



Predicting Liver Cirrhosis Stage Based on Clinical and Biochemical Features

Introduction

- Liver cirrhosis is a chronic liver disease where healthy liver tissue is replaced with scar tissue (fibrosis).
- This scarring blocks blood flow through the liver and impairs its function.
- It often results from long-term damage due to:
 - Chronic alcohol abuse
 - Hepatitis B/C infections
 - Non-alcoholic fatty liver disease (NAFLD)
- Symptoms may include fatigue, jaundice, abdominal swelling, and confusion.
- Cirrhosis is a progressive disease and can lead to liver failure or liver cancer if untreated



Dataset Overview



DATASET: CIRRHOSIS.CSV



RECORDS: ~312 PATIENTS,
20 FEATURES.



TARGET VARIABLE: 'STAGE'
(SEVERITY OF CIRRHOSIS).

Stages of Liver Disease



Healthy Liver

100% of Liver
Function



Fatty Liver

75-80% of Liver
Function



Liver Fibrosis

20-30% of Liver
Function

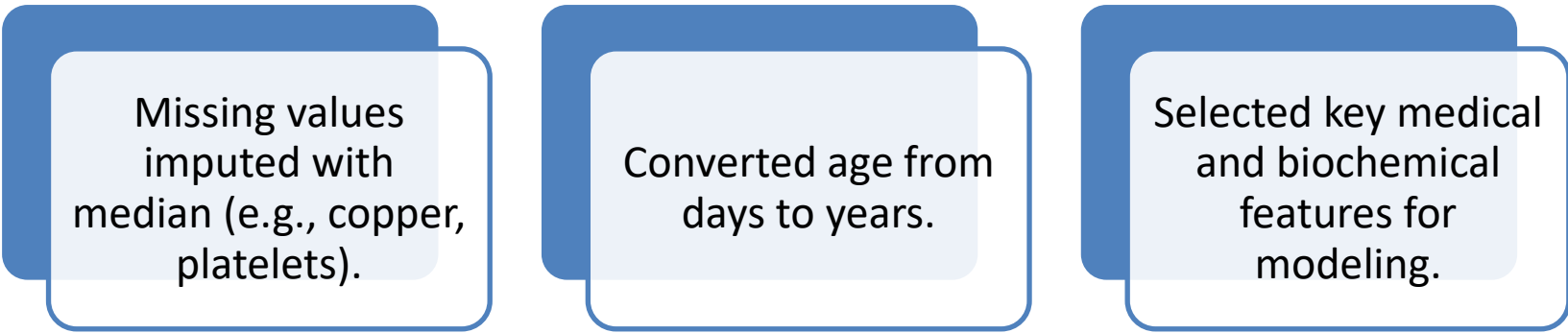


Cirrhosis

2-7% of Liver
Function



Data Cleaning & Preprocessing



Missing values
imputed with
median (e.g., copper,
platelets).

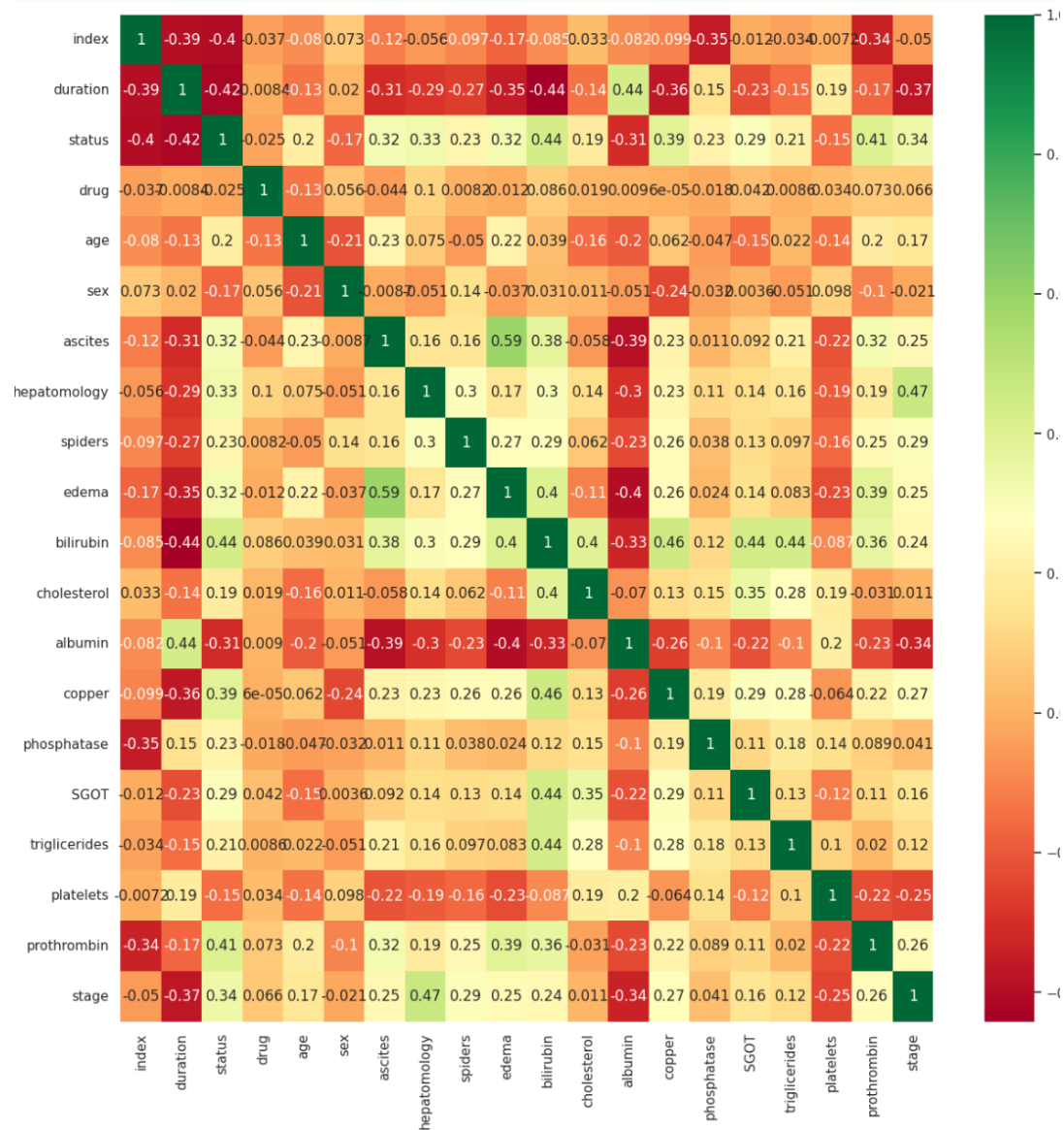
Converted age from
days to years.

Selected key medical
and biochemical
features for
modeling.

Exploratory Data Analysis

- Correlation heatmap shows relationship between features.
- Strong correlation between bilirubin and stage.
- Identified most predictive variables.



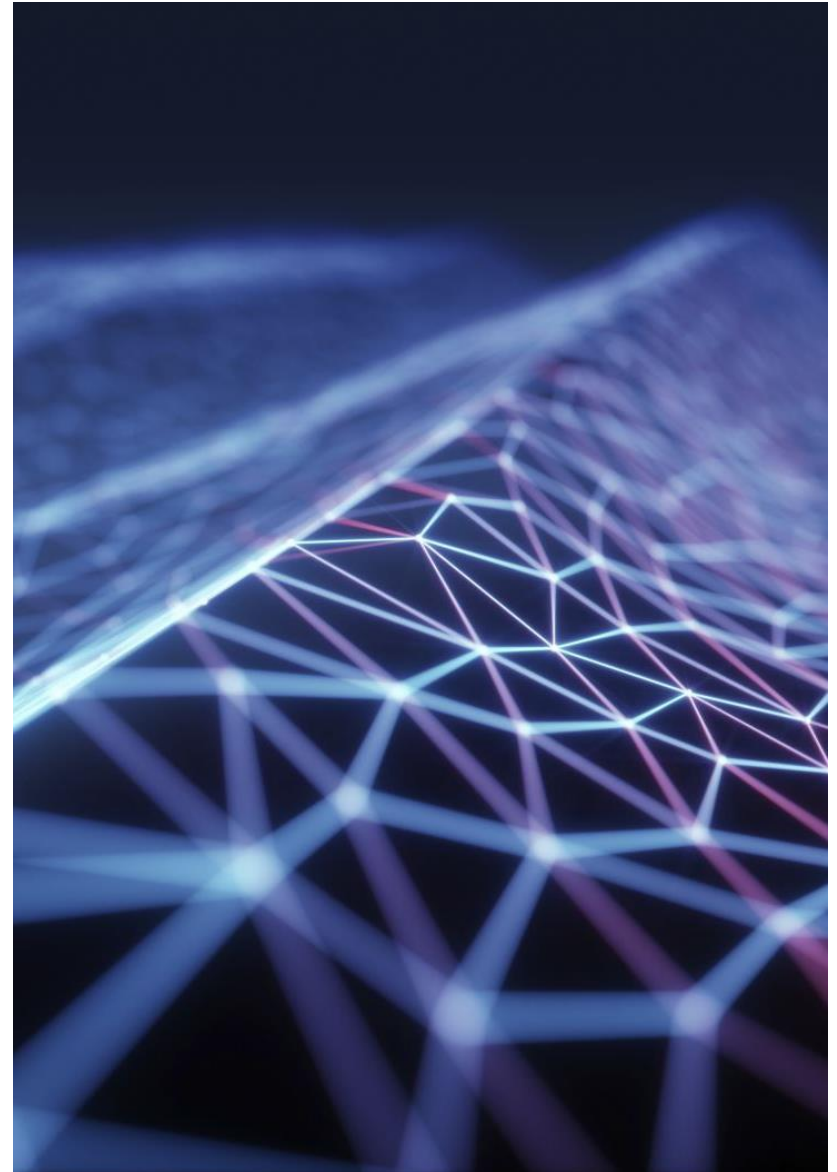


Model Preparation

X: Selected features like age, bilirubin, copper, etc.

y: Target variable 'stage'.

Train-test split: 80% training, 20% testing.



Modeling Approach



Used custom-built neural
model(ANN) with ReLU activation.

Evaluation Metrics



ACCURACY TO EVALUATE
PREDICTIONS.



R^2 SCORE USED TO
MEASURE REGRESSION
PERFORMANCE.



POTENTIAL USE OF
CONFUSION MATRIX (NOT
SHOWN HERE).

Results & Interpretation

Model successfully predicted some stages accurately.

Stage misclassifications occurred in borderline cases.

```
[231] # Accuracy  
      acc_train= accuracy(y_train, y_train_pred)  
      acc_test = accuracy(y_test, y_test_pred)  
      acc_train
```

```
np.float64(0.41767068273092367)
```

```
[232] acc_test
```

```
np.float64(0.3968253968253968)
```

Train Scores:
Precision: 0.45

Test Scores:
Precision: 0.4

Challenges & Improvements



SMALL DATASET SIZE
LIMITS ACCURACY.



INVESTING MORE TIME
ON THIS DATASET IS
REQUIRED



MISSING DATA
IMPUTATION MAY
INTRODUCE BIAS.



FUTURE: USE RANDOM
FORESTS, XGBOOST, OR
DEEP LEARNING.

Conclusion



Machine Learning
can support early
diagnosis.



Accurate staging
helps guide
treatment decisions.



Further research can
enhance model
robustness.