Quiz-01

- Due Sep 1 at 11:59pm
- Points 10
- Questions 10
- Available Aug 30 at 12:01am - Sep 1 at 11:59pm
- Time Limit None
- Allowed Attempts 3

Instructions

Intro and Universal Approximators

This quiz covers lectures 1 and 2. Several of the questions invoke concepts from the hidden slides in the slide deck, which were not covered in class. So please go over the slides before answering the questions.

You will have three attempts for the quiz. Questions will be shuffled and you will not be informed of the correct answers until after the deadline. While you may discuss the concepts underlying the questions with others, you must solve all questions on your own - see course policy.

Take the Quiz Again

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	78 minutes	7.5 out of 10

Correct answers are hidden.

Score for this attempt: 7.5 out of 10
Submitted Aug 31 at 3:26pm
This attempt took 78 minutes.

Question 1
1 / 1 pts
According to the university's academic integrity policies and 11785 course policies, which of the following practices are **NOT** allowed in this course? Select all that apply.

- [ ] Helping another student debug their code
- [x] Discuss quiz solutions with another student before the deadline
- [ ] Asking a TA for help debugging your code
- [x] Posting code in a public post on piazza

- [ ] Discussing concepts from class with another student

Question 2

1 / 1 pts

Match the corresponding terms and definitions introduced in Lecture 1.

The McCulloch and Pitts model

is a Logical Calculus of the

Alexander Bain

is known for his Connectio

Lawrence Kubie

modeled the memory as a

Hebbian Learning

is a proposed mechanism t

Marvin Minsky and Seymour Papert

Their mechanism known as

One of David Hartley's Observations

Our brain represents comp

Frank Rosenblatt

made the first algorithmica

Associationism Theory by Aristotle

These are his four laws: Th

PartialQuestion 3

0.5 / 1 pts

We sometimes say that neural networks are connectionist machines as opposed to von Neumann machines. Which of the following describe why we make this distinction? (select all that apply)

Slide: lec 1, “Connectionist Machines”.

- [x] A von Neumann machine has a general purpose architecture with a processing unit that is distinct from the memory that holds the programs and data. A connectionist machine makes no distinction between processing unit and the program.
- [x] It is possible to create hardware implementations of von Neumann machines (e.g. CPU's) as well as software implementations (e.g. virtual machines). However, connectionist machines can only be implemented in software (e.g. neural networks in Python).
- [ ] Because of its flexibility, a von Neumann machine is capable of computing any Boolean function of a given number of Boolean inputs, whereas connectionist machines, no matter how complex, are fundamentally unable to model certain types of Boolean functions.
- [x] A von Neumann machine can be used for general-purpose computing by simply providing a different program, without changing the machine itself. A connectionist machine implements a specific program, and changing the program requires changing the machine.

IncorrectQuestion 4

0 / 1 pts

If the first layer of a network operating on real-valued inputs does not have sufficient neurons to capture all the basic features of the function (map), which of the following activation functions gives subsequent layers the best chance to capture the features missed by the first layer.

Hint: Which activation function has the most gradation

Slide: Lec 2, "Sufficiency of architecture" slides 145 - 148

- ( ) Threshold activation
- ( ) Relu activation
- ( ) Leaky Relu activation
- (•) Sigmoid activation

Lec 2 recording, “Sufficiency of architecture”, on slide comparing activation functions.

Question 5

1 / 1 pts

Which statements did David Hartley assert or imply in Observations on Man? (select all that apply)

Slide: lec 1, "Dawn of Connectionism"

- [x] The brain can receive and process sensory input.
- [ ] The brain is composed of neurons connected in a network.
- [x] Memories can be stored and can be linked to sensory input.
- [x] The brain records information as vibrations.

Question 6

1 / 1 pts

Although we haven't yet covered training of neural networks in the lectures, we can give you this advance bit of information: The number of required training inputs to train a network properly is monotonically related to the number of parameters.

In general, as the depth of a NN increases, at what rate does the number of training observations required to adequately train the network change? (Choose the most appropriate answer)

Slide: lec 2, “The challenge of depth”

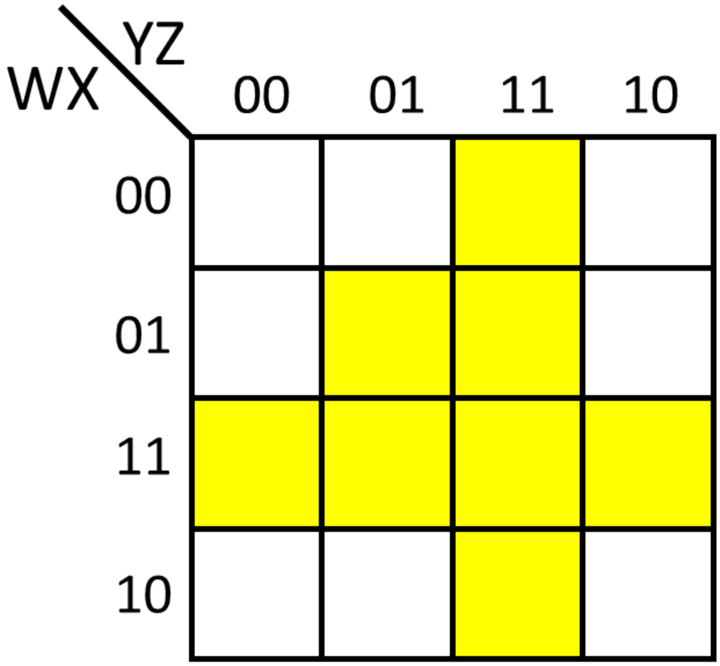
- (•) Decreases exponentially
- ( ) Increases exponentially
- ( ) Increases quadratically
- ( ) Decreases quadratically

In general, for a given function, deeper networks will require exponentially fewer parameters than shallower ones to model the function accurately (exactly, or with arbitrary precision). The number of required training inputs is monotonically related to the number of parameters.

IncorrectQuestion 7

0 / 1 pts

What is the fewest neurons needed (including any output layer neurons) for a network to implement the truth table shown by the following Karnaugh map? (numeric answer, int and float are both fine)



Slide: lec 2, “Reducing a Boolean Function”.

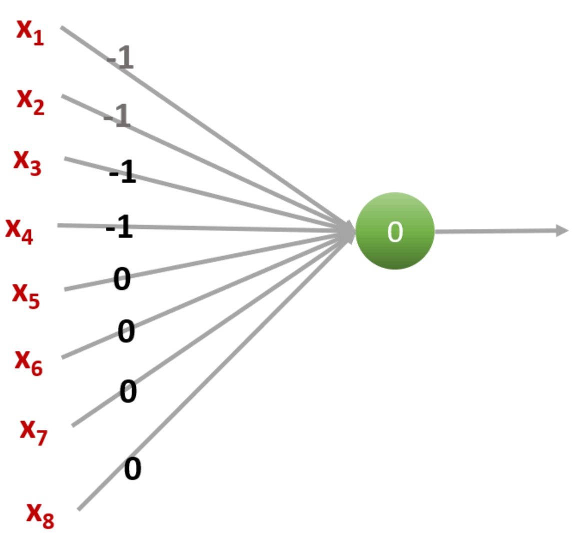
3

O = WX + YZ + XZ, one neuron for each summand and one output neuron (total 4).

Question 8

1 / 1 pts

Under which conditions will the perceptron graph below fire? Note that ~ is NOT. (select all that apply)



Slide: lec 2, "Perceptron as a Boolean gate"

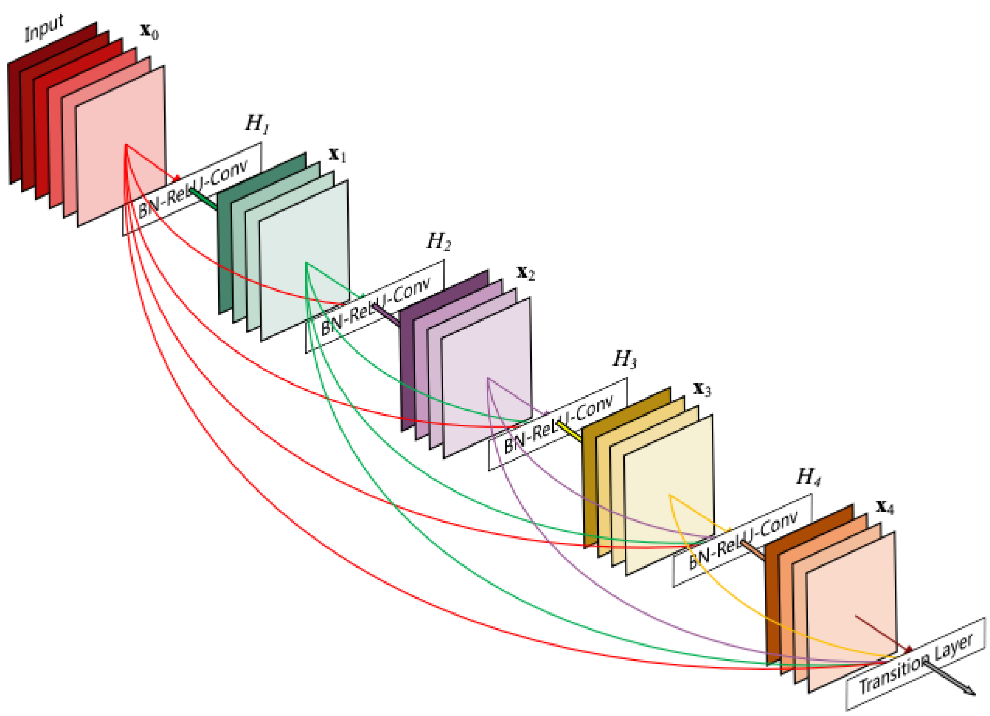
- [ ] x1 & x2 & x3 &x4
- [x] ~x1 & ~x2 & ~x3 & ~x4

- [ ] Never fires
- [x] fires only if x1, x2, x3, x4 are all 0, regardless of x5 .. x8

Question 9

1 / 1 pts

Densenet (2017) is a CNN network architecture that achieved SoTA results with fewer params than many of its contemporaries. Its main idea was the ‘dense block’: a block where each layer output is concatenated into the input of the downstream layers.



Above is a diagram of a dense-block. Each “H_i” consists of Batchnorm->ReLU->Convolution. Each H_i outputs a “x_i” (one stack of squares in the image) .

What is the depth of the dense block above? (numeric answer, int and float are both fine)

Assume that "BN+RELU+CONV" is a single layer. The square planes simply represent the data flowing down between layers and are not layers themselves.

(Note: for the definition of network depth, see the lecture 2 recording)

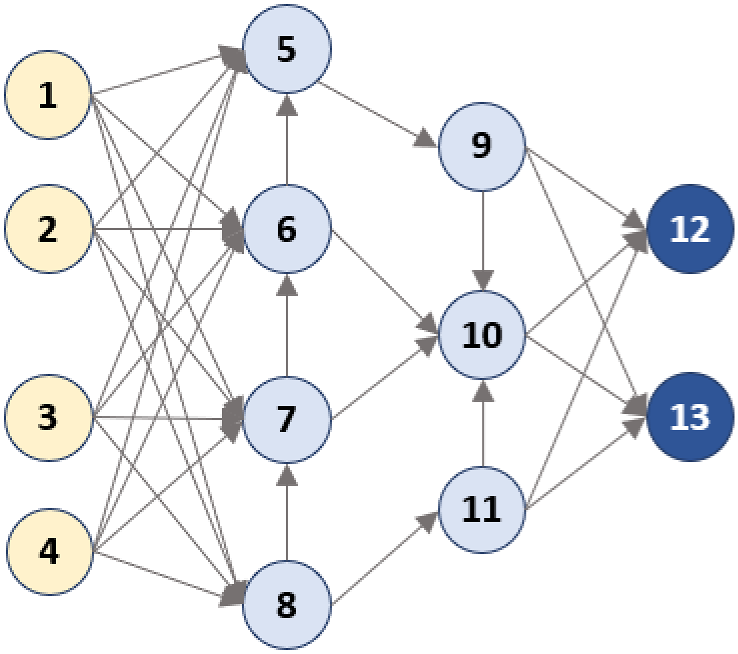
Slide: lec 2, "Deep Structures"

5

The longest path is Input->H_1->H_2->H_3->H_4->Transition: 5 edges.

Question 10

1 / 1 pts



If the yellow nodes are inputs (not neurons) and the dark blue nodes are outputs, which neurons are in layer 6?

Slide: lec 2, "What is a layer"

- 10, 12
- 9, 10, 11
- 10, 12, 13
- 10

Admittedly this one is tricky. Pay attention to the definitions. Definition of layer #: the neurons reachable (with the same number of edges) along the longest path from source to sink. If a node is reachable multiple times along the longest path, its layer # is the max of those reachable cases (i.e. if some node D was reachable along the longest path 3 times [1,2,3] along the longest path, its layer number is 3).

The longest path would be from any input -> 8 -> 7 -> 6 -> 5 -> 9 -> 10 -> any output. Node 8 would count as layer 1, as you traverse 1 edge along the longest path to get there. While it is possible to visit 5,6,7,8 via one edge, they are visited later along the longest path. Layer 2 would be 7 AND 11, as 11 is not visited later along the longest path, but is still reachable with the same number of edges.

But layer 6 is node 10 only. While 10 is reachable twice along the longest path (once via 3 edges once via 6 edges), remember that we choose the max. Also, 12 and 13 are visited AFTER 10 along the longest path, meaning they are part of layer 7.

So the sixth layer only contains node 10.

Quiz Score: 7.5 out of 10