1. Calculate probability $P(Y=1|x,w)$ given an x and w vector
    a. Also recalculate weights
2. Question regarding gradient boosting: 3 decision trees given, calculate the new target variable $(y_i - F(x_i))$. 2 examples where learning rate = 1, 1 example where learning rate = 0.5 (more complex)
3. Recommender system, collaborative filtering:
    a. Graph given with ratings of user x compared with 3 other users. Order users according to impact on predicted rating (highest absolute weight first)
    b. Given user x2 rates m 0.5 under his average, x3 rates m 0.5 under his average and x4 rates m 1 above his average. Do you expect the predicted rating to be above or below average and explain → Solve with utilizing impact of question a
4. Insight question
5. Given a set of transactions, create the <e>-projection using prefixspan (simple question)
6. Association rules
    a. calculate confidence of a rule
    b. calculate lift of a rule
    c. Which items could POSSIBLY be in the AH-projected tree of the data
        i. Answer
            1. None of the not-frequent items
            2. Not A or H since already taken out
            3. Not any item less frequent than A (since added as last)
7. Given L1 and L2 regularization and an empty graph, draw an estimation of the evolution of non-zero weights in function of the lamba for both regularizations (—> Illustration can be found in notebook of session 2)
8. Figure given with a lot of points with 5 points being bigger: Given A as initial cluster mean, which examples are likely to be picked next as cluster mean if we would use kmean++. Order the 4 points according to most likely
    a. Answer: Furthest point away from A is most likely to be picked as next cluster mean
9. Question regarding DB scan (clustering): Figure given with a lot of points and some being assigned a letter
    a. Q1: What is point A called according to DBscan
    b. Q2: What is point B called according to DBscan
    c. Q3: Would A and C be part of the same cluset?
        i. Explain density connected and see if this is the case
10. Given a 1D distribution of 10 examples (being + or -) on a x axis. We use a logistic regression classifier:

++   – – – – –   +++
_____
-1   0   1
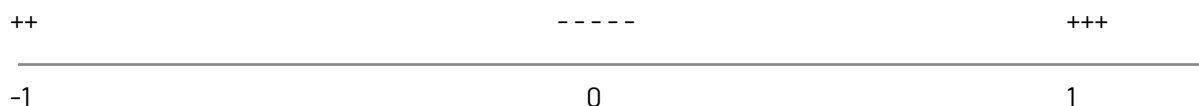
    a. What is the max accuracy that this classifier can achieve (0.8)
    b. Explain and draw decision surface (DS on right side of the - instances)
11. Given some frequent itemsets that are found using Toivonen:
    a. Give the itemsets that are going to be part of the negative border

i. Answer: All immediate supersets of the frequent sets for which ALL subsets are frequent → Generate using a similar way as Appriori Join and check if all subsets are frequent