In [1]:
```
from IPython.display import Image
Image(filename='logo.PNG', height=340, width=900)
```

Out[1]:

Random Forest Classification - SOCIAL NETWORKS DATASET

Importing Libraries

In [2]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

In [3]:
```
df = pd.read_csv('Social_Network_Ads.csv')
df.head()
```

Out[3]:

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

	User ID	Gender	Age	EstimatedSalary	Purchased
0	15624510	Male	19	19000	0
1	15810944	Male	35	20000	0
2	15668575	Female	26	43000	0
3	15603246	Female	27	57000	0
4	15804002	Male	19	76000	0

In [4]:
```
df.describe()
```

Out[4]:

	User ID	Age	EstimatedSalary	Purchased
count	4.000000e+02	400.000000	400.000000	400.000000
mean	1.569154e+07	37.655000	69742.500000	0.357500
std	7.165832e+04	10.482877	34096.960282	0.479864
min	1.556669e+07	18.000000	15000.000000	0.000000
25%	1.562676e+07	29.750000	43000.000000	0.000000
50%	1.569434e+07	37.000000	70000.000000	0.000000
75%	1.575036e+07	46.000000	88000.000000	1.000000
max	1.581524e+07	60.000000	150000.000000	1.000000

In [5]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 5 columns):
User ID            400 non-null int64
Gender             400 non-null object
Age                400 non-null int64
EstimatedSalary    400 non-null int64
Purchased          400 non-null int64
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
dtypes: int64(4), object(1)
memory usage: 15.8+ KB
```

In [6]:
```
df[df.isnull().any(axis=1)]
```

Out[6]:

User ID	Gender	Age	EstimatedSalary	Purchased

In [7]:
```
df.drop('User ID', axis = 1, inplace=True)
df.head()
```

Out[7]:

	Gender	Age	EstimatedSalary	Purchased
0	Male	19	19000	0
1	Male	35	20000	0
2	Female	26	43000	0
3	Female	27	57000	0
4	Male	19	76000	0

In [8]:
```
df=pd.get_dummies(df, drop_first=True)
```

In [9]:
```
cols = ['Age', 'EstimatedSalary','Gender_Male', 'Purchased']
df=df[cols]
df
```

Out[9]:

	Age	EstimatedSalary	Gender_Male	Purchased
0	19	19000	1	0
1	35	20000	1	0
2	26	43000	0	0
3	27	57000	0	0
4	19	76000	1	0

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

	Age	EstimatedSalary	Gender_Male	Purchased
...	...	...	...	...
395	46	41000	0	1
396	51	23000	1	1
397	50	20000	0	1
398	36	33000	1	0
399	49	36000	0	1

400 rows × 4 columns

Creating X and Y Variables

In [10]:
```
X = df.iloc[:,:-1].values
X
```

Out[10]:
```
array([[   19, 19000,     1],
       [   35, 20000,     1],
       [   26, 43000,     0],
       ...,
       [   50, 20000,     0],
       [   36, 33000,     1],
       [   49, 36000,     0]], dtype=int64)
```

In [11]:
```
y = df.iloc[:,-1].values
y
```

Out[11]:
```
array([0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1,
1,
       1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0,
       0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,
0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,
0,
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
       0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0,
0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0,
       0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0,
0,
       0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0,
0,
       0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0,
       0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0,
1,
       0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1,
0,
       1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1,
0,
       1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0,
1,
       0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0,
1,
       1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1,
1,
       0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1,
0,
       1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0,
1,
       0, 1, 0, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0,
1,
       1, 1, 0, 1], dtype=int64)
```

Splitting into Train & Test Data

In [12]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size =
0.25, random_state = 0)
```

In [13]:
```
len(X_train)
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

Out[13]: 300

In [14]:
```
X_train
```

Out[14]:
```
array([[    44,  39000,      0],
       [    32, 120000,      1],
       [    38,  50000,      0],
       [    32, 135000,      0],
       [    52,  21000,      0],
       [    53, 104000,      0],
       [    39,  42000,      1],
       [    38,  61000,      1],
       [    36,  50000,      0],
       [    36,  63000,      0],
       [    35,  25000,      0],
       [    35,  50000,      1],
       [    42,  73000,      1],
       [    47,  49000,      0],
       [    59,  29000,      0],
       [    49,  65000,      1],
       [    45, 131000,      0],
       [    31,  89000,      0],
       [    46,  82000,      0],
       [    47,  51000,      0],
       [    26,  15000,      1],
       [    60, 102000,      1],
       [    38, 112000,      0],
       [    40, 107000,      1],
       [    42,  53000,      0],
       [    35,  59000,      1],
       [    48,  41000,      1],
       [    48, 134000,      0],
       [    38, 113000,      0],
       [    29, 148000,      1],
       [    26,  15000,      0],
       [    60,  42000,      1],
       [    24,  19000,      1],
       [    42, 149000,      1],
       [    46,  96000,      0],
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
[    28,  59000,      1],
[    39,  96000,      1],
[    28,  89000,      1],
[    41,  72000,      1],
[    45,  26000,      1],
[    33,  69000,      0],
[    20,  82000,      0],
[    31,  74000,      1],
[    42,  80000,      1],
[    35,  72000,      0],
[    33, 149000,      0],
[    40,  71000,      1],
[    51, 146000,      0],
[    46,  79000,      1],
[    35,  75000,      1],
[    38,  51000,      1],
[    36,  75000,      0],
[    37,  78000,      0],
[    38,  61000,      1],
[    60, 108000,      0],
[    20,  82000,      0],
[    57,  74000,      1],
[    42,  65000,      1],
[    26,  80000,      1],
[    46, 117000,      1],
[    35,  61000,      1],
[    21,  68000,      0],
[    28,  44000,      0],
[    41,  87000,      1],
[    37,  33000,      0],
[    27,  90000,      1],
[    39,  42000,      1],
[    28, 123000,      1],
[    31, 118000,      0],
[    25,  87000,      1],
[    35,  71000,      0],
[    37,  70000,      1],
[    35,  39000,      1],
[    47,  23000,      1],
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
[    35, 147000,      0],
[    48, 138000,      0],
[    26,  86000,      1],
[    25,  79000,      1],
[    52, 138000,      0],
[    51,  23000,      1],
[    35,  60000,      0],
[    33, 113000,      0],
[    30, 107000,      1],
[    48,  33000,      1],
[    41,  80000,      0],
[    48,  96000,      0],
[    31,  18000,      1],
[    31,  71000,      0],
[    43, 129000,      1],
[    59,  76000,      0],
[    18,  44000,      0],
[    36, 118000,      1],
[    42,  90000,      0],
[    47,  30000,      0],
[    26,  43000,      0],
[    40,  78000,      1],
[    46,  59000,      1],
[    59,  42000,      0],
[    46,  74000,      0],
[    35,  91000,      1],
[    28,  59000,      0],
[    40,  57000,      1],
[    59, 143000,      1],
[    57,  26000,      0],
[    52,  38000,      0],
[    47, 113000,      0],
[    53, 143000,      0],
[    35,  27000,      1],
[    58, 101000,      0],
[    45,  45000,      0],
[    23,  82000,      0],
[    46,  23000,      1],
[    42,  65000,      1],
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
[    28,  84000,      0],
[    38,  59000,      1],
[    26,  84000,      0],
[    29,  28000,      0],
[    37,  71000,      0],
[    22,  55000,      0],
[    48,  35000,      0],
[    49,  28000,      1],
[    38,  65000,      0],
[    27,  17000,      0],
[    46,  28000,      1],
[    48, 141000,      1],
[    26,  17000,      0],
[    35,  97000,      0],
[    39,  59000,      0],
[    24,  27000,      0],
[    32,  18000,      1],
[    46,  88000,      1],
[    35,  58000,      1],
[    56,  60000,      1],
[    47,  34000,      1],
[    40,  72000,      0],
[    32, 100000,      1],
[    19,  21000,      0],
[    25,  90000,      1],
[    35,  88000,      1],
[    28,  32000,      1],
[    50,  20000,      0],
[    40,  59000,      1],
[    50,  44000,      0],
[    35,  72000,      1],
[    40, 142000,      0],
[    46,  32000,      0],
[    39,  71000,      0],
[    20,  74000,      1],
[    29,  75000,      1],
[    31,  76000,      1],
[    47,  25000,      1],
[    40,  61000,      1],
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
[    34, 112000,      1],
[    38,  80000,      0],
[    42,  75000,      0],
[    47,  47000,      0],
[    39,  75000,      0],
[    19,  25000,      1],
[    37,  80000,      0],
[    36,  60000,      1],
[    41,  52000,      1],
[    36, 125000,      1],
[    48,  29000,      0],
[    36, 126000,      0],
[    51, 134000,      0],
[    27,  57000,      0],
[    38,  71000,      1],
[    39,  61000,      0],
[    22,  27000,      0],
[    33,  60000,      0],
[    48,  74000,      1],
[    58,  23000,      0],
[    53,  72000,      1],
[    32, 117000,      0],
[    54,  70000,      1],
[    30,  80000,      1],
[    58,  95000,      0],
[    26,  52000,      0],
[    45,  79000,      1],
[    24,  55000,      1],
[    40,  75000,      1],
[    33,  28000,      0],
[    44, 139000,      0],
[    22,  18000,      1],
[    33,  51000,      0],
[    43, 133000,      0],
[    24,  32000,      0],
[    46,  22000,      0],
[    35,  55000,      1],
[    54, 104000,      0],
[    48, 119000,      0],
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
[    35,  53000,      1],
[    37, 144000,      1],
[    23,  66000,      0],
[    37, 137000,      0],
[    31,  58000,      1],
[    33,  41000,      0],
[    45,  22000,      0],
[    30,  15000,      1],
[    19,  19000,      1],
[    49,  74000,      1],
[    39, 122000,      1],
[    35,  73000,      1],
[    39,  71000,      1],
[    24,  23000,      1],
[    41,  72000,      0],
[    29,  83000,      0],
[    54,  26000,      0],
[    35,  44000,      0],
[    37,  75000,      1],
[    29,  47000,      0],
[    31,  68000,      0],
[    42,  54000,      1],
[    30, 135000,      1],
[    52, 114000,      0],
[    50,  36000,      0],
[    56, 133000,      1],
[    29,  61000,      1],
[    30,  89000,      1],
[    26,  16000,      1],
[    33,  31000,      1],
[    41,  72000,      0],
[    36,  33000,      1],
[    55, 125000,      0],
[    48, 131000,      0],
[    41,  71000,      0],
[    30,  62000,      0],
[    37,  72000,      1],
[    41,  63000,      0],
[    58,  47000,      0],
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
[    30, 116000,      0],
[    20,  49000,      1],
[    37,  74000,      1],
[    41,  59000,      1],
[    49,  89000,      1],
[    28,  79000,      1],
[    53,  82000,      0],
[    40,  57000,      1],
[    60,  34000,      1],
[    35, 108000,      1],
[    21,  72000,      1],
[    38,  71000,      1],
[    39, 106000,      1],
[    37,  57000,      0],
[    26,  72000,      0],
[    35,  23000,      0],
[    54, 108000,      0],
[    30,  17000,      1],
[    39, 134000,      1],
[    29,  43000,      1],
[    33,  43000,      1],
[    35,  38000,      1],
[    41,  45000,      1],
[    41,  72000,      0],
[    39, 134000,      0],
[    27, 137000,      0],
[    21,  16000,      0],
[    26,  32000,      1],
[    31,  66000,      1],
[    39,  73000,      0],
[    41,  79000,      1],
[    47,  50000,      0],
[    41,  30000,      0],
[    37,  93000,      0],
[    60,  46000,      0],
[    25,  22000,      1],
[    28,  37000,      0],
[    38,  55000,      0],
[    36,  54000,      0],
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
[    20,  36000,      0],
[    56, 104000,      0],
[    40,  57000,      1],
[    42, 108000,      0],
[    20,  23000,      0],
[    40,  65000,      1],
[    47,  20000,      1],
[    18,  86000,      0],
[    35,  79000,      1],
[    57,  33000,      0],
[    34,  72000,      0],
[    49,  39000,      0],
[    27,  31000,      0],
[    19,  70000,      1],
[    39,  79000,      0],
[    26,  81000,      1],
[    25,  80000,      1],
[    28,  85000,      0],
[    55,  39000,      1],
[    50,  88000,      0],
[    49,  88000,      1],
[    52, 150000,      1],
[    35,  65000,      0],
[    42,  54000,      1],
[    34,  43000,      1],
[    37,  52000,      1],
[    48,  30000,      0],
[    29,  43000,      1],
[    36,  52000,      1],
[    27,  54000,      0],
[    26, 118000,      0]], dtype=int64)
```

In [15]:
```
from sklearn.preprocessing import StandardScaler
sc_X = StandardScaler()
X_train = sc_X.fit_transform(X_train)
X_test = sc_X.transform(X_test)
```

Building the Model

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

In [16]:
```
from sklearn.ensemble import RandomForestClassifier

# Creating a RAndom Forest Classifier
clf = RandomForestClassifier(n_estimators = 40, n_jobs= 2, random_state
= 0)

# Training the Classifier
clf.fit(X_train, y_train)
```

Out[16]:
```
RandomForestClassifier(bootstrap=True, class_weight=None, criterion='gi
ni',
                       max_depth=None, max_features='auto', max_leaf_no
des=None,
                       min_impurity_decrease=0.0, min_impurity_split=No
ne,
                       min_samples_leaf=1, min_samples_split=2,
                       min_weight_fraction_leaf=0.0, n_estimators=40, n
_jobs=2,
                       oob_score=False, random_state=0, verbose=0,
                       warm_start=False)
```

In [17]:
```
# Applying the trained data to predict test data
y_pred = clf.predict(X_test)
y_pred
```

Out[17]:
```
array([0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0,
1,
       0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,
0,
       1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0,
1,
       0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0,
1,
       1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1], dtype=int64)
```

In [18]:
```
y_test
```

Out[18]:
```
array([0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0,
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

```
       1,
              0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,
       0,
              1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0,
       1,
              0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 0,
       1,
              1, 0, 0, 1, 0, 0, 0, 1, 0, 1, 1, 1], dtype=int64)
```

In [19]:
```
y_pred
```

Out[19]:
```
array([0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0,
       1,
              0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0,
       0,
              1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0,
       1,
              0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0,
       1,
              1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1], dtype=int64)
```

Accuracy Metrics

In [20]:
```
from sklearn import metrics
from sklearn.metrics import confusion_matrix
```

In [21]:
```
cm = confusion_matrix(y_test, y_pred)
```

In [22]:
```
cm
```

Out[22]:
```
array([[64,  4],
       [ 3, 29]], dtype=int64)
```

In [23]:
```
metrics.accuracy_score(y_test, y_pred)
```

Out[23]:
```
0.93
```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD

In [24]:
```
pd.crosstab(y_test, y_pred, rownames = ['Actual Status'], colnames = [
'Predicted Status'])
```

Out[24]:

Actual Status \ Predicted Status	0	1
0	64	4
1	3	29

Predicting a Value

In [25]:
```
clf.predict([[1., 45., 160000.]])
```

Out[25]:
```
array([1], dtype=int64)
```

In [26]:
```
X_test[12]
```

Out[26]:
```
array([-0.11157634, -0.42281668,  1.02020406])
```

In [ ]:
```

```

Create PDF in your applications with the Pdfcrowd HTML to PDF API PDFCROWD