Lab Class Bonus Questions

The following exercises can be completed optionally to earn additional points towards the exam admission. Each exercise is worth 10 points; submissions will be graded only for students who need the points.

Exercise 1 : Concept Learning

The set of possible examples is given by all points of the x-y plane with integer coordinates from the interval $[1, 10]$. The hypothesis space is given by the set of all rectangles. A rectangle is defined by the points (x_1, y_1) and (x_2, y_2) (bottom left and upper right corner). Hypotheses are written as $\langle x_1, y_1, x_2, y_2\rangle$, and assign a point (x, y) to the value 1, if $x_1 \leq x \leq x_2$ and $y_1 \leq y \leq y_2$ hold, with arbitrary, but fixed integer values for x_1, y_1, x_2, y_2 from the interval $[1, 10]$.

(a) For the setting described above, formulate the most general hypothesis g_0.

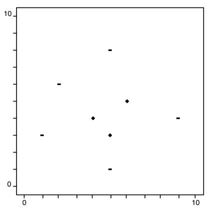
(b) Clarify for yourself how the "more-general" relation $\geq_g$ works in this setting, and check all that apply:

- ☐ $\langle 1, 2, 3, 4\rangle \geq_g \langle 1, 1, 4, 4\rangle$
- ☐ $\langle 2, 3, 6, 7\rangle \geq_g \langle 3, 4, 5, 7\rangle$
- ☐ $\langle 1, 1, 2, 8\rangle \geq_g \langle 1, 1, 3, 3\rangle$
- ☐ $\langle 3, 3, 9, 9\rangle \geq_g \langle 1, 1, 1, 1\rangle$

(c) Given a hypothesis $h = \langle 2, 3, 5, 7\rangle$, and an example $x = (2, 7)$ with $c(x) = 0$, determine two hypotheses h_1 and h_2 such that both are minimal specializations of h, and both are consistent with x.

Hint: for the correct answers h_i, there must not exist any hypothesis h' consistent with x where $h \geq_g h'$ and $h' \geq_g h_i$.

(d) Given the following training set:



No.	1	2	3	4	5	6	7	8
Point (x, y)	(5,3)	(9,4)	(1,3)	(5,8)	(4,4)	(5,1)	(6,5)	(2,6)
Class	1	0	0	0	1	0	1	0

Use the Candidate-Elimination algorithm to determine the set of the most general hypotheses G and the set of the most specific hypotheses S. Specify the hypotheses from G and S as $\langle x_1, y_1, x_2, y_2\rangle$ and draw them on the chart.

Hint: pay particular attention, when determining minimal specializations of hypotheses in G, regarding the criteria for keeping the specialized hypotheses.

(e) What happens if an additional example $x_9 = (1, 8)$ with $c(x_9) = 1$ is added?

Exercise 2 : Decision Trees and Statistical Learning

Given the following training set with dogs data:

Color	Fur	Size	Class
brown	ragged	small	well-behaved
black	ragged	big	dangerous
black	smooth	big	dangerous
black	curly	small	well-behaved
white	curly	small	well-behaved
white	smooth	small	dangerous
red	ragged	big	well-behaved

(a) Use the ID3 algorithm to determine a decision tree, where the attributes are to be chosen with the maximum average information gain *iGain*:

$$\mathit{iGain}(D, A) \equiv H(D) - \sum_{a \in A} \frac{|D_a|}{|D|} \cdot H(D_a) \quad \text{with} \quad H(D) = -p_\oplus \log_2(p_\oplus) - p_\ominus \log_2(p_\ominus)$$

Thoroughly explain the steps of your calculations.

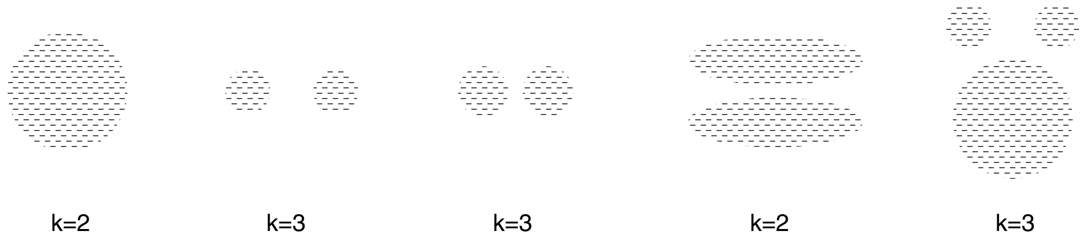
(b) Instead of the decision tree, determine the parameters for a Naïve Bayes classifier on the same dataset.

(c) Classify the new example (Color=black, Fur=ragged, Size=small) using both above classifiers.

Exercise 3 : k-Means

(a) Given k equally sized classes with n objects each. What is the probability that by a random selection with replacement of k objects exactly one object from each class is chosen?

(b) Given are the data sets shown below. Illustrate the clustering results of k-means for the specified values of k and mark the positions of the respective centroids. Assume that the squared error objective function is employed. If more than one solutions appear plausible, differentiate between them in terms of a global or a local minimum.



(c) The Voronoi diagram for a set M of k points in a plane is a partition of *all points of the plane* into k regions, such that every point of the plane is assigned to its closest point in M (see the figure below). Explain the relationship between Voronoi diagrams and k-Means clusters.