1.1. Give three computer applications for which machine learning approaches seem appropriate and three for which they seem inappropriate. Pick applications that are not already mentioned in this chapter, and include a one-sentence justification for each.

Ans.

Machine learning: Face recognition, handwritten recognition, credit card approval.

Not machine learning: calculate payroll, execute a query to database, use WORD.

2.1. Explain why the size of the hypothesis space in the EnjoySport learning task is 973. How would the number of possible instances and possible hypotheses increase with the addition of the attribute WaterCurrent, which can take on the values Light, Moderate, or Strong? More generally, how does the number of possible instances and hypotheses grow with the addition of a new attribute A that takes on k possible values?

Ans.

Since all occurrence of "ϕ" for an attribute of the hypothesis results in a hypothesis which does not accept any instance, all these hypotheses are equal to that one where attribute is "ϕ". So the number of hypothesis is 4*3*3*3*3*3 +1 = 973.

With the addition attribute Watercurrent, the number of instances = 3*2*2*2*2*2*3 = 288, the number of hypothesis = 4*3*3*3*3*3*4 +1 = 3889.

Generally, the number of hypothesis = 4*3*3*3*3*3*(k+1)+1.

2.3. Consider again the EnjoySport learning task and the hypothesis space H described in Section 2.2. Let us define a new hypothesis space H' that consists of all pairwise disjunctions of the hypotheses in H. For example, a typical hypothesis in H' is (?, Cold, High, ?, ?, ?) v (Sunny, ?, High, ?, ?, Same) Trace the CANDIDATE-ELIMINATATION algorithm for the hypothesis space H' given the sequence of training examples from Table 2.1 (i.e., show the sequence of S and G boundary sets.)

Ans.

S0= (ϕ,ϕ,ϕ,ϕ,ϕ,ϕ) v (ϕ,ϕ,ϕ,ϕ,ϕ,ϕ)

G0 = (?, ?, ?, ?, ?, ?) v (?, ?, ?, ?, ?, ?)

Example 1: <Sunny, Warm, Normal, Strong, Warm, Same, Yes>

S1=(*Sunny, Warm, Normal, Strong, Warm, Same)* v (ϕ,ϕ,ϕ,ϕ,ϕ,ϕ)

G1 = (?, ?, ?, ?, ?, ?) v (?, ?, ?, ?, ?, ?)

Example 2: <Sunny, Warm, High, Strong, Warm, Same, Yes>

S2= {(*Sunny, Warm, Normal, Strong, Warm, Same)* v *(Sunny, Warm, High, Strong, Warm, Same*),

(Sunny, Warm, ? , Strong, Warm, Same) v (ϕ,ϕ,ϕ,ϕ,ϕ,ϕ)}

G2 = (?, ?, ?, ?, ?, ?) v (?, ?, ?, ?, ?, ?)

Example 3: <Rainy, Cold, High, Strong, Warm, Change, No>

S3={(*Sunny, Warm, Normal, Strong, Warm, Same)* v *(Sunny, Warm, High, Strong, Warm, Same*),

(Sunny, Warm, ? , Strong, Warm, Same) v (ϕ,ϕ,ϕ,ϕ,ϕ,ϕ)}

G3 = {(Sunny, ?, ?, ?, ?, ?) v *(?, Warm, ?, ?, ?, ?),*

(Sunny, ?, ?, ?, ?, ?) v *(?, ?, ?, ?, ?,* Same*),*

(?, Warm, ?, ?, ?, ?) v *(?, ?, ?, ?, ?,* Same*)}*

Example 4: <Sunny, Warm, High, Strong, Cool, Change, Yes>

S4= {~~(Sunny, Warm, ?, Strong, ?, ?) v (Sunny, Warm, High, Strong, Warm, Same)~~,

~~(Sunny, Warm, Normal, Strong, Warm, Same) v (Sunny, Warm, High, Strong, ?, ?)~~,

~~(*Sunny, Warm,* ? *, Strong,* ? *,* ? *)* v (ϕ,ϕ,ϕ,ϕ,ϕ,ϕ),~~

(*Sunny, Warm,* ? *, Strong, Warm, Same)* v *(Sunny, Warm, High, Strong, Cool, Change*)}

G4 = {(Sunny, *?, ?,* ?, ?, ?) v *(?, Warm, ?,* ?, ?, *?),*

(Sunny, *?, ?,* ?, ?, ?) v *(?, ?, ?,* ?, ?, Same*),*

(?, Warm, ?, ?, ?, *?)* v *(?, ?, ?,* ?, ?, Same*)}*

2.4. Consider the instance space consisting of integer points in the x, y plane and the set of hypotheses H consisting of rectangles. More precisely, hypotheses are of the form $a \leq x \leq b$, $c \leq y \leq d$, where a, b, c, and d can be any integers.

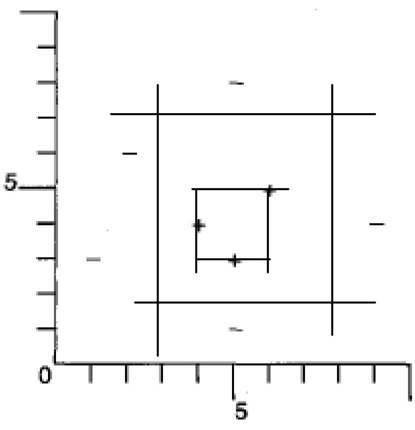
(a) Consider the version space with respect to the set of positive (+) and negative (-) training examples shown below. What is the S boundary of the version space in this case? Write out the hypotheses and draw them in on the diagram.

(b) What is the G boundary of this version space? Write out the hypotheses and draw them in.

(c) Suppose the learner may now suggest a new x, y instance and ask the trainer for its classification. Suggest a query guaranteed to reduce the size of the version space, regardless of how the trainer classifies it. Suggest one that will not.

(d) Now assume you are a teacher, attempting to teach a particular target concept (e.g., $3 \leq x \leq 5$, $2 \leq y \leq 9$). What is the smallest number of training examples you can provide so that the CANDIDATE-ELIMINATION algorithm will perfectly learn the target concept?

Ans. (a) S= (4,6,3,5) (b) G=(3,8,2,7) (c) e.g., (7,6), (5,4) (d) 4 points: (3,2,+), (5,9,+), (2,1,-),(6,10,-)



2.6. Complete the proof of the version space representation theorem (Theorem 2.1).

Proof: Every member of $VS_{H,D}$ satisfies the right-hand side of expression.

Let h be an arbitrary member of $VS_{H,D}$, then h is consistent with all training examples in D. Assuming h does not satisfy the right-hand side of the expression, it means $\neg(\exists s \in S)\exists(g \in G)(g \geq h \geq s) = \neg(\exists s \in S)\exists(g \in G)\ (g \geq h) \wedge (h \geq s)$. Hence, there does not exist g from G so that g is more general or equal to h or there does not exist s from S so that h is more general or equal to s.

If the former holds, it leads to an inconsistence according to the definition of G. If the later holds, it leads to an inconsistence according to the definition of S. Therefore, h satisfies the right-hand side of the expression. □ (Notes: since we assume the expression is not fulfilled, this can be only be if S or G is empty, which can only be in the case of any inconsistent training examples, such as noise or the concept target is not member of H.)