

1. Table 1 shows the Body Mass and Running Speed of different animals.

Table 1

| Animal | Body Mass (Kg) | Running Speed (Km/hr) |
|------------|----------------|-----------------------|
| Rhino | 1400 | 45 |
| Horse | 400 | 70 |
| Pronghorn | 50 | 100 |
| Giraffe | 1000 | 60 |
| wildebeest | 300 | 90 |
| Cheetah | 60 | 110 |

a) Identify the type of relation that exists between the body mass and the running speed of the animals. May roughly plot the values to check the relationship.

b) Analyze the data and train a suitable machine learning approach for predicting the running speed of animals.

c) Now, predict the running speed for the rabbit, if the body mass is 50 kg.

2. Consider the truth table for XOR gate given below and apply the activation function f as given for the variables X_1 , X_2 and Target/Output variable Z_1 . Assume the learning rate = 1.5, threshold (θ) = 1 and initial weights with bias = 0. Show how the perceptron model is trained on the data given for 3 epochs.

| X_1 | X_2 | Z_1 |
|-------|-------|-------|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

The activation function

$$f(y_{in}) = \begin{cases} 1 & \text{if } y_{in} \geq \theta \\ 0 & \text{if } y_{in} < \theta \end{cases}$$

3. Consider each of the following scenarios :

1. We collect a set of data on the top 500 firms in India. For each firm we record profit, number of employees, industry and the CEO salary. We are interested in understanding which factors affect CEO salary.
2. An Insurance company is considering launching a new product and wish to know whether it will be a success or a failure. The Insurance company collects data on 50 similar products that were previously launched. For each product they have recorded whether it was a success or failure, price charged for the product, market budget, competition price, and ten other variables.
3. We are interested in predicting the % change in the Tumor in relation to the weekly changes in different Medicines. Hence we collect weekly data for the complete year. For each week we record the % change in Tumor, the % change in the Medicine dosage, the % change in the Medicine 2 dosage, and the % change in the Medicine3 dosage.

a) Explain whether each scenario given is a classification or regression problem.

| | | |
|----|--|--|
| | | b) Indicate whether we are most interested in inference or prediction. Also, mention one ML model for the task and how the task is achieved. |
| 4. | | <p>You have a movie dataset. Dataset has various features like average_review_audience, number_of_reviews, release_date, boxOffice_success etc. collected from different sources for different movies along with movie_category. You are required to develop a machine learning model that predicts the movie_category from all other features. There are 5 categories to which a movie can belong to. You are required to train a classifier that predicts movies category, but the input data has non-linear relationship with the movie_category. Assume that your ML library has only a binary classifier that can work fine on the linearly separable data.</p> <p>a. How do you address this problem of non-linear relation? Explain.</p> <p>b. How can you use the binary classifier for predicting 5 classes? Explain.</p> |