end
matrices
process
findings
EDA
intro

A Classification Model for Predicting Liver Patients

BY NAIF ALZAHRANI

The Idea

- Liver is responsible for up to 500 separate vital functions.
  - Fighting infections
  - Transforming food into energy
  - Dispose of waste
- Risk from a range of factors
  - Poor diet
  - Alcohol
  - Genetic factors
- Liver damage
  - Cirrhosis
  - Hepatitis
  - Cancer
- Early diagnosis of hepatitis offers wider treatment options

Features
1

1. Age
2. Gender
3. Total Bilirubin
4. Direct Bilirubin
5. Alkaline Phosphatase
6. Alanine Aminotransferase
7. Aspartate Aminotransferase
8. Total Proteins
9. Albumin
10. Albumin and Globulin Ratio

Target
2

Result

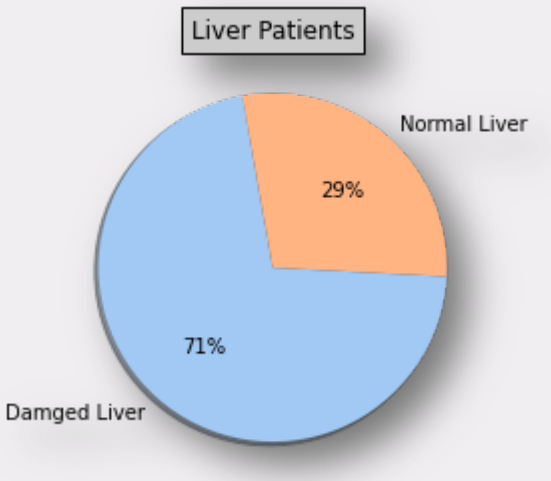
(1) Liver Patient
(2) non-Liver Patient.

Preparation
3

Cleaning

1. Fixing Columns
2. Dropping Nulls
3. Data Scaling
4. Pandas Profile

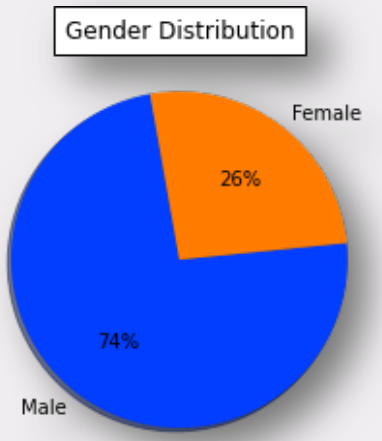
Balance



Records

30691 Rows
21917 LD
8774 None

Gender



01

k-nearest neighbors

The maximum score is 0.97 at K = 1

	precision	recall	f1-score
1	0.96	0.97	0.97
2	0.92	0.91	0.91
accuracy			0.95
macro avg	0.94	0.94	0.94
weighted avg	0.95	0.95	0.95

03

Random Forest

Scored 0.83 at depth 5

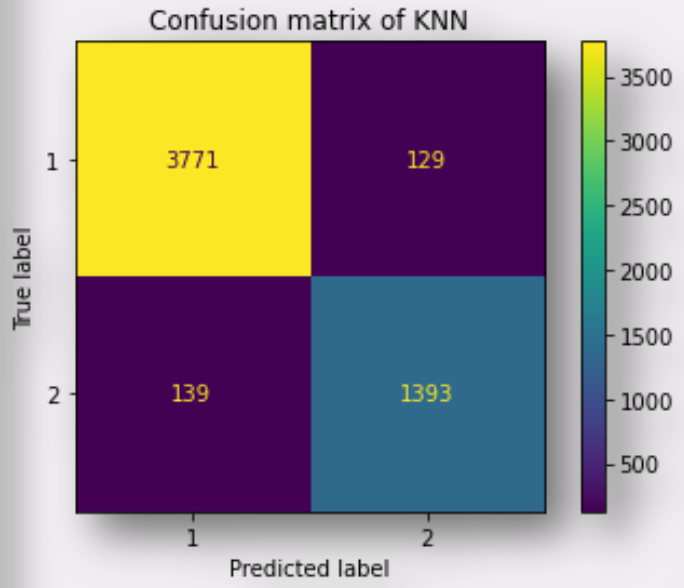
	precision	recall	f1-score
1	0.81	0.99	0.89
2	0.94	0.42	0.58
accuracy			0.83
macro avg	0.88	0.70	0.74
weighted avg	0.85	0.83	0.80

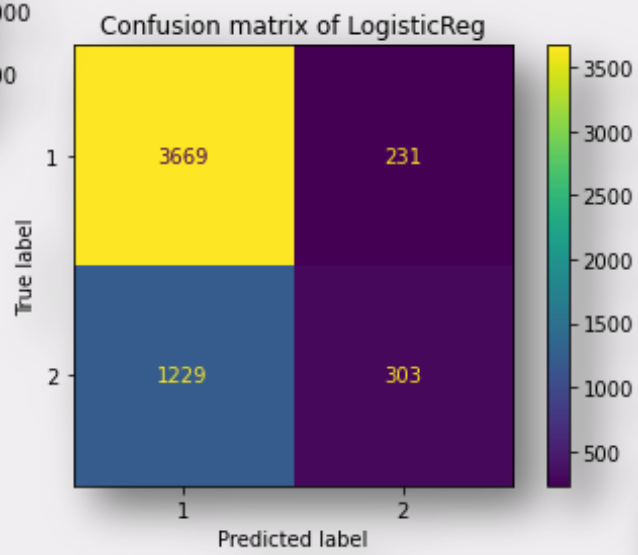
02

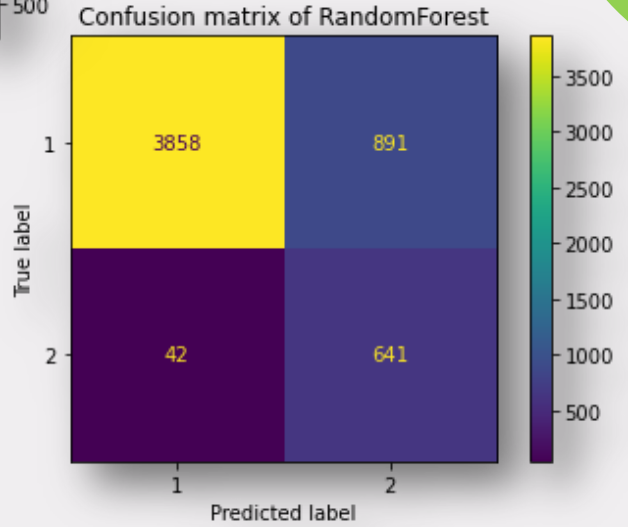
Logistic Regression

Scored 0.73

	precision	recall	f1-score
1	0.75	0.94	0.83
2	0.57	0.20	0.29
accuracy			0.73
macro avg	0.66	0.57	0.56
weighted avg	0.70	0.73	0.68


Confusion matrix of KNN
3771
129
139
1393
True label
Predicted label
1
2
3500
3000
2500
2000
1500
1000
500


Confusion matrix of LogisticReg
3669
231
1229
303
True label
Predicted label
1
2
3500
3000
2500
2000
1500
1000
500


Confusion matrix of RandomForest
3858
891
42
641
True label
Predicted label
1
2
3500
3000
2500
2000
1500
1000
500

Conclusion:

- KNN
- Implementation of such model and Early diagnosis.

Thoughts:

- Articulate the problem early.
- Check your data quality.
- Understand the data.
- Prepare your data.
- Apply your ML algrithms.