```
In [76]: import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns

         import sklearn
         from sklearn.model_selection import train_test_split
         from sklearn.neighbors import KNeighborsRegressor
         from sklearn.linear_model import LinearRegression
         from sklearn.tree import DecisionTreeRegressor
         from sklearn.metrics import mean_squared_error,mean_absolute_error,r2_score

         from sklearn.metrics import confusion_matrix,accuracy_score,classification_repor
         import warnings
         warnings.filterwarnings('ignore')
```

```
In [10]: df=pd.read_csv('loan.csv')
         df
```

Out[10]:

	age	campaign	pdays	previous	no_previous_contact	not_working	job_admin.
0	56	1	999	0	1	0	0
1	57	1	999	0	1	0	0
2	37	1	999	0	1	0	0
3	40	1	999	0	1	0	1
4	56	1	999	0	1	0	0
...	...	...	...	...	...	...	...
41183	73	1	999	0	1	1	0
41184	46	1	999	0	1	0	0
41185	56	2	999	0	1	1	0
41186	44	1	999	0	1	0	0
41187	74	3	999	1	1	1	0

41188 rows × 60 columns

```
In [6]: df.describe()
```

Out[6]:

	age	campaign	pdays	previous	no_previous_contact	no
count	41188.00000	41188.000000	41188.000000	41188.000000	41188.000000	411
mean	40.02406	2.567593	962.475454	0.172963	0.963217	
std	10.42125	2.770014	186.910907	0.494901	0.188230	
min	17.00000	1.000000	0.000000	0.000000	0.000000	
25%	32.00000	1.000000	999.000000	0.000000	1.000000	
50%	38.00000	2.000000	999.000000	0.000000	1.000000	
75%	47.00000	3.000000	999.000000	0.000000	1.000000	
max	98.00000	56.000000	999.000000	7.000000	1.000000	

8 rows × 60 columns

```
In [9]: df.shape
```

Out[9]: (41188, 60)

```
In [10]: df.columns
```

Out[10]: Index(['age', 'campaign', 'pdays', 'previous', 'no_previous_contact',
       'not_working', 'job_admin.', 'job_blue-collar', 'job_entrepreneur',
       'job_housemaid', 'job_management', 'job_retired', 'job_self-employed',
       'job_services', 'job_student', 'job_technician', 'job_unemployed',
       'job_unknown', 'marital_divorced', 'marital_married', 'marital_single',
       'marital_unknown', 'education_basic.4y', 'education_basic.6y',
       'education_basic.9y', 'education_high.school', 'education_illiterate',
       'education_professional.course', 'education_university.degree',
       'education_unknown', 'default_no', 'default_unknown', 'default_yes',
       'housing_no', 'housing_unknown', 'housing_yes', 'loan_no',
       'loan_unknown', 'loan_yes', 'contact_cellular', 'contact_telephone',
       'month_apr', 'month_aug', 'month_dec', 'month_jul', 'month_jun',
       'month_mar', 'month_may', 'month_nov', 'month_oct', 'month_sep',
       'day_of_week_fri', 'day_of_week_mon', 'day_of_week_thu',
       'day_of_week_tue', 'day_of_week_wed', 'poutcome_failure',
       'poutcome_nonexistent', 'poutcome_success', 'Loan_Status_label'],
      dtype='object')

```
In [11]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 41188 entries, 0 to 41187
Data columns (total 60 columns):
 #   Column                        Non-Null Count  Dtype
---  ------                        --------------  -----
 0   age                           41188 non-null  int64
 1   campaign                      41188 non-null  int64
 2   pdays                         41188 non-null  int64
 3   previous                      41188 non-null  int64
 4   no_previous_contact           41188 non-null  int64
 5   not_working                   41188 non-null  int64
 6   job_admin.                    41188 non-null  int64
 7   job_blue-collar               41188 non-null  int64
 8   job_entrepreneur              41188 non-null  int64
 9   job_housemaid                 41188 non-null  int64
 10  job_management                41188 non-null  int64
 11  job_retired                   41188 non-null  int64
 12  job_self-employed             41188 non-null  int64
 13  job_services                  41188 non-null  int64
 14  job_student                   41188 non-null  int64
 15  job_technician                41188 non-null  int64
 16  job_unemployed                41188 non-null  int64
 17  job_unknown                   41188 non-null  int64
 18  marital_divorced              41188 non-null  int64
 19  marital_married               41188 non-null  int64
 20  marital_single                41188 non-null  int64
 21  marital_unknown               41188 non-null  int64
 22  education_basic.4y            41188 non-null  int64
 23  education_basic.6y            41188 non-null  int64
 24  education_basic.9y            41188 non-null  int64
 25  education_high.school         41188 non-null  int64
 26  education_illiterate          41188 non-null  int64
 27  education_professional.course 41188 non-null  int64
 28  education_university.degree   41188 non-null  int64
 29  education_unknown             41188 non-null  int64
 30  default_no                    41188 non-null  int64
 31  default_unknown               41188 non-null  int64
 32  default_yes                   41188 non-null  int64
 33  housing_no                    41188 non-null  int64
 34  housing_unknown               41188 non-null  int64
 35  housing_yes                   41188 non-null  int64
 36  loan_no                       41188 non-null  int64
 37  loan_unknown                  41188 non-null  int64
 38  loan_yes                      41188 non-null  int64
 39  contact_cellular              41188 non-null  int64
 40  contact_telephone             41188 non-null  int64
 41  month_apr                     41188 non-null  int64
 42  month_aug                     41188 non-null  int64
 43  month_dec                     41188 non-null  int64
 44  month_jul                     41188 non-null  int64
 45  month_jun                     41188 non-null  int64
 46  month_mar                     41188 non-null  int64
 47  month_may                     41188 non-null  int64
 48  month_nov                     41188 non-null  int64
 49  month_oct                     41188 non-null  int64
 50  month_sep                     41188 non-null  int64
 51  day_of_week_fri               41188 non-null  int64
 52  day_of_week_mon               41188 non-null  int64
 53  day_of_week_thu               41188 non-null  int64
 54  day_of_week_tue               41188 non-null  int64
```

```
 55   day_of_week_wed              41188 non-null   int64
 56   poutcome_failure             41188 non-null   int64
 57   poutcome_nonexistent         41188 non-null   int64
 58   poutcome_success             41188 non-null   int64
 59   Loan_Status_label            41188 non-null   int64
dtypes: int64(60)
memory usage: 18.9 MB
```

In [12]: `df.describe()`

Out[12]:

	age	campaign	pdays	previous	no_previous_contact	no
count	41188.00000	41188.000000	41188.000000	41188.000000	41188.000000	411
mean	40.02406	2.567593	962.475454	0.172963	0.963217	
std	10.42125	2.770014	186.910907	0.494901	0.188230	
min	17.00000	1.000000	0.000000	0.000000	0.000000	
25%	32.00000	1.000000	999.000000	0.000000	1.000000	
50%	38.00000	2.000000	999.000000	0.000000	1.000000	
75%	47.00000	3.000000	999.000000	0.000000	1.000000	
max	98.00000	56.000000	999.000000	7.000000	1.000000	

8 rows × 60 columns

In [15]: `df['Loan_Status_label'].value_counts()`

Out[15]:
```
Loan_Status_label
0    36548
1     4640
Name: count, dtype: int64
```

In [20]:
```
print(round(len(df[df['Loan_Status_label']==0])/len(df),2))
print(round(len(df[df['Loan_Status_label']==1])/len(df),2))
```
```
0.89
0.11
```

In [21]: `df.isnull().sum()`

```
Out[21]:  age                            0
          campaign                       0
          pdays                          0
          previous                       0
          no_previous_contact            0
          not_working                    0
          job_admin.                     0
          job_blue-collar                0
          job_entrepreneur               0
          job_housemaid                  0
          job_management                 0
          job_retired                    0
          job_self-employed              0
          job_services                   0
          job_student                    0
          job_technician                 0
          job_unemployed                 0
          job_unknown                    0
          marital_divorced               0
          marital_married                0
          marital_single                 0
          marital_unknown                0
          education_basic.4y             0
          education_basic.6y             0
          education_basic.9y             0
          education_high.school          0
          education_illiterate           0
          education_professional.course  0
          education_university.degree    0
          education_unknown              0
          default_no                     0
          default_unknown                0
          default_yes                    0
          housing_no                     0
          housing_unknown                0
          housing_yes                    0
          loan_no                        0
          loan_unknown                   0
          loan_yes                       0
          contact_cellular               0
          contact_telephone              0
          month_apr                      0
          month_aug                      0
          month_dec                      0
          month_jul                      0
          month_jun                      0
          month_mar                      0
          month_may                      0
          month_nov                      0
          month_oct                      0
          month_sep                      0
          day_of_week_fri                0
          day_of_week_mon                0
          day_of_week_thu                0
          day_of_week_tue                0
          day_of_week_wed                0
          poutcome_failure               0
          poutcome_nonexistent           0
          poutcome_success               0
```
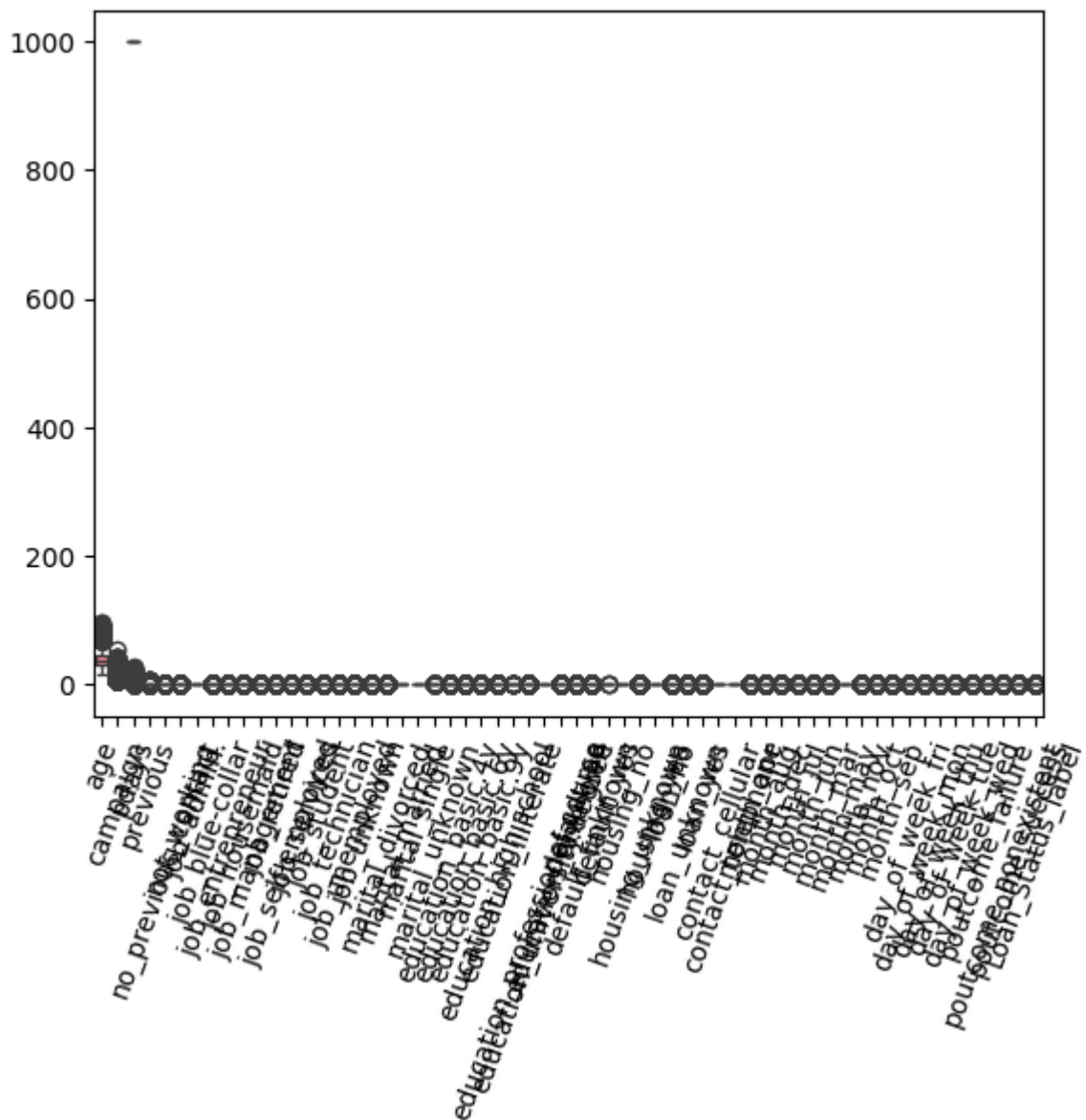
```
Loan_Status_label                    0
dtype: int64
```

In [34]:
```python
sns.boxplot(df)
plt.xticks(rotation =70)
plt.show()
```


In [38]:
```python
Q1 = df.quantile(0.25)
Q3 = df.quantile(0.75)
IQR = Q3 - Q1
print(f'IQR : \n\n{IQR}')
```

```
IQR :

age                              15.0
campaign                          2.0
pdays                             0.0
previous                          0.0
no_previous_contact               0.0
not_working                       0.0
job_admin.                        1.0
job_blue-collar                   0.0
job_entrepreneur                  0.0
job_housemaid                     0.0
job_management                    0.0
job_retired                       0.0
job_self-employed                 0.0
job_services                      0.0
job_student                       0.0
job_technician                    0.0
job_unemployed                    0.0
job_unknown                       0.0
marital_divorced                  0.0
marital_married                   1.0
marital_single                    1.0
marital_unknown                   0.0
education_basic.4y                0.0
education_basic.6y                0.0
education_basic.9y                0.0
education_high.school             0.0
education_illiterate              0.0
education_professional.course     0.0
education_university.degree       1.0
education_unknown                 0.0
default_no                        0.0
default_unknown                   0.0
default_yes                       0.0
housing_no                        1.0
housing_unknown                   0.0
housing_yes                       1.0
loan_no                           0.0
loan_unknown                      0.0
loan_yes                          0.0
contact_cellular                  1.0
contact_telephone                 1.0
month_apr                         0.0
month_aug                         0.0
month_dec                         0.0
month_jul                         0.0
month_jun                         0.0
month_mar                         0.0
month_may                         1.0
month_nov                         0.0
month_oct                         0.0
month_sep                         0.0
day_of_week_fri                   0.0
day_of_week_mon                   0.0
day_of_week_thu                   0.0
day_of_week_tue                   0.0
day_of_week_wed                   0.0
poutcome_failure                  0.0
poutcome_nonexistent              0.0
```

```
        poutcome_success              0.0
        Loan_Status_label             0.0
        dtype: float64
```

In [41]:
```python
lower_bound= Q1 - 1.5*IQR
upper_bound= Q3 +1.5*IQR
print(f'Lower bound for outliers: \n\n{lower_bound}\n\n')
print(f'Upper bound for outliers :\n\n{upper_bound}n\n')
```

```
Lower bound for outliers:

age                              9.5
campaign                        -2.0
pdays                          999.0
previous                         0.0
no_previous_contact              1.0
not_working                      0.0
job_admin.                      -1.5
job_blue-collar                  0.0
job_entrepreneur                 0.0
job_housemaid                    0.0
job_management                   0.0
job_retired                      0.0
job_self-employed                0.0
job_services                     0.0
job_student                      0.0
job_technician                   0.0
job_unemployed                   0.0
job_unknown                      0.0
marital_divorced                 0.0
marital_married                 -1.5
marital_single                  -1.5
marital_unknown                  0.0
education_basic.4y               0.0
education_basic.6y               0.0
education_basic.9y               0.0
education_high.school            0.0
education_illiterate             0.0
education_professional.course    0.0
education_university.degree     -1.5
education_unknown                0.0
default_no                       1.0
default_unknown                  0.0
default_yes                      0.0
housing_no                      -1.5
housing_unknown                  0.0
housing_yes                     -1.5
loan_no                          1.0
loan_unknown                     0.0
loan_yes                         0.0
contact_cellular                -1.5
contact_telephone               -1.5
month_apr                        0.0
month_aug                        0.0
month_dec                        0.0
month_jul                        0.0
month_jun                        0.0
month_mar                        0.0
month_may                       -1.5
month_nov                        0.0
month_oct                        0.0
month_sep                        0.0
day_of_week_fri                  0.0
day_of_week_mon                  0.0
day_of_week_thu                  0.0
day_of_week_tue                  0.0
day_of_week_wed                  0.0
poutcome_failure                 0.0
poutcome_nonexistent             1.0
```

```
poutcome_success                    0.0
Loan_Status_label                   0.0
dtype: float64


Upper bound for outliers :

age                                69.5
campaign                            6.0
pdays                             999.0
previous                            0.0
no_previous_contact                 1.0
not_working                         0.0
job_admin.                          2.5
job_blue-collar                     0.0
job_entrepreneur                    0.0
job_housemaid                       0.0
job_management                      0.0
job_retired                         0.0
job_self-employed                   0.0
job_services                        0.0
job_student                         0.0
job_technician                      0.0
job_unemployed                      0.0
job_unknown                         0.0
marital_divorced                    0.0
marital_married                     2.5
marital_single                      2.5
marital_unknown                     0.0
education_basic.4y                  0.0
education_basic.6y                  0.0
education_basic.9y                  0.0
education_high.school               0.0
education_illiterate                0.0
education_professional.course       0.0
education_university.degree         2.5
education_unknown                   0.0
default_no                          1.0
default_unknown                     0.0
default_yes                         0.0
housing_no                          2.5
housing_unknown                     0.0
housing_yes                         2.5
loan_no                             1.0
loan_unknown                        0.0
loan_yes                            0.0
contact_cellular                    2.5
contact_telephone                   2.5
month_apr                           0.0
month_aug                           0.0
month_dec                           0.0
month_jul                           0.0
month_jun                           0.0
month_mar                           0.0
month_may                           2.5
month_nov                           0.0
month_oct                           0.0
month_sep                           0.0
day_of_week_fri                     0.0
day_of_week_mon                     0.0
```

```
day_of_week_thu          0.0
day_of_week_tue          0.0
day_of_week_wed          0.0
poutcome_failure         0.0
poutcome_nonexistent     1.0
poutcome_success         0.0
Loan_Status_label        0.0
dtype: float64n
```

In [42]: `df.shape`

Out[42]: (41188, 60)

In [43]: `df[((df<lower_bound) | (df>upper_bound)).any(axis=1)]`

Out[43]:

	age	campaign	pdays	previous	no_previous_contact	not_working	job_admin.
0	56	1	999	0	1	0	0
1	57	1	999	0	1	0	0
2	37	1	999	0	1	0	0
3	40	1	999	0	1	0	1
4	56	1	999	0	1	0	0
...	...	...	...	...	...	...	...
41183	73	1	999	0	1	1	0
41184	46	1	999	0	1	0	0
41185	56	2	999	0	1	1	0
41186	44	1	999	0	1	0	0
41187	74	3	999	1	1	1	0

41188 rows × 60 columns

In [44]:
```
print(lower_bound)
print(upper_bound)
```

```
age                             9.5
campaign                       -2.0
pdays                         999.0
previous                        0.0
no_previous_contact             1.0
not_working                     0.0
job_admin.                     -1.5
job_blue-collar                 0.0
job_entrepreneur                0.0
job_housemaid                   0.0
job_management                  0.0
job_retired                     0.0
job_self-employed               0.0
job_services                    0.0
job_student                     0.0
job_technician                  0.0
job_unemployed                  0.0
job_unknown                     0.0
marital_divorced                0.0
marital_married                -1.5
marital_single                 -1.5
marital_unknown                 0.0
education_basic.4y              0.0
education_basic.6y              0.0
education_basic.9y              0.0
education_high.school           0.0
education_illiterate            0.0
education_professional.course   0.0
education_university.degree    -1.5
education_unknown               0.0
default_no                      1.0
default_unknown                 0.0
default_yes                     0.0
housing_no                     -1.5
housing_unknown                 0.0
housing_yes                    -1.5
loan_no                         1.0
loan_unknown                    0.0
loan_yes                        0.0
contact_cellular               -1.5
contact_telephone              -1.5
month_apr                       0.0
month_aug                       0.0
month_dec                       0.0
month_jul                       0.0
month_jun                       0.0
month_mar                       0.0
month_may                      -1.5
month_nov                       0.0
month_oct                       0.0
month_sep                       0.0
day_of_week_fri                 0.0
day_of_week_mon                 0.0
day_of_week_thu                 0.0
day_of_week_tue                 0.0
day_of_week_wed                 0.0
poutcome_failure                0.0
poutcome_nonexistent            1.0
poutcome_success                0.0
Loan_Status_label               0.0
```

```
dtype: float64
age                             69.5
campaign                         6.0
pdays                          999.0
previous                         0.0
no_previous_contact              1.0
not_working                      0.0
job_admin.                       2.5
job_blue-collar                  0.0
job_entrepreneur                 0.0
job_housemaid                    0.0
job_management                   0.0
job_retired                      0.0
job_self-employed                0.0
job_services                     0.0
job_student                      0.0
job_technician                   0.0
job_unemployed                   0.0
job_unknown                      0.0
marital_divorced                 0.0
marital_married                  2.5
marital_single                   2.5
marital_unknown                  0.0
education_basic.4y               0.0
education_basic.6y               0.0
education_basic.9y               0.0
education_high.school            0.0
education_illiterate             0.0
education_professional.course    0.0
education_university.degree      2.5
education_unknown                0.0
default_no                       1.0
default_unknown                  0.0
default_yes                      0.0
housing_no                       2.5
housing_unknown                  0.0
housing_yes                      2.5
loan_no                          1.0
loan_unknown                     0.0
loan_yes                         0.0
contact_cellular                 2.5
contact_telephone                2.5
month_apr                        0.0
month_aug                        0.0
month_dec                        0.0
month_jul                        0.0
month_jun                        0.0
month_mar                        0.0
month_may                        2.5
month_nov                        0.0
month_oct                        0.0
month_sep                        0.0
day_of_week_fri                  0.0
day_of_week_mon                  0.0
day_of_week_thu                  0.0
day_of_week_tue                  0.0
day_of_week_wed                  0.0
poutcome_failure                 0.0
poutcome_nonexistent             1.0
poutcome_success                 0.0
```

```
Loan_Status_label              0.0
dtype: float64
```

In [45]: `((df<lower_bound) | (df>upper_bound)).sum()`

```
Out[45]:  age                            469
          campaign                      2406
          pdays                         1515
          previous                      5625
          no_previous_contact           1515
          not_working                   3609
          job_admin.                       0
          job_blue-collar               9254
          job_entrepreneur              1456
          job_housemaid                 1060
          job_management                2924
          job_retired                   1720
          job_self-employed             1421
          job_services                  3969
          job_student                    875
          job_technician                6743
          job_unemployed                1014
          job_unknown                    330
          marital_divorced              4612
          marital_married                  0
          marital_single                   0
          marital_unknown                 80
          education_basic.4y            4176
          education_basic.6y            2292
          education_basic.9y            6045
          education_high.school         9515
          education_illiterate            18
          education_professional.course 5243
          education_university.degree      0
          education_unknown             1731
          default_no                    8600
          default_unknown               8597
          default_yes                      3
          housing_no                       0
          housing_unknown                990
          housing_yes                      0
          loan_no                       7238
          loan_unknown                   990
          loan_yes                      6248
          contact_cellular                 0
          contact_telephone                0
          month_apr                     2632
          month_aug                     6178
          month_dec                      182
          month_jul                     7174
          month_jun                     5318
          month_mar                      546
          month_may                        0
          month_nov                     4101
          month_oct                      718
          month_sep                      570
          day_of_week_fri               7827
          day_of_week_mon               8514
          day_of_week_thu               8623
          day_of_week_tue               8090
          day_of_week_wed               8134
          poutcome_failure              4252
          poutcome_nonexistent          5625
          poutcome_success              1373
```

```
Loan_Status_label                4640
dtype: int64
```

In [46]:
```python
outliers = df[((df<lower_bound)|(df>upper_bound)).any(axis =1)]
print(outliers.shape)
```

(41188, 60)

In [47]:
```python
df_no_outliers=df[~((df<lower_bound)|(df>upper_bound)).any(axis=1)]
df_no_outliers.shape
```

Out[47]: (0, 60)

In [48]:
```python
df.corr()
```

Out[48]:

	age	campaign	pdays	previous	no_previous_cc
age	1.000000	0.004594	-0.034369	0.024365	-0.0
campaign	0.004594	1.000000	0.052584	-0.079141	0.0
pdays	-0.034369	0.052584	1.000000	-0.587514	0.9
previous	0.024365	-0.079141	-0.587514	1.000000	-0.5
no_previous_contact	-0.034292	0.052569	0.999992	-0.587462	1.0
not_working	0.207547	-0.017529	-0.112797	0.104066	-0.1
job_admin.	-0.102584	0.011745	-0.025398	0.018888	-0.0
job_blue-collar	-0.024191	-0.001775	0.065335	-0.054845	0.0
job_entrepreneur	0.031213	-0.002203	0.019246	-0.013239	0.0
job_housemaid	0.085403	0.004226	-0.001649	-0.011569	-0.0
job_management	0.062040	-0.009135	0.000254	0.006735	0.0
job_retired	0.440772	-0.006847	-0.072084	0.065113	-0.0
job_self-employed	-0.001356	0.006361	0.014307	-0.011231	0.0
job_services	-0.065731	0.002383	0.030573	-0.011885	0.0
job_student	-0.199749	-0.024657	-0.096367	0.104671	-0.0
job_technician	-0.064383	0.001557	0.004575	-0.017142	0.0
job_unemployed	-0.004426	-0.000200	-0.023085	0.008426	-0.0
job_unknown	0.047773	0.002625	-0.011419	-0.003345	-0.0
marital_divorced	0.166125	0.005872	0.011711	-0.003066	0.0
marital_married	0.271266	0.002332	0.031614	-0.043420	0.0
marital_single	-0.411703	-0.007624	-0.042015	0.048485	-0.0
marital_unknown	0.001062	0.009873	-0.005989	0.009095	-0.0
education_basic.4y	0.244079	0.004000	0.005804	-0.017601	0.0
education_basic.6y	0.009897	-0.000991	0.023238	-0.021493	0.0
education_basic.9y	-0.038320	-0.005278	0.036963	-0.026418	0.0
education_high.school	-0.106550	0.000195	0.005522	0.014347	0.0
education_illiterate	0.017007	-0.002188	-0.002086	-0.002613	-0.0
education_professional.course	0.002054	0.002554	-0.003493	-0.007631	-0.0
education_university.degree	-0.071138	-0.000950	-0.036958	0.025419	-0.0
education_unknown	0.069485	0.002162	-0.022014	0.022641	-0.0
default_no	-0.165019	-0.032923	-0.080072	0.102533	-0.0
default_unknown	0.165001	0.033007	0.080047	-0.102604	0.0

	age	campaign	pdays	previous	no_previous_cc
default_yes	0.001891	-0.003803	0.001668	0.002766	0.0
housing_no	0.001755	0.010959	0.010131	-0.021677	0.0
housing_unknown	-0.001092	-0.000396	0.002033	0.003769	0.0
housing_yes	-0.001414	-0.010800	-0.010719	0.020446	-0.0
loan_no	0.007224	-0.004831	-0.000737	0.000504	-0.0
loan_unknown	-0.001092	-0.000396	0.002033	0.003769	0.0
loan_yes	-0.007198	0.005294	-0.000086	-0.002143	-0.0
contact_cellular	-0.007021	-0.077368	-0.117970	0.212848	-0.1
contact_telephone	0.007021	0.077368	0.117970	-0.212848	0.1
month_apr	0.014285	-0.058519	-0.009629	0.079384	-0.0
month_aug	0.065309	0.015009	-0.003176	-0.048852	-0.0
month_dec	0.046563	-0.009819	-0.076425	0.059557	-0.0
month_jul	-0.043563	0.098385	0.049695	-0.116785	0.0
month_jun	-0.009532	0.070144	0.015957	-0.069905	0.0
month_mar	0.008390	-0.015244	-0.078769	0.072733	-0.0
month_may	-0.067523	-0.032704	0.070690	-0.012534	0.0
month_nov	0.030045	-0.079028	-0.016946	0.081050	-0.0
month_oct	0.053008	-0.049126	-0.128677	0.127754	-0.1
month_sep	0.040889	-0.033953	-0.154629	0.157651	-0.1
day_of_week_fri	0.007148	0.026641	0.013782	0.004404	0.0
day_of_week_mon	0.019027	0.014405	0.001031	-0.002012	0.0
day_of_week_thu	-0.021356	0.005141	-0.008186	0.001512	-0.0
day_of_week_tue	0.019137	-0.025125	-0.007300	0.000090	-0.0
day_of_week_wed	-0.023671	-0.021088	0.001021	-0.003929	0.0
poutcome_failure	-0.002537	-0.068890	0.006354	0.682608	0.0
poutcome_nonexistent	-0.016376	0.087645	0.491353	-0.878776	0.4
poutcome_success	0.035626	-0.050893	-0.950700	0.524045	-0.9
Loan_Status_label	0.030399	-0.066357	-0.324914	0.230181	-0.3

60 rows × 60 columns

```
In [53]: df.corr()['Loan_Status_label']
```

```
Out[53]:  age                              0.030399
          campaign                        -0.066357
          pdays                           -0.324914
          previous                         0.230181
          no_previous_contact             -0.324877
          not_working                      0.121246
          job_admin.                       0.031426
          job_blue-collar                 -0.074423
          job_entrepreneur                -0.016644
          job_housemaid                   -0.006505
          job_management                  -0.000419
          job_retired                      0.092221
          job_self-employed               -0.004663
          job_services                    -0.032301
          job_student                      0.093955
          job_technician                  -0.006149
          job_unemployed                   0.014752
          job_unknown                     -0.000151
          marital_divorced                -0.010608
          marital_married                 -0.043398
          marital_single                   0.054133
          marital_unknown                  0.005211
          education_basic.4y              -0.010798
          education_basic.6y              -0.023517
          education_basic.9y              -0.045135
          education_high.school           -0.007452
          education_illiterate             0.007246
          education_professional.course   0.001003
          education_university.degree      0.050364
          education_unknown                0.021430
          default_no                       0.099344
          default_unknown                 -0.099293
          default_yes                     -0.003041
          housing_no                      -0.011085
          housing_unknown                 -0.002270
          housing_yes                      0.011743
          loan_no                          0.005123
          loan_unknown                    -0.002270
          loan_yes                        -0.004466
          contact_cellular                 0.144773
          contact_telephone               -0.144773
          month_apr                        0.076136
          month_aug                       -0.008813
          month_dec                        0.079303
          month_jul                       -0.032230
          month_jun                       -0.009182
          month_mar                        0.144014
          month_may                       -0.108271
          month_nov                       -0.011796
          month_oct                        0.137366
          month_sep                        0.126067
          day_of_week_fri                 -0.006996
          day_of_week_mon                 -0.021265
          day_of_week_thu                  0.013888
          day_of_week_tue                  0.008046
          day_of_week_wed                  0.006302
          poutcome_failure                 0.031799
          poutcome_nonexistent            -0.193507
          poutcome_success                 0.316269
```

```
       Loan_Status_label                    1.000000
       Name: Loan_Status_label, dtype: float64
```

In [11]:
```python
x=df.drop(columns=['Loan_Status_label'],axis=1)
y=df['Loan_Status_label']
```

In [56]:
```python
x
```

Out[56]:

	age	campaign	pdays	previous	no_previous_contact	not_working	job_admin.
0	56	1	999	0	1	0	0
1	57	1	999	0	1	0	0
2	37	1	999	0	1	0	0
3	40	1	999	0	1	0	1
4	56	1	999	0	1	0	0
...	...	...	...	...	...	...	...
41183	73	1	999	0	1	1	0
41184	46	1	999	0	1	0	0
41185	56	2	999	0	1	1	0
41186	44	1	999	0	1	0	0
41187	74	3	999	1	1	1	0

41188 rows × 59 columns

In [57]:
```python
y
```

Out[57]:
```
0        0
1        0
2        0
3        0
4        0
        ..
41183    1
41184    0
41185    0
41186    1
41187    0
Name: Loan_Status_label, Length: 41188, dtype: int64
```

In [17]:
```python
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=42
```

In [18]:
```python
x_train
```

Out[18]:

	age	campaign	pdays	previous	no_previous_contact	not_working	job_admin.
12556	40	2	999	0	1	0	0
35451	31	4	999	0	1	0	1
30592	59	6	999	1	1	1	0
17914	43	5	999	0	1	0	0
3315	39	2	999	0	1	0	1
...	...	...	...	...	...	...	...
6265	58	2	999	0	1	1	0
11284	37	1	999	0	1	0	0
38158	35	1	4	1	0	0	1
860	40	2	999	0	1	0	0
15795	29	2	999	0	1	0	1

32950 rows × 59 columns

In [62]: x_test

Out[62]:

	age	campaign	pdays	previous	no_previous_contact	not_working	job_admin.
32884	57	1	999	1	1	0	0
3169	55	2	999	0	1	0	0
32206	33	1	999	1	1	0	0
9403	36	4	999	0	1	0	1
14020	27	2	999	0	1	0	0
...	...	...	...	...	...	...	...
12322	27	1	999	0	1	0	0
23440	41	7	999	0	1	0	0
29431	46	12	999	1	1	0	0
16627	31	2	999	0	1	0	1
1871	59	2	999	0	1	0	1

8238 rows × 59 columns

In [ ]:
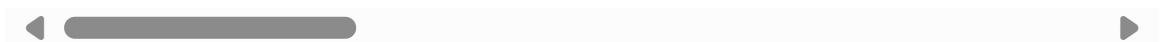
In [20]:
```python
from sklearn.preprocessing import StandardScaler
```

```
In [24]:    sc = StandardScaler()

            x_train_sc= sc.fit_transform(x_train)
            x_test_sc=sc.fit_transform(x_test)
```

```
In [22]:    x_train
```

Out[22]:

	age	campaign	pdays	previous	no_previous_contact	not_working	job_admin.
12556	40	2	999	0	1	0	0
35451	31	4	999	0	1	0	1
30592	59	6	999	1	1	1	0
17914	43	5	999	0	1	0	0
3315	39	2	999	0	1	0	1
...	...	...	...	...	...	...	...
6265	58	2	999	0	1	1	0
11284	37	1	999	0	1	0	0
38158	35	1	4	1	0	0	1
860	40	2	999	0	1	0	0
15795	29	2	999	0	1	0	1

32950 rows × 59 columns

```
In [23]:    x_train_sc
```

```
Out[23]:    array([[-1.66930454e-03, -2.06241614e-01,  1.94660673e-01, ...,
                    -3.40481710e-01,  3.98373809e-01, -1.84965343e-01],
                   [-8.64094846e-01,  5.13675879e-01,  1.94660673e-01, ...,
                    -3.40481710e-01,  3.98373809e-01, -1.84965343e-01],
                   [ 1.81900684e+00,  1.23359337e+00,  1.94660673e-01, ...,
                     2.93701532e+00, -2.51020518e+00, -1.84965343e-01],
                   ...,
                   [-4.80794606e-01, -5.66200360e-01, -5.14775262e+00, ...,
                    -3.40481710e-01, -2.51020518e+00,  5.40641820e+00],
                   [-1.66930454e-03, -2.06241614e-01,  1.94660673e-01, ...,
                    -3.40481710e-01,  3.98373809e-01, -1.84965343e-01],
                   [-1.05574497e+00, -2.06241614e-01,  1.94660673e-01, ...,
                    -3.40481710e-01,  3.98373809e-01, -1.84965343e-01]],
                  shape=(32950, 59))
```

```
In [26]:    x_test_sc
```

```
Out[26]:  array([[ 1.63562386, -0.56486335,  0.19840223, ...,  2.98952736,
                   -2.53145677, -0.18861365],
                 [ 1.44262285, -0.19952267,  0.19840223, ..., -0.33450104,
                    0.39502946, -0.18861365],
                 [-0.68038832, -0.56486335,  0.19840223, ...,  2.98952736,
                   -2.53145677, -0.18861365],
                 ...,
                 [ 0.57411828,  3.45388416,  0.19840223, ...,  2.98952736,
                   -2.53145677, -0.18861365],
                 [-0.87338934, -0.19952267,  0.19840223, ..., -0.33450104,
                    0.39502946, -0.18861365],
                 [ 1.82862488, -0.19952267,  0.19840223, ..., -0.33450104,
                    0.39502946, -0.18861365]], shape=(8238, 59))
```

```python
In [27]:  lr = LogisticRegression()
          lr.fit(x_train_sc,y_train)
```

Out[27]: ▼ LogisticRegression ⓘ ⍰

          ▶ Parameters

```python
In [28]:  lr.score(x_train,y_train)
```

Out[28]: 0.8911380880121396

```python
In [29]:  lr.score(x_test,y_test)
```

Out[29]: 0.8880796309783928

```python
In [30]:  print(f'Training Accuracy :{round(lr.score(x_train,y_train),2)*100}%')
          print(f'Test Accuracy     :{round(lr.score(x_test,y_test),2)*100}%')
```

```
Training Accuracy :89.0%
Test Accuracy     :89.0%
```

In [ ]:

```python
In [37]:  knn = KNeighborsRegressor(n_neighbors=7)
          knn.fit(x_train,y_train)
```

Out[37]: ▼ KNeighborsRegressor ⓘ ⍰

          ▶ Parameters

```python
In [45]:  print(f'Training Accuracy :{round(knn.score(x_train,y_train),2)*100}%')
          print(f'Test Accuracy     :{round(knn.score(x_test,y_test),2)*100}%')
```

```
Training Accuracy :31.0%
Test Accuracy     :5.0%
```

```python
In [48]:  dtree = DecisionTreeRegressor(max_depth=9)
          dtree.fit(x_train,y_train)
```

```
Out[48]:  ▼ DecisionTreeRegressor  ⓘ ⓘ

          ▶ Parameters
```

```
In [49]:  print(f'Training Accuracy:{round(dtree.score(x_train,y_train),2)*100}%')
          print(f'Test Accuracy:{round(dtree.score(x_test,y_test),2)*100}%')
```

```
Training Accuracy:25.0%
Test Accuracy:11.0%
```

```
In [62]:  lr = LinearRegression()
          lr.fit(x_train,y_train)
```

```
Out[62]:  ▼ LinearRegression  ⓘ ⓘ

          ▶ Parameters
```

```
In [54]:  lr.coef_
```

```
Out[54]:  array([[ 0.00201149, -0.08745845, -0.00212815,  0.16239052,  0.00085786,
                   0.47454794,  0.05473844, -0.15383082, -0.07670865, -0.0420984 ,
                  -0.06475622,  0.26319342, -0.04324801, -0.13699808,  0.24708782,
                  -0.05459497, -0.0357333 , -0.01857775, -0.15697032, -0.00646726,
                   0.09798183,  0.00392923, -0.00727861, -0.00354071, -0.11126238,
                  -0.05912369,  0.00692359, -0.041064  ,  0.07190036,  0.08191892,
                   0.31694894, -0.37661627, -0.00185919, -0.05130146,  0.01889007,
                  -0.02911512, -0.0091884 ,  0.01889007, -0.07122818,  0.39499941,
                  -0.45652593,  0.32585489, -0.66403745,  0.15324375, -0.49246554,
                   0.30350426,  0.66817654, -0.60034406, -0.65798655,  0.53600784,
                   0.36651981, -0.02030538, -0.20352329,  0.03823005,  0.04925475,
                   0.07481735, -0.0494228 , -0.03076258,  0.01865887]])
```

```
In [55]:  x.columns
```

```
Out[55]:  Index(['age', 'campaign', 'pdays', 'previous', 'no_previous_contact',
                 'not_working', 'job_admin.', 'job_blue-collar', 'job_entrepreneur',
                 'job_housemaid', 'job_management', 'job_retired', 'job_self-employed',
                 'job_services', 'job_student', 'job_technician', 'job_unemployed',
                 'job_unknown', 'marital_divorced', 'marital_married', 'marital_single',
                 'marital_unknown', 'education_basic.4y', 'education_basic.6y',
                 'education_basic.9y', 'education_high.school', 'education_illiterate',
                 'education_professional.course', 'education_university.degree',
                 'education_unknown', 'default_no', 'default_unknown', 'default_yes',
                 'housing_no', 'housing_unknown', 'housing_yes', 'loan_no',
                 'loan_unknown', 'loan_yes', 'contact_cellular', 'contact_telephone',
                 'month_apr', 'month_aug', 'month_dec', 'month_jul', 'month_jun',
                 'month_mar', 'month_may', 'month_nov', 'month_oct', 'month_sep',
                 'day_of_week_fri', 'day_of_week_mon', 'day_of_week_thu',
                 'day_of_week_tue', 'day_of_week_wed', 'poutcome_failure',
                 'poutcome_nonexistent', 'poutcome_success'],
                dtype='object')
```

```
In [63]:  pd.DataFrame(lr.coef_, index=x.columns,columns=['Coefficients/weights'])
```

	Coefficients/weights
age	0.000468
campaign	-0.004046
pdays	-0.002870
previous	0.038687
no_previous_contact	2.543796
not_working	0.042511
job_admin.	0.006092
job_blue-collar	-0.008288
job_entrepreneur	-0.007919
job_housemaid	-0.003928
job_management	-0.004909
job_retired	0.022905
job_self-employed	-0.007474
job_services	-0.007328
job_student	0.044514
job_technician	0.001932
job_unemployed	-0.024909
job_unknown	-0.010688
marital_divorced	-0.010373
marital_married	-0.003805
marital_single	0.006759
marital_unknown	0.007420
education_basic.4y	-0.014540
education_basic.6y	-0.012719
education_basic.9y	-0.020613
education_high.school	-0.016809
education_illiterate	0.089458
education_professional.course	-0.016766
education_university.degree	-0.002353
education_unknown	-0.005658
default_no	0.042344
default_unknown	0.010154
default_yes	-0.052498

	Coefficients/weights
housing_no	-0.000328
housing_unknown	0.002211
housing_yes	-0.001883
loan_no	-0.000511
loan_unknown	0.002211
loan_yes	-0.001700
contact_cellular	0.036052
contact_telephone	-0.036052
month_apr	-0.026038
month_aug	-0.127250
month_dec	0.121278
month_jul	-0.112827
month_jun	-0.052237
month_mar	0.215443
month_may	-0.109137
month_nov	-0.129622
month_oct	0.127199
month_sep	0.093191
day_of_week_fri	-0.002314
day_of_week_mon	-0.015840
day_of_week_thu	0.005647
day_of_week_tue	0.004898
day_of_week_wed	0.007608
poutcome_failure	-0.065673
poutcome_nonexistent	-0.010402
poutcome_success	0.076076

In [64]: `lr`

Out[64]:
```
▼ LinearRegression  ⓘ ⓘ

▶ Parameters
```

In [68]:
```python
y_pred_train=lr.predict(x_train)
y_pred_test=lr.predict(x_test)
```

```
In [65]: x_train[:3]
```

Out[65]:

	age	campaign	pdays	previous	no_previous_contact	not_working	job_admin.
12556	40	2	999	0	1	0	0
35451	31	4	999	0	1	0	1
30592	59	6	999	1	1	1	0

3 rows × 59 columns

```
In [66]: y_train[:3]
```

```
Out[66]: 12556    0
         35451    0
         30592    0
         Name: Loan_Status_label, dtype: int64
```

```
In [69]: y_pred_train[:3]
```

```
Out[69]: array([-0.03354994,  0.09632921,  0.13188504])
```

```
In [70]: x_test[:3]
```

Out[70]:

	age	campaign	pdays	previous	no_previous_contact	not_working	job_admin.
32884	57	1	999	1	1	0	0
3169	55	2	999	0	1	0	0
32206	33	1	999	1	1	0	0

3 rows × 59 columns

```
In [71]: y_test[:3]
```

```
Out[71]: 32884    0
         3169     0
         32206    0
         Name: Loan_Status_label, dtype: int64
```

```
In [72]: y_pred_test
```

```
Out[72]: array([0.08423694, 0.01119656, 0.07370406, ..., 0.13130264, 0.13318661,
                0.05738234], shape=(8238,))
```

```
In [78]: def training_evaluation(actual,predicted):
             mse =mean_squared_error(actual,predicted)
             rmse=np.sqrt(mean_squared_error(actual,predicted))
             mae=mean_absolute_error(actual,predicted)
             r2=r2_score(actual,predicted)

             print(f'MSE    :{mse}')
```

```
        print(f'RMSE    :{rmse}')
        print(f'MAE     :{mae}')
        print(f'R2_score  :{r2}')
        print()

        plt.scatter(actual,predicted,color='r',label='Actual price')
        plt.plot([actual.min(),actual.max()],[predicted.min(),predicted.max()],color
        plt.title('Best fit line - Training data')
        plt.xlabel('Actual price')
        plt.ylabel('Predicted price')
        plt.legend()
        plt.show()
```

In [79]: `training_evaluation(y_train,y_pred_train)`

```
MSE     :0.08261930553215971
RMSE    :0.28743574157045904
MAE     :0.1667332223085798
R2_score  :0.17214830947233095
```

```
---------------------------------------------------------------------------
AttributeError                            Traceback (most recent call last)
Cell In[79], line 1
----> 1 training_evaluation(y_train,y_pred_train)

Cell In[78], line 14, in training_evaluation(actual, predicted)
     11 print()
     13 plt.scatter(actual,predicted,color='r',label='Actual price')
---> 14 plt.plot([actual.min(),actual.max()],[predicted.min(),predicted.max()],co
lors=    ,lw=3,label=              )
     15 plt.title('Best fit line - Training data')
     16 plt.xlabel('Actual price')

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\pyplo
t.py:3838, in plot(scalex, scaley, data, *args, **kwargs)
   3830 @_copy_docstring_and_deprecators(Axes.plot)
   3831 def plot(
   3832     *args: float | ArrayLike | str,
   (...)   3836     **kwargs,
   3837 ) -> list[Line2D]:
-> 3838     return gca().plot(
   3839         *args,
   3840         scalex=scalex,
   3841         scaley=scaley,
   3842         **({    : data} if data is not None else {}),
   3843         **kwargs,
   3844     )

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\axes
\_axes.py:1777, in Axes.plot(self, scalex, scaley, data, *args, **kwargs)
   1534 """
   1535 Plot y versus x as lines and/or markers.
   1536
   (...)   1774 (``'green'``) or hex strings (``'#008000'``).
   1775 """
   1776 kwargs = cbook.normalize_kwargs(kwargs, mlines.Line2D)
-> 1777 lines = [*self._get_lines(self, *args, data=data, **kwargs)]
   1778 for line in lines:
   1779     self.add_line(line)

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\axes
\_base.py:297, in _process_plot_var_args.__call__(self, axes, data, return_kwarg
s, *args, **kwargs)
    295     this += args[0],
    296     args = args[1:]
--> 297 yield from self._plot_args(
    298     axes, this, kwargs, ambiguous_fmt_datakey=ambiguous_fmt_datakey,
    299     return_kwargs=return_kwargs
    300 )

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\axes
\_base.py:546, in _process_plot_var_args._plot_args(self, axes, tup, kwargs, retu
rn_kwargs, ambiguous_fmt_datakey)
    544     return list(result)
    545 else:
--> 546     return [l[0] for l in result]

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\axes
\_base.py:539, in <genexpr>(.0)
    534 else:
```

```
    535     raise ValueError(
    536         f"label must be scalar or have the same length as the input "
    537         f"data, but found {len(label)} for {n_datasets} datasets.")
--> 539 result = (make_artist(axes, x[:, j % ncx], y[:, j % ncy], kw,
    540                      {**kwargs,         : label})
    541         for j, label in enumerate(labels))
    543 if return_kwargs:
    544     return list(result)

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\axes
\_base.py:338, in _process_plot_var_args._make_line(self, axes, x, y, kw, kwargs)
    336 kw = {**kw, **kwargs}  # Don't modify the original kw.
    337 self._setdefaults(self._getdefaults(kw), kw)
--> 338 seg = mlines.Line2D(x, y, **kw)
    339 return seg, kw

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\line
s.py:421, in Line2D.__init__(self, xdata, ydata, linewidth, linestyle, color, gap
color, marker, markersize, markeredgewidth, markeredgecolor, markerfacecolor, mar
kerfacecoloralt, fillstyle, antialiased, dash_capstyle, solid_capstyle, dash_join
style, solid_joinstyle, pickradius, drawstyle, markevery, **kwargs)
    417 self.set_markeredgewidth(markeredgewidth)
    419 # update kwargs before updating data to give the caller a
    420 # chance to init axes (and hence unit support)
--> 421 self._internal_update(kwargs)
    422 self.pickradius = pickradius
    423 self.ind_offset = 0

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\artis
t.py:1233, in Artist._internal_update(self, kwargs)
   1226 def _internal_update(self, kwargs):
   1227     """
   1228     Update artist properties without prenormalizing them, but generating
   1229     errors as if calling `set`.
   1230
   1231     The lack of prenormalization is to maintain backcompatibility.
   1232     """
-> 1233     return self._update_props(
   1234         kwargs, {cls.__name__}

   1235         {prop_name!r} )

File ~\AppData\Local\Programs\Python\Python313\Lib\site-packages\matplotlib\artis
t.py:1206, in Artist._update_props(self, props, errfmt)
   1204             func = getattr(self, f"set_{k}", None)
   1205             if not callable(func):
-> 1206                 raise AttributeError(
   1207                     errfmt.format(cls=type(self), prop_name=k),
   1208                     name=k)
   1209         ret.append(func(v))
   1210 if ret:

AttributeError: Line2D.set() got an unexpected keyword argument 'colors'
```
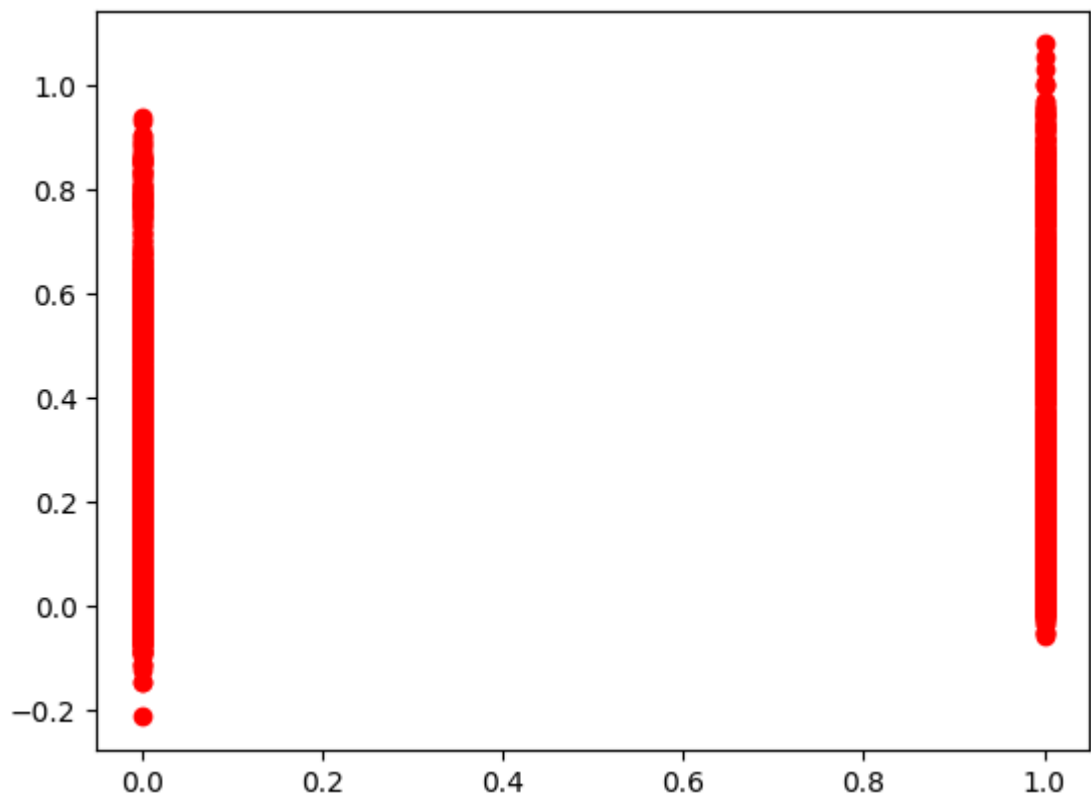
In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: