Forward Selection:

Code ready to Write quality

```
import pandas as pd

dataset_name = "/content/drive/MyDrive/winequality-white.csv"
data_frame = ...
```

fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality
6.7	0.27	0.41	20.7	0.035	45.0	170.0	1.0010	3.03	0.48	9	6
6.6	0.29	0.37	1.6	0.049	14.0	132.0	0.9940	3.30	0.49	9.5	6
8.3	0.42	0.62	19.25	0.040	41.0	172.0	1.0002	2.98	0.67	9.7	5
7.2	0.21	0.32	8.5	0.058	47.0	186.0	0.9956	3.19	0.40	9.9	6
7.2	0.23	0.32	8.5	0.058	47.0	186.0	0.9956	3.19	0.40	9.9	6

```
data.quality.value_counts()
```

```
6    2198
5    1457
7     880
4     163
8     175
3      20
9       5
Name: quality, dtype: int64
```

```
# split into dependent and independent variables
X = data.iloc[:,0:11]
y = data.iloc[:,11]
```

```
# splitting into train or test
from sklearn.model_selection import train_test_split

X_train,X_test,y_train,y_test = train_test_split(X,y,train_size = 0.8,random_state = 10)
```

Random Forest

```
# importing the model
from sklearn.ensemble import RandomForestClassifier
model = RandomForestClassifier()
model.fit(X_train,y_train)
```

```
RandomForestClassifier
RandomForestClassifier()
```

```
y_pred = model.predict(X_test)
```

```
y_pred
```

```
array([5, 5, 6, 6, 5, 5, 6, 6, 5, 6, 5, 6, 5, 5, 5, 6, 5, 6, 5, 5, 5, 5,
       6, 5, 6, 5, 5, 6, 6, 5, 6, 6, 6, 5, 5, 5, 5, 6, 5, 6, 5, 5, 5, 6,
       5, 6, 5, 5, 6, 5, 6, 5, 5, 5, 6, 5, 6, 5, 5, 5, 6, 5, 6, 5, 5, 5,
       5, 5, 5, 5, 5, 6, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5,
       5, 5, 5, 5, 5, 5
```