In [91]:
```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LinearRegression
from sklearn.linear_model import LogisticRegression
import optuna
from sklearn.metrics import accuracy_score
```

In [49]:
```python
df = pd.read_csv(r"D:\download\archive (6)\WineQT.csv")
```

In [50]:
```python
df.head()
```

Out[50]:

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality	Id
0	7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.9978	3.51	0.56	9.4	5	0
1	7.8	0.88	0.00	2.6	0.098	25.0	67.0	0.9968	3.20	0.68	9.8	5	1
2	7.8	0.76	0.04	2.3	0.092	15.0	54.0	0.9970	3.26	0.65	9.8	5	2
3	11.2	0.28	0.56	1.9	0.075	17.0	60.0	0.9980	3.16	0.58	9.8	6	3
4	7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.9978	3.51	0.56	9.4	5	4

In [51]:
```python
df.columns
```

Out[51]:
```
Index(['fixed acidity', 'volatile acidity', 'citric acid', 'residual sugar',
       'chlorides', 'free sulfur dioxide', 'total sulfur dioxide', 'density',
       'pH', 'sulphates', 'alcohol', 'quality', 'Id'],
      dtype='object')
```

In [52]:
```python
df.dtypes
```

Out[52]:
```
fixed acidity           float64
volatile acidity        float64
citric acid             float64
residual sugar          float64
chlorides               float64
free sulfur dioxide     float64
total sulfur dioxide    float64
density                 float64
pH                      float64
sulphates               float64
alcohol                 float64
quality                   int64
Id                        int64
dtype: object
```

In [53]:
```python
df.shape
```

Out[53]:
```
(1143, 13)
```

In [54]:
```python
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1143 entries, 0 to 1142
Data columns (total 13 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   fixed acidity         1143 non-null   float64
 1   volatile acidity      1143 non-null   float64
 2   citric acid           1143 non-null   float64
 3   residual sugar        1143 non-null   float64
 4   chlorides             1143 non-null   float64
 5   free sulfur dioxide   1143 non-null   float64
 6   total sulfur dioxide  1143 non-null   float64
 7   density               1143 non-null   float64
 8   pH                    1143 non-null   float64
 9   sulphates             1143 non-null   float64
 10  alcohol               1143 non-null   float64
 11  quality               1143 non-null   int64  
 12  Id                    1143 non-null   int64  
dtypes: float64(11), int64(2)
memory usage: 116.2 KB
```

In [55]:
```python
df.isna().any()
```

Out[55]:
```
fixed acidity           False
volatile acidity        False
citric acid             False
residual sugar          False
chlorides               False
free sulfur dioxide     False
total sulfur dioxide    False
density                 False
pH                      False
sulphates               False
alcohol                 False
quality                 False
Id                      False
dtype: bool
```

Exploratory Data Analysis

In [56]:
```python
col_list = list(df.columns.values)
col_list
```

Out[56]:
```
['fixed acidity',
 'volatile acidity',
 'citric acid',
 'residual sugar',
 'chlorides',
 'free sulfur dioxide',
 'total sulfur dioxide',
 'density',
 'pH',
 'sulphates',
 'alcohol',
 'quality',
 'Id']
```

In [57]:
```python
col_list.remove("Id")
col_list.remove("quality")

col_list
```

Out[57]:
```
['fixed acidity',
 'volatile acidity',
 'citric acid',
 'residual sugar',
 'chlorides',
 'free sulfur dioxide',
 'total sulfur dioxide',
 'density',
 'pH',
 'sulphates',
 'alcohol']
```

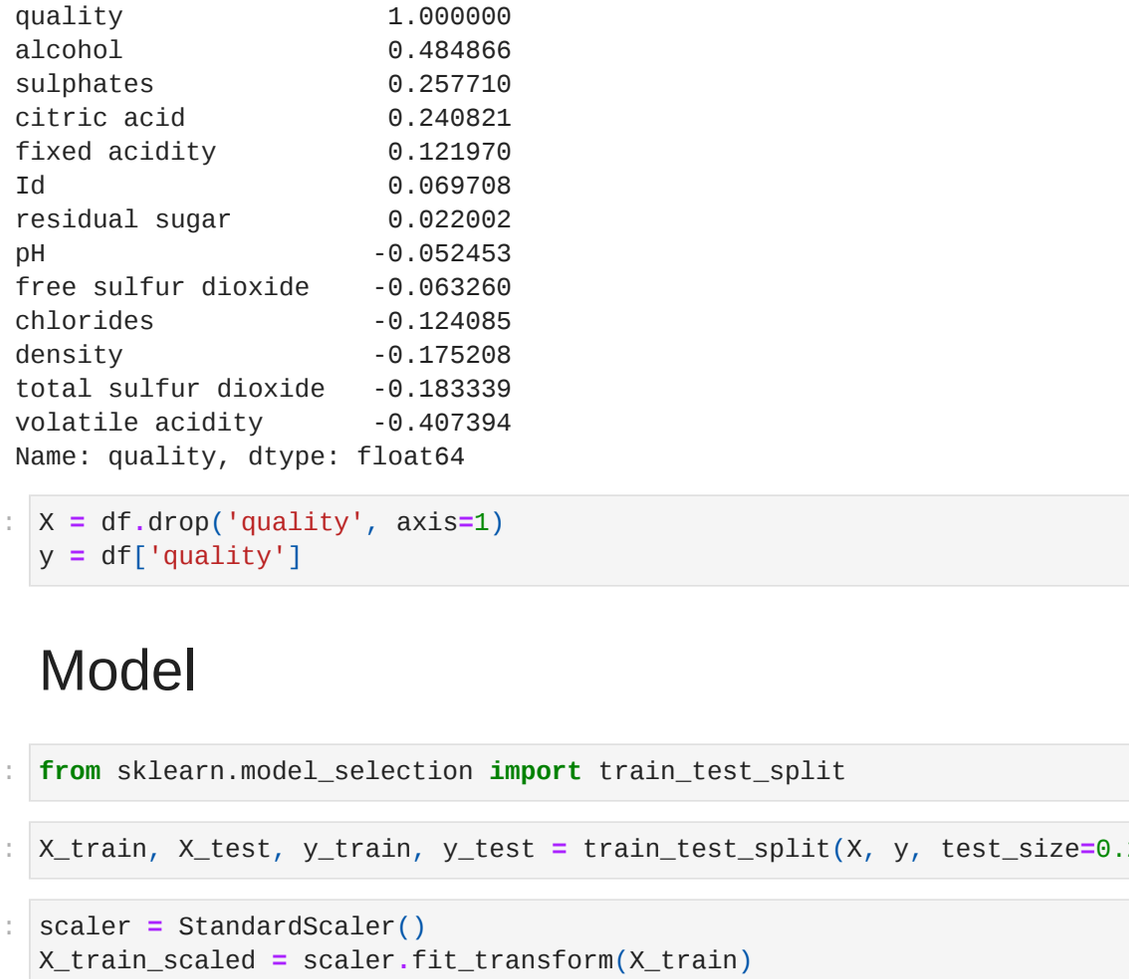
In [60]:
```python
import warnings
warnings.filterwarnings('ignore')
from scipy import stats
fig,ax = plt.subplots(11,3,figsize=(25,60))
for index,i in enumerate(col_list):
    sns.distplot(df[i],ax=ax[index,0],color='blue')
    sns.boxplot(df[i],ax=ax[index,1],color='gree')
    stats.probplot(df[i],plot=ax[index,2])

fig.tight_layout()
plt.suptitle("Uni-Variate Analysis of continuous variables")
```

Out[60]:
```
Text(0.5, 0.98, 'Uni-Variate Analysis of continuous variables')
```

In [61]:
```python
df['quality'].value_counts().plot(kind='bar', title='Quality')
```

Out[61]:
```
<Axes: title={'center': 'Quality'}>
```



In [80]:
```python
print(correlation_matrix['quality'].sort_values(ascending=False))
```

```
quality                 1.000000
alcohol                 0.484866
sulphates               0.257710
citric acid             0.240821
fixed acidity           0.121970
Id                      0.069708
residual sugar          0.022002
pH                     -0.052453
free sulfur dioxide    -0.063260
chlorides              -0.124085
density                -0.175208
total sulfur dioxide   -0.183339
volatile acidity       -0.407394
Name: quality, dtype: float64
```

In [69]:
```python
X = df.drop('quality', axis=1)
y = df['quality']
```

Model

In [70]:
```python
from sklearn.model_selection import train_test_split
```

In [71]:
```python
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

In [72]:
```python
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.fit_transform(X_test)
```

In [85]:
```python
from imblearn.over_sampling import SMOTE
smote = SMOTE(random_state=42)
X_train_resampled, y_train_resampled = smote.fit_resample(X_train_scaled, y_train)
```

In [89]:
```python
def objective_lr(trial):
    C = trial.suggest_loguniform('C', 1e-3, 1e3)

    model = LogisticRegression(C=C, random_state=42)
    model.fit(X_train_resampled, y_train_resampled)
    y_pred = model.predict(X_test_scaled)

    return -accuracy_score(y_test, y_pred)
```

In [90]:
```python
study_lr = optuna.create_study(direction='minimize')
study_lr.optimize(objective_lr, n_trials=100)
```

```
[I 2023-08-23 10:04:23,462] A new study created in memory with name: no-name-7311c9b0-1290-4d0b-aed8-7c61c59dd65b
[I 2023-08-23 10:04:23,621] Trial 0 finished with value: -0.4104803493449782 and parameters: {'C': 0.7381427445200661}. Best is trial 0 with value: -0.4104803493449782.
[I 2023-08-23 10:04:23,663] Trial 1 finished with value: -0.4104803493449782 and parameters: {'C': 0.2390960613028371}. Best is trial 0 with value: -0.4104803493449782.
[I 2023-08-23 10:04:23,720] Trial 2 finished with value: -0.4104803493449782 and parameters: {'C': 1.3527510162704521}. Best is trial 0 with value: -0.4104803493449782.
[I 2023-08-23 10:04:23,765] Trial 3 finished with value: -0.4366812227074236 and parameters: {'C': 311.227136063635}. Best is trial 3 with value: -0.4366812227074236.
[I 2023-08-23 10:04:23,819] Trial 4 finished with value: -0.4410480349344978 and parameters: {'C': 96.78859327246558}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:23,866] Trial 5 finished with value: -0.4410480349344978 and parameters: {'C': 134.62942846184913}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:23,920] Trial 6 finished with value: -0.4104803493449782 and parameters: {'C': 1.6258563767966203}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:23,973] Trial 7 finished with value: -0.4366812227074236 and parameters: {'C': 452.545117935432}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,023] Trial 8 finished with value: -0.4366812227074236 and parameters: {'C': 5.657420764562677}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,076] Trial 9 finished with value: -0.4366812227074236 and parameters: {'C': 262.53209759084393}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,093] Trial 10 finished with value: -0.28820960698689957 and parameters: {'C': 0.001778889888994725}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,147] Trial 11 finished with value: -0.4366812227074236 and parameters: {'C': 42.3046562498302}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,213] Trial 12 finished with value: -0.4410480349344978 and parameters: {'C': 53.07733542970679}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,270] Trial 13 finished with value: -0.4410480349344978 and parameters: {'C': 39.20766907314035}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,324] Trial 14 finished with value: -0.43231441048034935 and parameters: {'C': 729.9528289216984}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,377] Trial 15 finished with value: -0.4366812227074236 and parameters: {'C': 13.910751173417893}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,431] Trial 16 finished with value: -0.4410480349344978 and parameters: {'C': 110.22938702214952}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,485] Trial 17 finished with value: -0.4366812227074236 and parameters: {'C': 639.4277713111557}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,545] Trial 18 finished with value: -0.4410480349344978 and parameters: {'C': 11.139746113092315}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,614] Trial 19 finished with value: -0.4366812227074236 and parameters: {'C': 122.19110570860134}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,683] Trial 20 finished with value: -0.4366812227074236 and parameters: {'C': 857.1261176520713}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,743] Trial 21 finished with value: -0.4410480349344978 and parameters: {'C': 64.81142117476604}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,816] Trial 22 finished with value: -0.4410480349344978 and parameters: {'C': 142.40093973968894}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,873] Trial 23 finished with value: -0.4366812227074236 and parameters: {'C': 34.528114945517494}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,935] Trial 24 finished with value: -0.4410480349344978 and parameters: {'C': 137.05385807379568}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:24,996] Trial 25 finished with value: -0.4410480349344978 and parameters: {'C': 27.816942845145707}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,054] Trial 26 finished with value: -0.4410480349344978 and parameters: {'C': 6.0294649292622875}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,115] Trial 27 finished with value: -0.43231441048034935 and parameters: {'C': 219.7490151564448}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,178] Trial 28 finished with value: -0.4279475982532751 and parameters: {'C': 68.3718738651336}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,237] Trial 29 finished with value: -0.43231441048034935 and parameters: {'C': 323.09128388290645}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,301] Trial 30 finished with value: -0.4410480349344978 and parameters: {'C': 19.746434873234133}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,365] Trial 31 finished with value: -0.4366812227074236 and parameters: {'C': 57.97169591509047}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,426] Trial 32 finished with value: -0.43231441048034935 and parameters: {'C': 32.22834447840863}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,491] Trial 33 finished with value: -0.4410480349344978 and parameters: {'C': 90.51568647203396}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,553] Trial 34 finished with value: -0.4366812227074236 and parameters: {'C': 918.7706171648769}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,625] Trial 35 finished with value: -0.4366812227074236 and parameters: {'C': 239.16544726250117}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,694] Trial 36 finished with value: -0.42358078602620086 and parameters: {'C': 2.8901320904631183}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,767] Trial 37 finished with value: -0.4410480349344978 and parameters: {'C': 9.66954496160761}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,832] Trial 38 finished with value: -0.4410480349344978 and parameters: {'C': 19.866970095945266}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,892] Trial 39 finished with value: -0.4410480349344978 and parameters: {'C': 300.3927076243865}. Best is trial 4 with value: -0.4410480349344978.
[I 2023-08-23 10:04:25,952] Trial 40 finished with value: -0.44541484716157204 and parameters: {'C': 47.55993975202472}. Best is trial 40 with value: -0.44541484716157204.
[I 2023-08-23 10:04:26,007] Trial 41 finished with value: -0.4410480349344978 and parameters: {'C': 60.71643524262415}. Best is trial 40 with value: -0.44541484716157204.
[I 2023-08-23 10:04:26,070] Trial 42 finished with value: -0.4366812227074236 and parameters: {'C': 36.60388894360864}. Best is trial 40 with value: -0.44541484716157204.
[I 2023-08-23 10:04:26,123] Trial 43 finished with value: -0.4366812227074236 and parameters: {'C': 439.1770649465081}. Best is trial 40 with value: -0.44541484716157204.
[I 2023-08-23 10:04:26,178] Trial 44 finished with value: -0.4104803493449782 and parameters: {'C': 0.7000095791417738}. Best is trial 40 with value: -0.44541484716157204.
[I 2023-08-23 10:04:26,233] Trial 45 finished with value: -0.44541484716157204 and parameters: {'C': 148.46477074303922}. Best is trial 40 with value: -0.44541484716157204.
[I 2023-08-23 10:04:26,294] Trial 46 finished with value: -0.4410480349344978 and parameters: {'C': 182.9495186353782}. Best is trial 40 with value: -0.44541484716157204.
[I 2023-08-23 10:04:26,353] Trial 47 finished with value: -0.4366812227074236 and parameters: {'C': 387.31240219400513}. Best is trial 40 with value: -0.44541484716157204.
[I 2023-08-23 10:04:26,414] Trial 48 finished with value: -0.4497816593886463 and parameters: {'C': 99.21519297775365}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,476] Trial 49 finished with value: -0.4366812227074236 and parameters: {'C': 115.46135655334533}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,537] Trial 50 finished with value: -0.4366812227074236 and parameters: {'C': 458.56435707102986}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,602] Trial 51 finished with value: -0.4410480349344978 and parameters: {'C': 73.58649748733922}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,663] Trial 52 finished with value: -0.4366812227074236 and parameters: {'C': 200.2788088272584}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,725] Trial 53 finished with value: -0.4366812227074236 and parameters: {'C': 19.31825499286287}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,784] Trial 54 finished with value: -0.44541484716157204 and parameters: {'C': 106.59084996363067}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,855] Trial 55 finished with value: -0.44541484716157204 and parameters: {'C': 106.64822615103826}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,923] Trial 56 finished with value: -0.44541484716157204 and parameters: {'C': 89.16823840432664}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:26,980] Trial 57 finished with value: -0.4497816593886463 and parameters: {'C': 104.59775573851961}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,043] Trial 58 finished with value: -0.4410480349344978 and parameters: {'C': 174.29811362017236}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,103] Trial 59 finished with value: -0.44541484716157204 and parameters: {'C': 605.322235224299}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,159] Trial 60 finished with value: -0.4366812227074236 and parameters: {'C': 48.17208905362385}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,220] Trial 61 finished with value: -0.4410480349344978 and parameters: {'C': 102.41802787176485}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,283] Trial 62 finished with value: -0.4410480349344978 and parameters: {'C': 96.55836052279399}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,340] Trial 63 finished with value: -0.4410480349344978 and parameters: {'C': 154.1357677706352}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,398] Trial 64 finished with value: -0.4410480349344978 and parameters: {'C': 296.31010811769926}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,463] Trial 65 finished with value: -0.4366812227074236 and parameters: {'C': 44.935214763221275}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,527] Trial 66 finished with value: -0.4410480349344978 and parameters: {'C': 73.16390566903921}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,588] Trial 67 finished with value: -0.44541484716157204 and parameters: {'C': 153.71738069417808}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,649] Trial 68 finished with value: -0.4410480349344978 and parameters: {'C': 23.756253435815577}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,719] Trial 69 finished with value: -0.4366812227074236 and parameters: {'C': 38.69611578432412}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,779] Trial 70 finished with value: -0.44541484716157204 and parameters: {'C': 87.67807800609941}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,842] Trial 71 finished with value: -0.4366812227074236 and parameters: {'C': 684.1229164129188}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,910] Trial 72 finished with value: -0.4366812227074236 and parameters: {'C': 499.95419630352836}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:27,974] Trial 73 finished with value: -0.4410480349344978 and parameters: {'C': 262.5127841384856}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,038] Trial 74 finished with value: -0.4410480349344978 and parameters: {'C': 627.2467565203034}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,101] Trial 75 finished with value: -0.4366812227074236 and parameters: {'C': 890.2576616856232}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,157] Trial 76 finished with value: -0.4410480349344978 and parameters: {'C': 196.98915218851664}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,210] Trial 77 finished with value: -0.4366812227074236 and parameters: {'C': 338.5255290619542}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,270] Trial 78 finished with value: -0.4366812227074236 and parameters: {'C': 113.42008456534448}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,324] Trial 79 finished with value: -0.4366812227074236 and parameters: {'C': 48.25026058779677}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,382] Trial 80 finished with value: -0.4410480349344978 and parameters: {'C': 264.91520318692767}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,438] Trial 81 finished with value: -0.44541484716157204 and parameters: {'C': 165.01970957280605}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,494] Trial 82 finished with value: -0.44541484716157204 and parameters: {'C': 143.1178660299481}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,554] Trial 83 finished with value: -0.44541484716157204 and parameters: {'C': 535.8228043376868}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,622] Trial 84 finished with value: -0.4410480349344978 and parameters: {'C': 71.76663211194828}. Best is trial 48 with value: -0.4497816593886463.
```

```
[I 2023-08-23 10:04:28,676] Trial 85 finished with value: -0.4410480349344978 and parameters: {'C': 27.89494033235841}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,733] Trial 86 finished with value: -0.44541484716157204 and parameters: {'C': 127.46072050469203}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,786] Trial 87 finished with value: -0.4366812227074236 and parameters: {'C': 422.4686384535481}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,838] Trial 88 finished with value: -0.4366812227074236 and parameters: {'C': 222.51369154068533}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,893] Trial 89 finished with value: -0.4410480349344978 and parameters: {'C': 89.31280514167366}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:28,950] Trial 90 finished with value: -0.4410480349344978 and parameters: {'C': 70.30948005170568}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,008] Trial 91 finished with value: -0.4366812227074236 and parameters: {'C': 56.21177522350271}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,062] Trial 92 finished with value: -0.4410480349344978 and parameters: {'C': 86.90924279043828}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,126] Trial 93 finished with value: -0.4410480349344978 and parameters: {'C': 128.8735161249421}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,192] Trial 94 finished with value: -0.43231441048034935 and parameters: {'C': 365.39073428390617}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,261] Trial 95 finished with value: -0.4366812227074236 and parameters: {'C': 220.53881635051187}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,322] Trial 96 finished with value: -0.4410480349344978 and parameters: {'C': 42.19052099521887}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,376] Trial 97 finished with value: -0.44541484716157204 and parameters: {'C': 91.25685395927937}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,438] Trial 98 finished with value: -0.44541484716157204 and parameters: {'C': 163.62495478899328}. Best is trial 48 with value: -0.4497816593886463.
[I 2023-08-23 10:04:29,497] Trial 99 finished with value: -0.4410480349344978 and parameters: {'C': 109.20858495900454}. Best is trial 48 with value: -0.4497816593886463.
```

In [76]:

```
print("Linear Regression - Best trial:")
print(study_lr.best_trial.params)
print("Accuracy:", -study_lr.best_value)
```

```
Linear Regression - Best trial:
{'C': 92.7111324133625}
Accuracy: 0.4497816593886463
```

In [77]:

```
print("X_train_scaled shape:", X_train_scaled.shape)
print("y_train_resampled shape:", y_train_resampled.shape)
print("X_test_scaled shape:", X_test_scaled.shape)
```

```
X_train_scaled shape: (914, 12)
y_train_resampled shape: (2322,)
X_test_scaled shape: (229, 12)
```

In [94]:

```
best_lr_params = study_lr.best_params
best_lr_model = LogisticRegression(**best_lr_params, random_state=42)
best_lr_model.fit(X_train_resampled, y_train_resampled)
```

Out[94]:

```
▾ LogisticRegression
LogisticRegression(C=99.21519297775365, random_state=42)
```

In [93]:

```
final_lr_preds = best_lr_model.predict(X_test_scaled)
```

In [47]:

```
from sklearn.metrics import classification_report
lr_classification_report = classification_report(y_test, final_lr_preds)
print("Linear Regression Classification Report:\n", lr_classification_report)
```

```
Linear Regression Classification Report:
               precision    recall  f1-score   support

           3       0.00      0.00      0.00         0
           4       0.03      0.17      0.05         6
           5       0.64      0.49      0.56        96
           6       0.60      0.43      0.50        99
           7       0.34      0.38      0.36        26
           8       0.11      1.00      0.20         2

    accuracy                           0.45       229
   macro avg       0.29      0.41      0.28       229
weighted avg       0.57      0.45      0.50       229
```

In [ ]: