```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.tree import DecisionTree
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
import warnings
warnings.filterwarnings('ignore')
```

```
data = pd.read_csv('/kaggle/input/bank-marketing-dataset/bank.csv')
```

```
data.head()
```

	marital_status	age	policy_holder	policy_duration	policy_premium	policy_premium_month	policy_premium_year	insured_sex	insured_occupation	insured_education_level	insured_hobbies	insured_relationship	number_of_vehicles_involved	property_damage	bodily_injury	witnesses	police_report_available	total_claim_amount	injury_claim	property_claim	vehicle_claim	auto_make	auto_model	auto_year	insured_zip	capital-gains	capital-loss	incident_hour_of_the_day	number_of_claims_involved	bodily_injuries	authorities_contacted	incident_type	incident_severity	insured_hobbies	insured_occupation	insured_education_level	property_damage	police_report_available	fraud_reported
0	married	48	521585	17-10-2014	31	500804	6000	male	exec-managerial	JD	chess	husband	1	YES	NO	2	YES	70610	7050	14100	49410	saab	92x	2004	466132	NaN	-700000	15	1	1	1	Single Vehicle Collision	Minor Damage	chess	exec-managerial	JD	YES	YES	N
1	married	62	608232	18-06-2015	34	433677	6000	male	other-service	MD	dirt_bike	husband	1	NO	NO	0	NO	63400	6340	12680	44380	mercedes	E400	2007	468176	5060700	0	18	1	1	1	Multi-vehicle Collision	Minor Damage	dirt_bike	other-service	MD	NO	NO	N
2	married	29	610706	10-05-2011	3	448404	5070	female	sales	PhD	camping	husband	3	NO	NO	2	NO	107700	10770	21540	75390	dodge	RAM	2007	430632	0	0	10	3	2	1	Parked Car	Minor Damage	camping	sales	PhD	NO	NO	N
3	married	41	610616	06-06-2014	10	603988	5040	female	tech-support	PhD	movies	husband	1	YES	YES	1	YES	80520	8052	16104	56364	audi	A5	2006	431202	0	0	17	1	0	1	Single Vehicle Collision	Minor Damage	movies	tech-support	PhD	YES	YES	N
4	married	44	610077	12-12-2015	9	548827	5070	male	craft-repair	JD	hiking	husband	1	NO	NO	0	NO	67500	6750	13500	47250	ford	F150	2004	608420	0	-65000	22	1	0	1	Multi-vehicle Collision	Minor Damage	hiking	craft-repair	JD	NO	NO	N
500	married	29	610616	10-12-2014	3	394796	5040	female	tech-support	JD	movies	husband	1	NO	NO	2	NO	108480	10848	21696	75936	saab	92x	2004	608062	0	0	10	1	2	1	Parked Car	Minor Damage	movies	tech-support	JD	NO	NO	N
501	married	62	608232	18-06-2015	34	433677	6000	male	other-service	MD	dirt_bike	husband	1	NO	NO	0	NO	63400	6340	12680	44380	mercedes	E400	2007	468176	5060700	0	18	1	1	1	Multi-vehicle Collision	Minor Damage	dirt_bike	other-service	MD	NO	NO	N
502	married	41	610616	06-06-2014	10	603988	5040	female	tech-support	PhD	movies	husband	1	YES	YES	1	YES	80520	8052	16104	56364	audi	A5	2006	431202	0	0	17	1	0	1	Single Vehicle Collision	Minor Damage	movies	tech-support	PhD	YES	YES	N

1000 rows × 40 columns

```
data.info()
```

```
<class'.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 40 columns):
 #   marital_status               1000 non-null   object
 #   age                          1000 non-null   int64
 #   policy_holder                1000 non-null   int64
 #   policy_duration              1000 non-null   object
 #   policy_premium               1000 non-null   int64
 #   policy_premium_month         1000 non-null   int64
 #   policy_premium_year          1000 non-null   int64
 #   insured_sex                  1000 non-null   object
 #   insured_occupation           1000 non-null   object
 #   insured_education_level      1000 non-null   object
 #   insured_hobbies              1000 non-null   object
 #   insured_relationship         1000 non-null   object
 #   number_of_vehicles_involved  1000 non-null   int64
 #   property_damage              1000 non-null   object
 #   bodily_injury                1000 non-null   object
 #   witnesses                    1000 non-null   int64
 #   police_report_available      1000 non-null   object
 #   total_claim_amount           1000 non-null   int64
 #   injury_claim                 1000 non-null   int64
 #   property_claim               1000 non-null   int64
 #   vehicle_claim                1000 non-null   int64
 #   auto_make                    1000 non-null   object
 #   auto_model                   1000 non-null   object
 #   auto_year                    1000 non-null   int64
 #   insured_zip                  1000 non-null   int64
 #   capital-gains                1000 non-null   int64
 #   capital-loss                 1000 non-null   int64
 #   incident_hour_of_the_day     1000 non-null   int64
 #   number_of_claims_involved    1000 non-null   int64
 #   bodily_injuries              1000 non-null   int64
 #   authorities_contacted        1000 non-null   object
 #   incident_type                1000 non-null   object
 #   incident_severity            1000 non-null   object
 #   insured_hobbies              1000 non-null   object
 #   insured_occupation           1000 non-null   object
 #   insured_education_level      1000 non-null   object
 #   property_damage              1000 non-null   object
 #   police_report_available      1000 non-null   object
 #   fraud_reported               1000 non-null   object
dtypes: int64: 17, object: 23
memory usage: 320.0 KB
```

```
data.describe()
```

	age	policy_holder	policy_premium	policy_premium_month	policy_premium_year	insured_sex	insured_occupation	insured_education_level	insured_hobbies	insured_relationship	number_of_vehicles_involved	property_damage	bodily_injury	witnesses	police_report_available	total_claim_amount	injury_claim	property_claim	vehicle_claim	auto_make	auto_model	auto_year	insured_zip	capital-gains	capital-loss	incident_hour_of_the_day	number_of_claims_involved	bodily_injuries	authorities_contacted	incident_type	incident_severity	insured_hobbies	insured_occupation	insured_education_level	property_damage	police_report_available	fraud_reported
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000	1000.000000
mean	38.948000	608349.000000	1256.406000	104.594000	15072847.680000	5000.000000	500804.000000	6000.000000	433677.000000	6000.000000	1.000000	0.000000	0.992000	1.000000	0.000000	52761.094000	7433.420000	7699.560000	37628.114000	NaN	NaN	2005.103000	601062.356000	25126.100000	-26093.828000	52.105000	1.487000	1.839000	1.000000	1.992000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.247000
std	9.139096	610616.000000	243.187091	2.000000	2918186.400305	0.000000	448404.000000	5040.000000	603988.000000	5040.000000	0.000000	0.000000	0.089443	0.000000	0.000000	26401.339831	4880.951353	4807.730159	18886.252893	NaN	NaN	6015.814116	71296.095991	27877.736340	28104.096686	6.015861	0.592962	1.018172	0.000000	0.089443	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.431048
min	19.000000	610077.000000	433.330000	100.000000	5196641.000000	5000.000000	448404.000000	5040.000000	608420.000000	5040.000000	1.000000	0.000000	0.992000	1.000000	0.000000	70610.000000	7050.000000	14100.000000	49410.000000	NaN	NaN	1995.000000	430632.000000	0.000000	-65000.000000	0.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000
25%	32.000000	608232.000000	1091.000000	103.000000	13000000.000000	5000.000000	448404.000000	5040.000000	608420.000000	5040.000000	1.000000	0.000000	0.992000	1.000000	0.000000	63400.000000	6340.000000	12680.000000	44380.000000	NaN	NaN	2000.000000	468176.000000	0.000000	-70000.000000	12.000000	1.000000	1.000000	1.000000	1.992000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000
50%	38.000000	608232.000000	1256.000000	104.000000	15000000.000000	5000.000000	448404.000000	5040.000000	608420.000000	5040.000000	1.000000	0.000000	0.992000	1.000000	0.000000	70610.000000	7050.000000	14100.000000	49410.000000	NaN	NaN	2005.000000	608420.000000	0.000000	-70000.000000	15.000000	1.000000	1.000000	1.000000	1.992000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000
75%	44.000000	608232.000000	1422.000000	106.000000	17000000.000000	5000.000000	448404.000000	5040.000000	608420.000000	5040.000000	1.000000	0.000000	0.992000	1.000000	0.000000	80520.000000	8052.000000	16104.000000	56364.000000	NaN	NaN	2006.000000	608420.000000	0.000000	-70000.000000	22.000000	1.000000	1.000000	1.000000	1.992000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000
max	62.000000	610616.000000	1433.000000	108.000000	17000000.000000	5000.000000	448404.000000	5040.000000	608420.000000	5040.000000	1.000000	0.000000	0.992000	1.000000	0.000000	108480.000000	10848.000000	21696.000000	75936.000000	NaN	NaN	2007.000000	608420.000000	0.000000	-70000.000000	23.000000	1.000000	1.000000	1.000000	1.992000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000

```
data.isnull()
```

```
marital_status               False
age                          False
policy_holder                False
policy_duration              False
policy_premium               False
policy_premium_month         False
policy_premium_year          False
insured_sex                  False
insured_occupation           False
insured_education_level      False
insured_hobbies              False
insured_relationship         False
number_of_vehicles_involved  False
property_damage              False
bodily_injury                False
witnesses                    False
police_report_available      False
total_claim_amount           False
injury_claim                 False
property_claim               False
vehicle_claim                False
auto_make                    False
auto_model                   False
auto_year                    False
insured_zip                  False
capital-gains                False
capital-loss                 False
incident_hour_of_the_day     False
number_of_claims_involved    False
bodily_injuries              False
authorities_contacted        False
incident_type                False
incident_severity            False
insured_hobbies              False
insured_occupation           False
insured_education_level      False
property_damage              False
police_report_available      False
fraud_reported               False
dtype: bool
```

```
data.isnull.sum()
```

```
marital_status               0
age                          0
policy_holder                0
policy_duration              0
policy_premium               0
policy_premium_month         0
policy_premium_year          0
insured_sex                  0
insured_occupation           0
insured_education_level      0
insured_hobbies              0
insured_relationship         0
number_of_vehicles_involved  0
property_damage              0
bodily_injury                0
witnesses                    0
police_report_available      0
total_claim_amount           0
injury_claim                 0
property_claim               0
vehicle_claim                0
auto_make                    0
auto_model                   0
auto_year                    0
insured_zip                  0
capital-gains                0
capital-loss                 0
incident_hour_of_the_day     0
number_of_claims_involved    0
bodily_injuries              0
authorities_contacted        0
incident_type                0
incident_severity            0
insured_hobbies              0
insured_occupation           0
insured_education_level      0
property_damage              0
police_report_available      0
fraud_reported               0
dtype: int64
```

```
data.dropna()
```

```
marital_status               NaN
age                          NaN
policy_holder                NaN
policy_duration              NaN
policy_premium               NaN
policy_premium_month         NaN
policy_premium_year          NaN
insured_sex                  NaN
insured_occupation           NaN
insured_education_level      NaN
insured_hobbies              NaN
insured_relationship         NaN
number_of_vehicles_involved  NaN
property_damage              NaN
bodily_injury                NaN
witnesses                    NaN
police_report_available      NaN
total_claim_amount           NaN
injury_claim                 NaN
property_claim               NaN
vehicle_claim                NaN
auto_make                    NaN
auto_model                   NaN
auto_year                    NaN
insured_zip                  NaN
capital-gains                NaN
capital-loss                 NaN
incident_hour_of_the_day     NaN
number_of_claims_involved    NaN
bodily_injuries              NaN
authorities_contacted        NaN
incident_type                NaN
incident_severity            NaN
insured_hobbies              NaN
insured_occupation           NaN
insured_education_level      NaN
property_damage              NaN
police_report_available      NaN
fraud_reported               NaN
dtype: int64
```

```
data.dropna()
```

```
marital_status               0
age                          0
policy_holder                0
policy_duration              0
policy_premium               0
policy_premium_month         0
policy_premium_year          0
insured_sex                  0
insured_occupation           0
insured_education_level      0
insured_hobbies              0
insured_relationship         0
number_of_vehicles_involved  0
property_damage              0
bodily_injury                0
witnesses                    0
police_report_available      0
total_claim_amount           0
injury_claim                 0
property_claim               0
vehicle_claim                0
auto_make                    0
auto_model                   0
auto_year                    0
insured_zip                  0
capital-gains                0
capital-loss                 0
incident_hour_of_the_day     0
number_of_claims_involved    0
bodily_injuries              0
authorities_contacted        0
incident_type                0
incident_severity            0
insured_hobbies              0
insured_occupation           0
insured_education_level      0
property_damage              0
police_report_available      0
fraud_reported               0
dtype: int64
```

```
data.duplicated()
```

```
0    False
1    False
2    False
3    False
4    False
dtype: bool
```

```
data.drop_duplicates()
```

	marital_status	age	policy_holder	policy_duration	policy_premium	policy_premium_month	policy_premium_year	insured_sex	insured_occupation	insured_education_level	insured_hobbies	insured_relationship	number_of_vehicles_involved	property_damage	bodily_injury	witnesses	police_report_available	total_claim_amount	injury_claim	property_claim	vehicle_claim	auto_make	auto_model	auto_year	insured_zip	capital-gains	capital-loss	incident_hour_of_the_day	number_of_claims_involved	bodily_injuries	authorities_contacted	incident_type	incident_severity	insured_hobbies	insured_occupation	insured_education_level	property_damage	police_report_available	fraud_reported
0	married	48	521585	17-10-2014	31	500804	6000	male	exec-managerial	JD	chess	husband	1	YES	NO	2	YES	70610	7050	14100	49410	saab	92x	2004	466132	NaN	-700000	15	1	1	1	Single Vehicle Collision	Minor Damage	chess	exec-managerial	JD	YES	YES	N
1	married	62	608232	18-06-2015	34	433677	6000	male	other-service	MD	dirt_bike	husband	1	NO	NO	0	NO	63400	6340	12680	44380	mercedes	E400	2007	468176	5060700	0	18	1	1	1	Multi-vehicle Collision	Minor Damage	dirt_bike	other-service	MD	NO	NO	N
2	married	29	610706	10-05-2011	3	448404	5070	female	sales	PhD	camping	husband	3	NO	NO	2	NO	107700	10770	21540	75390	dodge	RAM	2007	430632	0	0	10	3	2	1	Parked Car	Minor Damage	camping	sales	PhD	NO	NO	N
3	married	41	610616	06-06-2014	10	603988	5040	female	tech-support	PhD	movies	husband	1	YES	YES	1	YES	80520	8052	16104	56364	audi	A5	2006	431202	0	0	17	1	0	1	Single Vehicle Collision	Minor Damage	movies	tech-support	PhD	YES	YES	N
4	married	44	610077	12-12-2015	9	548827	5070	male	craft-repair	JD	hiking	husband	1	NO	NO	0	NO	67500	6750	13500	47250	ford	F150	2004	608420	0	-65000	22	1	0	1	Multi-vehicle Collision	Minor Damage	hiking	craft-repair	JD	NO	NO	N

1000 rows × 40 columns

```
data.shape
```

```
(1000, 40)
```

```
data.columns
```

```
Index(['marital_status', 'age', 'policy_holder', 'policy_duration',
       'policy_premium', 'policy_premium_month', 'policy_premium_year',
       'insured_sex', 'insured_occupation', 'insured_education_level',
       'insured_hobbies', 'insured_relationship',
       'number_of_vehicles_involved', 'property_damage', 'bodily_injury',
       'witnesses', 'police_report_available', 'total_claim_amount',
       'injury_claim', 'property_claim', 'vehicle_claim', 'auto_make',
       'auto_model', 'auto_year', 'insured_zip', 'capital-gains',
       'capital-loss', 'incident_hour_of_the_day',
       'number_of_claims_involved', 'bodily_injuries',
       'authorities_contacted', 'incident_type', 'incident_severity',
       'insured_hobbies', 'insured_occupation', 'insured_education_level',
       'property_damage', 'police_report_available', 'fraud_reported',
       'auto_year_0', 'auto_year_1', 'auto_year_2', 'auto_year_3',
       'auto_year_4', 'auto_year_5', 'auto_year_6', 'auto_year_7',
       'auto_year_8', 'auto_year_9', 'auto_year_10', 'auto_year_11',
       'auto_year_12', 'auto_year_13', 'auto_year_14', 'auto_year_15',
       'auto_year_16', 'auto_year_17', 'auto_year_18', 'auto_year_19',
       'auto_year_20', 'auto_year_21', 'auto_year_22', 'auto_year_23',
       'auto_year_24', 'auto_year_25', 'auto_year_26', 'auto_year_27',
       'auto_year_28', 'auto_year_29', 'auto_year_30', 'auto_year_31',
       'auto_year_32', 'auto_year_33', 'auto_year_34', 'auto_year_35',
       'auto_year_36', 'auto_year_37', 'auto_year_38', 'auto_year_39',
```