

A Review on COVID Mortality Prediction with Machine Learning Methods

Revathy V. S.¹, Sandra Sunny², Dr. Neetha Thomas³
Santhigiri College of Computer Sciences, Vazhithala, Kerala^{1,2,3}

Abstract: *More than a year has passed since the report of the first case of coronavirus disorder 2019 (COVID), and increasing deaths hold to occur. Minimizing the time required for useful resource allocation and medical choice making, together with triage, desire of air flow modes and admission to the intensive care unit is essential. system learning strategies are acquiring an increasingly more sought-after role in predicting the outcome of COVID sufferers. in particular, the use of baseline gadget learning techniques is swiftly growing in COVID mortality prediction, in view that a mortality prediction model could unexpectedly and effectively help medical choice-making for COVID sufferers at approaching threat of demise. latest research reviewed predictive fashions for SARS-CoV-2 prognosis, severity, period of sanatorium live, extensive care unit admission or mechanical ventilation modes results; however, systematic opinions centered on prediction of COVID mortality outcome with machine gaining knowledge of methods are missing within the literature. the present evaluation appeared into the studies that carried out gadget mastering, inclusive of deep studying, methods in COVID mortality prediction as a result trying to gift the prevailing posted literature and to offer feasible causes of the pleasant effects that the research received. The have a look at also mentioned hard components of cutting-edge research, supply in guidelines for destiny developments.*

Keywords: Covid, Machine Learning, Deep Learning, CNN, Imaging, Mortality Prediction, etc.

I. INTRODUCTION

Coronavirus Disease (Covid-19) Is an Infectious Disease Caused by The Sars-Cov-2 Virus. Most People Infected with The Virus Will Experience Mild to Moderate Respiratory Illness and Recover Without Requiring Special Treatment. Despite the invention of various vaccine formulas from completely different pharmaceutical corporations, several issues associated with mass production and distribution across the globe still persist. This issue is attended by political and economic constraints that will more limit immunizing agent access. For these reasons, pandemic containment may be an onerous task, leading to increased deaths. At the time this manuscript is written, SARS-CoV-2 numbers according by the globe Health Organization worldwide include: nearly 173,005,553 folks infected with SARS-CoV-2; over three,727,605 death cases and around one,900,955,505 immunizing agent doses administered.

The aim of this review is the ways to predict COVID mortality by: (1) summarizing the present printed literature on baseline ML- and DL-based COVID mortality prognosis systems supported medical evaluations, laboratory exams and pc imaging (CT); (2) presenting relevant info including the kind of information used, the information ripping technique, the projected cc methodology and analysis metrics; (3) providing potential explanations of the simplest results obtained; (4) discussing difficult aspects of current studies, providing suggestions for future developments.

II. LITERATURE REVIEW METHODS

This systematic overview considers the ML and DL implemented to COVID mortality prediction. We carried out a we strive to shed a few mild on extraordinary traits of these studies in phrases of:

- (i) records supply,
- (ii) elegance of functions,

- (iii) carried out functions ranking method,
- (iv) applied ML technique,
- (v) metrics evaluated for performance assessment.

1. Data Source

We emphasized the study location and whether the dataset of each study was public or private, single site or multi center.

2. Data Partitioning

We centered on the kind of model validation that every study would split data into train and take a look at tests notably we have a tendency to selected to report the quantity of subjects used for the train and take a look at set, and therefore the corresponding range of survived and non-survived subjects.

3. Class of Features

We anticipated to gather papers with each clinical and imaging features. we included hand-made extracted capabilities with radiomic evaluation and the features found out with using convolutional neural networks (CNN). clinical capabilities comprise demographic (e.g., age, sex, race), comorbidities (e.g., diabetes, heart disorder), symptoms (e.g., cough, fever), important signs (e.g., coronary heart charge, oxygen saturation), laboratory values (e.g., glucose, creatinine, hemoglobin), ailment treatment and scientific course (e.g., artificial ventilation, duration in health center, tablets). clinical features can be classified in binary and continuous functions (numerical values).

4. Implemented Features Ranking Method

To build a reliable model for fixing type, the feature set ought to incorporate as a good deal beneficial records as feasible, and some of features as small as possible. it is important to filter out the irrelevant and redundant functions with the aid of deciding on a subset of relevant functions to keep away from over-fitting and address the problem of dimensionality. feature ranking (or selection or discount) strategies are a great approach for features space dimensionality reduction. Feature ranking improves features understanding and reduces the computational cost, increasing the efficiency of the classification. SHAP and LASSO algorithm are widely used methods for feature selection in survival studies.

5. Implemented Machine Learning(ML) Techniques

With the aim of distinctive we tend to targeted on the prediction technique utilized in every work. Since in literature there are several implementable and customizable algorithms, we tend to expected to search out many and totally different ways utilized within the works enclosed during this review. However, we tend to expected to search out techniques as a result of one of the subsequent four categories, in line with the characterized basic algorithm: (I) regression (ii) classifier (iii) neural network and (iv) ensemble

III. LITERATURE REVIEW RESULTS

3.1. Data Source

Table 1. Studies Data Sources, Samples and Validation Characteristics.

Centers	Location	Survived and Non-Survived Sample Size	Type of Validation	Sample Size Train	Sample Size Test	Online Available Dataset
Five centers	United states	2392 survived, 1323 non survived (deaths up to 3 gg:140; 5 gg: 281; 7 gg: 393; 10 gg: 509)	Internal, External, Internal Prospectively, Prospectively Merged	1514 (deaths up to 3 gg: 40; 5 gg: 74; 7 gg: 112; 10 gg: 182) (center1)	external: 2201 (deaths up to 3 gg: 135; 5 gg: 276; 7 gg: 382; 10 gg: 494) (center 2, 3, 4, 5; time1) prospectively merge; 383 (deaths up to 3 gg: 3; 5 gg: 5; 7 gg: 11; 10 gg: 15) (all five center merged; time2)	NO
Single center	United Kingdom	275 survived, 93 non survived	Internal	318	80	NO
Single center	United Kingdom	275 survived, 93 non survived	Internal	318	80	NO
Four centers	Korea	299 survived, 214 non survived	Internal, External	361 (195 survived, 212 non survived) (center1)	external: 106 (survived 104, non survived 2) (center2,3,4)	NO
Thirty centers	Italy	3182 survived, 712 non survived, 41.5% of whom resident in Central/Southern Italian regions (15.6% death north italy; 6.4% center-south);	Merged	2725 (all thirty centers merged)	1169 (all thirty centers merged)	NO
Single center	United states	355 survived, 43 non survived	Internal	318	80	NO
Two multicentric dataset open source	China	Cohort1: 28428 survived, 530 non survived; Cohort2: 1325 survived, 123 non survived	Two double merged	Training (1): 8687 (chort1); training (2): 434 (chort2)	double merge (1): first 14190, secondly 6081 (cohort1), double merge (2): first 710, secondly 304 (chort2)	YES
Single center	China	2737 survived, 259 non survived	Internal	2339 (center1)	internal: 585 (center1); external: 72 (70 survived, 2 non survived) (center2)	NO
Single center	China	142 survived; 39 non survived	Internal	154	27	NO
Two multicentric dataset open source	China	662 survived, 57 non survived (169933 slices)	Merged			YES
Single center	United states	2985 survived, 505 non survived	Internal	2793	698	NO
Thirty three centers	Italy, Spain, United States	2302 survived; 760 non survived	Merged, Three External	2755 (25 centers merged)	merge: 760 (1:25 centers), external (1): 323 (center26), external (2): 219 (27:32 centers), external (3): 323 (center35)	NO
Multicentric database	Korea	7772 survived, 228 non survived the dataset was splitted according to the ratio 7:3	Merged	5600	2400	NO
Two centers	China	1198 survived; 72 non survived	Internal, External	554 (513 survival, 41 non survival) (center1)	Internal: 233 (217 survival, 16 non survival) (center1) External: 286 (279 survival, 7 non survival) (center2)	NO
Five centers	United states	35:9 survived; 510 non survived	Five Internal	Training (1): 463 (center1), training (2): 1151 (center2), training (3): 524 (center3), training (4): 378 (center4), training (5): 340 (center2)	Internal (1): 148 (center1), internal (2): 493 (center2), internal (3): 225 (center3), internal (4): 162 (center4), internal (5): 145 (center2)	NO
Two centers	China	148 survived, 99 non survived	Internal, External	183 (115 survived, 68 non survived) (center1)	external: 64 (33 survived, 31 non survived) (center2)	NO
Single center	China	(1) 288 Survived, 187 non survived, (2) 189 survived, 162 non survived	Internal Prospectively Internal	Training (1): 375 (time1), training (2): 246	internal prospectively (1): 110 (time2), internal (2): 105	NO
Single center	United states	3226 survived, 1087 non survived	Internal	3468	845	NO
Two centers	Iran, United States	193 patients	External	105 (center1)	88 (center1)	NO
Multicentric database	United States	5308 patients	Internal Merged Prospectively	3597 (2909 survived, 688 non survived)	1711	Researcher affiliated with Mass General Brigham may apply for access
Multicentric database	United States	648 survived, 87 non survived	Merged			NO
Single center	Italy	266 survived, 75 non survived	Internal			NO

In table 1 we show the type of model validation that every have a look at used to split information into educate and take a look at agencies, indicating the number of topics and the corresponding variety of survived and non-survived topics. inner validation changed into performed in 15/24 studies. A total of 7/24 studies performed external validation, particularly 4 of these combined data from different sites producing a single database and 3 of these used multisite publicly available epidemiological datasets. A total of 2/24 studies implemented internal prospective validation and 3/24 studies implemented a prospective merged validation.

Features Type			Features Class					Features Selection			Dimension Reduction	
Binaries	Continuous	Images	Demographics	Comorbities	Syntoms	Vital Signs	Laboratory	Treatment	SHAP	LASSO		Others
Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	73 to 10
Yes	No	No	Yes	Yes	Yes	No	No	Yes	Yes	No	No	22
Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	73 to 30
Yes	No	No	Yes	Yes	No	No	Yes	No	No	No	Yes	12
No	Yes	No	No	No	No	No	Yes	No	Yes	No	Yes	26 to 5
Yes	No	No	Yes	Yes	Yes	No	No	Yes	No	No	No	No
Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	1224 to 83
Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	56 to 5
Yes	Yes	Yes	No	No	No	No	Yes	No	No	No	No	No
Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No	No	34 to 20

Features Type			Features Class					Features Selection			Dimension Reduction	
Binaries	Continuous	Images	Demographics	Comorbities	Syntoms	Vital Signs	Laboratory	Treatment	SHAP	LASSO		Others
Yes	Yes	No	Yes	Yes	No	Yes	No	No	No	Yes	No	34 to 8
Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	No	No	22
Yes	No	No	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Not specified
Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No	48 to 6
Yes	No	No	Yes	Yes	No	No	No	No	No	No	No	No
Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	20 to 4
Yes	Yes	No	Yes	No	Yes	No	Yes	No	No	No	Yes	75 to 3
No	Yes	No	Yes	Yes	No	No	No	Yes	Yes	No	No	48 to 10
No	No	Yes	No	No	No	No	No	No	No	No	No	No
Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	Not specified
Yes	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	No	No	109 to 10
Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	No	No	Yes	Not specified
Yes	Yes	No	Yes	No	No	Yes	Yes	No	No	No	Yes	142

Machine Learning Technique	Metrics										
	Accuracy	AUC-ROC	AU-PRC	Sensitivity	Specificity	PPV	NPV	F1-Score	MCC	Balanced Accuracy	k-Fold
Ensemble (XGBoost)	int val (3 days): 97.6%; ext val (3 days): 93.6%; int val prosp (3 days): 97.1%; merged val prosp (3 days): 94.2%	int val (3 days): 89%, ext val (3 days): 87.7%; int val prosp (3 days): 96.2%; merged val prosp (3 days): 87.9%	int val (3 days): 44.5%, ext val (3 days): 44.4% int val prosp (3 days): 55.1%; merged val prosp (3 days): 13.1%	int val (3 days): 44.2%, ext val (3 days): 37% int val prosp (3 days): 50%; merged val prosp (3 days): 33.3%	int val (3 days): 99.1%, ext val (3 days): 93.6%, int val prosp (3 days): 97.1%; merged val prosp (3 days): 94.2%	No	No	int val (3 days): 49.8%; ext val (3 days): 41.7% int val prosp (3 days): 28.6%; merged val prosp (3 days): 14.3%	No	No	10
ANN	int val: 86.25%	int val: 90.12%	No	int val: 87.5%	int val: 85.9%	int val: 60.87%	int val: 96.49%	No	No	No	10
Cox Regressor	int val: 83.75%	int val: 86.9%	No	int val: 50%	int val: 96.6%	int val: 84.6%	int val: 83.6%	No	No	No	10
Ensemble (DNN + RF)	int val: 93% ext val: 92%	No	No	int val: 92% ext val: 100%	int val: 93%, ext val: 100%	No	No	No	No	int val: 93%, ext val: 96%	100
Ensemble (RF)	merged val: 83.4%	No	No	merged val: 95.2%	merged val: 30.8%	No	No	merged val: 90.4%	No	No	10
SVM	No	int val: 93%	int val: 76%	int val: 91%	int val: 91%	No	No	No	No	No	Unclear
Auto encoder	No	No	No	No	No	No	No	No	No	No	Unclear
Ensemble (GBDT)	int val (severe): 88.9%, int val (non-severe): 92.4%, int val(total): 79.9%	int val (severe): 94.1%, int val (non-severe): 93.2%, int val(total): 91.8%	No	int val (severe): 89.9%, int val (non-severe): 94% int val(total): 77.4%	int val (severe): 78.8%, int val (non-severe): 61.9%, int val(total): 90.3%	int val (severe): 43.2%, int val (non-severe): 35.1%, int val(total): 48.3%	int val (severe): 97.8%, int val (non-severe): 97.9%, int val(total): 97.2%	No	No	No	5
DNN	No	int val: 96.8%	No	No	No	No	No	No	No	No	5
PLS	merged val: 78.73%	merged val: 85.6%	No	merged val: 88.24%	merged val: 78.26%	merged val: 16.67%	merged val: 99.26%	No	merged val: 52.36%	No	10
Ensemble (CatBoost)	int val: 80.3%	int val: 85%	No	No	No	int val: 79%	int val: 81.6%	No	No	No	Unclear
Ensemble (LR, SVM, GBDT, NN)	int val: 96.21%, ext val 1: 97.6% ext val 2: 92.46%	int val: 92.4%, ext val 1: 95.5%, ext val 2: 87.9%	No	No	No	No	No	No	No	No	10
Ensemble (XGBoost)	merged val: 85.02%, ext val 1: 74.92%, ext val 2: 286.76%, ext val 3: 61.3%	merged val: 90.19%, ext val 1: 87.45%, ext val 2: 91.62%, ext val 3: 80.66%	No	No	merged val: 86.58%, ext val 1: 74.23%, ext val 2: 87.43%, ext val 3: 58.12%	merged val: 66.3%, ext val 1: 25.74%, ext val 2: 48.94%, ext val 3: 24.18%	merged val: 93.02%, ext val 1: 97.3%, ext val 2: 97.09%, ext val 3: 94.71%	No	No	No	Unclear
SVM	No	merged val: 96.2%	No	merged val: 92.0%	merged val: 91.8%	merged val: 25.7%	merged val: 99.7%	No	No	merged val: 91.9%	10
Ensemble (XGBoost)	int val: 99.1% ext val: 99.7%	No	No	int val: 87.5% ext val: 85.7%	No	int val: 99.1% ext val: 99.7%	No	No	No	No	500

Federate learning (MLP)	int val: 78%	int val: 83.6%	int val: 27.6%	int val: 80.5%	int val: 70.2%	No	No	int val: 32.8%	No	No	490
LR	No	int val: 89.5%, ext val: 88.1%	No	int val: 89.2% ext val: 83.9%	int val: 68.7% ext val: 79.1%	No	No	No	No	No	10
Ensemble (XGBoost)	No	No	No	int val: 95% ext val prosp: 98%	No	int val: 95% ext val prosp: 91%	No	int val: 95% ext val prosp: 94%	No	No	500
Ensemble XGBoost	No	int val: 90.3%	int val: 79.1%	int val: 83.8%	int val: 83.6%	int val: 60.9%	int val: 94.4%	No	No	No	10
LR	No	ext val: 73.6%	No	No	No	No	No	No	No	No	Unclear
Ensemble (RF)	No	No	No	No	No	No	No	int val: 87%	No	No	Unclear
LightGBM	No	merged val: 88%	No	No	No	No	No	No	No	No	10
Ensemble (RF)	No	int val: 84%	No	int val: 78.8%	int val: 77.4%	No	No	No	No	No	Unclear
Ensemble GBDT	merged val prosp: 96%	merged val prosp: 99%	No	merged val prosp: 24%	merged val prosp: 97%	merged val prosp: 90%	merged val prosp: 98%	No	No	No	Unclear

IV. DISCUSSION

Few studies tried COVID survival analysis with applied mathematical strategies. We set to focus our review on mortality prediction through Machine Learning (ML) techniques that are able to fit nonlinear and complicated interaction effects between predictors significantly, ML improved predictability compared to other Applied mathematical strategies on prediction of survival, in numerous sensible domains. Variability in dataset dimensions, experimental methods and options decisions limit the comparison of the chosen studies.

A. Datasets.

The studies enclosed during this review share many limitations. First, the quantity of patients offered for testing may well be thought of tiny, poignant the importance of the results. To boot, deceased cases area unit usually a minority compared to the those alive.

The few datasets that area unit publicly offered area unit subject to the doable risk of institution bias thanks to the shortage of data regarding exclusioncriteria.an extra bias may be related to the due to impossibility of knowing whether or not patients area unit SARS-CoV-2 really truly SARS-CoV-2 positive due to the unclear definition of patients accomplishment .In addition, most studies were blind to patients United Nations agency were admitted for clinically suspected SARS-CoV-2 and tested positive for the virus but died due to unrelated morbidities. Since imbalance problems characterize the SARS-CoV-2 mortality rate 3.6%. An unbalanced knowledge choice must completely or negatively have an effect on thetraining and testing method. So representative sample is needed for a stable model. Nevertheless, these sensible results may be useful because ofthe adopted methods (Neural networks, SVM, Ensemble algorithms) that are gloriousfrom previous literature to realize high performance on unbalanced datasets adjusted with oversampling or under sampling techniques.

B. Demography

Although the enforced models square measure representative of hospitalized patients with confirmed SARS-CoV-2 infection and relative outcome among the geographic remit of the study site, caution ought to be used oncegeneralizing to different populations. notably, results might not be generalized to populations with completely different geographical and socio-economic conditions, variations in national health service or insurance-based health expenses. A incorporate information and a prospective validation may well be helpful in an exceedingly target population. generalization. Furthermore, caution ought to be exercised in management practices changes or evolution of COVID pathological process.

C. Accuracy

Looking at the performance measures of the developed models, only a few achieved ACC > 90% on at least one validation technique. The highest accuracy for internal (99.1%) and external (99.7%) validation was achieved by Guan et al. with XG Boost. In terms of ML methods, ensemble learning was high performing (ACC > 90%) among studies. Moreover, studies that compared non-ensemble and ensemble learners showed best performance with ensemble models. moreover, ensemble models are less prone to overfitting issues compared to individual classifiers. Cross-validation is important to achieve higher accurate results with a limited amount of data.

D. Models Validation

External validation may be a rigorous key step before scattering the prediction model during a clinical setting. Since the aim of the according prognosticative models is to tell patients and carriers about a mortality outcome, it is essential that predictions ought to be well-calibrated on a target population. In this context, associate external validation might contribute to increase this target population and to generalize the model.

E. Clinical Features Predictive Ability

Sample size is regularly imbalanced with a relative minority of COVID positive mortalities, it is probably useful to create a international database for the generalization of consequences and the maximum important extracted functions, with a nicely-balanced variety of survivors and non-survivors. in the end, it is essential to notice that the SARS-CoV-2 pandemic is uncommon and evolving. therefore, a actual-time update of model prediction capabilities could be required.

F. Images Features Predictive Ability

Only 2 of the studies had info relating to radiologic pictures. Imaging might also be a vital prognostic issue. The results obtained from Ning et al. and Fang et al. in terms of accuracy (78% reportable by Ning et al.) and AUC-ROC (85% and event three.6% respectively) area unit worse than others that used clinical options solely. this might rely upon a lower range of deceased patients (8% of total subjects in Ning et al. and not specific in Fang et al.). Moreover, the mortality analysis of Ning et al. enclosed laboratory options and imaging options solely and Fang et al. enclosed solely score options associated with CT images.

On the opposite hand, Ning et al. reportable an honest MCC price (53%) that would be related to the combination of the metric capacity unit technique and pictures options. consistent with recent studies] chest CT pictures play a vital role within the designation, watching and severity stratification of COVID. However, further studies with a dataset of clinical options and pictures ought to be created to completely exploit the benefits of group action clinical and imaging options. though completely different studies used X-rays for predicting mortality, with each radiologist-assessed and AI-assessed disease severity scores, in our data, there aren't any studies that applied these predictors to millilitre ways within the analysis of mortality. additional studies might judge the quality of this application

V. CONCLUSION

In this review paper the authors have constructed the systematic review specially considers this state of art in ML & DL as applied to covid mortality prediction from the studies the author has found best practice for developing optimal machine learning models and he found that

1. The use of a high-quality data set with a large balanced number of samples.
2. The implementation of an ensemble of different ML methodologies.
3. Clinical features should include different features class type including Age, CRP, LDH values etc.

REFERENCES

- [1] S. M. Metev and V. P. Veiko, Laser Assisted Microtechnology, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

- [2] J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. *Lecture Notes in Statistics*. Berlin, Germany: Springer, 1989, vol. 61.
- [3] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," *IEEE Electron Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.
- [4] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, "High resolution fiber distributed measurements with coherent OFDR," in *Proc. ECOC'00*, 2000, paper 11.3.4, p. 109.
- [5] R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "High-speed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.
- [6] (2002) The IEEE website. [Online]. Available: <http://www.ieee.org/>
- [7] M. Shell. (2002) IEEEtran homepage on CTAN, www.ctan.org/tex-archive/macros/latex/contrib/supported
- [8] FLEXChip Signal Processor (MC68175/D), Motorola, 1996.
- [9] "PDCA12-70 data sheet," Opto Speed SA, Mezzovico, Switzerland.
- [10] A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP," M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.
- [11] J. Padhye, V. Firoiu, and D. Towsley, "A stochastic model of TCP Reno congestion avoidance and control," Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.
- [12] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997.