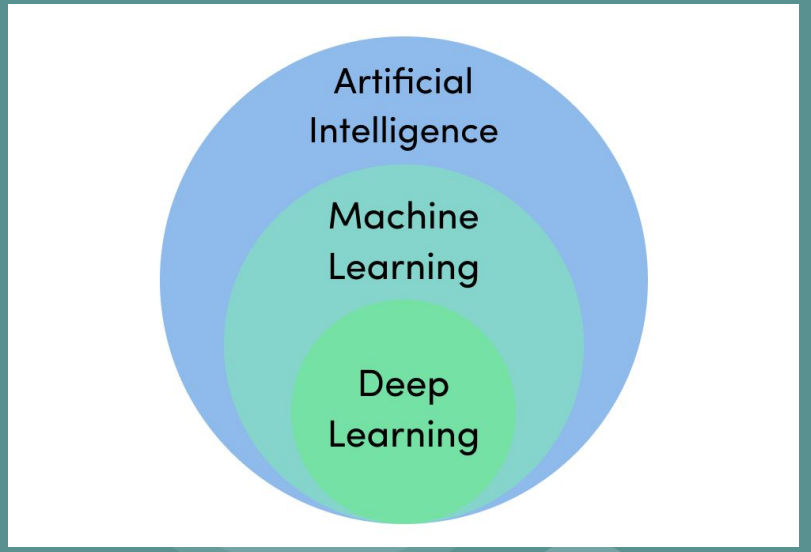
Deep Learning/AI vs. Machine Learning

Showdown between Dr. Chung and Dr. Blossom

Good Morning

Deep Learning/AI vs. Machine Learning



As Presented by Team 8


Zoe


Farres


Zack

Agenda

- Background
- Textbook Reference
- What is Deep Learning
- Differences between Deep Learning/AI and Machine Learning
- SVM
  - Introduction to SVM
  - Problem 9.7.8

Timeline of Neural Networks and SVMs

Textbook Crosswalk

An Introduction to Statistical Learning

10.1 Single Layer Neural Networks
10.2 Multilayer Neural Networks
10.3 Convolutional Neural Networks
10.4 Document Classification
10.5 Recurrent Neural Networks
10.6 When to Use Deep Learning
10.7 Fitting a Neural Network
10.8 Interpolation and Double Descent
10.9 Lab: Deep Learning
10.10 Exercises

Deep Learning with Python

1. What is deep learning
2. The mathematical building blocks of neural networks
3. Introduction to Keras and TensorFlow
4. Getting started with neural networks Classification and regression
5. Fundamentals of machine learning
6. The universal workflow of machine learning
7. Working with Keras: A deep dive
8. Introduction to deep learning for computer vision
9. Advanced deep learning for computer vision
10. Deep learning for time series
11. Deep learning for text
12. Generative deep learning
13. Best practices for the real world
14. Conclusions

Textbook Crosswalk

An Introduction to Statistical Learning

10.1 Single Layer Neural Networks
10.2 Multilayer Neural Networks
10.3 Convolutional Neural Networks
10.4 Document Classification
10.5 Recurrent Neural Networks
10.6 When to Use Deep Learning
10.7 Fitting a Neural Network
10.8 Interpolation and Double Descent
10.9 Lab: Deep Learning
10.10 Exercises

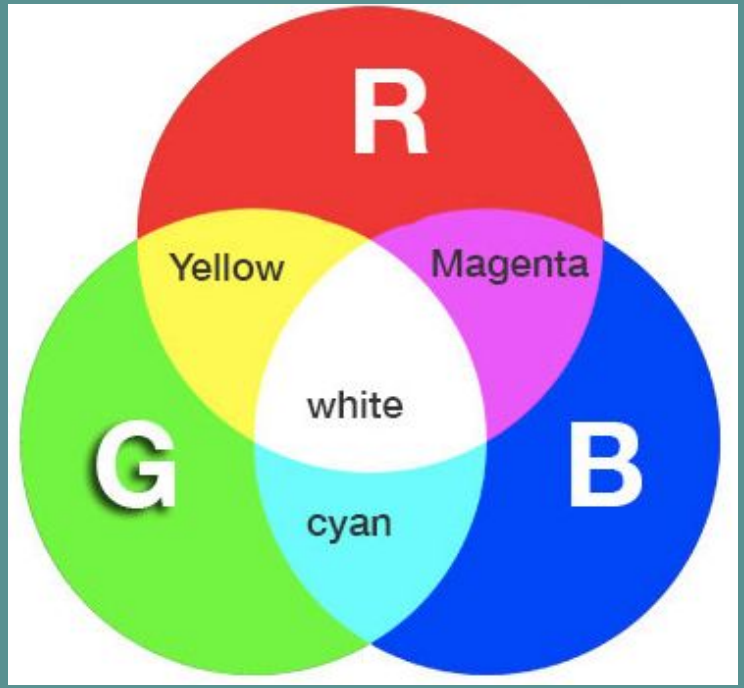
Color in Code

Color	Colorname	Hex	(R,G,B)
	Black	#000000	(0,0,0)
	White	#FFFFFF	(255,255,255)
	Red	#FF0000	(255,0,0)
	Lime	#00FF00	(0,255,0)
	Blue	#0000FF	(0,0,255)

Deep Learning with Python

1. What is deep learning
2. The mathematical building blocks of neural networks
3. Introduction to Keras and TensorFlow
4. Getting started with neural networks Classification and regression
5. Fundamentals of machine learning
6. The universal workflow of machine learning
7. Working with Keras: A deep dive
8. Introduction to deep learning for computer vision
9. Advanced deep learning for computer vision
10. Deep learning for time series
11. Deep learning for text
12. Generative deep learning
13. Best practices for the real world
14. Conclusions

Color In Code



Color	Colorname	Hex	(R,G,B)
	Black	#000000	(0,0,0)
	White	#FFFFFF	(255,255,255)
	Red	#FF0000	(255,0,0)
	Lime	#00FF00	(0,255,0)
	Blue	#0000FF	(0,0,255)
	Yellow	#FFFF00	(255,255,0)
	Cyan	#00FFFF	(0,255,255)
	Magenta	#FF00FF	(255,0,255)
	Silver	#C0C0C0	(192,192,192)
	Gray	#808080	(128,128,128)
	Maroon	#800000	(128,0,0)
	Olive	#808000	(128,128,0)
	Green	#008000	(0,128,0)
	Purple	#800080	(128,0,128)
	Teal	#008080	(0,128,128)
	Navy	#000080	(0,0,128)

Selected Chapter 10 Topics

Topic 1

10.1

Single Neural Networks

Topic 2

10.2

Multilayer Neural Networks

Topic 3

10.3

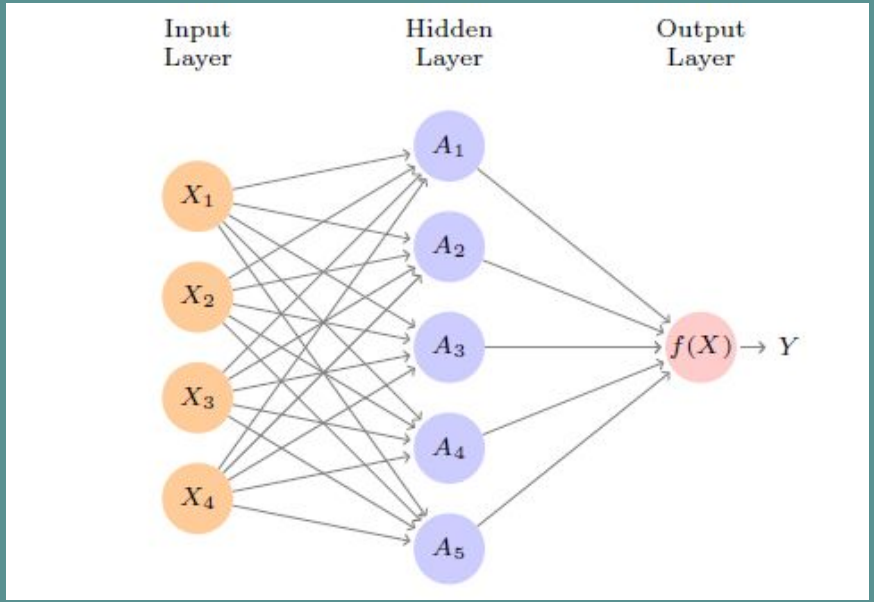
Convolutional Neural Networks

Single Layer Neural Networks (10.1)

- We have built nonlinear prediction models in earlier chapters, using trees, boosting and generalized additive models.
- What distinguishes neural networks from these previous models is the structure of the model

Recognized terms from Artificial Intelligence:

- ReLU (rectified linear unit) activation function
- Sigmoid activation function
- Keras
- Tensorflow



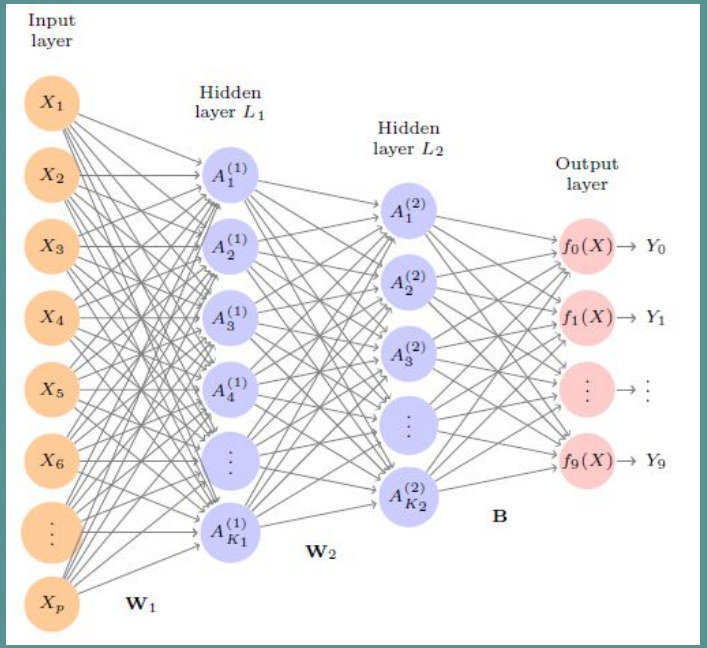
Simple feed-forward neural network for modeling a quantitative response using p = 4 predictors

Multilayer Neural Networks (10.2)

- Modern neural networks typically have more than one hidden layer
- In theory a single layer with a large number of units has ability to approximate most functions
- The task of discovering a good solution is made much easier with multiple layers each of modest size.

Recognized terms from Artificial Intelligence:

- MNIST handwritten digit dataset
- weights
- Softmax activation function
- dropout

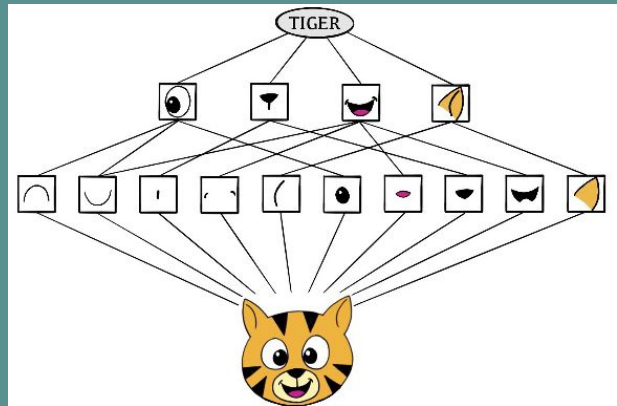


Neural network diagram with two hidden layers and multiple outputs. The input layer has p = 784 units, the two hidden layers K1 = 256 and K2 = 128 units respectively, and the output layer 10 units. Along with intercepts (referred to as biases in the deep-learning community) this network has 235,146 parameters (referred to as weights)

Convolutional Neural Networks (the Layman's Version)

- Identification of low-level features such as small edges, patches of color
- Low-level features combined to form higher-level features such as parts of ears or eyes
- Absence or presence of higher level features contribute to probability of a given class

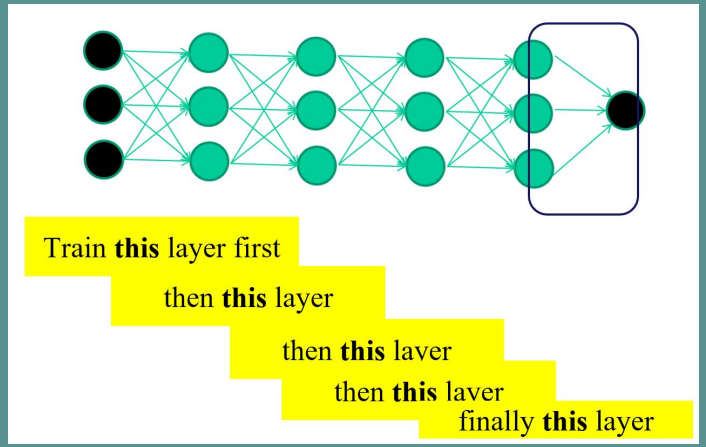
- How does it work? (still Layman's Version)
  - Leveraging two types of hidden layers: Convolution and Pooling
  - Convolution layers search for instances of small patterns
  - Pooling layers downsample these to select a prominent subset
  - State of the Art results come from use of many convolutional and pooling layers



Schematic showing how a convolutional neural network classifies an image of a tiger. The network takes in the image and identifies local features. It then combines the local features in order to create compound features, which in this example include eyes and ears. These compound features are used to output the label "tiger".

What is Deep Learning?

- A subsection of Machine Learning, also referred to Deep Neural Networks
- Neural Network model that uses several layers of nodes between input and output
- Within each set of layers, they identify features and process in a series of stages
- Final layer is trained to predict class based on output from each of the previous layers
- Ultimately, Deep Learning *is* Machine Learning
    - But with some nuanced changes



Deep Learning

- For this program, we use Python
- Requires Neural Networks
  - Feature Engineering
- Models generally takes longer to run
  - Generally larger datasets
- Stronger computing power needed
- Black box/harder to interpret
  - Harder to debug

Machine Learning

- For this program, we use R
- Supervised vs. Unsupervised Learning
  - Many different classifiers to use
  - Different combinations to test
- May have an idea as to which features produce the best results
- Easier to interpret

DL Functions

General Keras Commands:

- .sequential()
- .add()
- Dense()
- .compile()
- .fit()

Primary Library/Package

- Keras/TensorFlow
  - Keras.models
  - keras.layers
- scikit learn
- numpy/pandas

ML functions

General Model Commands:

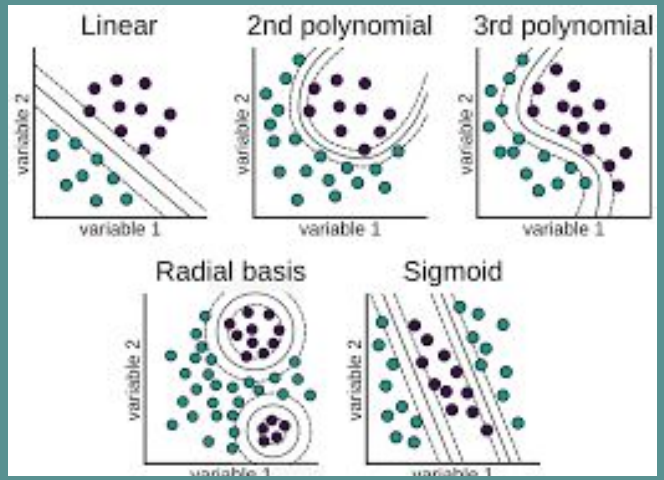
- lm()
- lda()
- qda()
- glm()
- svm()
- tree()

Common Libraries/Packages

- Caret
- E1071
- Dpylr
- Tree
- MLtools

Applied Question: SVM model

1. Supervised learning algorithms (labeled)
2. Parametric(for linear decision boundary), Non-parametric(non-linear decision boundaries)
3. classify observations into binary groups by drawing a hyperplane in the N-dim space



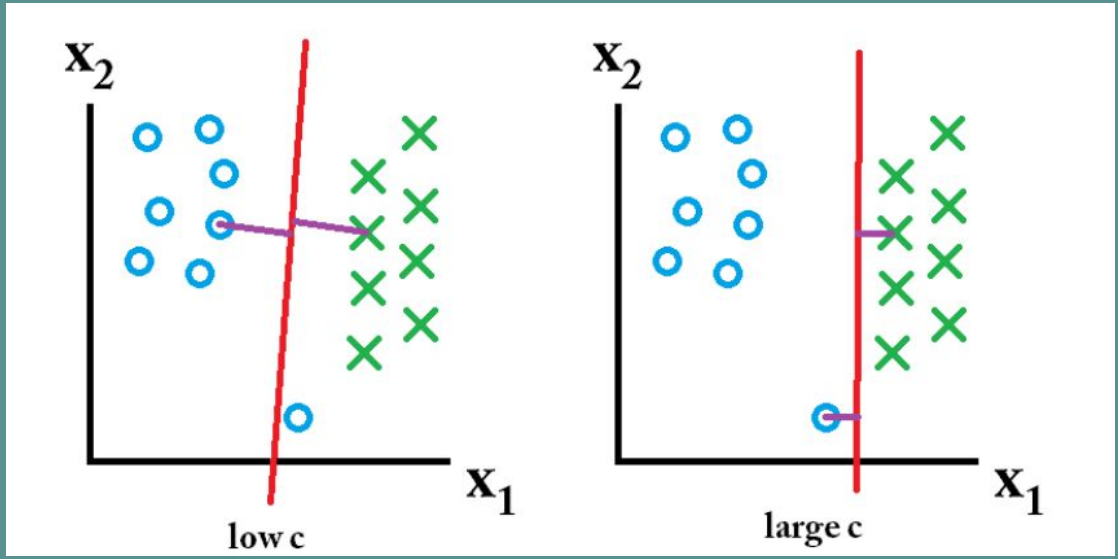
Decision Boundary (set by the kernel parameter)

TradeOff

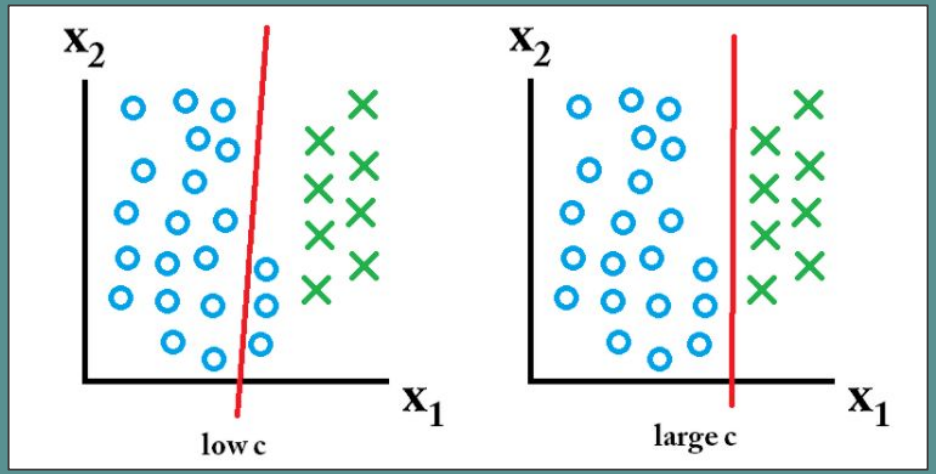
Bias: Maximizing the Minimum Margin (def)

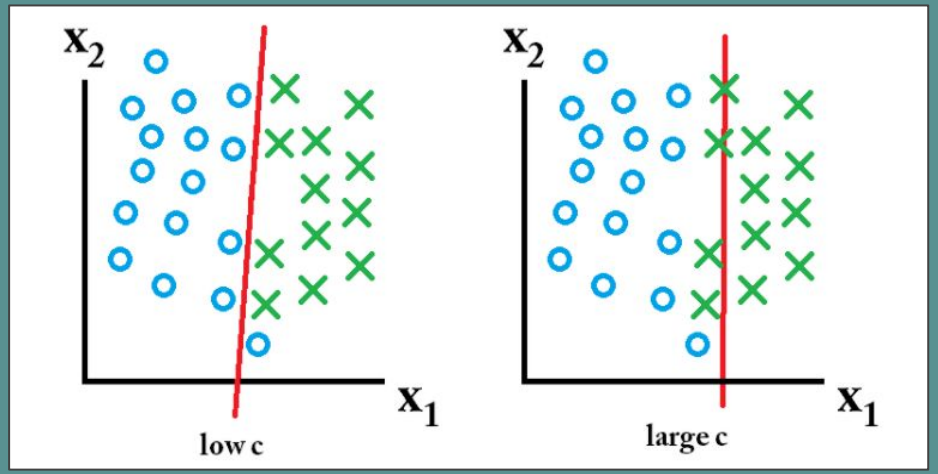
VS

Variance: Correctly Separating Observations into Classes



Test dataset





C (penalty) is inversely correlated with cost (budget)

A summary of the R script

1. Install necessary packages
2. Split the data into training and testing
3. Build the SVM model with linear/radial/polynomial approaches
    a. Find the best cost by cross validation
    b. Calculate test and training error
4. Compare the three models and select the best model

Poll Question 1

Question: Which of the following is a component of a CNN?

Multiple Choice Answers:

A. Feature Detector

B. Classifier

C. Cost Function

D. Activation Function

E. All of the Above

Poll Question 2

Question:

SVM supports _____ classification.

Multiple Choice Answers:

A. Multiclass

B. Binary

C. Both

D. None

Poll Question 3

Question:

The cost parameter in the SVM means

Multiple Choice Answers:

A. The number of cross-validations to be made
B. The kernel to be used
C. The tradeoff between misclassification and the margin of the model
D. None of the above

Poll Question 4

Question:

A popular type of artificial neural network. They are used to process sequences of data, text, genomes, handwriting, and more. These artificial neural networks make use of backpropagation for the training requirements.

Multiple Choice Answers:

A. concurrent neural network

B. convolutional neural networks

C. recurrent neural network

D. evolutional neural network

Poll Question 5

Question:

What is the name for variables used to determine the structure of a neural network. They are also used to understand parameters, such as the learning rate and the number of hidden layers, and more, present in the neural network.

Multiple Choice Answers:

A.ultraparameters

B. superparameters

C. primeparameters

D. hyperparameters

Poll Question 6

Question:

An approach for classification that was developed in the computer science community in the 1990s and that has grown in popularity since then. This approach has been shown to perform well in a variety of settings, and are often considered one of the best “out of the box” classifiers.

Multiple Choice Answers:

A.Tableau

B. Convolutional Neural Nets

C. Tree Based Methods

D. Support Vector Machine