1. In this question we consider the algorithm for learning conjunctions in the online model, as discussed in class. What happens if we run this algorithm on a noisy dataset where no conjunction is consistent with the data? Illustrate your answer with an example.

2. Consider learning classifiers in a problem where each example is a two-dimensional point with integer values and each classifier is given by a rectangle on the grid where the grid is bounded between $(0, 0)$ and $(M, M)$ where $M$ is a large number. Every classifier is given by its bottom-left point and top-right point and classifies all the points inside the rectangle as positive and points outside as negative. For example the rectangle $((1, 17), (5, 55))$ classifies the points $(1, 12)$ and $(6, 75)$ as negative and the point $(3, 22)$ as positive.

   (a) Give a simple upper bound on the number of classifiers in this class.

   (b) You are given 4 labeled examples: $(2, 7)$ is negative, $(10, 93)$ is positive, $(82, 33)$ is positive, $(84, 55)$ is negative. What concepts are consistent with this dataset?

   (c) Considering the PAC analysis given by Eq (7.2) in the Mitchell textbook: How many examples are needed to guarantee with probability $(1 - \delta) = 0.99$ that the error of a consistent learner is at most $\epsilon = 0.01$? Please explain your answer.

   (d) What is the bound if we change $\epsilon$ to be 0.001? and what is the bound if we change $\delta$ to be 0.001?

   (e) Can you give an efficient algorithm (polynomial in the number of examples; independent of $M$) that finds a consistent concept and can therefore be used with such guarantees?

   (f) How many examples are needed for the same problem but for the agnostic PAC analysis given by Eq (7.3) in the textbook for $\epsilon = \delta = 0.01$?

   (g) Can you give a simple (but less efficient) algorithm for this case (polynomial in the number of examples and $M$)?

3. Consider a dataset for frequent set mining as in the following table where we have 6 binary features and each row represents a transaction.

\[
\begin{array}{cccccc}
0 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 1
\end{array}
\]

   (a) Illustrate the first three levels of the level-wise algorithm (set sizes 1, 2 and 3) for support threshold of 3 transactions, by identifying candidate sets and calculating their support. What are the maximal frequent sets discovered in the first 3 levels?

   (b) Pick one of the maximal sets and use the algorithm on page 3 (top right) of the slides handout to check if any of its subsets are association rules with frequency at least 0.3 and confidence at least 0.6. Please explain your answer and show your work.