



PREDICTING STUDENT ADMISSIONS



DATA IN
ACTION

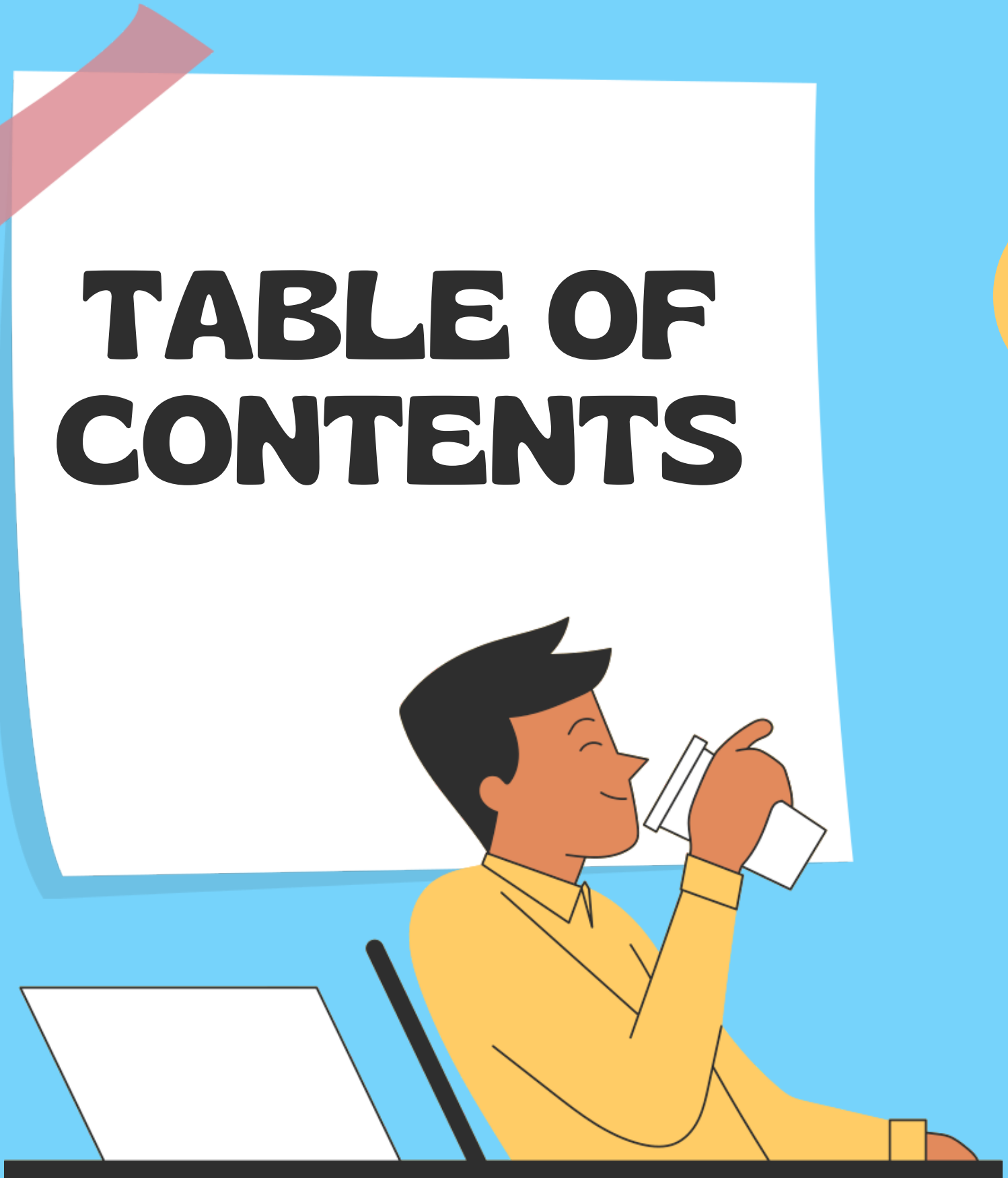


TABLE OF CONTENTS

1

ABOUT OUR DATA SET

2

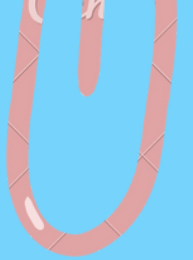
SUPERVISED LEARNING

3

IMPLEMENTING THE MODELS

4

SUMMARY AND INFERENCE

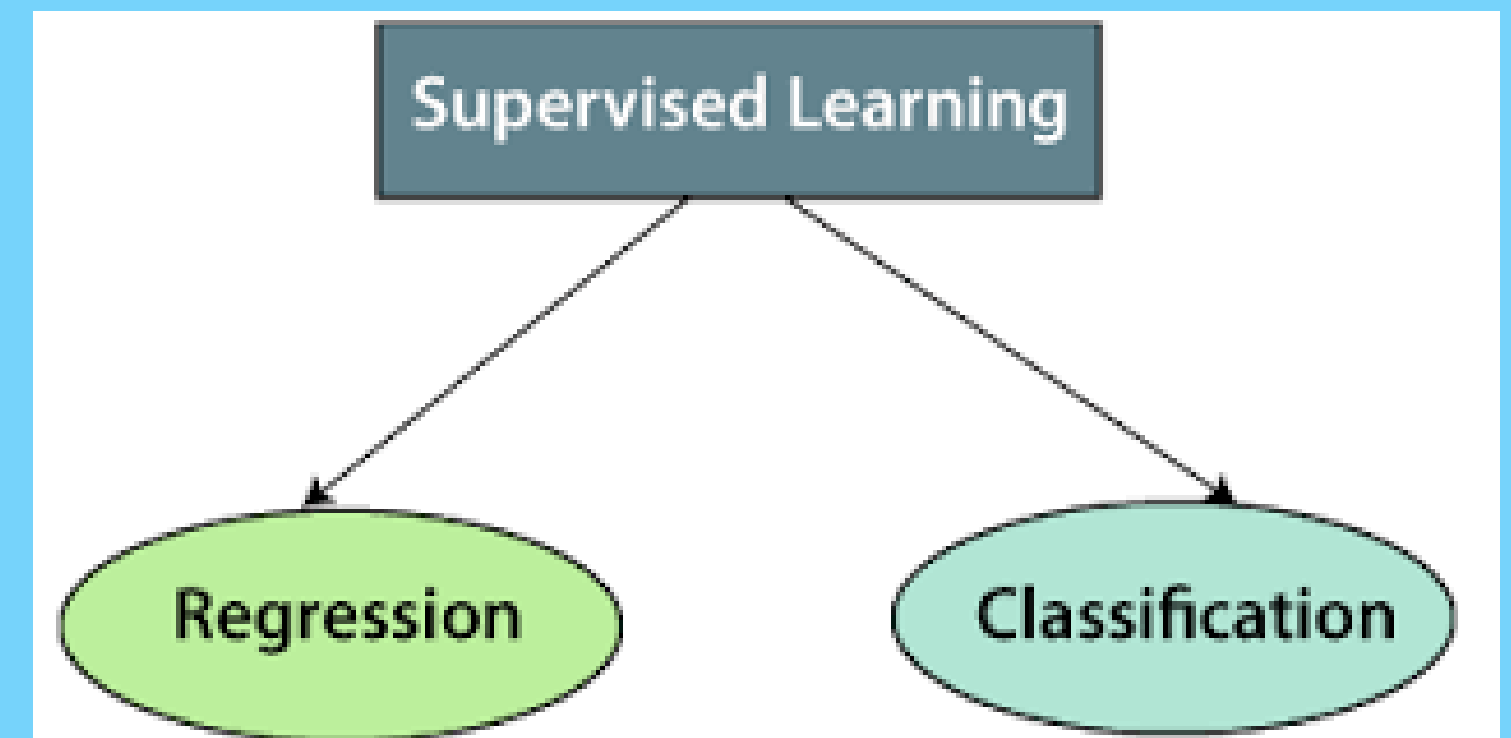


ABOUT OUR DATASET

THIS DATASET INCLUDES VARIOUS INFORMATION LIKE GRE SCORE, TOEFL SCORE, UNIVERSITY RATING, SOP (STATEMENT OF PURPOSE), LOR (LETTER OF RECOMMENDATION), CGPA AND CHANCE OF ADMIT. IN THIS DATASET, 400 ENTRIES ARE INCLUDED.

- **GRE SCORES (OUT OF 340)**
- **TOEFL SCORES (OUT OF 120)**
- **UNIVERSITY RATING (OUT OF 5)**
- **STATEMENT OF PURPOSE (SOP)**
- **LETTER OF RECOMMENDATION (LOR) STRENGTH (OUT OF 5)**
- **UNDERGRADUATE GPA (OUT OF 10)**
- **CHANCE OF ADMIT (RANGING FROM 0 TO 1).**

SUPERVISED LEARNING



- models are trained using labeled data.
- models need to find the mapping function to map the input variable (X) with the output variable (Y).

$$Y=F(X)$$

**LET'S
BEGIN!**



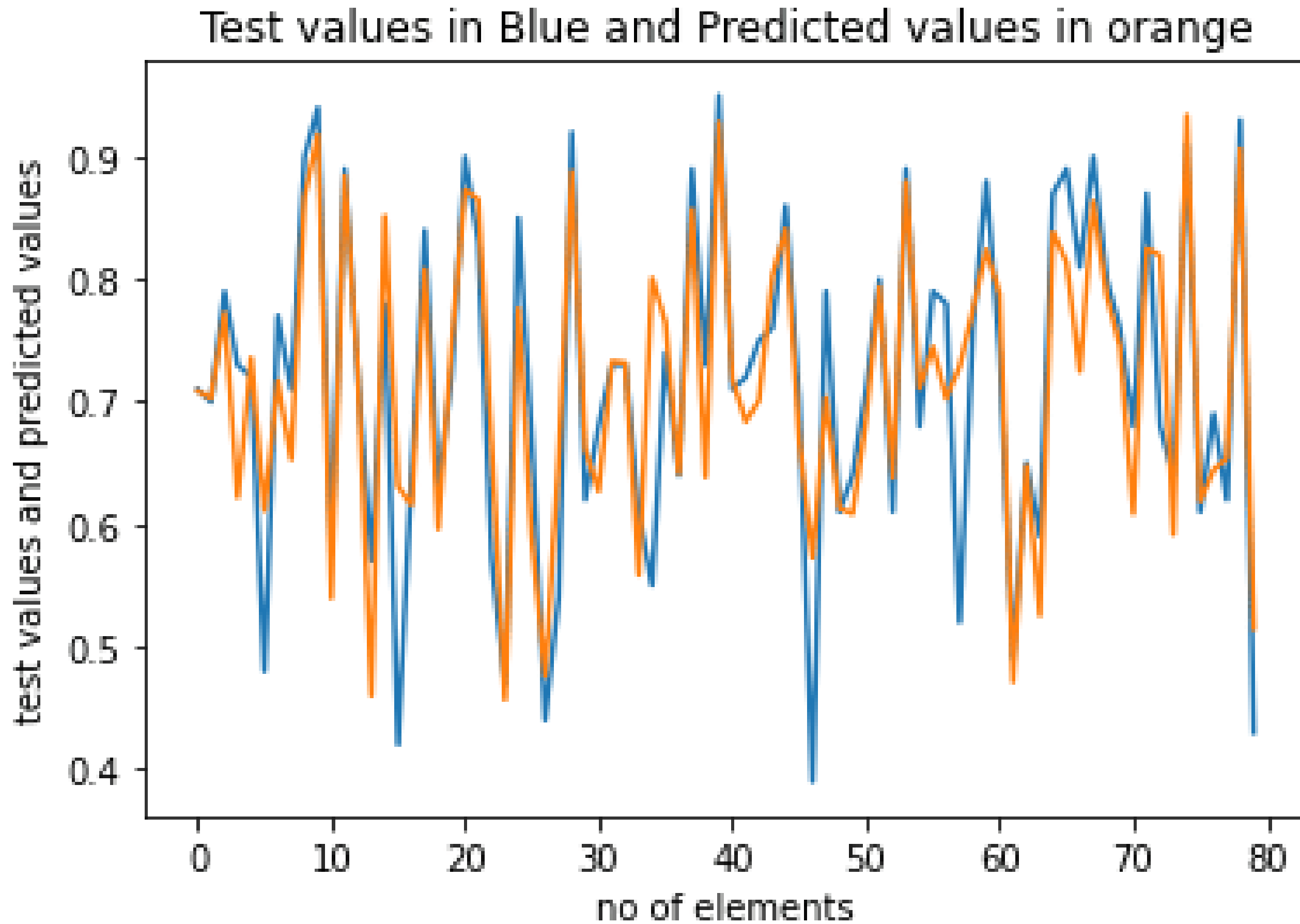


MULTILINEAR REGRESSION

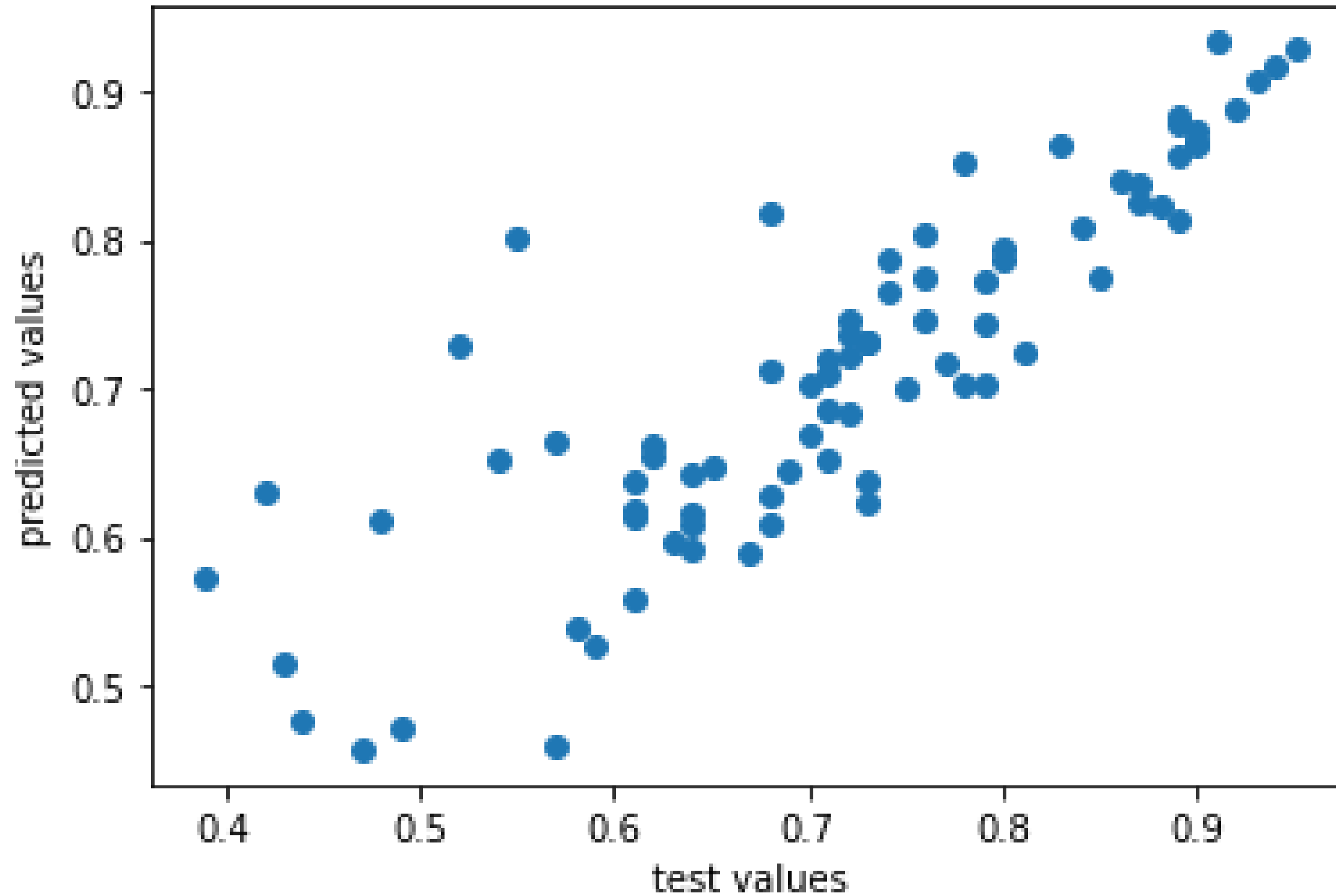
IMPORTING THE MODEL AND VISUALIZING IT

```
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor.fit(X_train, y_train)
y_pred = regressor.predict(X_test)
```

VISUALIZATION



INFERENCE



the mean squared error is 0.004901434759480476



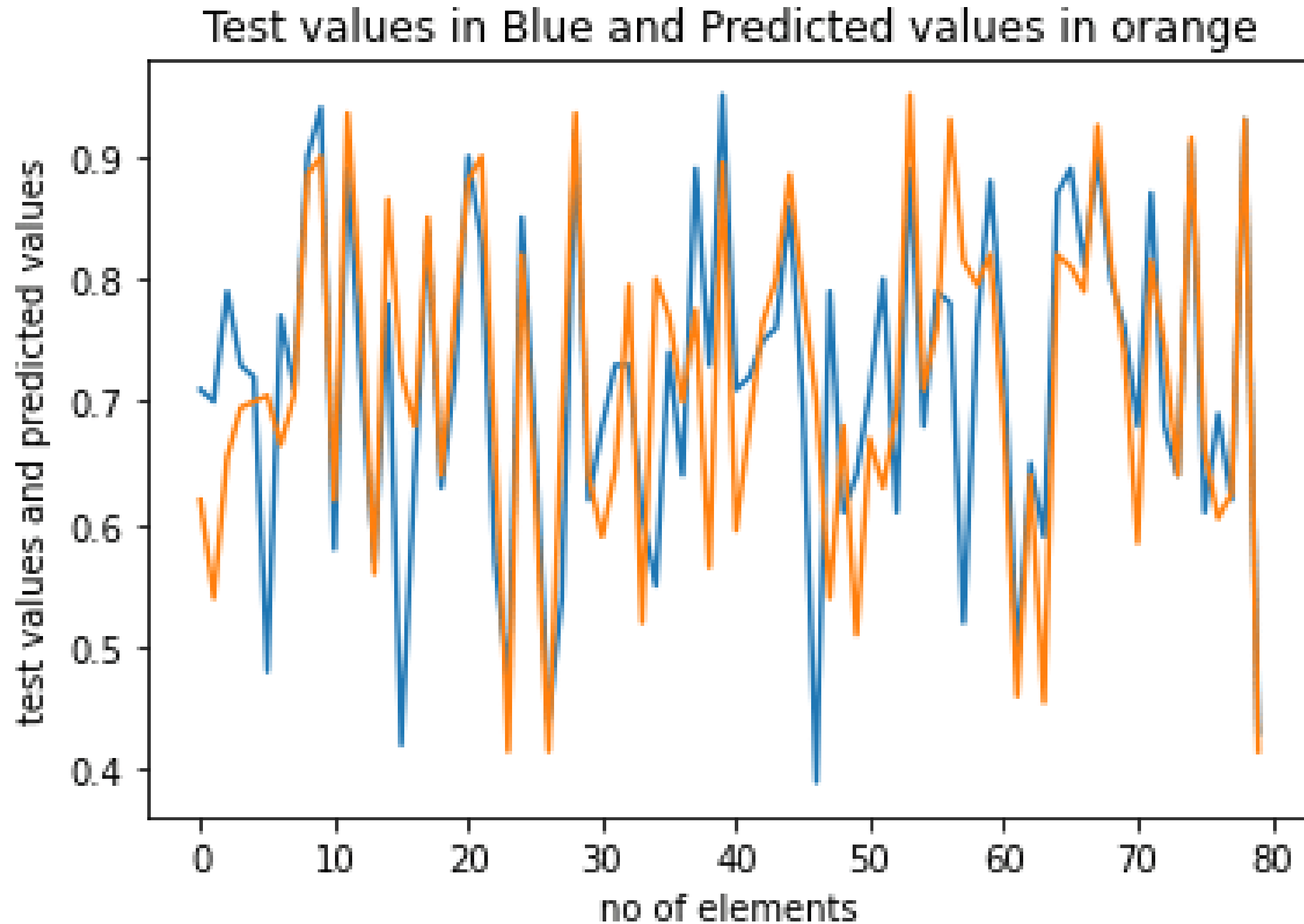
-
-
-
-

KNN REGRESSION

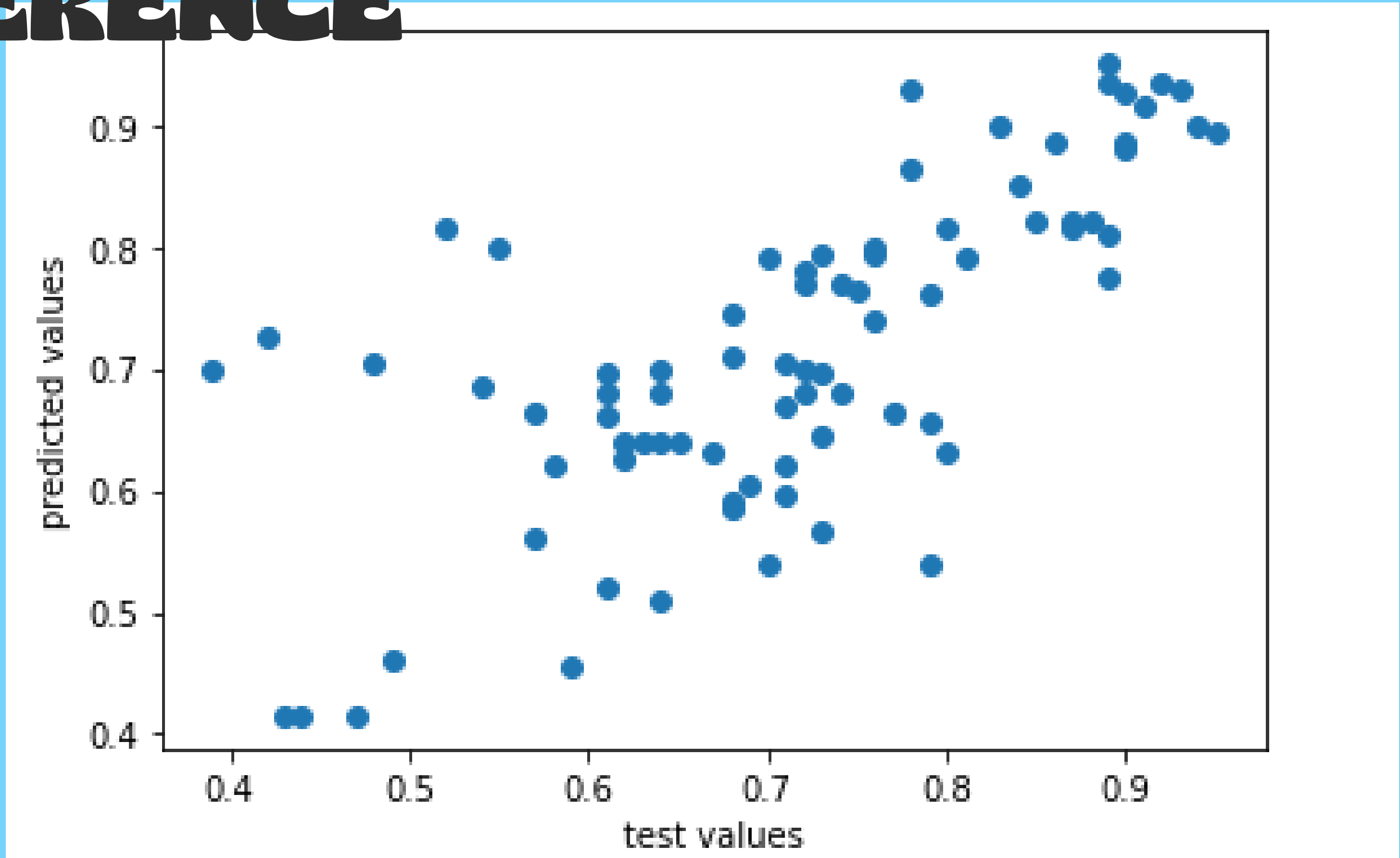
IMPORTING THE MODEL

```
from sklearn.neighbors import KNeighborsRegressor
KNneigh = KNeighborsRegressor(n_neighbors=2)
KNneigh.fit(X_train,y_train)
knn= KNneigh.predict(X_test)
```

VISUALIZATION



INFERENCE



mean squared error is 0.010400937499999

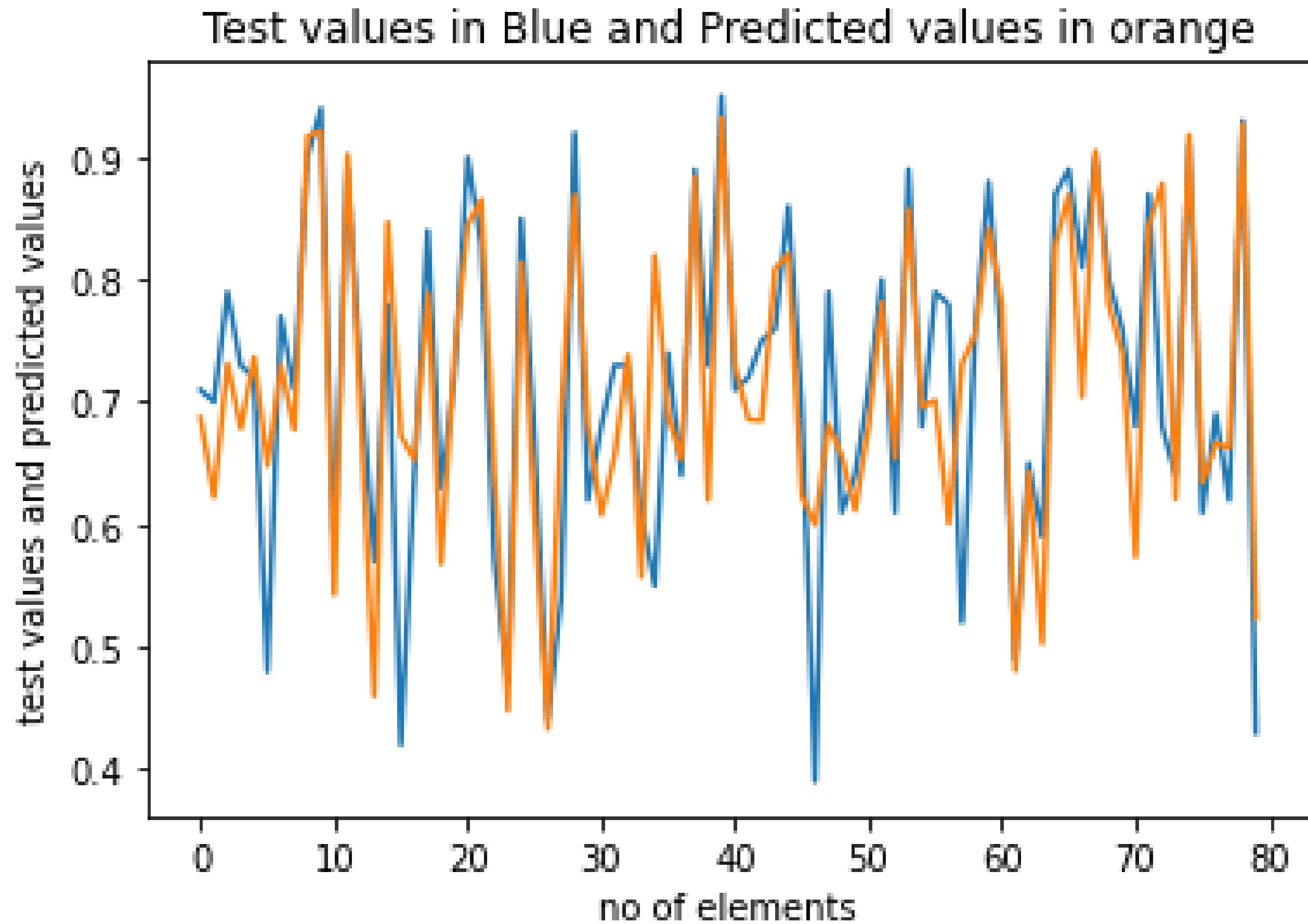


• RANDOM FOREST • REGRESSION

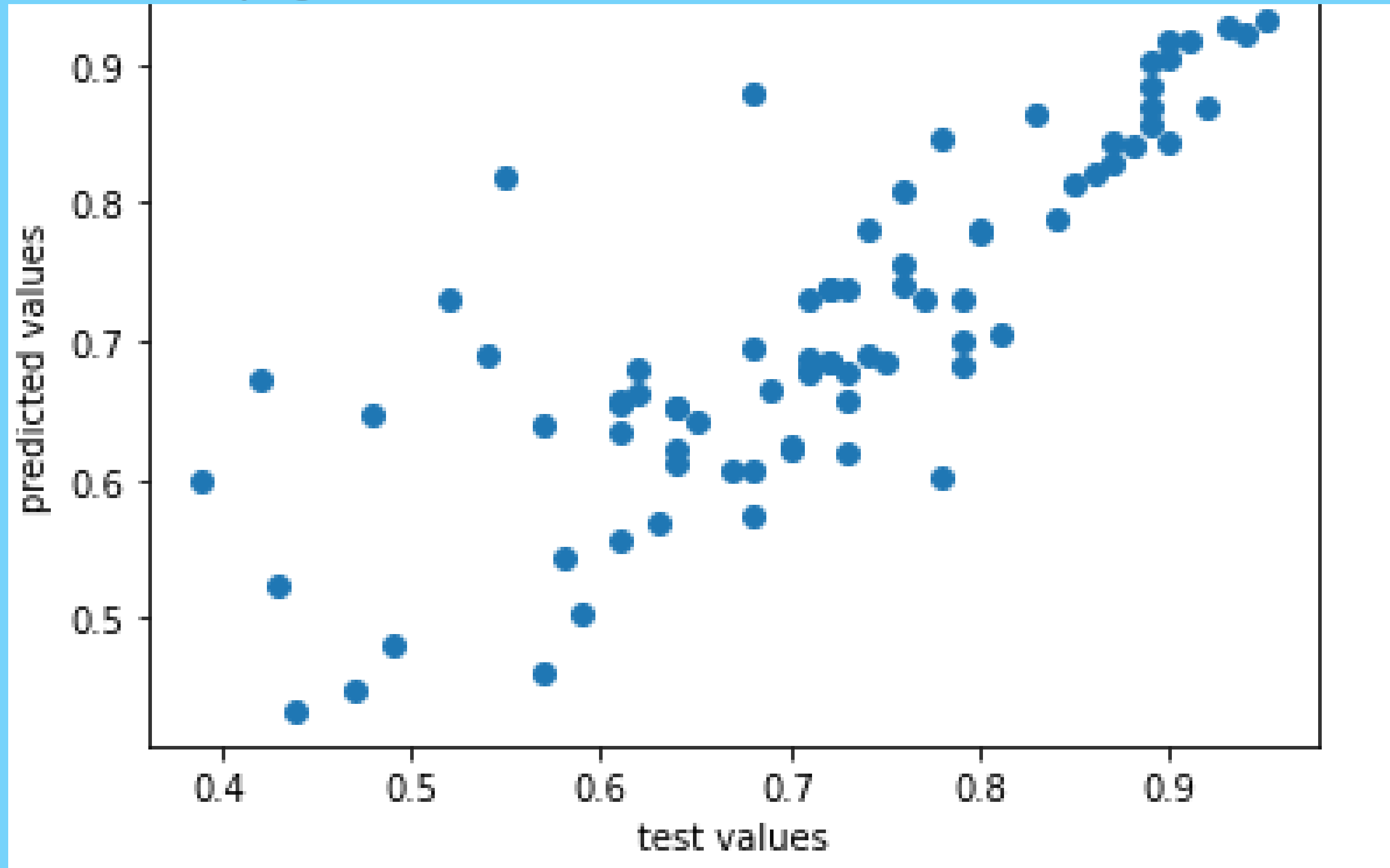
IMPORTING THE MODEL

```
RF = RandomForestRegressor(n_estimators = 100)
RF.fit(X_train,y_train)
randfor= RF.predict(X_test)
```

VISUALIZATION



INFERENCE



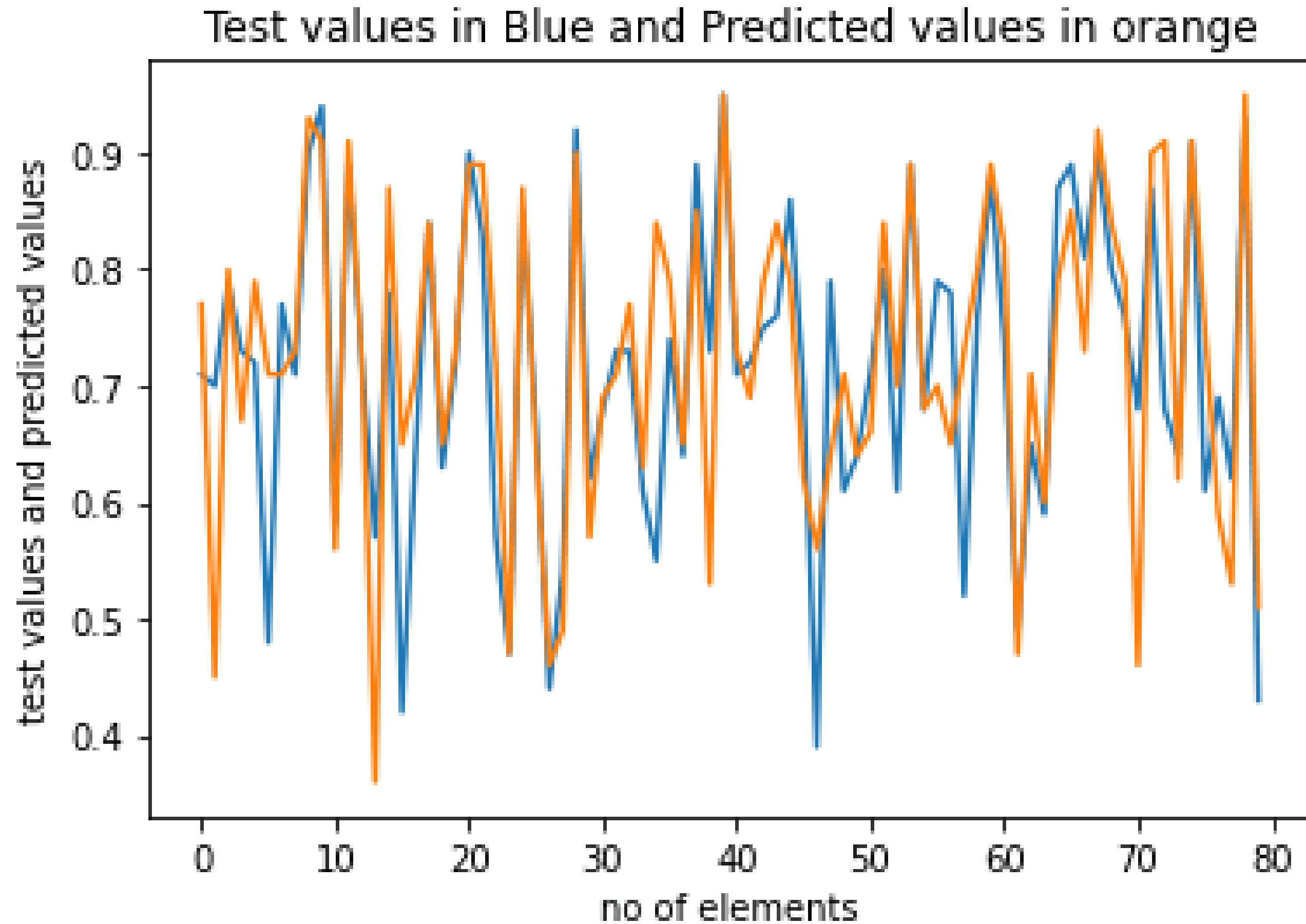
the mean squared error is 0.006644478624999991

- **DECISION TREE**
- **REGRESSION**

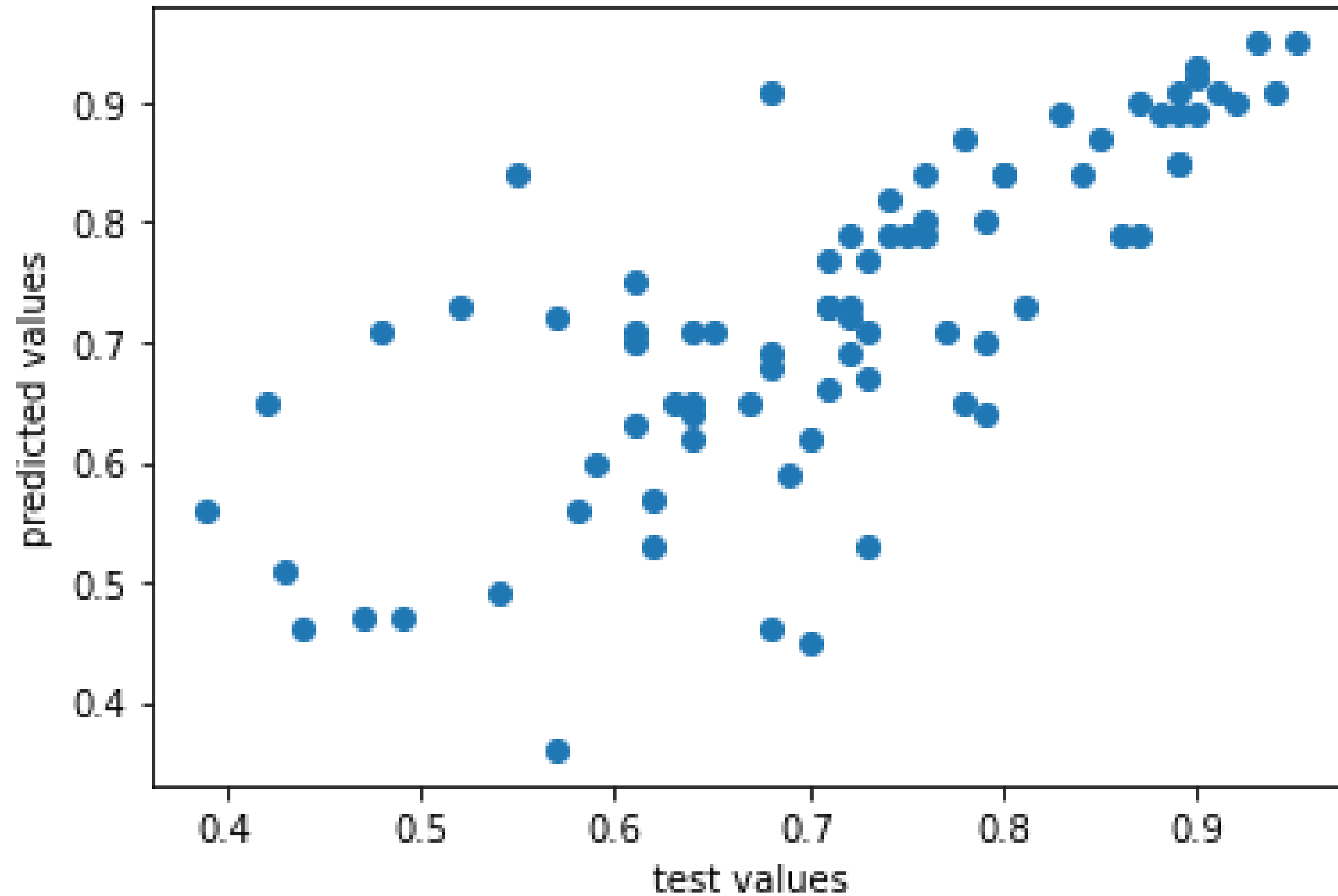
IMPORTING THE MODEL

```
DTR = DecisionTreeRegressor()  
DTR.fit(X_train,y_train)  
dectree= DTR.predict(X_test)
```

VISUALIZATION



INFERENCE



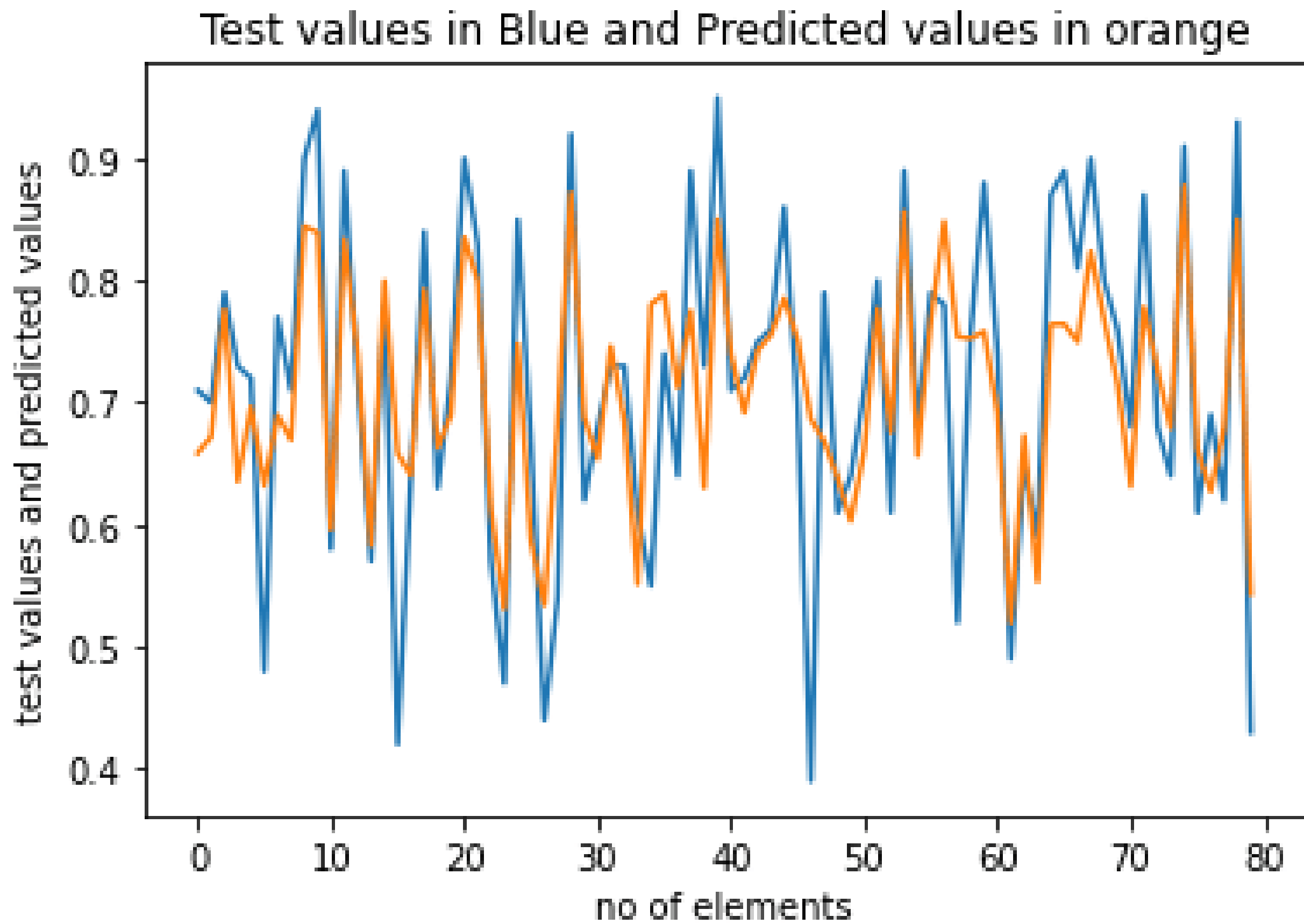
the mean squared error is 0.0093525

- # SUPPORT VECTOR REGRESSION

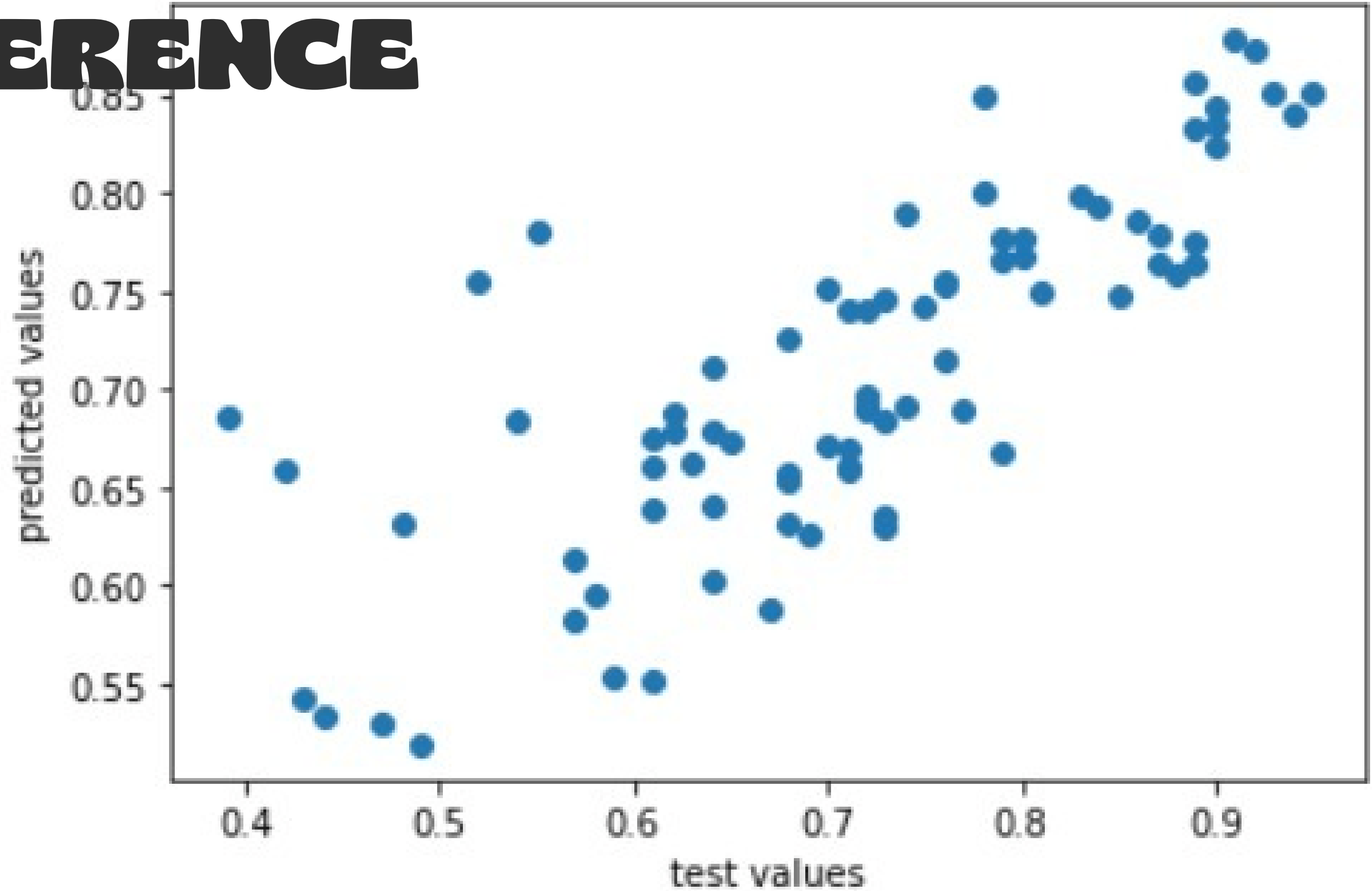
IMPORTING THE MODEL

```
from sklearn.svm import SVR  
regressor=SVR(kernel='rbf')  
regressor.fit(X_train,y_train)
```

VISUALIZATION

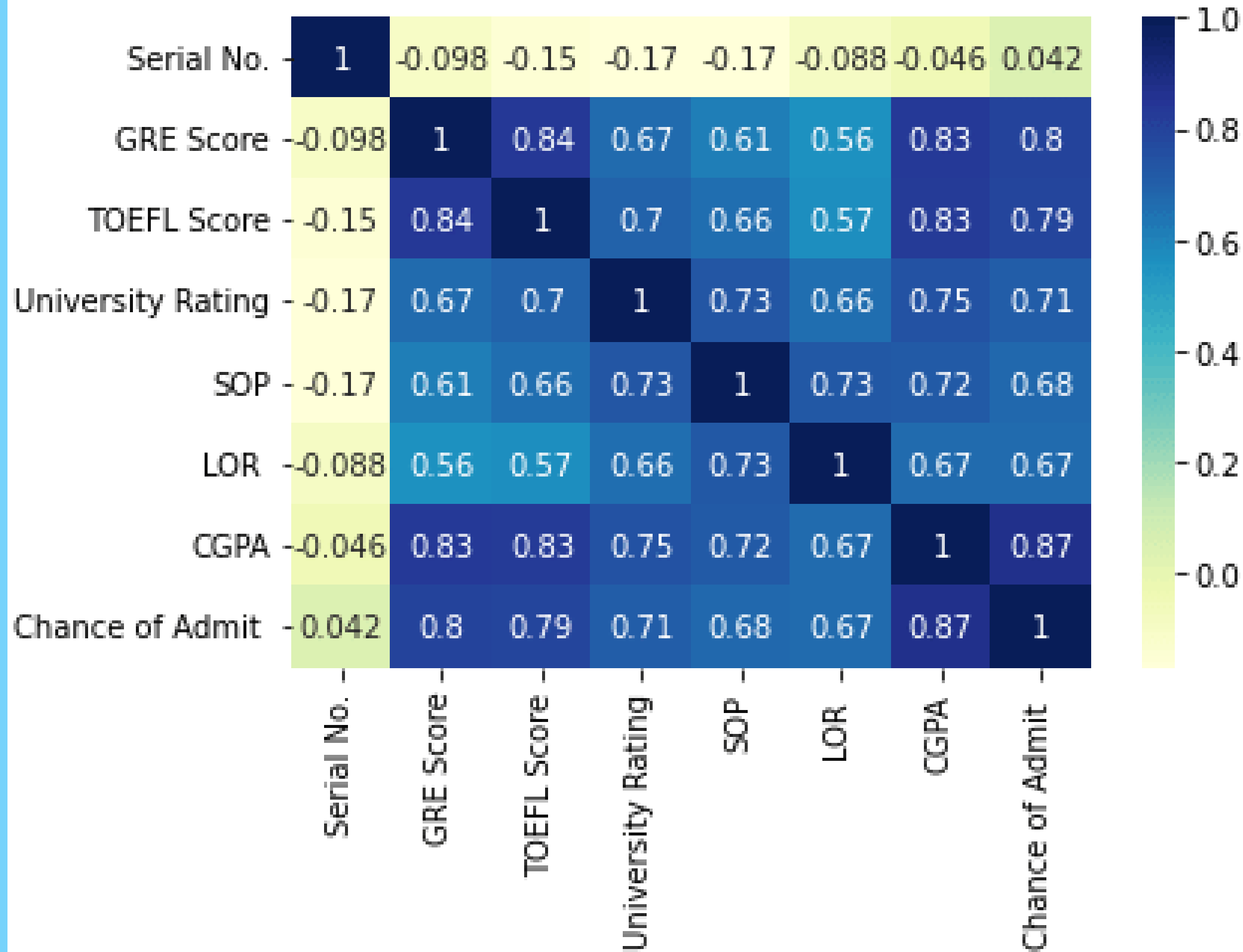


INFERENCE



the mean squared error is 0.00719784640281392

CORRELATION MATRIX





CONCLUSION

MEAN SQUARED ERROR OF MODELS:

- MULTILINEAR REGRESSION: 0.0049
- KNN REGRESSION: 0.0104
- RANDOM FOREST TREES: 0.0066
- DECISION TREES: 0.0093
- SUPPORT VECTOR REGRESSION: 0.007