KMeans [25]

Consider a data set:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 |
| X1 | 1.2 | 1.6 | 0.5 | 0.7 |
| X2 | 1.4 | 0.5 | 0.5 | 0.2 |

1. [5] Suppose the random initial cluster IDs are assigned as cl = [1, 1, 1, 0]. Find the cluster centroids.

A screenshot of a black background with white text

Description automatically generated

A black background with white text

Description automatically generated

1. [5] Calculate the distance of each point to each centroid.

A screenshot of a black box

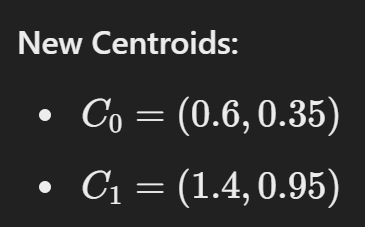
Description automatically generated

1. [5] According to the calculated distances between the point and the centroid, reassign the cluster IDs.

A black background with white text

Description automatically generated

1. [5] Calculate the new centroid.



1. [5] Calculate the within-cluster variances and the total within-cluster variance.

A black background with white text

Description automatically generated

Neural Network [25]

Let the weight matrices w1 be and w2 be

1. [3] Draw the network diagram.

x\_1 x\_2

| |

| |

(0.16) (0.02)

| |

v v

h\_1 ------ <-- Hidden Layer

^ ^

(0.10) (0.70)

| |

| |

h\_2 <-----

h\_1 ---> (0.05)

\

--> y (Output)

/

h\_2 --->(0.33)

1. [5] Perform the forward propagation using the sigmoid function as a non-linear activtion. Report the accuracy rate.

A screenshot of a computer

Description automatically generated

1. [10] Then, perform the backward propagation. Report the new updated w1 and w2 when the learning rate value of 0.1 is used.

A black background with white text

Description automatically generated

1. [5] Reperform the forward propagation and predict the class.

A screenshot of a black screen

Description automatically generated

1. [2] Is the neural network algorithm efficient? Explain why. If you are going to use other than a neural network, which classification model would you prefer?

Neural networks can be quite computationally intensive and often require substantial training data and careful tuning of hyperparameters to achieve good performance. For a small dataset or simpler classification tasks, this overhead may not be efficient.

If not using a neural network, a more traditional and often more efficient choice would be a model like Logistic Regression or a Support Vector Machine (SVM). These models are generally quicker to train, easier to interpret, and can perform well with fewer data points.

Graphical Model [25]

|  |  |  |  |
| --- | --- | --- | --- |
|  | X1 | X2 | Y |
| 1 | 1 | 1 | 1 |
| 2 | 1 | 0 | 1 |
| 3 | 0 | 1 | 1 |
| 4 | 0 | 0 | 0 |
| 5 | 1 | 0 | 1 |
| 6 | 1 | 1 | 1 |
| 7 | 0 | 0 | 0 |
| 8 | 1 | 0 | 0 |
| 9 | 1 | 1 | 1 |
| 10 | 0 | 1 | 1 |

Consider the data set.

1. [5] Draw a graphical representation of the data table if the joint probability of p(x1,x2,Y) = p(Y|x1, x2)p(x2|x1)p(x1).
2. [5] Construct probability tables of p(x1), p(x2|x1), and p(Y|x1,x2)
3. [3] Calculate the probability of p(Y=1).
4. [12] Calculate the probability of p(x1=1|Y=1).

Decision Tree [25]

Using the same data table above, classify Y.

1. [5] Calculate the entropy of Y, H(Y).
2. [10] Calculate the entropies, H(Y|x1) and H(Y|x2).
3. [5] Calculate the information gain of x1 and x2.
4. [5] Determine the best split and draw a tree diagram. What is the accuracy of classification?