

Covid 19 Pandemic Review Using Machine Learning Algorithm

¹Roshan K. Ambatkar, ²Dr. Ashish B. Sasankar

^{*1}Research Scholar, Department of Electronics and Computer Science
Rastrasanta Tukdji Maharaj Nagpur University, Nagpur, Maharashtra, India

ambatkar.roshan@gmail.com

²Principal, Indraprasth New Arts, Commerce and Science College Wardha, Maharashtra, India

ashishdigital14@gmail.com

ABSTRACT

The around the world spread of the COVID-19 broad has actuated around the world endeavors to contain and supervise the disease. One significant locale of explore that has picked up basic thought is the utilize of machine learning (ML) to help recognize, predict, and fight COVID-19. ML has risen as a lively and rapidly progressing field in this setting. In show disdain toward of the reality that there's as of presently a significant whole of ask almost writing on the subject, the speedy advancement inside the number of dispersions and ML applications related to COVID-19 demands ceaseless study and evaluation. This article presents a comprehensive review of the first afterward considers approximately on ML calculations utilized in COVID-19-related errands. It centers on the potential comes about of ML in two essential ranges: conclusion and figure. Given the criticalness and earnestness of COVID-19 results, the capacity to create taught figures utilizing available clinical and examine data has gotten to be pivotal. To support this objective, this overview clarifies the sorts of calculations utilized, the datasets connected, and the sorts of assignments tended to. A number of crucial commitments have been recognized from considers conducted between January 2020 and January 2021. Most of the ML approaches utilized in these applications drop underneath the category of managed learning calculations. Be that because it may, in show disdain toward of promising disclosures, various proposed models have not be that as it may been executed in real-world clinical settings and remain at the exploratory organize. expressive and prognostic capabilities outlined by ML frameworks alter well with comes around point by point inside the supporting composing. By the by, a key obstacle in various current applications is the utilize of imbalanced datasets, which can show slant and impact the precision of the models.

Keyword: - Machine learning, COVID-19, Feature selection, Artificial intelligence, Diagnosis, Prognosis.

1 INTRODUCTION

The coronavirus may be a expansive RNA disease known to be display inside the mid-1960s and a short time later. You're reliable for causing coordinate afflictions from smooth ailments and colds to coordinate ailments. Two well-known coronaviruses are overpowering coronaviruses related with strongly clutter (SARS-COV) and Center Eastern coronavirus (MERS-COV). SARS-COV was recognized in 2003 when it to start with appeared up in Guangdong Zone, southern China. Mers-Cov came to Saudi Arabia in 2012. In December 2019, a present day coronavirus infection was detailed in a Chinese city in Uhan, Hubei. On January 7, 2020, the unused disease was recognized as Covid-19. Signs consolidate fever, dry hack, muscle torment, gastrointestinal side impacts and asymptoms. From December 2020 to Walk 2020, the world experienced a gigantic spread of Covid 19 illnesses, and the World Prosperity Organization (WH) clarified the far reaching. On January 22, 2021, it point by point over 96 million COVID claims and 2 million COVID 19 cases around the world

The around the world country has been influenced by the malady, driving to a gathering of measures, counting arrive closures, period squares and travel controls. Visit side impacts of Covid19 disease are as a run the appear smooth, but in numerous patients the defilement can cause legitimate to goodness, directly and after that unsafe complications. Overseeing different Covid 19 cases may be a major challenge that has overwhelmed prosperity work environments around the world. Be that since it may, there's not satisfactory data around the defilement be that because it may. Since the creation of the Covid 19 contamination, examiners from assorted regions have looked into the cutting edge illness. Machine learning (ML) may be a division of manufactured encounters (AI), centering on the period of frameworks that can be learned and moved forward from diagrams without being unequivocally modified. ML has been utilized in different locales, checking healthcare and medicinal computer science. ML employments key investigate introduction to initiate it and battle Covid-19. In affiliation with COVID-19, different inquire nearly boundaries have been shown for

the application and movement of ML calculations. A quick look for PubMed 94,609 disseminations related to Covid-19 from January 2021.

Different ponder work has been passed on to utilize ML in Covid-19 inquire around. Agbehadji et al. In rundown, we utilize methods to recognize and track COVID-19 utilizing tremendous information stages, AI models and nature-inspired calculations. Bullock et al. It clarifies how KI considers the challenges of Covid 19 with a combination of criteria, counting atomic, helpful and epidemiological applications. Naud highlighted the veritable potential applications of AI interior the battle against Covid-19. Different applications are being talked roughly, checking mishandle and want, affirmation and assess, treatment and immunizations, and social control. To anticipate and attest Covid-19, we conducted a ponder examination of the utilize of ML, noteworthy learning, coherent and quantifiable approaches. Lalmuanawma et al. We assessed existing ML improvements for Covid-19 screening, surveying, choosing, abuse and unfaltering movement. As of late, one of Tayarani-NS. In this way, AI approaches have been utilized to combat the distant coming to, checking clinical applications, Covid-199 imaging, pharmaceutical and epidemiological inquire roughly. A wide number of AI progresses have been looked into interior the composing, counting critical learning, ML, made neural systems (ANNs), and developmental glucose look at. Spread rundown for Covid-19 datasets is turned on. Wu et al. It has tended to the utilize of colossal information progression to anticipate and coordinate Covid-19 in China. Be that since it may, this field of inquire around is remarkably lively, and the number of related conveyances is quickly amplifying. At this point, different ML calculations are effectively open in information examination program such as Weka. This proposes it can be effectively utilized without specialized information.

The reason of this paper is to highlight the nonstop endeavors to utilize show up day ML calculations inside the center of this wide. The peculiarity of this think around is compared to as of by and by scattered arrange. Commonly a center on a really fundamental level open for ML illustrative and prescient models, and is vital inner parts an hour (internal parts an hour). These approaches discussion to speedier, cheaper symptomatic choices to switch comparable transcriptional polymerase chain reaction tests (RT-PCRs), in show up despise toward of decreased execution. When optimizing choices, prioritize the assignment of compelled resources at the beat, and the unfavorable result of truly care unit (ICU) supporting or passing is as of by and by uncommonly basic.

The main contributions of this work can be summarized as follows.

- (i) We review the recent ML algorithms in this field and focus on their potential in two main applications: diagnosis of COVID-19 and prediction of mortality risk and severity, using simple clinical and laboratory data.
- (ii) We analyze the main features that were found to be the most relevant to these applications.
- (iii) Open issues and future lines of research are highlighted based on the findings of this survey.

The rest of this paper comprises of the taking after methods: Locale 2 contains a brief depiction of numerous fundamental ML calculations for the preeminent afterward and most afterward craftsmanship. Part 4 looks at the preeminent afterward composing on the utilize of ML for the affirmation and appraise of the respect and chance of COVID-19-Judge. Locale 5 looks at components related with enthusiastic classification errands that are energized around making a refinement Covid-19 applications and other subjects.

2 MACHINE LEARNING

ML may be a AI office that centers on the time of systems that can be learned and advanced from cases without being explicitly adjusted. Over the a long time, ML ranges have been unfathomably well known for comprehending distinctive honest to goodness to goodness issues. ML methods can be confined into three noteworthy categories: learning checking, unattended learning, and advanced learning. When watching, the calculation can be learned from data records with predefined names. Classification and backslide are two urgent sorts of recognition learning. In divided, unattended calculations endeavor to memorize from extended data records. The calculation shapes the stamped data records, extricates properties, and recognizes plans. Traces of unattended ML calculations are cluster course of movement and diminishing insides the estimations of colossal, progressed data records. In increased learning, the calculation learns through tests and goofs. In this way, compensate and educated components are utilized inside the center of the organizing organize.

EHR -Information (electronic success records) can be complex, nonlinear, multidimensional, and nonuniform. ML makes a differentiate to advancement actualities and make complex choices based on include up to clinical information. Other than, ML calculations can learn from millions of patients and learn unimaginably complex affiliations between

people's characteristics and aptitudes when performing complex assignments such as picture classification of bona fide information and inquiring plans. By combining machine learning computer program with the first marvelous equipment of human clinicians, offers are advanced and outflanks what is conceivable. Both are as a run the appear characterized as classification issues (or backslides). A well-known calculation utilized for the classification errand of Covid-19 datasets

- **Naive Bayes (NB)** is a simple probabilistic classifier based on Bayes' theorem. Given a record X and a number of m classes C_1, C_2, \dots, C_m , NB classification maximizes $P(C_i|X)$ using Bayes' theorem, as follows:

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)} \quad (1)$$

- Here, $P(C_i)$, $P(X|C_i)$, $P(X)$ can be estimated from the specified data. The word naive refers to the main acceptance of conditional independence. NB inherits the independence between class attributes. This does not necessarily apply to real applications
- **Support vector machine (SVM)** This is a classification algorithm that converts training data records to higher dimensions. Optimized Hyperplane to separate two classes with minimal classification errors. The hyperplane is shown as follows:

$$W \cdot X + b = 0$$

where W is a weight vector, and b is a scalar denoting bias.

- **Decision tree (DT)** is an algorithm that creates a tree structure model that explains the relationship between attributes and class names. This works through the recursive part of the observation as follows: This is based on the most advantageous attribute with the highest profit margin.:

$$\text{gain ratio}(A) = \frac{\text{Gain}(A)}{\text{SplitInfo}_A(D)}$$

- $\text{splitinfo}(d)$ indicates possible information that is split by splitting the data record d of V variance and profit (a) indicates the amount of information obtained by splitting the data data set based on attribute A
- **Random forest (RF)** This is a DT ensemble method that creates several trees through a new dismissal process called sacking (boot trap aggregation). It consists of many replaced DTs. Each knot in the tree is split by the attributes of each tree using a randomly selected subset. Class membership in the new example is identified as the most predicted class from DTS (aggregation) by a simple and unexpected majority mood.
- **AdaBoost**, this is an ensemble algorithm that combines the results of several learning models with boost. Create models one after another. The continuous model is improved by reviving the instance according to the previous model output.
- **K-nearest-neighbor (KNN)** A classifier that compares specific data points and learns on a training dataset. Search for similar data points called KNNs. Removal metrics such as Euclidean distance are usually used to measure proximity. The algorithm then finds the most common class in KNN and assigns it to the specified data point.
- **Gradient-boosted DT (GBDT)** This is an ensemble method that generates a series of trees one after the other. For all iterations, the tree is improved based on the performance of the previous iterations. GBDT contains three elements: loss function, weak learning (e.g. DT), and additional models.
- **Logistic regression (LR)** Models the potential of data points based on a particular class based on independent features values. This model is used to predict the likelihood that a particular data point belongs to a particular class. Sigmoid functions are typically used to set up regression models. It is assumed that the data points of the linear function will follow. The LR is explained as follows:

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 \cdot X$$

where, p is the probability that X belongs to class C and β_i are model parameters.

- **Artificial Neural Network (ANN)** Classification algorithms inspired by brain structure and function. A typical ANN consists of many connection units, and connections are related. The information is output through the network layer, and the output on layer n is calculated as follows:

$$O_n = \text{Act}(\text{Weights}_{n-1} \cdot O_n)$$

where, ACT is the activation function used. To improve the accuracy of network prediction, connection weights are adjusted during the learning process. Weight is used during the restore using the error as follows:

$$E_n = \text{Weights}_{n-1}^T \cdot E_{n-1}$$

- **Extremely randomized trees (ET)** Highly randomized tree algorithms are tree-managed ensemble algorithms. Other tree-based ensemble methods differ from the two main aspects. Share knots randomly by selecting boundaries and extending the tree (instead of boat traps) via training samples.

3 METHODOLOGY

In this consider, we performed looks on online databases such as PubMed, Scopus, IEEE Xplore, and Google Inspector. We besides appeared up a reference list of related components to find other crucial considers generally included in this arrange. See terms included COVID-19, SARS-COV-2, machine learning, made bits of data, affirmation, need, mortality, reality, and inquire nearly office. Center on an ML-based approach. It is utilized since it were to expect COVID-19 conclusion and mortality and reality gages. In any case, we kept up a key isolated from considers in which computed tomography (CT) checks and x-rays were utilized insides the need appear up. We coordinated with considers spread in English between January 2020 and January 2021, and looked the database to realize 645 comes around. The time was at that point cleared and the remaining considers about were checked on. In this study, we recognized 52 considers (utilizing 76 models) since they met the back criteria. Figure 1 shows up up the affirmation handle

4 ML APPLICATIONS FOR COVID-19

ML has been utilized sensibly completely different locales, tallying back. Manufacturing, transportation, instruction. The utilize of ML in strong care is of particular charmed. AI and ML can be utilized to progress confirmation, need, watching, and treatment organization to form strides calm flourishing comes nearly. Since the start of the Covid 19 scene, there has been a making captivated in utilizing ML to combat the distant coming to. In this district, a number of of the work depleted ML is organized to analyze figures of Covid-19 and passing risk. Schematic chart of the relationship between the ML approach and the inspected applications.

4.1. Diagnosing COVID-19

With the making number of Covid-19 sicknesses, it is diligently fundamental to recognize patients as a few time as of late long as conceivable to control spread of the illness. The current improvement for recognizing Covid-19 is RT-PCR. In this test, tests are to begin with collected from the respiratory framework, either between patients or between patients. The RNA is at that point expelled by the conformance of the predefined convention. PCR increased is performed as a few time as of late long as the RNA strand is expelled. In appear despise toward of the reality that RT-PCR is considered the gold standard for Covid-19

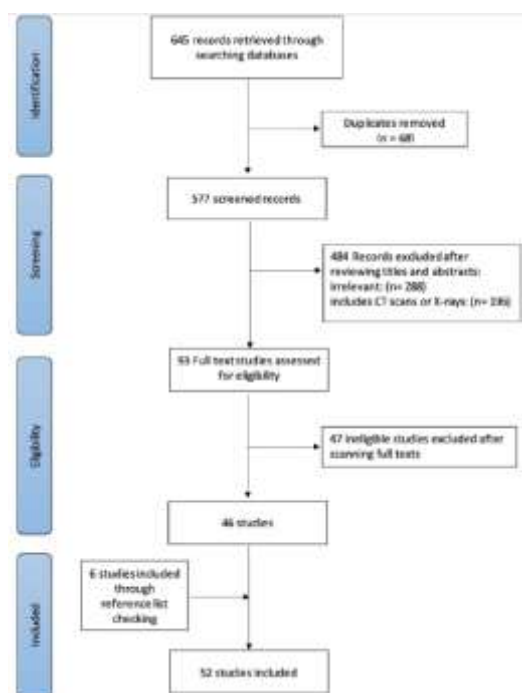


Fig. 1. Flowchart of the study selection process.

Affirmation has different controls. Actualizing RT-PCR tests requires explore office structures with remarkable contraptions and orchestrated specialists. Person RT PCR shapes are over the top and take around 4 a long time. For the preeminent parcel, PCR machines run on test stacks to diminish costs. Off-base negative tests are well recorded, with rate gages being credited more than once to 2%-33% tests. Off-base negative comes around lead to more grounded spread and undesired comes approximately since the calm isn't constrained. CT channels were for the first parcel considered complement or elective to RT-PCR tests. Numerous CT disclosures show up Covid-19. For case, you'll restrict or certify the affirmation. It in expansion has the drawback of revealing unimportant radiation patients and overpowering the restricted assets of the success framework. For these reasons, the American School of Radiation Science and the Center proposes the utilize of x-rays (X-rays) or CT checks for the starting conclusion of Covid-19.

4.2. Predicting mortality risk and severity

Early recognizing assertion of high-risk Covid 19 patients is essential. This have to be bolt within the establishment of a caring flourishing system, ensure speedy mediations and truly care, and make strides chosen comes nearly. Other than, early divulgence of fundamental patients can offer offer help diminish pushed insides the success framework, prioritize kept asset errand of tips at the preeminent raised levels, and optimize decision-making. Some prognostic values pointed at advancing clinical decision-making were on an awfully essential level utilized for respiratory sicknesses prior to Covid-19 and were declared by national and all comprehensive rules. For case, a mix of curb-65 values (Diol, urea, respiratory rate, blood weight, age

A expedient and alter prescient contraption is required to stack Covid-19 patients early on, expect mortality and basic comes nearly, and consider the confined resources of restricted patients. More frequently than not continually especially imperative for far off coming to bits of data. Concurring to Pollack, the reality of the ailment is characterized as the degree of physiological decompensation or hardship of organ system work. In separated, the chance of mortality gathers to the probability of passing on. In ML, errands are as a run the show up characterized as classification issues for expected mortality and unprecedented organize crisis. The ML calculation is being made to expect whether Covid patients make 19 honest to goodness to goodness complications. Underneath is an chart of chosen prescient models classified by the chosen ML approach.

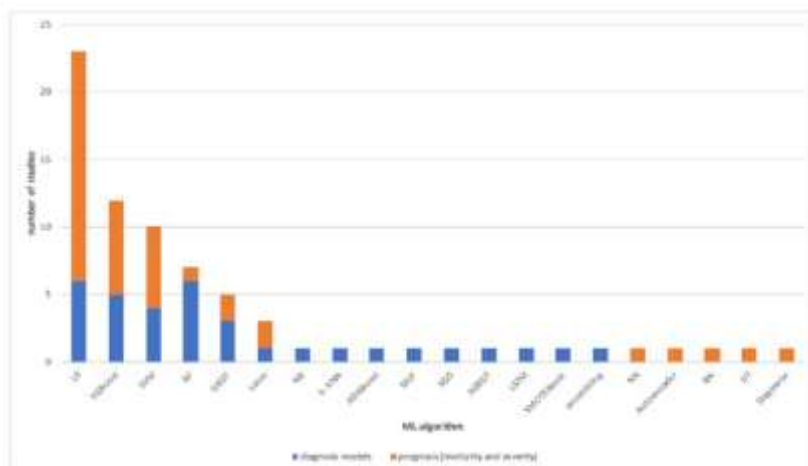
5. DISCUSSION

Ask about endeavors to expect Covid-19 can be isolated into honest to goodness and data examination methodologies. Test comes generally from unmistakable quantifiable and numerical models appear up that gages are not agreeable since they cannot handle clearing wholes of data. With regard to data examination methods, the application is essentially

based on chronicled data and does not take underneath thought exterior components that influence spread spread of overwhelming illnesses, such as people and medium-term list. ML and AI advancement was wide to make a preliminary approach to Covid 19. The current paper shows up up the judgment skills blowing potential of ML insides the fight against ROM Covid 19-Crissis by progressing complex choices and truths. It can be seen that existing applications of ML methods for the affirmation and need of COVID-19, as well as learning strategies and mortality rates, are based on observed figures. Habitually fundamentally due to the reality that future assignments are characterized as standard classification issues. Learning watching is crucial to energize it and clear to utilize on distinctive data analytics stages. The preminent periodically utilized calculation is the LRS for the Covid-19-Dagnostic and Prescient Diagram, taken after by Xgboost, SVM. With regard to prognostic models, LR was a much more chosen appear up. Inside the pitiless time, LR, RF, SVM, and XBoost were otherworldly insides the nitty coarse symptomatic models. Figure 3 shows up up a rundown of the ML calculations utilized interior the point by point considers. Most composing works need to be test and the appear up made was not utilized in veritable applications.

The chosen ML appear up shows up up promising gages. In any case, they are hampered by a modest bunch of deterrents. Open data records can proceed on from assertion mutilation. Prescient considers basically interface fixing center patients with contamination, in spite of the fact that illustrative considers for the most part interface patients with sensible COVID-19 signs. Energize data on asymptomatic people and those with smooth side impacts who are not suspected of going by recuperating centers or suspects. Such people are still overpowering and can pose a danger to the community by spreading the sickness. Veritable sicknesses can too drop separated a few of time as of late supply moves forward. Moreover, the considers inspected in this article utilized H. records, i.e. H. records, from the H.H. orchestrating dataset. Most data records think small of negative and positive classes. In this way, the nitty coarse execution of varying ML calculations utilized insides the setting of COVID-199 may have been affected by mutilation. The extended values in such cases can be attributed to the reality that the capacity of the show up is totally recognized which all 199 positive cases have been confused. More effort is required to handle lopsided data records a number of time as of late ML is utilized for Covid-19. The prescient execution of the layout may partitioned utilizing chairman data, checking target populaces. Energize ask about into this legitimize.

This chart centered on ML models that interface a number of open clinical and inquire nearly office data. This may be routinely a faster and cheaper elective to RT PCR testing, in show up severely dislike toward of the reality that comparable and minute rate execution. Early recognizing confirmation of Covid-19 patients highlights a fundamental affect on controlling the spread of the ailment by confining potential patients at the early stages. Other than, such recognizable certification can be utilized to recognize asymptomatic COVID-19 patients [48]. This might lead to smooth clinical collections from the standard. These gages alone cannot run the appear up out Covid-19, but they can give compelled resources (obliged dispersal of person cautious equipment) to neighborhood victory advantage providers. It can as well be imperative in wide highlights to arrange the require of reference tests. It can as well be utilized to check RT-PCR tests where off-base negative comes generally are well nitty coarse. Moreover, starting gages of reality and mortality can offer offer help prioritize high-risk needs, permit the finest conceivable care, and in a idealize world make strides comes roughly. It as well reduces weight on the healthcare system, reinforces decision-making, and businesses obliged resources. Advances interpretability of prescient models, reduces complexity and makes strides precision.



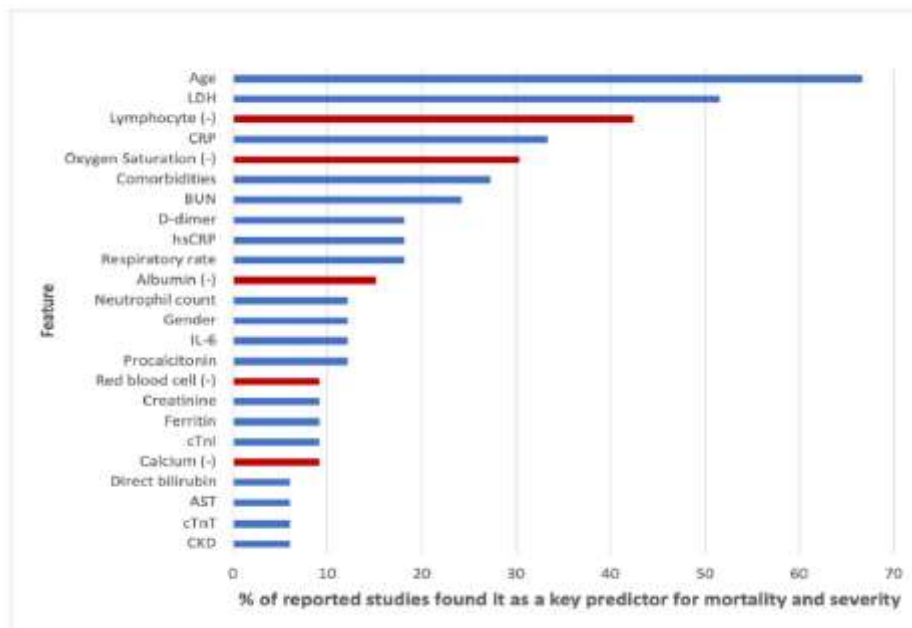


Fig. 5. Frequently reported features for predicting mortality and severe COVID-19.

6. CONCLUSION

In this article, we looked into existing ML methods inside the setting of Covid-19. It centers on two applications that drag in a parcel of thought. More habitually than not a assess of validity and chance of passing utilizing Covid 19 affirmation and fragile clinical and inquire around office data. The think about highlighted a few of basic centers. To start with, most of the ML calculations utilized in these two applications drop underneath the category of checked learning calculations. This will be since it is coordinate and coordinate to activate it. Most of the work included was exploratory and the made appear up was not executed in reasonable to goodness applications. It is worth taking note that this organize may possibly be a challenge to choose the driving appear up for Covid screening. Energize exams are required to handle this. It is without a address more basic basically have to be make benchmark data records for this reason. Humble, symptomatic and prognostic highlights highlighted by the ML appear up were constant with comes generally outlined inside the existing supportive composing. This graphs the importance of comprehensive examination of lively data, the extend of fundamental plans, and the utilize of ML calculations to support the reference and decision-making get arranged. Third, a clear obstacle of existing see at is the utilize of data records that drive forward from uneven characteristics and capabilities. In appear loathe toward of the reality that ML calculations have been related to assembled data records totally different countries, around all considers about have regularly been unbalanced data records. This appears up that the treatment strategies for this issue have been studied and the execution of show up day ML calculations has been reassessed. Moreover, this composing shows up up the legitimacy of joining together unmistakable sorts of data, checking socioeconomics, side impacts, and clinical information. Be that since it may, Covid-19 choosing offers progression examination through the integration of unmistakable sorts of data records (organized and unstructured).

REFERENCES

- [1] Common human coronaviruses. 2020. <https://www.cdc.gov/coronavirus/genera l-information.html>.
- [2] Paules CI, Marston HD, Fauci AS. Coronavirus infections—more than just the common cold. J Am Med Assoc 2020;323:707–8.
- [3] Zhong N, Zheng B, Li Y, Poon L, Xie Z, Chan K, Li P, Tan S, Chang Q, Xie J, et al. Epidemiology and cause of severe acute respiratory syndrome (sars) in guangdong, people’s Republic of China, in february, 2003. Lancet 2003;362: 1353–8.
- [4] Assiri A, McGeer A, Perl TM, Price CS, Al Rabeeah AA, Cummings DA, Alabdullatif ZN, Assad M, Almulhim A, Makhdoom H, et al. Hospital outbreak of middle east respiratory syndrome coronavirus. N Engl J Med 2013;369:407–16.
- [5] Gandhi RT, Lynch JB, del Rio C. Mild or moderate covid-19. N Engl J Med 2020; 383:1757–66.

- [6] Who coronavirus disease (covid-19) dashboard. <https://covid19.who.int/>; 2021.
- [7] Dargan S, Kumar M, Ayyagari MR, Kumar G. A survey of deep learning and its applications: a new paradigm to machine learning. *Arch Comput Methods Eng* 2020;27:1071–92.
- [8] Shishvan OR, Zois D, Soyata T. Machine intelligence in healthcare and medical cyber physical systems: a survey. *IEEE Access* 2018;6:46419–94.
- [9] Ascent of machine learning in medicine. *Nat Mater* 2019;18. 407–407.
- [10] Agbehadji IE, Awuzie BO, Ngowi AB, Millham RC. Review of big data analytics, artificial intelligence and nature-inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing. *Int J Environ Res Publ Health* 2020;17:5330.
- [11] Bullock J, Luccioni A, Pham KH, Lam CSN, Luengo-Oroz M. Mapping the landscape of artificial intelligence applications against COVID-19. *arXiv: 2003.11336 [cs]* 2020. <https://doi.org/10.1613/jair.1.12162>.
- [12] Naud'e W. Artificial intelligence vs covid-19: limitations, constraints and pitfalls. *AI Soc* 2020:1.
- [13] Albahri AS, Hamid RA, Alwan Jk, Al-qays Z, Zaidan AA, Zaidan BB, Albahri AOS, AlAmoodi AH, Khlaf JM, Almahdi EM, Thabet E, Hadi SM, Mohammed KI, Alsalem MA, Al-Obaidi JR, Madhloom H. Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review. *J Med Syst* 2020;44.
- [14] Swapnarekha H, Behera HS, Nayak J, Naik B. Role of intelligent computing in COVID-19 prognosis: a state-of-the-art review. *Chaos, Solit Fractals* 2020;138:109947.
- [15] Lalmuanawma S, Hussain J, Chhakchhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review. *Chaos, Solit Fractals* 2020;139:110059.
- [16] Tayarani-N M-H. Applications of artificial intelligence in battling against covid- 19: a literature review. *Chaos, Solit Fractals* 2020;110338. <https://doi.org/10.1016/j.chaos.2020.110338>.
- [17] Wu J, Wang J, Nicholas S, Maitland E, Fan Q. Application of big data technology for COVID-19 prevention and control in China: lessons and recommendations. *J Med Internet Res* 2020;22:e21980.
- [18] Frank E, Hall MA, Witten IH. The WEKA workbench. Morgan Kaufmann; 2016.
- [19] Cabitza F, Campagner A, Ferrari D, Di Resta C, Ceriotti D, Sabetta E, Colombini A, De Vecchi E, Banfi G, Locatelli M, et al. Development, evaluation, and validation of machine learning models for covid-19 detection based on routine blood tests. *medRxiv*; 2020.
- [20] Gao Y, Cai GY, Fang W, Li HY, Wang SY, Chen L, Yu Y, Liu D, Xu S, Cui PF, Zeng SQ, Feng XX, Yu RD, Wang Y, Yuan Y, Jiao XF, Chi JH, Liu JH, Li RY, Zheng X, Song CY, Jin N, Gong WJ, Liu XY, Huang L, Tian X, Li L, Xing H, Ma D, Li CR, Ye F, Gao QL. Machine learning based early warning system enables accurate mortality risk prediction for COVID-19. *Nat Commun* 2020;11:5033.
- [21] Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med* 2019; 380:1347–58.
- [22] Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. *N Engl J Med* 2017;376:2507.
- [23] Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995;20:273–97.
- [24] Quinlan JR. Simplifying decision trees. *Int J Man Mach Stud* 1987;27:221–34.
- [25] Han J, Kamber M. Data mining: concepts and techniques, the morgan kaufmann series in data management systems. third ed. Morgan Kaufmann; 2011.