Heart Disease Prediction

Python notebook using data from Heart Disease UCI · 233 views · 9mo ago · beginner, classification, data cleaning


```
print(acc)
```

```
85.24590163934425
```

In [52]:

```
neigh = KNeighborsClassifier(n_neighbors=13 )
neigh.fit(X_train, y_train)
KNNpredicted = neigh.predict(X_test)
acc=sum(KNNpredicted==y_test)/len(y_test)*100
print(acc)
```

```
50.81967213114754
```

In [53]:

```
svc = svm.SVC(kernel='linear').fit(X_train, y_train)
pred_target=svc.predict(X_test)
acc=sum(pred_target==y_test)/len(y_test)*100
print(acc)
```

```
86.88524590163934
```

In [54]:

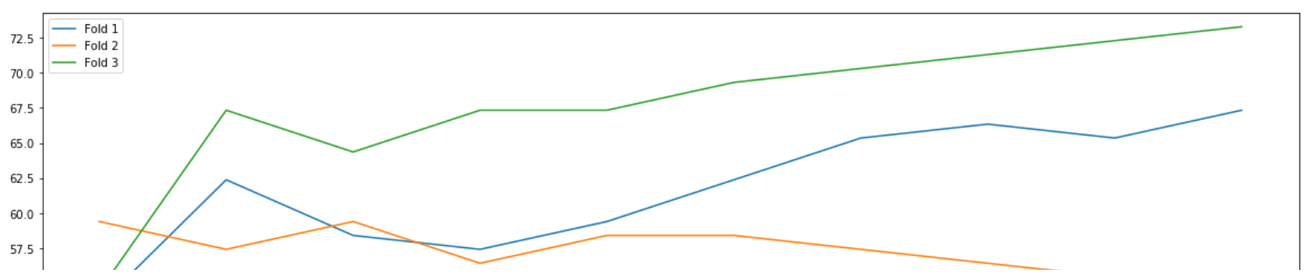
```
predicted = {}
key =0
predicted[1]=[]
predicted[2]=[]
predicted[3]=[]
kfold = KFold(3,True,1)
for train, test in kfold.split(X):
    X_train, X_test = X.iloc[train], X.iloc[test]
    y_train, y_test = y[train], y[test]
    for k in range(1,20,2):
        neigh = KNeighborsClassifier(n_neighbors=k )
        neigh.fit(X_train, y_train)
        KNNpredicted = neigh.predict(X_test)
        acc=sum(KNNpredicted==y_test)/len(y_test)*100
        key = key+1 if k==1 else key
        predicted[key].append(  acc )
```

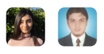
In [55]:

```
X = list(range(1,20,2))
plt.figure(figsize=(20,5))
plt.plot(X,predicted[1],label ="Fold 1")
plt.plot(X,predicted[2],label ="Fold 2")
plt.plot(X,predicted[3],label ="Fold 3")
plt.legend(loc="upper left")
```

Out[55]:

```
<matplotlib.legend.Legend at 0x7f73297839e8>
```





Data

Data Sources

Heart Disease UCI

heart.csv 14 columns

Heart Disease UCI

https://archive.ics.uci.edu/ml/datasets/Heart+Disease

Last Updated: 2 years ago (Version 1)

About this Dataset

Context

This database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. In particular, the Cleveland database is the only one that has been used by ML