शिक्षा मंत्रालय
MINISTRY OF
EDUCATION
सत्यमेव जयते

INDIAN INSTITUTE OF TECHNOLOGY
JODHPUR

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NPTEL

PMRF
Prime Minister's Research Fellowship

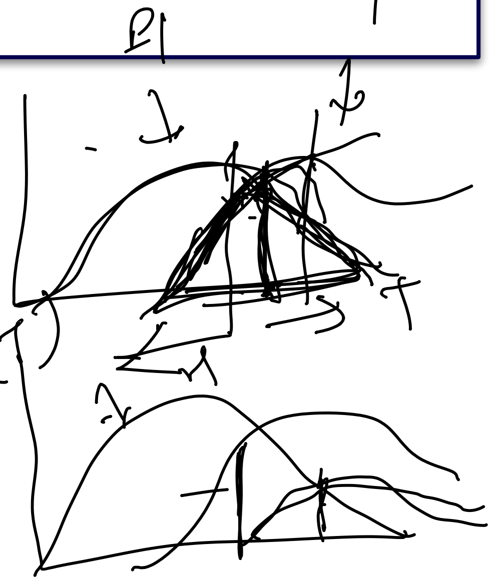
Week 3 - Live Session

Data Mining

Swapnil S. Mane
mane.1@iitj.ac.in
PMRF Research Scholar

Q1. In a multiclass classification problem, Bayes classifier assigns an instance to the class corresponding to:

a) Maximum aposteriori probability
b) Maximum apriori probability
c) Lowest aposteriori probability
d) Lowest apriori probability



Q2. Which of the following is incorrect about Naive Bayes:

a) Attributes can be nominal or numeric

b) Attributes are statistically dependent on one another given the class value.

c) Attributes are equally likely.

d) All of the above.

Q3. A fair coin is tossed n times. The probability that the difference between the number of heads and tails is (n-3) is:

a) 2^{-n}
b) 0
c) C(n,n-3)*2-n
d) 2^{-n+3}

h = heads
$n-h$ = tails

$h-(n-h) = n-3$

$h = (2n-3)/2 = n - 3/2$

Q4. Three companies supply bulbs. The percentage of bulbs supplied by them and the probability of them being defective is given below:

$P(B/D) = 2.$

Company	% of bulbs supplied	Probability of defective
A	60	0.01
B	30	0.02
C	10	0.03

Given that the bulb is defective probability that it is supplied by B is:

a) 0.1
b) 0.2
c) 0.3
d) 0.4

$$P(B/D) * P(D) = P(D/B) * P(B) \Rightarrow P(B/D) = \frac{P(D/B) * P(B)}{P(D)}$$

$$P(D/B) * P(B) = 0.02 * 0.3 = 0.006$$

$$P(D) = P(D/A)\,P(A) + P(D/A) \cdot P(A) + P(D/C) * P(C)$$

$$= 0.01 * 0.6 + 0.02 * 0.3 + 0.03 * 0.1$$

$$P(D) = 0.015$$

$$= 0.006 / 0.015 = 0.4$$

Q5. If $P(Z \cap X) = 0.2$, $P(X) = 0.3$, $P(Y) = 1$ then $P(Z|X \cap Y)$ is:

$$P(Z | X \cap Y) = \frac{P(Z \cap X \cap Y)}{P(X \cap Y)}$$

a) 0
b) 2/3
c) Not enough data.
d) None of the above

$P(Y) = 1$

$P(A \cap Y) = P(A)$

$P(Z \cap X \cap Y) = P(Z \cap X)$

$P(X \cap Y) = P(X)$

$$P(Z | X \cap Y) = \frac{P(Z \cap X)}{P(X)} = \frac{0.2}{0.3} = 2/3$$

For questions 6-7, consider the following hypothetical data regarding the hiring of a person.

GPA	Effort	Confidence	Hire
Low	Some	Yes	No
Low	Lots	Yes	Yes
High	Lots	No	No
High	Some	No	Yes
High	Lots	Yes	Yes

Y = 3
N = 2

Q6. Using Naïve Bayes determine whether a person with GPA=High, Effort=Some, and Confidence=Yes be hired:

a) Yes
b) No
c) The example cannot be classified.
d) Both classes are equally likely

$p(F_1 F_2 F_3 / \text{Hired})$ –

$p(F_1 F_2 F_3 / \text{N Hired})$

$$p(\text{Class} / \text{Feat}) = \frac{p(F/C) \cdot p(C)}{p(F)}$$

$p(H) = \frac{3}{5}$

$p(N) = \frac{2}{5}$

$$p(C/F) \propto p(F/C) \cdot p(C)$$

→ ① $p(G=H, E=S, C=Y / \text{Hire}=Y) \cdot p(H) =$

② $p(G=H, E=S, C=Y / \text{Hire}=N) \cdot p(N)$

① $p(G=H/H) * p(E=S/H) * p(C=Y/H)$

$\frac{2}{3} * \frac{1}{3} * \frac{2}{3} = \frac{4}{27} * \frac{3}{5} = \frac{12}{135} = \frac{4}{45}$

Y

② $\frac{1}{2} * \frac{1}{2} * \frac{1}{2}$

$= \frac{1}{8} * \frac{2}{5}$

$= \frac{2}{40} = \frac{1}{20}$

NH

$\frac{4}{45} > \frac{1}{20}$

Q7. Using Naïve Bayes determine whether a person with Effort=lots, and Confidence=No be hired:

a) Yes ✓
b) No
c) The example cannot be classified
d) Both classes are equally likely

$0.133 \leftarrow P(H / E=L, C=N)$

$0.1 \leftarrow P(NH / E=L, C=N)$

$P(H / E=L, C=N) = \dfrac{P(E=L, C=N / H) * P(H)}{P(E=L, C=N)}$

$\Rightarrow P(E=L / H) * P(C=N / H) \quad = \left(\frac{2}{3} * \frac{1}{2}\right) \frac{3}{5} = 0.133$

$\left(P(E=L / NH) * P(C=N / NH)\right) * P(NH)$

$\left(\frac{1}{2} * \frac{1}{2}\right) * \frac{2}{5} = 0.1$