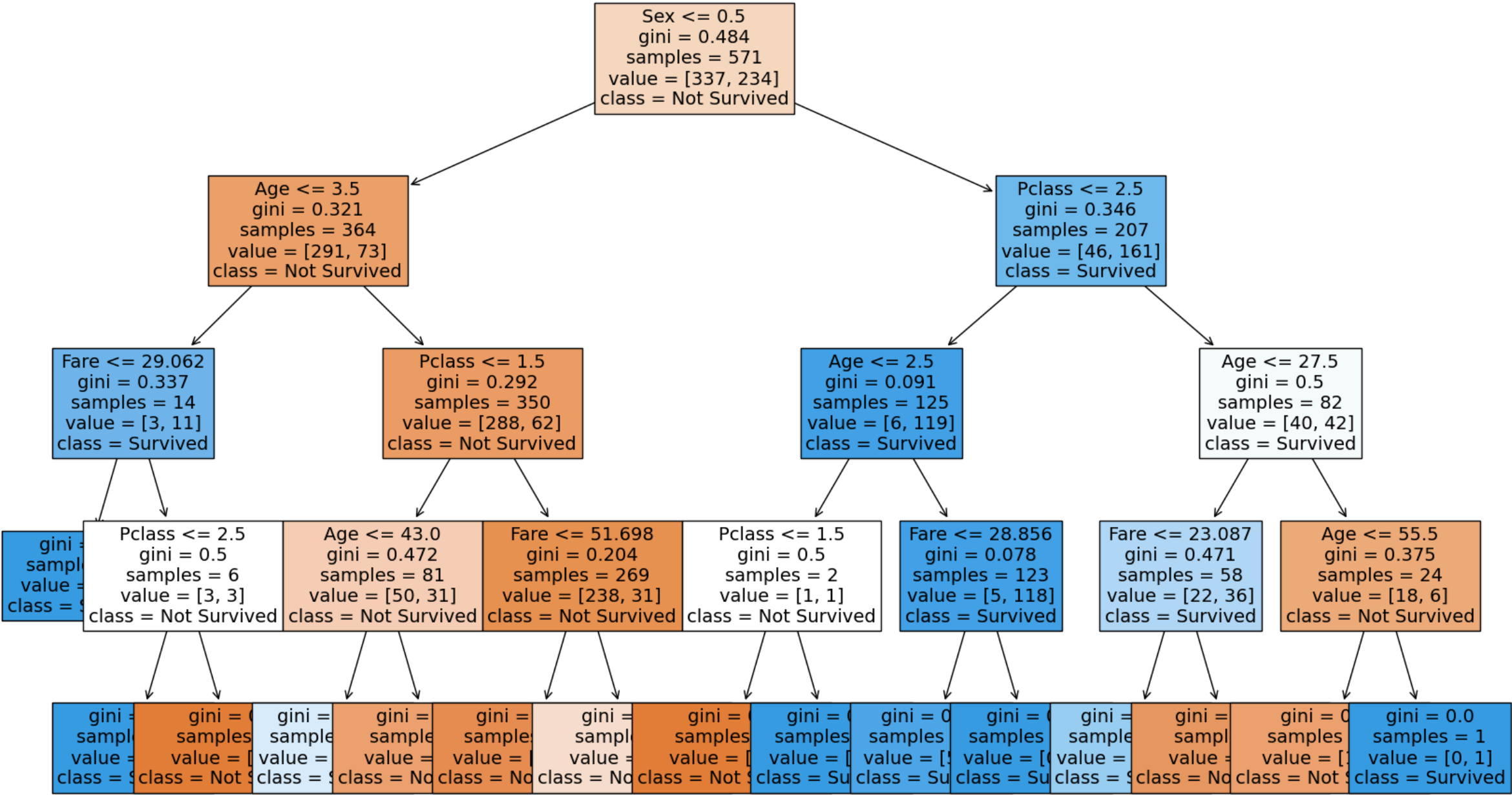


```
In [7]: import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier, plot_tree
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
import matplotlib.pyplot as plt
```

```
In [8]: file_path='titanic.csv'
df=pd.read_csv(file_path)
df=df[['Survived','Pclass','Sex','Age','Fare']]
df.dropna(subset=['Survived','Pclass','Sex','Age','Fare'],inplace=True)
df['Sex']=df['Sex'].map({'male':0,'female':1})
X=df[['Pclass','Sex','Age','Fare']]
y=df['Survived']
```

```
In [26]: X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=42)
clf=DecisionTreeClassifier(max_depth=4,random_state=42)
clf.fit(X_train,y_train)
plt.figure(figsize=(20,12))
plot_tree(clf,
          feature_names=X.columns,
          class_names=["Not Survived","Survived"],
          filled=True,
          fontsize=14)
plt.title("Decision tree - Titanic survival prediction",fontsize=20)
plt.show()
```

Decision tree - Titanic survival prediction



```
In [22]: y_pred=clf.predict(X_test)
print("Accuracy:",accuracy_score(y_test,y_pred))
print("Precision:",precision_score(y_test,y_pred))
print("Recall:",recall_score(y_test,y_pred))
print("F1-score:",f1_score(y_test,y_pred))
```

Accuracy: 0.7342657342657343
Precision: 0.6551724137931034
Recall: 0.6785714285714286
F1-score: 0.6666666666666666