In [1]:
```
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.datasets import load_wine
```

In [2]:
```
# 导入数据
wine = load_wine()
```

In [3]:
```
# 划分训练集和测试集
from sklearn.model_selection import train_test_split

Xtrain, Xtest, Ytrain, Ytest = train_test_split(wine.data, wine.target, test_size=0.3)
```

In [4]:
```
# 初始化、训练、预测
clf = DecisionTreeClassifier(random_state=0)
rfc = RandomForestClassifier(random_state=0)

clf = clf.fit(Xtrain, Ytrain)
rfc = rfc.fit(Xtrain, Ytrain)

score_c = clf.score(Xtest, Ytest)
score_r = rfc.score(Xtest, Ytest)
```

In [5]:
```
print(score_c)
print(score_r)
```

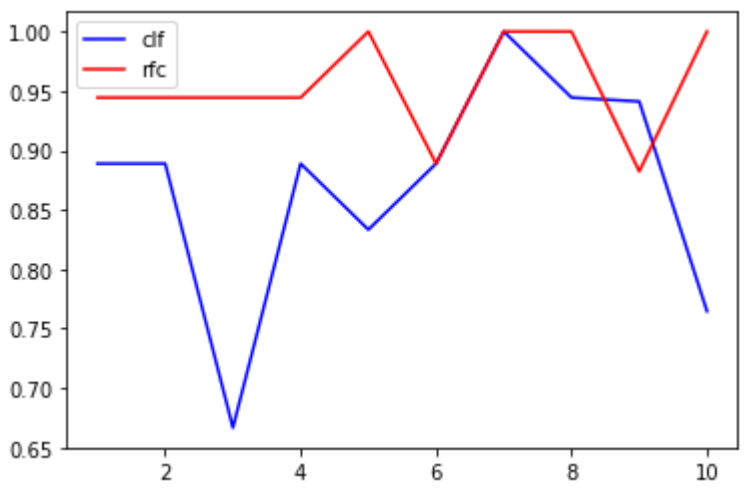
```
0.9074074074074074
1.0
```

In [8]:
```
# 交叉验证
from sklearn.model_selection import cross_val_score
import matplotlib.pyplot as plt

clf = DecisionTreeClassifier()
clf_s = cross_val_score(clf, wine.data, wine.target, cv=10)

rfc = RandomForestClassifier(n_estimators=10)
rfc_s = cross_val_score(rfc, wine.data, wine.target, cv=10)

plt.plot(range(1,11), clf_s, color='blue', label='clf')
plt.plot(range(1,11), rfc_s, color='red', label='rfc')
plt.legend()
plt.show()
```

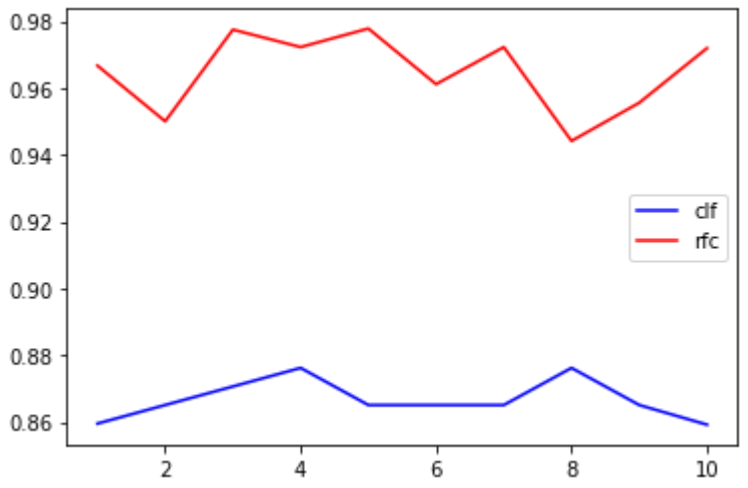


In [9]:
```
# 十组交叉验证
clf_s = []
rfc_s = []

for i in range(10):
  clf = DecisionTreeClassifier()
  clf_s.append(cross_val_score(clf, wine.data, wine.target, cv=10).mean())

  rfc = RandomForestClassifier(n_estimators=10)
  rfc_s.append(cross_val_score(rfc, wine.data, wine.target, cv=10).mean())

plt.plot(range(1,11), clf_s, color='blue', label='clf')
plt.plot(range(1,11), rfc_s, color='red', label='rfc')
plt.legend()
plt.show()
```



In [10]:
```
# n_estimators 的学习曲线
rfc_s = []

for i in range(200):
  rfc = RandomForestClassifier(n_estimators=i+1, n_jobs=-1)
  rfc_s.append(cross_val_score(rfc, wine.data, wine.target, cv=10).mean())

plt.plot(range(1,201), rfc_s)
plt.show()
```