

sml-assignment-6

September 18, 2024

```
[ ]: import pandas as pd
df=pd.read_csv('/content/Book1.csv')
df
```

```
[ ]:      study hours per week  class attendance %  previous exam score \
0                10                90                85
1                 5                60                70
2                15                95                90
3                 2                30                50
4                12                85                88
5                 8                74                65
6                14                92                93
7                 3                50                55
8                11                88                80
9                 6                65                60
```

```
      participation in group study  subission of assignments  pass or fail
0                1                100                1
1                0                80                0
2                1                100                1
3                0                40                0
4                1                90                1
5                0                70                0
6                1                95                1
7                0                60                0
8                1                85                1
9                0                75                0
```

```
[ ]: target='pass or fail'
print(df[target])
```

```
0    1
1    0
2    1
3    0
4    1
5    0
6    1
```

```

7    0
8    1
9    0
Name: pass or fail, dtype: int64

```

```

[ ]: x=df.drop(target,axis=1)
     print(x)

```

	study hours per week	class attendance %	previous exam score \
0	10	90	85
1	5	60	70
2	15	95	90
3	2	30	50
4	12	85	88
5	8	74	65
6	14	92	93
7	3	50	55
8	11	88	80
9	6	65	60

	participation in group study	subission of assignments
0	1	100
1	0	80
2	1	100
3	0	40
4	1	90
5	0	70
6	1	95
7	0	60
8	1	85
9	0	75

```

[ ]: w0=1
     w1=3
     w2=2
     w3=4
     w4=1
     w5=2
     z=w0*1+w1*x["study hours per week"]+w2*x["class attendance %"]+w3*x["previous_
     ↪exam score"]+w4*x["participation in group study"]+w5*x["subission of_
     ↪assignments"]
     print(z)
     yp=[]
     for i in range(10):
         if z[i]>0:
             yp.append(1)
             #print(yp)

```

```
else:
    yp.append(0)
    # print(yp)
```

```
0    752
1    576
2    797
3    347
4    740
5    573
6    790
7    450
8    701
9    539
dtype: int64
```

```
[ ]: print(yp)
```

```
[1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
```

```
[ ]: error=df[target]-yp
      print(error)
```

```
0    0
1   -1
2    0
3   -1
4    0
5   -1
6    0
7   -1
8    0
9   -1
Name: pass or fail, dtype: int64
```

```
[ ]: mean=error.mean()
      print(mean)
```

```
-0.5
```

```
[ ]: e_update=w0+0.01*mean
      print(e_update)
      e_update1=w1+0.01*mean
      print(e_update1)
      e_update2=w2+0.01*mean
      print(e_update2)
      e_update3=w3+0.01*mean
      print(e_update3)
```

```
e_update4=w4+0.01*mean
print(e_update4)
e_update5=w5+0.01*mean
print(e_update5)
```

```
0.995
2.995
1.995
3.995
0.995
1.995
```

```
[ ]: w0=e_update
      w1=e_update1
      w2=e_update2
      w3=e_update3
      w4=e_update4
      w5=e_update5
      z=w0*1+w1*x["study hours per week"]+w2*x["class attendance %"]+w3*x["previous_
      ↳exam score"]+w4*x["participation in group study"]+w5*x["subission of_
      ↳assignments"]
      print(z)
      yp1=[]
      for i in range(10):
          if z[i]>0:
              yp1.append(1)
          else:
              yp1.append(0)
```

```
0    750.565
1    574.920
2    795.490
3    346.385
4    738.615
5    571.910
6    788.520
7    449.155
8    699.670
9    537.965
dtype: float64
```

```
[ ]: print(yp1)
```

```
[1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
```

```
[ ]: error=df[target]-yp1
      print(error)
```

```

0    0
1   -1
2    0
3   -1
4    0
5   -1
6    0
7   -1
8    0
9   -1

```

Name: pass or fail, dtype: int64

```
[ ]: error_mean=error.mean()
      print(error_mean)
```

-0.5

```
[ ]: e_update=w0+0.01*error_mean
      print(e_update)
      e_update1=w1+0.01*error_mean
      print(e_update1)
      e_update2=w2+0.01*error_mean
      print(e_update2)
      e_update3=w3+0.01*error_mean
      print(e_update3)
      e_update4=w4+0.01*error_mean
      print(e_update4)
      e_update5=w5+0.01*error_mean
      print(e_update5)
```

```

0.99
2.99
1.9900000000000002
3.99
0.99
1.9900000000000002

```

```
[ ]: w0=e_update
      w1=e_update1
      w2=e_update2
      w3=e_update3
      w4=e_update4
      w5=e_update5
      z=w0*1+w1*x["study hours per week"]+w2*x["class attendance %"]+w3*x["previous_
      ↪exam score"]+w4*x["participation in group study"]+w5*x["subission of_
      ↪assignments"]
      print(z)
```

```

0    749.13
1    573.84
2    793.98
3    345.77
4    737.23
5    570.82
6    787.04
7    448.31
8    698.34
9    536.93
dtype: float64

```

```

[ ]: yp2=[]
for i in range(10):
    if z[i]>0:
        yp2.append(1)
    else:
        yp2.append(0)
print(yp2)

```

```
[1, 1, 1, 1, 1, 1, 1, 1, 1, 1]
```

```

[ ]: import pandas as pd
df=pd.read_csv('/content/diabetes_data_upload.csv')
df

```

```

[ ]:
    Age  Gender Polyuria Polydipsia sudden weight loss weakness Polyphagia \
0    40   Male      No        Yes           No      Yes      No
1    58   Male      No        No           No      Yes      No
2    41   Male     Yes        No           No      Yes     Yes
3    45   Male      No        No          Yes     Yes     Yes
4    60   Male     Yes     Yes           Yes     Yes     Yes
..    ...   ...      ...      ...           ...     ...     ...
515   39  Female     Yes     Yes           Yes      No     Yes
516   48  Female     Yes     Yes           Yes     Yes     Yes
517   58  Female     Yes     Yes           Yes     Yes     Yes
518   32  Female      No        No           No      Yes      No
519   42   Male      No        No           No      No      No

    Genital thrush visual blurring Itching Irritability delayed healing \
0                No                No      Yes           No      Yes
1                No                Yes      No           No      No
2                No                No      Yes           No      Yes
3                Yes                No      Yes           No      Yes
4                No                Yes     Yes           Yes     Yes
..                ...                ...      ...           ...     ...
515               No                No      Yes           No      Yes

```

516	No	No	Yes	Yes	Yes
517	No	Yes	No	No	No
518	No	Yes	Yes	No	Yes
519	No	No	No	No	No

	partial paresis	muscle stiffness	Alopecia	Obesity	class
0	No	Yes	Yes	Yes	Positive
1	Yes	No	Yes	No	Positive
2	No	Yes	Yes	No	Positive
3	No	No	No	No	Positive
4	Yes	Yes	Yes	Yes	Positive
..
515	Yes	No	No	No	Positive
516	Yes	No	No	No	Positive
517	Yes	Yes	No	Yes	Positive
518	No	No	Yes	No	Negative
519	No	No	No	No	Negative

[520 rows x 17 columns]

```
[ ]: y='class'
      print(df[y])
```

```
0      Positive
1      Positive
2      Positive
3      Positive
4      Positive
...
515     Positive
516     Positive
517     Positive
518     Negative
519     Negative
```

Name: class, Length: 520, dtype: object

```
[ ]: x=df.drop(y,axis=1)
      print(x)
```

	Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	Polyphagia	\
0	40	Male	No	Yes	No	Yes	No	
1	58	Male	No	No	No	Yes	No	
2	41	Male	Yes	No	No	Yes	Yes	
3	45	Male	No	No	Yes	Yes	Yes	
4	60	Male	Yes	Yes	Yes	Yes	Yes	
..	
515	39	Female	Yes	Yes	Yes	No	Yes	
516	48	Female	Yes	Yes	Yes	Yes	Yes	

517	58	Female	Yes	Yes	Yes	Yes	Yes
518	32	Female	No	No	No	Yes	No
519	42	Male	No	No	No	No	No

	Genital thrush	visual blurring	Itching	Irritability	delayed healing	\
0	No	No	Yes	No	Yes	
1	No	Yes	No	No	No	
2	No	No	Yes	No	Yes	
3	Yes	No	Yes	No	Yes	
4	No	Yes	Yes	Yes	Yes	
..	
515	No	No	Yes	No	Yes	
516	No	No	Yes	Yes	Yes	
517	No	Yes	No	No	No	
518	No	Yes	Yes	No	Yes	
519	No	No	No	No	No	

	partial paresis	muscle stiffness	Alopecia	Obesity
0	No	Yes	Yes	Yes
1	Yes	No	Yes	No
2	No	Yes	Yes	No
3	No	No	No	No
4	Yes	Yes	Yes	Yes
..
515	Yes	No	No	No
516	Yes	No	No	No
517	Yes	Yes	No	Yes
518	No	No	Yes	No
519	No	No	No	No

[520 rows x 16 columns]

```
[ ]: df['class']=df['class'].replace({'Negative':0,'Positive':1})
df['Gender'] = df['Gender'].replace({'Female': 0, 'Male': 1})
df['Polyuria'] = df['Polyuria'].replace({'No': 0, 'Yes': 1})
df['Polydipsia'] = df['Polydipsia'].replace({'No': 0, 'Yes': 1})
df['sudden weight loss'] = df['sudden weight loss'].replace({'No': 0, 'Yes': 1})
df['weakness'] = df['weakness'].replace({'No': 0, 'Yes': 1})
df['Polyphagia'] = df['Polyphagia'].replace({'No': 0, 'Yes': 1})
df['Genital thrush'] = df['Genital thrush'].replace({'No': 0, 'Yes': 1})
df['visual blurring'] = df['visual blurring'].replace({'No': 0, 'Yes': 1})
df['Itching'] = df['Itching'].replace({'No': 0, 'Yes': 1})
df['Irritability'] = df['Irritability'].replace({'No': 0, 'Yes': 1})
df['delayed healing'] = df['delayed healing'].replace({'No': 0, 'Yes': 1})
df['partial paresis'] = df['partial paresis'].replace({'No': 0, 'Yes': 1})
df['muscle stiffness'] = df['muscle stiffness'].replace({'No': 0, 'Yes': 1})
df['Alopecia'] = df['Alopecia'].replace({'No': 0, 'Yes': 1})
```



```
df['Obesity'] = df['Obesity'].replace({'No': 0, 'Yes': 1})
print(df)
```

	Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	\
0	40	1	0	1	0	1	
1	58	1	0	0	0	1	
2	41	1	1	0	0	1	
3	45	1	0	0	1	1	
4	60	1	1	1	1	1	
..	
515	39	0	1	1	1	0	
516	48	0	1	1	1	1	
517	58	0	1	1	1	1	
518	32	0	0	0	0	1	
519	42	1	0	0	0	0	

	Polyphagia	Genital thrush	visual blurring	Itching	Irritability	\
0	0	0	0	1	0	
1	0	0	1	0	0	
2	1	0	0	1	0	
3	1	1	0	1	0	
4	1	0	1	1	1	
..	
515	1	0	0	1	0	
516	1	0	0	1	1	
517	1	0	1	0	0	
518	0	0	1	1	0	
519	0	0	0	0	0	

	delayed healing	partial paresis	muscle stiffness	Alopecia	Obesity	\
0	1	0	1	1	1	
1	0	1	0	1	0	
2	1	0	1	1	0	
3	1	0	0	0	0	
4	1	1	1	1	1	
..	
515	1	1	0	0	0	
516	1	1	0	0	0	
517	0	1	1	0	1	
518	1	0	0	1	0	
519	0	0	0	0	0	

	class
0	1
1	1
2	1
3	1

```

4          1
..         ...
515        1
516        1
517        1
518        0
519        0

```

[520 rows x 17 columns]

```

[ ]: w0=1
      w1=3
      w2=2
      w3=4
      w4=1
      w5=2
      w6=3
      w7=1
      w8=2
      w9=1
      w10=3
      w11=1
      w12=2
      w13=1
      w14=3
      w15=1
      w16=2
      z=
      ↪w0*1+w1*df['Age']+w2*df['Gender']+w3*df['Polyuria']+w4*df['Polydipsia']+w5*df['sudden_
      ↪weight loss']+w6*df['weakness']+w7*df['Polyphagia']+w8*df['Genital_
      ↪thrush']+w9*df['visual_
      ↪blurring']+w10*df['Itching']+w11*df['Irritability']+w12*df['delayed_
      ↪healing']+w13*df['partial paresis']+w14*df['muscle_
      ↪stiffness']+w15*df['Alopecia']+w16*df['Obesity']
      print(z)
      yp=[]
      for i in range(520):
          if z[i]>0:
              yp.append(1)
          else:
              yp.append(0)

```

```

0          138
1          183
2          143
3          151
4          208
...

```

515	132
516	163
517	193
518	107
519	129

```
[ ]: print(yp)
```

[illegible]

```
[ ]: error=df['class']-yp
      c=0
      for i in range(520):
          if error[i]==0:
              c=c+1
      print(c/520)
```

0.6153846153846154

```
[ ]: mean=error.mean()  
      print(mean)
```

-0.38461538461538464

```
[ ]: e_update=w0+0.01*mean
print(e_update)
e_update1=w1+0.01*mean
print(e_update1)
```

```

e_update2=w2+0.01*mean
print(e_update2)
e_update3=w3+0.01*mean
print(e_update3)
e_update4=w4+0.01*mean
print(e_update4)
e_update5=w5+0.01*mean
print(e_update5)
e_update6=w6+0.01*mean
print(e_update6)
e_update7=w7+0.01*mean
print(e_update7)
e_update8=w8+0.01*mean
print(e_update8)
e_update9=w9+0.01*mean
print(e_update9)
e_update10=w10+0.01*mean
print(e_update10)
e_update11=w11+0.01*mean
print(e_update11)
e_update12=w12+0.01*mean
print(e_update12)
e_update13=w13+0.01*mean
print(e_update13)
e_update14=w14+0.01*mean
print(e_update14)
e_update15=w15+0.01*mean
print(e_update15)
e_update16=w16+0.01*mean
print(e_update16)

```

```

0.9961538461538462
2.996153846153846
1.9961538461538462
3.996153846153846
0.9961538461538462
1.9961538461538462
2.996153846153846
0.9961538461538462
1.9961538461538462
0.9961538461538462
2.996153846153846
0.9961538461538462
1.9961538461538462
0.9961538461538462
2.996153846153846
0.9961538461538462

```

1.9961538461538462

```
[ ]: w0=e_update
      w1=e_update1
      w2=e_update2
      w3=e_update3
      w4=e_update4
      w5=e_update5
      w6=e_update6
      w7=e_update7
      w8=e_update8
      w9=e_update9
      w10=e_update10
      w11=e_update11
      w12=e_update12
      w13=e_update13
      w14=e_update14
      w15=e_update15
      w16=e_update16
      z=
      ↪w0*1+w1*df['Age']+w2*df['Gender']+w3*df['Polyuria']+w4*df['Polydipsia']+w5*df['sudden_
      ↪weight loss']+w6*df['weakness']+w7*df['Polyphagia']+w8*df['Genital_
      ↪thrush']+w9*df['visual_
      ↪blurring']+w10*df['Itching']+w11*df['Irritability']+w12*df['delayed_
      ↪healing']+w13*df['partial paresis']+w14*df['muscle_
      ↪stiffness']+w15*df['Alopecia']+w16*df['Obesity']
      print(z)
      yp1=[]
      for i in range(520):
          if z[i]>0:
              yp1.append(1)
          else:
              yp1.append(0)
```

```
0      137.811538
1      182.753846
2      142.807692
3      150.796154
4      207.711538
...
515     131.819231
516     162.776923
517     192.738462
518     106.853846
519     128.830769
Length: 520, dtype: float64
```

```
[ ]: print(yp1)
```

[illegible]

```
[ ]: error=df['class']-yp1
      print(error)
```

```

0      0
1      0
2      0
3      0
4      0
..
515    0
516    0
517    0
518   -1
519   -1
Name: class, Length: 520, dtype: int64

```

```
[ ]: X_train=df.drop('class',axis=1)
      print(X_train)
```

	Age	Gender	Polyuria	Polydipsia	sudden weight loss	weakness	\
0	40	1	0	1	0	1	
1	58	1	0	0	0	1	
2	41	1	1	0	0	1	
3	45	1	0	0	1	1	
4	60	1	1	1	1	1	

..
515	39	0	1	1	1	1	0	
516	48	0	1	1	1	1	1	
517	58	0	1	1	1	1	1	
518	32	0	0	0	0	0	1	
519	42	1	0	0	0	0	0	

	Polyphagia	Genital thrush	visual blurring	Itching	Irritability	\
0	0	0	0	1	0	
1	0	0	1	0	0	
2	1	0	0	1	0	
3	1	1	0	1	0	
4	1	0	1	1	1	
..	
515	1	0	0	1	0	
516	1	0	0	1	1	
517	1	0	1	0	0	
518	0	0	1	1	0	
519	0	0	0	0	0	

	delayed healing	partial paresis	muscle stiffness	Alopecia	Obesity
0	1	0	1	1	1
1	0	1	0	1	0
2	1	0	1	1	0
3	1	0	0	0	0
4	1	1	1	1	1
..
515	1	1	0	0	0
516	1	1	0	0	0
517	0	1	1	0	1
518	1	0	0	1	0
519	0	0	0	0	0

[520 rows x 16 columns]

```
[ ]: Y_train=df['class']
      print(Y_train)
```

0	1
1	1
2	1
3	1
4	1
..	
515	1
516	1
517	1
518	0

```
519      0
Name: class, Length: 520, dtype: int64
```