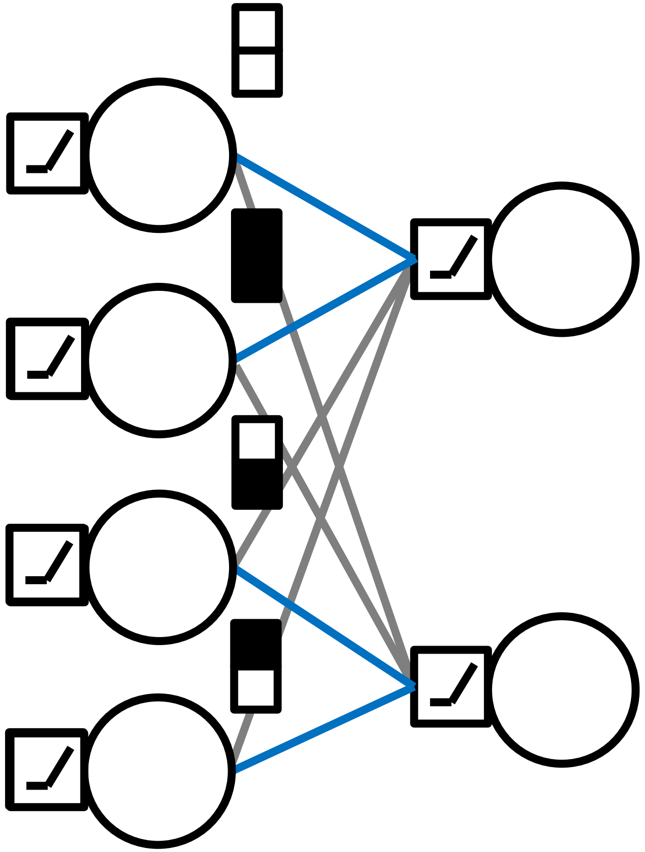
FULLY CONNECTED LAYER

- You already know about these!
- Utilize extracted features to make final prediction

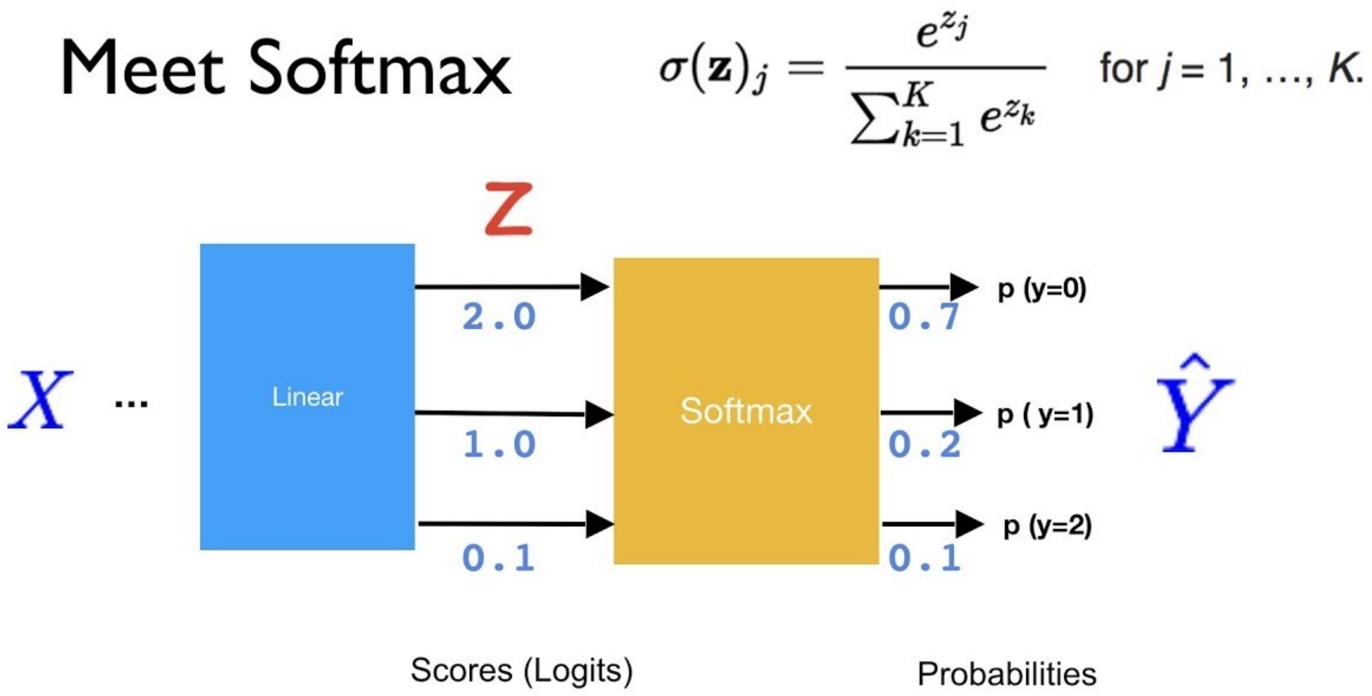


FULLY CONNECTED LAYER

- You already know about these!
- Utilize extracted features to make final prediction



SOFTMAX



Picture from Sung Kim's YouTube video on Softmax Classifier

OUTLINE

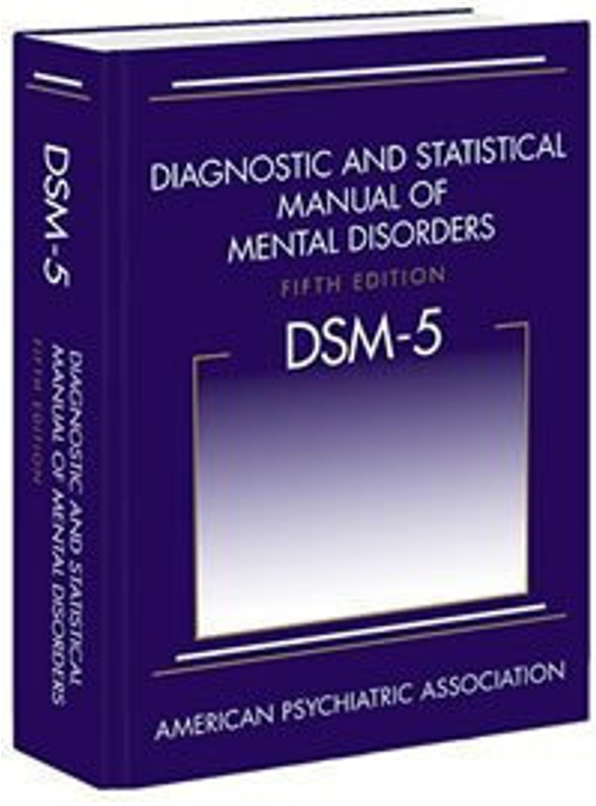
- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
  - Neural Network in Medical Imaging
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

OUTLINE

- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
  - Neural Network in Medical Imaging
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

DIAGNOSIS

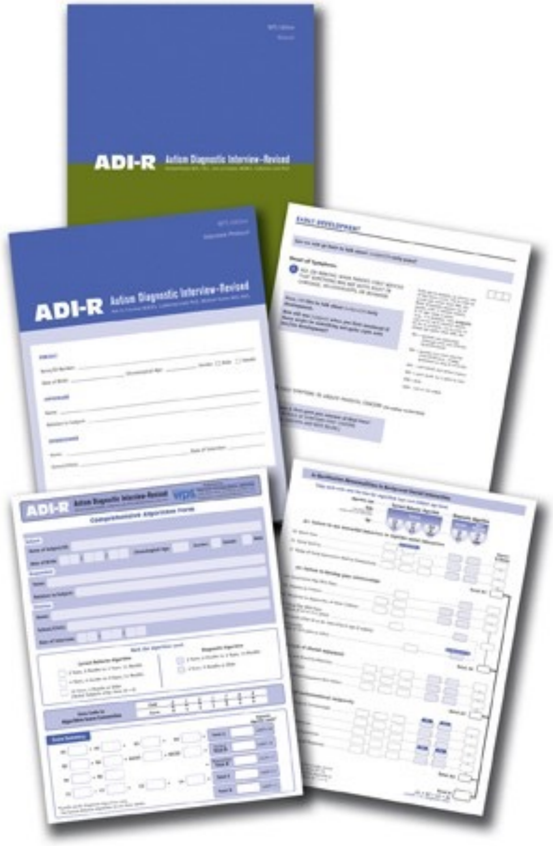
- Criteria from DSM-5
- Autism Diagnostic Interview-Revised (ADI-R)
- Autism Diagnostic Observation Schedule (ADOS)
- ADI-R and ADOS assess same clinical criteria
- Reliable specificity and sensitivity for ASD (Falkmer et al., 2013)
- ADOS has higher sensitivity than ADI-R (Randall et al., 2018)



Picture from etsy.com

DIAGNOSIS

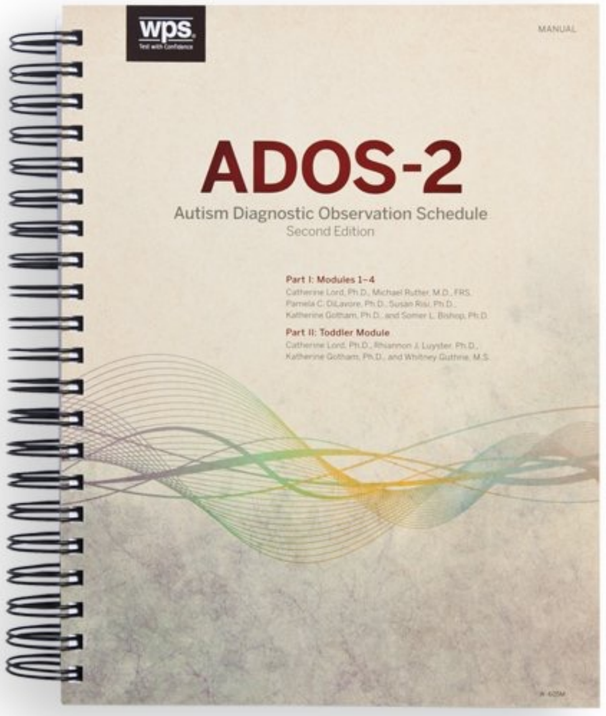
- Criteria from DSM-5
- Autism Diagnostic Interview-Revised (ADI-R)
- Autism Diagnostic Observation Schedule (ADOS)
- ADI-R and ADOS assess same clinical criteria
- Reliable specificity and sensitivity for ASD (Falkmer et al., 2013)
- ADOS has higher sensitivity than ADI-R (Randall et al., 2018)



Picture from stoeltingco.com

DIAGNOSIS

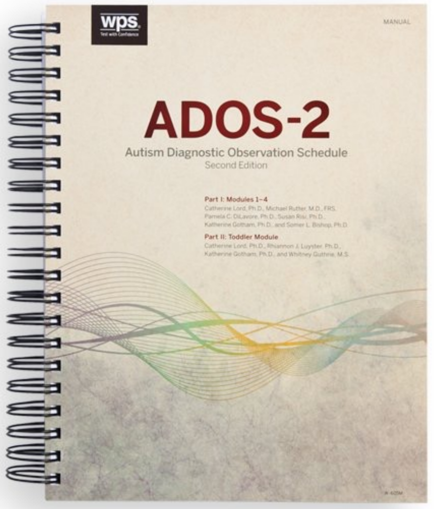
- Criteria from DSM-5
- Autism Diagnostic Interview-Revised (ADI-R)
- Autism Diagnostic Observation Schedule (ADOS)
- ADI-R and ADOS assess same clinical criteria
- Reliable specificity and sensitivity for ASD (Falkmer et al., 2013)
- ADOS has higher sensitivity than ADI-R (Randall et al., 2018)

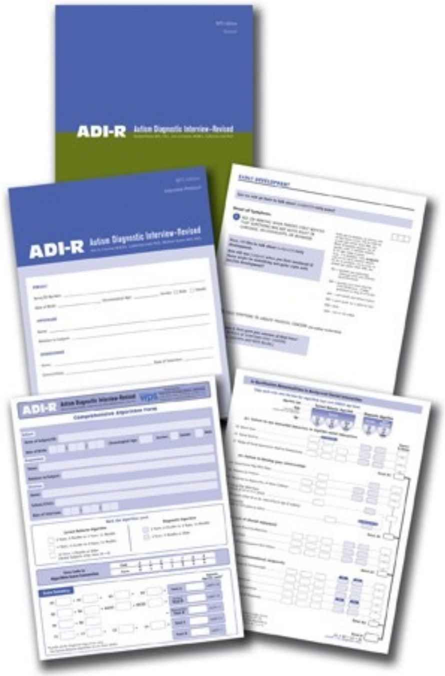


Picture from wpspublish.com

DIAGNOSIS

- Criteria from DSM-5
- Autism Diagnostic Interview-Revised (ADI-R)
- Autism Diagnostic Observation Schedule (ADOS)
- ADI-R and ADOS assess same clinical criteria
- Reliable specificity and sensitivity for ASD (Falkmer et al., 2013)
- ADOS has higher sensitivity than ADI-R (Randall et al., 2018)

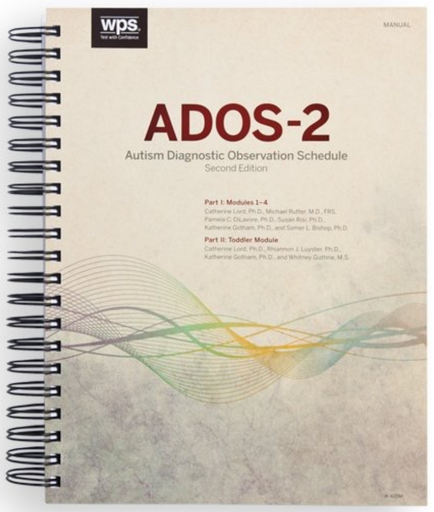


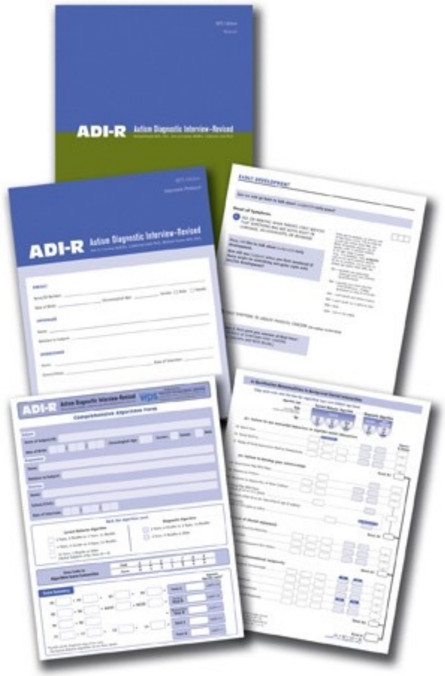


Pictures from wpspublish.com and stoeltingco.com

DIAGNOSIS

- Criteria from DSM-5
- Autism Diagnostic Interview-Revised (ADI-R)
- Autism Diagnostic Observation Schedule (ADOS)
- ADI-R and ADOS assess same clinical criteria
- Reliable specificity and sensitivity for ASD (Falkmer et al., 2013)
- ADOS has higher sensitivity than ADI-R (Randall et al., 2018)

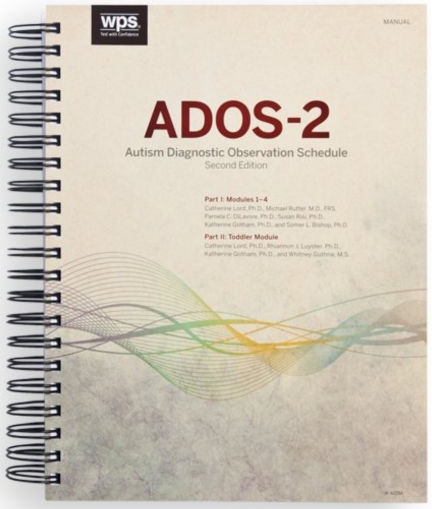




Pictures from wpspublish.com and stoeltingco.com

DIAGNOSIS

- Criteria from DSM-5
- Autism Diagnostic Interview-Revised (ADI-R)
- Autism Diagnostic Observation Schedule (ADOS)
- ADI-R and ADOS assess same clinical criteria
- Reliable specificity and sensitivity for ASD (Falkmer et al., 2013)
- ADOS has higher sensitivity than ADI-R (Randall et al., 2018)



Pictures from wpspublish.com and stoeltingco.com

EARLY DIAGNOSIS AND TREATMENT

- Current behavioral assessments require children to be 12 months and older (Esler et al. 2015)
- Improve quality of life for individuals and families (Elder et al., 2017)
- Estimated healthcare savings of $208,500/child (Chasson et al., 2007)
- Substantial sustained gains in IQ, language, academic performance, and adaptive behavior (Myers et al., 2007)

Picture from nashp.org

EARLY DIAGNOSIS AND TREATMENT

- Current behavioral assessments require children to be 12 months and older (Esler et al. 2015)
- Improve quality of life for individuals and families (Elder et al., 2017)
- Estimated healthcare savings of $208,500/child (Chasson et al., 2007)
- Substantial sustained gains in IQ, language, academic performance, and adaptive behavior (Myers et al., 2007)

Picture from nashp.org

EARLY DIAGNOSIS AND TREATMENT

- Current behavioral assessments require children to be 12 months and older (Esler et al. 2015)
- Improve quality of life for individuals and families (Elder et al., 2017)
- Estimated healthcare savings of $208,500/child (Chasson et al., 2007)
- Substantial sustained gains in IQ, language, academic performance, and adaptive behavior (Myers et al., 2007)

Picture from nashp.org

EARLY DIAGNOSIS AND TREATMENT

- Current behavioral assessments require children to be 12 months and older (Esler et al. 2015)
- Improve quality of life for individuals and families (Elder et al., 2017)
- Estimated healthcare savings of $208,500/child (Chasson et al., 2007)
- Substantial sustained gains in IQ, language, academic performance, and adaptive behavior (Myers et al., 2007)

Picture from nashp.org

FUNCTIONAL DIFFERENCES IN ASD

- Majority of changes seen in ASD (Rudie et al., 2013)
- Decreased functional connectivity throughout the brain
  - Resting-state network (Cherkassky et al., 2006)
  - Anterior-posterior connections (Cherkassky et al., 2006)
  - Fronto-parietal connections (Just et al., 2007)
  - Within functional systems (Rudie et al., 2013)
  - Interhemispheric connections (Anderson et al., 2011)
- No deactivation of resting-state network (Kennedy et al., 2006)

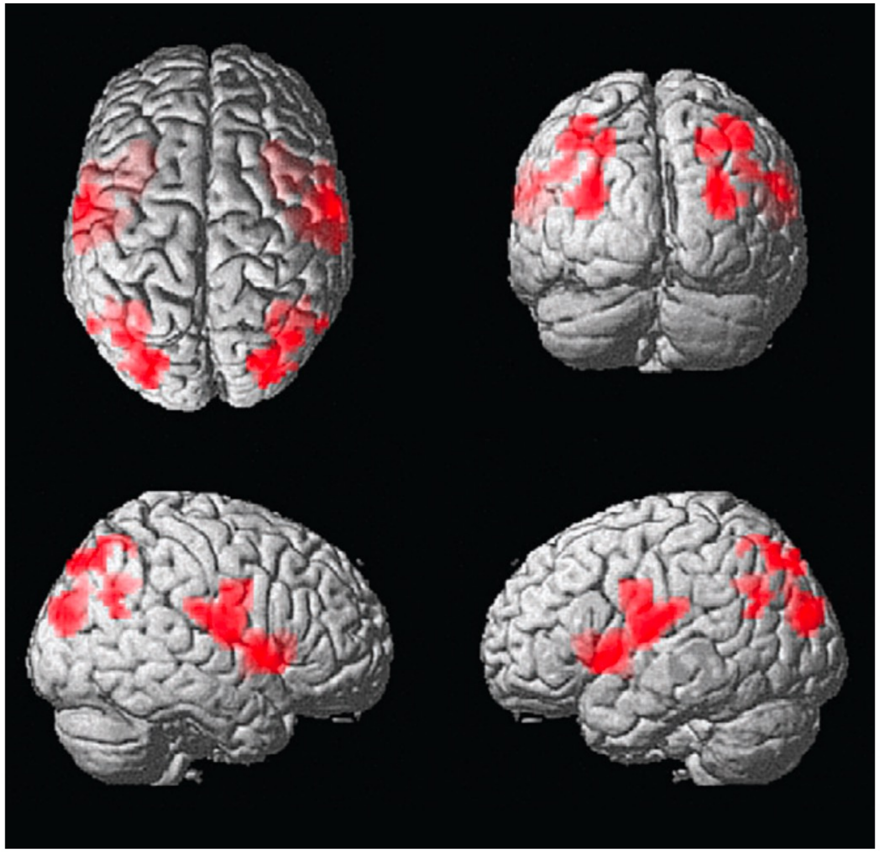


Figure from Anderson et al., 2011

FUNCTIONAL DIFFERENCES IN ASD

- Majority of changes seen in ASD (Rudie et al., 2013)
- Decreased functional connectivity throughout the brain
  - Resting-state network (Cherkassky et al., 2006)
  - Anterior-posterior connections (Cherkassky et al., 2006)
  - Fronto-parietal connections (Just et al., 2007)
  - Within functional systems (Rudie et al., 2013)
  - Interhemispheric connections (Anderson et al., 2011)
- No deactivation of resting-state network (Kennedy et al., 2006)

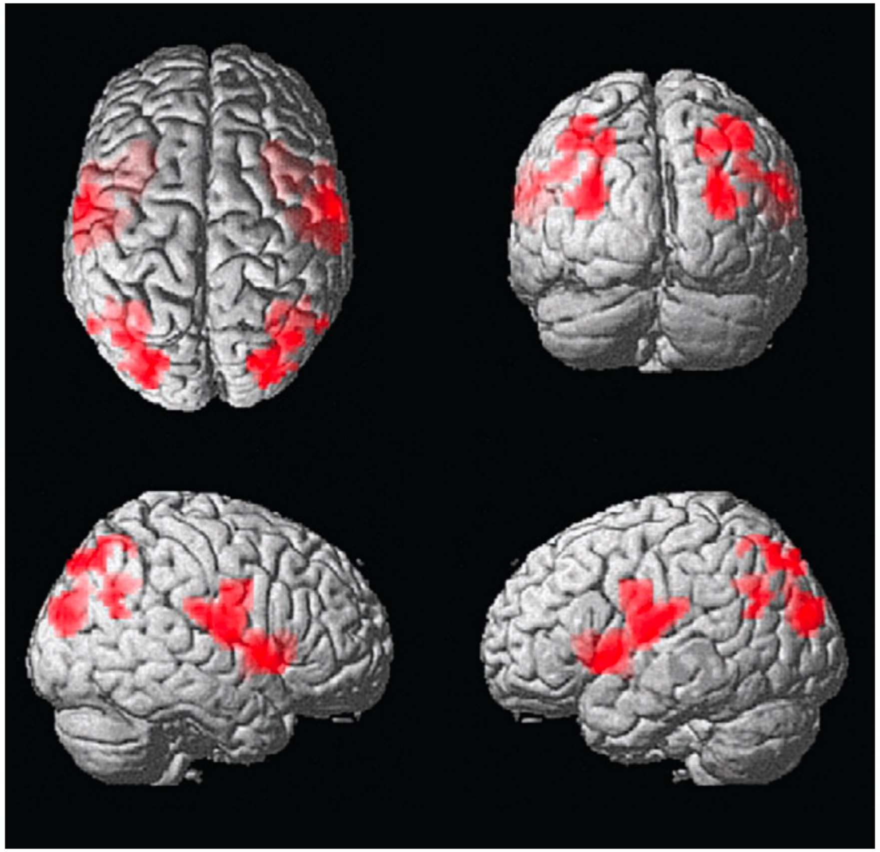


Figure from Anderson et al., 2011

FUNCTIONAL DIFFERENCES IN ASD

- Majority of changes seen in ASD (Rudie et al., 2013)
- Decreased functional connectivity throughout the brain
  - Resting-state network (Cherkassky et al., 2006)
  - Anterior-posterior connections (Cherkassky et al., 2006)
  - Fronto-parietal connections (Just et al., 2007)
  - Within functional systems (Rudie et al., 2013)
  - Interhemispheric connections (Anderson et al., 2011)
- No deactivation of resting-state network (Kennedy et al., 2006)

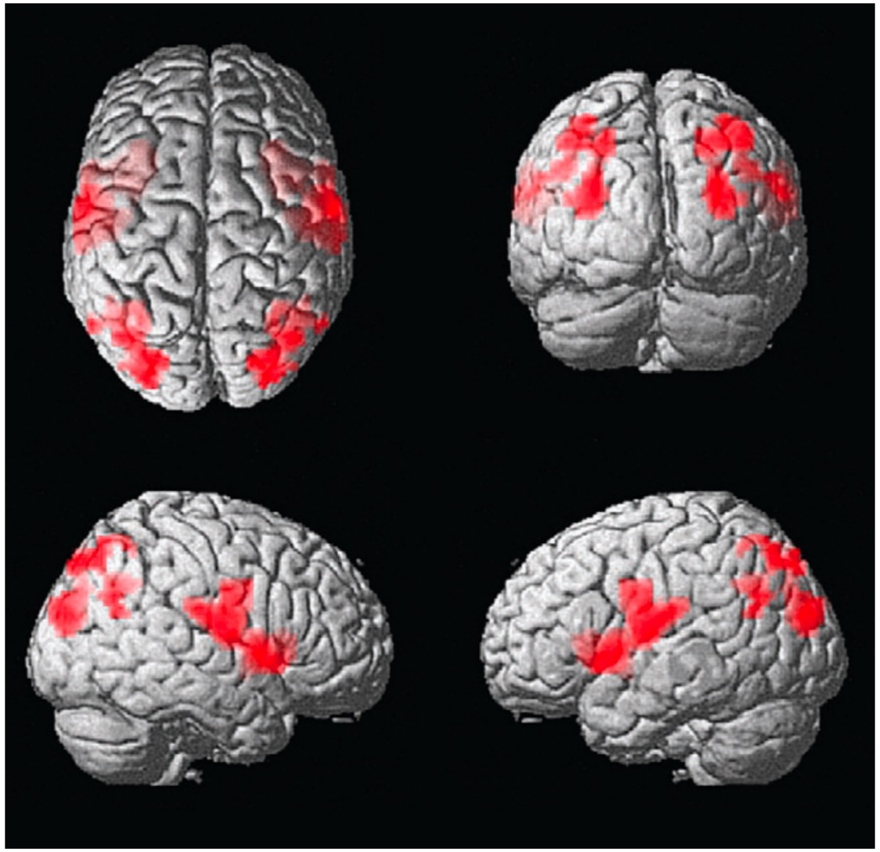


Figure from Anderson et al., 2011

FUNCTIONAL DIFFERENCES IN ASD

- Majority of changes seen in ASD (Rudie et al., 2013)
- Decreased functional connectivity throughout the brain
  - Resting-state network (Cherkassky et al., 2006)
  - Anterior-posterior connections (Cherkassky et al., 2006)
  - Fronto-parietal connections (Just et al., 2007)
  - Within functional systems (Rudie et al., 2013)
  - Interhemispheric connections (Anderson et al., 2011)
- No deactivation of resting-state network (Kennedy et al., 2006)

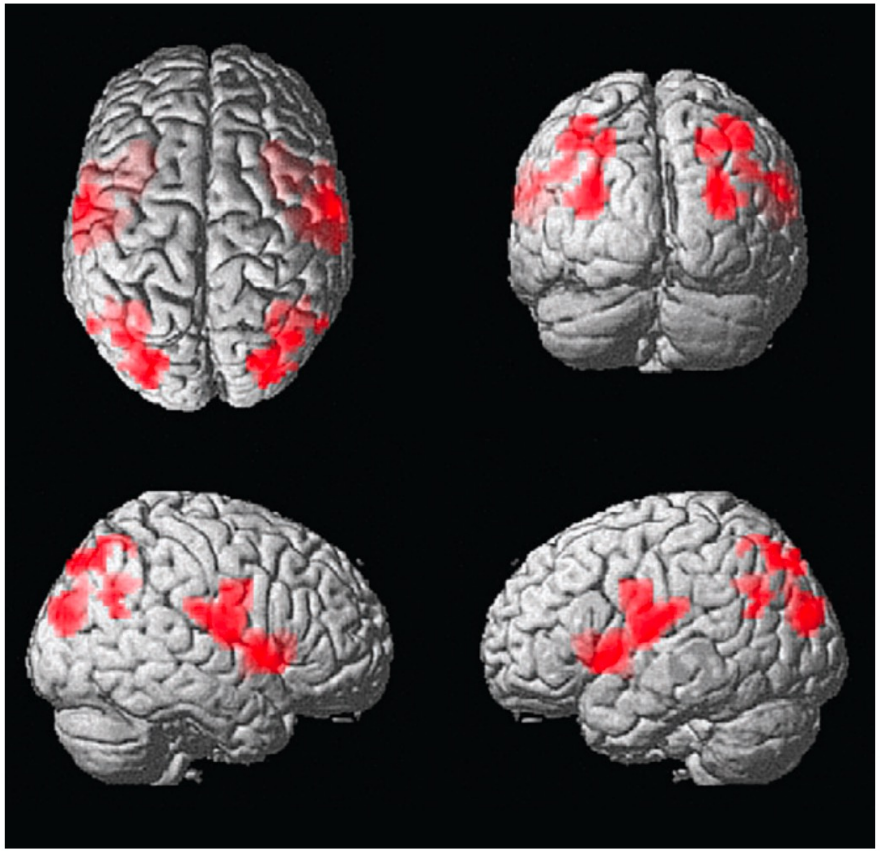


Figure from Anderson et al., 2011

FUNCTIONAL DIFFERENCES IN ASD

- Majority of changes seen in ASD (Rudie et al., 2013)
- Decreased functional connectivity throughout the brain
  - Resting-state network (Cherkassky et al., 2006)
  - Anterior-posterior connections (Cherkassky et al., 2006)
  - Fronto-parietal connections (Just et al., 2007)
  - Within functional systems (Rudie et al., 2013)
  - Interhemispheric connections (Anderson et al., 2011)
- No deactivation of resting-state network (Kennedy et al., 2006)

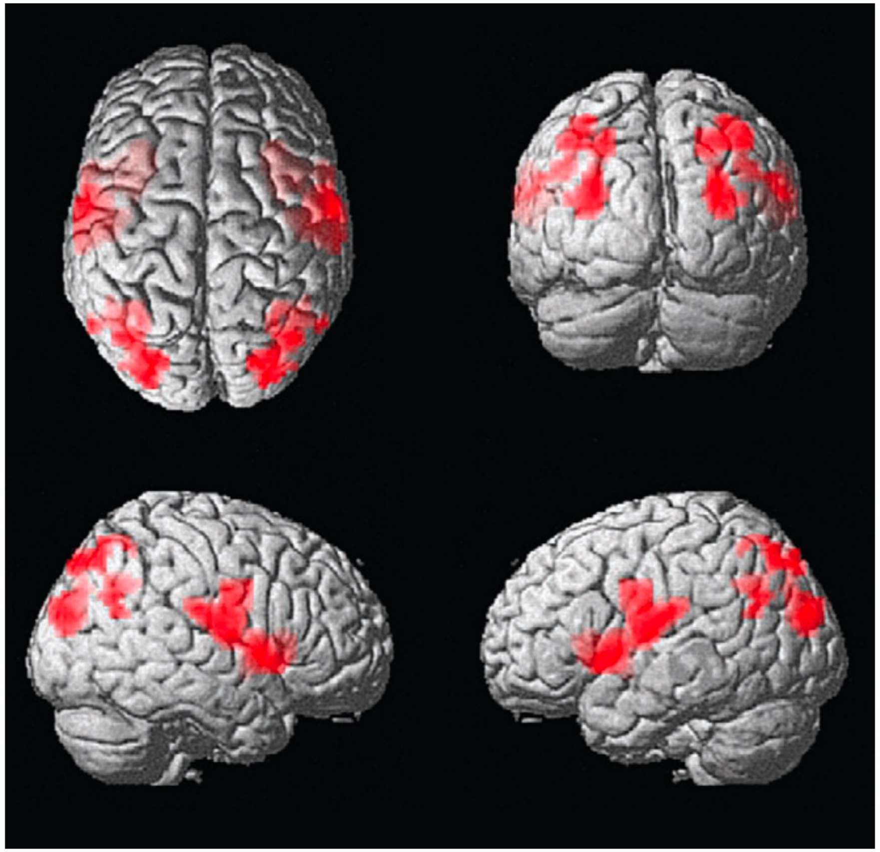


Figure from Anderson et al., 2011

FUNCTIONAL DIFFERENCES IN ASD

- Majority of changes seen in ASD (Rudie et al., 2013)
- Decreased functional connectivity throughout the brain
  - Resting-state network (Cherkassky et al., 2006)
  - Anterior-posterior connections (Cherkassky et al., 2006)
  - Fronto-parietal connections (Just et al., 2007)
  - Within functional systems (Rudie et al., 2013)
  - Interhemispheric connections (Anderson et al., 2011)
- No deactivation of resting-state network (Kennedy et al., 2006)

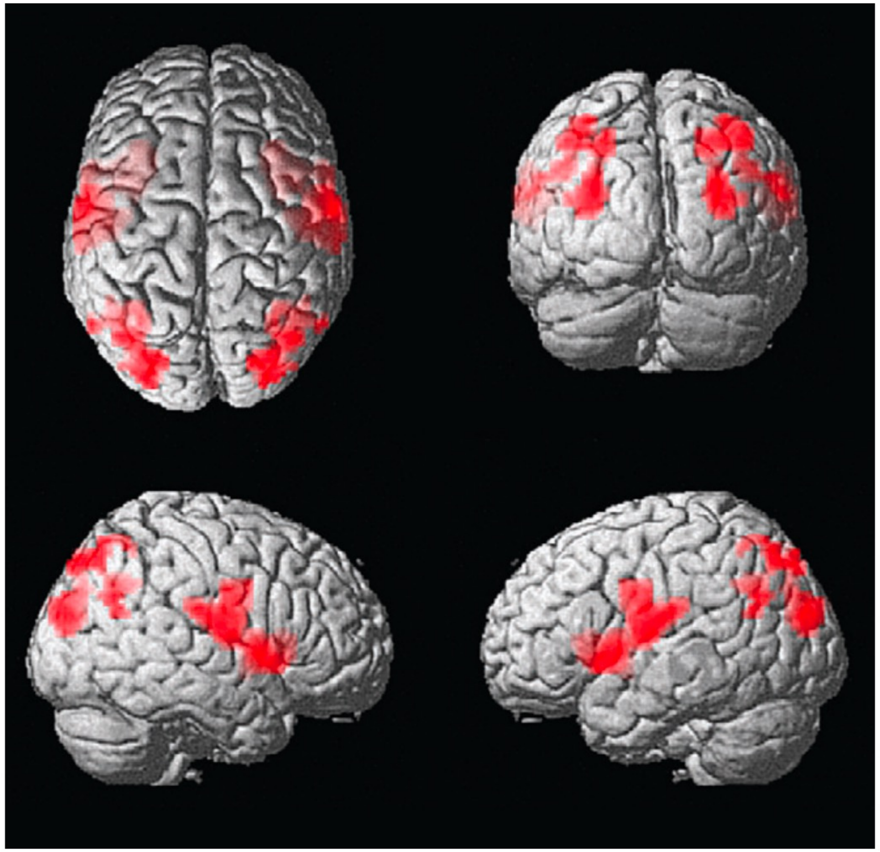


Figure from Anderson et al., 2011

FUNCTIONAL DIFFERENCES IN ASD

- Majority of changes seen in ASD (Rudie et al., 2013)
- Decreased functional connectivity throughout the brain
  - Resting-state network (Cherkassky et al., 2006)
  - Anterior-posterior connections (Cherkassky et al., 2006)
  - Fronto-parietal connections (Just et al., 2007)
  - Within functional systems (Rudie et al., 2013)
  - Interhemispheric connections (Anderson et al., 2011)
- No deactivation of resting-state network (Kennedy et al., 2006)

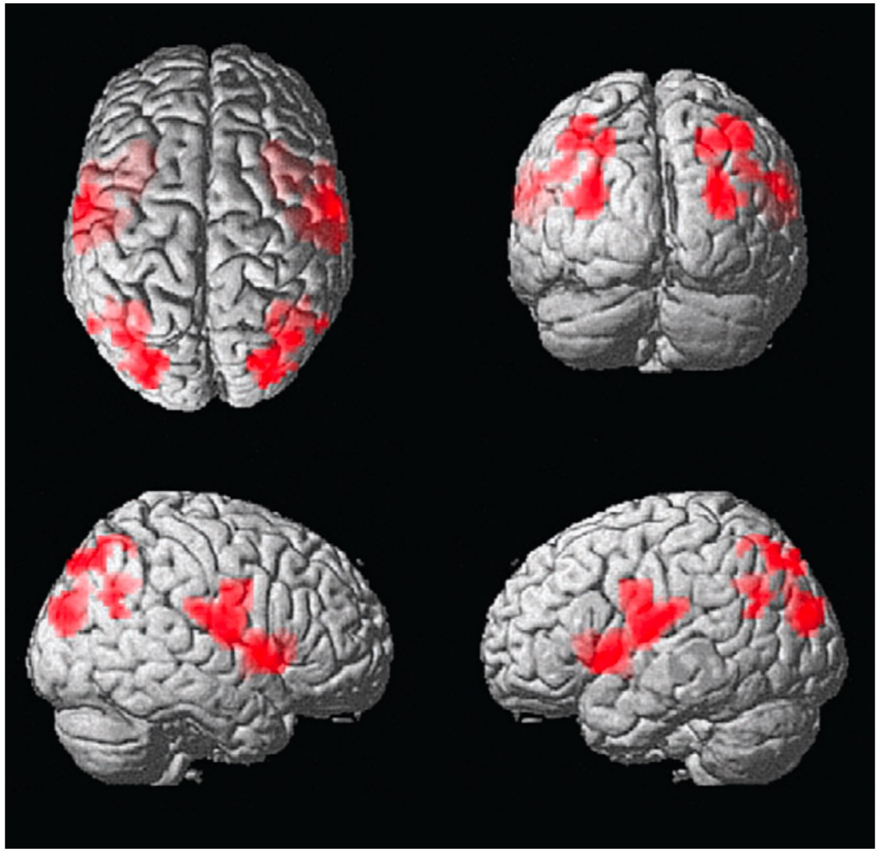


Figure from Anderson et al., 2011

FUNCTIONAL DIFFERENCES IN ASD

- Majority of changes seen in ASD (Rudie et al., 2013)
- Decreased functional connectivity throughout the brain
  - Resting-state network (Cherkassky et al., 2006)
  - Anterior-posterior connections (Cherkassky et al., 2006)
  - Fronto-parietal connections (Just et al., 2007)
  - Within functional systems (Rudie et al., 2013)
  - Interhemispheric connections (Anderson et al., 2011)
- No deactivation of resting-state network (Kennedy et al., 2006)

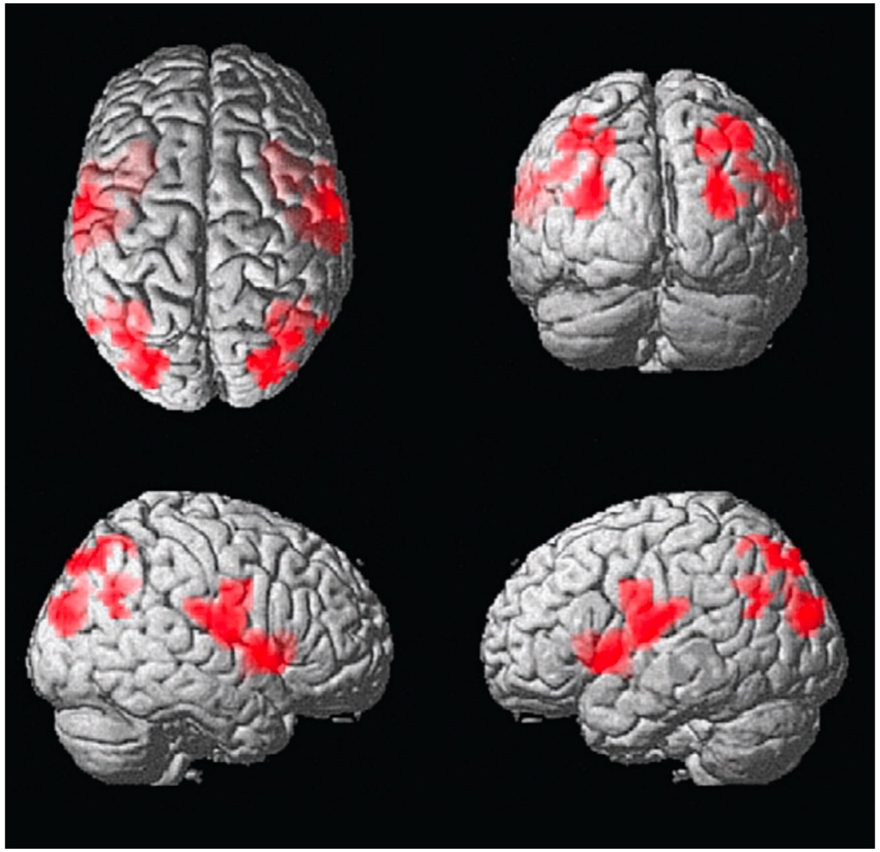


Figure from Anderson et al., 2011

STRUCTURAL DIFFERENCES IN ASD

- Reduced corpus callosum volume (Anderson et al., 2011; Chen et al., 2011; Pagnozzi et al., 2018)
- Increased frontal lobe volume (Amaral et al., 2008; Chen et al., 2011; Pagnozzi et al., 2018)
- Increased cortical thickness in frontal and parietal lobes (Chen et al., 2011; Pagnozzi et al., 2018)
- great heterogeneity otherwise (Amaral et al., 2008)

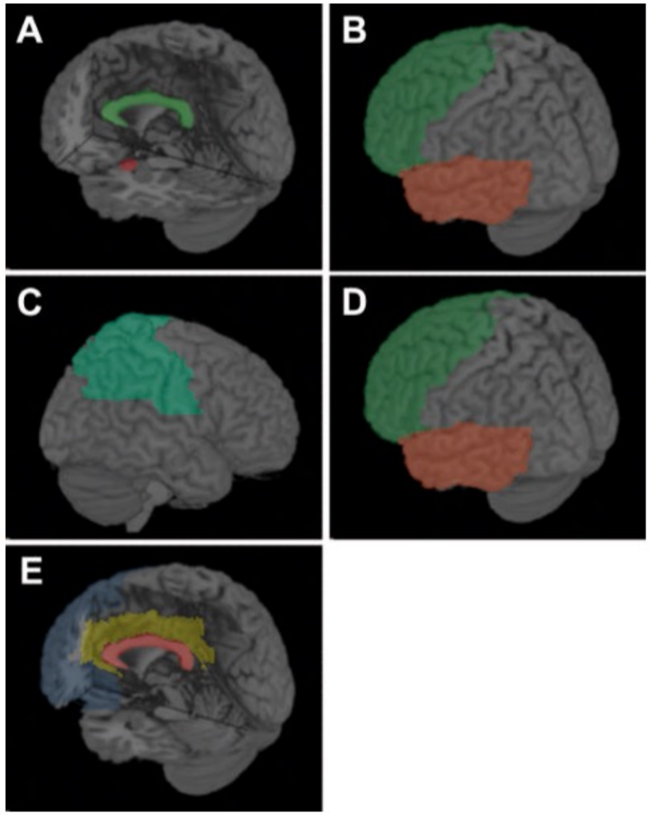


Figure from Chen et al., 2011

STRUCTURAL DIFFERENCES IN ASD

- Reduced corpus callosum volume (Anderson et al., 2011; Chen et al., 2011; Pagnozzi et al., 2018)
- Increased frontal lobe volume (Amaral et al., 2008; Chen et al., 2011; Pagnozzi et al., 2018)
- Increased cortical thickness in frontal and parietal lobes (Chen et al., 2011; Pagnozzi et al., 2018)
- great heterogeneity otherwise (Amaral et al., 2008)

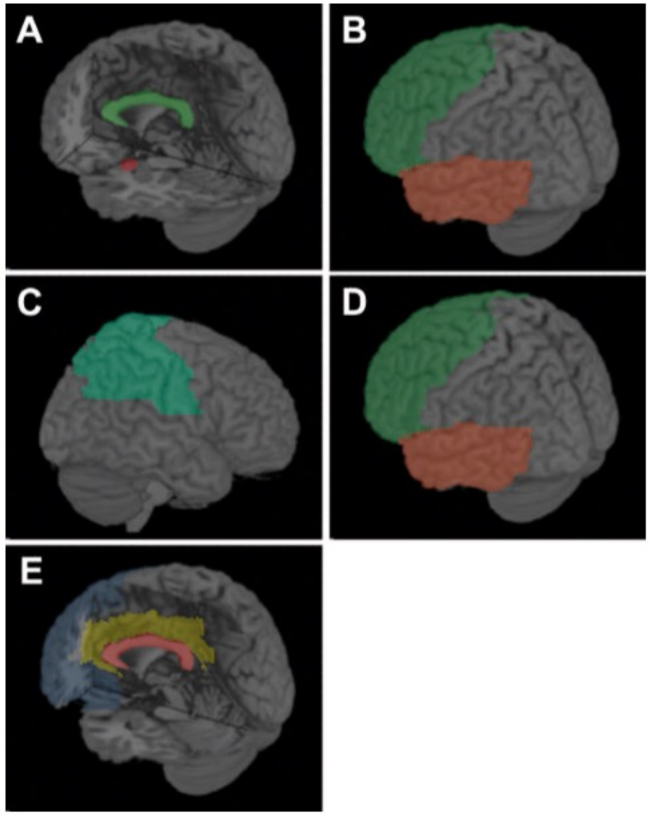


Figure from Chen et al., 2011

STRUCTURAL DIFFERENCES IN ASD

- Reduced corpus callosum volume (Anderson et al., 2011; Chen et al., 2011; Pagnozzi et al., 2018)
- Increased frontal lobe volume (Amaral et al., 2008; Chen et al., 2011; Pagnozzi et al., 2018)
- Increased cortical thickness in frontal and parietal lobes (Chen et al., 2011; Pagnozzi et al., 2018)
- great heterogeneity otherwise (Amaral et al., 2008)

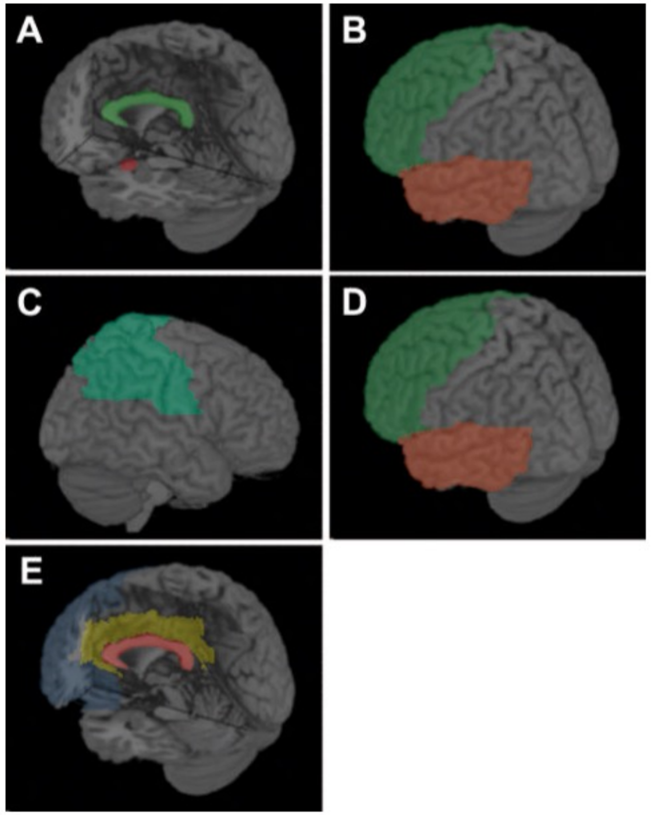


Figure from Chen et al., 2011

STRUCTURAL DIFFERENCES IN ASD

- Reduced corpus callosum volume (Anderson et al., 2011; Chen et al., 2011; Pagnozzi et al., 2018)
- Increased frontal lobe volume (Amaral et al., 2008; Chen et al., 2011; Pagnozzi et al., 2018)
- Increased cortical thickness in frontal and parietal lobes (Chen et al., 2011; Pagnozzi et al., 2018)
- great heterogeneity otherwise (Amaral et al., 2008)

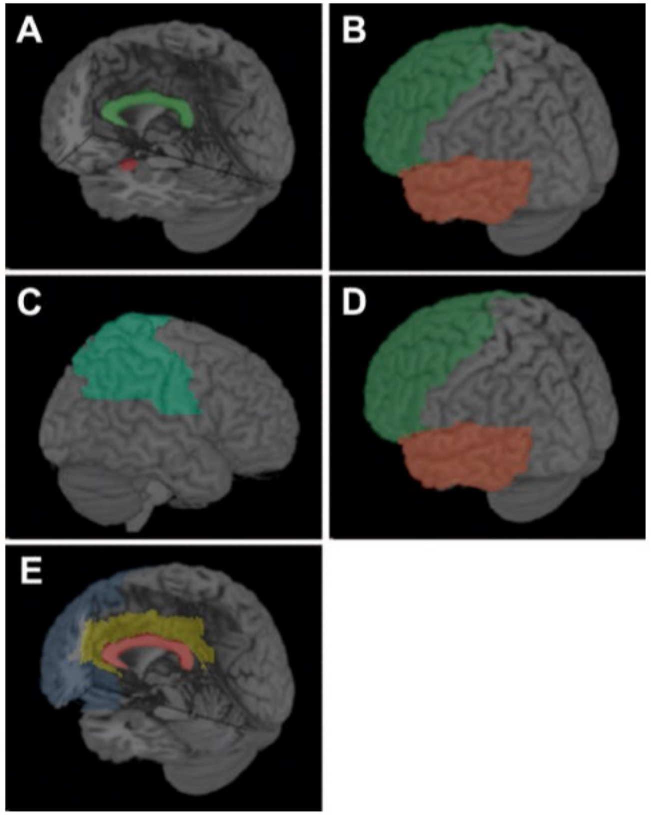


Figure from Chen et al., 2011

ASD DIAGNOSIS USING MACHINE LEARNING

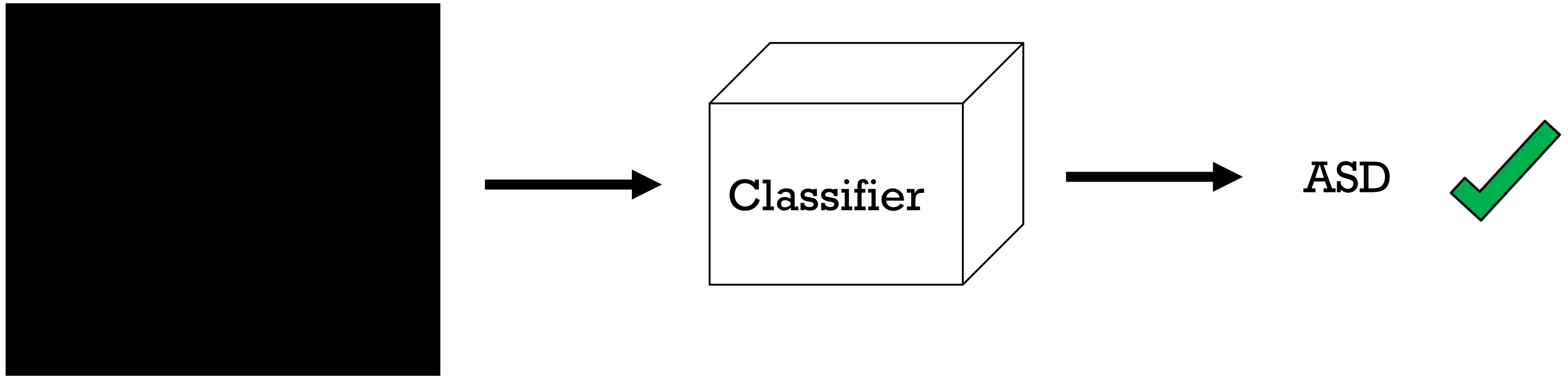
- Accuracy of 70% using 3-layer fully connected NN with functional connectivity features from fMRI
  - Autism Brain Imaging Data Exchange (ABIDE)

ASD DIAGNOSIS USING MACHINE LEARNING

- Accuracy of 70% using 3-layer fully connected NN with functional connectivity features from fMRI
  - Autism Brain Imaging Data Exchange (ABIDE)

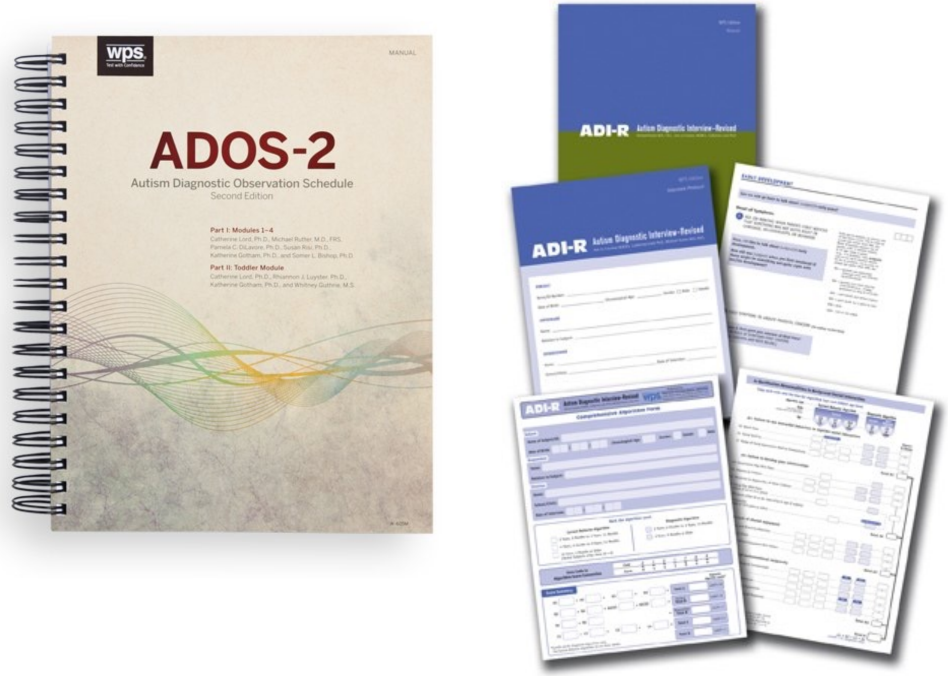
ASD DIAGNOSIS USING MACHINE LEARNING

- Aim 1: classify ASD from neuroimaging data using the following classification methods:
  - Nearest neighbor (NN)
  - Naïve-Bayes
  - Linear discriminant analysis (LDA)
  - Residual neural networks (RNNs)



ASD DIAGNOSIS USING MACHINE LEARNING

- Aim 2. link behavioral assessments scores to neuroimaging differences

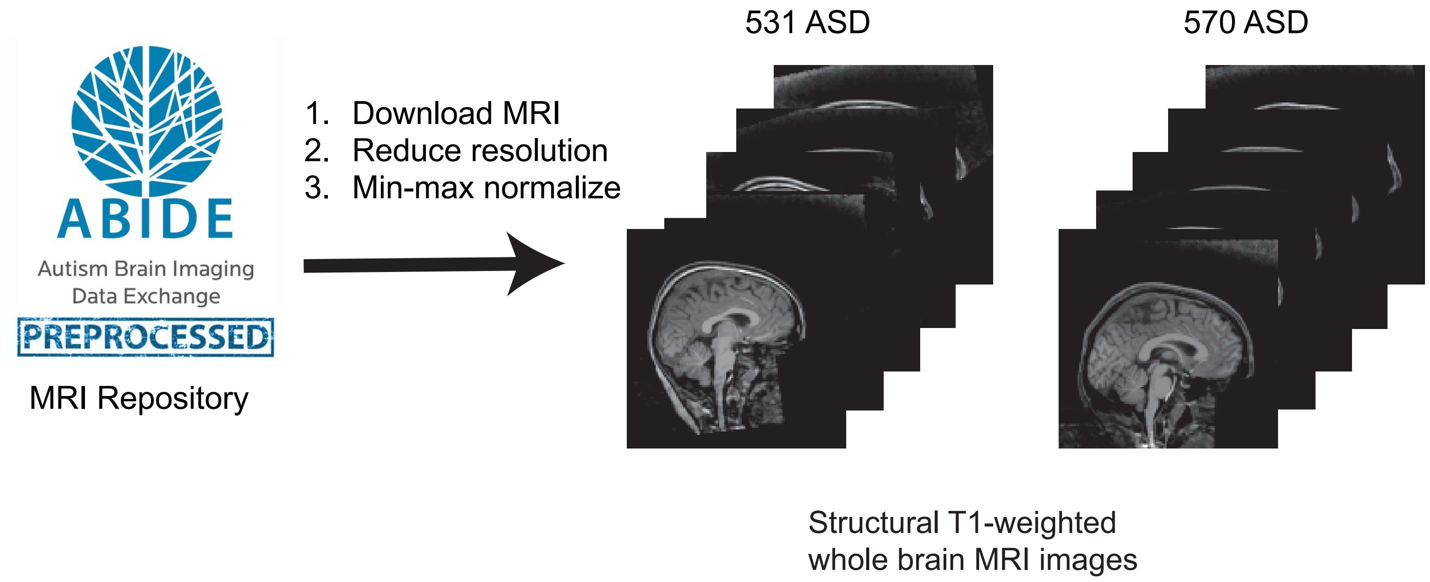


Pictures from wpspublish.com and stoeltingco.com

OUTLINE

- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
  - Neural Network in Medical Imaging
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

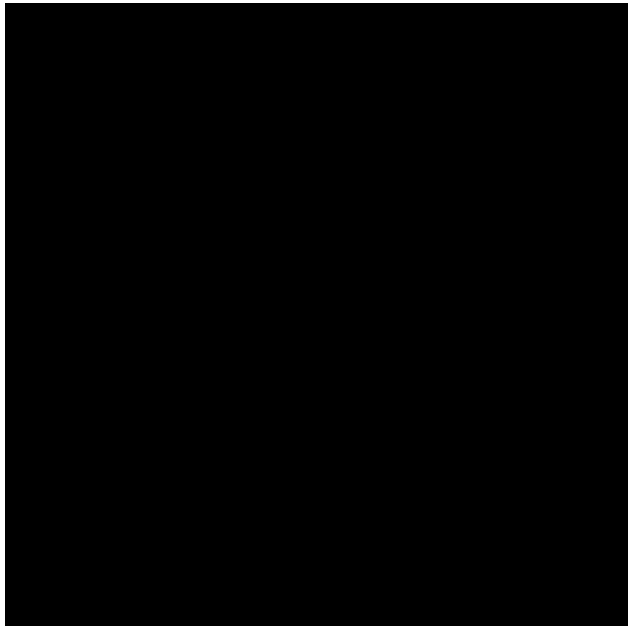
DOWNLOAD AND PREPROCESS



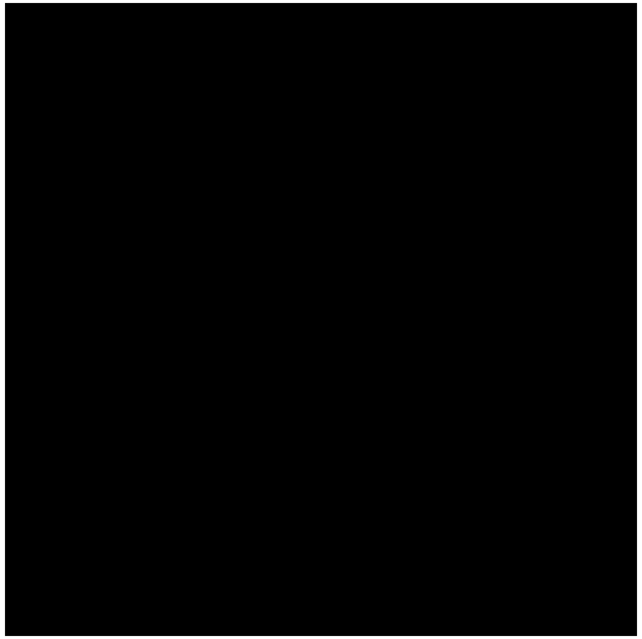
DOWNLOAD AND PREPROCESS

Cortical Thickness

Structural



Functional



CLASSIFIERS

- NN, Naïve-Bayes, LDA
- PCA to 1100 features
- 8-fold cross validation
- Ensembles
- Distance to bound analysis

CLASSIFIERS

- NN, Naïve-Bayes, LDA
- PCA to 1100 features
- 8-fold cross validation
- Ensembles
- Distance to bound analysis

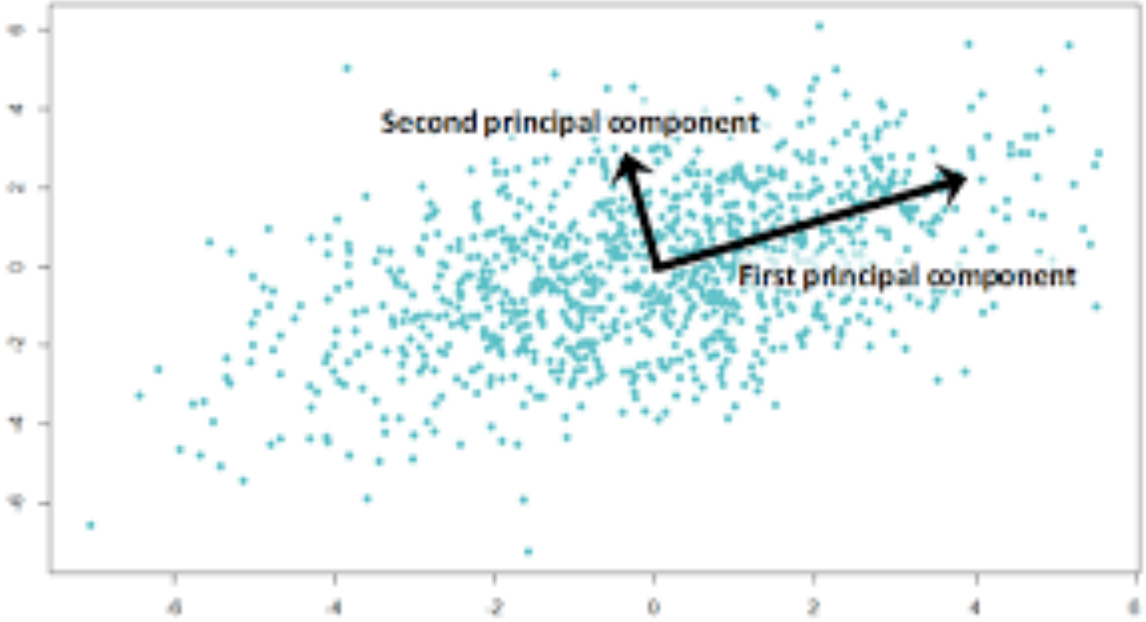
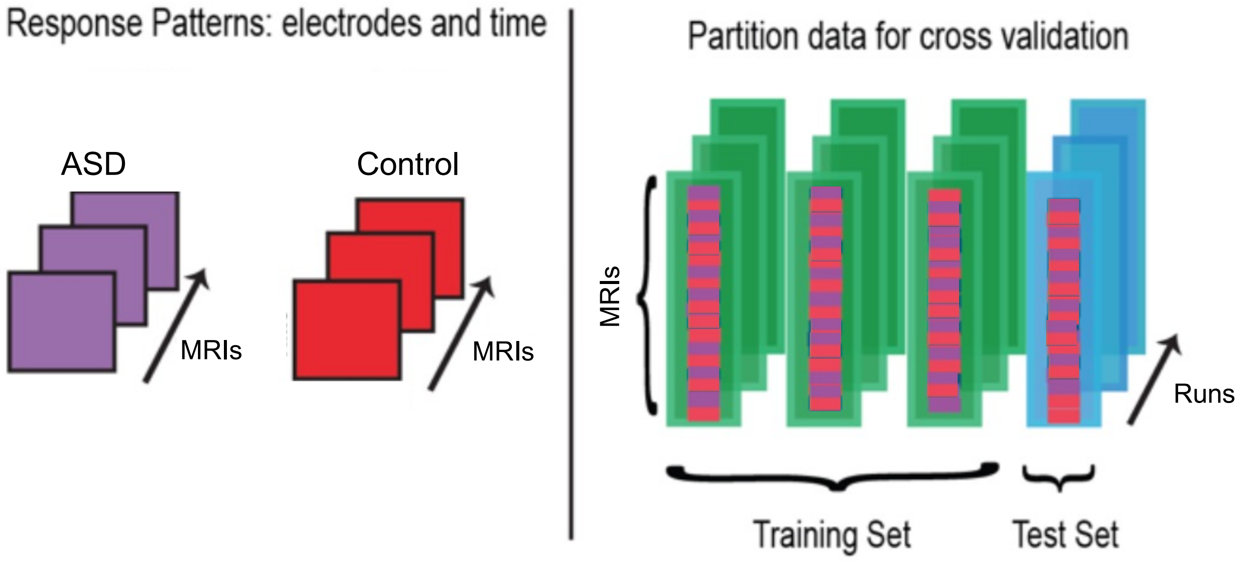


Figure from analyticsvidhya.com

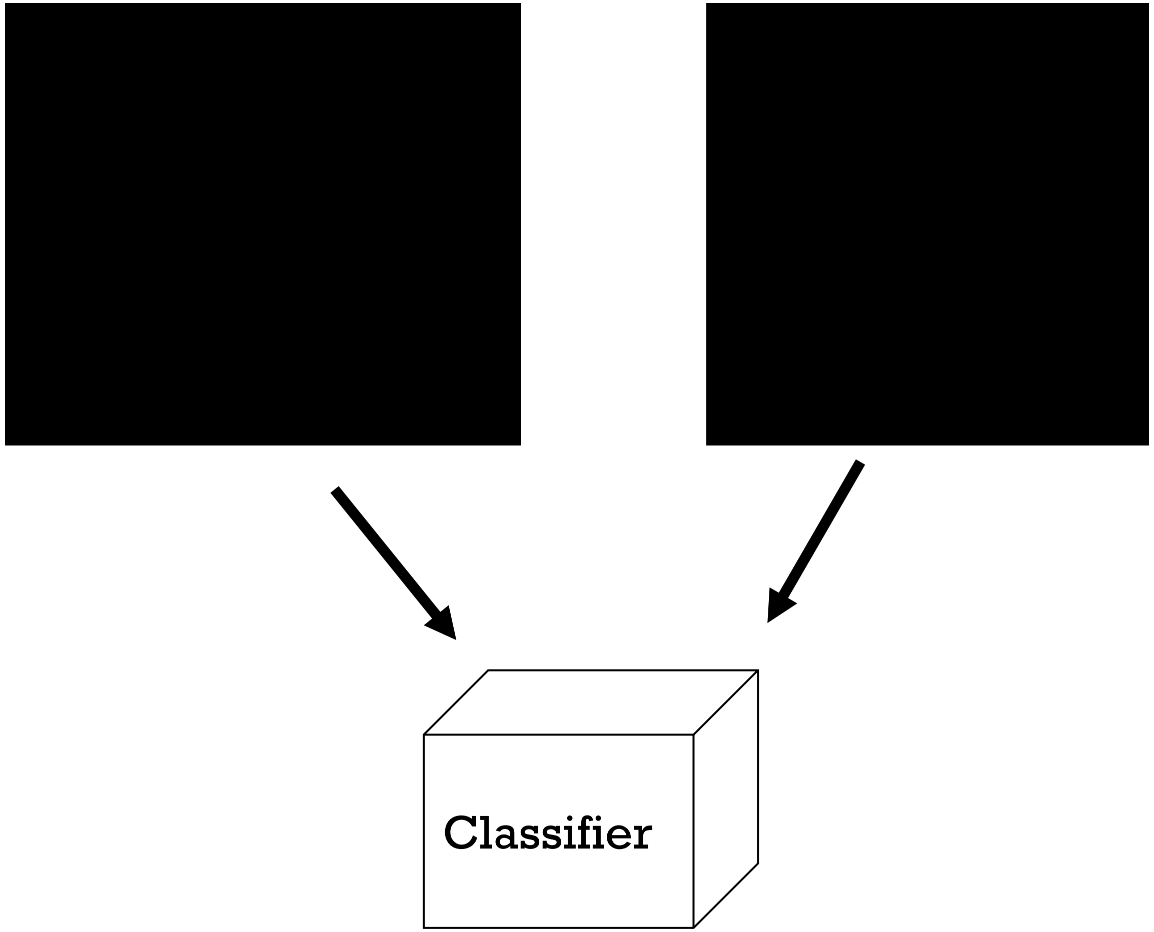
CLASSIFIERS

- NN, Naïve-Bayes, LDA
- PCA to 1100 features
- 8-fold cross validation
- Ensembles
- Distance to bound analysis



CLASSIFIERS

- NN, Naïve-Bayes, LDA
- PCA to 1100 features
- 8-fold cross validation
- Ensembles
- Distance to bound analysis



CLASSIFIERS

- NN, Naïve-Bayes, LDA
- PCA to 1100 features
- 8-fold cross validation
- Ensembles
- Distance to bound analysis

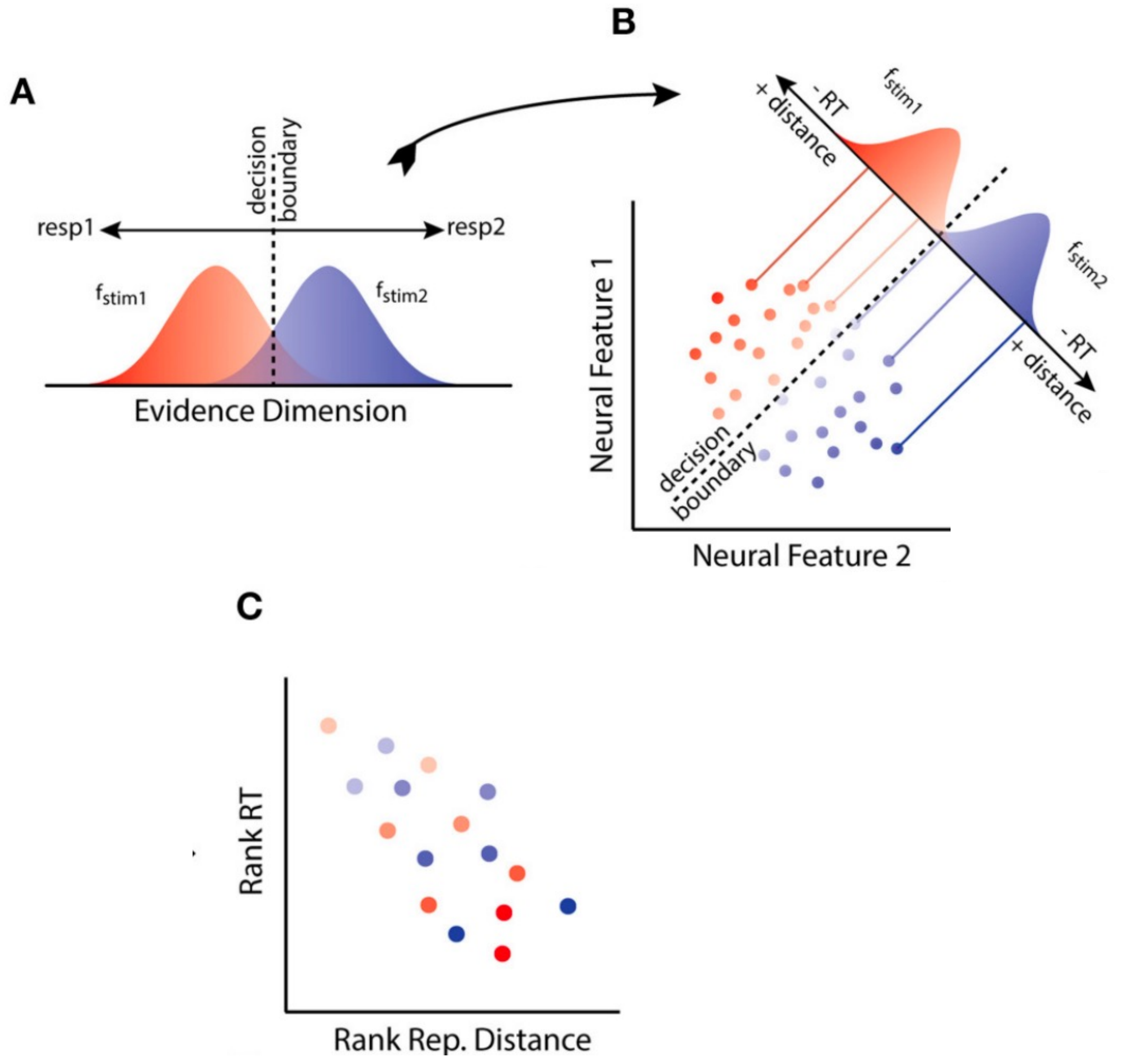
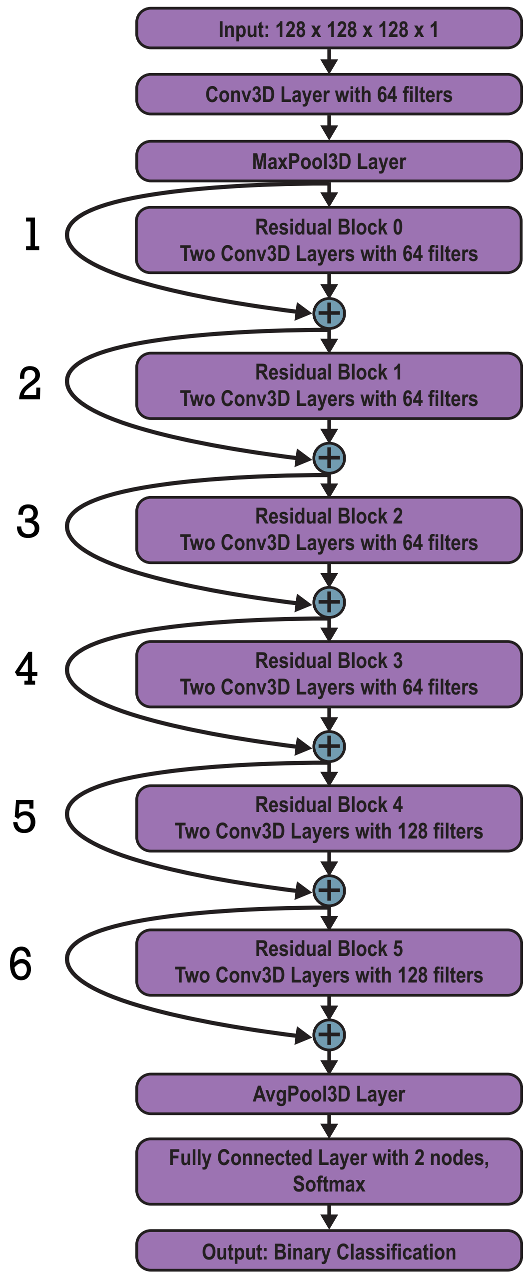


Figure from Ritchie et al., 2016

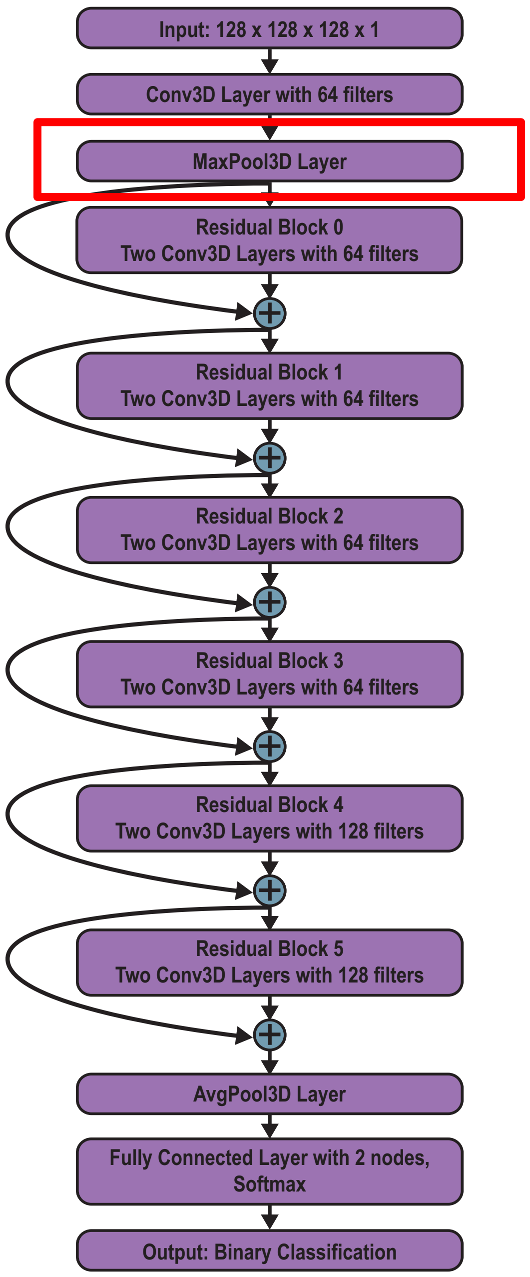
ASDNET

- 6 residual blocks
- Max pooling before
- Average pooling after
- Fully connected at end
- Softmax
- 5-fold cross validation



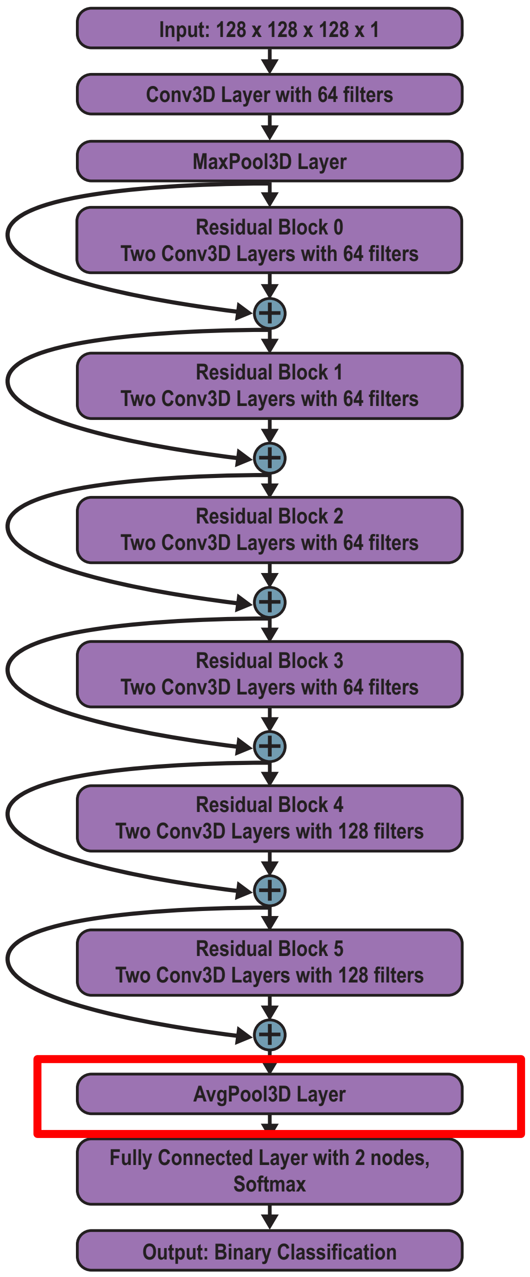
ASDNET

- 6 residual blocks
- Max pooling before
- Average pooling after
- Fully connected at end
- Softmax
- 5-fold cross validation



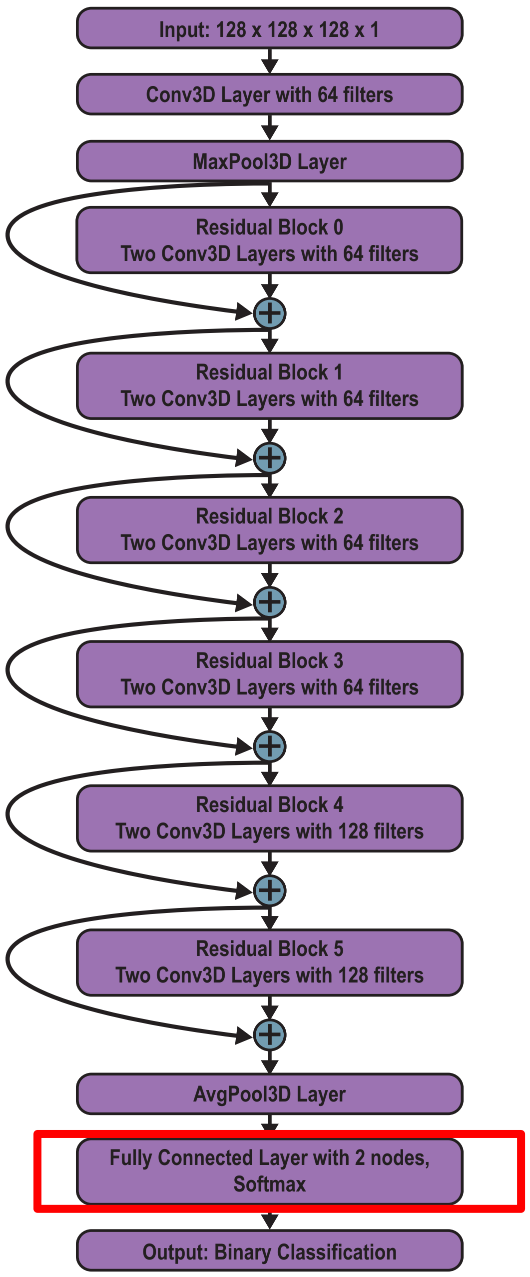
ASDNET

- 6 residual blocks
- Max pooling before
- Average pooling after
- Fully connected at end
- Softmax
- 5-fold cross validation



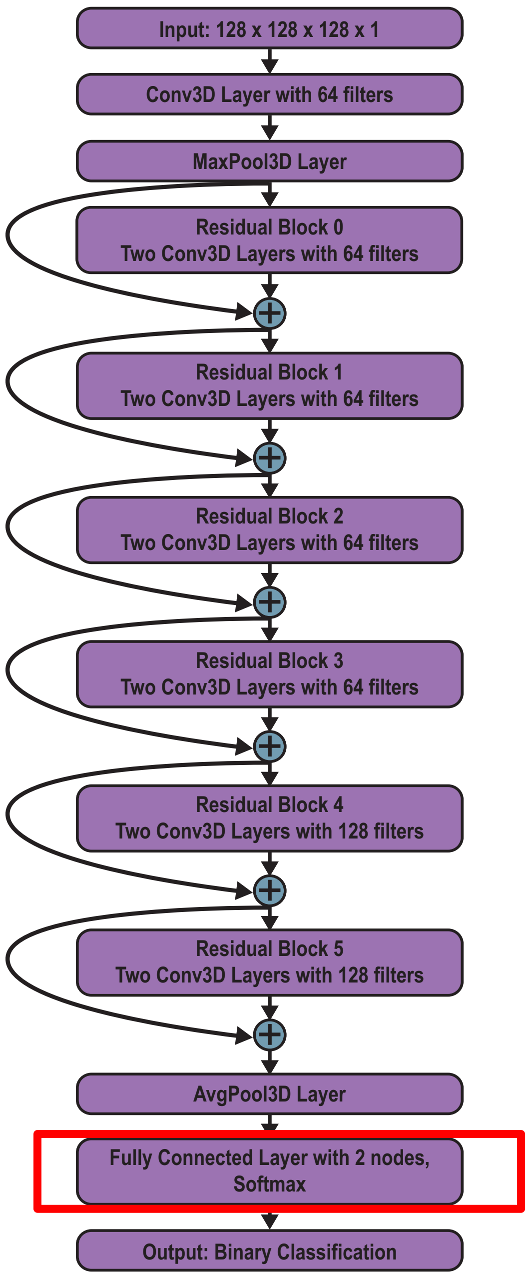
ASDNET

- 6 residual blocks
- Max pooling before
- Average pooling after
- Fully connected at end
- Softmax
- 5-fold cross validation



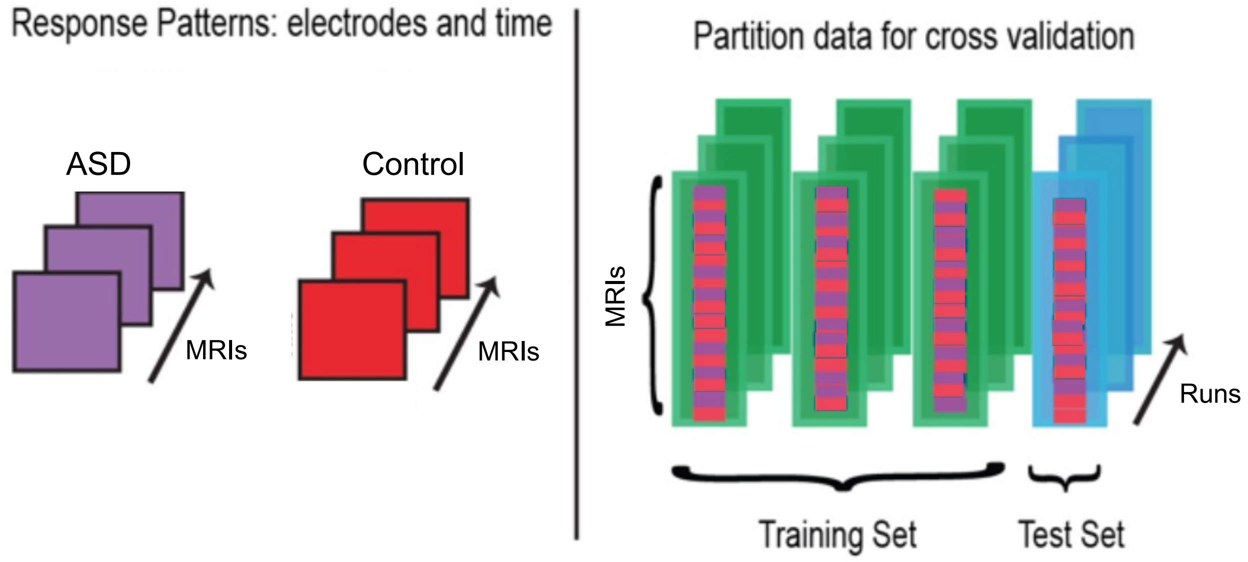
ASDNET

- 6 residual blocks
- Max pooling before
- Average pooling after
- Fully connected at end
- Softmax
- 5-fold cross validation



ASDNET

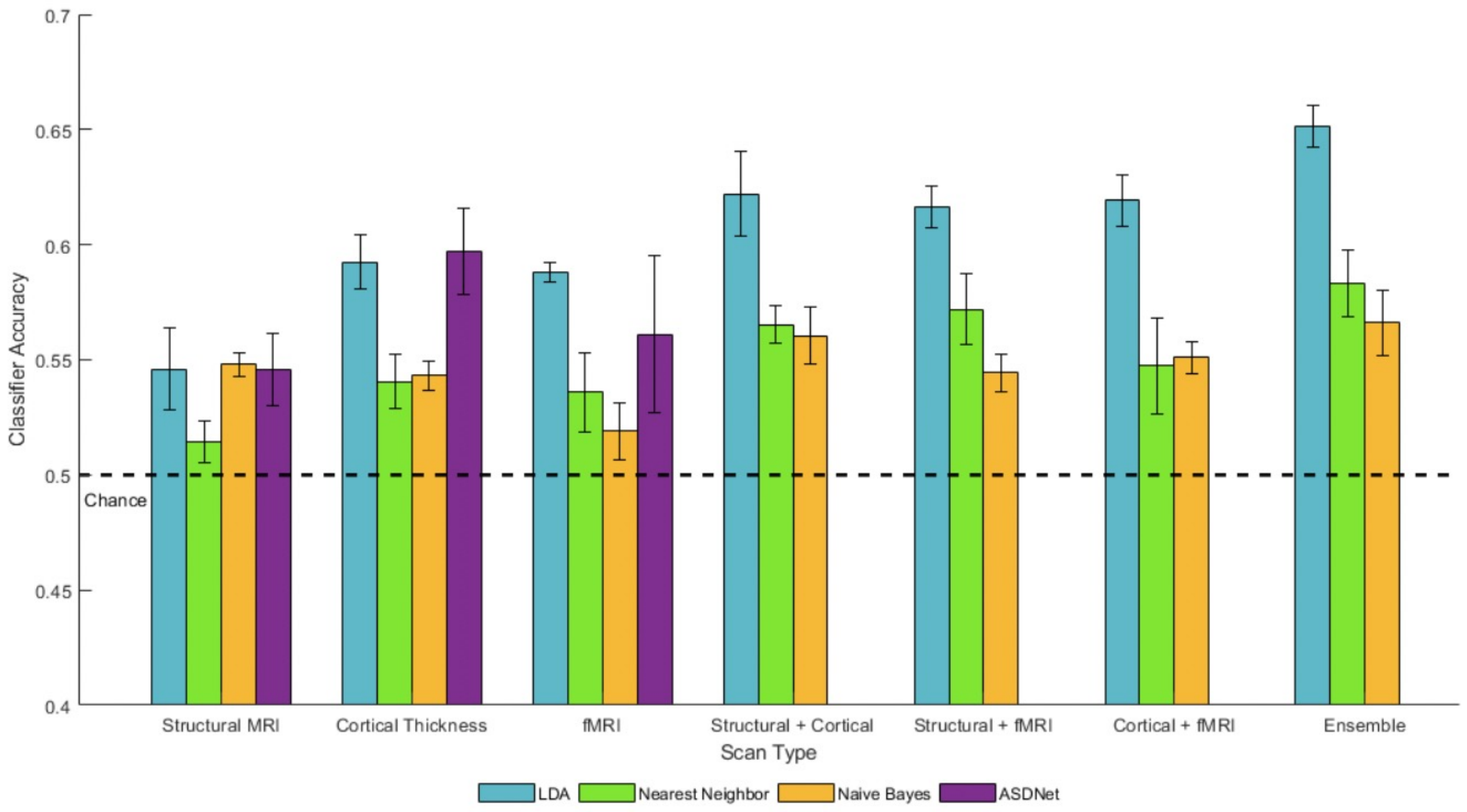
- 6 residual blocks
- Max pooling before
- Average pooling after
- Fully connected at end
- Softmax
- 5-fold cross validation



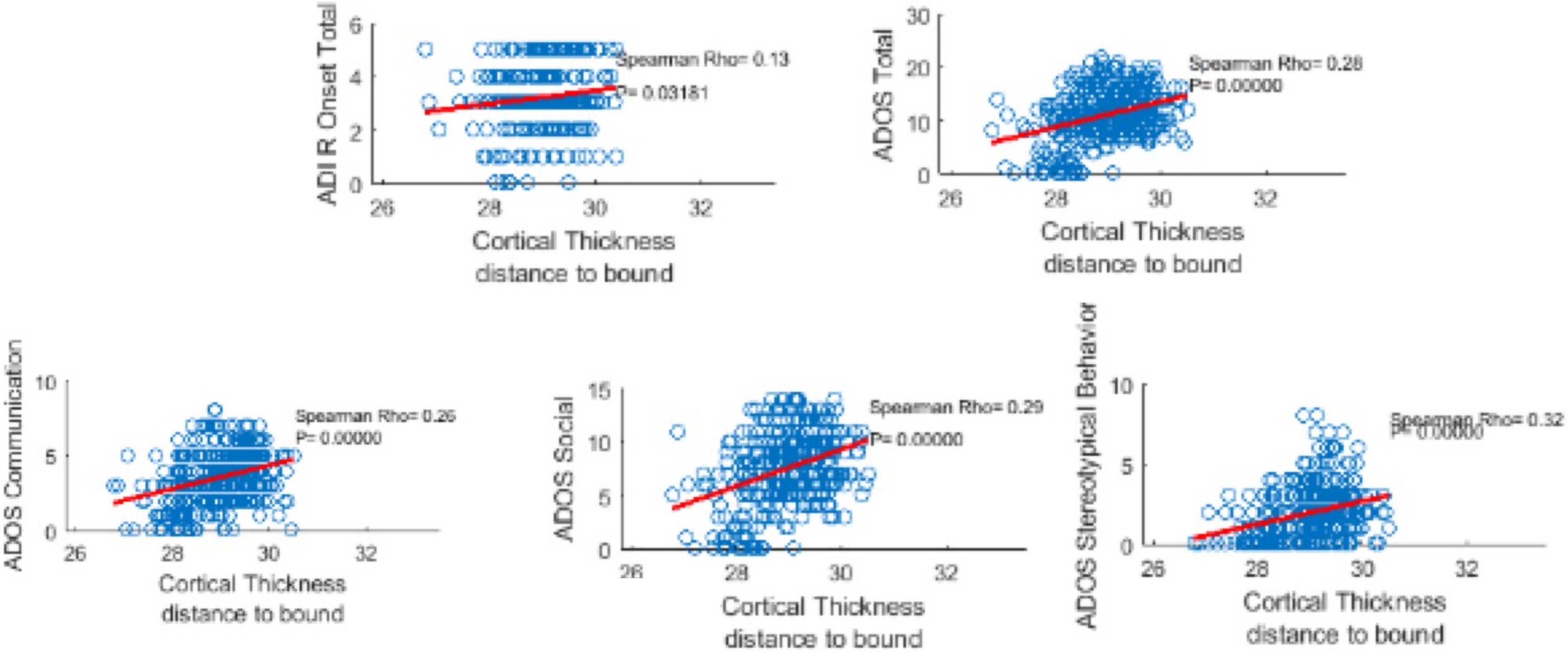
OUTLINE

- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
  - Neural Network in Medical Imaging
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

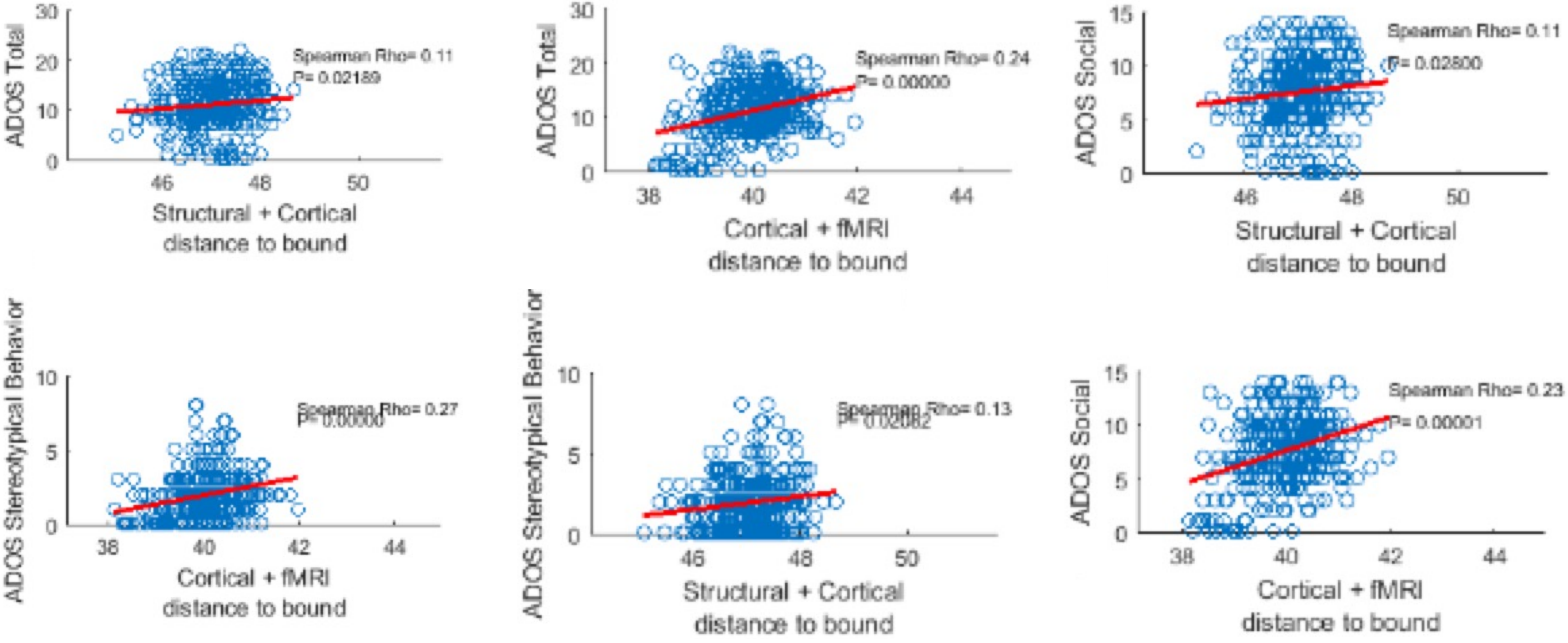
CLASSIFICATION PERFORMANCE



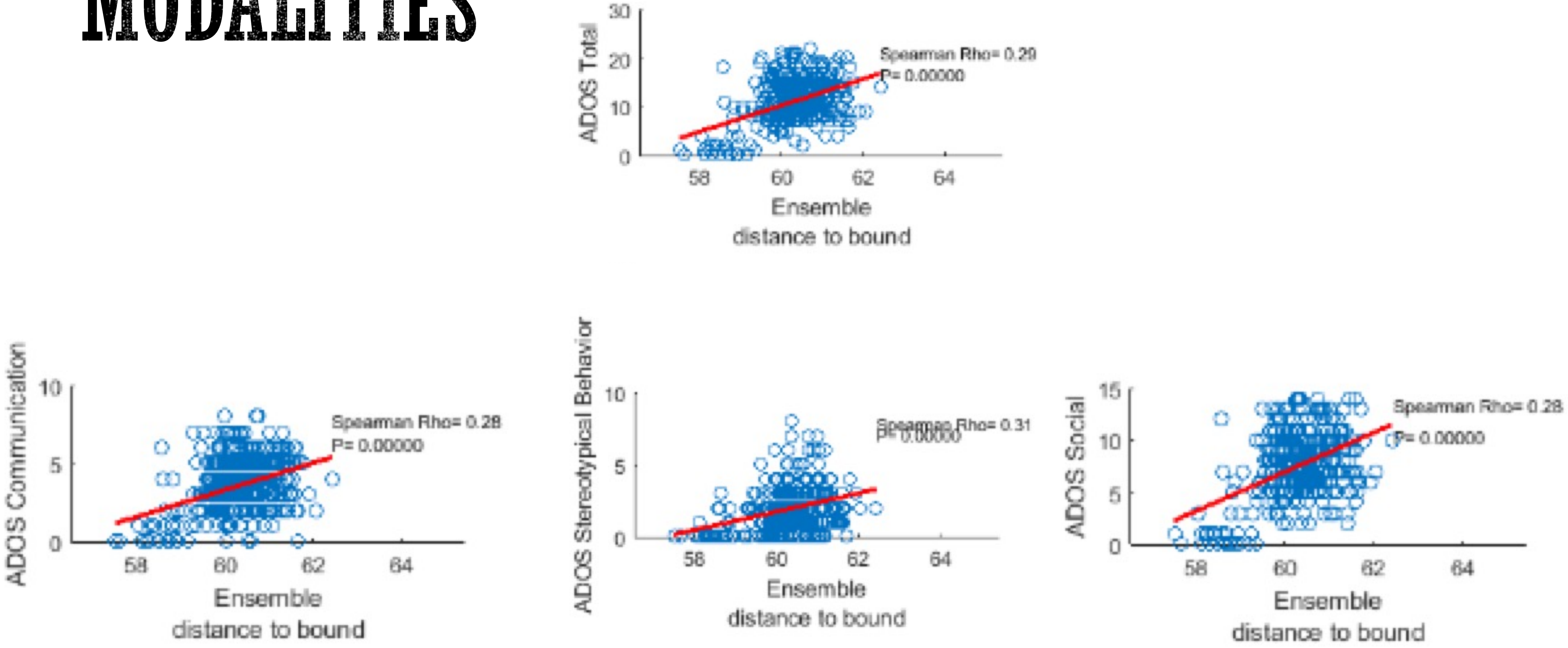
DISTANCE TO BOUND – ONE IMAGING MODALITY



DISTANCE TO BOUND – TWO IMAGING MODALITIES



DISTANCE TO BOUND – ALL IMAGING MODALITIES



OUTLINE

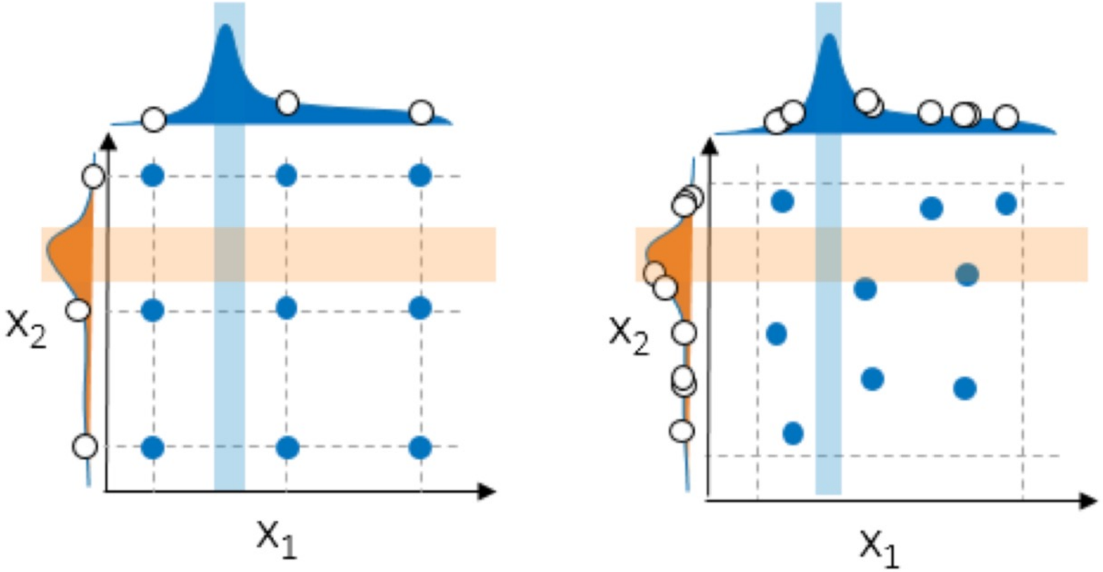
- Chapter 1: Introduction to Classification and Machine Learning
  - Classification
  - A Few Basic Classifiers
  - Artificial Neural Networks
  - Neural Network in Medical Imaging
- Chapter 2: Classification of Autism Spectrum Disorder Using Machine Learning
  - Introduction
  - Methods
  - Results
  - Discussion

FUTURE DIRECTIONS

- Improve accuracy of ASDNet further
  - Hyperparameter tuning
    - Learning rate, activation function, etc.
- Distance to bound with ASDNet
- Ensembles for ASDNet

FUTURE DIRECTIONS

- Improve accuracy of ASDNet further
  - Hyperparameter tuning
    - Learning rate, activation function, etc.
- Distance to bound with ASDNet
- Ensembles for ASDNet

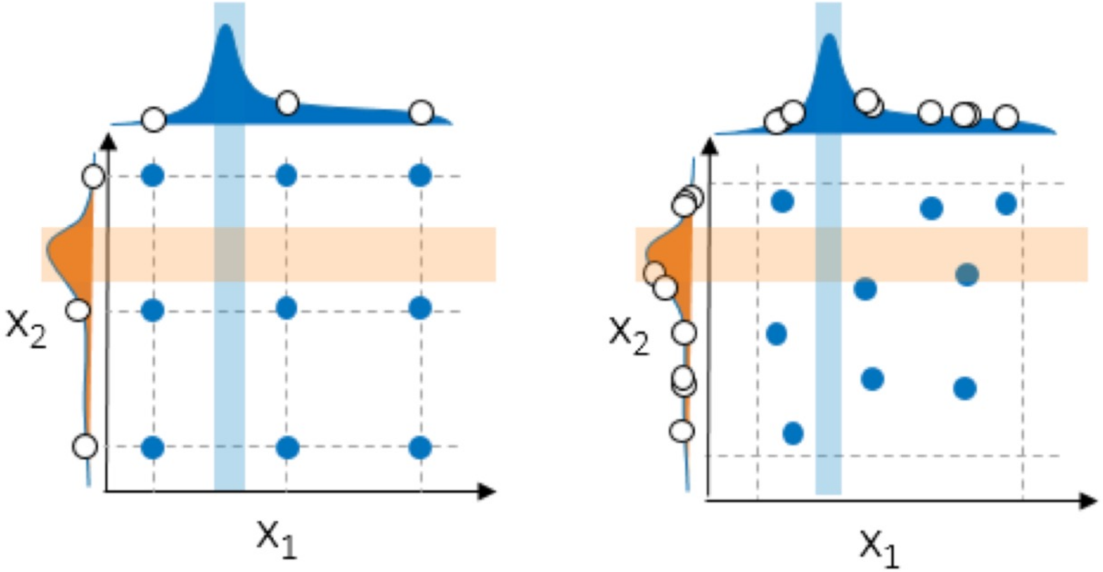


(a) Standard Grid Search (b) Random Search

Figure from support.sas.com

FUTURE DIRECTIONS

- Improve accuracy of ASDNet further
  - Hyperparameter tuning
    - Learning rate, activation function, etc.
- Distance to bound with ASDNet
- Ensembles for ASDNet



(a) Standard Grid Search (b) Random Search

Figure from support.sas.com

FUTURE DIRECTIONS

- Improve accuracy of ASDNet further
  - Hyperparameter tuning
    - Learning rate, activation function, etc.
- Distance to bound with ASDNet
- Ensembles for ASDNet

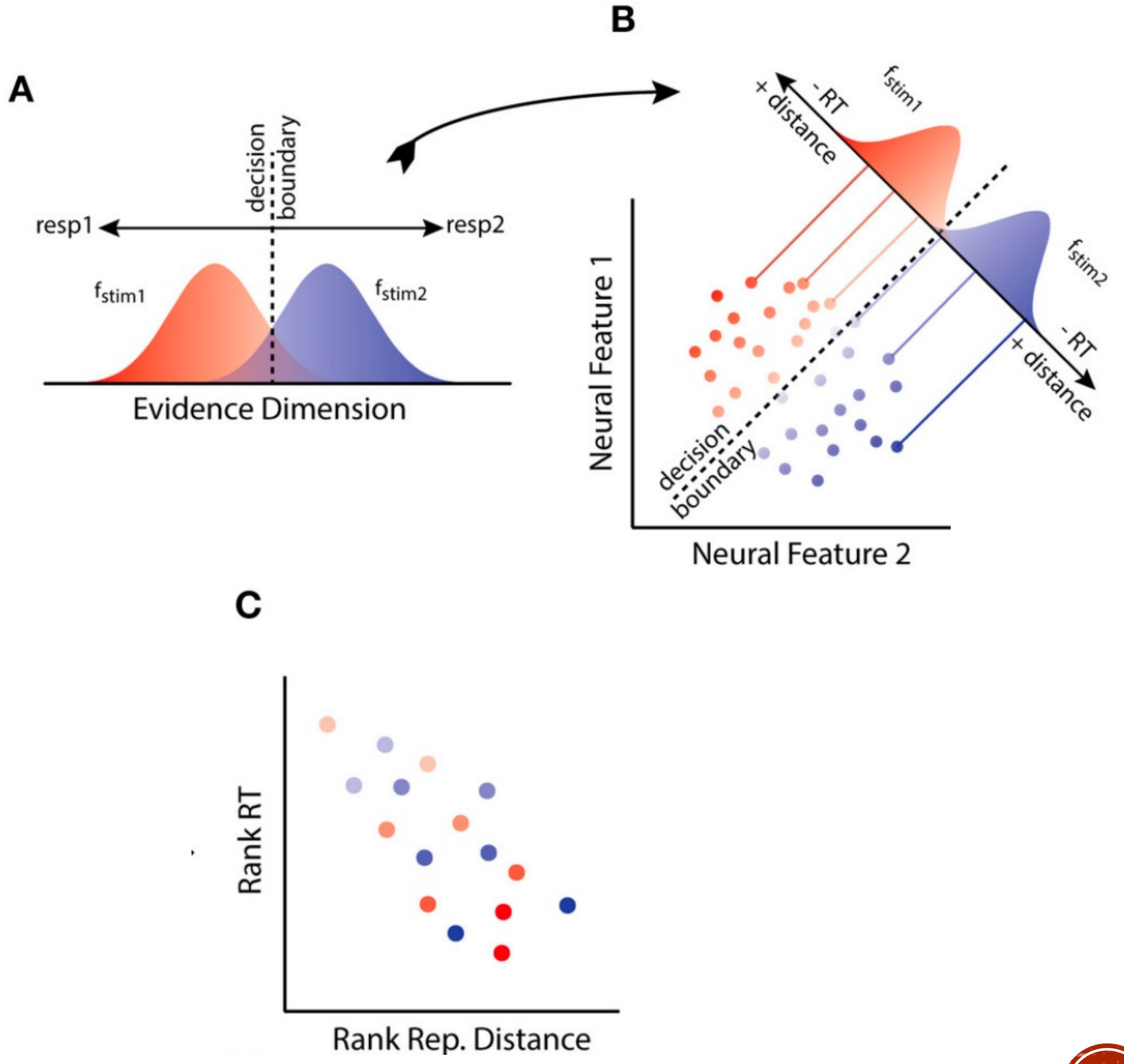
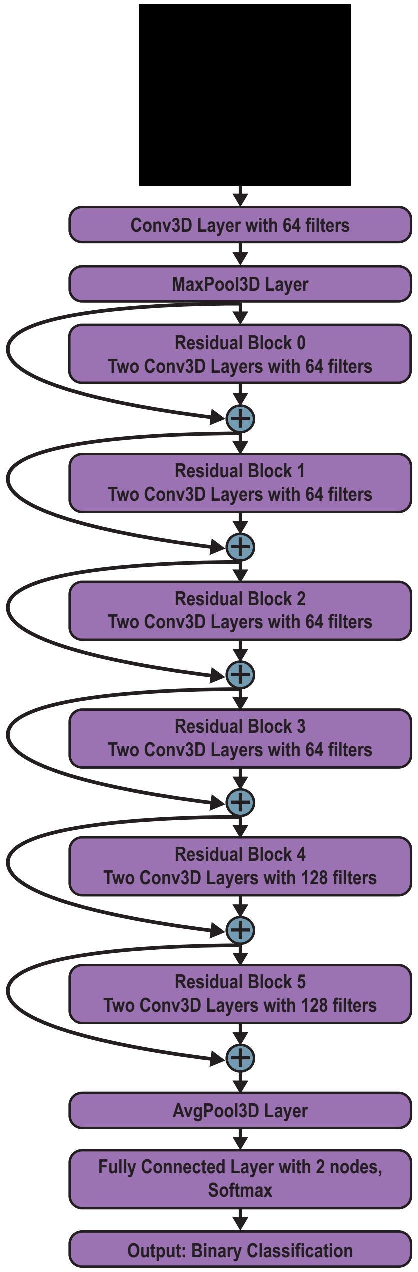
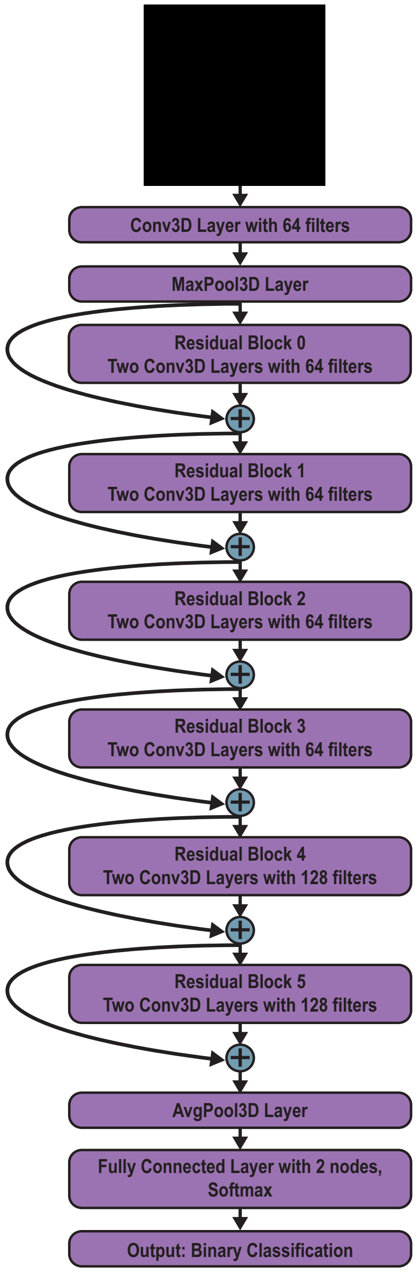


Figure from Ritchie et al., 2016

FUTURE DIRECTIONS

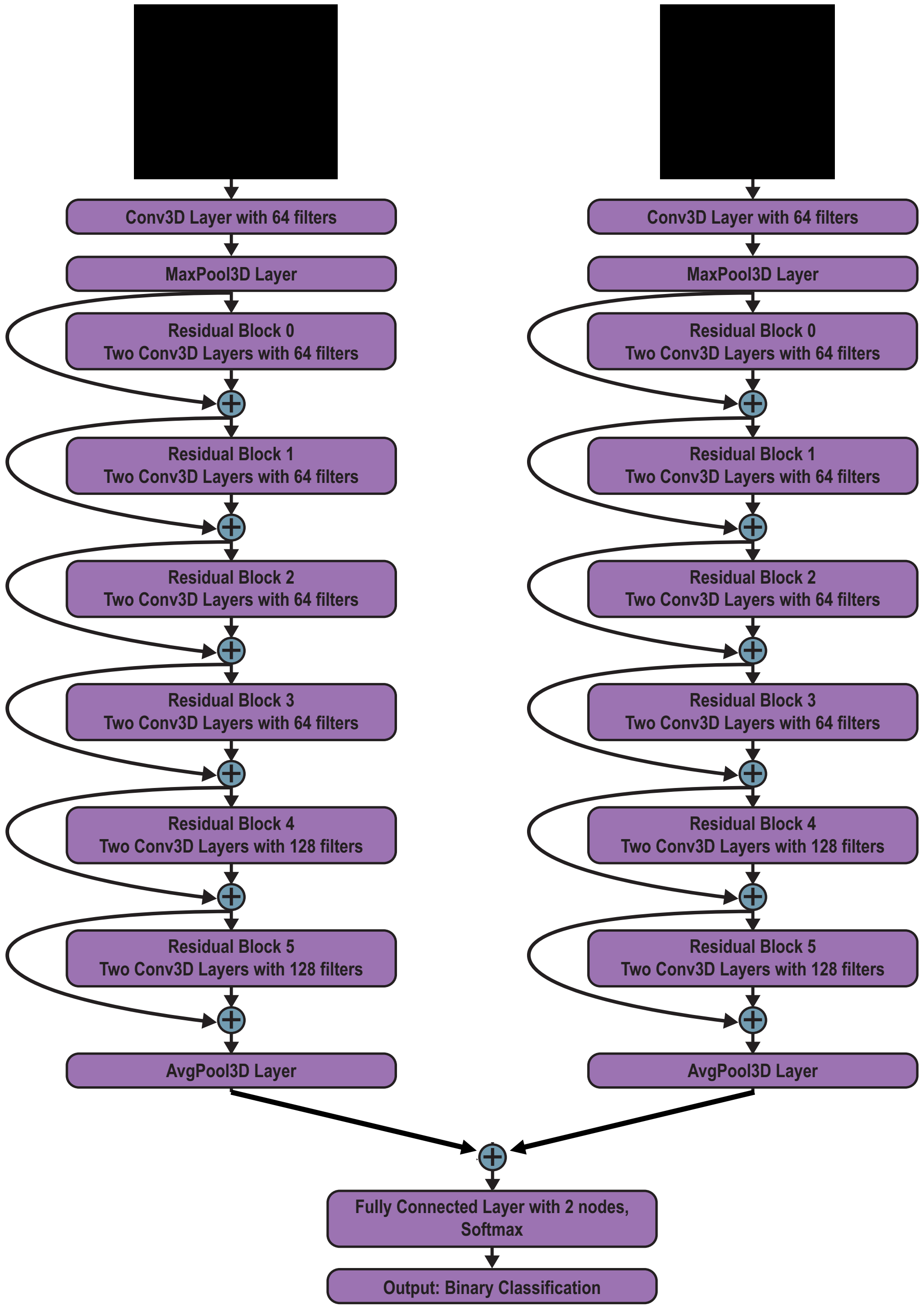
- Improve accuracy of ASDNet further
  - Hyperparameter tuning
    - Learning rate, activation function, etc.
- Distance to bound with ASDNet
- Ensembles for ASDNet



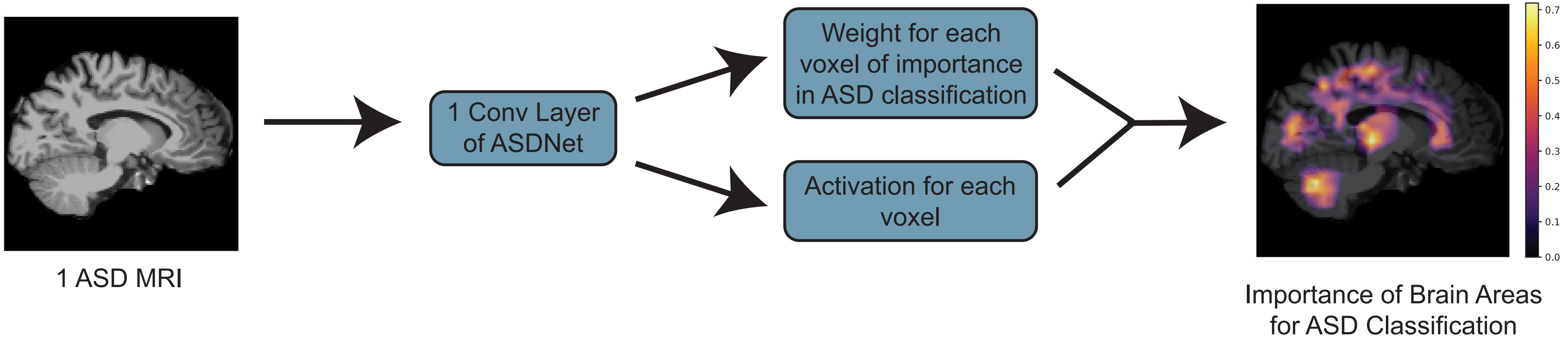


FUTURE DIRECTIONS

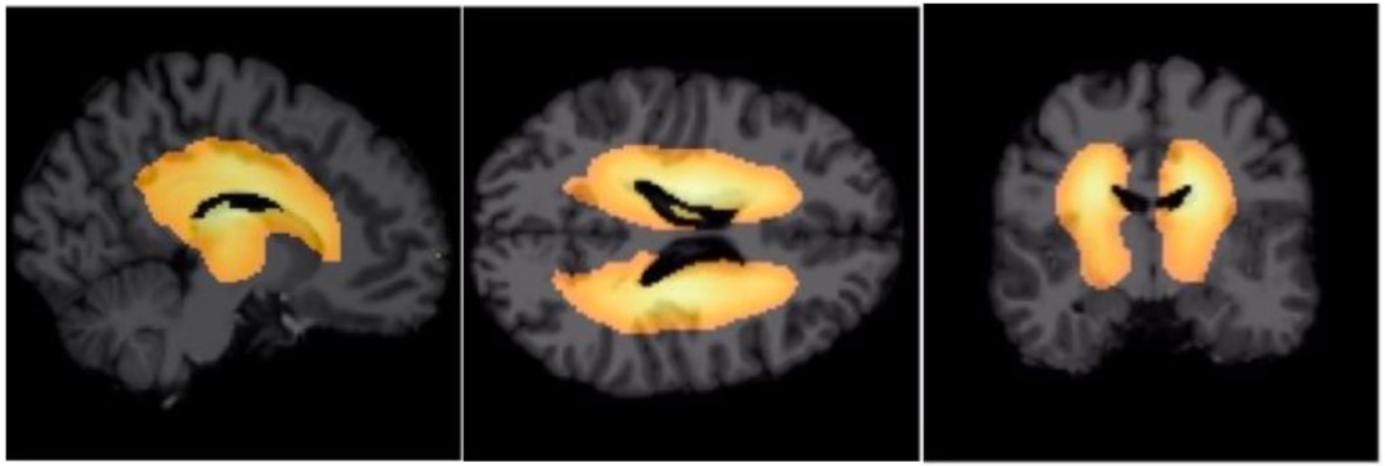
- Improve accuracy of ASDNet further
  - Hyperparameter tuning
    - Learning rate, activation function, etc.
- Distance to bound with ASDNet
- Ensembles for ASDNet



FUTURE DIRECTIONS – VISUALIZATION WITH 3D-GRAD-CAM

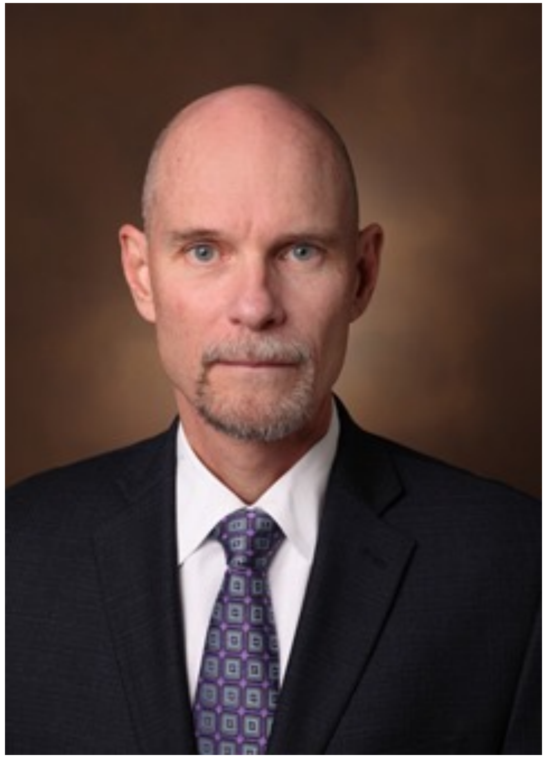


FUTURE DIRECTIONS – VISUALIZATION WITH 3D-GRAD-CAM



ACKNOWLEDGMENTS

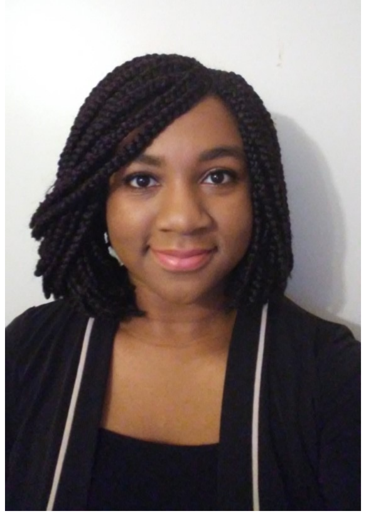
ACKNOWLEDGMENTS



ACKNOWLEDGMENTS

ACKNOWLEDGMENTS

ACKNOWLEDGMENTS



ACKNOWLEDGMENTS

THANKS FOR YOUR ATTENTION!

QUESTIONS??