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 fact.
   (b) A data point.
   (c) A description.
   (d) All of the above.
   (e) None of the above, the answer is:
2. (5 points) The central tendency of the data is measured by:
   (a) standard deviation
   (b) variance
   (c) interquartile range measure
   (d) median
   (e) None of the above, the answer is:
3. (5 points) The spread of the data is measured by:
   (a) geometric mean
   (b) variance
   (c) arithmetic mean
   (d) Pearson correlation
   (e) None of the above, the answer is:
4. (5 points) You are training a machine learning model and want to evaluate its performance properly:
   (a) Splitting your data into training and test sets is sufficient to estimate the model's generalization performance in all cases.
   (b) Cross-validation provides a way to use the entire dataset for both training and evaluation, reducing variance in performance estimates.
   (c) The validation set is used to test the final model performance on unseen data.
   (d) Increasing the size of the training set decreases both bias and variance simultaneously, regardless of model complexity.
   (e) Testing the model on the training set gives a reliable estimate of how it will perform on new data.

5. (5 points) The process of computing $P(x)$ from $P(x|y,z)P(y|z)P(z)$ is called

   (a) Bootstrapping

   (b) Marginalizing

   (c) Generalizing

   (d) Specifizing

   (e) None of the above, answer is:

6. (5 points) Suppose we have $P(A, B, C, D, E, F)$, where each of the $A, \ldots, F$ has values from 1 to 1000. Write a pseudo-code loop to find $P(F)$.

   (e) Answer is:

7. (5 points) Fair coin flipping game: We start with \$1. Heads we win 50%, tails we lose 50%. After 3 rounds, with a fair coin, the *median* value we will have:

   (e) Answer is:

8. (5 points) Fair coin flipping game: We start with \$1. Heads we win 50%, tails we lose 50%. After 3 rounds, with a fair coin, the *geometric mean* value we will have:

   (e) Answer is:

9. (5 points) Fair coin flipping game: We start with \$1. Heads we win 50%, tails we lose 50%. After 3 rounds, with a fair coin, the *arithmetic mean* value we will have:

   (e) Answer is:

10. (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.

    (e) None of the above, answer is:

11. (5 points) Conditional probability $P(y|x)$ differs from likelihood $P(y|x)$:

    (a) They both sum to 1.

    (b) Likelihood tells us the probability of y given x.

    (c) Probability $P(y|x)$ is a function of y, while likelihood $P(y|x)$ is a function of x.

    (d) Probability $P(y|x)$ is a function of x, while likelihood $P(y|x)$ is a function of y.

    (e) None of the above, answer is:

12. (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) $P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|-A)P(-A)}$

    (e) None of the above, answer is:

13. (5 points) Which one of these is correct?

    (a) $P(A,B,C) = P(A|B)P(B|C)P(C)$

    (b) $P(A,B,C) = P(A|C)P(C|B)P(B)$

    (c) $P(A,B,C) = P(A|B,C)P(B|C)P(C)$

    (d) $P(A,B,C) = P(A|B)P(A|C)P(B)P(C)$

    (e) None of the above, answer is:

14. (5 points) If $P(x|y) = P(x,y)/P(y)$ 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:

15. (5 points) Customers return 10% of items purchased online. 30% of all items are on sale. Of the returned items, 60% were originally purchased on sale. A customer purchases an item on sale; use Bayes rule to find probability it will be returned.

Answer is :

16. (5 points) Continuing from previous question. Of the returned items, 90% are fragile, while of the non-returned items only 10% are fragile. A customer purchases a fragile item; use Bayes rule to find probability it will be returned.

Answer is :

17. (5 points) Continuing from previous question. A customer purchases a fragile item on sale. Use Bayes rule to find probability it will be returned.

Answer is :

18. (5 points) Continuing from previous question. A customer purchases a fragile item on sale. Use Naive Baye to find probability it will be returned.

Answer is :

19. (5 points) A soda company produces 1 million bottles of drinks. 90% of the bottles are "diet" sodas, and 10% are regular sodas.

- Drinking a diet soda carries a 0.5% chance of causing a mild reaction, and 5% of those reactions are severe enough to require hospitalization. Diet sodas never cause weight gain.
- Regular sodas have a 5% chance of causing weight gain, and 10% of those who gain weight develop a serious health complication.

A quick calculation shows:

- Expected hospitalizations from diet soda:

$$1{,}000{,}000 \times 0.90 \times 0.005 \times 0.05 = 225$$

- Expected serious health complications from regular soda:

$$1{,}000{,}000 \times 0.10 \times 0.05 \times 0.10 = 50$$

One might conclude: "Drinking diet soda is more dangerous than drinking regular soda!" What is wrong with this analysis?

(e) Answer:

20. (5 points) Consider a model predicting house prices:

- A simple linear model might consistently underpredict or overpredict prices for certain types of houses.
- A very complex model with many parameters can fit the training data almost perfectly but performs poorly on new data.

Which of the following statements best describes the bias-variance tradeoff?

(a) High-bias models tend to underfit the data, while high-variance models tend to overfit the data.

(b) High-bias models tend to overfit the data, while high-variance models tend to underfit the data.

(c) High-bias models always have lower training error than high-variance models.

(d) High-variance models are preferred because they minimize error on new data.

(e) Bias and variance are unrelated; only the training data size matters.