Comp 6321 - Machine Learning - Assignment 3

Federico O'Reilly Regueiro

November 10h, 2016

Question 1: Midterm preparation question

Propose an adequate learning algorithm.

1.a 1000 samples, 6-dimensional continuous space, classify ≈100 examples.

1.b Clasifier for children in special-ed, justified to the board before it's implemented.

1.c Binary classification of 1 million bits (empirical preference rate for others), very large data-set. Frequent updates.

1.d 40 attributes, discrete and continuous, some have noise; only about 50 labeled observations.

Question 2: Properties of entropy

2.a Compute the following for $p(0,0) = 1/3, p(0,1) = 1/3, p(1,0) = 0, p(1,1) = 1/3$.

i $H[x]$

ii $H[y]$

iii $H[y|x]$

iv $H[x|y]$

v $H[x, y]$

vi $I[x, y]$

2.b

2.c

Question 3: Kernels

3.a

3.b

3.c

3.d

3.e

Question 4: Nearest neighbour vs decision trees

Question 5: Bayes rate

5.a

5.b

5.c

Question 6: Implementation