Lab4

1. State the difference and similarity of GD and SGD, give each two situitions, one should use GD, the other should use SGD. Explain why. (10)

Write your answer here

In [2]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn import svm
from sklearn.ensemble import RandomForestClassifier
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
```

In [9]:
```
STAB = pd.read_csv("smart_grid_stability_augmented.csv")
STAB = STAB.drop('stab', axis = 1)
STAB.head(5)
```

Out[9]:

	tau1	tau2	tau3	tau4	p1	p2	p3	p4	g1
0	2.959060	3.079885	8.381025	9.780754	3.763085	-0.782604	-1.257395	-1.723086	0.650456
1	9.304097	4.902524	3.047541	1.369357	5.067812	-1.940058	-1.872742	-1.255012	0.413441
2	8.971707	8.848428	3.046479	1.214518	3.405158	-1.207456	-1.277210	-0.920492	0.163041
3	0.716415	7.669600	4.486641	2.340563	3.963791	-1.027473	-1.938944	-0.997374	0.446209
4	3.134112	7.608772	4.943759	9.857573	3.525811	-1.125531	-1.845975	-0.554305	0.797110

2. Seperate the STAB dataset to train and test, stabf is y, the others are x, make test dataset 30% of the total dataset (10)

In [8]:
```
#Write your code here.
```

3. Apply Random Forest to this dataset, predict stabf, show the accuracy. (10)

In [10]:
```
#Write your code here.
```

4. Apply SVM to this dataset, predict stabf, show the accuracy. (10)

In [11]:
```
#Write your code here.
```

In [14]:
```
wine = pd.read_csv("Wine.csv")
wine.head(5)
```

Out[14]:

	Alcohol	Malic_Acid	Ash	Ash_Alcanity	Magnesium	Total_Phenols	Flavanoids	Nonflavanoi
0	14.23	1.71	2.43	15.6	127	2.80	3.06	
1	13.20	1.78	2.14	11.2	100	2.65	2.76	
2	13.16	2.36	2.67	18.6	101	2.80	3.24	
3	14.37	1.95	2.50	16.8	113	3.85	3.49	
4	13.24	2.59	2.87	21.0	118	2.80	2.69	

In [15]:
```
X=wine.drop("Customer_Segment",axis=1).values
y=wine["Customer_Segment"].values
```

5. Standartize the variables X (10)

In [16]:
```
#Write your code here.
```

6. Use PCA to reduce standartized X to 2 dimensions (10)

In [17]:
```
#Write your code here.
```

7. Use LDA to reduce standartized X to 2 dimensions (10)

In [18]:
```
#Write your code here.
```

8. Apply logistic regression to standartized X, PCA X and LDA X to predict Y, compare the result. Remember to split train and test sets. Set test sets to 20% of whole data. (10)

In [19]:
```
#write your code here
```

9. Apply feature select to the STAB data.(10)

In [20]:
```
#write your code here
```

10. Apply SVM to the edited STAB dataset, compare the result with the previous one.(Remember to split the dataset)(10)

In [21]:
```
#write your code here
```