

CSCI/ARTI 8950 Machine Learning

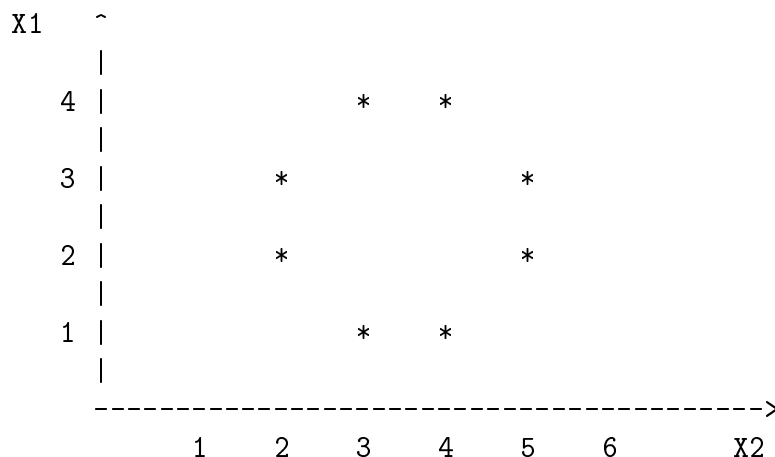
Assignment Number 6: Due 4/21/2011 (in class)

1. [15 points][FIN] Consider the following training set of samples for machine learning:

Example	Sentence	Label
1	This test is easy	+
2	This problem is easy	+
3	Two is more than three	-
4	Swimming is easy	-

Apply the Naive Bayes method to the above examples. You should:

- Identify the vocabulary.
 - For every word W_i in the vocabulary, compute $P(W_i \parallel +)$ and $P(W_i \parallel -)$. (Hint, many probabilities are equal)
 - Use the Naive Bayes classifier to classify the following sentences:
 - This problem is more than easy
 - No test is easy
2. [10 points][FIN]
- Propose a lazy version of the back-propagation algorithm for training neural networks. What are the advantages and disadvantages of your algorithm, compared to the original back-propagation algorithm?
 - Propose an eager version of the nearest neighbour algorithm for classification. What are the advantages and disadvantages of your algorithm, compared to the original nearest neighbour algorithm?
3. [15 points][FIN] Consider the following diagram of a set of 8 instances for machine learning:



- (a) Consider the hypothesis space H_1 consisting of all possible **circles** in the plane (i.e. each hypothesis h in H_1 is a circle which classifies all points in it as positive and all points outside it as negative). Does H_1 shatter the given set of instances? Briefly justify your answer.
 - (b) Consider the hypothesis space H_2 consisting of all possible **rectangles** in the plane. Does H_2 shatter the given set of instances? Briefly justify your answer.
 - (c) Based **only** on your answers to parts (a) and (b) above, what can you conclude about the VC dimensions of H_1 and H_2 ?
4. **[10 points]** The exclusive OR function (XOR) takes two binary (bit) inputs X_1 and X_2 and produces one binary output which is 1 if X_1 and X_2 are different and 0 otherwise.
- (a) Give a decision tree which correctly computes the XOR function for any values of X_1 and X_2 .
 - (b) Give a classifier system in the format used by the GIL program which correctly computes the XOR function for any values of X_1 and X_2 (i.e. matches the value pairs that result in a 1 and does not match the others).