Logistic Regression

February 20, 2025

1 Implementation of Logistic Regression on StudyHours Vs ExamResult Dataset

[ ]:
```
import numpy as np
```

[ ]:
```
def Sig(z):
  return (1/(1+np.exp(-z)))
```

[ ]:
```
import pandas as pd
```

[ ]:
```
df_hs = pd.read_csv('/content/hours_results.csv')
df_hs
```

[ ]:

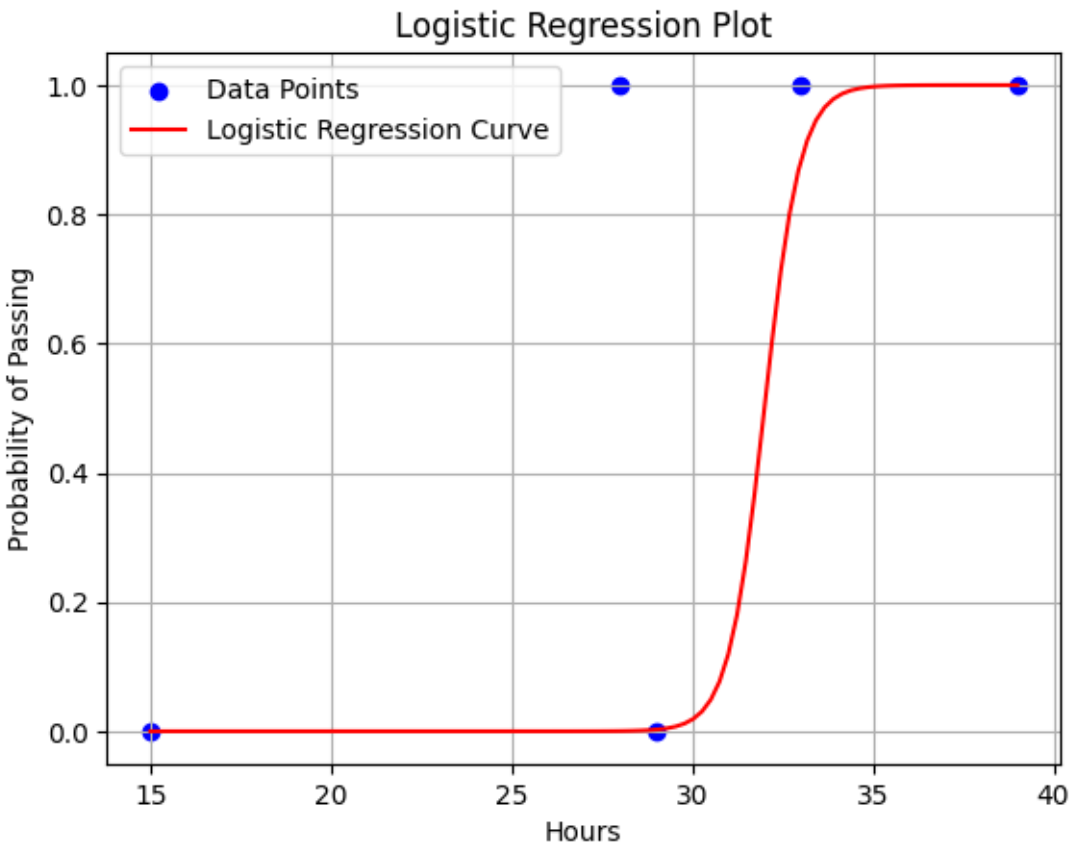
	Hours	Result
0	29	0
1	15	0
2	33	1
3	28	1
4	39	1

[ ]:
```
a = -64
b = 2
predicted=list()
for i in range(len(df_hs)):
  x=df_hs.iloc[i,0]
  # print(x)
  z = a+(b*x)
  tmp = Sig(z)
  print(tmp)
  if tmp>0.5:
    predicted.append(1)
  else:
    predicted.append(0)

df_hs['predicted']=predicted
df_hs
```

```
0.0024726231566347743
1.71390843154201e-15
0.8807970779778823
0.0003535501304664781
0.9999991684719722
```

[ ]:
```
   Hours  Result  predicted
0     29       0          0
1     15       0          0
2     33       1          1
3     28       1          0
4     39       1          1
```

[ ]:
```
import matplotlib.pyplot as plt
# Assuming 'df_hs' is your DataFrame and 'hours' and 'results' are your columns

# Extract data for plotting
X = df_hs['Hours'].values.reshape(-1, 1)  # Reshape to a 2D array
y = df_hs['Result'].values

# Create a scatter plot of the data points
plt.scatter(X, y, color='blue', label='Data Points')

# Generate a range of x-values for the logistic regression curve
x_range = np.linspace(X.min(), X.max(), 100)

# Calculate predicted values for the x_range using your logistic regression␣
 ↪model
# Instead of using the previously calculated 'predicted' list,
# we need to predict the probabilities for the new x_range values
predicted_probs = [Sig(a + b * x) for x in x_range] # Assuming a and b are your␣
 ↪model parameters

# Plot the logistic regression curve using the calculated probabilities
plt.plot(x_range, predicted_probs, color='red', label='Logistic Regression␣
 ↪Curve')

# Add labels and a legend
plt.xlabel('Hours')
plt.ylabel('Probability of Passing')
plt.title('Logistic Regression Plot')
plt.legend()
plt.grid()

# Display the plot
plt.show()
```


Logistic Regression Plot
Data Points
Logistic Regression Curve
Probability of Passing
Hours
1.0
0.8
0.6
0.4
0.2
0.0
15
20
25
30
35
40