ML 4375 – Intro to Machine Learning – Summer 2018 – Mazidi – Exam 2 Review

Test format:

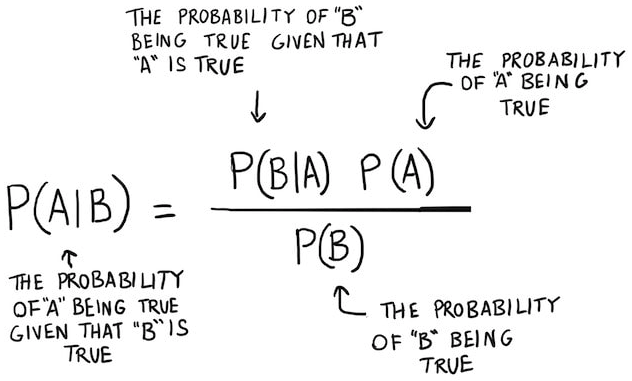
* 20 multiple-choice or T/F questions (3 points each)
  + ***you may circle more than one answer choice for the MC***
  + emphasis on these algorithms:
    - Naïve Bayes
    - SVM
    - Decision Trees
    - Neural Networks
  + Know how they work; special characteristics; bias/variance; regression/classification
* 10 multiple-choice or T/F calculation questions (4 points each)

Terminology and Concepts:

* bias and variance
* discriminative v. generative classifier
* regression, classification
* entropy, information gain, entropy and how to calculate them (formulas provided)

Formulas Provided on Exam:

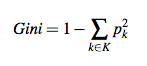
Bayes’ Theorem:



Entropy:

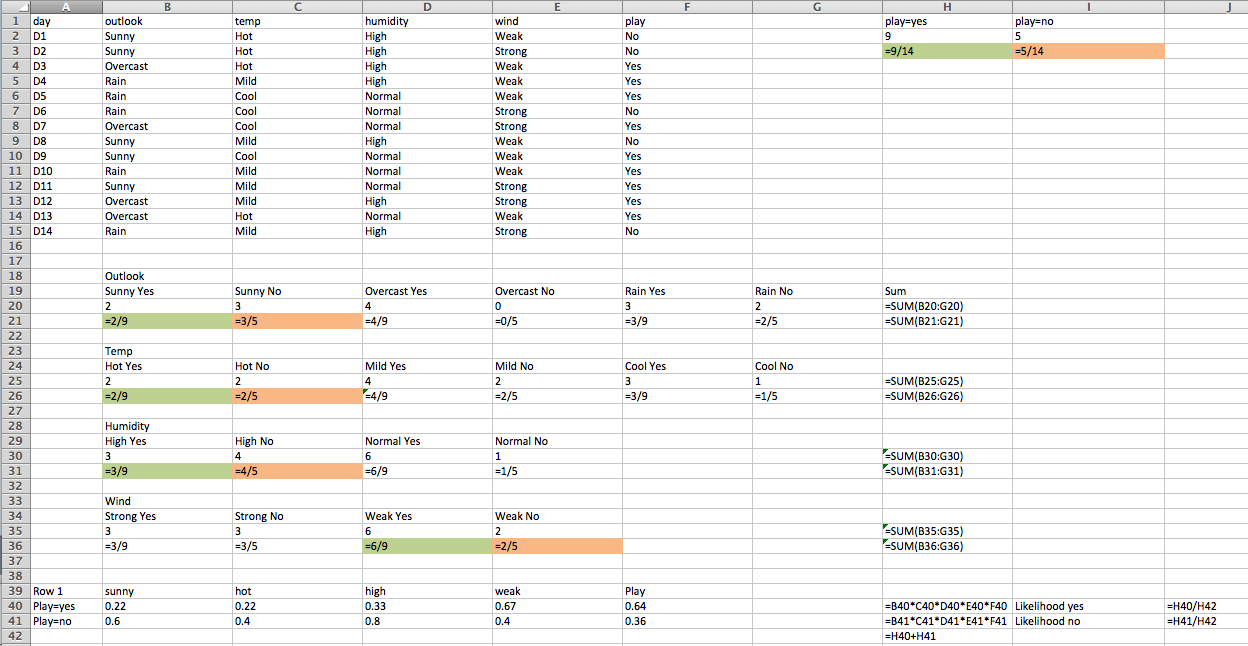


Information Gain:



Gini Index:

Practice

Naïve Bayes

1. What is the likelihood P(sunny|play)?

2/9

1. What is the likelihood P(strong|play)?

3/9

1. What is P(play|sunny)?

2/9\*9/14 / [2/9\*9/14 + 7/9\*9/14]

1. What is P(play|strong)?

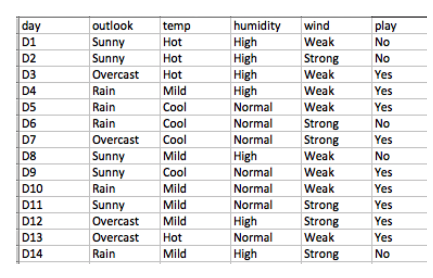
3/9\*9/14 / [3/9\*9/14 + 6/9\*9/14]

1. What is P(play|sunny,strong)?

2/9\*3/9\*9/14 / [2/9\*3/9\*9/14 + 7/9\*6/9\*9/14]

Practice

Decision Trees



1. What is H(play)?

-9/14 \* log 9/14 – 5/14 \* log 5/14 = .94

1. What is H(strong)?

-3/6 \* log 3/6 – 3/6 \* log 3/6 = 1

What is H(weak)?

-6/8 \* log 6/8 – 2/8 \* log 2/8 = .811

1. What is IG(wind)?

.94 – 8/14 \* .811 - 6/14 \* 1 = 0.048

1. What is Gini(play)?

1 – (9/14)^2 – (5/14)^2 = 0.459