Quiz 2

After the thorough experiments, scientists found that kinetic energy K can be calculated from the mass of the object m and its velocity v using the formula: $K = \frac{1}{2}mv^2$. Let's consider this as a model infer from the experiments in Machine Learning.

1. What are the features of this model?

Answer: m and v

2. What are the target attribute of this model?

Answer: K

3. What are the coefficients of the model?

Answer: 1/2

4. Is the model Linear? Non-linear? Exponential?

Answer: Non-linear or Polynomial of degree 2

Quiz 3

You are given 10 training data points: $(x_1, y_1), (x_2, y_2), ...(x_{10}, y_{10})$. Here y_i are target attributes and a function $f(x_i)$ is created as a model to predict the target attribute y_i. (Please use L2 norm for norm calculation.)

1. What is the error function(loss function) in predicting each data?

Answer: Loss function $(y_i - f(x_i))^2$

2. What is the cost function (total error) of all training data?

Answer:

$$\sqrt{\sum_{i=1}^{i=10} y_i - f(x_i)^2}$$

Note: it is ok if square root is not in the formula and using the squared error.

Quiz 4

1. Calculate the entropy of the probability vector [1/3, 2/3].

Answer:

$$\frac{1}{3}log_2(3) + \frac{2}{3}log_2(\frac{3}{2}) = 0.9183$$

2. After calculating the information gain, the value student got is -0.1 for using Hair color feature while -0.5 for using Height feature. Which feature should be used to split the data? Is there anything wrong with this decision?

Answer: There seems to be calculation error. Gain cannot be negative.

Quiz 5

Create multi-layer nerualnet to solve:

$$(x_1 \wedge \bar{x_2}) \vee (\bar{x_1} \wedge x_2)$$

Answer

t_1**:**

$$\begin{aligned} x_1 + (1 - x_2) &\geq 1.5 \\ -0.5x_0 + x_1 - x_2 &\geq 0 \end{aligned}$$

t_2**:**

$$\begin{aligned} (1 - x_1) + x_2 &\geq 1.5 \\ -0.5x_0 - x_1 + x_2 &\geq 0 \end{aligned}$$

$t_1 \vee t_2$**:**

$$\begin{aligned} t_1 + t_2 &\geq 0.5 \\ -0.5t_0 + t_1 + t_2 &\geq 0 \end{aligned}$$

We can draw the multi-layer neuralnet using calculated weights from above with 3 inputs x_0, x_1,x_2. All these nodes fully connected to hidden layer with nodes t_1 and t_2. Output layer has one node O and it is connected to t_0, t_1 and t_2. The weights will be according to the calculated values above and threshold function is applied to t_1, t_2 and O.

Quiz 6

You are given a perceptron (Fully connected layer in deep learning) with 1 hidden layer (2 layer perceptron). Input layer has 1000 nodes, hidden layer has 100 nodes and output layer has 10 nodes. Additional bias connections are also added as needed to achieve the additional flexibility.

Each batch has 32 examples. There are total 1024000 examples. All the weights will be updated after each batch.

1. Calculate the number of weights one need to update after each batch?

Answer: The number of total weights to update in first layer is 100,100 and the second layer is 1010. Therefore, total is 101,110.

2. How many times all the weights will be updated in each epoch?

Answer: $\frac{1024000}{32} = 3200$