Dataset is skewed i.e the number of rows with label 2 are almost double the number of rows

with label 4. Using this dataset for knn,(k=4) we get an accuracy of 98.5632183908046.

After removing many rows with label 2 to reduce the skewness, so that the number of rows

having label 2 and label 4 are equal, the accuracy obtained (for k=4) is 98.68421052631578

```
98.68421052631578
divya@divya-HP-Notebook:~/Desktop/ml_project$ python3 knn.py
229 229
[1294261, 4, 10, 8, 5, 4, 1, 10, 1, 1, 4]
98.68421052631578
divya@divya-HP-Notebook:~/Desktop/ml_project$ python3 knn.py
458 241
[1071084, 3, 3, 2, 2, 3, 1, 1, 2, 3, 2]
98.5632183908046
divya@divya-HP-Notebook:~/Desktop/ml_project$ 
```