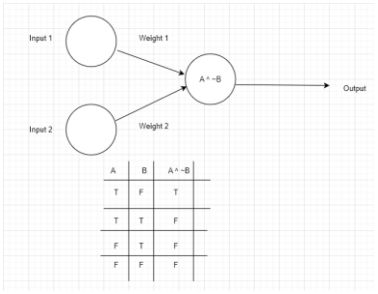
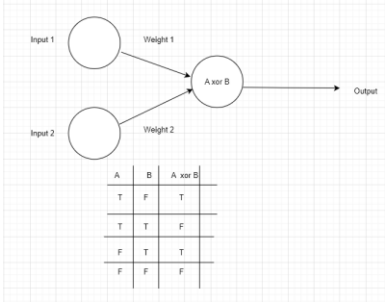
CS7641 Machine Learning - Problem Set 1

Mlukacsko3@gatech.edu / Michael Lukacsko

Question 2. Design a two-input perceptron that implements the boolean function $A \wedge \neg B$. Design a two-layer network of perceptrons that implements $A \oplus B$ (where $\oplus$ is XOR).

- $A \wedge \neg$B answer
- 

A	B	A ^ ¬B
T	F	T
T	T	F
F	T	F
F	F	F

  - A two-input perceptron (image above) the implements the function $A \wedge \neg B$ can be done as follows:
    - Set the weight of input 1 (A) to 1 and the weight of input 2 (B) to -1.
    - Use a threshold/step-based activation function that returns 1 if the weighted sum of inputs is greater than or equal to 0. Otherwise, return 0.
  - The output of this perceptron will be 1 if A is 1 and B is 0, otherwise the output is a 0. When B is 0, the weighted sum of inputs is 1* A + (-1) * B = A - B, which is positive when A is 1 and negative when A is 0, so the threshold function returns 1 only when A is 1 and B is 0. When B is 1, the weighted sum of inputs is 1 * A + (-1) * B = A - B, which is always negative, so the threshold function returns 0 for all inputs.
- $A \oplus B$ (where $\oplus$ is XOR) answer
- 

A	B	A xor B
T	F	T
T	T	F
F	T	T
F	F	F

  - A two-input perceptron (image above) the implements the function A xor B can be done as follows:
    - Set the weight of input 1 (A) to 1 and the weight of input 2 (B) to 1.
    - Use a threshold/step-based activation function that returns 1 if the weighted sum of inputs is greater than or equal to 1. Otherwise, return 0.
  - The output of this perceptron will be 1 if A and B have different values (A xor B = 1), otherwise it will be 0. When A and B have different values, the weighted sum of inputs is 1 * A + 1 * B = A + B, which is greater than or equal to 1. When A and B have the same

value, the weighted sum of inputs is 2 * A (or 2 * B), which is less than 1, so the threshold function returns 0.

Question 5: Suggest a lazy version of the eager decision tree learning algorithm ID3. What are the advantages and disadvantages of your lazy algorithm compared to the original eager algorithm?

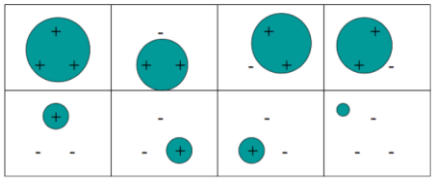
- Answer:
  One alternative/lazy version of the eager decision tree learning algorithm ID3 algorithm is called "Iterative Dichotomiser 3 with a Lazy Decision Tree" (IDL). The main difference between IDL and the original ID3 algorithm is that IDL defers computation of the decision tree until a query is received, whereas ID3 builds the tree immediately.
- Advantages of IDL over ID3 are:
  - Faster training
    - IDL does not build the tree immediately, so it takes less time to train than ID3.
  - Reduced memory consumption.
    - IDL does not store the entire decision tree in memory, which can be advantageous for large datasets with many attributes.
  - Improved accuracy
    - IDL can handle noisy or incomplete data better than ID3 because it does not commit to a fixed tree structure. Instead, IDL builds a decision tree on-the-fly as needed, which can improve the accuracy of the model.
- Disadvantages of IDL over ID3 are:
  - Slower query times
    - IDL takes longer to make predictions than ID3 because it must construct a decision tree at query time.
  - Reduced interpretability.
    - IDL does not produce a fixed decision tree, which can make it difficult to interpret the model and understand how the decision was made.
  - Increased computational complexity.
    - IDL may require more computation than ID3 for a given query because it must construct a decision tree at query time.
- As a result, IDL can be useful in situations where training speed and memory consumption are a concern or when dealing with noisy or incomplete data. However, when interpretability and query speed are more important, it may not be the best choice.

Question 7: Give the VC dimension of these hypothesis spaces, briefly explaining your answers:

1. An origin-centered circle (2D),

2. An origin-centered sphere (3D)

- Answer: Origin-centered circle 2D



  - The image above shows a circle that can shatter any set of three points that are not collinear. When a 4th point is added, it is not possible to shatter all 4 points, hence, the VC dimension is exactly 3 because there exists a set of three points that can be shattered by a circle but no set of 4.
  - Where the lines are collinear, the VC dimension is only 2. This is because no matter how the circle is positioned, it will always enclose all three points or none of them.
- Answer: Origin-centered sphere (3D)
  - The VC dimension of an origin-centered sphere in 3D is 4. This is because a 3D origin-centered sphere cannot shatter any set of five points that form a convex polyhedron. Therefore, the VC dimension is 4.