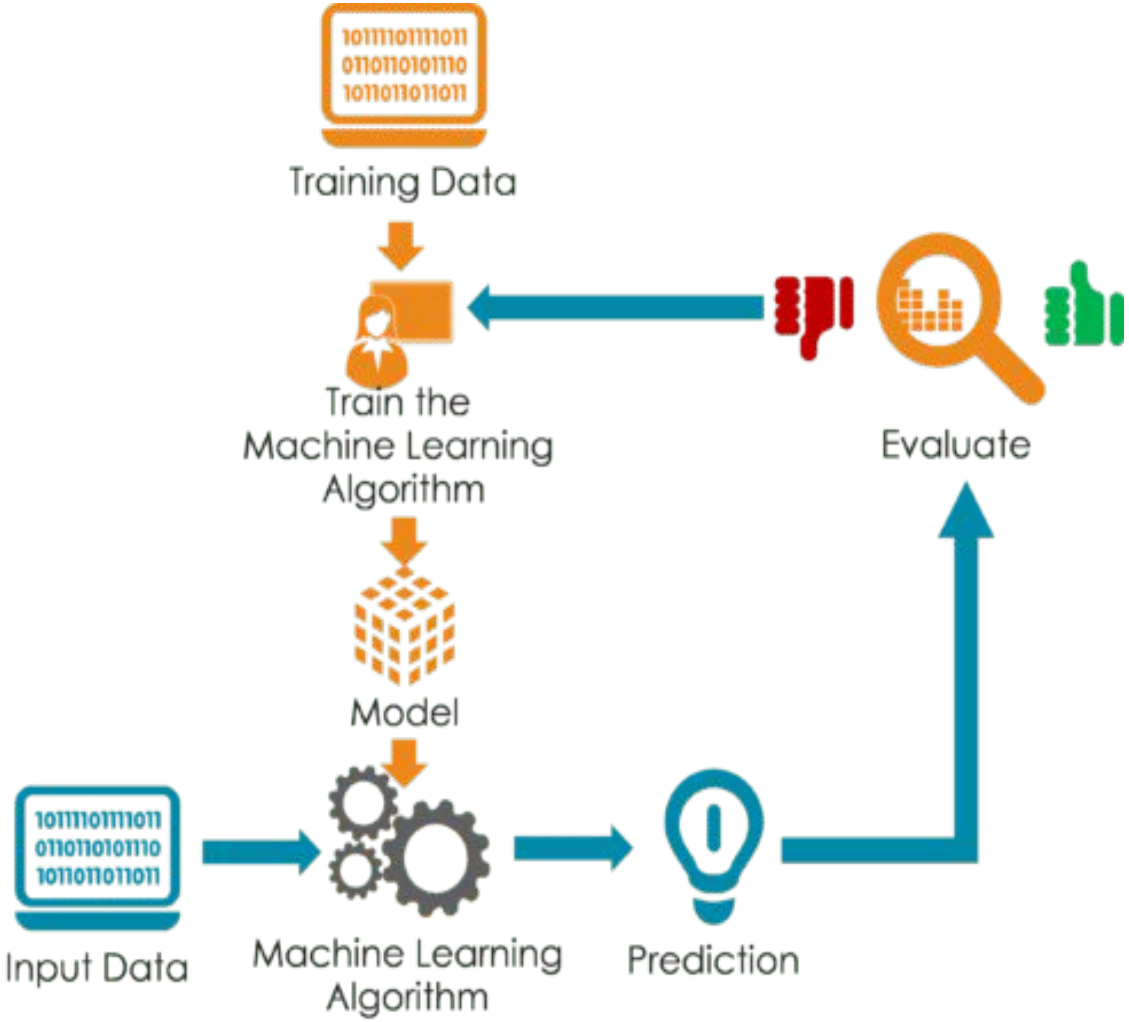
Heart Disease Prediction Using Machine Learning

By

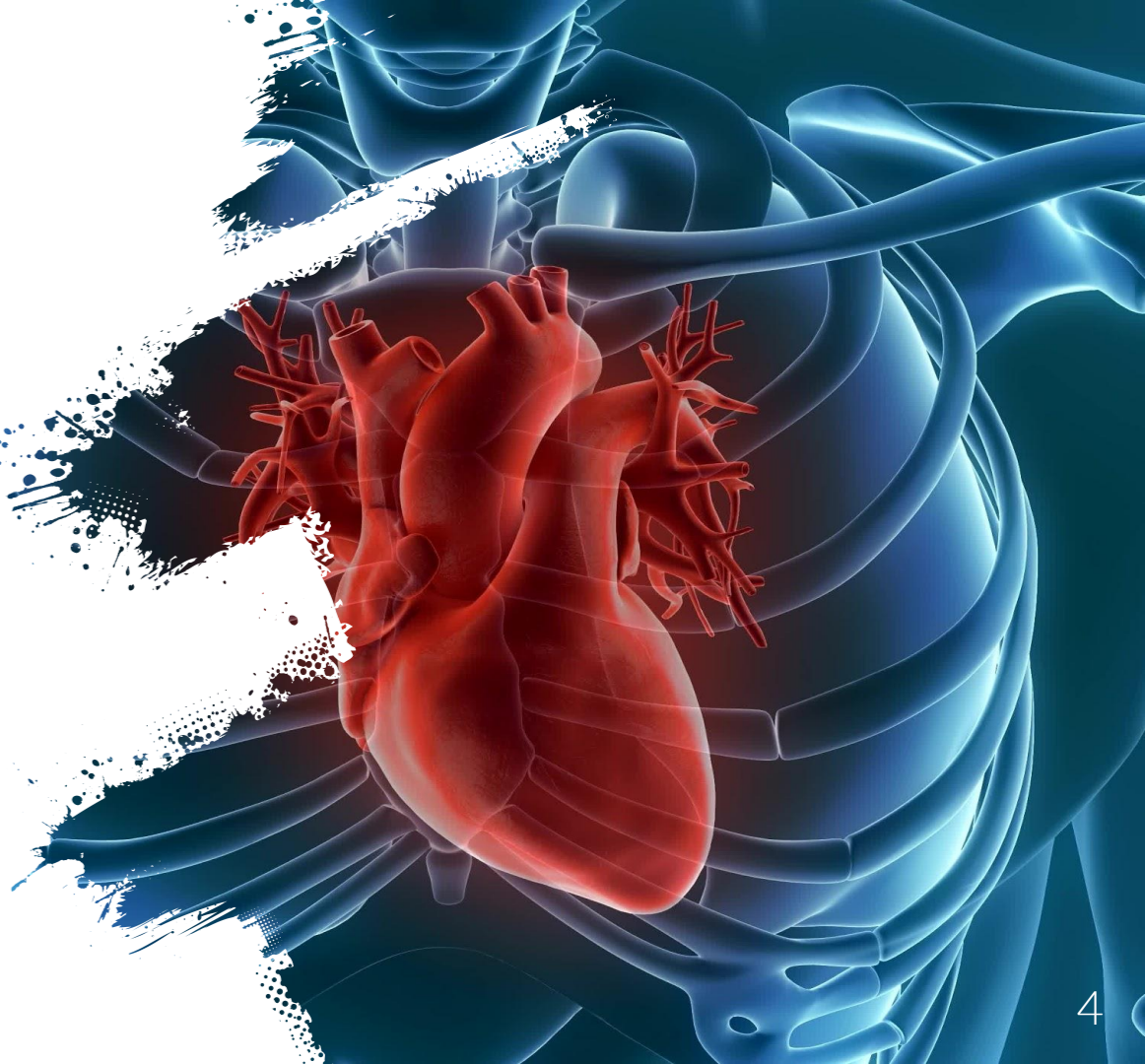
Shivam Vatshayan

INTRODUCTION

- What is Machine Learning ?
- Goal of the Project
- What is Supervised Learning ?


Training Data
Train the Machine Learning Algorithm
Evaluate
Model
Input Data
Machine Learning Algorithm
Prediction

The Working of the Project



Dataset

heart

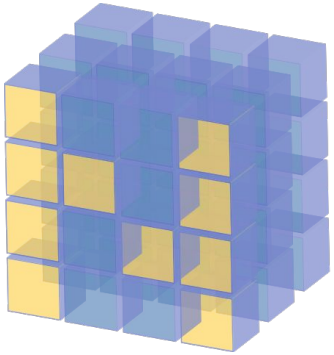
age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
44	1	1	120	263	0	1	173	0	0	2	0	3	1
52	1	2	172	199	1	1	162	0	0.5	2	0	3	1

3 step process

Split the dataset

Train the dataset

Compare the Algos

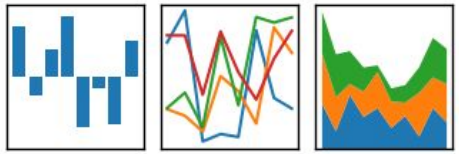


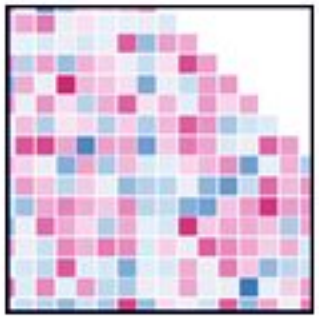
scikit learn

NumPy

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$

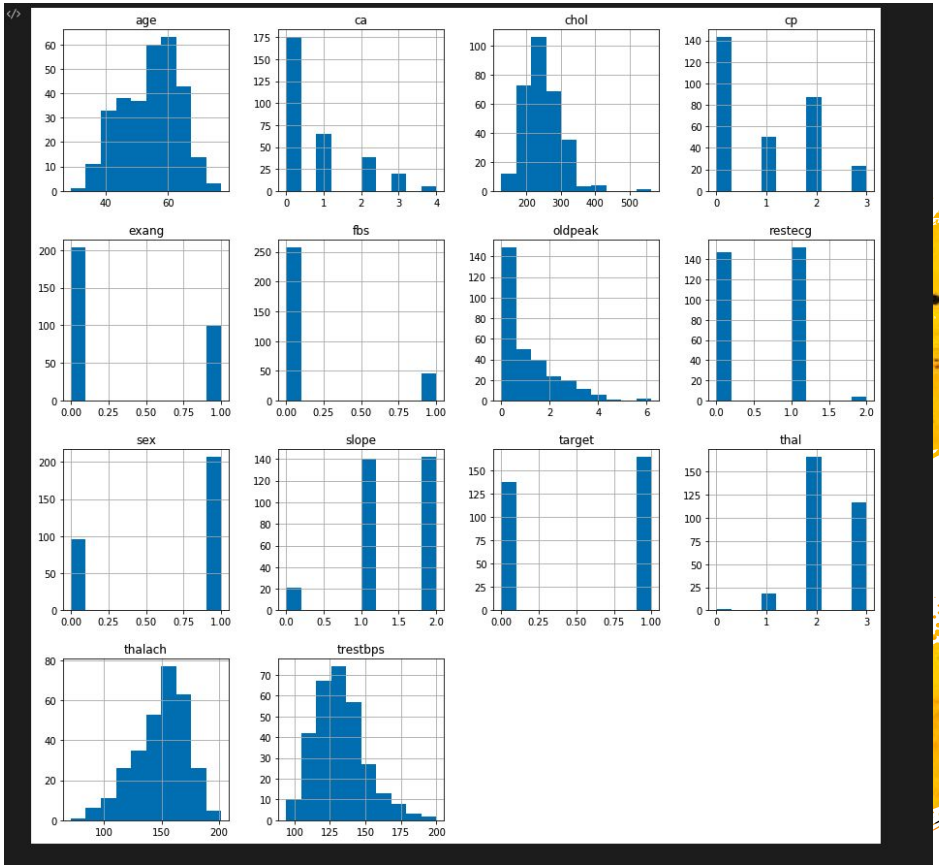




Seaborn

matplotlib

Data visualization



KNeighbours clasifier

```
# Training the knn classifier model with k value as 12
knn_classifier = KNeighborsClassifier(n_neighbors=12)
cvs_scores = cross_val_score(knn_classifier, X, y, cv=10)
print("KNeighbours Classifier Accuracy with K=12 is: {}%".format(round(cvs_scores.mean(), 4)*100))
```

```
KNeighbours Classifier Accuracy with K=12 is: 85.07000000000001%
```

Random Forest

```
# Training the random forest classifier model with n value as 90
forest_classifier = RandomForestClassifier(n_estimators=90)
cvs_scores = cross_val_score(forest_classifier, X, y, cv=5)
print("Random Forest Classifier Accuracy with n_estimators=90 is: {}%".format(round(cvs_scores.mean(), 4)*
```

```
Random Forest Classifier Accuracy with n_estimators=90 is: 83.82%
```

Decision Tree Classifier

```
# Training the decision tree classifier model with max_depth value as 3
decision_classifier = DecisionTreeClassifier(max_depth=3)
cvs_scores = cross_val_score(decision_classifier, X, y, cv=10)
print("Decision Tree Classifier Accuracy with max_depth=3 is: {}%".format(round(cvs_scores.mean(), 4)*100)
```

```
Decision Tree Classifier Accuracy with max_depth=3 is: 78.77%
```

•Conclusion

We have best three algorithms for predicting the heart disease through given datasets. We are also getting best accuracy. It is one of the new system that use three famous algorithm for prediction of diseases especially heart.


thank
you