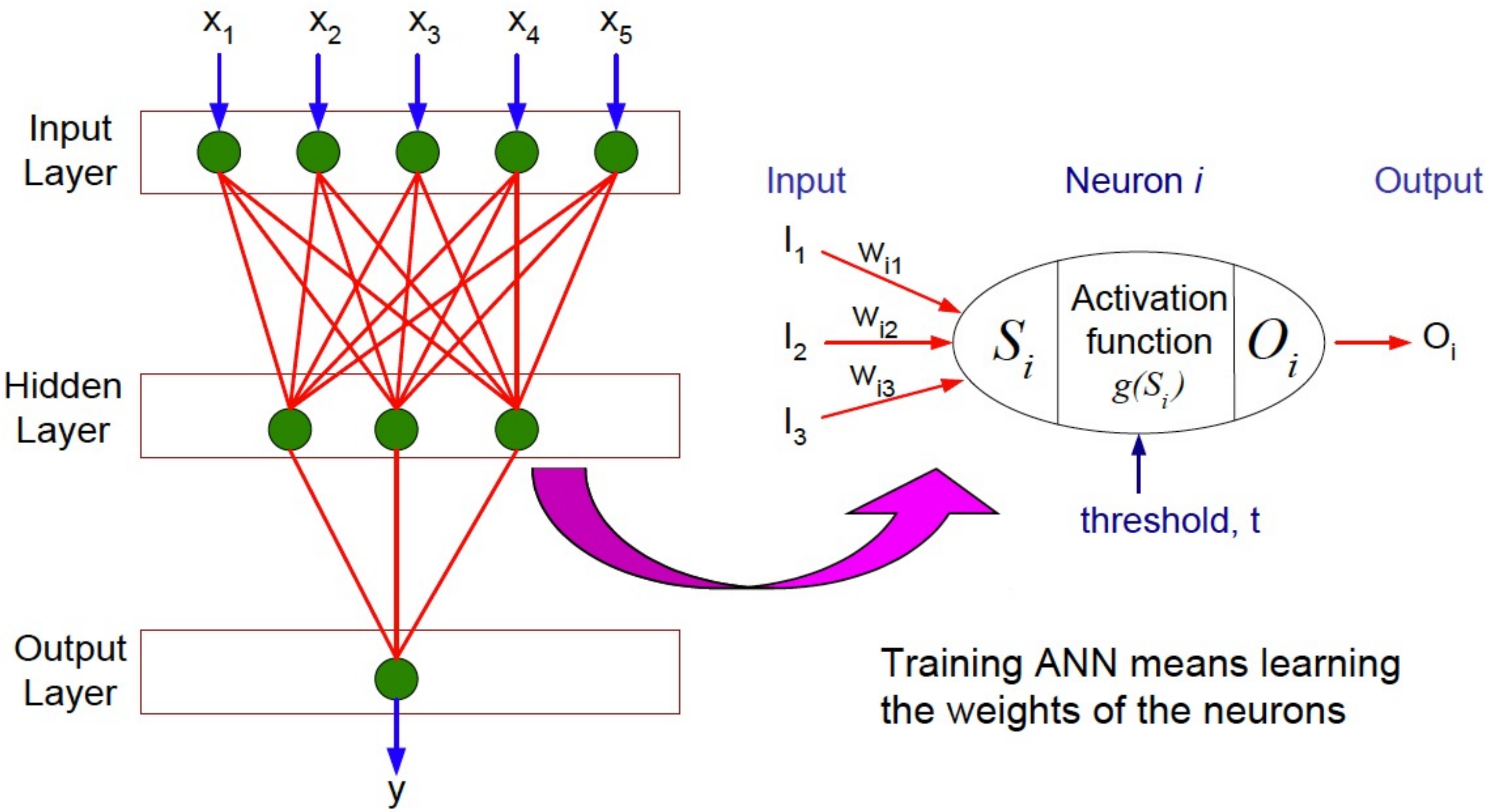
Neural Networks & Deep Learning

Exercise 4

UNIVERSITY OF MANNHEIM
School of Business Informatics and Mathematics

Neural Networks
General Structure



Training ANN means learning the weights of the neurons

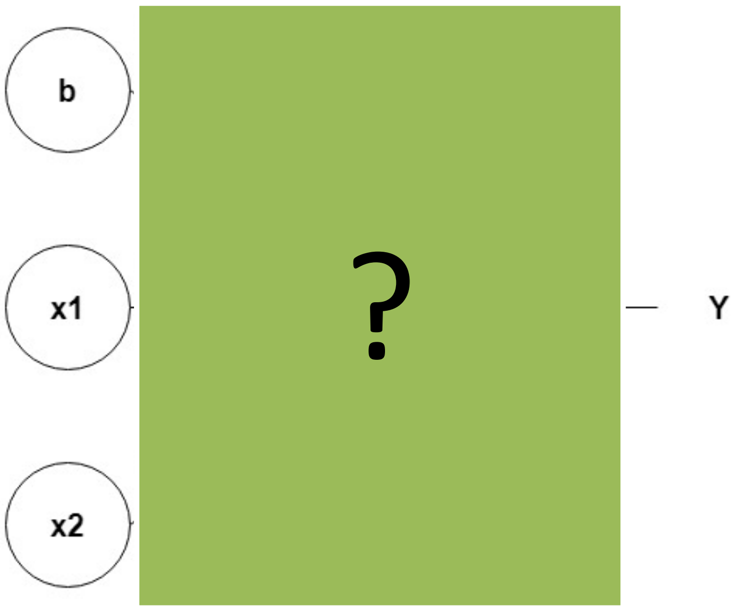
Neural Networks
Representing Logics

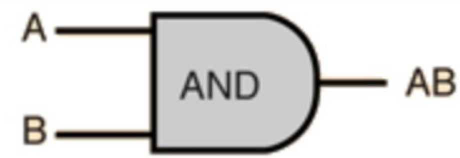
TODO

Try to represent the following logic gates with Neural Networks

- AND
- OR
- NOT
- NAND
- XOR

Neural Networks
Representing Logics: AND

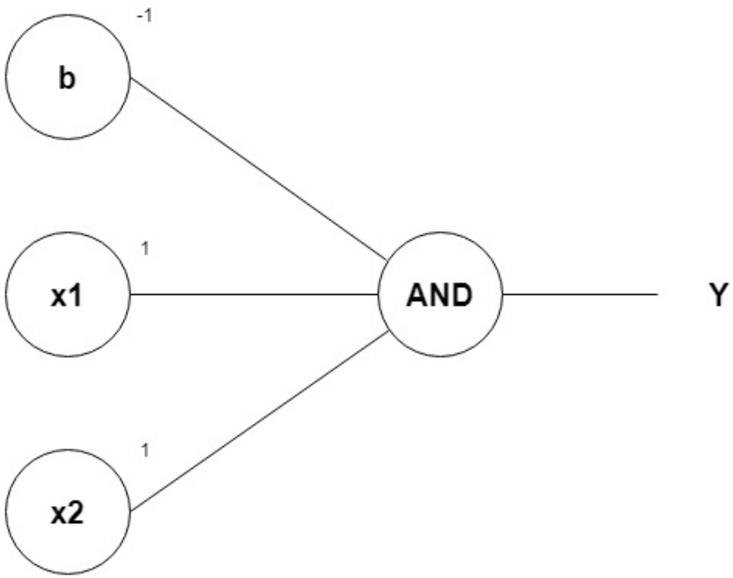


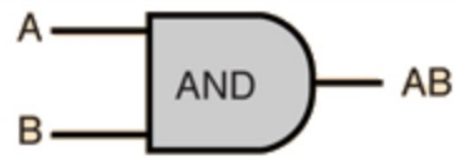


A	B	Out
0	0	0
0	1	0
1	0	0
1	1	1

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks
Representing Logics: AND



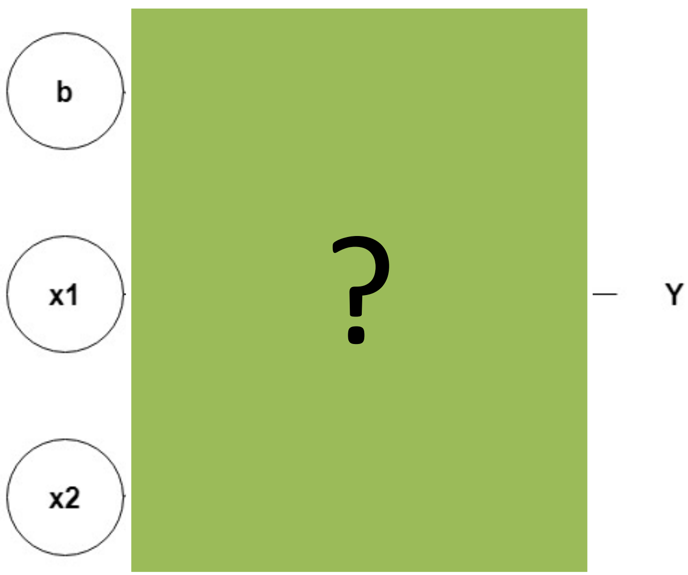


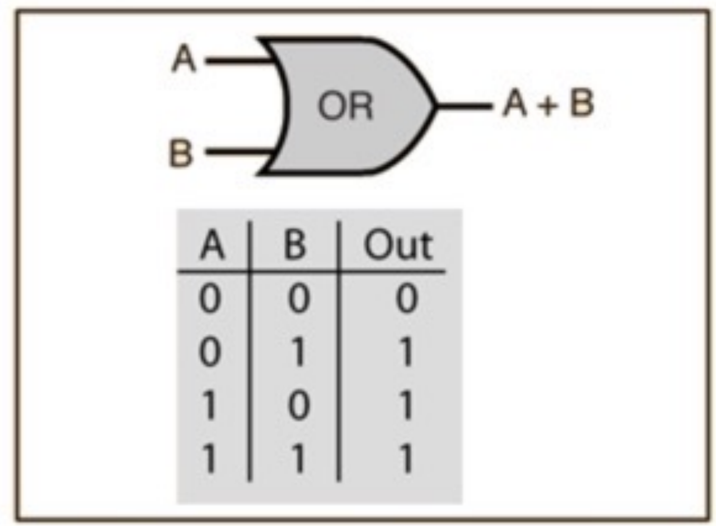
A	B	Out
0	0	0
0	1	0
1	0	0
1	1	1

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks

Representing Logics: OR

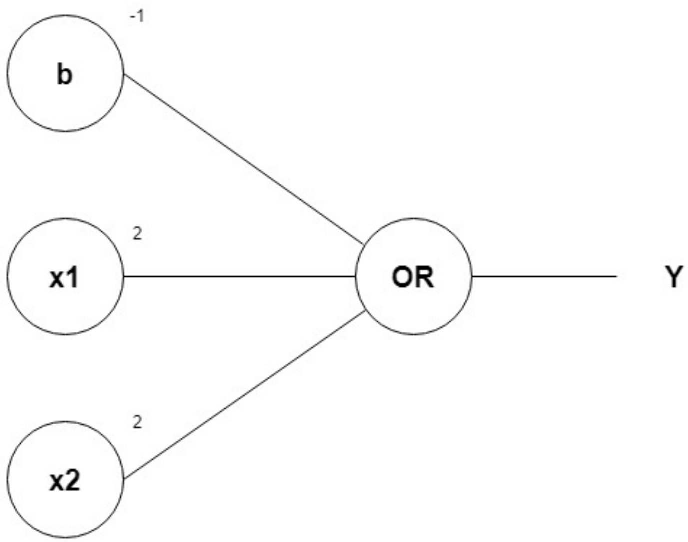


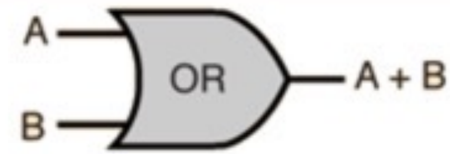


A	B	Out
0	0	0
0	1	1
1	0	1
1	1	1

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks
Representing Logics: OR

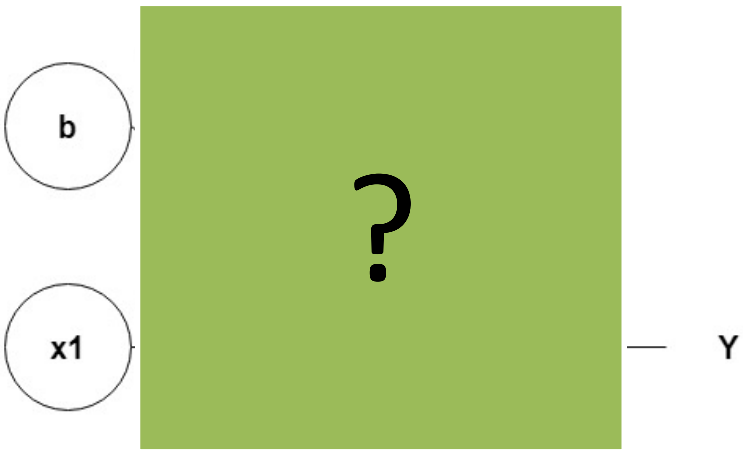


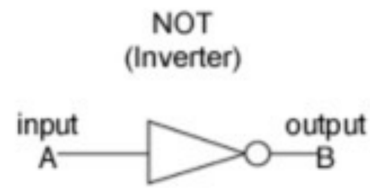


A	B	Out
0	0	0
0	1	1
1	0	1
1	1	1

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks
Representing Logics: NOT

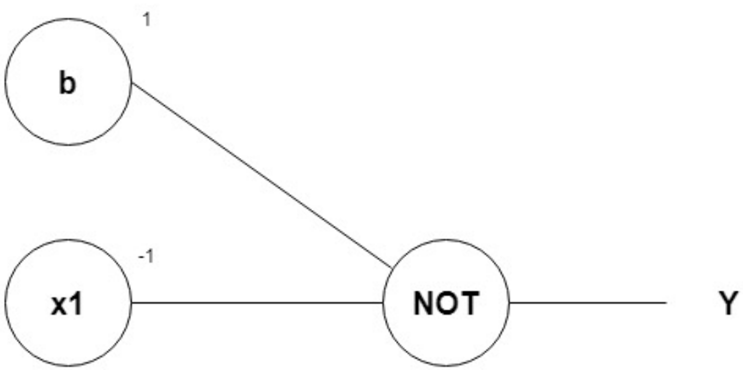


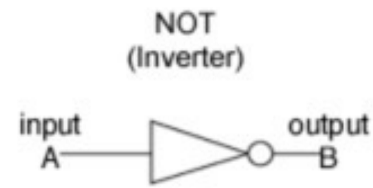


A	B
0	1
1	0

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks
Representing Logics: NOT

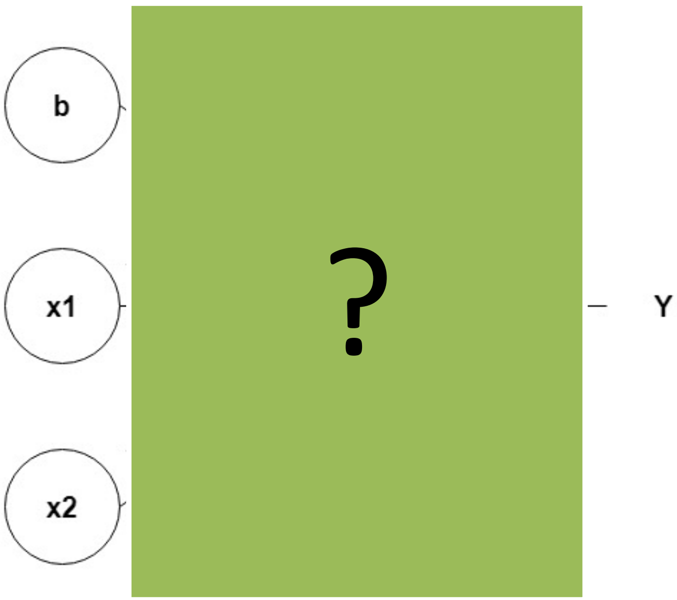


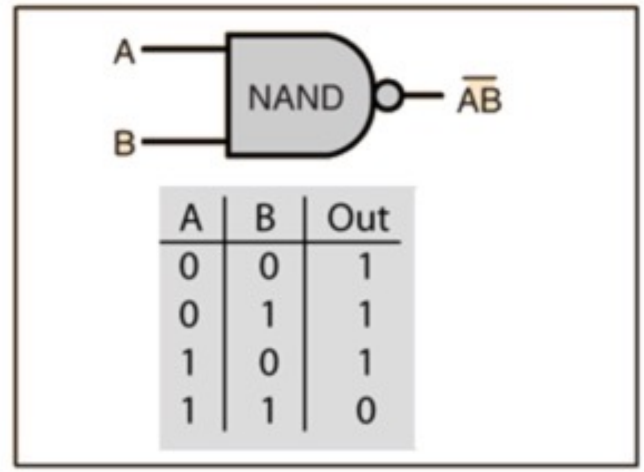


A	B
0	1
1	0

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks
Representing Logics: NAND

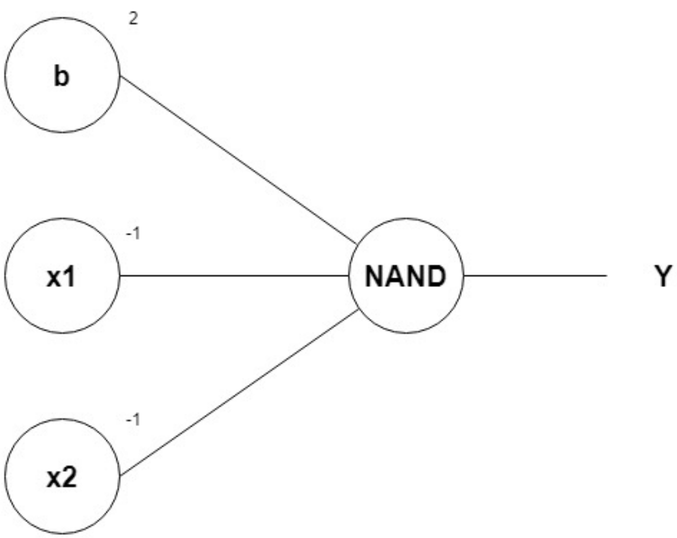




A	B	Out
0	0	1
0	1	1
1	0	1
1	1	0

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks
Representing Logics: NAND

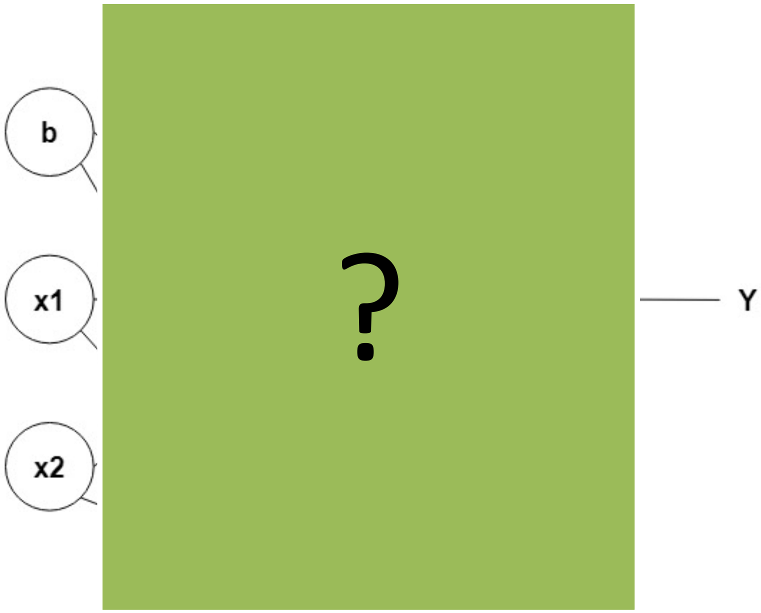


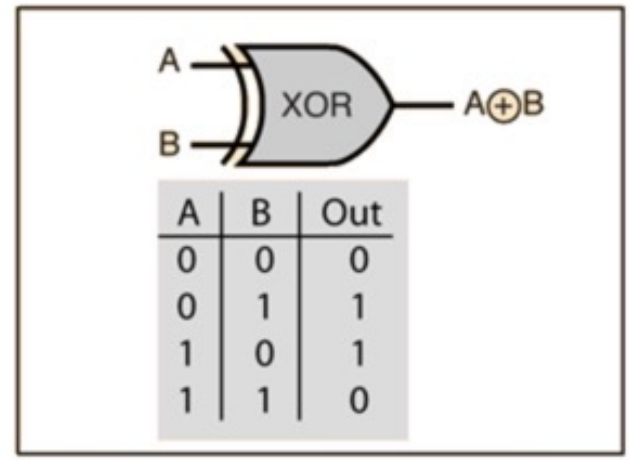
A	B	Out
0	0	1
0	1	1
1	0	1
1	1	0

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks

Representing Logics: XOR



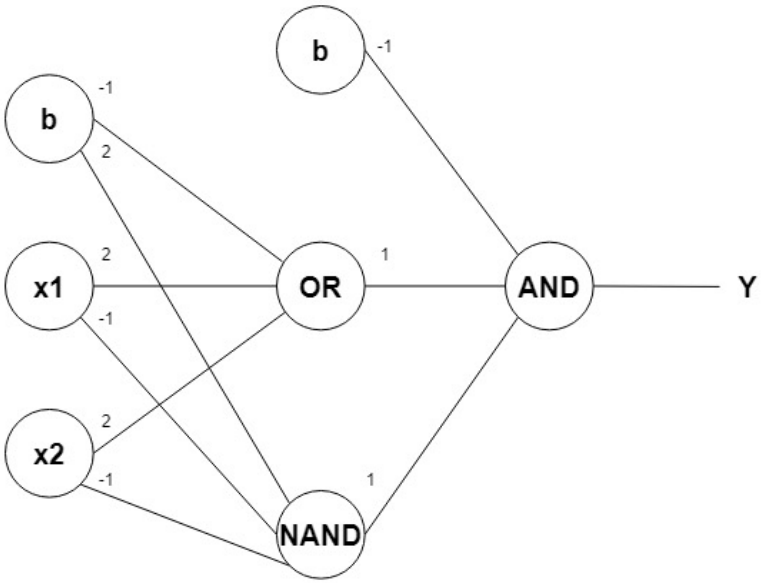


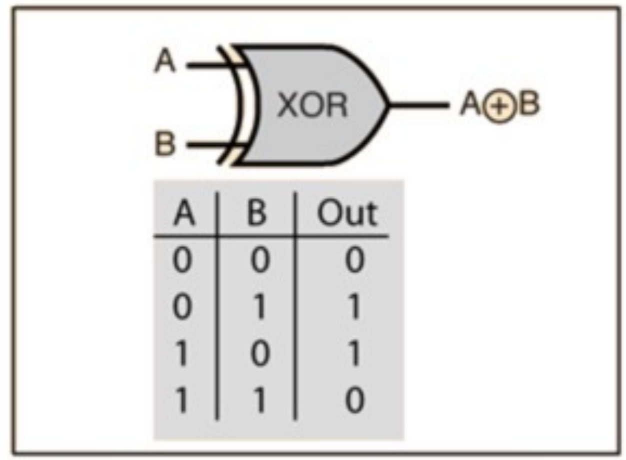
A	B	Out
0	0	0
0	1	1
1	0	1
1	1	0

Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks

Representing Logics: XOR





A	B	Out
0	0	0
0	1	1
1	0	1
1	1	0

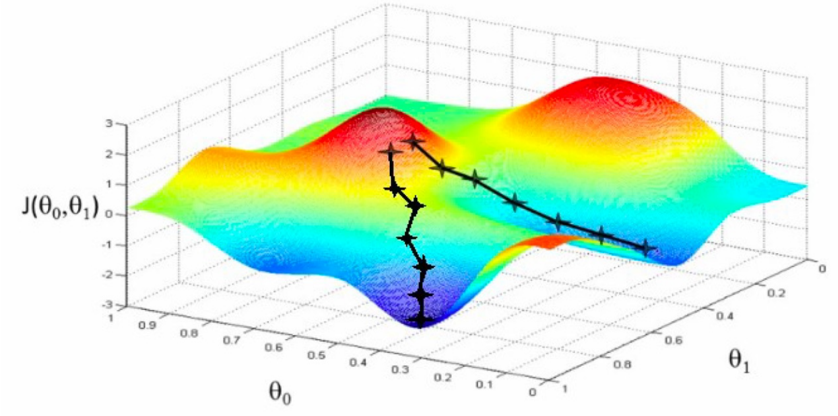
Images taken from https://medium.com/@stanleydukor/neural-representation-of-and-or-not-xor-and-xnor-logic-gates-perceptron-algorithm-b0275375fea1

Neural Networks
TRUE or FALSE?

„When the optimization of a Neural Network converges, we always end up with the optimal solution.“

Neural Networks
TRUE or FALSE?

„When the optimization of a Neural Network converges, we always end up with the optimal solution.“



FALSE!

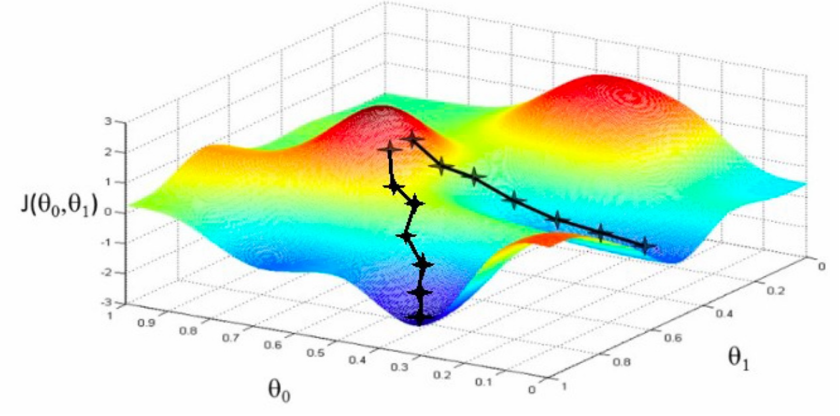
Neural Networks
TRUE or FALSE?

„A higher learning rate typically leads to faster convergence, but increases the risk of getting stuck in local optima.“

Neural Networks

TRUE or FALSE?

„A higher learning rate typically leads to faster convergence, but increases the risk of getting stuck in local optima.“



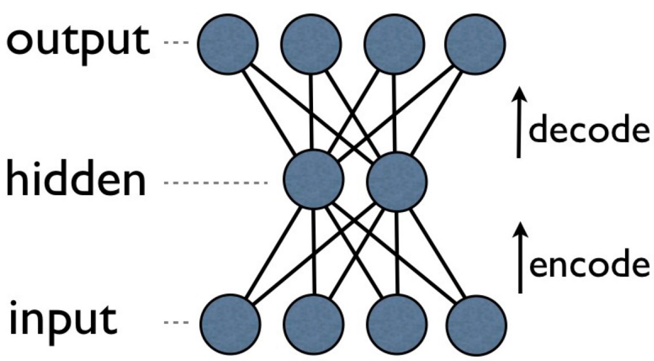
FALSE! It can lead to faster convergence (depending on the data) but it does not increase the risk of getting stuck in local optima.

Neural Networks
TRUE or FALSE?

„Auto-Encoders can be used to remove noise from a given input.“

Neural Networks
TRUE or FALSE?

„Auto-Encoders can be used to remove noise from a given input.“



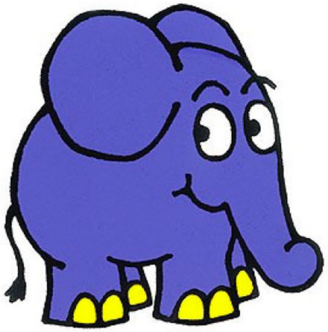
TRUE!

Neural Networks
TRUE or FALSE?

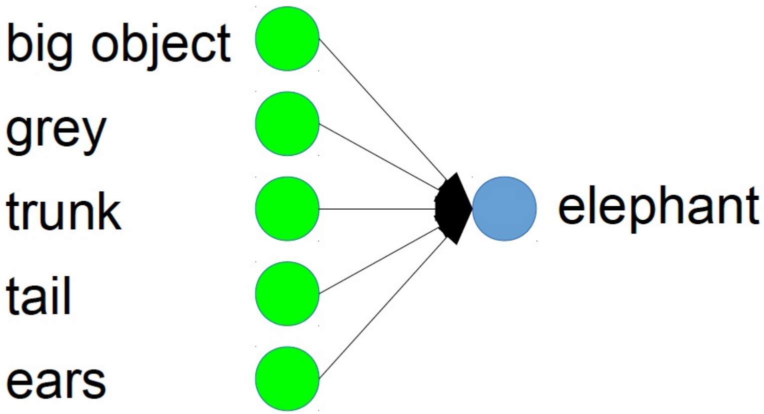
„*Dropout* is a mechanism used during the training of Neural Networks to drastically reduce training time."

Neural Networks

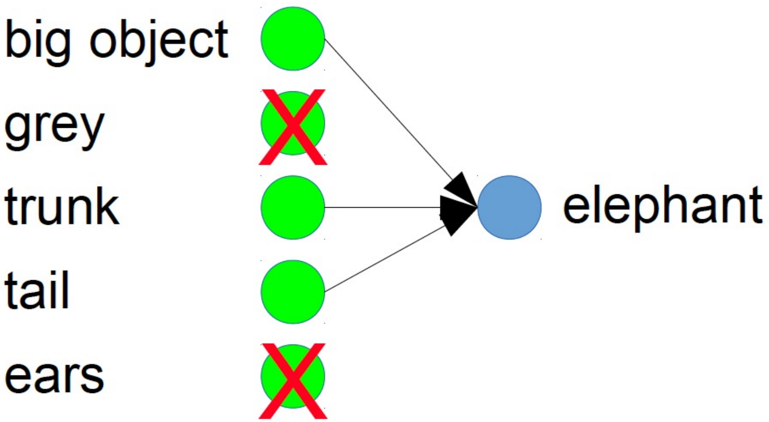
TRUE or FALSE?



„*Dropout* is a mechanism used during the training of Neural Networks to drastically reduce training time.“



FALSE!

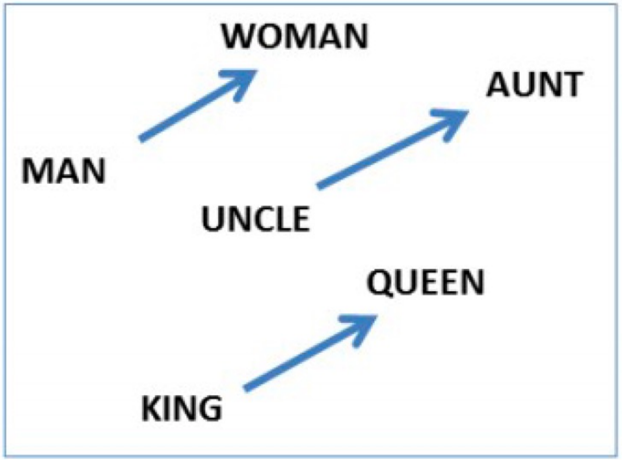


Neural Networks
TRUE or FALSE?

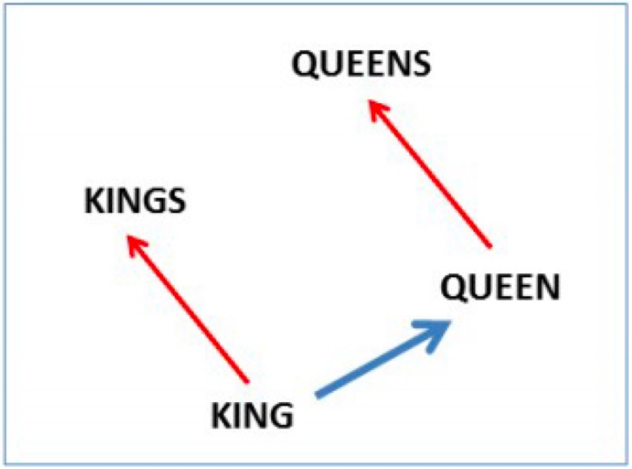
„Word2Vec converts every word to a vector in a latent vector space to make computations possible.“

Neural Networks
TRUE or FALSE?

„Word2Vec converts every word to a vector in a latent vector space to make computations possible.“



TRUE!



Word2Vec
Task

You have a corpus of 1,000 product reviews that you shall use to train a sentiment predictor. As the number of unique words in these product reviews is 150 and hence rather small, you are unsure whether to use One-Hot-Encoding or Word2Vec to encode the words in your product reviews.

TODO

1) Name some advantages for both of the approaches.
2) Which one would you prefer?
3) Would it change your opinion if you knew that your sentiment predictor will never encounter words that are not in the training corpus?

Word2Vec
Solution

One-Hot Encoding	Word2Vec
(+) Simple encoding scheme where it is clear what is meant by a given vector	(-) Word is represented by a vector of latent features without clear meaning
(+) No training data necessary to encode words; any word or token can be encoded	(-) Needs a lot of training data or a pre-trained model has to be used. How to deal with unknown words?
(-) Vector size is equivalent to the size of the vocabulary	(+) Vector size is fixed (typically set to something between 100 and 1000)
(-) No semantic relation between words	(+) Similar words are in similar “regions”

Especially because of the last point, Word2Vec is expected to be the better solution for this problem. But if the vocabulary size is fixed, we might not need the semantic similarity between vectors as this similarity could be encoded in the model (if the training data is expressive enough).