

```
import pandas as pd
```

```
df = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/titanic.csv")
df.head()
```

	PassengerId	Name	Pclass	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Survived
0	1	Braund, Mr. Owen Harris	3	male	22.0	1	0	A/5 21171	7.2500	NaN	0
1	2	Cumings, Mrs. John Bradley (Florence Briggs T. B.)	1	female	38.0	1	0	PC 17599	71.2833	C85	1

```
df.drop(['PassengerId', 'Name', 'SibSp', 'Parch', 'Ticket', 'Cabin', 'Embarked'], axis='columns', inplace=True)
```

```
df.head()
```

	Pclass	Sex	Age	Fare	Survived
0	3	male	22.0	7.2500	0
1	1	female	38.0	71.2833	1
2	3	female	26.0	7.9250	1
3	1	female	35.0	53.1000	1
4	3	male	35.0	8.0500	0

```
inputs = df.drop('Survived', axis='columns')
target = df.Survived
```

```
inputs.Sex = inputs.Sex.map({'male': 1, 'female': 2})
```

```
inputs.Age[:10]
```

```
0    22.0
1    38.0
2    26.0
3    35.0
4    35.0
5     NaN
6    54.0
7     2.0
```

```
8    27.0
9    14.0
Name: Age, dtype: float64
```

```
inputs.Age = inputs.Age.fillna(inputs.Age.mean())
```

```
inputs.head()
```

	Pclass	Sex	Age	Fare
0	3	1	22.0	7.2500
1	1	2	38.0	71.2833
2	3	2	26.0	7.9250
3	1	2	35.0	53.1000
4	3	1	35.0	8.0500

```
from sklearn.model_selection import train_test_split
```

```
X_train, X_test, y_train, y_test = train_test_split(inputs,target,test_size=0.2)
```

```
len(X_train)
```

```
712
```

```
len(X_test)
```

```
179
```

```
from sklearn import tree
model = tree.DecisionTreeClassifier()
```

```
model.fit(X_train,y_train)
```

```
DecisionTreeClassifier(ccp_alpha=0.0, class_weight=None, criterion='gini',
                        max_depth=None, max_features=None, max_leaf_nodes=None,
                        min_impurity_decrease=0.0, min_impurity_split=None,
                        min_samples_leaf=1, min_samples_split=2,
                        min_weight_fraction_leaf=0.0, presort='deprecated',
                        random_state=None, splitter='best')
```

```
model.score(X_test,y_test)
```

```
0.7988826815642458
```

✓ 0s completed at 10:40

● ✕