Machine Learning

22AIE213

Assignment-2

Team 17:

Y Nishanth (AIE22165)

Meghraj G K (AIE22129)

Akshat Kuttan (AIE22104)

Pseudocode:

function euclidean\_distance(vec1, vec2):

return sqrt(sum((vec1 - vec2)^2))

function manhattan\_distance(vec1, vec2):

return sum(abs(vec1 - vec2))

function knn\_classifier(training\_data, test\_instance, k):

distances = [euclidean\_distance(test\_instance, x) for x in training\_data]

sorted\_indices = argsort(distances)

k\_nearest\_neighbors = sorted\_indices[:k]

return k\_nearest\_neighbors

function label\_encoding(data):

unique\_labels = unique(data)

label\_dict = {label: i for i, label in enumerate(unique\_labels)}

encoded\_data = [label\_dict[label] for label in data]

return encoded\_data

function one\_hot\_encoding(data):

unique\_labels = unique(data)

encoded\_data = zeros((len(data), len(unique\_labels)))

for i, label in enumerate(data):

encoded\_data[i, unique\_labels.index(label)] = 1

return encoded\_data

function unit\_tests():

# (Tests for Euclidean distance, Manhattan distance, k-NN classifier, label encoding, and one-hot encoding)

print("All unit tests passed!")

# Run unit tests

unit\_tests()

Description:

1. Data Loading and Processing:

* Loads data from an ARFF file using scipy.
* Converts the ARFF data to a numpy array.
* Splits the data into training and test sets.

1. Distance Calculations:

* Defines functions for Euclidean and Manhattan distance calculations.

1. k-NN Classifier:

* Implements a k-NN classifier using both Euclidean and Manhattan distances.
* Predicts labels for the test set using the implemented classifier.

1. Evaluation:

* Evaluates the performance of the k-NN classifier using accuracy, confusion matrix, and classification report for both Euclidean and Manhattan distances.

1. Label Encoding and One-Hot Encoding:

* Performs label encoding on the target variable.
* Performs one-hot encoding on the target variable.

1. Results:

* Prints the evaluation metrics for both distance metrics.
* Prints predicted classes for each test instance.