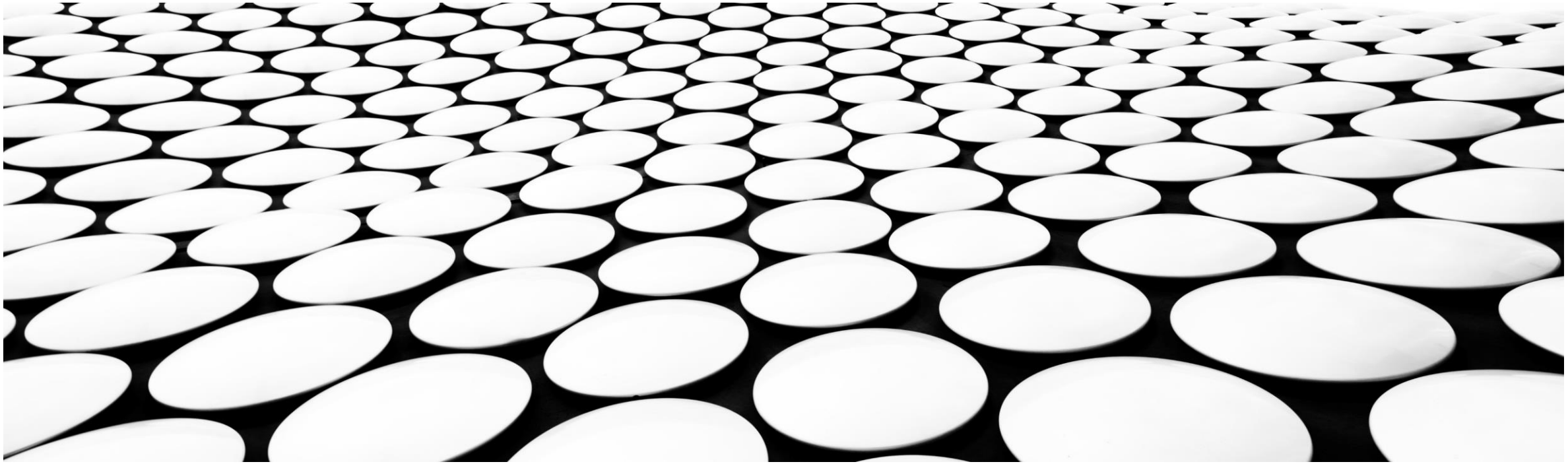
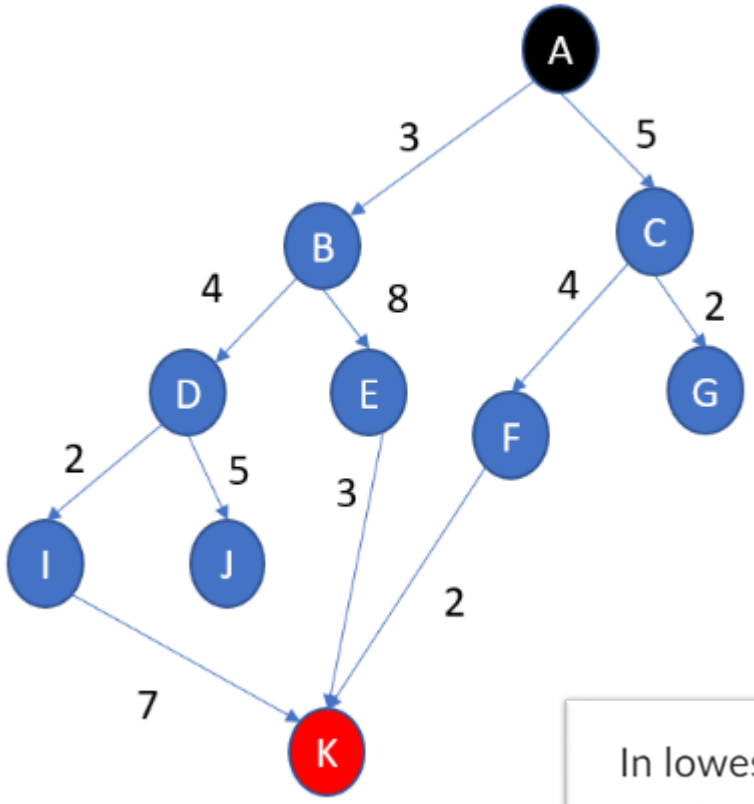
EXAMPLE OF QUESTIONS – QUIZ1



Week 2
Blind and heuristic searches

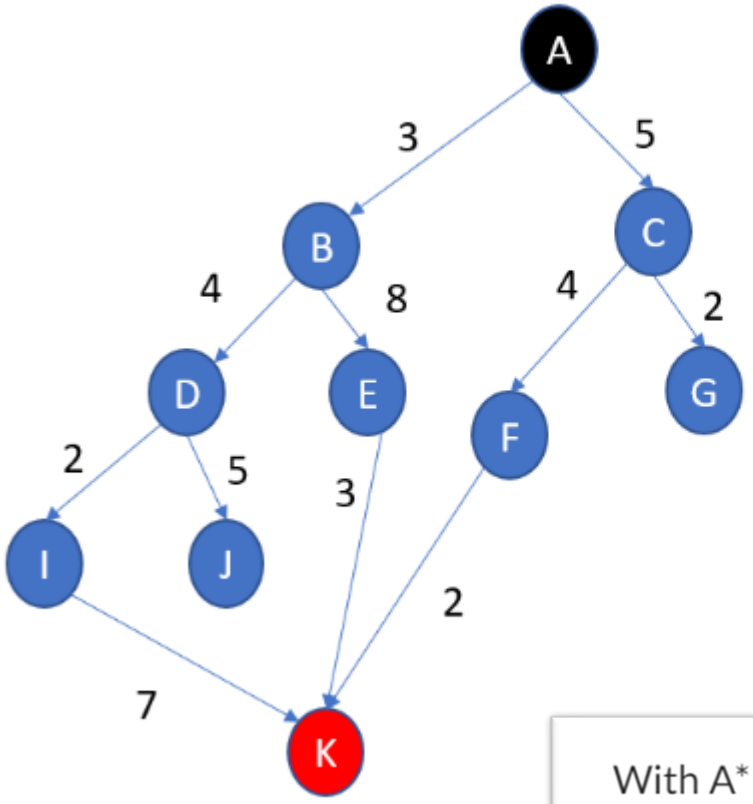


Start Node: A
End Node: K

Answer is TRUE
Both would evaluate to 7

In lowest-cost first search, the cost evaluation for node D and node G has the same result.

- ◯ True
- ◯ False



Distance to K

H(A) = 12
H(B) = 7
H(C) = 9
H(D) = 4
H(E) = 3
H(F) = 2
H(G) = 20
H(I) = 7
H(J) = 20

Answer is TRUE
At node D, accumulated cost is 7 and estimated reminder is given by h(D) = 4. The A* adds both.

With A* search, at node D, the cost considered to decide whether to open it will be 11.

- ◯ True
- ◯ False

Week 2
Constraint Satisfaction Problems

Answer is B. In the TSP seen in class, we talked about visiting all cities with minimum cost.

In the Travelling Salesman Problem (TSP), the optimization refers to:

- ◯ Minimizing the distance between two cities
- ◯ Minimizing the number of km to reach all cities
- ◯ Maximizing the number of km to reach all cities
- ◯ Maximizing the number of cities visited

Answer is B.

In the knapsack problem, the hard constraints could refer to:

- ◯ the weight of each object to put in the bag
- ◯ the maximum weight that the bag can carry
- ◯ the value of the bag
- ◯ the value of the individual objects to put in the bag

Week 2
Greedy and Randomized Searches

Answer is D.

In this algorithm, we build a solution by going mostly "greedy", but sometimes doing something else.

- ◯ Random Restart Algorithm
- ◯ Greedy Algorithm
- ◯ Random Modification Algorithm
- ◯ Random Step Algorithm

Answer is B.

In this algorithm, we build an initial solution, and then make changes to part of this solution to find a better one.

- ◯ Random Step Algorithm
- ◯ Random Modification Algorithm
- ◯ Random Restart Algorithm
- ◯ Greedy Algorithm

Week 3
Population-Based Algorithms

Answer is C.

In a genetic algorithm, the population size:

- ◯ doubles at each generation
- ◯ decreases of one individual at each generation
- ◯ is fixed through generations
- ◯ increases of one individual at each generation

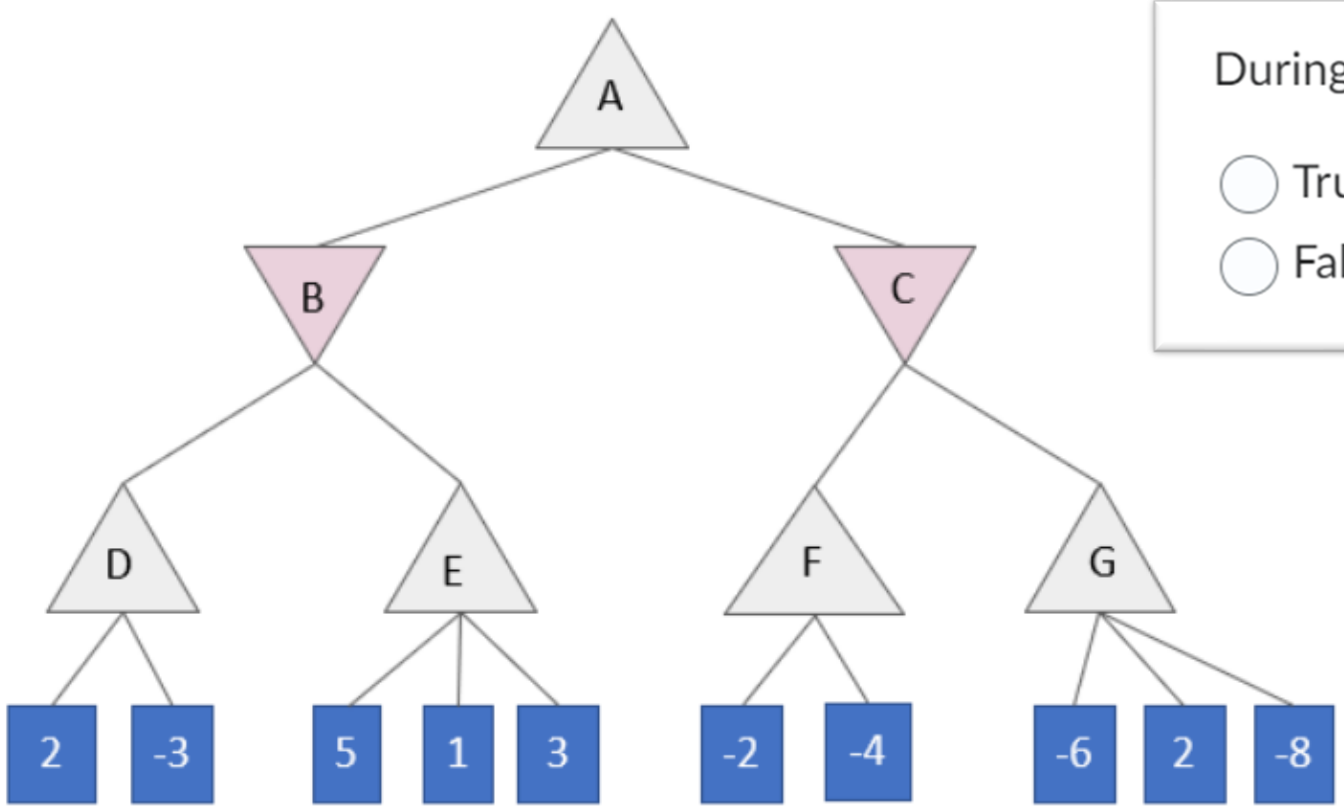
Answer is C.

In genetic algorithms, the operators are:

- alternation and cross-over
- modification et mutation
- mutation and cross-over
- mutation and alternation

Week 3
Adversarial Searches

The game in this figure, in which it is the maximizer's turn to play, will be used for the questions in this section. Assume the exploration goes from left to right.



During Minimax (no pruning), the value of C is -8.

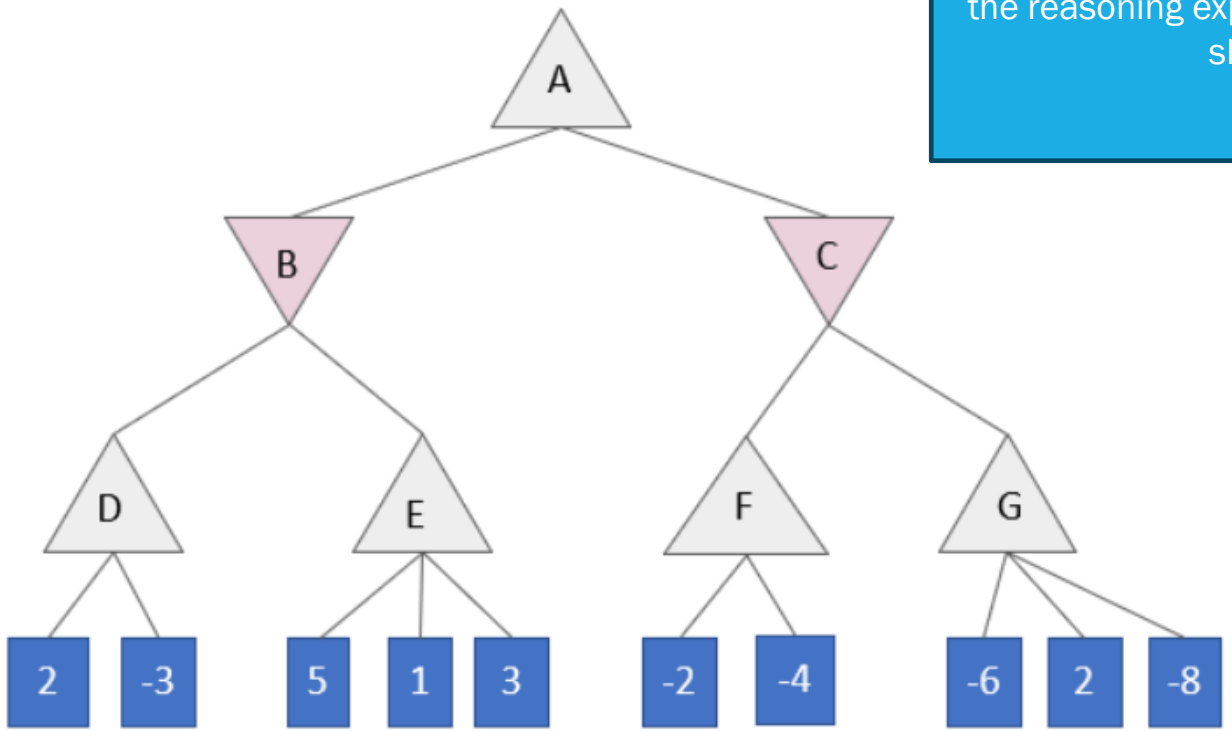
- ◯ True
- ◯ False

Answer is False. Value of C (minimiser node) is -2. Maximizer at node F would choose -2. Maximizer at node G would choose 2. So minimizer at node C would choose -2.

With alpha-beta pruning, the beta value for B, after exploring the children of D and the children of E, is 5.

- ◯ True
- ◯ False

Answer is False. Value for B is 2. (see the reasoning explained on previous slide)



Week 4
Introduction to ML

The goal of Supervised Machine Learning is to make good predictions on test data.

- ◯ True
- ◯ False

Answer is True. That is exactly what the goal of SML is.

A system with only 2 features cannot learn to separate data into 3 classes.

- ◯ True
- ◯ False

Answer is False.

The process of one-hot encoding can transform continuous features into discrete features.

- ◯ True
- ◯ False

Answer is False. Attention… I think we misread in class and said True. Discrete features are like (good, average, bad) which can be transformed into a vector of 3 features, and we would always just have one « hot », so either (1,0,0) or (0,1,0) or (0,0,1).

Annotation is the name of the automatic process for labeling data before training.

- ◯ True
- ◯ False

Answer is False. That process is manual.

A 4-fold cross-validation means that the training of the learner will be done 4 times, each time with 3/4 of the training data available.

- ◯ True
- ◯ False

Answer is True.

Week 4
Naive Bayes

Instance	Raining?	Temperature	Tired?	Activity
S1	Yes	Hot	No	Golf
S2	Yes	Warm	Yes	Bike
S3	Yes	Cold	No	TV
S4	No	Hot	No	Golf
S5	No	Warm	No	Golf
S6	Yes	Cold	No	TV
S7	Yes	Hot	Yes	TV
S8	No	Hot	Yes	TV
S9	No	Hot	No	Golf
S10	Yes	Cold	No	Bike

Evaluate P(Raining=Yes|Bike)

- ◯ 2/6
- ◯ 1/6
- ◯ 0/1
- ◯ 1/1

Answer is 1/1. There are 2 times that Bike is selected, and both times Rain is Yes.

Instance	Raining?	Temperature	Tired?	Activity
S1	Yes	Hot	No	Golf
S2	Yes	Warm	Yes	Bike
S3	Yes	Cold	No	TV
S4	No	Hot	No	Golf
S5	No	Warm	No	Golf
S6	Yes	Cold	No	TV
S7	Yes	Hot	Yes	TV
S8	No	Hot	Yes	TV
S9	No	Hot	No	Golf
S10	Yes	Cold	No	Bike

Evaluate P(Tired=Yes|Golf)

- ◯ 3/4
- ◯ 1/3
- ◯ 0/4
- ◯ 1/4

Answer is 0/4. There are 4 times that Golf is selected and among those, Tired is always No.

We need to classify a test sample in which Raining=No, Temperature=Hot and Tired=No. According to a Bayes Classifier, the class order (from most probable to least probable) would be:

- ◯ Bike-Golf-TV
- ◯ TV-Bike-Golf
- ◯ Golf-TV-Bike
- ◯ Bike-TV-Golf
- ◯ Golf-Bike-TV
- ◯ TV-Golf-Bike

Answer is Golf-TV-Bike. Use Bayes Rules to calculate P(Activity|Sample) for each activity.

Instance	Raining?	Temperature	Tired?	Activity
S1	Yes	Hot	No	Golf
S2	Yes	Warm	Yes	Bike
S3	Yes	Cold	No	TV
S4	No	Hot	No	Golf
S5	No	Warm	No	Golf
S6	Yes	Cold	No	TV
S7	Yes	Hot	Yes	TV
S8	No	Hot	Yes	TV
S9	No	Hot	No	Golf
S10	Yes	Cold	No	Bike

Week 4
Evaluation

Test	Activity (Gold Standard)	Activity Prediction
S1	Golf	Bike
S2	Bike	Bike
S3	TV	TV
S4	Golf	Bike
S5	Golf	Bike
S6	TV	TV
S7	TV	TV
S8	TV	Golf
S9	Golf	Golf
S10	Bike	Golf

The system's recall for the class Golf is:

- 0,5
- 0,75
- 0,25
- 1,0

Answer is 0.25. Among the 4 times that Golf is in the Gold Standard, the system predicted it correctly only once.

Test	Activity (Gold Standard)	Activity Prediction
S1	Golf	Bike
S2	Bike	Bike
S3	TV	TV
S4	Golf	Bike
S5	Golf	Bike
S6	TV	TV
S7	TV	TV
S8	TV	Golf
S9	Golf	Golf
S10	Bike	Golf

The system's precision for the class TV is:

- ◯ 0,5
- ◯ 0,75
- ◯ 1,0
- ◯ 0,25

Answer is 1.0. Among the 3 times that the system predicted TV, it was correct 3 times.

Week 5
Intro to NN

We are training a ***linear regression*** model using gradient descent. The training has now reached the state of the weights shown in the figure.

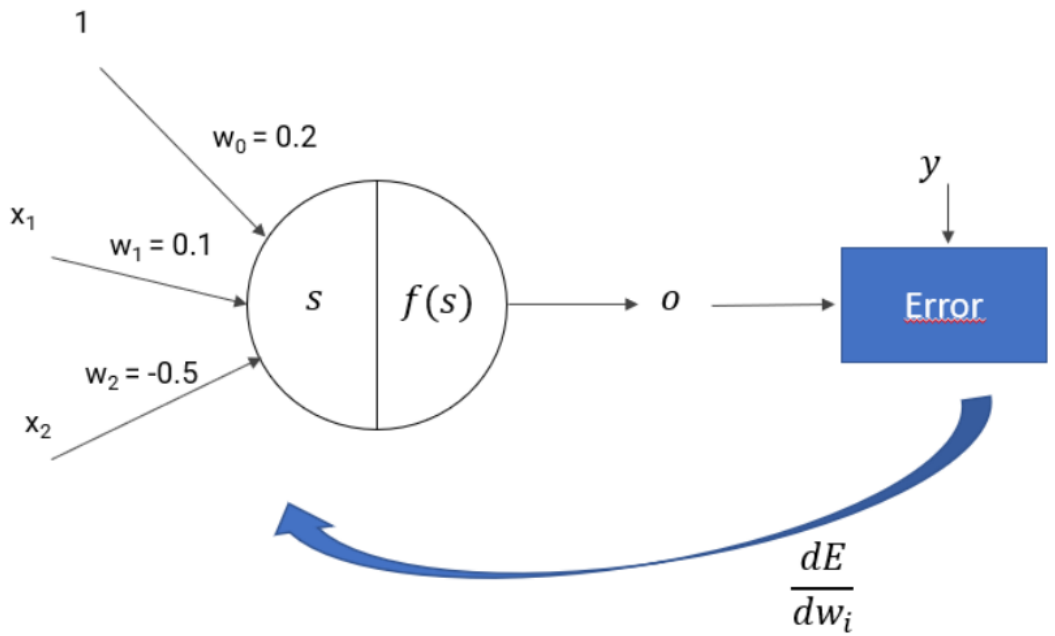
We have a training sample S in the training set (y is the target):

S = [x_1, x_2, y] = [1, 1, 0].

After adjusting the weight w_0, assuming an L2 error, and a learning parameter alpha of 0.1, what would be the new value for w_0?

- 0,22
- 0,198
- 0,202
- 0,18

Answer is 0.22. You need to apply the gradient descent algorithm for linear regression (slide 54 in the lecture on Intro to NN).



The error used for binary classification is:

- ◯ Log-Loss error
- ◯ L1 error
- ◯ L2 error
- ◯ Lmax error

Answer is A.

The Perceptron uses the function below as a non-linear transformation.

- ◯ tanh function
- ◯ sigmoid function
- ◯ step function
- ◯ ReLu function

Answer is B.

The XOR affair highlighted the fact that:

- ◯ The Perceptron cannot learn fast enough XOR types of problems.
- ◯ Gradient descent cannot be used with the Perceptron.
- ◯ The Perceptron cannot perform linear class separation with more than one class.
- ◯ The Perceptron cannot perform non-linear class separation.

Answer is D. (C is also true but it is not what the XOR affair was about)