International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 2, February, 2026, pp. 499-510
Journal homepage: http://iieta.org/journals/isi

A Hybrid CNN–LSTM Framework for Diabetes Risk Prediction in COVID-19 Patients

Raisa M. Mulla[1], K. R. Desai[2*]

[1] Department of Electronics and Telecommunication Engineering, Shivaji University, Kolhapur 416004, India
[2] Department of Electronics and Telecommunication Engineering, Bharati Vidyapeth's College of Engineering, Kolhapur 416004, India

Corresponding Author Email: kamalakar.desai@bharatividyapeeth.edu

https://doi.org/10.18280/isi.310218

Received: 11 November 2025
Revised: 15 January 2026
Accepted: 21 February 2026
Available online: 28 February 2026

Keywords:
COVID-19, diabetes risk prediction, hybrid deep learning, convolutional neural network, long short-term memory, Synthetic Minority Over-sampling Technique, feature selection, risk stratification

ABSTRACT

Accurate prediction of diabetes risk in patients with COVID-19 is essential for early intervention and improved clinical outcomes. However, the complex interactions between clinical and demographic factors pose significant challenges for traditional predictive models. This study proposes a hybrid deep learning framework that integrates convolutional neural networks (CNN) and long short-term memory (LSTM) to capture both spatial and sequential patterns in patient data. To enhance model performance, Synthetic Minority Over-sampling Technique (SMOTE) is employed to address class imbalance, and Random Forest–based feature selection is applied to identify the most relevant predictors. The proposed model is further optimized through hyperparameter tuning to improve generalization capability. Experimental results demonstrate that the hybrid CNN–LSTM model achieves superior performance, with a macro recall and F1-score of 0.89 on the balanced dataset, outperforming conventional machine learning and deep learning approaches. These findings indicate that the proposed framework provides an effective and reliable solution for diabetes risk prediction in COVID-19 patients, supporting clinical decision-making and risk stratification.

1. INTRODUCTION

The COVID-19 pandemic is a consequence of the novel coronavirus, SARS-CoV-2, which has a detrimental effect on communities and health systems worldwide. The virus first appeared at the end of 2019 and swiftly expanded over the world, killing a significant number of people and inflicting severe illness [1]. The patients who are exposed to risk of infection are those who are chronically sick like diabetes or a metabolic condition where they have high levels of blood glucose. Diabetes comes with impaired immunity, which exposes vulnerable persons to infectious diseases [2]. Since diabetes is a disorder that impacts about 422 million individuals worldwide, determining the interplay between COVID-19 and this prevalent disorder is extremely important.

Additional comorbidities, including obesity and cardiovascular diseases, as well as inadequate immune functions and chronic inflammation, can also increase this level of risk [3]. As such, predictive models are required to determine high-risk individuals and determine the risk of diabetes among COVID-19 patients to help verify that adverse outcomes are avoided.

Several approaches have been proposed to improve diabetic patient care during the COVID-19 epidemic. Diabetes risk classification and prediction using ML and DL techniques have shown a great deal of promise for improving patient outcomes. The most typical forms of risk assessment are based on the ML algorithms, i.e., on Logistic Regression (LR), Decision Trees (DT), and Support Vector Machines (SVM) that process clinical data to reveal the key predictors [4, 5]. By contrast, DL models, specifically CNNs and RNNs are more appropriate in working with complex data, such as medical images and patient records in sequence [6, 7]. These computational techniques show high potential to assist in managing diabetes risks among COVID-19 patients and allow them to be provided with a personalized treatment approach.

This study evaluates the risk of diabetes in COVID-19 patients using a predictive modeling technique. A wide range of ML algorithms are used to compute on a large data set of demographics, clinical characteristics, diagnostic outcomes, and comorbidities. The primary objective is to develop a useful model that may identify individuals who are more likely to experience unfavorable consequences. By identifying the factors linked to poor COVID-19 results in diabetics, this article should help a healthcare professional make a therapeutic choice and manage a patient.

Additionally, the Fine-Tuned Hybrid CNN-LSTM model, a novel hybrid deep learning model that combines the benefits of two deep learning models, will be presented in this work as a diabetes risk prediction. The recommended framework will be made up of four key steps namely, preprocessing data and features, model fusion, and classification. This paper's primary conclusions can be summed up as follows:

- To utilize the data balancing procedures and feature selection with the help of the Random Forest (RF) to enhance the quality of the dataset so that the model was effective at detecting the relevant features and could avoid biases in predictions.

- • The traditional ML and DL models are evaluated on the common classification tasks in terms of such metrics as accuracy, precision, recall, and F1-score on publicly available datasets.
- A hybrid CNN-LSTM model that combines CNN and LSTM was recently proposed with the goal of increasing the accuracy of diabetes risk prediction in COVID-19 patients. By incorporating CNNs to learn geographical information and LSTMs to learn sequential patterns, this system enhances the quality of submission prediction.
- The second step is to adjust the hybrid CNN-LSTM model's hyper-parameters so that it can identify the different factors that contribute to diabetes risk in COVID-19 patients and get the highest level of prediction power. This tuning process is more equipped against noise and maximizes generalization capabilities of the model when extended even to clinical practices.

The paper's outline is as follows: A brief synopsis of related works from the past is given in Section 2. Section 3 outlined the planned deployment of the proposed fine-tuned hybrid CNN-LSTM model, and Section 4 included a detailed description of the suggested method of forecasting patients' risk of developing diabetes. The experimental design and findings are described in depth in Section 4, the work is summarized and possible future directions are suggested in Section 5, and the conclusion is covered in Section 6.

2. LITERATURE REVIEW

Reviews of the pertinent data indicate that diabetes is linked to a higher risk of death and severe COVID-19 illness. Hyperglycemia, comorbidities, and therapeutic options are also emphasized. Results thus support the importance of tight control of glycaemia and individualization of treatment if those patients' results are to be further improved.

In their hospital-based study of 218 patients with type 2 diabetes mellitus (T2DM), Elsayed et al. [8] used multivariate logistic regression to identify male gender, higher BMI, elevated CRP levels, advanced age, and poor glycemic control as significant predictors of severe COVID-19 outcomes. The ROC-AUC for this analysis was 0.85. Similarly, studies by Lipski et al. [9] and Kabootari et al. [10] employed machine learning approaches and reported advanced age, male sex, elevated inflammatory markers, poor glycemic control, and cardiovascular comorbidities as major predictors of mortality. On the other hand, Khadem et al. [11] used gradient boosting models and neural networks to emphasize the significance of blood glucose levels and diabetes status by including SHAP-based interpret ability. In addition, McGurnaghan et al. [12] found that severe COVID-19 outcomes were more common in diabetic individuals, with an adjusted odds ratio of 1.395. This elevated risk was further supported by the presence of diabetes-related comorbidities and inadequate glycemic control.

These findings were supported using comprehensive reviews and cohort studies by Schlesinger et al. [13] and Kristan et al. [14], which indicated that a greater number of patients were found to be at high risk of ICU hospitalization and mortality due to high HbA1c levels, long-term diabetes, age, obesity, high blood pressure, and elevated CRP with reported AUC-ROC values of up to 0.86. Moreover, Ramesh et al. [15] came up with a cost-effective model that comprises of biomarkers using NLR and CRP with an AUC of 0.82, which illustrates the usefulness of simple clinical markers in the early detection of risks, especially where the resources are limited.

The models developed by Orioli et al. [16] and Ortega et al. [17] utilize Logistic Regression to effectively categorize risks for COVID-19 hospitalized patients with diabetes. The authors emphasize that the most significant predictors of mortality include advanced age, male sex, obesity, high body mass index, inadequate glycemic control, and cardiovascular history. Mamtani et al. [18] demonstrated that admission hyperglycemia significantly increased the risk of hospital mortality even in non-diabetic patients (hazard ratio = 1.78), demonstrating the need to monitor glucose levels in all hospitalized patients, even though their study is not limited to diabetic cohorts. Another study by Collard et al. [19] indicated a strong correlation between cardiovascular disease, obesity, and hypertension with intensive care unit admission and mortality. Pazoki et al. [20] found that COVID-19 patients with diabetes who had poor glycemic control and high inflammatory markers were more likely to die.

Clinically interpretable risk scoring systems were also proposed in a number of studies. Lasbleiz et al. [21] have proposed the DIAB score whose sensitivity (77.7) and specificity (89.2) are high in terms of predicting hospitalization, and Bello-Chavolla et al. [22] proposed the CORONADO score with an AUC of 0.81 and a large cohort of Mexicans. These findings, taken together, indicate that risk assessment and clinical management approaches aimed at minimizing adverse outcomes among COVID-19 patients with diabetes should be conducted early and individually.

The study by Khodabakhsh et al. [23] developed machine learning models to predict in-hospital mortality among COVID-19 patients with diabetes using clinical data from 2,824 patients. The results showed that the decision tree model achieved high predictive performance (around 87% accuracy), identifying key risk factors such as age, ICU admission, and CPR status. Another similar study uses; age, male gender, lymphopenia, high levels of lactate dehydrogenase, and comorbidities hypertension, cardiovascular disease were found by Yu et al. [24] to be dominant predictors of COVID-19 mortality, which supports the initial classification of clinical risks. Putting emphasis on metabolic factors, Carrasco-Sánchez et al. [25] found that admission hyperglycemia was a significant predictor of mortality (OR = 1.75) when the blood glucose level rose by 50 mg/dL irrespective of diabetic status. In the same vein, Zhang et al. [26] performed an analysis on 258 diabetic patients and identified age, male gender, high body mass index, high levels of CRP and IL-6 as good predictors of a negative outcome with an AUC of 0.82. An extended evaluation by Apicella et al. [27] also elaborated on the worsening of COVID-19 by chronic inflammation and hyperglycemia, and the importance of closer glycemic regulation.

A number of studies suggested risk forecasting instruments of diabetic citizens. Sourij et al. [28] have created a straightforward mortality risk score based on 1,308 patients with diabetes or prediabetes and four primary variables, which are high blood glucose, advanced age, male gender, and hypertension as the primary factors, with favorable results (AUC-ROC = 0.88). Similarly, Liu et al. [29] found a significantly greater odds of ICU admission (OR = 2.5), mortality (OR = 2.8) among diabetic patients with an AUC of 0.85. Similar conclusions about the importance of hyperglycemia, lack of glycemic control, and inflammatory markers (CRP, IL-6) were made due to narrative and cohort studies by Singh and Khunti [30], Zhang et al. [31], and Guo et al. [32], and supported the necessity of aggressive metabolic management in COVID-19 patients with diabetes.

There is a notable discovery in the scientific literature about the effect of diabetes on the COVID-19 results in terms of mortality and disease severity. In sum, all these studies are important in emphasizing the need for proactive management of diabetes and other associated comorbid conditions in lessening the unfavorable outcomes seen in COVID-19.

3. METHODOLOGY

The methodology of diabetes risk prediction in COVID-19 patients begins with loading a publicly available dataset that includes the clinical features of a patient like age, sex, glucose level, status of diabetes, and so on. Imputation of missing values, normalization of numerical characteristics, and encoding of categorical variables are the earliest steps in preprocessing. As a first step in data preparation, we fill in missing values, standardize numerical characteristics, and use label encoding to encode categorical variables.

The outliers are either transformed or removed in order to minimize bias. In order to handle class imbalance, SMOTE is utilized to balance diabetic and non-diabetic classes. Feature selection is done using RF-based methods, which have the most influential features to reduce dimensionality in such a way that model efficiency is improved. The suggested hybrid DL model takes the best features of both LSTM and CNN networks and uses them to generate a spatial and sequential pattern in the input data. This hybrid architecture allows the model to potentially uncover complex relationships and time-dependent interactions among patient characteristics.

Methods like Grid Search are used for hyper-parameter tuning, which optimizes model architecture, batch sizes, and learning rates. Finally, a number of performance indicators show that the proposed model is helpful in predicting the risk of diabetes in COVID-19 patients. The proposed system architecture is shown in Figure 1, and the following section delves into the details of each stage.

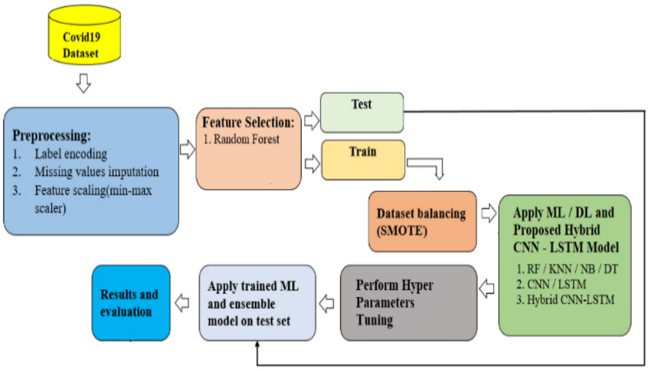


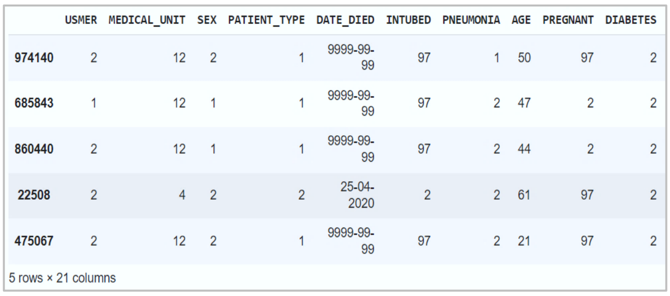
Figure 1. Proposed system architecture

3.1 Dataset description

The Mexican government made available on Kaggle the dataset that was used in this investigation [33]. It has anonymized-clinical and demographic data of COVID-19 tested persons. The data consists of 1,048,576 records about patients characterized by 21 variables, including symptoms, comorbidity, treatment, and outcomes. The binary variables are coded (1 Yes and 2 No) and 97 and 99 represent missing or unavailable information, Table 1 shows the dataset feature and its description.

Table 1. Dataset feature and description

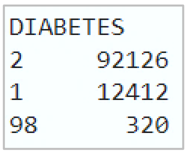
Feature Name	Description
Sex	Gender of the patient (1 = Female, 2 = Male)
Age	Age of the patient
Classification	COVID-19 test result (1–3: COVID-19 positive; ≥4: negative or inconclusive)
Patient Type	Type of care received (1 = Home care, 2 = Hospitalized)
Pneumonia	Presence of pneumonia
Pregnancy	Pregnancy status
Diabetes	Presence of diabetes
COPD	Presence of chronic obstructive pulmonary disease
Asthma	Presence of asthma
Inmsupr	Immunosuppression status
Hypertension	Presence of hypertension
Cardiovascular	Presence of cardiovascular disease
Renal Chronic	Presence of chronic kidney disease
Other Disease	Presence of other diseases
Obesity	Obesity status
Tobacco	Tobacco use
USMR	Level of medical unit providing care
Medical Unit	Type of healthcare institution
Intubed	Whether the patient was intubated
ICU	Admission to intensive care unit
Date Died	Date of death (9999-99-99 if alive)



	USMER	MEDICAL_UNIT	SEX	PATIENT_TYPE	DATE_DIED	INTUBED	PNEUMONIA	AGE	PREGNANT	DIABETES
974140	2	12	2	1	9999-99-99	97	1	50	97	2
685843	1	12	1	1	9999-99-99	97	2	47	2	2
860440	2	12	1	1	9999-99-99	97	2	44	2	2
22508	2	4	2	2	25-04-2020	2	2	61	97	2
475067	2	12	2	1	9999-99-99	97	2	21	97	2

5 rows × 21 columns

Figure 2. Dataset sample



```
DIABETES
2      92126
1      12412
98       320
```

Figure 3. Label counts (without balancing)

The first thing that needs to be done is to load the dataset that contains information about diabetes, which has normally been saved in CSV format. Among the many characteristics that may be included in this dataset are the patient's age, glucose levels, body mass index, blood pressure, and outcomes There is a possibility that the dataset includes information about COVID-19 infection, which is essential for the investigation being conducted. Reading the CSV file into a data structure, like a pandas Dataframe in Python, simplifies data manipulation and analysis during the dataset loading process. Figure 2 illustrates the dataset sample while Figure 3 illustrates the diabetes data count with 1 indicted diabetic patients, 2 is non diabetic and 98 related to others or having not identified.

3.2 Dataset pre-processing

Important pre-processing operations are performed to prepare the raw data for analysis. At this point, you should deal with missing values in one of two ways: either fill them in using measurement statistics like the mean or median, or remove rows and columns with too much missing data. It is necessary to identify any outliers that have the potential to bias the study and then either convert them or delete them. Data normalization or standardization is followed to make sure that the model does not get disproportionately influenced by characteristics having different scales. This may require encoding categorical data into numerical values, and feature scaling is done so that all characteristics fall into a scale comparable across one another. Dataset pre-processing comprises the following steps:

- Drop Rows with Label “98”.
- Handle missing values, and noise in the dataset.
- Normalize or scale numerical features.
- Min-Max Scalar.
- Encode categorical variables and labels.

3.2.1 Drop rows with label “98”

This is generally done when the label "98" in data preprocessing indicates an invalid or irrelevant category. In such cases, the removal of such rows helps the model not learn a pattern from erroneous or extraneous data points, which otherwise may impair the model's performance and predictions.

3.2.2 Handle missing values, and noise

High-quality data is very important for building reliable models; having missing values or outliers, or noisy data points will lead to biased results. When a value is missing, an imputation method that takes the feature's mean, median, or mode into account fills in the blanks. To solve the issue of missing values, mean values are utilized in this case. Likewise, noise in data-smoothening unexpected or meaningless variations-is achieved by filtering techniques. Removing these discrepancies makes the dataset more accurate reflection of real-world situations and reduces the chances of overfitting.

3.2.3 Normalize or scale numerical features

In cases when variables are on different scales or range, then standardization or scaling of numerical features must be done. This is because features with higher values may dominate others at the time of model training. One common technique for normalization is called Min-Max Scaling, in which features are transformed to fall within a scale in between 0 and 1. To accomplish this, we use the formula:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

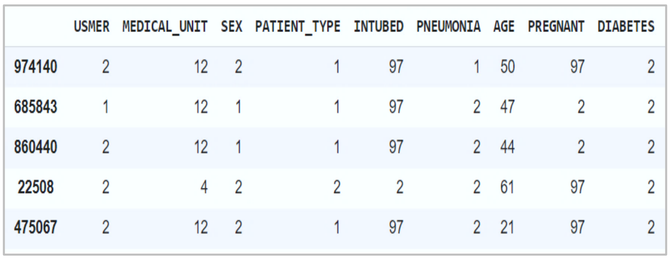
where, X_{min} and X_{max} are the feature's minimum and maximum values, respectively, and X is the initial value. Learning algorithms are made more stable and converge faster with Min-Max scaling, which makes sure that all numerical features are on the same scale.

3.2.4 Encode categorical variables and labels

Since ML models can only take numerical input, it is necessary to convert categorical variables to numbers. Encoding describes this procedure. For ordinal variables, which have an inherit order, we proceed with Ordinal Encoding - simply replacing each level of a categorical variable by a unique integer. For nominal (non-ordered) categories we should not rely on different weights and in such situations, it is appropriate to use one-hot encoding, where binary columns are created for each category to let the model know that it shouldn't assume any order of categories. Labels in the target variable are also encoded using Label Encoding if they represent discrete categories.

The correct interpretation of the categorical data by ML algorithms is dependent on this conversion, which in turn allows the model to successfully detect patterns and relationships. The Pre-processed dataset samples after encoding categorical variable is shown in Figure 4.

3.3 Applying feature selection techniques



	USMER	MEDICAL_UNIT	SEX	PATIENT_TYPE	INTUBED	PNEUMONIA	AGE	PREGNANT	DIABETES
974140	2	12	2	1	97	1	50	97	2
685843	1	12	1	1	97	2	47	2	2
860440	2	12	1	1	97	2	44	2	2
22508	2	4	2	2	2	2	61	97	2
475067	2	12	2	1	97	2	21	97	2

Figure 4. Pre-process dataset samples

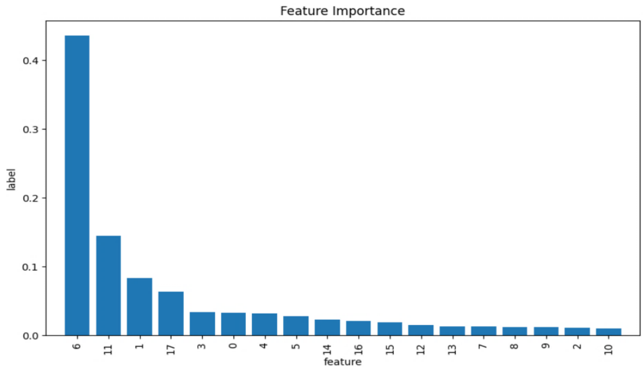


Figure 5. Feature importance score

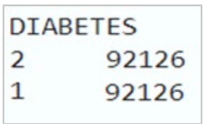
Table 2. Selected features list

Features
USMER
MEDICAL_UNIT
PATIENT_TYPE
INTUBED
PNEUMONIA
AGE
HYPERTENSION
OBESITY
TOBACCO

Identification and selection of the most pertinent features from the dataset that make a significant contribution to the prediction of the diabetic risk factor are considered to be part of the feature selection process. This is accomplished by the utilization RF feature importance model. The dimensionality of the dataset reduced through the use of feature selection, which also helps eliminate redundant or irrelevant data. As a result, the model's performance can be enhanced and calculation times can be reduced. Figure 5 shows the features ranking (graph) from high to low. We have selected features whose threshold is greater than 0.02 as shown in Table 2.

3.4 Applying data sampling technique

SMOTE is one of the data sampling approaches that is utilized here as the dataset is imbalanced (for example, there are a significantly greater number of patients who do not have diabetes than diabetic patients) (as shown in Figure 3). By taking this step, we ensure that the model does not develop a bias towards the class that constitutes the majority. It is possible to achieve a balanced dataset with the help of proper sampling, which ultimately results in more reliable model predictions, we used SMOTE for balancing the samples, the balance sample label count is shown in Figure 6.



```
DIABETES
2      92126
1      92126
```

Figure 6. Label counts (with smote balancing)

3.4.1 Train – test split of dataset

Data cleansing and feature selection are followed by the creation of a training set and a testing set from the dataset. Two common ratios for the sets are 80:20 and 70:30. We opted on a 70:30 split for our studies. By dividing the data in this way, we may access the model on unobserved data and have a good idea of how it will perform in the actual world. To eliminate any possibility of bias and to guarantee that both sets are representative of the entire dataset, the split must to be carried out in a random fashion.

3.4.2 Machine learning models for classification

ML models [33-35] play a critical role in predicting diabetes risk among COVID-19 patients by leveraging clinical and demographic data. These models examine patient data for trends that suggest an elevated risk of diabetes, especially in the presence of COVID-19. Based on the labeled data used to determine the risk class—which includes blood glucose readings, age, BMI, comorbidities, inflammatory markers—the algorithms can now forecast and categorize individual patients into diabetic and non-diabetic groups. Various categorization methods allow these models to distinguish between high-risk and low-risk patients. The goal is to generate early warnings in order to implement targeted interventions and personalized therapies. Here the diabetic risk factor is predicted using a number of different ML models, including the following:

Decision Tree:

Decision trees organize datasets hierarchically by routinely dividing them into subsets according to feature values that maximize class separation. Each division is identified using criteria such as Gini Impurity or Information Gain. A node's Gini Impurity is determined as follows:

$$Gini = 1 - \sum_{i=1}^{C} P_i^2$$

In this case, the graph shows the correlation between the two variables, where pi is the likelihood that class I will be present at the node and C is the total count of classes at the node. The tree is constructed until a stopping requirement, like the minimum number of samples per node or the maximum depth, is satisfied. Deep trees are useful, but they can overfit, capturing noise and decreasing generalization.

Random Forest:

In order to generate numerous decision trees, RF is a combination method that makes use of bootstrapped samples and randomly selected subsets of features for each split. For classification tasks, the model uses majority voting, and for regression tasks, it uses averaging to aggregate the results from individual trees. The final output Y for classification is determined as:

$$Y = Mode(T_1, T_2, \ldots T_n)$$

where, T_n is the prediction from the n th decision tree in the ensemble. This reduces overfitting by averaging out errors across trees, making the model more robust and accurate.

Logistic Regression:

Logistic regression (LR) is a statistical method that models the probability of a binary outcome using the logistic (sigmoid) function. The logistic function is used to do this. Given the input features X and coefficients γ, the following expression can be used to express the probability P of the positive class:

$$P(Y = 1|X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n)}}$$

where, the feature coefficients are denoted by β_i and the intercept by β_0. A threshold (often 0.5) is applied to determine the anticipated class. For linearly separable data, the model is effective and interpretable since it employs Maximum Likelihood Estimation (MLE) to determine the optimal parameters.

AdaBoost (Adaptive Boosting):

The iterative ensemble AdaBoost method can be used to produce a powerful classifier. In this method, many weak learners are combined including shallow decision trees to achieve the classifier. The model assigns weights to samples which have been misclassified during each iteration and this means the successive learner will focus on the difficult to classify cases. The prediction given by the final model is the weighted sum of the output of every weak learner:

$$F(x) = sign\left(\sum_{m=1}^{M} a_m h_m(x)\right)$$

This is where M is the total number of iterations, a_m is the weight provided to the weak learner, and h_m (x) is the m^{th} weak classifier. The effectiveness of increasingly precise classifiers in reaching the ultimate decision boundary is shown by computing the weights α_m according to the accuracy of each learner.

3.5 Deep learning models for classification

A subfield of machine learning known as "deep learning," these algorithms automatically learn intricate data patterns by means of artificial neural networks. These models are built to be able to classify incoming data into preset categories based on features learned in order to achieve classification tasks. Since it can model non-linear and detail-rich relationships and high-dimensional data, deep learning would best fit the management of complicated datasets. Deep learning excels in this area. These types of models are used in a subfield of machine learning known as "DL," and they automatically discover complicated patterns in data by means of artificial neural networks.

3.5.1 Convolutional neural network

CNNs [36, 37] differ from conventional neural networks in that they employ convolutional layers as filters on the input data to identify hierarchical traits and local patterns. CNNs prove to be very useful in image and sequence classification processes since they can follow spatial data and features of images and find intricate patterns. These are filters that slide over the input, and they perform a mathematical operation popularly referred to as convolution, which assists in identifying features of a picture, including edges and tendencies of sequential data. The adaptation for use on structured or sequential data is done by applying a convolution operation along a single dimension. This makes it fit for such tasks as time-series analysis or even classification of genomic sequences. Utilizing a network of interconnected processes, CNN automatically extracts characteristics from unprocessed input data for application in result regression or object classification. Included in CNNs are pooling layers, completely linked or thick layers, and convolutional layers. To define the output feature y for 1D Convolution:

$$y(t) = \sum_{i=0}^{k-1} x(t+i) \cdot w(i) + b$$

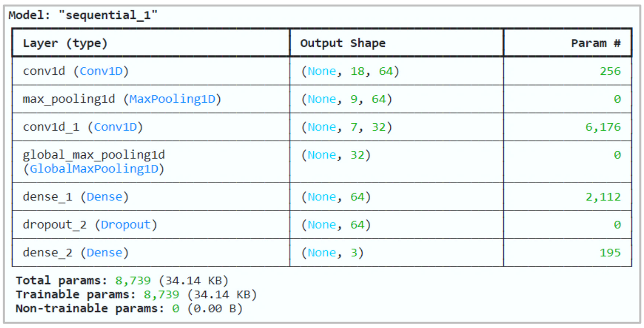
where:

x(t): Input sequence at position t
w(i): Kernel weight at position i
k: Size of the kernel
b: Bias term
y(t): Output at position t

This method is repeated over the whole input sequence, yielding a new feature map that highlights helpful patterns in the data. To lower the dimensionality, pooling is typically used after convolutional layers, and fully connected layers are used for final classification. Figure 7 shows the CNN architecture used to classify patients and forecast their risk of developing diabetes.

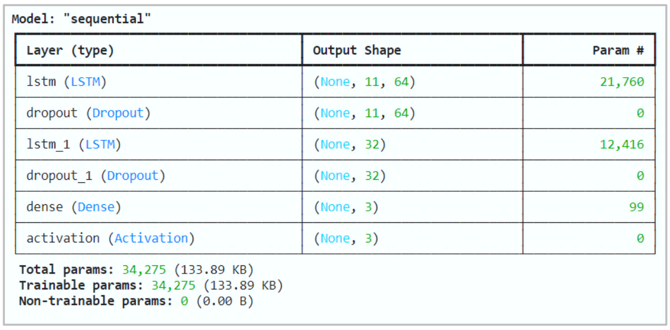
3.5.2 Long short-term memory

RNNs come in the form of LSTMs [38], which are specifically made to efficiently detect temporal dependencies in a sequence. These were created to solve the issue of disappearing gradients that conventional RNNs have. Input, forget, and output gates are essential components of the LSTM cell architecture, which allows the cell to efficiently control the flow of information. LSTMs are optimal for language modeling, voice recognition, and time series prediction due to their architecture and, among other things, their long-term memory capacity. LSTMs have the ability to learn and reflect intricate dependencies between sequential pattern pieces, unlike ordinary RNNs that struggle to replicate such dependencies due to their limited memory. Figure 8 shows the LSTM structure.



```
Model: "sequential_1"
Layer (type)                          Output Shape            Param #
conv1d (Conv1D)                       (None, 18, 64)              256
max_pooling1d (MaxPooling1D)          (None, 9, 64)                 0
conv1d_1 (Conv1D)                     (None, 7, 32)             6,176
global_max_pooling1d                  (None, 32)                    0
(GlobalMaxPooling1D)
dense_1 (Dense)                       (None, 64)                2,112
dropout_2 (Dropout)                   (None, 64)                    0
dense_2 (Dense)                       (None, 3)                   195
 Total params: 8,739 (34.14 KB)
 Trainable params: 8,739 (34.14 KB)
 Non-trainable params: 0 (0.00 B)
```

Figure 7. CNN architecture
Note: CNN - Convolutional Neural Network.



```
Model: "sequential"
Layer (type)                          Output Shape            Param #
lstm (LSTM)                           (None, 11, 64)           21,760
dropout (Dropout)                     (None, 11, 64)                0
lstm_1 (LSTM)                         (None, 32)               12,416
dropout_1 (Dropout)                   (None, 32)                    0
dense (Dense)                         (None, 3)                    99
activation (Activation)               (None, 3)                     0
 Total params: 34,275 (133.89 KB)
 Trainable params: 34,275 (133.89 KB)
 Non-trainable params: 0 (0.00 B)
```

Figure 8. LSTM architecture
Note: LSTM - Long Short Term Memory.

3.5.3 Proposed hybrid CNN+ LSTM model

Incorporating the greatest aspects of both convolutional and recurrent neural networks, the hybrid CNN-LSTM model provides a powerful tool for assessing sequences across applications. These two powerful deep learning approaches allow our model to be particularly well-suited for dealing with time-series signals, text sequences and physiological data. The CNN part is a strong feature extractor that implicitly learns and stores spatial patterns of raw input sequences. The LSTM layers, which are good at capturing long-term patterns in sequential information and sequence dependence, are then fed with these extracted characteristics. The result is a flexible model that can perform both complex feature extraction and temporal sequence modeling. Here, A 1D-CNN is used to capture local patterns and feature interactions across clinical attributes. The extracted features are then passed to an LSTM layer, which models dependencies within the learned feature representations, the details of model is provided in Model Architecture Overview section.

3.6 Model architecture overview

3.6.1 1D convolutional neural network layer for feature extraction

The final set of features selected through feature importance analysis is used as the input to the 1D-CNN layer. Typically, one or more 1D convolutional layers are incorporated in the first layers of this model to preprocess data from the input sequences. These are essentially filters that slide over the input sequence and create a convolution responsible for recognizing local patterns and identifying high-level features. Typically, each convolutional layer is succeeded by an activation function (e.g., ReLU) to introduce non-linearity, followed by a pooling layer or subsampling technique (MaxPooling) to reduce spatial resolution. This architecture must be considered what CNN should play as the primary role, in that it just compresses the raw data into a feature set with more structured format by extracting some important patterns such as spikes, trends and anomalies.

3.6.2 Flattening layer

A flattening layer was employed subsequent to the 1D convolutional layers to diminish the dimensionality of the feature maps from 2D to 1D. By serving as an important intermediary between the architecture's convolutional and recurrent parts, this transformation prepares spatially-derived features in a pacy way for upcoming LSTM layers to consider temporally. The 'flatten' operation keeps anything we have learned thus far, but restructures the previously learnt features into a sequence that RNN layers can analyses - this in effect enables the model to process spatial feature extraction and temporal pattern recognition together.

3.6.3 LSTM layer for temporal dependencies

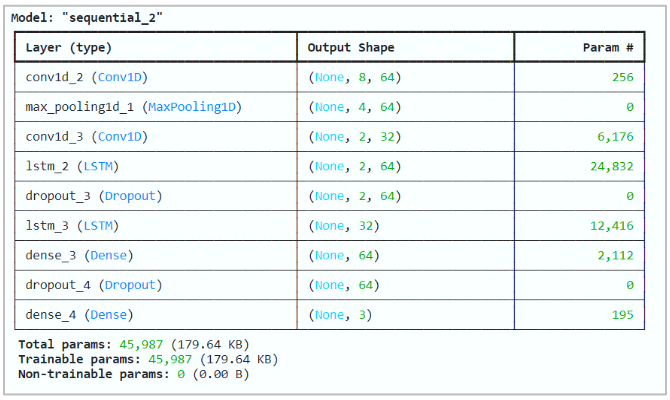
The flattened feature vector is subsequently input into the LSTM layer(s) to acquire temporal correlations among the sequence data. LSTMs regulate information dissemination through memory cells and gates (input, forget, and output gates). The network is able to remember meaningful patterns over time while ignoring less significant ones because of this. For instance, in diabetes risk prediction, the LSTM could capture real-world similitudes and deviations of health metrics in a span of several days or weeks, discovering patterns that are linked to worsening (or improving) conditions. Input, forget, and output gates, along with memory cells, govern the dissemination of data in LSTMs. This enables the network to acquire and preserve significant patterns while eliminating less critical ones.

3.6.4 Fully connected (dense) layer

The final output of the LSTM layer encodes the learned temporal patterns, and for classification, one or more dense layers are employed to interpret this output. The LSTM model decodes high order sequence information that these deep layers integrate and translate into class predictions. Each dense layer learns complex, non-linear data patterns by weighting its inputs and activates these weights (most often using ReLU for the hidden layers). The hybrid architecture's classification pipeline is completed by the last dense layer, which generates probabilistic outputs for target classes using an activation function such as softmax or sigmoid.

3.6.5 Output layer

The sigmoid function is employed for binary classification, whereas the softmax activation function is employed for multi-class classification. These two functions are utilized by the output layer. A network's outputs can be normalized into probability distributions across many classes using the softmax function, and for binary outcomes, independent probabilities can be generated using the sigmoid function. This probabilistic output enables precise class predictions, such as stratifying COVID-19 patients into high, medium, or low diabetes risk categories. The complete architecture of our hybrid CNN-LSTM model, illustrating these components and their interactions, is presented in Figure 9.



```
Model: "sequential_2"
┌──────────────────────────────────┬─────────────────────┬──────────┐
│ Layer (type)                     │ Output Shape        │  Param # │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ conv1d_2 (Conv1D)                │ (None, 8, 64)       │      256 │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ max_pooling1d_1 (MaxPooling1D)   │ (None, 4, 64)       │        0 │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ conv1d_3 (Conv1D)                │ (None, 2, 32)       │    6,176 │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ lstm_2 (LSTM)                    │ (None, 2, 64)       │   24,832 │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ dropout_3 (Dropout)              │ (None, 2, 64)       │        0 │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ lstm_3 (LSTM)                    │ (None, 32)          │   12,416 │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ dense_3 (Dense)                  │ (None, 64)          │    2,112 │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ dropout_4 (Dropout)              │ (None, 64)          │        0 │
├──────────────────────────────────┼─────────────────────┼──────────┤
│ dense_4 (Dense)                  │ (None, 3)           │      195 │
└──────────────────────────────────┴─────────────────────┴──────────┘
 Total params: 45,987 (179.64 KB)
 Trainable params: 45,987 (179.64 KB)
 Non-trainable params: 0 (0.00 B)
```

Figure 9. Proposed hybrid CNN-LSTM architecture
Note: CNN + LSTM - Convolutional Neural Network + Long Short Term Memory.

3.6.6 Hyper parameter tuning and model optimization for diabetic risk prediction

After Applying Proposed CNN-LSTM model, to enhance the performance of model hyper-parameters tuning and (L2 regularization) optimization of parameters is performed. The process of hyper parameter tuning involves optimizing the parameters that regulate the training process of a model.

By preventing the use of excessively high weight values, L2 regularization helps to decrease model overfitting. The loss function is then amended to include a penalty term that grows in square root of the model weights' size. Objective Function with L2 Regularization:

$$J(\theta) = L(\theta) + \lambda \sum_{j=1}^{n} \theta_j^2$$

where,

$L(\theta)$ = original loss (e.g., cross-entropy, MSE)
θ_j = model parameters
λ = regularization strength (hyperparameter)
$\sum \theta_j^2$ = L2 norm (squared)

Table 3. Hyper parameter description

Parameters	Value
Learning Rate	0.001
Batch Size	1000
Number of Epochs	10
Dropout Rate	0.5
Optimizer	adam
Activation Function	Relu, softmax
Loss Function	categorical_crossentropy
Number of Neurons/Units	128

The hyper-parameters utilized by the suggested hybrid CNN-LSTM model are displayed in Table 3. Now that the models have been trained and optimized, they can be used to predict the diabetic risk factor using fresh patient data. With the help of this prediction, it is possible to determine who is more likely to get diabetes, which paves the way for early intervention and individualized treatment strategies. Healthcare practitioners should be able to comprehend and take action based on the outcomes of the prediction, which will provide them with insights into the elements that are driving the risk.

3.7 Performance analysis

A variety of indicators, including Accuracy, Precision, Recall, F1-Score, Accuracy and Loss Graph, and Confusion Matrix, are used to evaluate the models' performance in the final stage.

4. RESULT ANALYSIS

4.1 Experimental setup

Table 4 describes the main elements and layout of the experiments, the detailed environment settings, the hardware resources, and the software utilization that assure the reproducibility and performance efficiency.

Table 4. Experimental setup used

Component	Description
Programming Language	Python 3.7
Development Environment	Google Colab
Hardware Configuration	Google Colab Environment with NVIDIA Tesla K80 GPU
Software Libraries	TensorFlow 2.4.1, Keras 2.4.3, Scikit-learn 0.24, Pandas 1.1.5, NumPy 1.19.5, Matplotlib 3.3.4, Seaborn 0.11.1
Data Processing Tools	Pandas for data handling and pre-processing, Scikit-learn for feature selection and scaling
Model Training Platform	Google Colab with GPU support for faster training
Operating System	Windows 10

4.2 Dataset used

The Kaggle COVID-19 dataset by Meir Nizri [33] includes rich data that can be used for studying the COVID-19 cases. It consists of several features that are patient demographics, symptoms, results of blood examination and a test on diabetes. This dataset is useful for predictive modeling of the virus spread, disease epidemiology research and factors influencing patient outcome.

4.3 Results and comparative analysis

By comparing its outcomes with those of other ML/DL models, we were able to establish that the hybrid CNN-LSTM model accurately predicts the diabetes-induced risk of COVID-19 patients. To validate the models, we looked at their precision, recall F1 score, accuracy loss curve graph, and classification accuracy. Furthermore, the influence of imbalanced data and data balancing methods (like SMOTE) was evaluated in a comparison between results obtained with original and balanced datasets. Our suggested hybrid CNN-LSTM model outperformed competing models on recall and F1-score measures, as shown in the performance evaluation. Furthermore, this demonstrates that it is able to incorporate spatial patterns in addition to temporal occurrences, which is an essential component in accurately predicting the likelihood of developing diabetes.

4.3.1 Without data balancing technique

Figure 10(a) and Figure 10(b) show the accuracy and loss curves, respectively, that show how well the hybrid CNN-LSTM model performed when data balancing wasn't used. We see a steep ascent in our model's accuracy on the training set in the early epochs of the accuracy graph, which levels off at a little over 88%. The validation accuracy behaves similarly, yet falls slightly lower, finding a decreased effect. This means model is a little over fitted, but it is learning well. The loss curve backs up this finding; when both the validation loss and the training loss fall off sharply in the beginning epochs and settle at around 0.30, they follow each other closely. The model may not be fully generalizing to new data, though, as indicated by the modest difference between training and validation losses. This could be due to the training data's class imbalance.

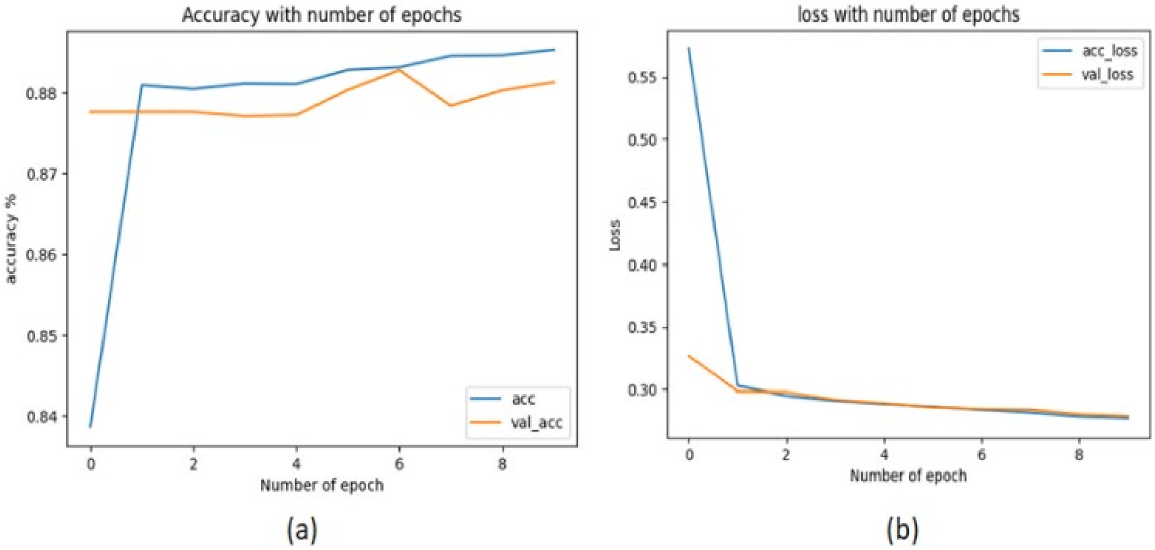


Figure 10. (a) Accuracy curve (b) Loss curve of hybrid CNN- LSTM model on imbalance dataset
Note: CNN + LSTM - Convolutional Neural Network + Long Short Term Memory.

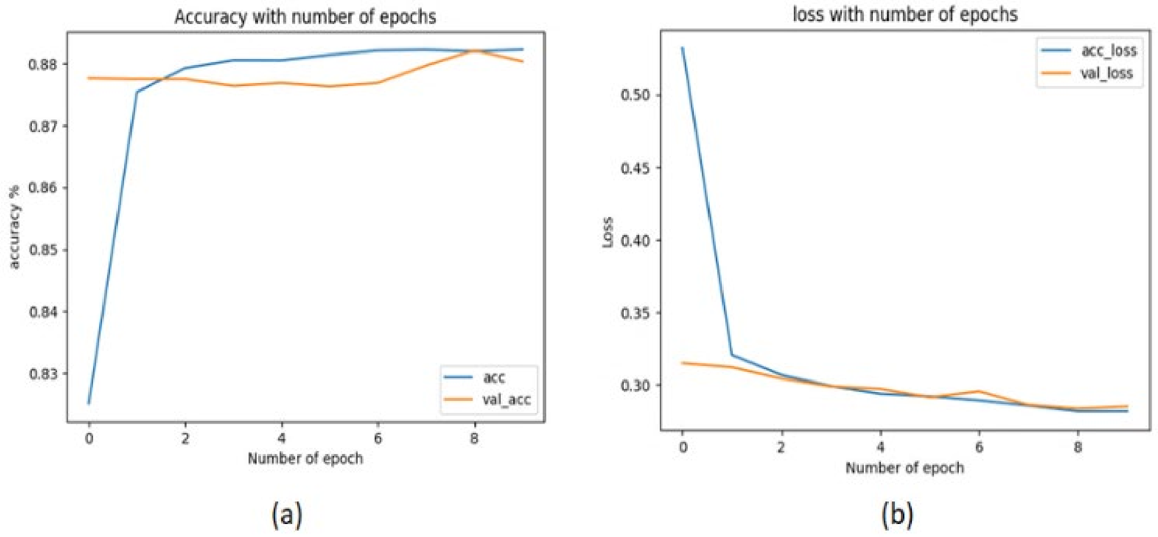


Figure 11. (a) Accuracy curve (b) Loss curve of hybrid CNN- LSTM model on balance dataset
Note: CNN + LSTM - Convolutional Neural Network + Long Short Term Memory.

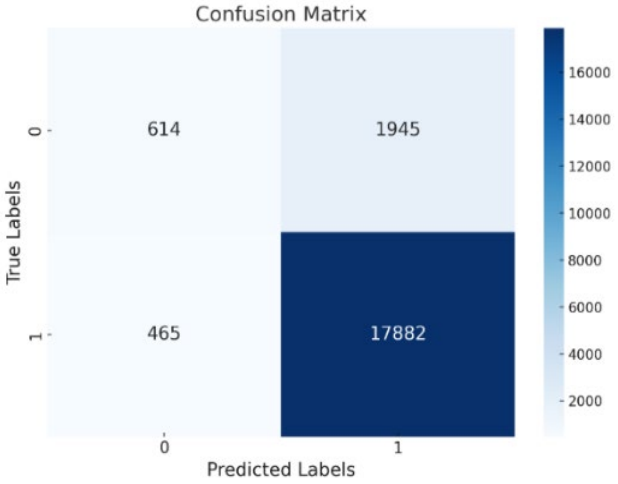
4.3.2 With data balancing technique

Figure 11(a) and Figure 11(b) show the accuracy and loss curves, respectively, that show how well the hybrid CNN-LSTM model performed when data balancing is used. It is possible to gain significant information regarding the performance of the Hybrid CNN-LSTM model by analyzing the accuracy and loss curves of the model that was trained without the use of data balancing algorithms. This is demonstrated by the accuracy curve, which demonstrates that the training accuracy of the model progressively grows until it approaches 90%. There appears to be a little generalization gap, since the validation accuracy levels out at slightly lower levels than the training accuracy.

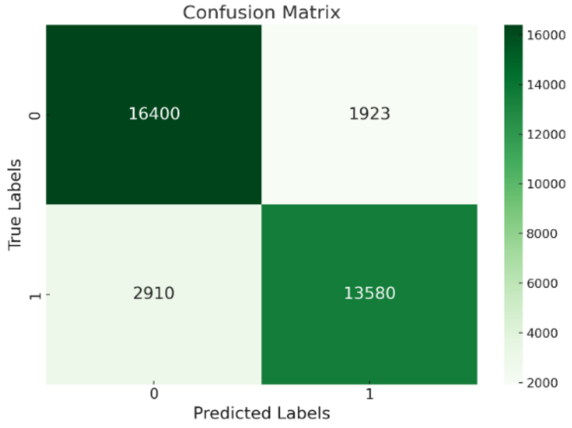
4.3.3 Confusion matrix

The confusion matrices in Figure 12 illustrate how this imbalance may result in a model that does well on the majority class but poorly on the minority class. Without making use of any data balancing strategies, the confusion matrices that have been presented exhibit the outcomes of proposed hybrid CNN-LSTM model. These findings indicate that there is a considerable class imbalance, in which the models mostly forecast the class that includes the majority of individuals (label 1). This is evident from the very low count of genuine negatives (found in the top left quadrant of each matrix) relative to the substantial number of true positives (placed in the bottom right quadrant of each matrix). This imbalance indicates that, without data balancing, the models exhibit bias towards the majority class. Figure 12(a) illustrates the corresponding confusion matrix of the proposed hybrid CNN-LSTM model applied to an imbalanced dataset.

The model has trouble predicting the minority class in earlier confusion matrix results, then this could be the reason why the model is biased towards the majority class. The confusion matrices (shown in Figure 12(b)) that were displayed after the application of data balancing approaches demonstrated a discernible enhancement in the model's capacity to make more accurate predictions about both classes. A superior distribution of genuine positives (label 1 accurately predicted) and true negatives (label 0 correctly predicted) is demonstrated by the findings. Because of the data balancing strategy, the bias towards the majority class has been effectively addressed, which has enabled the model to perform more evenly across both groups. When it comes to applications that require precise detection of both classes, this balanced performance is absolutely necessary. This ensures that neither class is misclassified to a disproportionate degree.



(a)



(b)

Figure 12. (a) CNN -LSTM confusion matrix on imbalance dataset; (b) CNN -LSTM confusion matrix on balance dataset
Note: CNN + LSTM - Convolutional Neural Network + Long Short Term Memory.

4.4 Comparative analysis graphs

The performance of different ML, DL and proposed hybrid CNN-LSTM algorithms with and without data balancing strategies is shown in the Table 5 and related graphs are shown in Figure 13 and Figure 14. It is clear that models that have been trained using data balancing strategies typically show improved accuracy and F1-score for all algorithms. For instance, when data balance is used, the accuracy of the hybrid CNN-LSTM model, Random Forest (RF), and Logistic Regression (LR) all somewhat improves. Data balance improves the accuracy of the AdaBoost and Decision Tree (DT) models greatly, increasing them from roughly 74% to approximately 88%. The graph shows that even while most models have good baseline accuracy even in the absence of data balancing, applying data balancing approaches provides more dependable and consistent performance while reducing the possibility of bias towards any certain class. Figure 15 shows the respective ROC AUC curve of all models.

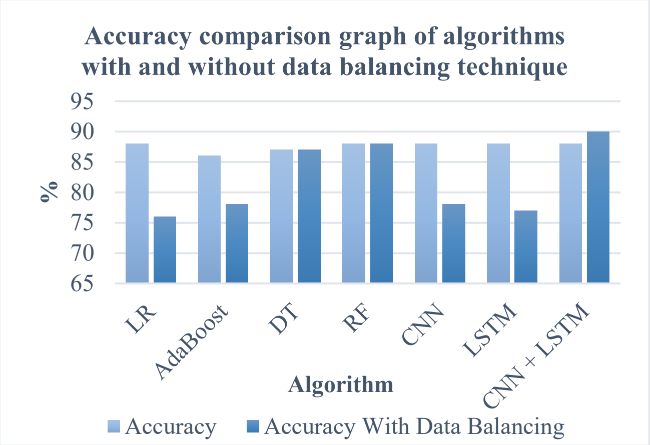


Figure 13. Accuracy comparison graph of algorithms with and without data balancing technique
Note: LR – Logistic Regression, DT – Decision Tree, RF – Random Forest, CNN – Convolutional Neural Network, LSTM – Long Short Term Memory, CNN + LSTM - Convolutional Neural Network + Long Short Term Memory.

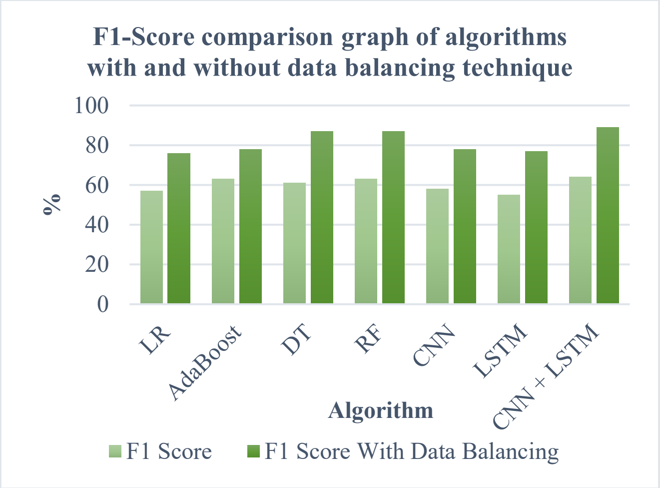


Figure 14. F1-Score comparison graph of algorithms with and without data balancing technique
Note: LR – Logistic Regression, DT – Decision Tree, RF – Random Forest, CNN – Convolutional Neural Network, LSTM – Long Short Term Memory, CNN + LSTM - Convolutional Neural Network + Long Short Term Memory.

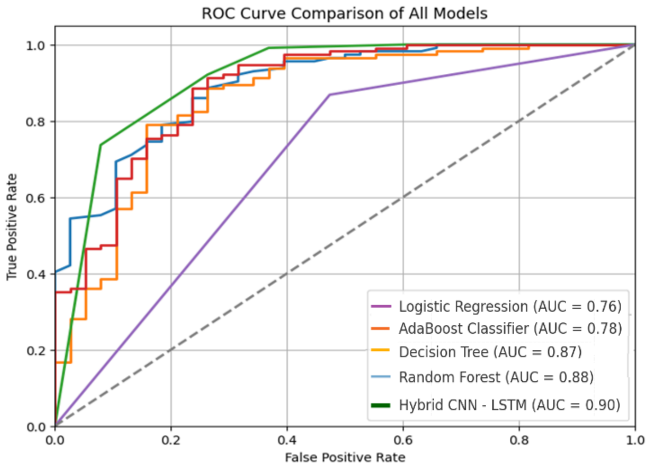


Figure 15. ROC AUC curve of all models
Note: AUC – ROC: Area Under the Receiver Operating Characteristic Curve.

Table 5. Performance comparison of models with and without data balancing

Algorithms	Accuracy	Precision (Macro)	Recall (Macro)	F1-Score (Macro)
Without Data Balancing				
Logistic Regression	88	67	55	57
AdaBoost	86	64	62	63
Decision Tree	87	66	60	61
Random Forest	88	71	61	63
CNN	88	72	56	58
LSTM	88	74	54	55
Hybrid CNN - LSTM	88	74	61	64
With Data Balancing				
Logistic Regression	76	76	76	76
AdaBoost	78	78	78	78
Decision Tree	87	87	87	87
Random Forest	87	88	87	87
CNN	78	77	77	78
Hybrid CNN - LSTM	90	89	89	89

5. CONCLUSION

Finding a trustworthy model to predict the likelihood of diabetes in COVID-19 patients is crucial. We present hybrid CNN-LSTM model having better performance perspective than traditional ML and other DL models because the proposed hybrid model captures spatial and temporal dependencies well. Balancing of the class distribution using SMOTE greatly improved model performances, enabling them to make reliable predictions in imbalanced settings. Moreover, Random Forest feature selection decreased the number of features and contributed to a more interpretable model, with still higher accuracy. The hybrid CNN-LSTM model yielded a superior recall performance of 89% and macro F1-scores of 89%, compared to Random Forest, Logistic Regression and AdaBoost, which demonstrates high robustness in tracing cutoff-risk population. The results suggested that a well-designed system implementing data balancing, feature selection and advanced modelling approaches is essential towards accurate risk prediction as well as the hospital response to patient's treatment in both the acute phase of the ongoing pandemic as well as during remission.

REFERENCES

[1] Hao, Y.J., Wang, Y.L., Wang, M.Y., Zhou, L., Shi, J.Y., Cao, J.M., Wang, D.P. (2022). The origins of COVID-19 pandemic: A brief overview. Transboundary and Emerging Diseases, 69(6): 3181-3197. https://doi.org/10.1111/tbed.14732

[2] Khawandanah, J. (2019). Double or hybrid diabetes: A systematic review on disease prevalence, characteristics and risk factors. Nutrition & Diabetes, 9(1): 33. https://doi.org/10.1038/s41387-019-0101-1

[3] Sharma, P., Behl, T., Sharma, N., Singh, S., Grewal, A.S., Albarrati, A., Albratty, M., Meraya, A.M., Bungau, S. (2022). COVID-19 and diabetes: Association intensify risk factors for morbidity and mortality. Biomedicine & Pharmacotherapy, 151: 113089. https://doi.org/10.1016/j.biopha.2022.113089

[4] Tasin, I., Nabil, T.U., Islam, S., Khan, R. (2023). Diabetes prediction using machine learning and explainable AI techniques. Healthcare Technology Letters, 10(1-2): 1-10. https://doi.org/10.1049/htl2.12039

[5] Zou, Q., Qu, K., Luo, Y., Yin, D., Ju, Y., Tang, H. (2018). Predicting diabetes mellitus with machine learning techniques. Frontiers in Genetics, 9: 515. https://doi.org/10.3389/fgene.2018.00515

[6] Abdelhafez, H.A., Amer, A.A. (2024). Machine learning techniques for diabetes prediction: A comparative analysis. Journal of Applied Data Sciences, 5(2): 792-807. https://doi.org/10.47738/jads.v5i2.219

[7] Fakhar, M.H., Baig, M.Z., Ali, A., Rana, M.T.A., Khan, H., Afzal, W., Farooq, H.U., Albouq, S. (2024). A deep learning-based architecture for diabetes detection, prediction, and classification. Engineering, Technology & Applied Science Research, 14(5): 17501-17506. https://doi.org/10.48084/etasr.8354

[8] Elsayed, A.M., Elsayed, M.S., Mansour, A.E., Mahedy, A.W., Araby, E.M., Morsy, M.H., Abd Elmoniem, R.O. (2024). Predictors of COVID-19 outcome in type 2 diabetes mellitus: A hospital-based study. The Egyptian Journal of Internal Medicine, 36(1): 2. https://doi.org/10.1186/s43162-023-00266-5

[9] Lipski, D., Radziemski, A., Wasiliew, S., Wyrwa, M., Szczepaniak-Chicheł, L., Stryczyński, Ł., Olasińska-Wiśniewska, A., Urbanowicz, T., Perek, B., Tykarski, A., Komosa, A. (2024). Assessment of COVID-19 risk factors of early and long-term mortality with prediction models of clinical and laboratory variables. BMC Infectious Diseases, 24(1): 685. https://doi.org/10.1186/s12879-024-09592-7

[10] Kabootari, M., Tirtashi, R.H., Hasheminia, M., Bozorgmanesh, M., Khalili, D., Akbari, H., Roshandel, G., Hadaegh, F. (2022). Clinical features, risk factors and a prediction model for in-hospital mortality among diabetic patients infected with COVID-19: Data from a referral centre in Iran. Public Health, 202: 84-92. https://doi.org/10.1016/j.puhe.2021.11.007

[11] Khadem, H., Nemat, H., Eissa, M.R., Elliott, J., Benaissa, M. (2022). COVID-19 mortality risk assessments for individuals with and without diabetes mellitus: Machine learning models integrated with interpretation framework. Computers in Biology and Medicine, 144: 105361. https://doi.org/10.1016/j.compbiomed.2022.105361

[12] McGurnaghan, S.J., Weir, A., Bishop, J., Kennedy, S., et al. (2021). Risks of and risk factors for COVID-19 disease in people with diabetes: A cohort study of the total population of Scotland. The Lancet Diabetes & Endocrinology, 9(2): 82-93. https://doi.org/10.1016/S2213-8587(20)30405-8

[13] Schlesinger, S., Neuenschwander, M., Lang, A., Pafili, K., Kuss, O., Herder, C., Roden, M. (2021). Risk phenotypes of diabetes and association with COVID-19 severity and death: A living systematic review and meta-analysis. Diabetologia, 64(7): 1480-1491. https://doi.org/10.1007/s00125-021-05458-8

[14] Kristan, M.M., Kim, Y.K., Nelson, T., Moxley, M.C., Yip, T.C.F., Munir, K., Malek, R. (2021). Predictors of severe COVID-19 in patients with diabetes: A multicenter review. Endocrine Practice, 27(8): 842-849. https://10.1016/j.eprac.2021.05.011

[15] Ramesh, J., Reddy, S.S., Rajesh, M., Varghese, J. (2021). Evaluation of simple and cost-effective immuno-haematological markers to predict outcome in hospitalized severe COVID-19 patients, with a focus on diabetes mellitus-A retrospective study in Andhra Pradesh, India. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 15(3): 739-745. https://doi.org/10.1016/j.dsx.2021.03.025

[16] Orioli, L., Servais, T., Belkhir, L., Laterre, P.F., Thissen, J.P., Vandeleene, B., Maiter, D., Yombi, J.C., Hermans, M.P. (2021). Clinical characteristics and short-term prognosis of in-patients with diabetes and COVID-19: A retrospective study from an academic center in Belgium. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 15(1): 149-157. https://doi.org/10.1016/j.dsx.2020.12.020

[17] Ortega, E., Corcoy, R., Gratacòs, M., Cos Claramunt, F.X., Mata-Cases, M., Puig-Treserra, R., Real, J., Vlacho, B., Castelblanco, E., Domingo, P., Khunti, K., Franch-Nadal, J., Mauricio, D. (2021). Risk factors for severe outcomes in people with diabetes hospitalised for COVID-19: A cross-sectional database study. BMJ Open, 11(7): e051237. https://doi.org/10.1136/bmjopen-2021-051237

[18] Mamtani, M., Athavale, A.M., Abraham, M., Vernik, J., Amarah, A.R., Ruiz, J.P., Joshi, A.J., Itteera, M., Zhukovski, S.D., Madaiah, R.P., White, B.C., Hart, P., Kulkarni, H. (2021). Association of hyperglycaemia with hospital mortality in nondiabetic COVID-19 patients: A cohort study. Diabetes & Metabolism, 47(3): 101254. https://doi.org/10.1016/j.diabet.2021.101254

[19] Collard, D., Nurmohamed, N.S., Kaiser, Y., Reeskamp, L.F., Dormans, T., Moeniralam, H., Simsek, S., Douma, R., Eerens, A., Reidinga, A.C., Elbers, P.W.G., Beudel, M., Vogt, L., Stroes, E.S.G., van den Born, B.J.H. (2021). Cardiovascular risk factors and COVID-19 outcomes in hospitalised patients: A prospective cohort study. BMJ Open, 11(2): e045482. https://10.1136/bmjopen-2020-045482

[20] Pazoki, M., Keykhaei, M., Kafan, S., Montazeri, M., Mirabdolhagh Hazaveh, M., Sotoodehnia, M., Kazemian, S., Talebpour, M., Ashraf, H., Moharari, R.S., Majidi, F., Amlashi, N.H., Zare, S., Fathi, M., Azimbeik, Z., Ebrahimi, M. (2021). Risk indicators associated with in-hospital mortality and severity in patients with diabetes mellitus and confirmed or clinically suspected COVID-19. Journal of Diabetes & Metabolic Disorders, 20(1): 59-69. https://doi.org/10.1007/s40200-020-00701-2

[21] Lasbleiz, A., Cariou, B., Darmon, P., Soghomonian, A., Ancel, P., Boullu, S., Houssays, M., Romain, F., Lagier, J.C., Boucekine, M., Resseguier, N., Gourdy, P., Pichelin, M., Wargny, M., Dutour, A., Gaborit, B. (2020). Phenotypic characteristics and development of a hospitalization prediction risk score for outpatients with diabetes and COVID-19: The DIABCOVID study. Journal of Clinical Medicine, 9(11): 3726. https://doi.org/10.3390/jcm9113726

[22] Bello-Chavolla, O.Y., Bahena-López, J.P., Antonio-Villa, N.E., Vargas-Vázquez, A., González-Díaz, A., Márquez-Salinas, A., Fermín-Martínez, C.A., Jesús Naveja, J., Aguilar-Salinas, C.A. (2020). Predicting mortality due to SARS-CoV-2: A mechanistic score relating obesity and diabetes to COVID-19 outcomes in Mexico. The Journal of Clinical Endocrinology & Metabolism, 105(8): 2752-2761. https://doi.org/10.1210/clinem/dgaa346

[23] Khodabakhsh, P., Asadnia, A., Moghaddam, A.S., Khademi, M., Shakiba, M., Maher, A., Salehian, E. (2023). Prediction of in-hospital mortality rate in COVID-19 patients with diabetes mellitus using machine learning methods. Journal of Diabetes & Metabolic Disorders, 22(2): 1177-1190. https://doi.org/10.1007/s40200-023-01228-y

[24] Yu, C., Lei, Q., Li, W., Wang, X., Liu, W., Fan, X., Li, W. (2020). Clinical characteristics, associated factors, and predicting COVID-19 mortality risk: A retrospective study in Wuhan, China. American Journal of Preventive Medicine, 59(2): 168-175. https://doi.org/10.1016/j.amepre.2020.05.002

[25] Carrasco-Sánchez, F.J., López-Carmona, M.D., Martínez-Marcos, F.J., Pérez-Belmonte L.M., et al. (2021). Admission hyperglycaemia as a predictor of mortality in patients hospitalized with COVID-19 regardless of diabetes status: Data from the Spanish SEMI-COVID-19 Registry. Annals of Medicine, 53(1): 103-116. https://doi.org/10.1080/07853890.2020.1836566

[26] Zhang, N., Wang, C., Zhu, F., Mao, H., Bai, P., Chen, L.L., Zeng T, Peng, M.M., Qiu, K.L., Wang, Y., Yu, M., Xu, S., Zhao, J., Li, N., Zhou, M. (2020). Risk factors for poor outcomes of diabetes patients with COVID-19: A single-center, retrospective study in early outbreak in China. Frontiers in Endocrinology, 11: 571037. https://10.3389/fendo.2020.571037

[27] Apicella, M., Campopiano, M.C., Mantuano, M., Mazoni, L., Coppelli, A., Del Prato, S. (2020). COVID-19 in people with diabetes: Understanding the reasons for worse outcomes. The Lancet Diabetes & Endocrinology, 8(9): 782-792. https://doi.org/10.1016/S2213-8587(20)30238-2

[28] Sourij, H., Aziz, F., Bräuer, A., Ciardi, C., et al. (2021). COVID-19 fatality prediction in people with diabetes and prediabetes using a simple score upon hospital admission. Diabetes, Obesity and Metabolism, 23(2): 589-598. https://doi.org/10.1111/dom.14256

[29] Liu, Z., Bai, X., Han, X., Jiang, W., Qiu, L., Chen, S., Yu, X. (2020). The association of diabetes and the prognosis of COVID-19 patients: A retrospective study. Diabetes Research and Clinical Practice, 169: 108386. https://doi.org/10.1016/j.diabres.2020.108386

[30] Singh, A.K., Khunti, K. (2020). Assessment of risk, severity, mortality, glycemic control and antidiabetic agents in patients with diabetes and COVID-19: A narrative review. Diabetes Research and Clinical Practice, 165: 108266. https://doi.org/10.1016/j.diabres.2020.108266

[31] Zhang, Y., Cui, Y., Shen, M., Zhang, J., Liu, B., Dai, M., Chen, L., Han, D., Fan, Y., Zeng, Y., Li, W., Lin, F., Li, S., Chen, X., Pan, P. (2020). Association of diabetes mellitus with disease severity and prognosis in COVID-19: A retrospective cohort study. Diabetes Research and Clinical Practice, 165: 108227. https://doi.org/10.1016/j.diabres.2020.108227

[32] Guo, W., Li, M., Dong, Y., Zhou, H., Zhang, Z., Tian, C., Qin, R., Wang, H., Shen, Y., Du, K., Zhao, L., Fan, H., Luo, S.S., Hu, D. (2020). Diabetes is a risk factor for the progression and prognosis of COVID-19. Diabetes/Metabolism Research and Reviews, 36(7): e3319. https://doi.org/10.1002/dmrr.3319

[33] COVID-19 Dataset. https://www.kaggle.com/datasets/meirnizri/covid19-dataset.

[34] Taoussi, C., Hafidi, I., Metrane, A. (2025). Machine learning and deep learning in health informatics: Advancements, applications, and challenges. Engineering Letters, 33(5): 1448-1461.

[35] Dhabliya, D., Khetani, V., Gandhi, Y., Tanjeja, A.K., Patil, C.D., Patil, S. (2025). Threat Intelligence Fusion with AI and Machine Learning. In: Bhateja, V., Patel, P., Tang, J. (eds) Evolution in Computational Intelligence. FICTA 2024 2024. Smart Innovation, Systems and Technologies, Springer, Singapore. https://doi.org/10.1007/978-981-96-2124-8_44

[36] Mane, D., Ashtagi, R., Suryawanshi, R., Kaulage, A.N., Hedaoo, A.N., Kulkarni, P.V., Gandhi, Y. (2024). Diabetic retinopathy recognition and classification using transfer learning deep neural networks. Traitement du Signal, 41(5): 2683-2691. https://doi.org/10.18280/ts.410541

[37] Shimbre, N., Solanki, R.K. (2025). ChestXFusionNet: A multimodal deep learning framework for predicting chest diseases from X-ray images and clinical data. In EPJ Web of Conferences, 328: 01059. https://doi.org/10.1051/epjconf/202532801059

[38] Ashtagi, R., Bidwe, R.V., Fukate, A., Kulkarni, O., Jadhav, P., Patil, S. (2025). Sentiment analysis on YouTube comments using long short-term memory (LSTM) networks. In 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), Goathgaun, Nepal, pp. 785-788. https://doi.org/10.1109/ICMCSI64620.2025.10883113