CISC 7700X Midterm Exam

Pick the best answer that fits the question. Not all of the answers may be correct. If none of the answers fit, write your own answer.

1. (5 points) A *model* is:
   (a) A description.
   (b) A fact.
   (c) A data point.
   (d) All of the above.

2. (5 points) Both mean and median measure:
   (a) The spread of the data.
   (b) The central tendency of the data.
   (c) The slope of the data.
   (d) The gradient of the data.

3. (5 points) Both standard deviation and interquartile range measure:
   (a) The central tendency of the data.
   (b) The slope of the data.
   (c) The spread of the data.
   (d) The gradient of the data.

4. (5 points) If $P(x, y) \neq P(x)P(y)$ then
   (a) x is more likely than y.
   (b) x is causes y.
   (c) x and y are independent.
   (d) x and y are not independent.
   (e) None of the above, answer is:

5. (5 points) The process of computing $P(x)$ from $P(x|y)P(y)$ is called
   (a) Marginalizing
   (b) Bootstrapping
   (c) Generalizing
   (d) Specifizing

6. (5 points) In Bayes rule: $P(x|y) = P(y|x)P(x)/P(y)$, the $P(y|x)$ is:
   (a) The prior probability.
   (b) The likelihood.
   (c) The posterior probability.

(d) The conditional probability of y given x.

7. (5 points) Conditional probability $P(y|x)$ differs from likelihood $P(y|x)$:
   (a) They're both the same.
   (b) They both sum to 1.
   (c) Probability $P(y|x)$ is a function of y, while likelihood $P(y|x)$ is a function of x.
   (d) Likelihood tells us the probability of y given x.

8. (5 points) Which one of these is correct?
   (a) $P(A|B) = \frac{P(B|A)P(A)}{\sum P(A,B)}$
   (b) $P(A|B) = P(B|A)P(A)P(B)$
   (c) $P(A|B) = P(A,B)/P(B|A)$
   (d) $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

9. (5 points) Which one of these is correct?
   (a) $P(A,B,C) = P(A|B,C)P(B,C)$
   (b) $P(A,B,C) = P(A|B)P(B|C)P(C)$
   (c) $P(A,B,C) = P(A|C)P(C|B)P(B)$
   (d) $P(A,B,C) = P(A|B)P(A|C)P(B)P(C)$

10. (5 points) If $P(x|y) \neq P(x,y)/P(y)$ then
    (a) x is more likely after y.
    (b) y is causes x.
    (c) x and y are independent.
    (d) x and y are not independent.
    (e) None of the above, answer is:

11. (5 points) From our past experience, we know it rains 1 in 5 days. When it rains, we observe 90% of the people carry umbrellas. When it's not raining, only 10% of the people carry an umbrella. We're in a basement (no windows), and we observe someone walking in with an umbrella. Using Bayes rule, what's the probability that it's raining?

Answer is :

12. (5 points) Continuing from previous question, we next observe someone wearing a rain-jacket. When it rains, we know about 90% of the people wear rain-jackets, and only about 10% of the people wear rain-jackets when it's not raining. Use the Bayes rule to find probability of rain given this additional evidence.

Answer is :

13. (5 points) Continuing from previous question, using Naive Bayes assumption, what's the probability of rain now that we've observed both umbrella *and* a rain-jacket?

Answer is :

14. (5 points) From past data, about 90% of people get RSV (Respiratory Syncytial Virus). Use the Bayes rule to find probability of getting RSV?

Answer is :

15. (5 points) Whenever someone has RSV, there is an 80% they will have a cough. Coughing show up in 40% from other causes. You notice a caugh. Use Bayes rule to find probability of RSV.

Answer is :

16. (5 points) Whenever someone has RSV, there is an 90% they will have a fever. Fever shows up in 40% from other causes. You notice a fever. Use Bayes rule to find probability of RSV.

Answer is :

17. (5 points) You notice a fever and a cough. Use Bayes rule to find probability of RSV.

Answer is :

18. (5 points) You notice a fever and a cough. Use Naive Bayes rule to find probability of RSV.

Answer is :

19. (5 points) The answer to previous question is:

    (a) The exact probability of rain given the evidence.

    (b) An overestimate.

    (c) An underestimate.

    (d) Would change if we first observed rain-jacket followed by umbrella.

20. (5 points) Given a sample of N data points, we discover that we can fit two models, a line: $y = w_0 + w_1 x$ and a polynomial:

$$y = w_0 + w_1 x + w_2 x^2 + w_3 x^3 + w_4 x^4 + w_5 x^5$$

    The polynomial fits our training dataset 'better'. Which is true:

    (a) We'd expect the line to have higher variance, but lower bias.

    (b) We'd expect both to have equivalent bias and variance.

    (c) We'd expect the line to have lower variance, but higher bias.

    (d) We'd expect the polynomial to perform better on other samples.