

## Prior Knowledge

Team ID	PNT2022TMID13268
Project Name	Smart Lender Applicant Credibility Prediction for Loan Approval

Prior-Knowledge :

Supervised Learning:

[https://www.youtube.com/watch?v=QeKshry8pWQ&t=24479s&ab\\_channel=GreatLearning](https://www.youtube.com/watch?v=QeKshry8pWQ&t=24479s&ab_channel=GreatLearning)

The screenshot shows a Jupyter Notebook interface with the title 'Naive+Bayesian+Pima+Diabetes+'. The code in the notebook is as follows:

```
In [8]: # invoking the NB Gaussian function to create the model
# fitting the model in the training data set
model = GaussianNB()
model.fit(X_train, y_train)

Out[8]: GaussianNB(priors=None)

In [9]: # make predictions
expected = y_test
predicted = model.predict(X_test)
# summarize the fit of the model
print(metrics.classification_report(expected, predicted))
print(metrics.confusion_matrix(expected, predicted))
```

The output of the code is a classification report and a confusion matrix:

	precision	recall	f1-score	support
0.0	0.80	0.79	0.79	147
1.0	0.64	0.65	0.65	84
avg / total	0.74	0.74	0.74	231

UnSupervised Learning:

[https://www.youtube.com/watch?v=D6gtZrsYi6c&t=1s&ab\\_channel=Simplilearn](https://www.youtube.com/watch?v=D6gtZrsYi6c&t=1s&ab_channel=Simplilearn)

The screenshot shows a Jupyter Notebook interface with the title 'Untitled2'. The code in the notebook is as follows:

```
kmeans_4.cluster_centers_
print(dict(zip(unique, counts)))

[[ 49.99263253  5.20447169]
 [180.34311782 10.52011494]
 [177.83509615 70.28846154]
 [ 50.40482436 32.36533958]]
{0: 2773, 1: 696, 2: 104, 3: 427}

In [77]: df_analyze['cluster']=kmeans_4.labels_
sns.set_style('whitegrid')
sns.lmplot('mean_dist_day', 'mean_over_speed_perc', data=df_analyze,
          hue='cluster', palette='coolwarm', size=6, aspect=1, fit_reg=False)

Out[77]: <seaborn.axisgrid.FacetGrid at 0x195865c0>
```

The output of the code is a scatter plot showing the relationship between 'mean\_dist\_day' and 'mean\_over\_speed\_perc' for different clusters. The plot is a white grid with a coolwarm color palette. The x-axis is labeled 'mean\_dist\_day' and the y-axis is labeled 'mean\_over\_speed\_perc'. The plot shows four distinct clusters of data points, each represented by a different color.

[https://www.youtube.com/watch?v=aWAnNHXIKww&t=2s&ab\\_channel=KrishNaik](https://www.youtube.com/watch?v=aWAnNHXIKww&t=2s&ab_channel=KrishNaik)



[https://www.youtube.com/watch?v=Ij4I\\_CvBnt0&t=2044s&ab\\_channel=edureka%21](https://www.youtube.com/watch?v=Ij4I_CvBnt0&t=2044s&ab_channel=edureka%21)

