iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** III **Month of publication:** March 2023

DOI: https://doi.org/10.22214/ijraset.2023.49750

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

Prediction of ICU Admission for Covid-19 Patients: A Machine Learning Approach Based on Complete Blood Count Data

Dr. J. Mary Dallfin Bruxella[1], T.S. Linda Mary Raint[2]
[1]*Assistant Professor,* [2]*M Sc Data Science, Department of Computer Science and Information Technology, KARE, Krishnankovil, Tamil Nadu*

Abstract: ***In this post, we talk about how to create prognostic Machine Learning (ML) models for COVID-19 progression. Specifically, we talk about how to forecast who willneed to be admitted to the intensive care unit (ICU) in the next five days.On the basis of 4995 Complete Blood Count (CBC) tests, we created three ML models.We propose three ML models that differ in terms of interpretability: two fully interpretable models and a blackbox one.We report an AUC of .81 and.83 for the interpretable models (the decision tree and logistic regression, respectively), and an AUC of .88 for the black-box model (an ensemble).This demonstrates how CBC data and machine learning techniques can be used to predict the cost-effectiveness of ICU admission for COVID-19 patients.In particular, because the CBC can be quickly obtained through routine blood tests, our models could be used in settings with limited resources and to quickly provide indications at triage and daily rounds.***
Keywords: Decision Tree, Support Vector Machine (SVM), K-Nearest Neighbors Algorithm (KNN), Machine Learning

I. INTRODUCTION

While positive outcomes for the diagnostic job [1]-[3] (i.e., the detection of COVID-19), the development of prognostic models, either to anticipate ICU admission or other outcomes (including mortality) or to stratify patients by risk, has so far trailed behind: Several investigations [4], [5] have discovered considerable bias or overfitting hazards in the current solutions. In order to overcome these limitations, the study describes a retrospective inquiry to develop prognostic Machine Learning (ML) models to forecast ICU admission, which may be seen as a proxy for sickness severity or a result of deteriorating circumstances. An important dataset of hematologic parameters was acquired from COVID-19 patients hospitalised to one of the largest teaching hospitals in Lombardy (Northern Italy), one of the most severely impacted regions of Italy. During the initial wave of the pandemic, Lombardy (Northern Italy), one of the most severely affected regions, collected a significant dataset of hematologic parameters from COVID-19 patients hospitalised at one of the major teaching hospitals. To be more specific, one of the most reliable datasets that has been made available for COVID-19 analysis thus far was used and processed by us [1]. We utilised this dataset to analyse COVID-19 because we were inspired by the encouraging results showing a significant correlation between blood test results and COVID-19 prognosis [6], [7]. We chose a small subset of characteristics from this dataset to establish the so-called Complete Blood Count (CBC), a simple, inexpensive blood test with a number of diagnostic and monitoring uses. To the best of our knowledge, this is the first attempt to do COVID-19 prognosis based just on CBC results using ML algorithms. To do this, we offer three models that were developed as extra decision assistance tools. Notwithstanding the black-box nature of the model's design and its low accuracy as a result of the combination of three models, it was chosen. One model has been chosen despite its poor clinical interpretability due to its black-box nature and low accuracy due to the assembling of three models. The other two models, a decision tree and a logistic regression, have been chosen for their explainability even though they are less accurate than the aforementioned ensemble model. In fact, these models can give doctors clearer cues to help them make decisions when managing and treating COVID-19

Architecture

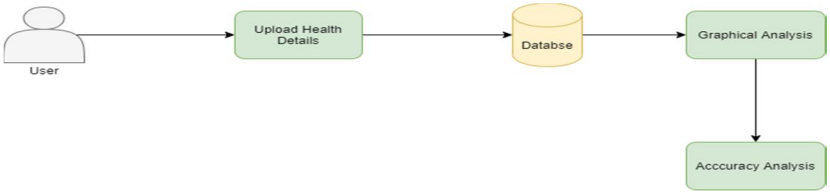


Fig 1: Architecture

©IJRASET: All Rights are Reserved

II. LITERATURE REVIEW

A. Machine Learning Models for Covid-19 Identification Based on Common Blood Tests are Created, Assessed, and Validated

Objectives The rRT-PCR test, the current industry standard for coronavirus illness detection (COVID-19), has a number of well-known drawbacks, including a lengthy turnaround time, a potential reagent shortage, false-negative rates of roughly 15-20%, and expensive equipment.Regular blood tests' hematochemical results might be a quicker and less expensive substitute.Methods The complete OSR dataset (72 features:complete blood count (CBC), biochemical, coagulation, hemogasanalysis, and CO-Oxymetry values, age, sex, and specific symptoms at triage) and two sub-datasets were used to develop machine learning (ML) models using hematochemical values from 1,624 patients (52% COVID-19 positive), admitted at San Raphael Hospital (OSR) from February to May 2020. COVID-specific and CBC dataset, 32 and 21 features respectively). For internal-external and external validation, 58 cases (50% COVID-19 positive) from another hospital and 54 negative patients gathered in 2018 at OSR were employed.Results The area under the receiver operating characteristic curve (AUC) for the algorithms ranged from 0.83 to 0.90 for the whole OSR dataset, from 0.83 to 0.87 for the COVID-specific dataset, and from 0.74 to 0.86 for the CBC dataset.The validations likewise produced positive outcomes, with AUC increasing from 0.75 to 0.78 and specificity increasing from 0.92 to 0.96.Conclusions ML can be used in blood tests as an addition or alternative to rRT-PCR for the quick and affordable identification of patients who are COVID-19 positive.This is especially helpful in underdeveloped nations or in nations where the spread of infectious diseases is on the rise.

B. Using Radiographic Characteristics To Diagnose With Covid-19: Problems and Prospects

The entire global medical sector is undercut by the coronavirus disease-2019 (COVID-19explosive)'s growth and widespread transmission.While 84 000 additional cases were verified on April 14, 2020, the limited medical resources globally impair diagnostic ability.The current gold standard for diagnosis is real-time reverse-transcription polymerase chain reaction (RT-PCR), however the false-negative rate is still a cause for worry.For initial screening and follow-up, radiographic technologies and tools, such as computed tomography (CT) and chest X-rays, were used.From these, the tools give a detailed diagnosis with specific pathologic findings for staging and treatment planning.

Despite the fact that radiographic imaging is thought to be less sensitive, several CT-positive patients were initially not ruled out by RT-PCR and later proven to be positive for COVID-19.Also, due to logistical challenges and the stress of providing healthcare, certain regions have reported a shortage of sampling kits and a delay in the turnaround time of PCR tests.In order to protect lives despite the crisis, we will talk about the difficulties and potential outcomes of using radiographic modalities for COVID-19 diagnosis in this review.

C. Machine Learning Is Used To Predict Sars-Cov-2infection In Routine Laboratory Blood Tests

There is an urgent need for accurate diagnostic methods to quickly identify SARS-CoV-2 positive persons for patient care management and staff protection. The most common diagnostic procedure is viral RNA detection by RT-PCR from nasopharyngeal swab specimens, although not all patient care facilities can provide results right once. Contrarily, routine laboratory testing is easily accessible and has a turn-around time (TAT) of typically 1-2 hours.

D. Asking The Right Questions For Covid-19 Machine Learning

The novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-), which caused COVID-19, has generated a crisis in the worl healthcare systems. In some nations, the rate at which medical resources are being used up has outpaced the availability of personal protective equipment and ventilators, the latter of which is now more necessary than ever because severe disease is characterised by life-threatening respiratory failure.

III. METHODOLOGY

A. Existing Methodology

More precisely, we processed and used one of the most trustworthy datasets for COVID-19 analysis to date [1] (which is available on the European open-access repository Zenodo1), driven by the encouraging findings of the substantial correlation between blood test data and COVID-19 prediction [6], [7].We took a small subset of attributes from this dataset to describe the so-called Complete Blood Count (CBC), a simple and low-cost common blood test that has a wide range of diagnostic and monitoring uses.

To the best of our knowledge, this is the first work to perform COVID-19 prognosis only based on CBC parameters usingML algorithms.

©IJRASET: All Rights are Reserved

We offer three models that have been designed as supplementary decision assistance tools in order to achieve this goal.Despite its limited clinical interpretability due to its black-box character and low accuracy due to the assembly of three models, one model has been chosen.While being less accurate than the previously described ensemble model, the other two models—a decision tree and a logistic regression—were chosen for their explainability.Certainly, these models can give doctors more easily interpreted information that will aid them in making decisions as they monitor and treat COVID-19 patients.

B. Proposed Methodology

As a concluding observation, we highlight a few significant general discrepancies between the suggested strategy and the works under consideration.First, regardless of the length of the hospital stay, all discussed models take the task of severity (either death or/and ICU admission) prediction into consideration.A case is deemed severe if any severe adverse outcome happens during the hospital stay.

While using this strategy might assist lessen the data imbalance, there is a chance that it will ignore significant confounding factors, such as therapy, or that it will require data that to be available right away after admission.In our method, we investigated prediction on a fixed 5-day horizon, which is still clinically significant and potentially valuable, in order to lessen theimpact of these confounding factors.

The fact that our suggested method is exclusively dependent on CBC data is a big benefit for the following reasons.CBC can be obtained during standard checkups; CBC data can be obtained quickly and inexpensively in comparison to that of other, more specialised biomarkers; CBC is less impacted by preanalytical (how specimens are collected, handled, and identified), analytical (which regards differences inthe testing methods in different laboratories or with different equipment), and biological variability (which is related to the fluctuations of biomarkers along patient's life) factors than other exams that are related to clinical chemistry, inflammatory markers, or coagulation parameters.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. System Modules

1) Dataset: 50 COVID-19 patients chest X-ray photos have been downloaded from the GitHub open source repository. This archive includes chest X-ray and CT images ofpatients with pneumonia, severe acute respiratory syndrome, COVID-19, Middle East respiratory syndrome, and acute respiratory distress syndrome (SARS).

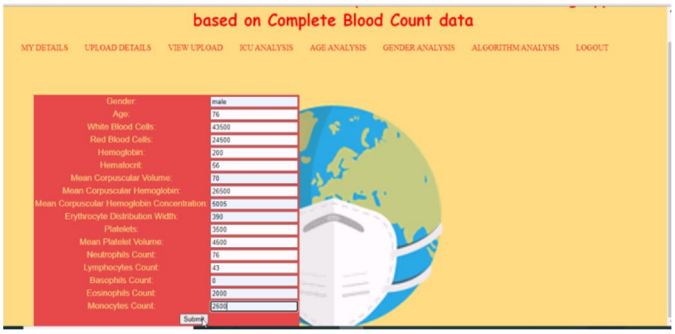


Fig 2: Patient Data

2) Training Models

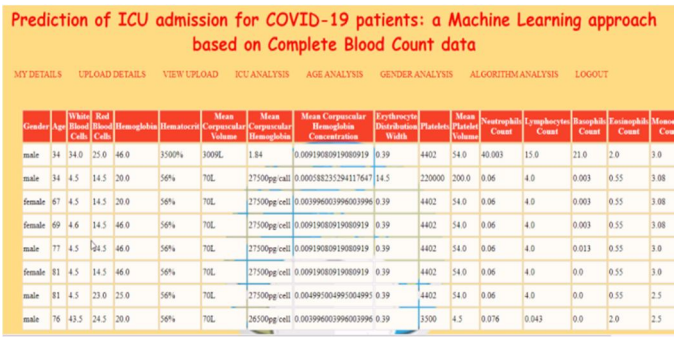


Fig 3: View Data

The classification, segmentation, and lesion identification of medical data are just a few applications where deep learning models have been successfully applied. Usingdeep learning models, image and signal data from medical imaging modalities such as computed tomography (CT), X-ray, and magnetic resonance imaging (MRI) are analysed.

3) *Training Phase:* Coronavirus illness patients have been predicted using chest X-ray pictures (COVID-19). Using chest X-ray images, popular pre-trained models like ResNet50, InceptionV3, and Inception ResNetV2 have been trained and tested.Training precision and loss metrics for the pre-trained models fold-3. Exact values that were projected. The accuracy, precision-recall trade-off, and AUC of the model were used to assess its performance after grouping all predicted values into the matrix.

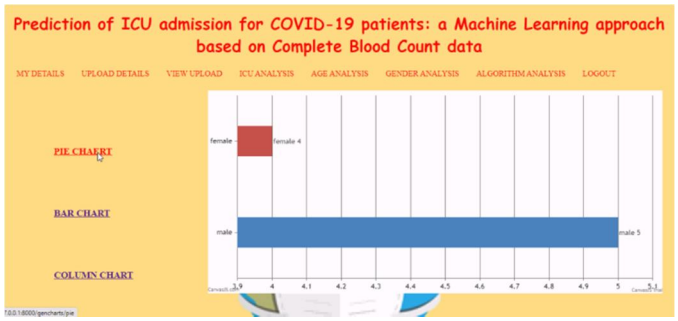


Fig 4: Gender Analysis

V. CONCLUSION

For the purpose of tackling the difficult task of determining if a COVID-19 patient will need to be transferred to the ICU within the next five days while they are in the hospital, we reported a retrospective study.Results from the suggested method, which was based on both interpretable and black-box models, were positive.Our techniques are also economical because they just rely on two demographic factors and the results of the CBC test, which is their greatest advantage in terms of acceptable accuracy.The execution of more COVID-specific tests (e.g., inflammatory markers, interleukins, and coagulation parameters [34]) on a regular basis is therefore not feasible in resource-limited settings, such as healthcare facilities managing a spike in unwell patients.In our upcoming study, we intend to externally evaluate our models using information from various hospitals and eras:As a result, the model may be tested in light of potential virus alterations as well as various patient treatment and therapy approaches.The latter ones cannot be ruled out in any existing predictivemodel, including ours, because they depend on the quantity of cases to be handled as well as the ongoing progress of knowledge regarding COVID-19 and its successful treatment (changing its prognosis).

REFERENCES

[1] F. Cabitza, A. Campagner, D. Ferrari, C. Di Resta, D. Ceriotti, E. Sabetta, A. Colombini, E. De Vecchi, G. Banfi, M. Locatelli et al. “Development, evaluation, and validation of machine learning models for covid-19 detection based on routine blood tests,” Clinical Chemistry and Laboratory Medicine (CCLM), vol. 59, no. 2, 2021.

[2] S.-G. Chen, J.-Y. Chen, Y.-P. Yang, C.-S. Chien, M.-L. Wang, and L.- T. Lin, “Use of radiographic features in covid-19 diagnosis: Challenges and perspectives,” Journal of the Chinese Medical Association, vol. 83, no. 7, p. 644, 2020.

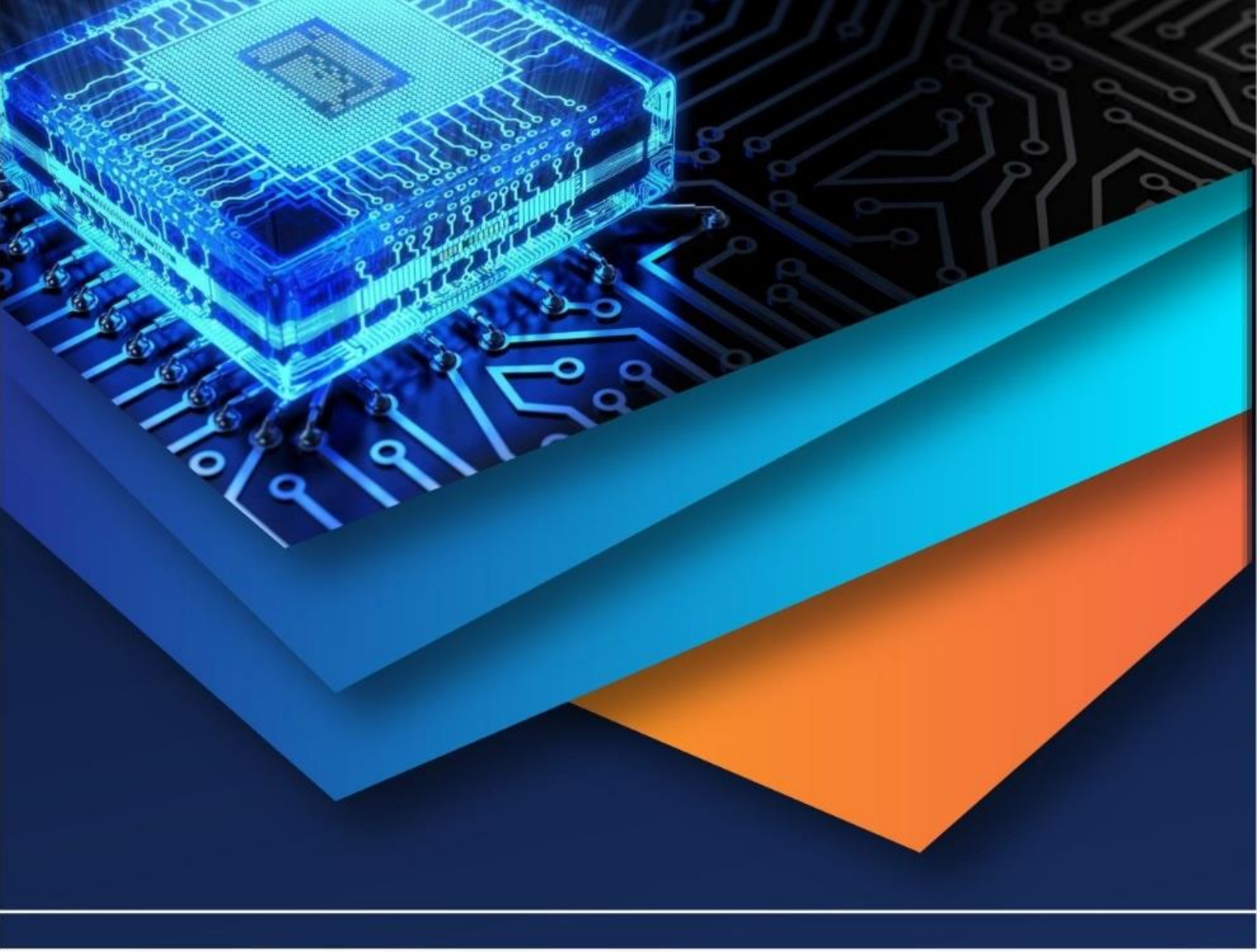
[3] H.S. Yang, Y. Hou, L. V. Vasovic, P. A. Steel, A. Chadburn, S. E. Racine-Brzostek, P. Velu, M. M. Cushing, M. Loda, R. Kaushal et al., “Routine laboratory blood tests predict sars-cov-2 infection using machine learning,” Clinical chemistry, vol. 66, no. 11, pp. 1396–1404, 2020.

[4] P. Bachtiger, N. S. Peters, and S. L. Walsh, “Machine learning for covid- 19—asking the right questions,” The Lancet Digital Health, vol. 2, no. 8, pp. e391–e392, 2020.

[5] L. Wynants, B. Van Calster, G. S. Collins, R. D. Riley, G. Heinze, E. Schuit, M. M. Bonten, D. L. Dahly, J. A. Damen, T. P. Debray et al., “Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal,” bmj, vol. 369, 2020.

[6] E. J. Favaloro and G. Lippi, “Recommendations for minimal laboratory testing panels in patients with covid-19: Potential for prognostic monitoring.” in Seminars in Thrombosis and Hemostasis, vol. 46, 2020, pp. 379–382.

[7] J. Linssen, A. Ermens, M. Berrevoets, M. Seghezzi, G. Previtali, H. Russcher, A. Verbon, J. Gillis, J. Riedl, E. de Jongh et al., “A novel haemocytometric covid-19 prognostic score developed and validated in an observational multicentre european hospital-based study,” Elife, vol. 9, p. e63195, 2020.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)