Machine Learning (ML) in Hindi

Week-1 Assignment

Maximum points: 20

Instructions

- Each question is a 1-point multiple choice question (MCQ) with only one correct answer until specified otherwise.

1. Which of the following best describes the relationship between artificial intelligence (AI), machine learning (ML), and deep learning (DL). ['*a>b*' symbolizes b is contained in a.]
    a. DL>ML>AI
    b. AI>DL>ML
    c. AI>ML>DL
    d. ML>DL>AI

Solution: (c)

2. Who among the following is not titled as one of the Godfathers of AI?
    a. Yann LeCun
    b. Arthur Samuel
    c. Yoshua Bengio
    d. Geoffrey Hinton

Solution: (b)

3. Which of the following is NOT a supervised learning task?
    a. Given the heights of students in the class, the task is to train a model to group students based on their heights.
    b. Given a set of labelled emails, the task is to train a model to predict whether a new incoming email is a spam.
    c. Given a set of labelled images, the task is to train a model to classify new images into one of the predetermined classes.
    d. Given a set of labelled text data, the task is to train a model to predict the sentiment of new text data as positive, negative, or neutral.

Solution: (a)

This is a clustering task, an unsupervised learning problem.

4. A phone monitoring machine classifies calls as incoming (I) and outgoing (O) calls. Similarly, the call can be classified as voice (V), or data (D) call. Based on statistics collected by the telephone company, the following probability model is deduced:

$$P[V] = 0.55$$
$$P[I] = 0.65$$
$$P[V \cap I] = 0.4$$

the respective probabilities of $P[V \cap O]$ and $P[V \cup I]$ are

   a. 0.17 and 0.8
   b. 0.15 and 0.8
   c. 0.17 and 0.95
   d. 0.15 and 0.95

Solution: (b)

$$P[V \cap O] = P[V] - P[V \cap \mathrm{I}] = 0.55 - 0.4 = 0.15$$
$$P[V \cup I] = P[V] + P[I] - P[V \cap I] = 0.55 + 0.65 - 0.4 = 0.8$$

5. Assume a hypothetical fair die of seven faces numbered 1 to 7. The expectation and variance of an event of rolling this die are, respectively,

   a. 4 and 4
   b. 7 and 7
   c. 4 and 7
   d. 7 and 4

Solution: (a)

Let X be the outcome when we roll a fair die of 7 faces.

$$X \in \{1,2,3,4,5,6,7\}$$

$$P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = P(X = 5) = P(X = 6) = P(X = 7) = \frac{1}{7}$$

Then, the expected value of X is

$$E[X] = 1\left(\frac{1}{7}\right) + 2\left(\frac{1}{7}\right) + 3\left(\frac{1}{7}\right) + 4\left(\frac{1}{7}\right) + 5\left(\frac{1}{7}\right) + 6\left(\frac{1}{7}\right) + 7\left(\frac{1}{7}\right) = \frac{28}{7} \; or \; 4$$

And

$$E[X^2] = 1^2\left(\frac{1}{7}\right) + 2^2\left(\frac{1}{7}\right) + 3^2\left(\frac{1}{7}\right) + 4^2\left(\frac{1}{7}\right) + 5^2\left(\frac{1}{7}\right) + 6^2\left(\frac{1}{7}\right) + 7^2\left(\frac{1}{7}\right) = \frac{140}{7} \; or \; 20$$

$$Var[X] = E[X^2] - (E[X])^2 = 20 - 4^2 = 4$$

6. A fair coin is tossed 10 times, and the outcome of each toss is either heads or tails. What is the probability of getting exactly 5 heads?
    a. 0.1
    b. 0.25
    c. 0.5
    d. 0.18

Solution: (b)

The probability of getting exactly 5 heads in 10 coin flips is calculated using the binomial distribution using the formula

$$P(X = x) = {}^{n}C_{r}p^{r}(1-p)^{n-r}$$

where,

- n: n is the number of trials
- r is the number of successes
- p is the probability of success on a single trial

Plugging in the values, we get

$$P(X = 5) = {}^{10}C_{5}\left(\frac{1}{2}\right)^{5}\left(\frac{1}{2}\right)^{5} = 0.246.$$

7. Creutzfeldt-Jakob disease is a rare, degenerative, fatal brain disorder. It affects about one person in every one million per year worldwide. Your team has developed a diagnostic test for it. It turns out that if a subject is suffering from the Creutzfeldt-Jakob disease, there is a 99% chance that your test will detect it. However, it is likely that for every 1000 healthy subjects tested, the test will produce a positive response to one of them. When you give this test kit to an independent doctor for validation, the doctor diagnosed her first patient as suffering from the disease as per your diagnostic test kit that she used. What is the probability that the patient is actually suffering from Creutzfeldt-Jakob disease?
    a. $1.1273e^{-4}$
    b. $6.4804e^{-4}$
    c. $9.8902e^{-4}$
    d. $5.778e^{-4}$

Solution: (c)

Let the disease be denoted by 'CJ' and the diagnostic test kit results by 'DKT'. Given:

$$P(CJ = 1) = 10^{-6}$$

$$P(DKT = 1|CJ = 1) = 0.99$$

$$P(DKT = 1|CJ = 0) = 10^{-3}$$

Applying Bayes' Theorem,

$$P(CJ = 1|DKT = 1) = \frac{P(DKT = 1|CJ = 1)P(CJ = 1)}{P(DKT = 1)}$$

$$= \frac{P(DKT = 1|CJ = 1)P(CJ = 1)}{P(DKT = 1|CJ = 1)P^{\cdot}(CJ = 1) + P(DKT = 1|CJ = 0)P^{\cdot}(CJ = 0))}$$

$$= 9.8902e^{-4}$$

8. Which of the following is NOT an example of Field?
    a. Set of Real numbers
    b. Set of Complex numbers
    c. Set of Rational numbers
    d. Set of Integer numbers

Solution: (d)

R, C and Q form a field, while Z does not form a field [Reason: Multiplicative Inverse].

9. Which of the following statements is TRUE about the rank of a matrix?
    a. Only the rank of a Null Matrix is zero.
    b. The rank of a matrix is always more than or equal to its minimum of number of rows and columns
    c. The rank of a matrix generated as a result of the product of a column vector and a row vector is 2.
    d. A full-rank matrix has a rank equal to the number of rows in the matrix.

Solution: (a)

- Only the rank of a Null Matrix is zero.
- The rank of a matrix is always less than or equal to its minimum of number of rows and columns in the matrix.
- The rank of a matrix generated as a result of the product of a column vector and a row vector is 1.
- A full-rank matrix has rank equal to the minimum of the number of rows and columns in the matrix.

10. The maximum number of non-zero singular value in the SVD of matrix $A = \begin{bmatrix} 1 & -1 & 2 \\ -3 & 3 & -6 \end{bmatrix}$ is
    a. 0
    b. 1
    c. 2
    d. 3

Solution: (b)

The maximum number of non-zero singular values is equal to the rank of the matrix. Here, one row is the multiple of the other. Hence, rank is equal to 1.

11. What are the eigenvalues for the eigenvalue decomposition of a 3×3 identity matrix?
    a. 1, 1, 1
    b. 1, 0, 0
    c. 1, 1, 0
    d. 0, 0, 0

Solution: (a)

12. Using singular value decomposition (SVD), find the singular values of the following matrix.

$$\begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}$$

    a. 6 and $\sqrt{6}$
    b. $\sqrt{6}$ and 4
    c. 16 and 6
    d. 4 and 6

Solution: (b)

$$U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

13. Which of the following options do not correspond to any one row of the V^T in the singular value decomposition (SVD) of the following matrix?

$$\begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}$$

a. $\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$
b. $\begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$
c. $\begin{bmatrix} -\frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$
d. $\begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$

Solution: (d)

$$U\Sigma V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

14. Which of the options is NOT a fit for the given blank in the following sentence?

"A sigma field is a field that is closed under any countable set of _____ of the events."

a. Unions
b. Integer number
c. Intersections
d. Combinations

Solution: (b)

"A sigma field is a field that is closed under any countable set of unions, intersections and combinations of the events."

15. Which of the following is NOT a matrix property?

a. A trace matrix is a square matrix $\mathbf{T}$ such that trace$(\mathbf{T}) = \mathrm{I}$
b. A unitary matrix is a square matrix $\mathbf{U}$ such that $\mathbf{U}^{\mathrm{H}} = \mathbf{U}^{-1}$
c. A normal matrix is a matrix $\mathbf{N}$ such that $\mathbf{N}\mathbf{N}^{\mathbf{H}} = \mathbf{N}^{\mathbf{H}}\mathbf{N}$.
d. A Hermitian matrix is a matrix $\mathbf{A}$ such that $\mathbf{A}^{\mathbf{H}} = \mathbf{A}$

Solution: (a)

There is no such thing as 'Trace Matrix'.

16. [Multiple select questions (MSQ)] Which of the following is a valid way to create a numpy array in Python?
    a. my_arr = [1, 2, 3]
    b. my_arr = numpy.array(1, 2, 3)
    c. my_arr = numpy.array([1, 2, 3])
    d. my_arr = numpy.array((1, 2, 3))

Solution: (c), (d)

Note: We understand that the above question was given as an MCQ. Students are advised to focus on the correct answers for now. The marking scheme of this question will be announced later.

17. What is the output of the following Python code? Assume that the required libraries have been imported.

```
a = [1, 2, 3]
b = [2, 3, 4]
c = [sum(x) for x in zip(a, b)]
print(c)
```

    a. [1, 2, 3, 2, 3, 4]

    b. [3, 5, 7]

    c. [2, 4, 6]

    d. [1, 2, 3, 4]

Solution: (b)

The zip() function creates an iterator that aggregates elements from each of the iterables passed to it. The sum() function adds the corresponding elements from a and b together, and the list comprehension creates a new list with these sums.

18. Which of the following Python code snippets would create a dictionary where the keys are the characters in a sentence and the values are the frequency of each character in the sentence? Assume that the required libraries have been imported.

a.

```
sentence = "the quick brown fox jumps over the lazy dog"
word_freq = {}
for word in sentence.split():
    word_freq[word] = word_freq.get(word, 0) + 1
```

b.

```
sentence = "the quick brown fox jumps over the lazy dog"
word_freq = dict.fromkeys(sentence.split(), 0)
for word in sentence.split():
    word_freq[word] += 1
```

c.

```
sentence = "the quick brown fox jumps over the lazy dog"
word_freq = {}
for word in sentence:
    if word not in word_freq:
        word_freq[word] = 0

    else:
        word_freq[word] += 1
```

d.

```
sentence = "the quick brown fox jumps over the lazy dog"
word_freq = {}
for word in sentence:
    if word not in word_freq:
        word_freq[word] = sentence.count(word)
```

Solution: (d)

19. Which numpy function can be used to create an array of evenly spaced numbers over a specified interval?
    a. np.linspace()
    b. np.arange()
    c. np.random.rand()
    d. np.zeros()

Solution: (a)

20. What is the error in the following Python code snippet? Assume that the required libraries have been imported.

```
import numpy as np
x = np.array([1, 2, 3])
y = np.array([4, 5])
z = np.concatenate([x, y])
print(z)
```

a. The concatenate method should be called with a tuple of arrays.
b. The shapes of the input arrays to the concatenate method should match.
c. The concatenate method should be called with axis=0.
d. There is no error.

Solution: (d)