dataset-code-my-project

March 30, 2024

```python
[4]: import numpy as np
     import pandas as pd
     import matplotlib.pyplot as plt
     from sklearn.datasets import make_classification

     # Generate synthetic data for subscription renewal prediction
     def generate_subscription_data(n_samples=1000, n_features=10, n_classes=2,
      ↪n_clusters_per_class=2):
         X, y = make_classification(n_samples=n_samples, n_features=n_features,
      ↪n_classes=n_classes,
                                    n_clusters_per_class=n_clusters_per_class,
      ↪random_state=42)

         # Convert to DataFrame
         columns = [f"Feature_{i+1}" for i in range(n_features)]
         data = pd.DataFrame(X, columns=columns)
         data['Renewal'] = y

         return data

     # Generate synthetic subscription data
     subscription_data = generate_subscription_data(n_samples=1000, n_features=10,
      ↪n_classes=2, n_clusters_per_class=2)

     # Display the first few rows of the dataset
     print(subscription_data.head())

     # Create a pie chart for subscription renewals
     renewal_counts = subscription_data['Renewal'].value_counts()
     labels = ['Renewed', 'Not Renewed']
     plt.pie(renewal_counts, labels=labels, autopct='%1.1f%%', startangle=140)
     plt.axis('equal')
     plt.title('Subscription Renewal Prediction')
     plt.show()
```
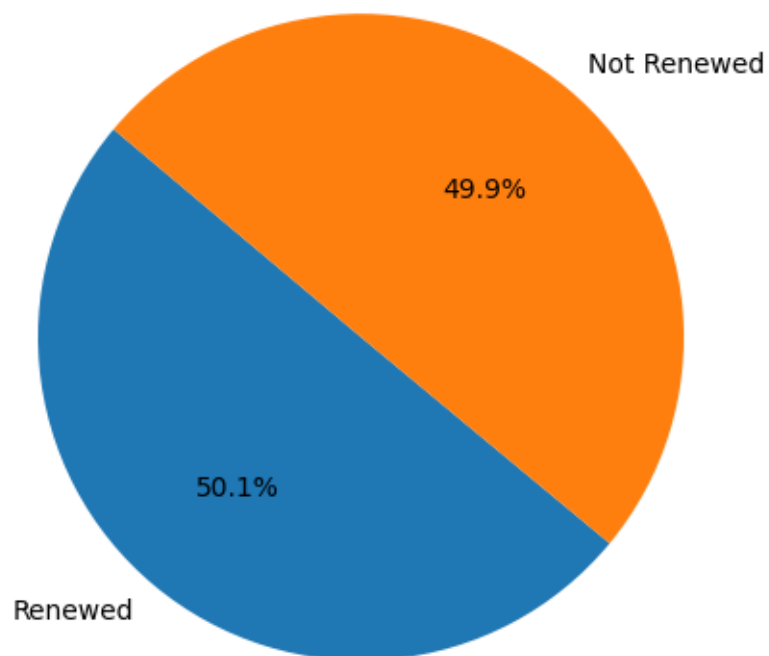
```
   Feature_1  Feature_2  Feature_3  Feature_4  Feature_5  Feature_6  \
0   0.964799  -0.066449   0.986768  -0.358079   0.997266   1.181890
1  -0.916511  -0.566395  -1.008614   0.831617  -1.176962   1.820544
```

```
2  -0.109484  -0.432774  -0.457649   0.793818  -0.268646  -1.836360
3   1.750412   2.023606   1.688159   0.006800  -1.607661   0.184741
4  -0.224726  -0.711303  -0.220778   0.117124   1.536061   0.597538

   Feature_7  Feature_8  Feature_9  Feature_10  Renewal
0  -1.615679  -1.210161  -0.628077    1.227274        0
1   1.752375  -0.984534   0.363896    0.209470        1
2   1.239086  -0.246383  -1.058145   -0.297376        1
3  -2.619427  -0.357445  -1.473127   -0.190039        0
4   0.348645  -0.939156   0.175915    0.236224        1
```

Subscription Renewal Prediction


```
[5]: import numpy as np
     import pandas as pd
     import matplotlib.pyplot as plt
     from sklearn.datasets import make_classification

     # Generate synthetic data for subscription renewal prediction
     def generate_subscription_data(n_samples=1000, n_features=10, n_classes=2,␣
      ↪n_clusters_per_class=2):
         X, y = make_classification(n_samples=n_samples, n_features=n_features,␣
      ↪n_classes=n_classes,
```

```python
                                    n_clusters_per_class=n_clusters_per_class,␣
 ↪random_state=42)

    # Convert to DataFrame
    columns = [f"Feature_{i+1}" for i in range(n_features)]
    data = pd.DataFrame(X, columns=columns)
    data['Renewal'] = y

    return data

# Generate synthetic subscription data
subscription_data = generate_subscription_data(n_samples=1000, n_features=10,␣
 ↪n_classes=2, n_clusters_per_class=2)

# Count the number of renewals and non-renewals
renewal_counts = subscription_data['Renewal'].value_counts()

# Plotting the bar graph
plt.bar(['Renewed', 'Not Renewed'], renewal_counts, color=['blue', 'red'])
plt.title('Subscription Renewal Prediction')
plt.xlabel('Subscription Renewal')
plt.ylabel('Number of Customers')
plt.show()
```
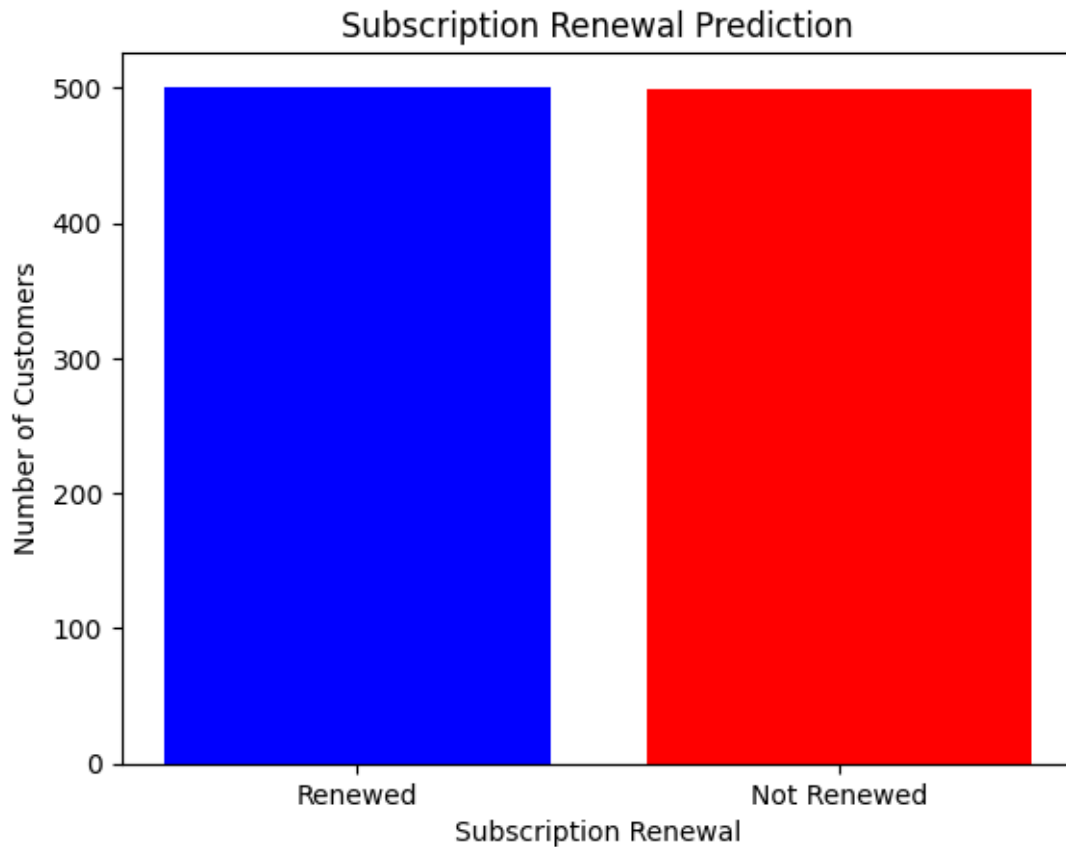
Subscription Renewal Prediction

```python
[6]: import matplotlib.pyplot as plt
     from matplotlib.patches import Rectangle

     # Define the positions of each block
     blocks = {
         'Start': (1, 5),
         'Gather Data': (1, 4),
         'Preprocess Data': (1, 3),
         'Feature Engineering': (1, 2),
         'Train Model': (1, 1),
         'Evaluate Model': (3, 3),
         'Deploy Model': (3, 2),
         'Monitor & Update': (3, 1),
         'End': (5, 3)
     }

     # Define the connections between blocks
     connections = [
         ('Start', 'Gather Data'),
         ('Gather Data', 'Preprocess Data'),
```

```python
    ('Preprocess Data', 'Feature Engineering'),
    ('Feature Engineering', 'Train Model'),
    ('Train Model', 'Evaluate Model'),
    ('Evaluate Model', 'Deploy Model'),
    ('Deploy Model', 'End'),
    ('Evaluate Model', 'Monitor & Update'),
    ('Monitor & Update', 'End')
]

# Create the plot
plt.figure(figsize=(8, 6))

# Plot blocks
for block, pos in blocks.items():
    plt.gca().add_patch(Rectangle(pos, 1, 0.5, color='lightblue', alpha=0.5))
    plt.text(pos[0] + 0.5, pos[1] + 0.25, block, ha='center', va='center')

# Plot connections
for connection in connections:
    start = blocks[connection[0]]
    end = blocks[connection[1]]
    plt.arrow(start[0] + 0.5, start[1], end[0] - start[0], end[1] - start[1],
              head_width=0.1, head_length=0.1, fc='k', ec='k')

# Set plot limits and remove axes
plt.xlim(0, 6)
plt.ylim(0, 6)
plt.axis('off')
plt.title('Customer Subscription Renewal Prediction Flowchart')
plt.show()
```

Customer Subscription Renewal Prediction Flowchart

Start

Gather Data

Preprocess Data

Evaluate Model

End

Feature Engineering

Deploy Model

Train Model

Monitor & Update