# LIVER-PATIENT-ANALYSIS

[![author](https://img.shields.io/badge/author-KHAGGA VENKATRA NARASIMHAM-ff69b4.svg?style=flat-square)](https://www.linkedin.com/in/KHAGGA VENKATRA NARASIMHAM/)

[![GitHub followers](https://img.shields.io/github/followers/KHAGGA VENKATRA NARASIMHAM21?style=social)](https://github.com/KHAGGA VENKATRA NARASIMHAM21?tab=followers)

[![GitHub watchers](https://img.shields.io/github/watchers/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?style=social)](https://github.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/watchers)

[![GitHub stars](https://img.shields.io/github/stars/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?style=social)](https://github.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/stargazers)

[![GitHub forks](https://img.shields.io/github/forks/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?style=social)](https://github.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/network/members)

![GitHub language count](https://img.shields.io/github/languages/count/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?style=flat-square)

![GitHub top language](https://img.shields.io/github/languages/top/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?logoColor=9cf&style=flat-square)

![GitHub repo size](https://img.shields.io/github/repo-size/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?logoColor=important&style=flat-square)

[![GitHub issues](https://img.shields.io/github/issues/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?style=flat-square)](https://github.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/issues?q=is%3Aopen+is%3Aissue)

[![GitHub closed issues](https://img.shields.io/github/issues-closed/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?style=flat-square)](https://github.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/issues?q=is%3Aissue+is%3Aclosed)

[![GitHub pull requests](https://img.shields.io/github/issues-pr/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?logoColor=yellow&style=flat-square)](https://github.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/pulls?q=is%3Aopen+is%3Apr)

[![GitHub closed pull requests](https://img.shields.io/github/issues-pr-closed/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS?logoColor=yellow&style=flat-square)](https://github.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/pulls?q=is%3Apr+is%3Aclosed)

[![LICENSE](https://img.shields.io/dub/l/vibe-d.svg?style=flat-square)](https://github.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/blob/master/LICENSE)

[![HitCount](http://hits.dwyl.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS.svg)](http://hits.dwyl.com/KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS)

### Objectives of Research

In India, delayed diagnosis of diseases is a fundamental problem due to a shortage of medical

professionals. A typical scenario, prevalent mostly in rural and somewhat in urban areas is:

1. A patient going to a doctor with certain symptoms.

2. The doctor recommending certain tests like blood test, urine test etc depending on the

symptoms.

3. The patient taking the aforementioned tests in an analysis lab.

4. The patient taking the reports back to the reports back to the hospital, where they are

examined and the disease is identified.

This project aims to reduce the time delay caused due to the unnecessary back and forth shuttling between the hospital and the pathology lab. Here a machine learning algorithm will be trained to predict a liver disease in patients using a data-set collected from North East of Andhra Pradesh, India.

### Problem Statement

The problem statement is formally defined as:

ÔÇÿGiven a dataset containing various attributes of 584 Indian patients, use the features

available in the dataset and define a supervised classification algorithm which can identify

whether a person is suffering from liver disease or not. This data set contains 416 liver patient

records and 167 non- liver patient records.The data set was collected from north east of Andhra

Pradesh, India. This data set contains 441 male patient records and 142 female patient records.

Any patient whose age exceeded 89 is listed as being of age "90"

### Team Members

- [@KHAGGA VENKATRA NARASIMHAM21](https://github.com/KHAGGA VENKATRA NARASIMHAM21) - \*\*KHAGGA VENKATRA NARASIMHAM\*\* (Project Head)

- [@Mallala Sivachand](https://github.com/AdityaSindol) - \*\* Mallala Sivachand \*\*

- [@Naidu Hemanthmanideep](https://github.com/keshavpola) - \*\*Naidu Hemanthmanideep\*\*

- [@Nalabothu Ravikishor](https://github.com/KaranDange) - \*\* Nalabothu Ravikishor \*\*

- [@Nidamanuri Naveenkumar](https://github.com/Veereshsg) - \*\* Nidamanuri Naveenkumar \*\*

<p align="center">

<img src="https://github.com/ KHAGGA VENKATRA NARASIMHAM21/Liver-Patient-Analysis/blob/master/Team%20Stark%20Group%20Pic.jpeg">

</p>

### Context

Patients with Liver disease have been continuously increasing because of excessive

consumption of alcohol, inhale of harmful gases, intake of contaminated food, pickles and

drugs. This dataset was used to evaluate prediction algorithms in an effort to reduce burden on

doctors.

### Content

This data set contains 416 liver patient records and 167 non liver patient records

collected from North East of Andhra Pradesh, India. The "Dataset" column is a class label used

to divide groups into liver patient (liver disease) or not (no disease). This data set contains 441

male patient records and 142 female patient records. Any patient whose age exceeded 89 is

listed as being of age "90".

### Columns:

\* Age of the patient

\* Gender of the patient

\* Total Bilirubin

\* Direct Bilirubin

\* Alkaline Phosphotase

\* Alamine Aminotransferase

\* Aspartate Aminotransferase

\* Total Protiens

\* Albumin

\* Albumin and Globulin Ratio

\* Dataset: field used to split the data into two sets (patient with liver disease, or no disease)

### Screenshots

­ƒæë \*\*Node Red Flow\*\*

![Noderedflow](https://github.com/ KHAGGA VENKATRA NARASIMHAM21/LIVER-PATIENT-ANALYSIS/blob/master/Screenshots/Node%20Red%20Flow.png)

­ƒæë \*\*ML Model UI\*\*

![MLModelUI](https://github.com/ KHAGGA VENKATRA NARASIMHAM 21/LIVER-PATIENT-ANALYSIS/blob/master/Screenshots/ML%20Modek%20UI.png)

### Conclusion

Initially, the dataset was explored and made ready to be fed into the classifiers. This

was achieved by removing some rows containing null values, transforming some columns

which were showing skewness and using appropriate methods (Label Encoding) to convert

the labels so that they can be useful for classification purposes. Performance metrics on which

the models would be evaluated were decided. The dataset was then split into a training and

testing set.

Firstly, a naive predictor and a benchmark model ('Logistic Regression') were run on

the dataset to determine the benchmark value of accuracy. The greatest difficulty in the

execution of this project was faced in two areas- determining the algorithms for training and

choosing proper parameters for fine-tuning. Initially, I found it very vexing to decide upon 3

or 4 techniques out of the numerous options available in sklearn.

This exercise made me realize that parameter tuning is not only a very interesting but

also a very important part of machine learning. I think this area can warrant further

improvement, if we are willing to invest a greater amount of time as well as computing power.