Task-B: Large variance features and their effect on linear classifier models

In [115…

```
# necessary libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LogisticRegression
from sklearn.linear_model import SGDClassifier
from sklearn.preprocessing import StandardScaler
from sklearn.preprocessing import MinMaxScaler
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot
init_notebook_mode(connected=True)
```

In [116…

```
# read the dataset
data = pd.read_csv('task_b.csv')
data=data.iloc[:,1:]
```

In [117…

```
# display the data
print(data.head())
print(data.shape)
```

```
            f1            f2        f3    y
0  -195.871045 -14843.084171  5.532140  1.0
1 -1217.183964  -4068.124621  4.416082  1.0
2     9.138451   4413.412028  0.425317  0.0
3   363.824242  15474.760647  1.094119  0.0
4  -768.812047  -7963.932192  1.870536  0.0
(200, 4)
```

In [118…

```
data.corr()['y'] # correlation of each feature wrt target column 'y'
```

Out[118…

```
f1    0.067172
f2   -0.017944
f3    0.839060
```

```
y     1.000000
Name: y, dtype: float64
```

In [119…

```
data.std() # standard deviation of each column
```

Out[119…

```
f1      488.195035
f2    10403.417325
f3        2.926662
y         0.501255
dtype: float64
```

In [120…

```
# obtain X & y data from the dataframe
X= data[['f1','f2','f3']].values
Y= data['y'].values
print(X.shape)
print(Y.shape)
```

```
(200, 3)
(200,)
```

What if our features are with different variance

*** As part of this task you will observe how linear models work in case of data having feautres with different variance**
*** from the output of the above cells you can observe that var(F2)>>var(F1)>>Var(F3)**

> **Task1**:
    1. Apply Logistic regression(SGDClassifier with logloss) on 'data' and check the feature importance
    2. Apply SVM(SGDClassifier with hinge) on 'data' and check the feature importance

> **Task2**:
    1. Apply Logistic regression(SGDClassifier with logloss) on 'data' after standardization
       i.e standardization(data, column wise): (column-mean(column))/std(column) and check the feature importance
    2. Apply SVM(SGDClassifier with hinge) on 'data' after standardization
       i.e standardization(data, column wise): (column-mean(column))/std(column) and check the

feature importance

Make sure you write the observations for each task, why a particular feature got more importance than others

TASK 1:

PART (A)

In [121…

```
# Apply Logistic regression on "data"
from sklearn.linear_model import SGDClassifier

# initialize the classifier
clf1 = SGDClassifier(loss = 'log',random_state = 60,eta0 = 0.001,
                     alpha = 0.001,learning_rate = 'constant',verbose = 2)
# fit the classifier
clf1.fit(X,Y)
```

```
-- Epoch 1
Norm: 17.49, NNZs: 3, Bias: -0.005500, T: 200, Avg. loss: 27246.549730
Total training time: 0.00 seconds.
-- Epoch 2
Norm: 23.01, NNZs: 3, Bias: 0.003500, T: 400, Avg. loss: 28182.684999
Total training time: 0.00 seconds.
-- Epoch 3
Norm: 10.30, NNZs: 3, Bias: 0.002500, T: 600, Avg. loss: 30746.160992
Total training time: 0.00 seconds.
-- Epoch 4
Norm: 5.46, NNZs: 3, Bias: 0.003500, T: 800, Avg. loss: 32577.348109
Total training time: 0.00 seconds.
-- Epoch 5
Norm: 14.86, NNZs: 3, Bias: 0.010500, T: 1000, Avg. loss: 29981.061135
Total training time: 0.00 seconds.
-- Epoch 6
Norm: 10.46, NNZs: 3, Bias: 0.014500, T: 1200, Avg. loss: 29193.908784
Total training time: 0.00 seconds.
Convergence after 6 epochs took 0.00 seconds
```

Out[121…

```
SGDClassifier(alpha=0.001, eta0=0.001, learning_rate='constant', loss='log',
              random_state=60, verbose=2)
```

In [122…
```
# obtain weights for the features
coef = clf1.coef_
print(coef) #-->2nd feature seems to be an important feature
```

```
[[ 2.4120961  10.05025104  1.63449413]]
```

OBSERVATION :

We can observe that the 2nd feature got the highest weight because , the logistic regression model is giving more importance to the 2nd feature as its having larger variance and its wide range of values are contributing more towards the fitted hyperplane.

PART (B)

In [123…
```
# Apply SVM on "data"
from sklearn.linear_model import SGDClassifier
clf2 = SGDClassifier(loss = 'hinge',random_state = 60,eta0 = 0.001,
                     alpha = 0.001,learning_rate = 'constant',verbose = 2)
clf2.fit(X,Y)
```

```
-- Epoch 1
Norm: 17.48, NNZs: 3, Bias: -0.006000, T: 200, Avg. loss: 26465.570447
Total training time: 0.00 seconds.
-- Epoch 2
Norm: 3.34, NNZs: 3, Bias: 0.005000, T: 400, Avg. loss: 27815.827754
Total training time: 0.00 seconds.
-- Epoch 3
Norm: 10.45, NNZs: 3, Bias: 0.013000, T: 600, Avg. loss: 28514.398449
Total training time: 0.00 seconds.
-- Epoch 4
Norm: 7.08, NNZs: 3, Bias: 0.015000, T: 800, Avg. loss: 32096.308856
Total training time: 0.00 seconds.
-- Epoch 5
Norm: 14.09, NNZs: 3, Bias: 0.019000, T: 1000, Avg. loss: 29461.537481
Total training time: 0.00 seconds.
-- Epoch 6
Norm: 12.66, NNZs: 3, Bias: 0.021000, T: 1200, Avg. loss: 27662.977856
Total training time: 0.00 seconds.
Convergence after 6 epochs took 0.00 seconds
```

Out[123…
```
SGDClassifier(alpha=0.001, eta0=0.001, learning_rate='constant',
              random_state=60, verbose=2)
```

In [124…
```
coef = clf2.coef_
print(coef) #--> 2nd feature seems to be an important feature
```

```
[[ 2.42301695 12.32208584  1.61741947]]
```

OBSERVATION :

We can observe that the 2nd feature got the highest weight and compared to logistic regression the svm model is heavily impacted by the 2nd feature and this happens because of large range of values in the 2nd feature which inherently dominate the loss function of svm.

TASK 2:

PART (A)

In [125…
```
# LET US FIRST STANDARDIZE THE DATA:
scaled_X = StandardScaler().fit_transform(X)
```

In [126…
```
# Apply Logistic regression on data after standardization:

clf1 = SGDClassifier(loss = 'log',random_state = 60,eta0 = 0.001,
                     alpha = 0.001,learning_rate = 'constant',verbose = 2)
# fit the model
clf1.fit(scaled_X,Y)
```

```
-- Epoch 1
Norm: 0.08, NNZs: 3, Bias: -0.000100, T: 200, Avg. loss: 0.676485
Total training time: 0.00 seconds.
-- Epoch 2
Norm: 0.16, NNZs: 3, Bias: -0.000042, T: 400, Avg. loss: 0.644393
Total training time: 0.00 seconds.
-- Epoch 3
Norm: 0.23, NNZs: 3, Bias: -0.000026, T: 600, Avg. loss: 0.615407
Total training time: 0.00 seconds.
-- Epoch 4
Norm: 0.30, NNZs: 3, Bias: -0.000126, T: 800, Avg. loss: 0.589189
Total training time: 0.00 seconds.
-- Epoch 5
Norm: 0.37, NNZs: 3, Bias: -0.000057, T: 1000, Avg. loss: 0.565412
Total training time: 0.00 seconds.
-- Epoch 6
```

```
Norm: 0.44, NNZs: 3, Bias: -0.000082, T: 1200, Avg. loss: 0.543824
Total training time: 0.00 seconds.
-- Epoch 7
Norm: 0.50, NNZs: 3, Bias: -0.000085, T: 1400, Avg. loss: 0.524182
Total training time: 0.00 seconds.
-- Epoch 8
Norm: 0.56, NNZs: 3, Bias: -0.000042, T: 1600, Avg. loss: 0.506270
Total training time: 0.00 seconds.
-- Epoch 9
Norm: 0.61, NNZs: 3, Bias: -0.000054, T: 1800, Avg. loss: 0.489894
Total training time: 0.01 seconds.
-- Epoch 10
Norm: 0.67, NNZs: 3, Bias: -0.000121, T: 2000, Avg. loss: 0.474882
Total training time: 0.01 seconds.
-- Epoch 11
Norm: 0.72, NNZs: 3, Bias: -0.000255, T: 2200, Avg. loss: 0.461084
Total training time: 0.01 seconds.
-- Epoch 12
Norm: 0.77, NNZs: 3, Bias: -0.000327, T: 2400, Avg. loss: 0.448373
Total training time: 0.01 seconds.
-- Epoch 13
Norm: 0.81, NNZs: 3, Bias: -0.000423, T: 2600, Avg. loss: 0.436634
Total training time: 0.02 seconds.
-- Epoch 14
Norm: 0.86, NNZs: 3, Bias: -0.000592, T: 2800, Avg. loss: 0.425766
Total training time: 0.02 seconds.
-- Epoch 15
Norm: 0.90, NNZs: 3, Bias: -0.000620, T: 3000, Avg. loss: 0.415682
Total training time: 0.02 seconds.
-- Epoch 16
Norm: 0.95, NNZs: 3, Bias: -0.000754, T: 3200, Avg. loss: 0.406311
Total training time: 0.02 seconds.
-- Epoch 17
Norm: 0.99, NNZs: 3, Bias: -0.000895, T: 3400, Avg. loss: 0.397578
Total training time: 0.02 seconds.
-- Epoch 18
Norm: 1.03, NNZs: 3, Bias: -0.001102, T: 3600, Avg. loss: 0.389422
Total training time: 0.02 seconds.
-- Epoch 19
Norm: 1.06, NNZs: 3, Bias: -0.001294, T: 3800, Avg. loss: 0.381786
Total training time: 0.02 seconds.
-- Epoch 20
Norm: 1.10, NNZs: 3, Bias: -0.001467, T: 4000, Avg. loss: 0.374627
Total training time: 0.02 seconds.
-- Epoch 21
```

```
Norm: 1.14, NNZs: 3, Bias: -0.001564, T: 4200, Avg. loss: 0.367903
Total training time: 0.02 seconds.
-- Epoch 22
Norm: 1.17, NNZs: 3, Bias: -0.001738, T: 4400, Avg. loss: 0.361578
Total training time: 0.02 seconds.
-- Epoch 23
Norm: 1.21, NNZs: 3, Bias: -0.002007, T: 4600, Avg. loss: 0.355613
Total training time: 0.02 seconds.
-- Epoch 24
Norm: 1.24, NNZs: 3, Bias: -0.002242, T: 4800, Avg. loss: 0.349984
Total training time: 0.02 seconds.
-- Epoch 25
Norm: 1.27, NNZs: 3, Bias: -0.002567, T: 5000, Avg. loss: 0.344660
Total training time: 0.02 seconds.
-- Epoch 26
Norm: 1.30, NNZs: 3, Bias: -0.002802, T: 5200, Avg. loss: 0.339620
Total training time: 0.03 seconds.
-- Epoch 27
Norm: 1.33, NNZs: 3, Bias: -0.003018, T: 5400, Avg. loss: 0.334841
Total training time: 0.03 seconds.
-- Epoch 28
Norm: 1.36, NNZs: 3, Bias: -0.003277, T: 5600, Avg. loss: 0.330302
Total training time: 0.03 seconds.
-- Epoch 29
Norm: 1.39, NNZs: 3, Bias: -0.003568, T: 5800, Avg. loss: 0.325990
Total training time: 0.03 seconds.
-- Epoch 30
Norm: 1.42, NNZs: 3, Bias: -0.003823, T: 6000, Avg. loss: 0.321884
Total training time: 0.03 seconds.
-- Epoch 31
Norm: 1.45, NNZs: 3, Bias: -0.004058, T: 6200, Avg. loss: 0.317969
Total training time: 0.03 seconds.
-- Epoch 32
Norm: 1.47, NNZs: 3, Bias: -0.004371, T: 6400, Avg. loss: 0.314234
Total training time: 0.03 seconds.
-- Epoch 33
Norm: 1.50, NNZs: 3, Bias: -0.004732, T: 6600, Avg. loss: 0.310666
Total training time: 0.03 seconds.
-- Epoch 34
Norm: 1.52, NNZs: 3, Bias: -0.005074, T: 6800, Avg. loss: 0.307255
Total training time: 0.03 seconds.
-- Epoch 35
Norm: 1.55, NNZs: 3, Bias: -0.005484, T: 7000, Avg. loss: 0.303987
Total training time: 0.03 seconds.
-- Epoch 36
```

```
Norm: 1.57, NNZs: 3, Bias: -0.005816, T: 7200, Avg. loss: 0.300859
Total training time: 0.04 seconds.
-- Epoch 37
Norm: 1.60, NNZs: 3, Bias: -0.006088, T: 7400, Avg. loss: 0.297860
Total training time: 0.04 seconds.
-- Epoch 38
Norm: 1.62, NNZs: 3, Bias: -0.006477, T: 7600, Avg. loss: 0.294979
Total training time: 0.04 seconds.
-- Epoch 39
Norm: 1.64, NNZs: 3, Bias: -0.006769, T: 7800, Avg. loss: 0.292213
Total training time: 0.04 seconds.
-- Epoch 40
Norm: 1.67, NNZs: 3, Bias: -0.007137, T: 8000, Avg. loss: 0.289554
Total training time: 0.04 seconds.
-- Epoch 41
Norm: 1.69, NNZs: 3, Bias: -0.007481, T: 8200, Avg. loss: 0.286994
Total training time: 0.04 seconds.
-- Epoch 42
Norm: 1.71, NNZs: 3, Bias: -0.007833, T: 8400, Avg. loss: 0.284531
Total training time: 0.04 seconds.
-- Epoch 43
Norm: 1.73, NNZs: 3, Bias: -0.008218, T: 8600, Avg. loss: 0.282156
Total training time: 0.04 seconds.
-- Epoch 44
Norm: 1.75, NNZs: 3, Bias: -0.008590, T: 8800, Avg. loss: 0.279865
Total training time: 0.04 seconds.
-- Epoch 45
Norm: 1.77, NNZs: 3, Bias: -0.009006, T: 9000, Avg. loss: 0.277654
Total training time: 0.04 seconds.
-- Epoch 46
Norm: 1.79, NNZs: 3, Bias: -0.009443, T: 9200, Avg. loss: 0.275519
Total training time: 0.04 seconds.
-- Epoch 47
Norm: 1.81, NNZs: 3, Bias: -0.009856, T: 9400, Avg. loss: 0.273457
Total training time: 0.04 seconds.
-- Epoch 48
Norm: 1.83, NNZs: 3, Bias: -0.010252, T: 9600, Avg. loss: 0.271462
Total training time: 0.04 seconds.
-- Epoch 49
Norm: 1.85, NNZs: 3, Bias: -0.010578, T: 9800, Avg. loss: 0.269532
Total training time: 0.04 seconds.
-- Epoch 50
Norm: 1.87, NNZs: 3, Bias: -0.010980, T: 10000, Avg. loss: 0.267665
Total training time: 0.04 seconds.
-- Epoch 51
```

```
Norm: 1.89, NNZs: 3, Bias: -0.011333, T: 10200, Avg. loss: 0.265856
Total training time: 0.05 seconds.
-- Epoch 52
Norm: 1.91, NNZs: 3, Bias: -0.011764, T: 10400, Avg. loss: 0.264105
Total training time: 0.05 seconds.
-- Epoch 53
Norm: 1.93, NNZs: 3, Bias: -0.012162, T: 10600, Avg. loss: 0.262405
Total training time: 0.05 seconds.
-- Epoch 54
Norm: 1.94, NNZs: 3, Bias: -0.012560, T: 10800, Avg. loss: 0.260755
Total training time: 0.05 seconds.
-- Epoch 55
Norm: 1.96, NNZs: 3, Bias: -0.012979, T: 11000, Avg. loss: 0.259156
Total training time: 0.05 seconds.
-- Epoch 56
Norm: 1.98, NNZs: 3, Bias: -0.013376, T: 11200, Avg. loss: 0.257603
Total training time: 0.05 seconds.
-- Epoch 57
Norm: 2.00, NNZs: 3, Bias: -0.013787, T: 11400, Avg. loss: 0.256094
Total training time: 0.05 seconds.
-- Epoch 58
Norm: 2.01, NNZs: 3, Bias: -0.014232, T: 11600, Avg. loss: 0.254627
Total training time: 0.05 seconds.
-- Epoch 59
Norm: 2.03, NNZs: 3, Bias: -0.014689, T: 11800, Avg. loss: 0.253201
Total training time: 0.05 seconds.
-- Epoch 60
Norm: 2.05, NNZs: 3, Bias: -0.015057, T: 12000, Avg. loss: 0.251813
Total training time: 0.05 seconds.
-- Epoch 61
Norm: 2.06, NNZs: 3, Bias: -0.015415, T: 12200, Avg. loss: 0.250463
Total training time: 0.05 seconds.
-- Epoch 62
Norm: 2.08, NNZs: 3, Bias: -0.015800, T: 12400, Avg. loss: 0.249149
Total training time: 0.05 seconds.
-- Epoch 63
Norm: 2.09, NNZs: 3, Bias: -0.016220, T: 12600, Avg. loss: 0.247869
Total training time: 0.05 seconds.
-- Epoch 64
Norm: 2.11, NNZs: 3, Bias: -0.016638, T: 12800, Avg. loss: 0.246622
Total training time: 0.05 seconds.
-- Epoch 65
Norm: 2.12, NNZs: 3, Bias: -0.017063, T: 13000, Avg. loss: 0.245407
Total training time: 0.05 seconds.
-- Epoch 66
```

```
Norm: 2.14, NNZs: 3, Bias: -0.017517, T: 13200, Avg. loss: 0.244222
Total training time: 0.05 seconds.
-- Epoch 67
Norm: 2.15, NNZs: 3, Bias: -0.017903, T: 13400, Avg. loss: 0.243067
Total training time: 0.05 seconds.
-- Epoch 68
Norm: 2.17, NNZs: 3, Bias: -0.018302, T: 13600, Avg. loss: 0.241938
Total training time: 0.06 seconds.
-- Epoch 69
Norm: 2.18, NNZs: 3, Bias: -0.018714, T: 13800, Avg. loss: 0.240838
Total training time: 0.06 seconds.
-- Epoch 70
Norm: 2.20, NNZs: 3, Bias: -0.019110, T: 14000, Avg. loss: 0.239764
Total training time: 0.06 seconds.
-- Epoch 71
Norm: 2.21, NNZs: 3, Bias: -0.019491, T: 14200, Avg. loss: 0.238715
Total training time: 0.06 seconds.
-- Epoch 72
Norm: 2.23, NNZs: 3, Bias: -0.019883, T: 14400, Avg. loss: 0.237689
Total training time: 0.06 seconds.
-- Epoch 73
Norm: 2.24, NNZs: 3, Bias: -0.020266, T: 14600, Avg. loss: 0.236689
Total training time: 0.06 seconds.
-- Epoch 74
Norm: 2.25, NNZs: 3, Bias: -0.020671, T: 14800, Avg. loss: 0.235710
Total training time: 0.06 seconds.
-- Epoch 75
Norm: 2.27, NNZs: 3, Bias: -0.021062, T: 15000, Avg. loss: 0.234753
Total training time: 0.06 seconds.
-- Epoch 76
Norm: 2.28, NNZs: 3, Bias: -0.021447, T: 15200, Avg. loss: 0.233818
Total training time: 0.06 seconds.
-- Epoch 77
Norm: 2.29, NNZs: 3, Bias: -0.021851, T: 15400, Avg. loss: 0.232902
Total training time: 0.06 seconds.
Convergence after 77 epochs took 0.06 seconds
```

Out[126…

```
SGDClassifier(alpha=0.001, eta0=0.001, learning_rate='constant', loss='log',
              random_state=60, verbose=2)
```

In [127…

```
# obtain weights of features
coef = clf1.coef_
print(coef) #--> Now f3 is useful in classifying the pts for which weight is large
```

```
[[-0.04112093  0.03238967  2.292158  ]]
```

OBSERVATION :

Here, we can infer that the 3rd feature is the useful feature in classifying the data points and this result is possible only because we gave equal importance by standardizing the features before modelling the data.

PART (B)

In [128…

```
# Apply SVM on "data"

clf2 = SGDClassifier(loss = 'hinge',random_state = 60,eta0 = 0.001,
                     alpha = 0.001,learning_rate = 'constant',verbose = 2)
clf2.fit(scaled_X,Y)
```

```
-- Epoch 1
Norm: 0.17, NNZs: 3, Bias: 0.000000, T: 200, Avg. loss: 0.931119
Total training time: 0.00 seconds.
-- Epoch 2
Norm: 0.34, NNZs: 3, Bias: 0.000000, T: 400, Avg. loss: 0.789376
Total training time: 0.00 seconds.
-- Epoch 3
Norm: 0.51, NNZs: 3, Bias: 0.000000, T: 600, Avg. loss: 0.647661
Total training time: 0.00 seconds.
-- Epoch 4
Norm: 0.66, NNZs: 3, Bias: -0.004000, T: 800, Avg. loss: 0.511441
Total training time: 0.00 seconds.
-- Epoch 5
Norm: 0.76, NNZs: 3, Bias: -0.001000, T: 1000, Avg. loss: 0.428559
Total training time: 0.00 seconds.
-- Epoch 6
Norm: 0.84, NNZs: 3, Bias: 0.007000, T: 1200, Avg. loss: 0.384094
Total training time: 0.00 seconds.
-- Epoch 7
Norm: 0.91, NNZs: 3, Bias: 0.011000, T: 1400, Avg. loss: 0.353919
Total training time: 0.00 seconds.
-- Epoch 8
Norm: 0.97, NNZs: 3, Bias: 0.011000, T: 1600, Avg. loss: 0.333159
Total training time: 0.00 seconds.
-- Epoch 9
Norm: 1.03, NNZs: 3, Bias: 0.011000, T: 1800, Avg. loss: 0.317311
Total training time: 0.00 seconds.
```

```
-- Epoch 10
Norm: 1.07, NNZs: 3, Bias: 0.013000, T: 2000, Avg. loss: 0.305072
Total training time: 0.00 seconds.
-- Epoch 11
Norm: 1.11, NNZs: 3, Bias: 0.015000, T: 2200, Avg. loss: 0.295650
Total training time: 0.00 seconds.
-- Epoch 12
Norm: 1.15, NNZs: 3, Bias: 0.016000, T: 2400, Avg. loss: 0.287621
Total training time: 0.01 seconds.
-- Epoch 13
Norm: 1.19, NNZs: 3, Bias: 0.016000, T: 2600, Avg. loss: 0.280065
Total training time: 0.01 seconds.
-- Epoch 14
Norm: 1.22, NNZs: 3, Bias: 0.015000, T: 2800, Avg. loss: 0.273688
Total training time: 0.01 seconds.
-- Epoch 15
Norm: 1.26, NNZs: 3, Bias: 0.014000, T: 3000, Avg. loss: 0.267849
Total training time: 0.01 seconds.
-- Epoch 16
Norm: 1.29, NNZs: 3, Bias: 0.015000, T: 3200, Avg. loss: 0.262268
Total training time: 0.01 seconds.
-- Epoch 17
Norm: 1.32, NNZs: 3, Bias: 0.015000, T: 3400, Avg. loss: 0.257131
Total training time: 0.01 seconds.
-- Epoch 18
Norm: 1.35, NNZs: 3, Bias: 0.016000, T: 3600, Avg. loss: 0.252824
Total training time: 0.01 seconds.
-- Epoch 19
Norm: 1.37, NNZs: 3, Bias: 0.015000, T: 3800, Avg. loss: 0.249281
Total training time: 0.01 seconds.
-- Epoch 20
Norm: 1.40, NNZs: 3, Bias: 0.014000, T: 4000, Avg. loss: 0.245838
Total training time: 0.01 seconds.
-- Epoch 21
Norm: 1.42, NNZs: 3, Bias: 0.012000, T: 4200, Avg. loss: 0.242494
Total training time: 0.01 seconds.
-- Epoch 22
Norm: 1.44, NNZs: 3, Bias: 0.009000, T: 4400, Avg. loss: 0.239718
Total training time: 0.01 seconds.
-- Epoch 23
Norm: 1.47, NNZs: 3, Bias: 0.006000, T: 4600, Avg. loss: 0.236984
Total training time: 0.02 seconds.
-- Epoch 24
Norm: 1.49, NNZs: 3, Bias: 0.003000, T: 4800, Avg. loss: 0.234253
Total training time: 0.02 seconds.
```

```
-- Epoch 25
Norm: 1.51, NNZs: 3, Bias: 0.001000, T: 5000, Avg. loss: 0.231647
Total training time: 0.02 seconds.
-- Epoch 26
Norm: 1.53, NNZs: 3, Bias: 0.000000, T: 5200, Avg. loss: 0.229203
Total training time: 0.02 seconds.
-- Epoch 27
Norm: 1.55, NNZs: 3, Bias: 0.000000, T: 5400, Avg. loss: 0.227043
Total training time: 0.02 seconds.
-- Epoch 28
Norm: 1.57, NNZs: 3, Bias: 0.000000, T: 5600, Avg. loss: 0.225045
Total training time: 0.02 seconds.
-- Epoch 29
Norm: 1.59, NNZs: 3, Bias: -0.001000, T: 5800, Avg. loss: 0.223116
Total training time: 0.02 seconds.
-- Epoch 30
Norm: 1.61, NNZs: 3, Bias: -0.002000, T: 6000, Avg. loss: 0.221124
Total training time: 0.02 seconds.
-- Epoch 31
Norm: 1.63, NNZs: 3, Bias: -0.004000, T: 6200, Avg. loss: 0.219346
Total training time: 0.02 seconds.
-- Epoch 32
Norm: 1.65, NNZs: 3, Bias: -0.005000, T: 6400, Avg. loss: 0.217517
Total training time: 0.02 seconds.
-- Epoch 33
Norm: 1.66, NNZs: 3, Bias: -0.008000, T: 6600, Avg. loss: 0.215781
Total training time: 0.02 seconds.
-- Epoch 34
Norm: 1.68, NNZs: 3, Bias: -0.008000, T: 6800, Avg. loss: 0.214389
Total training time: 0.02 seconds.
-- Epoch 35
Norm: 1.70, NNZs: 3, Bias: -0.010000, T: 7000, Avg. loss: 0.213015
Total training time: 0.02 seconds.
-- Epoch 36
Norm: 1.71, NNZs: 3, Bias: -0.011000, T: 7200, Avg. loss: 0.211739
Total training time: 0.02 seconds.
-- Epoch 37
Norm: 1.73, NNZs: 3, Bias: -0.012000, T: 7400, Avg. loss: 0.210512
Total training time: 0.02 seconds.
-- Epoch 38
Norm: 1.74, NNZs: 3, Bias: -0.013000, T: 7600, Avg. loss: 0.209288
Total training time: 0.02 seconds.
-- Epoch 39
Norm: 1.76, NNZs: 3, Bias: -0.014000, T: 7800, Avg. loss: 0.208093
Total training time: 0.03 seconds.
```

```
-- Epoch 40
Norm: 1.77, NNZs: 3, Bias: -0.014000, T: 8000, Avg. loss: 0.206976
Total training time: 0.03 seconds.
-- Epoch 41
Norm: 1.78, NNZs: 3, Bias: -0.014000, T: 8200, Avg. loss: 0.206049
Total training time: 0.03 seconds.
-- Epoch 42
Norm: 1.80, NNZs: 3, Bias: -0.016000, T: 8400, Avg. loss: 0.205249
Total training time: 0.03 seconds.
-- Epoch 43
Norm: 1.81, NNZs: 3, Bias: -0.016000, T: 8600, Avg. loss: 0.204515
Total training time: 0.03 seconds.
-- Epoch 44
Norm: 1.82, NNZs: 3, Bias: -0.015000, T: 8800, Avg. loss: 0.203853
Total training time: 0.03 seconds.
-- Epoch 45
Norm: 1.83, NNZs: 3, Bias: -0.015000, T: 9000, Avg. loss: 0.203155
Total training time: 0.03 seconds.
Convergence after 45 epochs took 0.03 seconds
```

Out[128…

```
SGDClassifier(alpha=0.001, eta0=0.001, learning_rate='constant',
              random_state=60, verbose=2)
```

In [129…

```
# obtain weights of features
coef = clf2.coef_
print(coef) #-->#--> Now f3 is useful in classifying the pts for which weight is large
```

```
[[-0.00476652  0.02253872  1.83053697]]
```

OBSERVATION :

Here, the same happens with SVM too where we can infer that the 3rd feature is the useful feature in classifying the data points and this result is possible only because we gave equal importance by standardizing the features before modelling the data.