Classification Assignment – CKD

Results:

1. Based on the data shared this falls under Domain – Machine Learning – Classification
2. Total – 399 rows and 25 columns of data present in CKD.csv
3. Converting the columns – sg , rbc, pc, pcc, ba,pe from string to numeric
4. Predicted the results using the below Classification Algorithms
    a. **Logistic Regression Algorithm – 98% Accuracy**
    b. **Naive Bayes – 94% Accuracy**
    c. **KNN Algorithm- 94% Accuracy**

Choosing Logistic Regression Algorithm as the best one for this Data set for CKD.csv , as this provides a prediction with 98% Accuracy

Prediction Results with Logistic Regression Algorithm

```
print("The confusion Matrix:\n",cm)
```

```
The confusion Matrix:
 [[51  0]
 [ 2 80]]
```

```
print("The report:\n",clf_report)
```

```
The report:
              precision    recall  f1-score   support

       False       0.96      1.00      0.98        51
        True       1.00      0.98      0.99        82

    accuracy                           0.98       133
   macro avg       0.98      0.99      0.98       133
weighted avg       0.99      0.98      0.99       133
```

- ✓ Overall Accuracy is 98% is predicted using **Logistic Regression Algorithm**

Prediction Results with Naive Bayes

```
print("The confusion Matrix:\n",cm)
```

```
The confusion Matrix:
 [[51  0]
 [ 8 74]]
```

```
print("The report:\n",clf_report)
```

```
The report:
              precision    recall  f1-score   support

       False       0.86      1.00      0.93        51
        True       1.00      0.90      0.95        82

    accuracy                           0.94       133
   macro avg       0.93      0.95      0.94       133
weighted avg       0.95      0.94      0.94       133
```

- ✓ Overall Accuracy is 94% is predicted using BernoulliNB
- ✓ Out of all Classification, Truly Classified displays as 1.00
- ✓ Out of all Classification, Correctly classified as true is 90%

Prediction Results with KNN Algorithm

```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
cm
```

```
array([[51,  0],
       [ 8, 74]])
```

```
from sklearn.metrics import classification_report
clf_report = classification_report(y_test, y_pred)
print(clf_report)
```

```
              precision    recall  f1-score   support

       False       0.86      1.00      0.93        51
        True       1.00      0.90      0.95        82

    accuracy                           0.94       133
   macro avg       0.93      0.95      0.94       133
weighted avg       0.95      0.94      0.94       133
```

- ✓ Overall Accuracy is 94% is predicted using **KNN Algorithm**