Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

Big Data Analysis and prediction of COVID-2019 Epidemic Using Machine Learning Models in Healthcare Sector

Sazzat Hossain[1], Mohammad Muzahidur Rahman Bhuiyan[2], Md Shafiqul Islam[3], Mohammad Moniruzzaman[4], Md Kamal Ahmed[5], Niropam Das[6], Abu Saleh Muhammad Saimon[7], Mia Md Tofayel Gonee Manik[8]

Abstract

A global pandemic of COVID-19 is underway. The world now thinks coronaviruses cause all illnesses, including this pandemic. The world is also seeing fast viral spread. COVID-19 spreads swiftly by respiratory droplets and contact, according to the WHO. Big data analytics technologies are essential for developing the information needed to make choices and take preventative action. The large volume of COVID-19 data available by multiple sources necessitates, however, a review of big data analysis's roles in containing the virus, highlighting the primary obstacles and future directions in COVID-19 data analysis, and offering a framework for relevant current applications and studies to aid in COVID-19 analysis research in the future. The aim of ML, DL methods in COVID-19 research and applications is the focus of this study. This endeavor use machine learning to evaluate and predict COVID-19 pandemic behavior to help control a pandemic. A study uses the Israeli Department of Health dataset to predict COVID-19 results using ML and big data analytics. The test results were analyzed using CNN-LSTM and Decision Tree classification algorithms. Model efficacy was measured by F1-score, recall, accuracy, and precision. CNN-LSTM was the most accurate, with 96.34% accuracy and strong predictive ability. Also, for the model explainability used LIME model. The findings may be used by various government agencies to implement remedial actions. It may be simpler to fight COVID-19 if methods for infectious illness forecasting were readily available.

Keywords: *Covid-19, Deep Learning, Machine Learning, Explainable AI, CNN-LSTM, Decision Tree, Big data.*

Introduction

The devastating 2019 coronavirus disease (COVID-19) pandemic caught health care systems throughout the globe off guard, and many countries lacked the resources necessary to react to and contain the epidemic [1]. Paramedics and volunteers conduct the war against COVID-19 at the forefront. The research community is also actively working to battle and ultimately eradicate this devastating epidemic. Plenty of chances have arisen to create solutions that rely on technology. Over the last year, research on Big Data in relation to COVID-19 has shown its usefulness in several areas, including case tracking, epidemic surveillance, monitoring of viral propagation and human movement, preventative measures, medical treatment, and medication development [2][3]. Big Data has been promoted by cutting-edge designs and technologies to address a wide range of real-world issues, including the inevitable use of such data in the

[1] School of Business, International American University, Los Angeles, CA 90010, USA. Email: sazzat786@gmail.com. ORCID ID: https://orcid.org/0009-0008-6325-5496
[2] College of Business, Westcliff University, Irvine, CA 92614, USA. Email: m.bhuiyan.466@westcliff.edu. ORCID ID: https://orcid.org/0009-0001-1774-9726
[3] Department of Computer Science, Maharishi International University, Fairfield, Iowa 52557, USA. Email: shafiqswh@gmail.com. ORCID ID: https://orcid.org/0009-0008-9067-4987
[4] Department of Computer Science, Maharishi International University, Fairfield, Iowa 52557, USA. Email: mohammad.moniruzzaman35@gmail.com.. ORCID ID: https://orcid.org/0009-0006-5981-4473
[5] School of Business, International American University, Los Angeles, CA 90010, USA. Email: kamalacademic88@gmail.com. ORCID ID: https://orcid.org/0009-0004-1003-6207
[6] School of Business, International American University, Los Angeles, CA 90010, USA. Email: niropomdas124@gmail.com. ORCID ID: https://orcid.org/0009-0004-6107-7025
[7] Department of Information Technology, Washington University of Science and Technology, Alexandria VA 22314, USA. Email: abus.student@wust.edu. ORCID ID: https://orcid.org/0009-0006-3147-1755
[8] College of Business, Westcliff University, Irvine, CA 92614, USA. Email: m.manik.407@westcliff.edu. ORCID ID: https://orcid.org/0009-0005-6098-5213

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

fight against the pandemic. Understanding public sentiment, anxiety, and reaction to COVID-19-related regulations using social media analysis helps with real-world problem-solving [4][5].

Predicting healthcare data has recently been shown to be an efficient use of machine algorithms. [6] A major objective of ML is to make computers capable of learning on their own and making appropriate adjustments to their behaviour without any help from humans [7]. A subset of the historical data that wasn't utilised for training is used to evaluate how well the trained ML model performs. The validation process is the term used to describe this procedure. The ML model is evaluated in this procedure according to performance metrics like accuracy. An accuracy of a ML model is determined by how well it predicts the features of unseen data. It is calculated by taking the entire number of features that need prediction and dividing it by the amount of features that have accurate predictions [8].

The integration of machine learning and big data has become crucial for making precise and timely decisions, improving the capacity to effectively respond to the crisis. The studies summarized below provide an overview of the methodologies and findings in this rapidly evolving field, demonstrating the important role of data-driven approaches in addressing the COVID-19 epidemic.

Contribution of the study

There are numerous contributions that this study provides to the domain of covid-19 prediction. This work mostly contributes to the following areas:

- **Dataset Utilization**: Utilized a large-scale dataset from the Israeli Department of Health, enhancing COVID-19 prediction models with detailed symptom and demographic data.
- **Methodological Advances**: Implemented advanced oversampling techniques like SMOTE, and model validation methods to improve accuracy and reliability.
- **ML models**: To applied machine learning based different classification models for the Big Data Analysis and prediction of COVID-2019 Epidemic.
- **Explainable AI:** Data imbalance was addressed using the SMOTE approach and explainable AI using the LIME framework for machine model interpretation.
- **Performance Evaluation**: Compared multiple models using comprehensive metrics, providing insights into their effectiveness for COVID-19 prediction.

Structure of paper

What follows is an outline of the rest of the paper. In Section II, present a literature review on Big Data Analysis and prediction of the COVID-2019 Epidemic using ML models, Section III presents methods and methodology, and Section IV results analysis and discussion. Conclusion and future work of this study present in section V.

Literature Review

Provide some earlier research on ML-based COVID-19 prediction in this section. Several authors have attempted to forecast coronavirus inspection by the integration of deep learning and ML methods. It was suggested how the occlusion approach may be used to identify COVID-19.

In[9] presents the use of large-scale data analysis and the leveraging of machine learning methods to forecast the COVID-19-related mortality rate. Following comparison research, the MLR. The algorithm has shown its superiority by performing very well. The proposed method gets a high R-Squared score of 0.987 and an amazing accuracy rate of 98.6%, demonstrating its remarkable adaptability to various contexts.

In [10], advocated a deep learning strategy that utilises RNNs and LSTM networks to forecast the likely COVID-19 case counts. While RNN models had a precision accuracy of 93.45%, LSTM models demonstrated a 98.58% accuracy rate. Malaysia, Morocco, and Saudi Arabia were compared in terms of coronavirus cases and fatalities as a consequence.

In [11], sought to predict COVID-19 patients' need for an ICU stay and eventual death by creating risk ratings based on their clinical parameters upon presentation. For the testing dataset, the risk score model

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

produced accurate predictions of ICU admission with an AUC0.74 ([95% CI, 0.63-0.85], p = 0.001) and death with an AUC0.83 ([95% CI, 0.73-0.92], p<0.001). Important independent clinical factors that predicted COVID-19-related ICU hospitalisation and death were discovered in this investigation.

In [12], according to the research, the new coronavirus epidemic began in late December 2019 and has since spread to over 7 million individuals, killing over 0.40 million over the globe. With an R-squared score of 0.9992, the anticipated number of instances closely matches the actual numbers. According to the results, two key elements that may assist to slow the rising rate of Covid-19 transmission are social distance and lockdown.

In [13], study aims to identify disease mortality biomarkers that might assist in healthcare system decision-making and logistical planning by mining a database of blood samples by 485 infected patients in Wuhan, China. To achieve this goal, we used machine learning methods to choose three biomarkers that could accurately anticipate a patient's death more than ten days in advance, with an accuracy rate of more than 90%.

In[14], develops a ML model to analyse enormous volumes of social media data. Curated annotated datasets, COVID-19 Rumour, NIR, and Zenodo, and Google and Polifact Fact Checked websites were used to create the ML model. A LightGBM Classifier, the fastest and most accurate model with a balanced accuracy score 0.82, identified 329107 tweets as 'Fake'.

There has been a lot of study on automated COVID-19 detection, according to the literature studies cited above. Many machine learning algorithms have also been developed to predict when COVID-19 would strike. An explanation of the machine learning methods' prediction was not advanced by the majority of the study. For immediate prediction, most of these experiments did not include the automated detection method into a website or mobile app. Table 1 provide the summary of the related work on big covid 1-19 data analysis using ML models.

Table 1: Studies Involving Big Data Analysis in COVID-19 using Machine Learning Research and various techniques

Reference	Methodology	Dataset	Performance	Limitations & Future Work
[9]	Multiple Linear Regression	COVID-19 data	Accuracy: 98.6%, R-Squared: 0.987	Focus on improving adaptability to various scenarios and expanding dataset for validation
[10]	RNN, LSTM	COVID-19 cases in Malaysia, Morocco, Saudi Arabia (up to Dec 3, 2020)	LSTM: 98.58% precision, RNN: 93.45% precision	Extend predictions beyond seven days and incorporate real-time data updates
[11]	Risk score based on clinical characteristics, ROC analysis	641 hospitalized patients with confirmed COVID-19	AUC: ICU admission 0.74 (95% CI, 0.63–0.85), mortality 0.83 (95% CI, 0.73–0.92)	Further validation with larger and more diverse datasets needed
[12]	Multiple linear regression, autocorrelation, autoregression	COVID-19 cases in India (up to June 6, 2020)	R-Squared: 0.9992	Explore the impact of additional factors like vaccination rates and new variants
[13]	Machine learning for biomarker identification	Blood samples from 485 patients in Wuhan, China	Accuracy: >90% for predicting mortality	Validate model with larger patient cohorts and investigate additional biomarkers
[14]	Light GBM Classifier, social media data analysis	Annotated datasets from multiple sources, 976087 South African COVID-19 tweets	Balanced accuracy: 0.82, identified 329107 tweets as 'Fake'	Enhance model with real-time data and expand to other regions to combat misinformation

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

Methodology

The methodology for developing the automatic COVID-19 prediction system involved several key steps, shows in figure 1. First, data collection was conducted using an open-source dataset containing individual symptom scores, basic patient information, and COVID-19 test outcomes for 2,742,596 patients. Next, data preprocessing was performed to clean and prepare the dataset for model training. This included removing records with missing age and gender information. The SMOTE oversampling approach was used to address a problem of class imbalance. Applying a stratified option to ensure evenly distributed classes, the dataset was thereafter splitted into a training set comprising 75% of the total and a testing set comprising 25%. In order to decipher the models, LIME-based explainable AI was used. Finally, various classification models, including CNN-LSTM and decision trees, were trained and evaluated to develop the prediction system. After the model predictions are made, performance evaluated to achieve the best outcome employing metrics like F1 score, recall, precision, and accuracy.

The following steps of data flow diagram is detailed description given in below:

Data Collection

In order to find trends and practical ways to study issue areas, data collection involves collecting information by many sources and then analysing it. The dataset used in this comparison analysis includes basic patient information, 2742,596 patients' test results for COVID-19, and individual outcomes of various symptoms.

Data Preprocessing

Preprocessing refers to the elimination of undesirable data from a dataset. This study includes initial investigation and data preparation prior to using the dataset for automated prediction model induction. The dataset often lacks patient information such as age and gender. Using average values to replace missing values might lead to bias in the overall dataset. To reduce data size, records with missing age and gender information were removed. Dropping missing values reduces the dataset size to 2186,227 rows. We removed the patient age feature from our dataset due to its poor association with COVID-19 findings, making it ineffective for training. The rows with unconfirmed COVID-19 findings have been removed. The dataset contains 2151,898 verified COVID-19 instances, with 1943,172 being negative and 208,726 being positive. To train machine learning models that need all features in numerical variables, all attributes are translated from categorical variables.

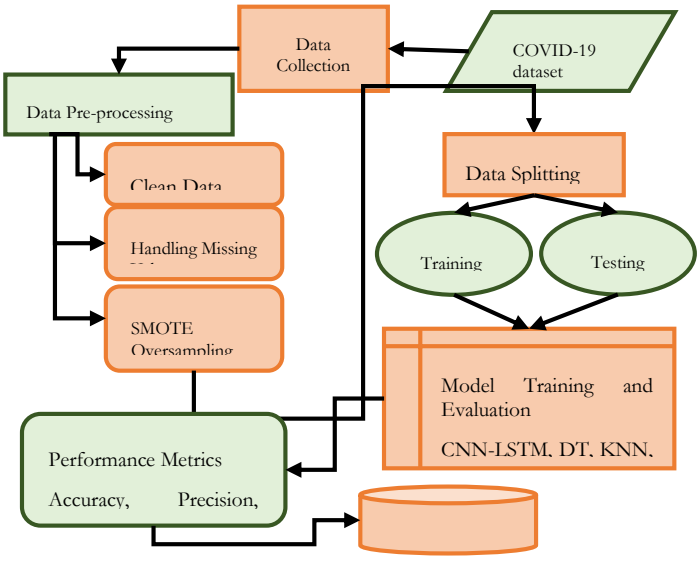

Figure 1: Methodology flow Diagram for big data analysis

SMOTE

We use SMOTE (Synthetic Minority Over-sampling Technique), which interpolates current samples to generate fake minority groups with comparable but not identical illness classes, to boost the dataset's diversity. It is well-known that this balancing strategy improves the outcomes produced by machine learning

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

algorithms [15]. To generate fake data, the SMOTE method employs the K-nearest neighbour pattern. It finds the K-nearest neighbours in the dataset after randomly selecting data points by a minority class. By picking the KNN from a randomly selected data, it creates synthetic data.

Classification Models

Some categorization models for the prediction of covid-19 are described in this section. These models used for comparative analysis.

1) Decision Tree

A decision tree representing f or a close approximation of it may be constructed from a collection of (x, f(x)) pairings. This process is known as decision tree learning. Although the set of pairings might theoretically be comprehensive when x is a finite domain, in practice, it is more common to draw from a (potentially infinite) domain X [16].

2) Hybrid Deep Learning Model

To train directional power flow prediction, two kinds of HDL were used, as mentioned before. Both the CNN-LSTM and the LSTM-CNN models are involved. See Figure 4 for an illustration of how CNN layers formed the basis of CNN-LSTM architecture. Features from the input dataset are what we're trying to extract. The CNN layers' outputs were sent to the LSTM layers to aid with sequence prediction, with a dense layer serving as the output.

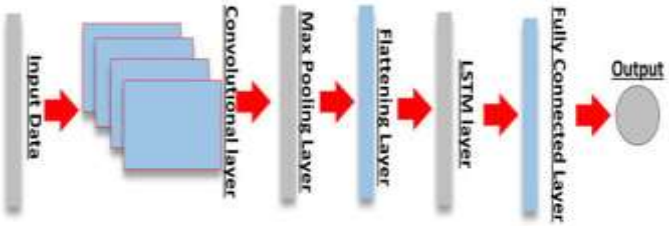

Figure 4: The structure of CNN-LSTM model.

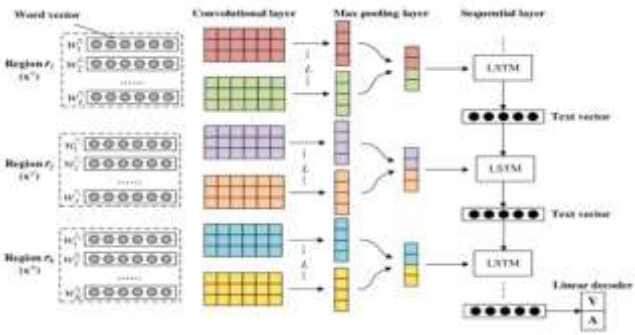

Figure 5 Layer structure of the CNN-LSTM model.

The operational sequences of a suggested COVID-19 forecasting system are shown in Fig. 5. The dataset was preprocessed and appropriate feature selection methods were used before training several ML models. The performance of different models is then contrasted in terms of different assessment indicators. Ultimately, the optimal model is used to develop the suggested automated COVID-19 prediction system.

LIME Based Explainability Model

In order to better comprehend the predictions made by machine learning models, explainable AI compiles a number of frameworks [17]. This study explains and predicts the predictions made by machine learning classifiers using the LIME-based explain ability technique. By making local approximations to the estimate locations, the LIME model is able to decipher the estimates that the machine learning model has generated[18]. The original LIME model has stability problems since it might provide various explanations when repeatedly assigned under same situations. In this research, we use an improved iteration of a LIME model to address the issue of instability

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

Results and Discussion

The simulated results of covide-19 prediction based on machine learning techniques in big data analytics discussed in this section. Results, dataset description, performance metrics, and classifier statistics are all part of this section, which displays the outcomes of the dataset assessment that was conducted for this research.

Dataset Description

The Israeli Department of Health provided the open-source dataset used in this investigation. In this dataset, you may find three different kinds of data. Included in the dataset are the following pieces of basic patient information: the date of the test, the patient's gender, and, if they are 60 or older, their age. Symptom indicators include the following: fever, sore throat, cough, shortness of breath, and headache. Furthermore, the dataset contains the patient's COVID-19 test result together with the patient's recent contact status with other COVID-19 patients.

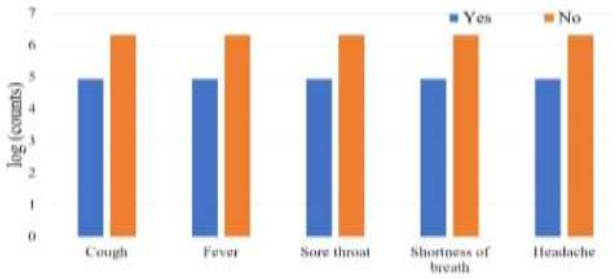

Figure 6: COVID-19 test results in terms of the features in the dataset.

Figure 6 illustrates that the dataset is severely skewed due to the large number of negative instances (9.3:1.0) and the relatively small number of positive occurrences.

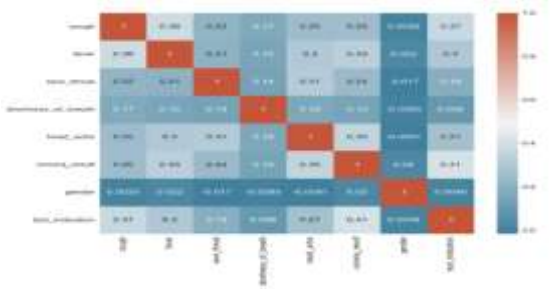

Figure 7: Correlation of various features of the used dataset.

Figure 7 displays the results of using the Pearson correlation index to quantify the linear dependency of different characteristics in the dataset. Additionally, we have resampled the dataset using the SMOTE oversampling approach to get a 1:1 ratio for the training data, which allows us to go deeper into the topic and provide a more balanced dataset for model training.

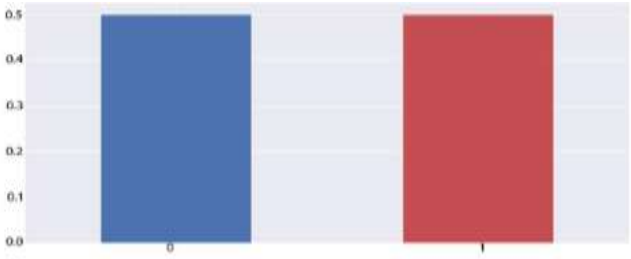

Figure 8: Training dataset after applying oversampling

After using the SMOTE oversampling approach, the training dataset is displayed in Figure 8. In figure x-axis displays a positive and negative class with balanced blue and red bars.

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

Performance Measures

Four distinct performance metrics— precision, accuracy, recall, and F1-score—have been utilized to assess an efficacy of every model. The following parameters are provided:

a) Confusion Matrix

A confusion matrix is one of a most well-known academic performance measures used to analyses the outcomes. The matrix's visual is shown in Figure 12. Four main qualities display the result data in the matrix, which is the combination of results from classifications. A true positive (TP) result is one in which the actual value matches the anticipated value of the classification. True negative (TN) principles are similar, only they center on zero. In the case of a false positive (FP), the expected value is 1 but the real value is 0, and in the case of a false negative (FN), the reverse is true.

b) Accuracy

Accuracy is the proportion of cases the model correctly categorized and the total error in class prediction. This measure summarizes the model's performance across classes. However, skewed data may misrepresent performance.

$$Acuracy = \frac{TP + TN}{TP + TN + FN + FP} \ldots . (4)$$

c) Precision

When all the data is classified into a class, precision is the percentage of cases that are correctly assigned to that class. It is calculated using the one-vs-all method for each class:

$$Precision = \frac{TP}{TP + FP} \ldots . (5)$$

d) Recall

Number of instances properly sorted into a class is assessed by recall or sensitivity. Recall is calculated using the one-vs-all approach, like precision.

$$Recall = \frac{TP}{TP + FN} \ldots . (6)$$

e) F1-score

The F1-score or F-measure is the weighted harmonic mean of recall and precision. When the dataset is heavily imbalanced, this measure is the most appropriate to use.

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} \ldots . (7)$$

Experiment Results

This section poses the findings of the Decision Tree and CNN-LSTM models applied to a large dataset for predicting covid-19 employing machine learning.

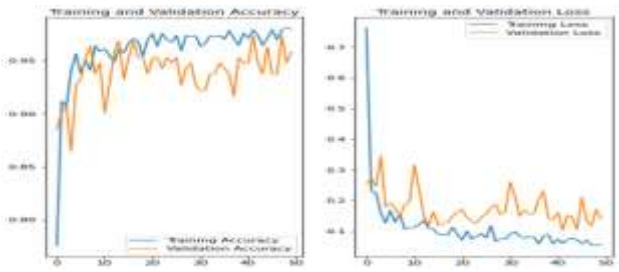

Figure 10: Training and validation/loss accuracy graphs of the CNN-LSTM model with SMOTE.

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

Figure 10 shows the accuracy/loss during training and validation as a function of a number of epochs utilized in the CNN-LSTM model approach. The COVID-19 symptoms are characterised by a progressive and temporal nature. With a validation accuracy of 96.34%, a CNN-LSTM model outperformed the competition, as expected.

Table 1: CNN-LSTM and DT models performance matrices for covid-19 data prediction

Model	Accuracy	Precision	Recall	F1-score
CNN-LSTM	96.34	98	98	98
Decision Tree	92.18	93	92	92

There are many performance measures that are shown in Table 1 for the CNN-LSTM and DT models. These measures include accuracy 96.34%, and F1-score, recall, and precision are 98% of CNN-LSTM model while on the DT model get 92/18%accuracy and 92%precision, recall and f1-score. Figure 16 provides a graphical representation of the value of these measures, which is shown next to the graph for the purpose of enabling better comprehension.

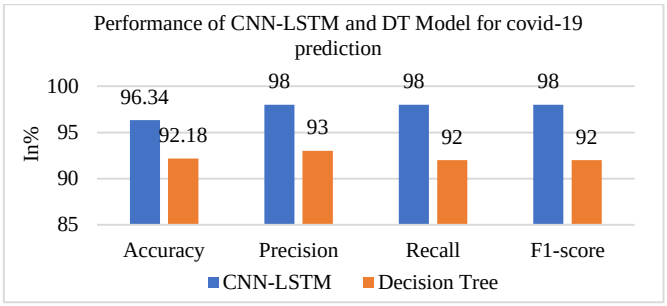

Figure 11: Bar graph of parameters Performance of CNN-LSTM and DT model for covid-19 health prediction

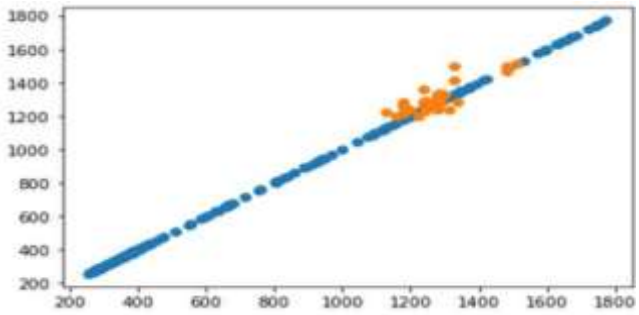

Figure 13: Actual and Forecast results by Decision Tree

Fig.13 shows that the projected value curve tendencies are extremely close to the real value one's, and that the numbers 200–1800 are labelled on an x-axis and a same range is on a y-axis of this scatter plot. It implies that the two variables are positively correlated with one another. this implies is that the y-axis values tend to grow in tandem with the x-axis values. the anticipated values comply the actual ones quite well.

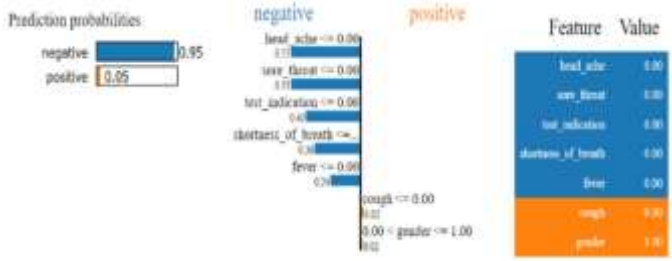

Figure 14: Interpretation of prediction results by the LIME approach

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

Figure 14 shows how the COVID-19 prediction might be interpreted using the LIME explainable AI paradigm. As an example, the blue bars show the treatment's history and symptoms that strongly suggest a bad forecast, whereas the orange bars express the opposite. Test indications, fever, headache, sore throat, and shortness of breath are the most important factors in generating the negative COVID-19 prognosis with 95% confidence in this instance, according to the explanation.

Comparative Analysis and Discussion

A comparative comparison using different models for COVID-19 prediction. In terms of performance measures, the following table 3 compares and contrasts different DL and ML models used for COVID-19 prediction.

Table 2: Comparison between various model for covid-19 prediction

Models	CNN-LSTM	Decision Tree	KNN [19]	ANN [20]
Accuracy	96.34	92.18	68	90.28
precision	98	93	66	87
Recall	98	92	67	93.22
F1-Score	98	92	65	89.89

The CNN-LSTM model outperformed the other models with an accuracy of 96.34%, proving to have superior precision (98%), recall (98%), and F1-score (98%). By comparison, the DT model's accuracy of 92.18% was quite good, albeit with somewhat lower precision (93%), recall (92%), and F1-score (92%). With an accuracy68%, precision66%, recall67%, and an F1-score65%, the KNN model considerably underperformed, suggesting it might not be the best choice for this task. The ANN model yielded a 90.28% accuracy rate, 87% precision, 93.22% recall, and an F1-score of 89.89%, indicating a moderate level of performance. All things considered, CNN-LSTM turned out to be the most successful model, with KNN performing the worst.

Conclusion and Future Scope

The COVID-19 pandemic has caused several issues in many areas of people's lives. ML and DL were brought forward as analytical tools to address pandemic-related problems. The work aims to predict whether a person has Covid-19 and increase awareness about their disease to help prevent its spread. The dataset includes symptoms, test dates and results, gender, age, and other important information. Data preparation methods include handling missing values, converting categorical characteristics, and employing SMOTE to balance uneven datasets. The CNN-LSTM model with SMOTE has the greatest classification accuracy and F1 score with 96% and 98% performance. After that, the LIME framework analyses prediction results using explainable AI. The CNN-LSTM model outperformed all measures, suggesting it might improve public health diagnostics and prognostics. Big data analytics and machine learning enable real-time illness trend monitoring, prediction, resource allocation, and pandemic public health responses.

The use of an open-source dataset consisting of patient data from a particular location is an apparent drawback of this work. One potential option for future research is to use a private dataset that includes more biomarkers that cover a wider range of characteristics and areas. Combining ML models with fuzzy logic frameworks and meta learning approaches may increase prediction accuracy. Improving the system's performance is possible via the use of feature selection using the wrapper approach. Research may enhance model interpretability and scalability for larger datasets and more complex pandemics. Social media and environmental data may enhance prediction systems. Ensemble learning and hybrid models may improve population and geographic forecast accuracy and robustness. Prediction models must be updated and modified to epidemiological data to be relevant in changing public health circumstances

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 14468 – 14477
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.6775

References

[1] C. J. Wang, C. Y. Ng, and R. H. Brook, "Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing," 2020. doi: 10.1001/jama.2020.3151.

[2] Republic Of Korea AI Strategy, "'Toward AI World Leader, beyond IT' National Strategy for Artificial intelligence," 2019.

[3] J. Cummings, G. Lee, A. Ritter, M. Sabbagh, and K. Zhong, "Alzheimer's disease drug development pipeline: 2019," Alzheimer's Dement. Transl. Res. Clin. Interv., 2019, doi: 10.1016/j.trci.2019.05.008.

[4] R. Xu, Y. Chen, E. Blasch, A. Aved, G. Chen, and D. Shen, "Hybrid blockchain-enabled secure microservices fabric for decentralized multi-domain avionics systems," 2020. doi: 10.1117/12.2559036.

[5] V. S. Tseng, J. Jia-Ching Ying, S. T. C. Wong, D. J. Cook, and J. Liu, "Computational Intelligence Techniques for Combating COVID-19: A Survey," 2020. doi: 10.1109/MCI.2020.3019873.

[6] F. Jiang, L. Deng, L. Zhang, Y. Cai, C. W. Cheung, and Z. Xia, "Review of the Clinical Characteristics of Coronavirus Disease 2019 (COVID-19)," 2020. doi: 10.1007/s11606-020-05762-w.

[7] M. S. Islam, N. Ab Ghani, and M. M. Ahmed, "A review on recent advances in deep learning for sentiment analysis: Performances, challenges and limitations," Compusoft, 2020.

[8] D. Gupta et al., "Optimized cuttlefish algorithm for diagnosis of Parkinson's disease," Cogn. Syst. Res., 2018, doi: 10.1016/j.cogsys.2018.06.006.

[9] R. M. Carrillo-Larco and M. Castillo-Cara, "Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: An unsupervised machine learning approach," Wellcome Open Res., 2020, doi: 10.12688/wellcomeopenres.15819.3.

[10] M. Maleki, M. R. Mahmoudi, M. H. Heydari, and K. H. Pho, "Modeling and forecasting the spread and death rate of coronavirus (COVID-19) in the world using time series models," Chaos, Solitons and Fractals, 2020, doi: 10.1016/j.chaos.2020.110151.

[11] Z. Zhao et al., "Prediction model and risk scores of ICU admission and mortality in COVID-19," PLoS One, 2020, doi: 10.1371/journal.pone.0236618.

[12] P. Nancy, S. Sridhar, R. Akiladevi, and V. Sudha, "Exploration on covid-19 data in india using machine learning for prediction of infected and death cases," Adv. Math. Sci. J., 2020, doi: 10.37418/amsj.9.7.26.

[13] L. Yan et al., "An interpretable mortality prediction model for COVID-19 patients," Nat. Mach. Intell., vol. 2, no. 5, pp. 283–288, 2020, doi: 10.1038/s42256-020-0180-7.

[14] Y. S. Jeong and J. H. Park, "Advanced big data analysis, artificial intelligence & communication systems," J. Inf. Process. Syst., 2019, doi: 10.3745/JIPS.02.0107.

[15] A. S. Hussein, T. Li, C. W. Yohannese, and K. Bashir, "A-SMOTE: A new preprocessing approach for highly imbalanced datasets by improving SMOTE," Int. J. Comput. Intell. Syst., 2019, doi: 10.2991/ijcis.d.191114.002.

[16] B. Kamiński, M. Jakubczyk, and P. Szufel, "A framework for sensitivity analysis of decision trees," Cent. Eur. J. Oper. Res., 2018, doi: 10.1007/s10100-017-0479-6.

[17] T. B. Plante et al., "Development and external validation of a machine learning tool to rule out COVID-19 among adults in the emergency department using routine blood tests: A large, multicenter, real-world study," J. Med. Internet Res., 2020, doi: 10.2196/24048.

[18] S. N. Malkanthi, N. Balthazaar, and A. A. D. A. J. Perera, "Lime stabilization for compressed stabilized earth blocks with reduced clay and silt," Case Stud. Constr. Mater., 2020, doi: 10.1016/j.cscm.2019.e00326.

[19] C. Hu et al., "Early prediction of mortality risk among patients with severe COVID-19, using machine learning," Int. J. Epidemiol., 2020, doi: 10.1093/ije/dyaa171.

[20] J. Zhou et al., "Clinical features predicting mortality risk in older patients with COVID-19," Curr. Med. Res. Opin., 2020, doi: 10.1080/03007995.2020.1825365.