

Creating model for the best value of K

```
test_score = {}
es = {}
train_score = {}
es = {}
for i in range(2, 4, 2):
    kf = KFold(n_splits
    = i) sum_train = 0
    sum_
    test
    = 0
    data
    = df
    for train, test in
        kf.split(data):
            train_data =
            data.iloc[train,:]
            test_data =
            data.iloc[test,:]
            x_train =
            train_data.drop(["prog"], axis=1)
            y_train = train_data['prognosis']
            x_test =
            test_data.drop(["prog"], axis=
            1) y_test =
            test_data["prognosis"]
            algo_model =
            dt.fit(x_train, y_train)
            sum_train +=
            dt.score(x_train, y_train)
            y_pred = dt.predict(x_test)
            sum_test +=
            accuracy_score(y_test, y_pred) average_test
            = sum_test/i
            average_train =
            sum_train/i
            test_scores[i] =
            average_test
            train_scores[i] =
            average_train
            print("kvalue: ", i)
```

- Saving the model in binary form:

```
from sklearn.externals import joblib
```

```
joblib.dump(dt, 'my_model_for_healthcare')
```