



International Conference on Machine Learning and Data Engineering (ICMLDE 2023)  
**Analyzing the Impact of COVID-19 on Liver Cancer: A  
Comprehensive Study using XGBoost Classifier and Feature  
Selection Techniques**

Rekha R Nair<sup>a</sup>, Tina Babu<sup>a</sup>, Kishore S<sup>b</sup>, Vishnu Vilashini S<sup>c</sup>

<sup>a</sup>Department of Computer Science and Engineering, Alliance University, Bengaluru, India

<sup>b</sup>Department of Computer Applications, Nitte Institute of Professional Education, Nitte University, Karnataka, India

<sup>c</sup>Department of Computer Applications, School of Engineering, Dayananda Sagar University, Bengaluru, Karnataka, India

---

**Abstract**

Healthcare systems around the world have faced challenges due to the Coronavirus disease (COVID-19) epidemic, which has taken resources and attention away from long-term diseases like liver cancer. To identify the effect of COVID-19 on liver cancer, proposed an effective research work with Extreme Gradient Boosting (XGBoost) classifier. The identification of the most suitable feature is executed through the utilization of SelectKBest, Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) techniques. A comparative analysis is carried out to ascertain the more precise feature selection, employing Decision Tree, K-Nearest Neighbors (KNN), CatBoost, LightGBM, and XGBoost algorithms. Further, the proposed work undertakes an analysis of the impact of COVID-19 on liver cancer, encompassing both the pandemic and pre-pandemic periods. In comparison to other feature selection techniques and models, SelectKBest feature selection with Xgboost classifier provided an accuracy of 0.92. The proposed research work provides valuable insights to healthcare professionals and policy makers, providing a better understanding of the challenges faced by liver cancer patients during the pandemic. By meticulously exploring feature selection techniques and employing advanced machine learning models, this study quantifiably demonstrates the effectiveness of the proposed methodology in addressing the impact of COVID-19 on liver cancer, contributing to a more comprehensive understanding of this critical issue.

© 2024 The Authors. Published by ELSEVIER B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the International Conference on Machine Learning and Data Engineering

*Keywords:* XGBoost Classifier; SelectKBest; Singular Value Decomposition; Principal Component Analysis; hepatocellular carcinoma

---

**1. Introduction**

Precise risk stratification of individuals in critical condition due to coronavirus disease 2019 (COVID-19) is imperative to effectively allocate resources, implement specific interventions, and enhance the likelihood of patient survival[1]. Ensuring an accurate and dependable diagnosis is of paramount importance to promptly administer med-

---

*E-mail address:* rekha.nair@alliance.edu.in

ical assistance to individuals displaying symptoms or confirmed infection, thereby assisting governmental organizations in curbing its transmission and safeguarding human lives[2, 3, 4]. The biomedical area has been significantly impacted by machine learning for the diagnosis and prognosis of liver disease [5, 6, 7, 8, 9]. Machine learning guarantees improved disease identification and prediction in the biomedical industry as well as increased decision-making objectivity [10, 11].

A key focus of this research is the evaluation of the precision of predictive models concerning liver cancer outcomes[12]. The evaluation of accuracy in this study is achieved through the utilization of feature selection techniques including PCA, SVD, and SelectKBest[13]. The primary objective is to pinpoint the most insightful attributes that play a pivotal role in predicting liver cancer outcomes within the backdrop of COVID-19. These dependable and precise predictive models stand to significantly augment decision-making within the domain of treating liver cancer patients impacted by the pandemic. This enhancement empowers healthcare practitioners to make well-informed decisions concerning patient care and to optimize strategies for treatment. This research aims to investigate the specific implications of the COVID-19 pandemic on the detection of liver cancer, with a focus on hepatocellular carcinoma(HCC)[14, 15, 16]. By analyzing a comprehensive dataset of HCC patients affected by COVID-19, the study seeks to identify the influence of the pandemic on liver cancer detection practices[17]. The research aims to evaluate the effectiveness of machine learning algorithms in detecting HCC at an early stage for evaluating existing strategies for liver cancer detection during and after the pandemic. The main objectives of this research are as follows:

- Implement various pre-processing techniques to obtain more accurate result.
- To evaluate Covid-19's effects on liver cancer during and before an epidemic.
- To detect the accuracy of predictive models with different feature extraction techniques like SelectKBest, SVD and PCA.

The subsequent section of the paper is outlined as follows. Techniques utilized for this work is explained in Section 2. The result and its analysis with various techniques is provided in Section 3 and conclusion of the proposed work with future findings is on Section 4.

## 2. Methodology

The proposed system comprises a range of components and processes designed to assess the influence of the COVID-19 pandemic on the occurrence of liver cancer. The system follows a systematic workflow to collect, pre-process, analyze, and evaluate data related to newly diagnosed liver cancer patients during the pandemic. The block diagram shown in Fig. 1 provides an overview of the proposed system.

### 2.1. Data acquisition

The Dataset "COVID-19 effect on Liver Cancer Prediction Dataset" is collected from Kaggle [17]. It comprises 450 rows and encompasses 27 columns, each serving as a potential feature for liver cancer prediction. The dataset's key attributes are Cancer (Cancer flag), Year, Month, Bleed (Spontaneous tumor hemorrhage), Mode Presentation, Age, Gender, Cirrhosis (Underlying liver disease), Tumor diameter in mm) etc.

### 2.2. Data preprocessing

It is crucial to ensure the quality and suitability of the dataset for analysis. This step involves several sub-processes.

- **Handling Missing Values**

For numerical data, missing values is imputed using statistical method median and categorical data, the mode is utilized.

- **Label Encoding**

Utilising the label encoding technique, categorical variables were transformed into numerical values once the missing values had been filled in.

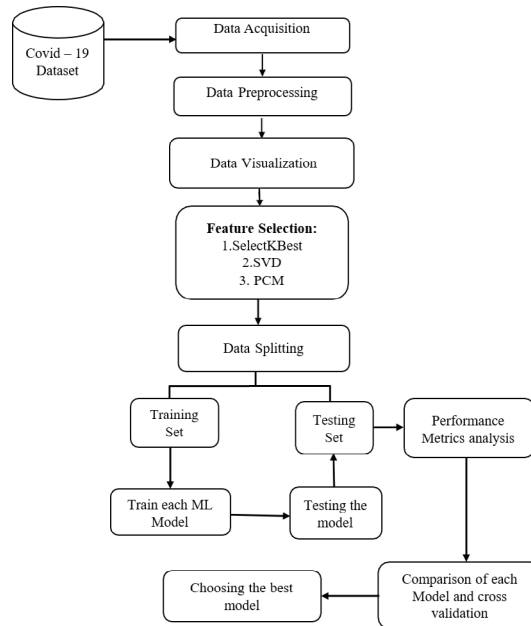


Fig. 1: Proposed Block diagram.

- **Outlier detection and removal**

Z score is used to detect and handle outliers in order to maintain the accuracy and reliability of the analysis after Label Encoding technique.

- **Normalization - Min-Max Scaling**

In order to ensure that all features contribute equally to the analysis and prevent features with huge value ranges from dominating the models, numerical data is rescaled using the min-max method to a range between 0 and 1.

### 2.3. Data visualization

It allows better understanding and analysis of data. One specific visualization technique employed in this study is pair plots, which allow for the comparison of various features with pandemic and pre-pandemic variables.

#### 2.3.1. Statistical Analysis

To assess the impacts of COVID-19 on liver cancer patients, a comprehensive statistical analysis conducted. This analysis aimed to evaluate any significant differences or associations between relevant variables using Proportion z-tests and chi-square tests. The statistical tests enabled to gain insights into the potential effects of COVID-19 on various aspects related to liver cancer.

### 2.4. Feature selection

The subset of variables that are most relevant and informative for predicting liver cancer during the COVID-19 pandemic were chosen using feature selection. The dimensionality of the dataset is decreased by choosing the most important features, which will improve model performance, interpretability, and computing efficiency. The following techniques are employed in this work:

- **SelectKBest:** This section analyses feature selection using SelectKBest and the F-test. The most informative features from the dataset that are significant to the target variable can be chosen using the SelectKBest algorithm with  $k$  value as 10.

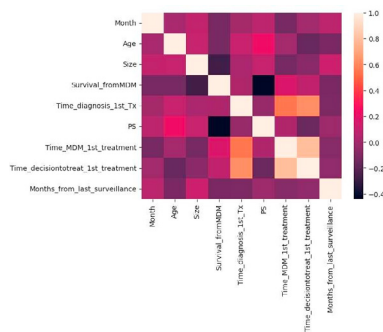


Fig. 2: Heat Map.

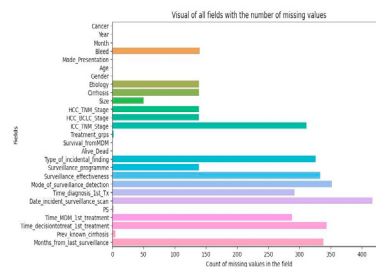


Fig. 3: Visualize the null values using a bar chart, illustrating the distribution of missing data values.

- **SVD:** The *n<sub>components</sub>* parameter is set to 10 in SVD, which indicates that the dimensionality of the data is reduced to 10 components.
- **PCA:** To maintain the top 10 principal components, the hyper-parameters used for PCA is *n<sub>components</sub>* = 10.

### 2.5. Model Training and Testing

The dataset is split into training (80%) and testing (20%) sets after preprocessing and feature selection. The models are trained on the preprocessed dataset and evaluated using appropriate techniques to assess their performance.

- **Training Set:** The training set consists of 80% of the preprocessed dataset.
- **Testing Set:** The training set consists of 20% of the preprocessed dataset.
- **Models Used:** The models utilized for comparative analysis are Decision Tree, KNN, CatBoost, LightGBM and XGBoost.

### 2.6. Experimental setup and Performance metrics

The authors utilized 8GB RAM windows operating system with intel processor. The software utilized is Jupyter Notebook with Python 3.7. Performance metrics such as Precision, F1 Score, Accuracy, and Recall are employed to assess the efficacy of machine learning models in predicting liver cancer outcomes amidst the backdrop of the COVID-19 pandemic K-fold cross-validation is applied to measure the performance with default K value[5].

## 3. Experiment Result and Analysis

### 3.1. Correlation Analysis

It helps to understand the degree of association between variables and identify potential patterns or dependencies. The correlations are visualized using a heatmap generated with the Seaborn library. Darker colors in heatmap indicate stronger correlations, while lighter colors suggest weaker or no correlation as shown in Fig. 2.

### 3.2. Visualizing Null Values

To evaluate the data quality and find potential gaps in the dataset, estimated the number of missing values for each field as shown in Fig. 3. The x-axis represents the count of missing values in each field, while the y-axis displays the field names.

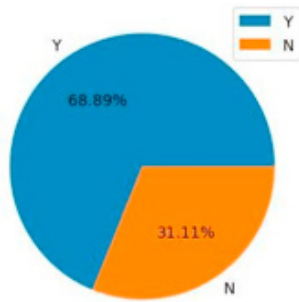


Fig. 4: Distribution of Cancer Patients in the Dataset.

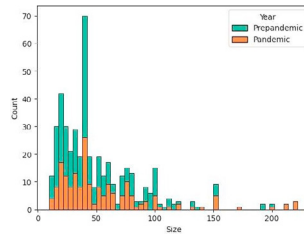


Fig. 5: Distribution of Tumor Sizes by Year of Diagnosis.

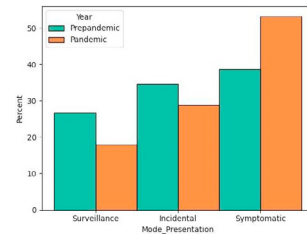


Fig. 6: Distribution of Cancer Presentation Modes during the COVID-19 Pandemic.

### 3.3. Data Visualization

#### 3.3.1. Cancer Patients Distribution in Dataset

A pie chart is created to visualize the distribution of cancer patients in the dataset as shown in the Fig. 4. The chart in the Fig. 4 shows that 69% of the data corresponds to cancer cases, while 31% represents non-cancer cases. This distribution indicates that the dataset is relatively balanced in terms of the presence of cancer cases.

#### 3.3.2. Effect of COVID-19 on Tumor Size

The histogram plot illustrates in Fig. 5 shows the distribution of tumor sizes in relation to the year of diagnosis. In the pre-pandemic period, a significant number of tumors were detected around the 50 mm range. However, during the pandemic, there was a noticeable decrease in the frequency of tumors within this size range. This observation suggests that the COVID-19 pandemic may have impacted the timely detection of tumors, potentially due to strained healthcare resources. Consequently, a higher proportion of larger tumors were identified during the pandemic period.

#### 3.3.3. Changes in Mode of Presentation of Cancer during the COVID-19 Pandemic

The distribution of various cancer presentation modes during the COVID-19 pandemic as compared to the pre-pandemic period is shown in the figure 6. It is evident that symptomatic tumors were more commonly detected during the pandemic, suggesting a potential change in the mode of cancer presentation. This could be attributed to factors such as delayed diagnosis, reduced access to medical facilities, and the prioritization of symptomatic cases during the COVID-19 pandemic.

#### 3.3.4. Impact of COVID-19 on Liver Cancer Incidence by Gender

The histogram plot in Fig. 7 illustrates how COVID-19 affects the incidence of liver cancer in both men and women. The x-axis represents the gender (Male or Female), and the y-axis represents the percentage of cases. The bars are grouped by the year, with different colors indicating the pandemic and pre-pandemic periods.

From the graph, it appears that during the pandemic, there was a higher increase in liver cancer cases among males compared to females.

#### 3.3.5. Did COVID increases the underlying liver diseases

The distribution of underlying liver disorders (Cirrhosis) throughout the pre-pandemic and pandemic periods is compared in the histogram in Fig.8.

A larger number of instances of underlying liver disease were found in the pre-pandemic period compared to the pandemic period, as shown by the graph in Fig.8. This suggests that the detection and diagnosis of underlying liver diseases may have been affected by the strain on the healthcare system during the pandemic.

#### 3.3.6. Impact of COVID on Previous Known Cirrhosis

Comparing known cases of cirrhosis from the pre-pandemic and pandemic periods is shown in Fig.9. The plot reveals that the number of known cirrhosis cases was higher in the pre-pandemic period compared to the pandemic

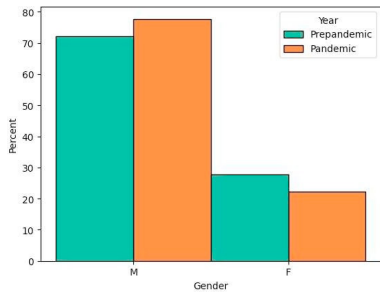


Fig. 7: Impact of COVID-19 on Liver Cancer Incidence by Gender.

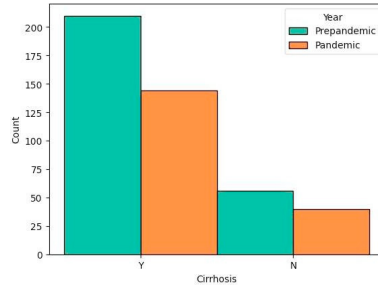


Fig. 8: Comparison of Underlying Liver Diseases (Cirrhosis) between Pre-Pandemic and Pandemic Periods.

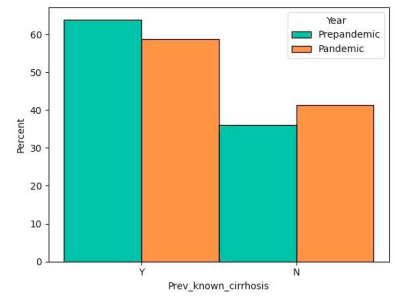


Fig. 9: Changes in Previous Known Cirrhosis Cases.

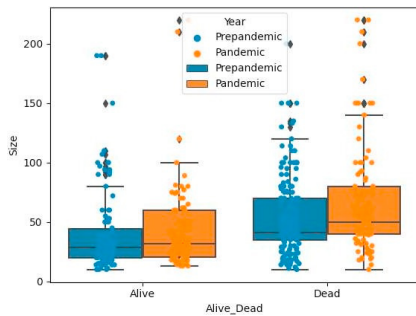


Fig. 10: Relationship between Tumor Size and Patient Survival.

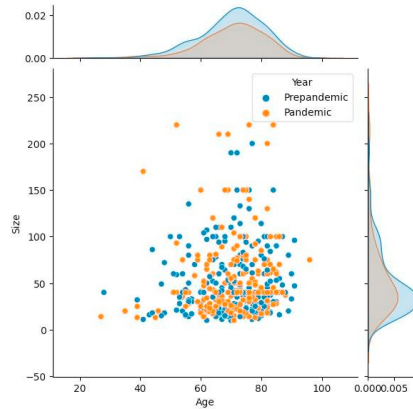


Fig. 11: Size vs Age Distribution.

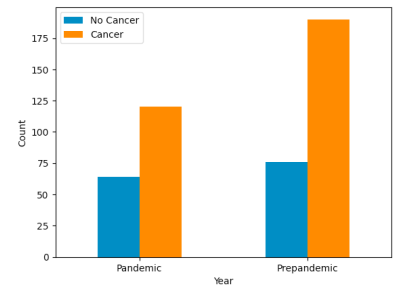


Fig. 12: Distribution of Cancer Cases in Pandemic and Pre-pandemic Periods.

period. Furthermore, there was an increase in the number of cases where cirrhosis was not known during the pandemic. This indicates that the detection and awareness of cirrhosis may have been impacted during the pandemic.

### 3.3.7. Impact of Tumor Size on Patient Survival

The stripplot and boxplot visualizations in Fig. 10 illustrates the association between tumor size and patient survival, considering both the pre-pandemic and pandemic periods. The stripplot reveals that patients who survived tend to have smaller tumors, with a concentration of data around 25 mm. On the other hand, patients who did not survive have a higher median tumor size, with the 50<sup>th</sup> percentile of data around 50 mm. Notably, the impact of the pandemic or pre-pandemic period does not appear to significantly influence this relationship.

### 3.3.8. Relationship between Tumor Size and Age

The jointplot visualizes the relationship between tumor size and patient age, considering both the pre-pandemic and pandemic periods. The plot shows the distribution of tumor sizes on the y-axis and patient ages on the x-axis. The hue parameter represents the different time periods, distinguishing between the pre-pandemic and pandemic data. The marginal ticks indicate the frequency of tumor sizes and ages separately. From the plot shown in Fig. 11, it can be observed that the distribution of tumor sizes varies across different age groups. However, the impact of the pandemic or pre-pandemic period on this relationship is not explicitly evident.

After analyzing the data, several key findings have been observed. The following are the conclusions based on the visualization:

- Tumors detected during the pre-pandemic period tended to be smaller in size compared to larger ones, suggesting variations in the stage of tumor detection between the two time periods.

- The number of symptomatic tumors increased during the pandemic due to delayed screenings, diagnosis and change of lifestyle, indicating a higher likelihood of tumors exhibiting noticeable symptoms during that time.
- There was a lower prevalence of underlying diseases and known cirrhosis cases during the pandemic period, potentially indicating changes in patient demographics or diagnostic patterns.
- Smaller tumors showed a higher possibility of patient survival, highlighting the importance of early detection and intervention.

### 3.4. Results of Statistical Analysis

- **Comparison result of cancer affected people pandemic and pre-pandemic.**

The null hypothesis (Ho) stated that the number of cancer cases during the pandemic is equal to the number of cancer cases during the prepandemic period. The alternative hypothesis (Ha) proposed that there is a difference. The cross-tabulation performed to examine the distribution of cancer cases based on the pandemic and prepandemic periods is shown in Table. 1 and Fig. 12.

The results of Proportion z tests: (-1.41, 0.16).

According to the findings, the p-value is less than 0.05 (the selected significance level), which indicates that the observed data strongly contradicts the null hypothesis. Reject the null hypothesis in this situation and reach the conclusion that the variables under investigation are significantly different or related.

- **Comparison result of incidence of cancer in males and females.**

To investigate the difference in cancer incidence between males and females, a cross-tabulation analysis was performed. The contingency table revealed the distribution of cancer cases among genders is shown in Table. 2. The results of Proportion z tests: (-3.79, 0.0002).

The test findings produced a p-value of 0.0002, which is less than the 0.05 specified level of significance. Strong evidence to reject the null hypothesis is indicated by this.

Cancer	N	Y
Year		
Pandemic	64	120
Prepandemic	76	190

Table 1: Result of Cancer Cases in Pandemic and Prepandemic Periods

Gender	F	M
Cancer		
N	52	88
Y	63	247

Table 2: Distribution result of cancer cases by gender.

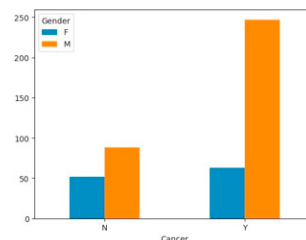


Fig. 13: Graphical result of cancer cases by gender.

- **Relationship between death and Hepatocellular Carcinoma\_tumor/node/metastas (HCC\_TNM\_Stages)**

To explore the relationship between death and HCC.TNM\_Stages, a cross-tabulation analysis was performed. The contingency table presented the distribution of death cases across different stages is displayed in Table. 3. To determine if the observed correlation was statistically significant, the chi-square test of independence was used. The test findings showed a significant association (p-value 0.05) between the HCC.TNM\_Stages and death, which provided evidence to reject the null hypothesis. Percentage calculations were performed to assess

Table 3: Distribution result of death by HCC TNM Stages.

HCC TNM Stages	Alive_Death	Alive Dead
I	126	140
II	32	26
IIIA+IIIB	20	66
IV	8	32

Table 4: Result of Cirrhosis by Year.

Year	Cirrhosis	N	Y
Pandemic	40	144	
Prepandemic	56	210	

Table 5: Result of Alive or Dead by Year

Year	Alive_Death	N	Y
Pandemic	84	100	
Prepandemic	102	164	

the proportion of deaths in each stage. The percentage of death in stage I is 0.32, stage II is 0.65, stage III is



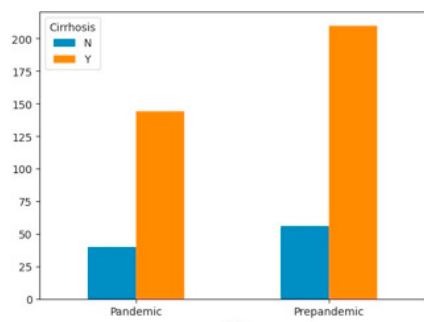


Fig. 14: Distribution of Cirrhosis by Year.

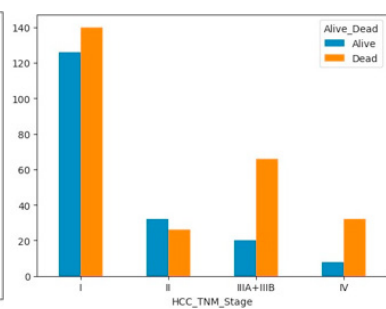


Fig. 15: Distribution of death by HCC\_TNM\_Stages.

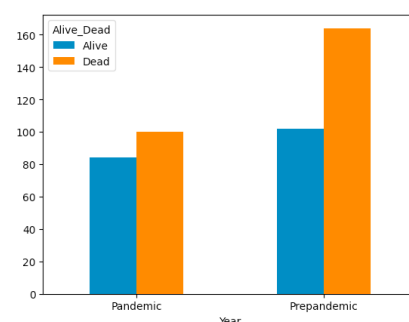


Fig. 16: Distribution of Alive or Dead by Year.

0.71 and stage IV is 0.8. The bar chart visualization in Fig. 15 effectively illustrates the distribution of deaths across different stages of HCC\_TNM. It is evident that the proportion of deaths increases as the HCC\_TNM stage advances.

● **Incidence of cancer result between males and females based on Cirrhosis**

To investigate the impact of the pandemic on cirrhosis, a cross-tabulation analysis was conducted. The contingency table displayed the distribution of cirrhosis cases between the pre-pandemic and pandemic periods as shown in Table. 4 and Fig.14. Proportions were calculated to compare the occurrence of cirrhosis in each period. The proportions Z-test was then performed to assess the statistical significance of any observed differences. The results of Proportion z tests: (0.31, 0.75).

The test results indicated that the p-value (0.75) is greater than the predetermined significance level of 0.05. Therefore, it cannot be concluded that there is a notable variation in the incidence of cirrhosis between the pandemic and pre-pandemic periods because there is insufficient information to support the alternative theory.

● **Results of a comparison of the death rates during the pre- and post-pandemic periods**

A cross-tabulation study was done to look at how death rates changed between the pre-pandemic and pandemic periods. The contingency table displayed the distribution of deceased and alive individuals for each period as shown in Table. 5 and Fig. 16. Proportions were calculated to determine the death rates in each period. The proportions Z-test was then performed to evaluate the statistical significance of any observed differences. The results of Proportion z tests: (-1.55, 0.12).

According to the calculations, there were 54.35% more deaths during the epidemic than there were before it, or 61.65%. The results of the proportions Z-test and the bar chart in Fig. 16 shows a p-value (0.121) greater than the predetermined significance level of 0.05. As a result, it is impossible to draw the conclusion that the death rates during the pre-pandemic and pandemic periods differ significantly from one another because there is little information to support the alternative hypothesis.

3.5. Results of model performance

The results of the machine learning models trained using three different feature selection techniques: SelectKBest, SVD (Singular Value Decomposition), and PCM (Pearson Correlation Matrix).

● **Performance result comparison using SelectKBest:**

The models evaluated using Accuracy, Precision, Recall, F1 Score to assess their effectiveness in predicting the target variable as shown in Table. 6 and Fig. 17. Among the models trained with the SelectKBest feature selection technique, the XGBoost model demonstrated the highest performance across multiple metrics.

● **Performance result comparison using SVD:**

The Table 7 and Fig.18 presents the performance metrics of the machine learning models trained with the SVD feature selection technique. Among the models trained with the SVD feature selection technique, the Decision Trees demonstrated the highest performance across multiple metrics.



Model	Accuracy	Precision	Recall	F1 Score
Decision Tree	0.91	0.91	0.91	0.91
KNN	0.67	0.68	0.67	0.67
CatBoost	0.77	0.80	0.81	0.81
LightGBM	0.87	0.88	0.87	0.87
XGBoost	<b>0.92</b>	<b>0.94</b>	<b>0.92</b>	<b>0.93</b>

Table 6: Comparison for model performance using SelectKBest

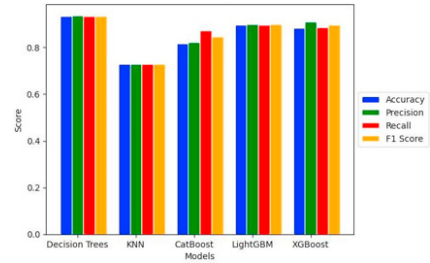


Fig. 17: Model Accuracy for SelectKBest

Model	Accuracy	Precision	Recall	F1 Score
Decision Tree	<b>0.93</b>	<b>0.94</b>	<b>0.93</b>	<b>0.94</b>
KNN	0.73	0.73	0.73	0.73
CatBoost	0.82	0.82	0.87	0.85
LightGBM	0.90	0.90	0.90	0.90
XGBoost	0.88	0.91	0.89	0.90

Table 7: Comparison for model performance using SVD

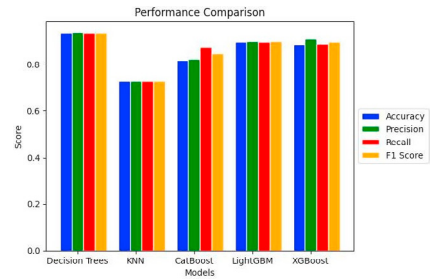


Fig. 18: Model Accuracy for SVD

• **Performance result comparison using PCA:**

The Table 8 below presents the performance metrics of the machine learning models trained with the PCA feature selection technique. Among the models trained with the PCA feature selection technique, the XGBoost model demonstrated the highest performance across multiple metrics.

Model	Accuracy	Precision	Recall	F1 Score
Decision Tree	0.88	0.88	0.88	0.883
KNN	0.75	0.75	0.75	0.75
CatBoost	0.85	0.84	0.91	0.87
LightGBM	0.88	0.88	0.88	0.88
XGBoost	<b>0.90</b>	<b>0.89</b>	<b>0.94</b>	<b>0.91</b>

Table 8: Comparison for model performance using PCA

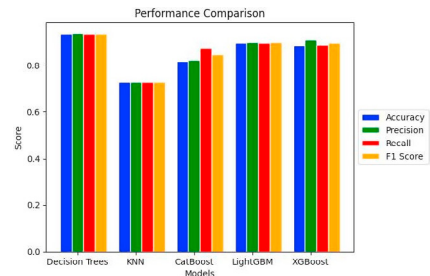


Fig. 19: Model Accuracy for PCA

From the bar chart shown in Fig. 19, it is evident that the XGBoost model outperforms the other models in terms of accuracy, reinforcing its status as the best performing model for the PCA feature selection technique.

Based on the results of the analysis using different feature selection techniques, the following conclusions are obtained:

- Using SelectKBest feature selection, the best performing model was Xgboost with an accuracy of 0.92. This indicates that the selected features using SelectKBest significantly improved the predictive performance of the Xgboost model.
- Utilizing SVD feature selection, the best performing model was Decision Trees with an accuracy of 0.93. This suggests that the reduced-dimensional features obtained through SVD were effective in capturing the underlying patterns and improving the accuracy of the Decision Trees model.

- The model with the highest accuracy using PCA feature selection was Xgboost, which had a 0.90 accuracy. This implies that the principal components derived from PCA contributed to enhancing the predictive capability of the Xgboost model.

#### 4. Conclusion

The COVID-19 pandemic has posed significant issues for the world's healthcare system, devoting crucial resources and attention away from chronic diseases like liver cancer. The proposed research work presents an effective research methodology by utilising the XGBoost classifier to clarify the effect of COVID-19 on liver cancer. The increase in symptomatic tumours during the pandemic indicates a higher likelihood that tumours may exhibit observable symptoms both during the pandemic and in the lead-up to it. The proposed work provided an accuracy of 91.97 % With SelectKBest feature selection on Xgboost classifier. The work can be improved with other feature selection methods and deep learning techniques.

#### References

- [1] Jonathan Montomoli, Luca Romeo, Sara Moccia, Michele Bernardini, Lucia Migliorelli, Daniele Berardini, Abele Donati, Andrea Carsetti, Maria Grazia Bocci, Pedro David Wendel Garcia, et al. Machine learning using the extreme gradient boosting (xgboost) algorithm predicts 5-day delta of sofa score at icu admission in covid-19 patients. *Journal of Intensive Medicine*, 1(02):110–116, 2021.
- [2] Marcos Antonio Alves, Giulia Zanon Castro, Bruno Alberto Soares Oliveira, Leonardo Augusto Ferreira, Jaime Arturo Ramírez, Rodrigo Silva, and Frederico Gadelha Guimarães. Explaining machine learning based diagnosis of covid-19 from routine blood tests with decision trees and criteria graphs. *Computers in Biology and Medicine*, 132:104335, 2021.
- [3] S Kishore, Rekha R Nair, Vishal Mehra, and Tina Babu. A generalized framework for brain tumor and pneumonia detection using streamlite application. In *2023 4th International Conference for Emerging Technology (INCET)*, pages 1–6. IEEE, 2023.
- [4] Xianbin Song, Jiangang Zhu, Xiaoli Tan, Wenlong Yu, Qianqian Wang, Dongfeng Shen, and Wenyu Chen. Xgboost-based feature learning method for mining covid-19 novel diagnostic markers. *Frontiers in Public Health*, 10:926069, 2022.
- [5] Tina Babu, Deepa Gupta, Tripty Singh, Shahin Hameed, Mohammed Zakariah, and Yousef Alotaibi. Robust magnification independent colon biopsy grading system over multiple data sources. *Computers, Materials Continua*, 69:99–128, 01 2021.
- [6] Tina Babu, Tripty Singh, Deepa Gupta, and Shahin Hameed. Colon cancer prediction on histological images using deep learning features and bayesian optimized svm. *Journal of Intelligent Fuzzy Systems*, 41:1–12, 03 2021.
- [7] Rekha R. Nair. Multi-sensor medical image fusion using pyramid-based dwt: a multi-resolution approach. *IET Image Processing*, 13:1447–1459(12), July 2019.
- [8] Rekha R. Nair and Tripty Singh. An optimal registration on shearlet domain with novel weighted energy fusion for multi-modal medical images. *Optik*, 225:165742, 2021.
- [9] Rekha R Nair, Tina Babu, and Tripty Singh. Multiresolution approach on medical image fusion by modified local energy. *Signal, Image and Video Processing*, pages 1–8, 2023.
- [10] Amrita Tripathi, Abhinandan Basavapattana, Rekha R Nair, and Tripty Singh. Visualization of covid bimodal scan using dnn. In *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 01–07. IEEE, 2021.
- [11] R Gayathri, Peeta Basa Pati, Tripty Singh, and Rekha R Nair. A framework for the prediction of diabettes mellitus using hyper-parameter tuned xgboost classifier. In *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–5. IEEE, 2022.
- [12] Anuj Kumar Dixit, Rekha R Nair, and Tina Babu. Analysis and classification of restaurants based on rating with xgboost model. In *2022 3rd International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, pages 1–6, 2022.
- [13] Daniel Asante Otchere, Tarek Omar Arbi Ganat, Jude Oghenerurie Ojoro, Bennet Nii Tackie-Otoo, and Mohamed Yassir Taki. Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. *Journal of Petroleum Science and Engineering*, 208:109244, 2022.
- [14] Mihir Gandhi, Wen-Huan Ling, Chien-Hung Chen, Joon Hyeok Lee, Masatoshi Kudo, Rawisak Chanwat, Simone I Strasser, Zhu Xu, Soh-Han Lai, and Pierce Kah-Hoe Chow. Impact of covid-19 on hepatocellular carcinoma management: a multicountry and region study. *Journal of hepatocellular carcinoma*, pages 1159–1167, 2021.
- [15] Davide Giuseppe Ribaldone, Gian Paolo Caviglia, Silvia Gaia, Emanuela Rolle, Alessandra Rizzo, Daniela Campion, Paola Rita Brunocilla, Giorgio Maria Saracco, and Patrizia Carucci. Effect of covid-19 pandemic on hepatocellular carcinoma diagnosis: results from a tertiary care center in north-west italy. *Current Oncology*, 29(3):1422–1429, 2022.
- [16] Natalia Salinas Parra, Heather M Ross, Adnan Khan, Marisa Wu, Risa Goldberg, Lokesh Shah, Sarah Mukhtar, Jacob Beiriger, Alexis Gerber, and Dina Halegoua-DeMarzio. Advancements in the diagnosis of hepatocellular carcinoma. *International Journal of Translational Medicine*, 3(1):51–65, 2023.
- [17] Fedesoriano. Covid-19 effect on liver cancer prediction dataset, 2023. <https://www.kaggle.com/datasets/fedesoriano/covid19-effect-on-liver-cancer-prediction-dataset>, Last accessed on 2023-08-20.