Data Science with Python

Day 04 - Activity

01. Use the given 'Pima Indians Diabetes' to answer the below questions
    1. Load the Pima Indians Diabetes dataset into scikit-learn.
    2. Perform necessary data preprocessing, such as handling missing values, normalizing or scaling the data, and splitting the dataset into training and testing sets.
    3. How can you implement a perceptron neural network with one hidden layer in scikit-learn, and train it on the Pima Indians Diabetes dataset using backpropagation with a specified number of epochs?
    4. Evaluate the performance of the model on the testing set using accuracy as the metric.
    5. Report the confusion matrix, precision, recall, and F1-score for the model.

02. Using the "***winequality-red.csv"*** dataset answer the following questions.
    1. Load the "winequality-red.csv" dataset and Separate the dataset into input (X) and output (y) variables.
    2. Preprocess the data using appropriate methodologies.
    3. Split the data into training and testing sets.
    4. Build and train the model using the Keras library
        i. Create a Sequential model, compile it with appropriate loss, optimizer, and metrics, and fit it to the training data.
    5. Evaluate the model by computing the classification report.
    6. Plot the training and validation loss curves using the appropriate plotting method.
    7. Calculate the precision for each class, and plot it to observe precision scores for different classes in the classification task.