Digitalization and Management Innovation IV
A. Lau and A.J. Tallón-Ballesteros (Eds.)

doi:10.3233/FAIA250734

Use of a Novel Clustered Bayesian 3D-CNN Model for Triage Classification of COVID-19 Mass Casualty Incidents

Liege CHEUNG [a], Jinrui CAO [b], Zequn LIN [b], Anthony C.H. SEK[b] and Adela S.M. LAU [b,1]
[a] *College of Liberal Arts & Sciences, University of Illinois Urbana-Champaign, USA*
[b] *Data Science Lab, Department of Statistics and Actuarial Science, School of Computing and Data Science, The University of Hong Kong, Hong Kong*
ORCiD IDs:
Liege Cheung https://orcid.org/0000-0001-7149-0455
Jinrui Cao https://orcid.org/0009-0008-1798-863X
Zequn Lin https://orcid.org/0009-0009-3021-8063
Anthony C.H. SEK https://orcid.org/0009-0000-4455-0790
Adela S.M. Lau https://orcid.org/0000-0001-5918-8309

Abstract. The first author previously conducted an empirical review of the constructs for determining triage recommendations for COVID mass casualty incidents (MCIs). This research aims to continue that previous study and to implement a novel clustered Bayesian 3D-CNN model for predicting COVID-19 MCI triage classification. An experimental research method is used here. Data were collected from the databases of the COVID Tracking Project, Oxford COVID-19 Government Response Tracker, and OxCGRT. The data samples were then integrated to construct a dataset with the three dominants of severity of infection (signs and symptoms of coronavirus), likelihood of spreading (attitudes towards pandemic spreading, personal behaviors, and government policy), and available resources (availability of physicians and medication, hospital vacancies, and evacuation assets) to train a 3D-CNN model on COVID19 test results, recovery histories and medication outcomes. The output of the model includes categories of triage recommendations entitled "immediate (hospital admission with immediate medical care)", "delayed (quarantine center and observation)", "minimal (home quarantine with medication)", "semi-minimal (home quarantine without medication)", and "no priority (no therapeutic resources)". Experimental results show that the accuracy rate, AUC, precision, recall and F1-score for the COVID19 MCI triage decisions are 82%, 85.8%, 0.835, 0.789, and 0.811, respectively. In addition, our model outperforms other triage decision models for emergency incidents. As a consequence of this study, medical professionals will be able to apply our model to their hospital data for triage decisions involving COVID-19 or other medical incidents.

Keywords. Bayesian network, 3D-CNN model, MCI triage classification, COVID-19

[1] Corresponding Author: Adela S.M. Lau, Department of Statistics and Actuarial Science, Run Run Shaw Building, The University of Hong Kong, Pokfulam Road, Hong Kong. Email: adelalau@hku.hk

1. Introduction

Although traditional statistical models provide straightforward implementation and clear interpretability, their reliance on strong parametric assumptions frequently leads to model misspecification, particularly when handling real-world epidemiological data. In an empirical study by Cheung et al. [1], the determinants of pandemic spread were found to include risk perception, hygiene behaviors, attitude towards pandemic prevention programs, culture, health education programs, government policies, technologies, information disclosure, and economic strategy, among others. Prediction of pandemic spread is a complex nonlinear problem that depends on many determinants, and traditional statistical methods cannot model nonlinear situations in a complex environment. Artificial intelligence (AI) can be used to model complex situations through nonlinear modeling; it requires no model assumptions or specific data structures, and can discover hidden patterns and relationships in the data. Thus, machine learning models should be employed for predictive modeling in complex environments.

The spread of a pandemic depends on specific factors such as the signs and symptoms of coronavirus, attitudes towards pandemic spreading and policy, and public resources, and conditional probability relationships exist among these factors and their determinants [1]. The spread therefore depends on combinations of these factors in this complex environment [1]. AI can be used to model both pandemic spread prediction and COVID-19 MCI triage classification. Methods that integrate traditional statistical models with AI methods can improve the accuracy of prediction and classification while minimizing computational complexity.

The following section will discuss the strengths and limitations of statistical and AI methods for prediction, and will discuss the determinants for predicting COVID-19 MCI triage classification. Sections 3 and 4 introduce the research method and the results, respectively. Section 5 contains a discussion of the results, and Section 6 concludes this study.

2. Literature Review

During the COVID-19 pandemic, accurate forecasting of the spread of infection was crucial in order to develop effective public health strategies. Researchers have employed both statistical and AI approaches to predict transmission patterns. Traditional statistical forecasting models have included autoregressive integrated moving average (ARIMA) [2] and susceptible-exposed-infectious-recovered (SEIR) models [3].

The basic idea of the ARIMA forecasting model is to predict future trends based on past time series data. A common assumption in time series models is that all the data are stationary, a mathematical term that refers to data whose statistical properties remain constant throughout the study period. The ARIMA model has been employed to predict the incidence and transmission dynamics of COVID-19 in the five countries with the highest reported caseloads [4]. Although ARIMA modelling yielded a high accuracy, the model failed to capture the complex nonlinear interactions affecting pandemic spread [4]. The study was constrained by the assumption that no vaccines or effective therapeutic interventions were available during the forecasting timeframe [4].

In a classic SEIR model, each individual is classified into one of four categories based on their epidemiological status: susceptible (S), exposed (E), infected (I), and recovered (R). To predict the spread of the pandemic, Huang and Qiao [5] implemented

the SEIR model to project infection trajectories, and estimated a pandemic peak occurring at approximately 14 days. The main limitations of their study were that they assumed unchanging infection rates, no introduction of new pharmaceutical interventions, and consistent medical management capacity throughout the analysis period [5]. In practice, the presence of super-spreaders may elevate COVID-19 transmission rates by up to 85% compared to typical infected individuals [6].

Although these methods benefit from the use of complete population datasets, without requiring additional sampling [7], they become less effective as the parameter complexity increases [8], and struggle with large datasets due to computational limitations [9]. Deep learning (DL) has emerged as a powerful alternative approach, with applications including long short-term memory (LSTM) networks [10], polynomial neural networks [11], and standard neural networks [12]. Comparative studies have demonstrated the superior computational efficiency of DL in processing time-series data [13].

Statistical methods can handle probabilistic relationships while AI captures complex details and nonlinear dynamics. Raddad et al. [14] developed a hybrid COVID-19 severity classification model that achieved an accuracy of 98% with a deep neural network (DNN) and 96% using XGBoost, which was validated using 5,654 patient records with 146 features. However, while AI achieves high classification accuracy, its low level of interpretability remains a fundamental limitation [15].

Bayesian networks offer distant advantages in terms of the graphical representation [16] of variables, although their performance remains contingent on the data dimensions and distributional assumptions [17]. In contrast, three-dimensional convolutional neural networks (3D-CNNs) have specific strengths, particularly for medical imaging analysis [18]. Their capabilities in regard to processing X-ray micro-topography, enabling rapid design iteration, and supporting probabilistic engineering design make them suitable for COVID-19 forecasting [19]. One important aspect is that 3D-CNNs can overcome the limitations of Bayesian networks in handling high-dimensional data. Hence, an integrated approach based on a Bayesian network and 3D-CNN model is recommended for predicting the spread of COVID-19.

In Cheung's empirical review study [20], the major determinants for the prediction of COVID-19 MCI triage classification included the severity of infection, likelihood of spreading and available resources. The determinants for the prediction of severity of infection rely on signs and symptoms of coronavirus, such as fever, cough, and shortness of breath [21]. Attitudes towards pandemic spread, personal behaviors, and government policy are the key determinants for predicting the likelihood of spreading. The availability of physicians and medication, hospital vacancies, and evacuation assets determine the available hospital resources. Hence, we adopt these factors when constructing our novel clustered Bayesian 3D-CNN model for predicting COVID-19 MCI triage classification based on COVID-19 test results, recovery history and medication outcome. The output of the network is divided into four categories of triage recommendations: "immediate (hospital admission with immediate medical care)", "delayed (quarantine center and observation)", "minimal (home quarantine with medication)", "semi-minimal (home quarantine without medication)", or "no priority (no therapeutic resources)".

3. Research Method

3.1 Data Selection

In this study, three publicly available COVID-19 datasets [22,23,24] were integrated to evaluate the severity of infection, likelihood of disease transmission, and availability of healthcare resources across 40 US states and territories, using a Bayesian network model for predicting the risk of COVID-19 positivity. These datasets were the COVID Racial Data Tracker (CRDT) [25], data from state-level testing and outcomes from the COVID Tracking Project [22], and the Oxford COVID-19 Government Response Tracker (OxCGRT) [24]. The selected variables were the three determinant categories of severity indicators, transmission factors, and resource metrics, which were augmented by a derived death rate metric and then used to simulate infection risk through a Bayesian network framework. The categories, units, data types, and descriptions of variables are summarized in Table 1.

The CRDT dataset [25] contains daily cumulative cases and deaths by race/ethnicity. We selected Cases_Total (cumulative confirmed and probable cases), Deaths_Total (cumulative confirmed and probable deaths), and positiveIncrease (daily new cases) as severity indicators, to capture the prevalence of infection and mortality burden. The States dataset, sourced from the COVID Tracking Project [22], offers daily metrics on testing and healthcare utilization; from this, we extracted hospitalizedCurrently (current hospitalizations with COVID-19), inIcuCurrently (current ICU patients), onVentilatorCurrently (current ventilator patients), and totalTestResultsIncrease (daily new tests) to reflect the real-time severity of disease and testing capacity, with information on both severity and transmission dynamics.

The OxCGRT dataset [24], filtered for US subnational data, provides government response policies and indices. We selected StringencyIndex_Average, GovernmentResponseIndex_Average, and ContainmentHealthIndex_Average as aggregate transmission factors, alongside containment policies (e.g., C1M_School.closing, C2M_Workplace.closing, C3M_Cancel.public.events), public health measures (H1_Public.information.campaigns, H6M_Facial.Coverings), healthcare policies (H4_Emergency.investment.in.healthcare, H8M_Protection.of.elderly_people), and economic support (E1_Income.support, E2_Debt.contract.relief) to model transmission risk and resource availability. A derived variable, death_rate (Deaths_Total/Cases_Total), was used to quantify case fatality as an additional metric of severity. The variable H5_Investment.in.vaccines was excluded, as it had zero values across all records, indicating no reported state-level vaccine investment. The period analyzed here was from April 12, 2020, to March 7, 2021, covering 330 days of the initial COVID-19 pandemic wave in the US. This period was selected based on the availability of data and the dynamic evolution of government responses.

We extracted two datasets (CRDT [22] and States [25]) from the COVID Tracking project [26], comprising daily records across 40 states and territories. Fifteen regions (including California and New York) were excluded due to incomplete data availability. The OxCGRT dataset was obtained from the OxCGRT GitHub repository [27], as this offered daily subnational policy data. We processed the data daily to capture rapid shifts in transmission patterns, healthcare burden, and intervention impacts.

The raw datasets were integrated through exact matching of the data and state fields, with all data values standardized to date/time format and state codes made consistent. To ensure data integrity, we filtered the dataset by requiring exact matches between

Cases_Total and ConfirmedCases and between Deaths_Total and ConfirmedDeaths in OxCGRT, thereby eliminating discrepancies. To fill the missing values in the columns entitled hospitalizedCurrently, inIcuCurrently, and onVentilatorCurrently, we implemented k-nearest neighbors (k=3) to preserve data patterns while completing all 2,184 daily records. The final dataset covered 40 states and territories, with approximately 54 observations per state, reflecting filtering constraints. We used data from April 12, 2020, to December 31, 2020 (1,934 samples) for model training and January 1, 2021, to March 7, 2021 (250 samples) for testing, supporting the Bayesian network's predictive analysis.

Finally, we mapped the selected variables to nine sub-factors to simulate patients' risk of infection: severity (X11: fever, X12: cough, X13: shortness of breath), transmission (X21: pandemic attitude, X22: personal behavior, X23: government policy), and resources (X31: physician availability, X32: hospital vacancy, X33: evacuation assets). These were normalized and discretized (threshold=0.15) to construct three Bayesian networks with noisy-OR conditional probability distributions, to estimate the probabilities for the severity $P(X_1)$, transmission $P(X_2)$, and resource $P(X_3)$ levels. A weighted COVID-19 positivity probability was computed as shown in (1), to gain insights into risk assessment and resource allocation.

$$R = 0.5P(X_1) + 0.4P(X_2) + 0.1\big(1 - P(X_3)\big) \quad (1)$$

In summary, the raw independent variables included seven severity indicators, 11 transmission factors, and four resource metrics, drawn from 2,184 daily samples from across 40 US states and territories. This selection supports a robust analysis of COVID-19 dynamics and facilitates the Bayesian network model's evaluation of infection risk.

Table 1. Summary of selected variables for Xs' input

Factor category	Sub-factors	Name	Units	Data type	Description
Identifiers		Date	Date	datetime64[ns]	Date of the record (daily granularity)
		State	N/A	object	State or region name within the USA
X1: Severity of infection		Cases_Total	Count	float64	Total confirmed COVID-19 cases reported in the region to date
		Deaths_Total	Count	float64	Total confirmed COVID-19 deaths reported in the region to date
	Fever	positiveIncrease	Count (daily)	float64	Number of new confirmed positive COVID-19 cases on that day
	Cough	hospitalizedCurrently	Count	float64	Number of COVID-19 patients currently hospitalized
	Shortness breath	inIcuCurrently	Count	float64	Number of COVID-19 patients currently in intensive care units (ICUs)
	Shortness breath	onVentilatorCurrently	Count	float64	Number of COVID-19 patients currently on ventilators
	Fever, cough, shortness breath	death_rate	Proportion (0-1)	float64	Death rate, calculated as Deaths_Total divided by Cases_Total
X2: Likelihood of spreading	Pandemic attitude	StringencyIndex_Average	Index (0-100)	float64	Average stringency of government containment and closure policies
	Pandemic attitude	GovernmentResponseIndex_Average	Index (0-100)	float64	Average government response index, including economic and health policies

	Pandemic attitude	ContainmentHealthIndex_Average	Index (0-100)	float64	Average index of containment and health system policies
		totalTestResultsIncrease	Count (daily)	float64	Number of new COVID-19 tests reported on that day
	Personal behavior	H1_Public.information.campaigns	Ordinal scale (0-2)	float64	Public information campaigns intensity (0: none, 1: limited, 2: extensive)
	Personal behavior	H6M_Facial.Coverings	Ordinal scale (0-4)	float64	Policy on face coverings/mask mandates (0: no policy to 4: required outside home at all times)
	Government policy	C1M_School.closing	Ordinal scale (0-3)	float64	School closure policy (0: no measures to 3: require closing all levels)
	Government policy	C2M_Workplace.closing	Ordinal scale (0-3)	float64	Workplace closure policy (0: no measures to 3: require all but essential workplaces to be closed)
	Government policy	C3M_Cancel.public.events	Ordinal scale (0-2)	float64	Cancellation of public events (0: no measures, 2: require cancellation)
	Government policy	C4M_Restrictions.on.gatherings	Ordinal scale (0-4)	float64	Restrictions on private gatherings (size limits)
	Government policy	C5M_Close.public.transport	Ordinal scale (0-2)	float64	Closure of public transport (0: no measures, 2: require closure)
	Government policy	C6M_Stay.at.home.requirements	Ordinal scale (0-3)	float64	Stay-at-home requirements (0: none, 3: require not leaving home, with minimal exceptions)
	Government policy	C7M_Restrictions.on.internal.movement	Ordinal scale (0-2)	float64	Restrictions on internal movement within the country/region
	Government policy	C8EV_International.travel.controls	Ordinal scale (0-4)	float64	International travel controls (0: no restrictions, 4: total border closure)
X3: Available resources	Physician availability	H4_Emergency.investment.in.healthcare	Ordinal scale (0-2)	float64	Emergency investment in healthcare system (0: none, 2: significant investment)
	Evacuation asset	H8M_Protection.of.elderly.people	Ordinal scale (0-2)	float64	Policies protecting elderly people (0: none, 2: strong protection measures)
	Evacuation asset	E1_Income.support	Ordinal scale (0-2)	float64	Income support policies for households (0: none, 2: generous support)
	Evacuation asset	E2_Debt.contract.relief	Ordinal scale (0-2)	float64	Debt or contract relief measures (0: none, 2: widespread relief)
	Hospital vacancies	Inverse_hospitalizedCurrently		float64	Indicator representing hospital bed vacancy, calculated as $1 - (hospitalizedCurrently_{normalized} + inIcuCurrently_{normalized} + onVentilatorCurrently_{normalized})/3$ A higher value indicates more vacant beds and less pressure on hospital resources. The original data are normalized to the range [0, 1] through min - max scaling.
	Hospital vacancies	Inverse_inIcuCurrently		float64	Indicator representing ICU bed vacancy, calculated using the same method as Inverse_hospitalizedCurrently,

	Hospital vacancies	Inverse_onVentilatorCurrently	float64	reflecting the vacancy status of ICU beds Indicator representing ventilator availability, calculated using the same method as Inverse_hospitalizedCurrently, reflecting the availability of ventilators

Table 2. Summary of selected variables for Ys' output

Y-Score Category	Name	Units	Datatype	Description
Y1: Immediate	Y1_Immediate_Score	Ordinal scale (1-4)	int64	Y1 score for immediate hospital admission with urgent medical care, based on high COVID probability and severe symptoms
	Y1_Probability_Distribution	Probability (0-1)	float64	CNN 4-class SoftMax probability distribution for Y1 immediate triage category, learned from spatial medical patterns
Y2: Delayed	Y2_Delayed_Score	Ordinal scale (1-4)	int64	Y2 score for delayed quarantine center observation, targeting moderate COVID probability cases
	Y2_Probability_Distribution	Probability (0-1)	float64	CNN 4-class SoftMax probability distribution for Y2 delayed triage category for quarantine center placement
Y3: Minimal	Y3_Minimal_Score	Ordinal scale (1-4)	int64	Y3 score for minimal home quarantine with medication support, for low-probability COVID cases with symptoms
	Y3_Probability_Distribution	Probability (0-1)	float64	CNN 4-class SoftMax probability distribution for Y3 minimal triage category for medicated home quarantine
Y4: Semi-minimal	Y4_Semi_Minimal_Score	Ordinal scale (1-4)	int64	Y4 score for semi-minimal home quarantine without medication, for very low-probability COVID cases
	Y4_Probability_Distribution	Probability (0-1)	float64	CNN 4-class SoftMax probability distribution for Y4 semi-minimal triage category for minimal medical intervention
Composite	Max_Y_Score	Ordinal scale (1-4)	int64	Maximum Y-score across all four categories from Bayesian network used to determine final triage priority
	Final_Triage_Category	Categorical (0-3)	int64	Final triage decision determined by CNN Y-score predictions: 0=immediate, 1=delayed, 2=minimal, 3=semi-minimal

3.2 Design of the Clustered Bayesian 3D-CNN model

Step 1: Data collection from various data sources
We first collected the patient symptom data, pandemic behavior factors, and healthcare resource factors, and stored them in the medical database for implementation of the clustered Bayesian 3D-CNN model (see Figure 1).
Step 2: Data preprocessing with Bayesian network
Data on infection factors (i.e., signs and symptoms such as fever indicators, cough, shortness of breath), pandemic spreading factors (i.e., pandemic attitude, personal behavior compliance, effectiveness of government policy), and healthcare resource availability factors (e.g., physician availability, hospital vacancy rates, evacuation asset readiness) were pre-processed. In particular, the X1 signs and symptoms input matrix had three layers to handle the three symptom factors (X11=fever, X12=cough,

X13=shortness of breath), with spatial dimensions of 10×10, representing the geographic distribution of patients measured at the state-day level for one temporal period. Similarly, the X2 pandemic attitudes input matrix had three layers, with spatial dimensions of 10×10, which stored the three pandemic behavioral factors (X21=pandemic attitude, X22=personal behavior, X23=government policy) for the same geographic regions. The X3 available resource matrix contained three layers to handle the healthcare resource factors (X31=physician availability, X32=hospital vacancy, X33=evacuation assets), with spatial dimensions 10×10, representing the regional healthcare infrastructure data.

Step 3: Model design

We expressed the three categories of medical explanatory variables as a clustered 3D matrix, and imported it into the three independent convolutional sub-models described above. To account for medical factors measured across different semantic domains and scales, three clusters are necessary. Otherwise, combining heterogeneous medical factors will constrain the choice of feature extraction patterns, thereby limiting the model's capacity to capture domain-specific medical relationships.

Step 4: Feature extraction with filters

In this step of training, we applied progressive filters **(8→16→32)** with matrix dimensions of 3×3×3 in each of the convolution layers to identify the most important medical patterns.

Step 5: Enhancement of model efficiency with activation function and max pooling

To determine the most significant medical patterns or features, we extracted the convoluted values from the three 3D-CNN models and passed them to their corresponding activation and max pooling functions with dimensions 2×2×2.

Step 6: Model training and optimization with Adam optimizer

Finally, the three 3D-CNN clusters generated hidden representations of the symptom severity, pandemic risk exposure, and healthcare resource availability. These formed a concatenated 96-dimensional feature vector that was used for training. In addition, to train the CNN with a backpropagation model for the Y-score prediction and four-class triage decision making, the COVID probability computed by the Bayesian network was passed to the CNN. The model predicted four separate Y-scores (Y1=immediate, Y2=delayed, Y3=minimal, Y4=semi-minimal) as four-class categorical distributions (probability distributions over severity levels 1 to 4). The Y-scores predicted by the CNN determined the final triage category directly, incorporating medical domain knowledge that Y-scores contain the necessary triage information. The entire medical decision system was trained through backpropagation, using a multi-task loss function (0.5×cross-entropy for triage + 0.5×cross-entropy for Y-scores), in which gradients were propagated backward through all prediction heads to jointly optimize the Y-score learning and four-class triage decisions. The Adam optimizer was used to update the network parameters iteratively, with learning rate scheduling and early stopping to ensure stable convergence while preserving the medical logic hierarchy of the Bayesian COVID→CNN Y-scores→Triage decisions. The details of the structural design of the clustered 3D-CNN model for COVID-19 medical triage are shown in Figure 1.

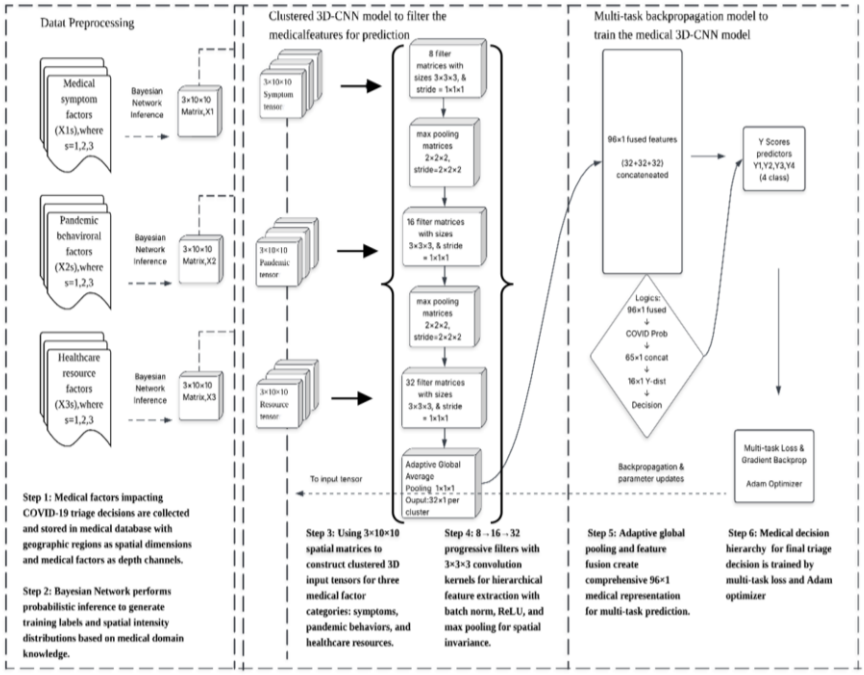

Figure 1. Clustered 3D-CNN model for triage decision prediction.

3.3 Data Pre-processing

The three symptom factors (i.e., $X_{1,i}$), three pandemic behavioral factors (i.e., $X_{2,j}$), and three healthcare resource factors (i.e., $X_{3,k}$) were pre-processed using **Bayesian Network inference** to predict the four Y-score triage categories (i.e., Y_m) for COVID-19 MCI triage decision making. The symptom input matrix has three layers corresponding to the three symptom factors, with spatial dimensions of 10×10, representing the patient distribution across geographic regions for each state-day period. Similarly, the pandemic input matrix has three layers and spatial dimensions of 10×10, storing three pandemic behavioral factors for the same geographic regions. The healthcare resource matrix contains three layers corresponding to the resource availability factors, with spatial dimensions of 10×10, representing the regional healthcare infrastructure capacity.
Symptom factor vector $X_{1,i}$ *= {fever indicator, cough presence, shortness of breath}*

where i = 1, 2, 3

Pandemic behavioral vector $X_{2,j}$ *= {pandemic attitude compliance, personal behaviour adherence, government policy effectiveness}*
where j = 1, 2, 3

Healthcare resource vector $X_{3,k}$ *= {physician availability, hospital vacancy rate, evacuation asset readiness}*

where k = 1, 2, 3

Triage categories $Y_m \in \{Y_1(Immediate), Y_2(Delayed), Y_3(Minimal), Y_4(Semi\text{-}minimal)\}$
where m = 1, 2, 3, 4

3.3.1 Data preprocessing for the Bayesian network

Unlike traditional normalization approaches, our method employs **Bayesian Network inference** with conditional probability distributions (CPDs) to capture medical domain knowledge and uncertainty at the data pre-processing stage.
The **probabilistic pre-processing pipeline** transforms raw medical categorical data into semantically informed spatial representations, as follows:
Raw Evidence → Bayesian Inference → Spatial Clustering →
Domain Normalization
This approach enables the clustered Bayesian 3D-CNN to learn from medically meaningful spatial patterns, rather than arbitrary categorical encodings, while maintaining inference consistency across the training and validation datasets.
Probabilistic Medical Inference Framework
The Bayesian network establishes causal medical relationships through a structured probabilistic model, a hierarchical medical decision framework. Given patient evidence **E = {X1, X2, X3}**, inference proceeds through the following intermediate medical constructs:
The formula for exposure risk assessment from pandemic behavioral factors is as follows:

$$P(Exposure_{risk}|X_{2,j}) = TabularCPD(exposure_{risk}|X_{21}, X_{22}, X_{23})$$

$$Risk_{score} = 0.4 \times X_{21} + 0.4 \times X_{22} + 0.2 \times X_{23}$$

$$Exposure_{risk} = 0.2 + Risk_{score} \times 0.6$$

where $TabularCPD$ Stands for the tabular conditional probability distribution.
The detection quality is calculated from the healthcare resource availability as follows:

$$P\left(Detection_{quality}\middle|X_3\right) = \text{TabularCPD}\left(detection_{quality}\middle|X_{31}, X_{32}, X_{33}\right)$$
$$Quality_{score} = 0.3 \times X_{31} + 0.4 \times X_{32} + 0.3 \times X_{33}$$
$$Detection_{quality} = 0.3 + Quality_{score} \times 0.5$$

The COVID probability integrates symptom evidence with exposure context, as follows:

$$P(COVID|X_1, Exposure_{risk}) = \text{TabularCPD}(covid|X_{11}, X_{12}, X_{13}, exposure_{risk})$$
$$Symptom_{score} = 0.2 \times X_{11} + 0.18 \times X_{12} + 0.25 \times X_{13}$$
$$COVID_{prob} = 0.15 + Symptom_{score} \times 0.6 + Exposure_{risk} \times 0.3$$

Medical Domain Normalization

Different medical parameters are evaluated using separate semantic domains and probabilistic scales, causing heterogeneous input characteristics. When modelling a medical neural network, the use of normalized spatial data typically accelerates the learning process and leads to faster convergence while preserving the medical semantic relationships.

In our clustered spatial representation, we perform **clipping normalization** to ensure numerical stability and preserve the medical semantic relationships, as follows:

$X' = clip(X, 0,1) = max\big(0, min(X, 1)\big)$

In this approach, all spatial intensity values are bounded within the range [0,1] while the relative relationships between active and inactive medical factors are preserved.

3.3.2 Model Design

Three clustered 3D-CNN models were developed to reflect the three medical factor categories (symptoms, pandemic behaviors, healthcare resources). A total of 60% of the dataset was used for training, 20% for validation, and 20% for testing. The dimensions of the three input matrices were 3×10×10, 3×10×10, and 3×10×10, and the predicted outputs were four-class categorical distributions for the scores Y1 to Y4 , with dimensions of 1×1 (COVID probability), and 1×1 (send-home decision). The structural details of the clustered 3D-CNN are shown in Figure 1.

Convolution Layer

The convolution layer takes as input a matrix with dimensions M rows, N columns, Q slices. In the first layer, there are **eight** filters (kernel), each with dimensions of m rows, n columns, and q slices, followed by **16** and **32** filters in the subsequent layers. The size of each convolution stride is $s_1 \times s_2 \times s_3$. Consequently, the output size of the convolution layer O(i,j,k) is $[(M–m)/s_1]+1$ rows, $[(N–n)/s_2]+1$ columns, and $[(Q–q)/s_3]+1$ slices. The equation used to compute the output of the convolution layer is shown below:

$$O(i,j,k) = \sum_{a=1}^{m}\sum_{b=1}^{n}\sum_{c=1}^{q} f(\, i+a-1, j+b-1, k+c-1)\, \mathrm{g(a,b,c)}$$

where i ranges from 1 to $[(M–m)/s_1]+1$, j ranges from 1 to $[(N–n)/s_2]+1$, and k ranges from 1 to $[(Q–q)/s_3]+1$.

Activation Function

To determine whether to pass onward information, and how much to transfer, the output of the convolution layer is fed to an activation function ReLU = max(x,0).

$$R(x) = \begin{cases} x & x > 0 \\ 0 & x \le 0 \end{cases}$$

where x is an element of the convolution result $O(i,j,k)$.

Batch Normalization Layer

Each convolution block includes batch normalization to stabilize medical feature learning:

$BN(x) = \gamma \times (x - \mu)/\sigma + \beta$

where μ and σ are the batch mean and standard deviation, respectively, and γ and β are learnable parameters.

Max Pooling Layer

A max pooling matrix has x rows, y columns and z slices, with a stride of size $t_1 \times t_2 \times t_3$. The output max pooling matrix M(u,v,w) has dimensions f×g×h (rows×columns×slices), where $f = [(M–x)/t_1]+1$, $g = [(N–y)/t_2]+1$, and $h = [(Q–z)/t_3]+1$. The equation used to compute the max pooling output is:

$M(u,v,w) = max_{c=0\ldots x-1,d=0\ldots y-1,e=0\ldots z-1} O(u+c,v+d,w+e)$
where u ranges from 1 to f, v ranges from 1 to g, and w ranges from 1 to h.

Adaptive Global Average Pooling Layer
Following the max pooling operations, adaptive global average pooling is applied to generate translation-invariant feature representations. For each channel i of the final feature map F with dimensions D×H×W, the adaptive global average pooling output GAP(i) is computed as:
$\text{GAP}(i) = \frac{1}{D\times H\times W} \times \sum_{d=1}^{D}\sum_{h=1}^{H}\sum_{w=1}^{W} F(i,d,h,w)$

Fully Connected Layer and Sigmoid Function
The outputs of the max pooling layers from the three medical clusters are concatenated into a feature input vector **with dimensions 96×1**, which serves as input to the fully connected layer. The aim is to train the clustered 3D-CNN model with backpropagation for medical triage prediction. Medical features are processed through multiple specialized prediction heads, where each medical task has its own dedicated pathway: the Y1–Y4 scores use four-class SoftMax classification, COVID probability uses sigmoid activation, and COVID-19 MCI triage decision uses sigmoid activation. This allows the model to learn task-specific medical patterns in a natural way.
For Y-scores, $\text{SoftMax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{4} e^{x_j}}$, where x_i is the raw logit/score for class *i* before activation, which ensures that each Y-score is a probability distribution over four severity levels (1–4).
For probabilities, $\text{Sigmoid}(x) = \frac{1}{1+e^{-x}}$, where x is the output from each medical prediction head, ensuring COVID and COVID-19 MCI triage decision probabilities are bounded between zero and one.

3.3.3 Model Training

Figure 2 illustrates the training process of the clustered 3D CNN model.

3.3.4 Complexity Analysis

The following 3D convolution formula shows how the computational complexity of the proposed medical clustered 3D-CNN model is computed, accounting for each convolution cluster:

$$K1 \times K2 \times K3 \times C_{in} \times D_{out} \times H_{out} \times W_{out} \times C_{out}$$

where the kernel size is $K1 \times K2 \times K3$; C_{in} and C_{out} are the number of input and output channels, respectively; and D, H, and W are the depth, height, and width of the 3D feature map.
For the medical symptom factors, the $3 \times 3 \times 3$ convolution with **eight filters** on a $3 \times 10 \times 10$ feature map with three input channels requires $3 \times 3 \times 3 \times 3 \times 3 \times 10 \times 10 \times 8 = 194{,}400$ multiply-accumulate operations. Similarly, for the first convolution layer, which processes the pandemic factors, the same 194,400 operations are required, and for the resource factors, another 194,400 operations are involved, giving a total of **583,200 operations** for the first convolution layer across all three clusters.

The subsequent layers with 8→16 and 16→32 channel progressions follow similar computational patterns, with reduced spatial dimensions due to pooling. Overall, the complexity of the clustered 3D-CNN model is $O(K^3NC)$, where $K = \max(K_1, K_2, K_3) = 3$ (kernel size), $C = \max(C_{in}, C_{out}) = 32$ (max channel count), and $N = \max(D_{out}, H_{out}, W_{out}) = 10$ (max feature map dimension).

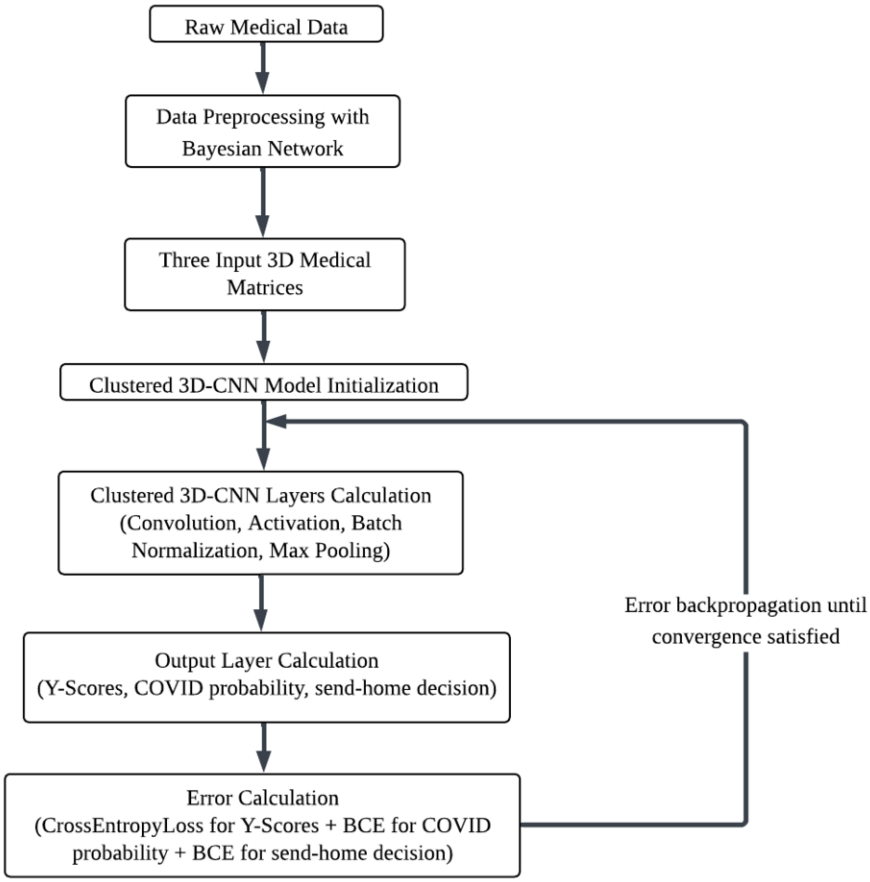

Figure 2. Training processes of the Clustered 3D-CNN model.

The evaluation metrics for medical triage performance are listed in Table 3.

Table 3 Evaluation metrics for medical triage performance

Evaluation metrics	Formula
Y-score accuracy	$\frac{1}{T}\sum_{i=1}^{T} 1[\hat{y}_i = y_i]$
Y-score cross-entropy loss	$-\frac{1}{T}\sum_{i=1}^{T}\sum_{j=1}^{4} y_{i,j}\log(\hat{y}(i,j))$

COVID Prob MAE	$\frac{1}{T}\sum_{i=1}^{T}\widehat{p_i} - p_i$
COVID-19 MCI triage classification accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$

4. Results and Discussion

Table 4 summarizes the results for the accuracy of COVID-19 prediction, triage decisions, the overall performance, and the significance analysis of the clustered Bayesian 3D-CNN model, a traditional machine learning ensemble, and a standard CNN.

Table 4. Results of COVID-19 triage predictions with Clustered Bayesian 3D-CNN, traditional ML ensemble, and Standard CNN.

Model	Average accuracy	Average AUC-ROC	Average precision	Average recall	F1 score	p-value	Cohen's d	95% CI
Clustered Bayesian 3D-CNN	0.820	0.858	0.835	0.789	0.811	-	-	-
Traditional machine learning ensemble	0.742	0.781	0.756	0.698	0.726	<0.001	0.73	[0.062, 0.094]
Standard CNN	0.734	0.776	0.748	0.691	0.719	<0.001	0.68	[0.071, 0.101]

We used McNemar's test comparing against the clustered Bayesian 3D-CNN (reference model). We also examined the effects using Cohen's conventions as well as a 95% confidence interval for the difference in accuracy from the reference model.

5. Discussion

The superior accuracy of the proposed clustered Bayesian 3D-CNN model in regard to triage classification (see Table 4) stems from its innovative integration of Bayesian inference with 3D CNNs. This architecture effectively captures the complex spatial-temporal and probabilistic relationships among heterogeneous forms of COVID-19 patient data, including symptoms, pandemic behavior, and healthcare resource availability. Unlike traditional machine learning models that assume conditional independence, our model leverages clustered data representations to preserve domain-specific parameter correlations, with a statistically significant although modest improvement in accuracy.

The statistically validated performance improvements translate directly to meaningful clinical benefits. The improvements in accuracy for our model of 7.8–8.6% represent approximately 78–86 correct triage decisions per 1000 patients, corresponding to substantial improvements in the efficiency of the emergency department. The improvement in precision of 0.079 over traditional machine learning reduces false

positive admissions by nearly 8%, potentially preventing unnecessary resource utilization during pandemic surges.

The effect sizes of 0.73 and 0.68 for Cohen's d exceed the 0.5 threshold that is typically considered as practically significant in medical decision support systems, thus confirming that our architectural innovations provide tangible clinical benefits beyond statistical significance. The consistent improvements across multiple metrics (accuracy, precision, recall, and AUC-ROC) represent robust performance advantages compared to single-metric optimization.

Tuning of the model involved iterative optimization of convolutional kernel sizes, channel depths, and batch normalization layers to improve feature extraction and reduce overfitting. We also applied max normalization and an 80/20 training-testing split, to ensure stability and generalizability across diverse patient profiles.

Key challenges included the management of heterogeneous data sources, taking into consideration the uncertainty inherent in medical data, and preventing overfitting of the model when trained on limited samples. These issues were addressed by using Bayesian pre-processing to embed causal knowledge and uncertainty, designing specialized CNN clusters for each data domain, and employing normalization and regularization techniques.

Overall, as shown in Table 3, our approach significantly outperformed the baseline models in terms of accuracy, AUC-ROC, precision, and recall, thus demonstrating its effectiveness for COVID-19 triage decision support.

However, in Almulihi's [28] study, focusing on traditional machine learning approaches for medical triage systems, our model's 78.9% recall for COVID-19 MCI triage decision is consistently lower than precision-optimized models. A comparative analysis of the recall results against a random forest-based model, an SVM-based model, and an ensemble NN-based model is provided in the last two rows of Table 4. This occurs mainly because simulated and real-world clinical datasets are used. In controlled clinical environments, patient presentations follow more predictable patterns, with standardized diagnostic protocols, while real-world emergency triage involves more complex patient interactions with comorbidities, symptom variability, and resource constraints [29]. Moreover, COVID-19 triage is more strongly influenced by the timing of symptom onset than symptom severity alone, due to the dynamic nature of viral progression and infectivity windows [29]. Notably, our test period covers the early pandemic phase in 2020–2021. The emergence of COVID-19 variants significantly influenced the clinical presentation and transmission patterns of patients worldwide; however, traditional models cannot adequately learn these evolving epidemiological scenarios from pre-pandemic training data.

6. Conclusion

This study has introduced a clustered Bayesian 3D-CNN model, thereby enhancing the interpretability and scalability of COVID-19 triage, using simulated patient data. As it includes a Bayesian network with 3D convolutional layers, our model can learn joint spatial-temporal and uncertainty patterns from mixed inputs such as symptom levels, public behaviors and the availability of health care tools, laying down a flexible engine that hospitals can scale up or down as demand shifts. The model achieves values for the accuracy of 82.0%, AUC-ROC of 85.8%, precision of 83.5%, recall of 78.9%, and F1-score of 81.1% (see Table 5). These values exceed the results attained by standard

benchmarks such as a full-learning ensemble, a typical 2D-CNN setup, an ARIMA pipeline, and an SVR approach. The recall rate still represents a notable shortfall, a discrepancy that is attributable to the noise inherent in field data and the rapid proliferation of novel variants. Hence, future versions of the model will need to target these strains more carefully. In general, the platform offers frontline clinicians a robust, data-driven methodology for expediting patient triage in time-critical scenarios.

Table 5. Comparison of our model with other models as benchmarks

Model Type	ARIMA-based model for emergency incident triage classification	SVR-based model for emergency incident triage classification	NN-based model for emergency incident triage classification	Clustered Bayesian 3D-CNN (ours) for COVID-19 MCI TRIAGE CLASSIFICATION
Accuracy	72.4%	–	75.8%	82.0%
COVID-19 MCI Triage F1-Score	68.2%	71.5%	69.7%	81.1%
COVID-19 MCI Triage AUC-ROC	0.742	0.758	0.771	0.858

References

[1] Cheung L, Lau ASM, Lam KF, Ng PY. A review of environmental factors for an ontology-based risk analysis for pandemic spread. 2024 April; 4(4):466–480, doi: 10.3390/covid4040031.

[2] Singh RK, Rani M, Bhagavathula AS, Sah R, Rodriguez-Morales AJ, Kalita H, Nanda C, Sharma S, Sharma YD, Rabaan AA, Rahmani J, Kumar P. Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model. JMIR Public Health and Surveillance. 2020 May;6(2):e19115, doi: 10.2196/19115.

[3] Yang Z, Zeng Z, Wang K, Wong SS, Liang W, Zanin M, Liu P, Cao X, G Z, Mai Z, Liang J, Liu X, Li S, Li Y, Ye F, Guan W, Yang Y, Li F, Luo S, Xie Y, Liu B, Wang Z, Zhang S, Zhang S, Wang Y, Zhong N, He J. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. Journal of Thoracic Disease. 2020 Mar;12(3):165, doi: 10.21037/jtd.202.

[4] Sahai AK, Rath N, Sood V, Singh MP. ARIMA modelling & forecasting of COVID-19 in top five affected countries. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 2020 Sep–Oct;14(5):1419–1427, doi: 10.1016/j.dsx.2020.07.042.

[5] Huang NE, Qiao F. A data driven time-dependent transmission rate for tracking an epidemic: a case study of 2019-nCoV. Sci Bull (Beijing). 2020 Feb;65(6):425-427, doi: 10.1016/j.scib.2020.02.005.

[6] Agrawal A, Bhardwaj R. Probability of COVID-19 infection by cough of a normal person and a super-spreader. Physics of Fluids. 2021 Mar;33(3), doi: 10.1063/5.0041596.

[7] Shinde GR, Kalamkar AB, Mahalle PN, Dey N, Chaki J, Hassanien AE. Forecasting models for coronavirus disease (COVID-19): a survey of the state-of-the-art. SN Computer Science. 2020 Jun;1(4):197, doi: 10.1007/s42979-020-00209-9.

[8] Santosh KC. COVID-19 prediction models and unexploited data. Journal of Medical Systems. 2020 Aug;44(9):170, doi: 10.1007/s10916-020-01645-z.

[9] Xu L, Magar R, Farimani AB. Forecasting COVID-19 new cases using deep learning methods. Computers in Biology and Medicine. 2022 May;144:105342, doi: 10.1016/j.compbiomed.2022.105342.

[10] Chimmula VKR, Zhang L. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons & Fractals. 2020 June;135:109864, doi: 10.1016/j.chaos.2020.109864.

[11] Fong SJ, Li G, Dey N, Crespo RG, Herrera-Viedma E. Finding an accurate early forecasting model from small dataset: a case of 2019-NCOV novel coronavirus outbreak. arXiv preprint. 2020 Mar; 2003.10776, doi: 10.48550/arXiv.2003.10776.

[12] Moftakhar L, Mozhgan SEIF, Safe MS. Exponentially increasing trend of infected patients with COVID-19 in Iran: a comparison of neural network and ARIMA forecasting models. Iranian Journal of Public Health. 2020 Oct;49(Suppl 1):92, doi: 10.18502/ijph.v49is1.3675.

[13] Naseem M, Akhund R, Arshad H, Ibrahim MT. Exploring the potential of artificial intelligence and machine learning to combat COVID-19 and existing opportunities for LMIC: a scoping review. Journal of Primary Care & Community Health. 2020 Sep;11:2150132720963634, doi: 10.1177/2150132720963634.
[14] Raddad Y, Hasasneh A, Abdallah O, Rishmawi C, Qutob N. Integrating statistical methods and machine learning techniques to analyze and classify COVID-19 symptom severity. Big Data and Cognitive Computing. 2024 Dec;8(12):192, doi: 10.3390/bdcc8120192.
[15] Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, Aviles-Rivero AI, Etmann C, McCague C, Beer L, Weir-McCall JR, Teng Z, Gkrania-Llotsas E, AIX-COVNET, Rudd JHF, Sala E, Schönlieb CB. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nature Machine Intelligence. 2021 Mar;3(3):199–217. doi: 10.1038/s42256-021-00307-0.
[16] Mascaro S, Wu Y, Woodberry O, Nyberg EP, Pearson R, Ramsay JA, Mace AO, Foley DA, Snelling TL, Nicholson AE. Modelling COVID-19 disease processes by remote elicitation of causal Bayesian networks from medical experts. BMC Medical Research Methodology. 2023 Mar;23(1):76, doi: 10.1186/s12874-023-01856-1.
[17] Kitson NK, Constantinou AC, Guo Z, Liu Y, Chobtham K. A survey of Bayesian network structure learning. Artificial Intelligence Review. 2023 Jan;56(8):8721–8814. doi: 10.1007/s10462-022-10351-w.
[18] Rao C, Liu Y. Three-dimensional convolutional neural network (3D-CNN) for heterogeneous material homogenization. Computational Materials Science. 2020 Nov;184:109850, doi: 10.1016/j.commatsci.2020.109850.
[19] Bougourzi F, Dornaika F, Nakib A, Distante C, Taleb-Ahmed A. 2D and 3D CNN-based fusion approach for COVID-19 severity prediction from 3D CT-scans. arXiv preprint. 2023 Mar;2303.08740, doi: 10.48550/arXiv.2303.08740.
[20] Cheung, L. Identifying triage determinants and using a novel Bayesian 3D-CNN model for COVID-19 mass casualty incidents (MCI) triage recommendation. Disaster Medicine and Public Health Preparedness. 2024 Oct;18:e144, doi:10.1017/dmp.2024.224.
[21] Li LQ, Huang T, Wang YQ, Wang ZP, Liang Y, Huang TB, Zhang HY, Sun W, Want Y. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. Journal of Medical Virology. 2020 Mar;92(6):577–583, doi: 10.1002/jmv.25757.
[22] The COVID Tracking Project. (2021). Data download. Retrieved from https://covidtracking.com/data/download.
[23] Thomas H, Noam A, Rafael G, Beatriz K, Anna P, Toby P, Samuel W, Emily C, Laura H, Saptarshi M, Helen T. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nature Human Behaviour. 2021 Mar;5(4):529–538, doi: 10.1038/s41562-021-01079-8.
[24] CRDT Dataset. Retrieved from https://github.com/Ethan-Lzq/Emphasizing-COVID-19-Triage/blob/main/datasets/CRDT%20Data%20-%20CRDT.csv.
[25] All-states-history dataset. Retrieved from https://github.com/Ethan-Lzq/Emphasizing-COVID-19-Triage/blob/main/datasets/all-states-history.csv.
[26] OxCGRT_compact_subnational dataset Retrieved from https://github.com/Ethan-Lzq/Emphasizing-COVID-19-Triage/blob/main/datasets/OxCGRT_compact_subnational_v1.csv
[27] The COVID Tracking Project dataset. Retrieved from 3rd raw data is from github https://github.com/OxCGRT/covid-policy-dataset/blob/main/data/OxCGRT_compact_subnational_v1.csv.
[28] Almulihi QA, Alquraini AA, Almulihi FAA, Alqarni OA, Almulihi SA, Almulihi MA, Alqarni MA, Alqarni AA, Alqarni SA, Alqarni FA. Applications of artificial intelligence and machine learning in emergency medicine triage—a systematic review. Medical Archives. 2024 Mar;78(3):198–206, doi:10.5455/medarh.2024.78.198-206.
[29] Brady WJ, Kurultak K, Koyfman A, Long B. Clinical update on COVID-19 for the emergency clinician: presentation and evaluation. American Journal of Emergency Medicine. 2022;50:43–51, doi: 10.1016/j.ajem.2022.01.028.