2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI)

A Distribution-based Regression for Real-time COVID-19 Cases Detection from Chest X-ray and CT Images

Nuha Zamzami[1], Pantea Koochemeshkian[2], and Nizar Bouguila[2]

[1]Department of Computer Science and Artificial Intelligence, University of Jeddah, Jeddah, Saudi Arabia
[2]Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montreal, QC., Canada
Email: nezamzami@uj.edu.sa,p_kooche@encs.concordia.ca,nizar.bouguila@concordia.ca

Abstract

The novel coronavirus (COVID-19) that started last December in Wuhan, Hubei Province, China has become a serious healthcare threat with over five million confirmed cases in 215 countries around the world as on May 20. The World Health Organization recommends a rapid diagnosis and immediate isolation of suspected cases. Thus, there is an imminent need to develop an automatic real-time detection system as a quick alternative diagnosis option to control the virus spread. In this work, we propose a regression model based on a flexible distribution called shifted-scaled Dirichlet for real-time detection of coronavirus pneumonia infected patient using chest X-ray radiographs. To derive the parameters of our proposed model, we adopt the maximum likelihood method, where we update the parameters based on the stochastic gradient descent. The experimental results demonstrate that our approach is highly effective for detecting COVID-19 cases and understand the infection on a real-time basis with high accuracy up to 97%.

1. Introduction

Coronaviruses (CoV) are a large family of viruses that cause diseases resulting from colds such as the Middle East Respiratory Syndrome (MERS-CoV) and Severe Acute Respiratory Syndrome (SARS-CoV). Coronavirus disease (COVID-19) is a new species discovered in 2019 due to contamination from animals to humans [22]. Signs of infection include respiratory symptoms, fever, cough, and dyspnea, wherein in more severe cases, the infection can cause pneumonia, severe acute respiratory syndrome, septic shock, multi-organ failure, and death [26]. Since the disease has been discovered last December, the World Health Organization (WHO) reported a total of approximately 5,164,133 confirmed cases, among which around 332,706 deaths and 2,059,735 were recovered as on May 2020. Figure 1 illustrates the rapid increase in the number of confirmed COVID-19 cases beginning 30 December 2019 and through 20 May 2020 based on WHO Coronavirus disease (COVID-2019) situation report No.121 on 20 May 2020 [1].

Several challenges are associated with COVID-19 in different disciplines, including health, society, and economy. For instance, the health system has come to the point of collapse, even in many developed countries due to the increasing demand for intensive care units simultaneously [31]. In the current public health emergency, tackling such challenges using real data, tools, and research is extremely significant. One of the significant matters to address is the early diagnosis of the infection in order to ensure receiving the appropriate treatment. In addition, identifying the patients or suspected cases quickly in order to isolate them is crucial to avoid the risk of infecting a larger population given the highly infectious nature of the virus [2].

In this work, we propose a novel regression model based on the shifted-scaled Dirichlet distribution for detecting COVID-19 cases from chest X-ray and CT images of patients which are positive or suspected of COVID-19 or other viral and bacterial pneumonias. The choice of regression techniques is justified by the fact that the regression-based models are highly efficient and scalable and these properties are appropriate for the concerned task. Moreover, given the novelty of the disease, there is very limited expertise in labeling the data specific to this new virus where datasets are just now being identified and annotated. Hence, there are not enough examples to achieve clinically meaningful

DOI 10.1109/IRI49571.2020.00023

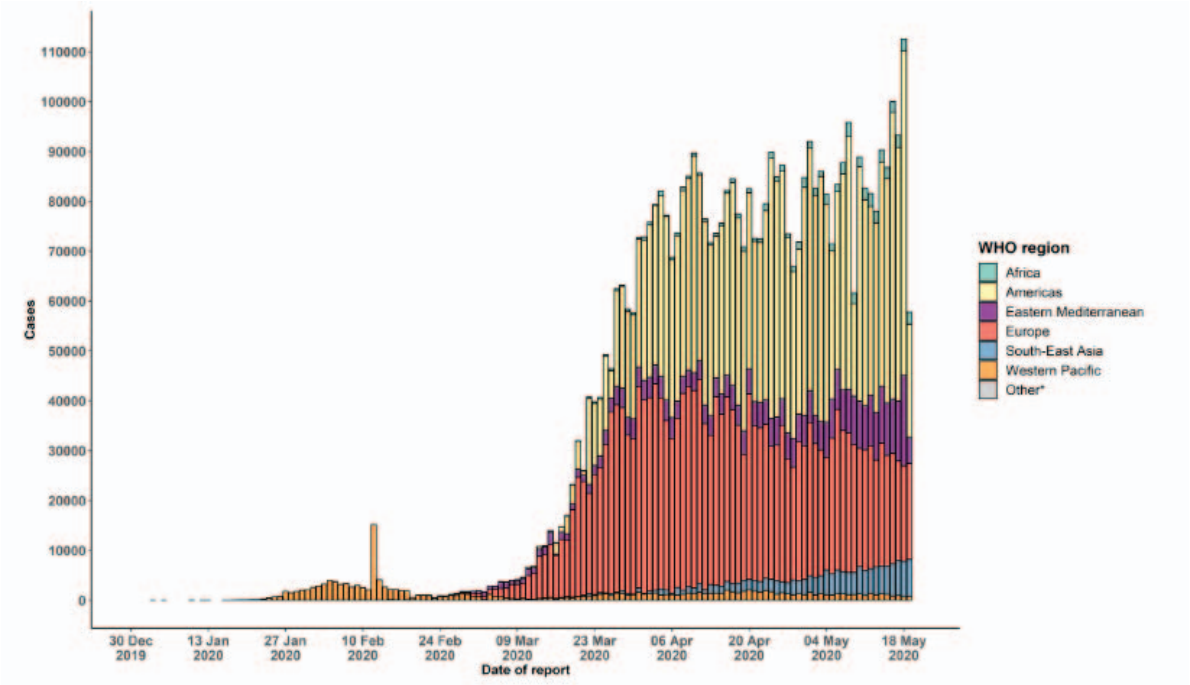

Figure 1. Number of confirmed COVID-19 cases, by date of report and WHO region [1].

learning at this early stage of data collection. It is our hypothesis that developing a model that can be updated incrementally as new data become available will support lifelong learning (*i.e.* the model could be improved each time new data are added).

The remainder of the present paper is organized as follows: Section 2 presents a review of the related literature. Section 3 discusses our shifted-scaled Dirichlet-based regression model with all the details about link functions as well as the approaches for estimating and updating the parameters. In Section 4, we present our experimental results on detecting COVID-19 cases both in offline and online fashions. Finally, we conclude this paper in section 5.

2. Related Works

The Centers for Disease Control and Prevention (CDC) and the WHO have issued interim guidelines to protect the population and to attempt to prevent the further spread of the novel corona virus from infected individuals. After the increased number of deaths, it becomes imperative that we develop novel models to attempt to control the rapidly spreading of the COVID-19 virus by reducing the time needed to identify positive or suspected cases and their rapid isolation. Artificial intelligence and deep learning techniques have emerged as powerful tools to transform medical care. For instance, several studies have promoted disease detection through artificial intelligence models and achieve superb results (e.g. [24, 34]). Such techniques can be useful tools in assisting diagnoses and decision making in treatment of COVID-19 [35].

As we enter the second quarter of the COVID-19 pandemic, few efforts to use of artificial intelligence based techniques for identifying the cases have been proposed. Among the attempts, the authors in [36] proposed to use a machine learning algorithm to improve possible COVID-19 case identification more quickly using a mobile phone–based web survey. Moreover, an automated natural language processing system using deep learning techniques has been proposed to extract clinically relevant information from massive electronic health records [25]. Furthermore, in [31], and [7] different convolutional neural network based models have been proposed for detecting coronavirus pneumonia infected patient using chest X-ray radiographs.

In this work, we propose a regression model to detect COVID-19 cases from chest x-ray and CT images. The main idea is to use a transformation function to link the linear or non-linear predictor to the mean response, to ensure that the mean is from an acceptable range. We propose to present the features as compositional data where vectors of parts of some whole which carry relevant information. That is, images are represented as proportions or percentages, *i.e.* normalized Histogram of Oriented Gradients (HOG), which are subject to a constant sum, , *i.e.*, $= 1$ or $= 100$. Their sample space is then represented by the

simplex. Such data occur in many applied fields: from geology and biology to election forecast, from medicine and psychology to economic studies. Because of the ineligibility of Gaussian distribution for real data, researchers tend to find a better distribution to replace it for better modelling [37, 17]. Here we try to find a better distribution for using it in the regression model. Historically, there are two different approaches to regression models that are commonly used in case of a compositional response variable with a system of covariates: Simplicial regression and Dirichlet regression [29]. The later one assumes that the response variable follows a Dirichlet distribution whose parameters are a log-linear function of a set of covariates [19, 21]. Although Dirichlet is widely used for modeling compositional data [10, 18, 9], researchers proposed to use some generalizations of the Dirichlet to improve the model flexibility and overcome the negative covariance and equal-confidence constraints of the Dirichlet [33]. Other distribution-based regression models for compositional data that have been proposed in the literature with excellent results include the generalized Dirichlet, Beta-Liouviel, and scaled Dirichlet [5, 6, 29].

The shifted-scaled Dirichlet (SSD) is a further generalization of the scaled one obtained after applying the perturbation and powering operations to a Dirichlet random composition [30]. By introducing another set of parameters, we can acquire many useful probability models. The shifted scaled Dirichlet, subsequently, keeps $2D + 1$ degrees of freedom, which grants it the flexibility for diverse real data applications [32, 20]. Recently, the shifted-scaled Dirichlet has shown to be a competitive modeling approach with higher flexibility and capability for categorizing compositional data [3, 4, 40].

3. The Proposed Approach

In this section, we present our shifted-scaled Dirichlet-based approach for COVID-19 cases detection. Given a set of chest X-ray and CT images $\{(X_i = [x_{i1}, x_{i2}, \ldots, x_{iD}]^T, y_i)\}$ with $i = 1, 2, \ldots, N$, each y_i can be described by a corresponding random variable Y_i, where $Y_i \sim \mathcal{SSD}(\alpha, \rho, \tau)$. $\mathcal{SSD}(\alpha, \rho, \tau)$ is the shifted-scaled Dirichlet distribution with the shape parameter α, the location parameter ρ and real scalar that controls the density spread τ. $Y_1, Y_2, \ldots, Y_N$ are i.i.d. random variables. In practice, we try to observe the distribution fitting results for the viral and bacterial pneumonias of each X-ray or CT image, and then the model will be updated incrementally by treating the new coming images as soon as it arrives in a temporal sequence.

3.1. Shifted-scaled Dirichlet-Based Model

The probability density function of a random vector of proportional data $\mathbf{X}_i = (x_{i1}, \ldots, x_{iD})$, follows a shifted-scaled Dirichlet distribution with parameters $\alpha = (\alpha_1, ..., \alpha_D) \in R_+^D, \rho = (\rho_1, \ldots, \rho_D) \in S^D, \tau \in R_+$, is given by [30] :

$$\mathcal{SSD}(\mathbf{X}|\theta) = \frac{\Gamma(\alpha_+)}{\prod_{d=1}^{D}\Gamma(\alpha_d)} \frac{1}{\tau^{D-1}} \frac{\prod_{d=1}^{D} \rho_d^{-(\alpha_d/\tau)} x_d^{(\alpha_d/\tau)-1}}{\Big(\sum_{d=1}^{D}\Big(\frac{x_d}{\rho_d}\Big)^{(1/\tau)}\Big)^{\alpha_+}} \tag{1}$$

where Γ is the Gamma function, and $\alpha_+ = \sum_{d=1}^{D} \alpha_d$. The link between the infection y_i and the features extracted from the X-ray or CT image x_i is based on relating the regression parameters to covariates. The inverse of any cumulative distribution function corresponding to a continuous distribution is called the link function [8, 41]. The parameters of Shifted-scaled Dirichlet distribution can be related to the m-dimensional covariates to form a regression line as follows:

$$\begin{aligned} \alpha_d &= g_1(\alpha_d x_1 + \alpha_d x_2 + \cdots + \alpha_d x_N),\ d = 1, \ldots, D \\ \rho_d &= g_2(\rho_1 x_1 + \rho_2 x_2 + \cdots + \rho_d x_N),\ d = 1, \ldots, D \\ \tau &= g_3(\tau x_1 + \tau x_2 + \cdots + \tau x_N) \end{aligned} \tag{2}$$

The link function $g(.)$ is selected as $g(\mu_i) = X_i^T\theta$ and $i = 1, \ldots, N$, where μ_i is the mean of X_i, and θ is a vector of the regression parameters. We have adopted the logit link to ensure that the expected value of Y_i is positive since we encode each type of infection with a positive value, such that:

$$logit(\mu_i) = \log\Big(\frac{\mu_i}{1-\mu_i}\Big) \tag{3}$$

Thus, we have the following link functions for the Shifted-scaled Dirichlet distribution:

$$\begin{aligned} g_1(\mu_i) &= X_i^T\alpha_d \\ g_2(\mu_i) &= X_i^T\rho_d \\ g_3(\mu_i) &= X_i^T\tau \end{aligned} \tag{4}$$

If we consider the final regression model to be a censored linear regression, the predicted target value $\hat{Y}$ is:

$$\hat{Y} = h_\eta X = \eta^T X \tag{5}$$

For our model, regression coefficient is $\eta = \tau\rho\alpha_d$.and $d = 1, \ldots, D$ is the dimension of the response vector. Next, we present how to derive the regression parameters based on a set of X-ray or CT images.

3.2. Deriving and Updating the Parameters

In this work, the simulation is based on two phases. In the first phase, the model is build from a training set which is considered as an offline phase, then each new x-ray or CT-image coming will be presented to the model as soon as they arrive and the model is updated accordingly in a real-time basis.

The offline phase: Here we utilize maximum likelihood estimation to derive a suitable regression parameters based on a set of independent x-ray images $\mathcal{X} = \{X_1, X2, \ldots, X_N\}$. Therefore, the complete log-likelihood is as follows:

$$p(\mathcal{X}|\Theta) = \prod_{i=1}^{N} \mathcal{SSD}(\mathbf{X}_i|\theta) \tag{6}$$

Obtaining the parameter estimates θ that maximizes the likelihood of the shifted-scaled Dirichlet-based regression model is based on taking the derivative of the complete log-likelihood function, and find θ that when the derivative is equal to zero. Thus, to derive the suitable parameters, the problem can be represented as follows:

$$\Theta^* = \arg\max_{\Theta} \mathcal{L}(\mathcal{X}, \Theta) \tag{7}$$

The first derivative with respect to α_d, ρ_d, and τ are calculated respectively in Eqs.(8), (9), (10) considering ρ_d constraints ($0 \leq \rho_d \leq 1\,; \sum_{d=1}^{D} \rho_d = 1$).

$$\begin{aligned}\frac{\partial \mathcal{L}(\mathcal{X}|\Theta)}{\partial \alpha_d} = \sum_{i=1}^{N} g_1'(\mu_i)\Big(\Psi(\alpha_+) - \Psi(\alpha_d) + \\ \frac{\log(X_{id}) - \log(\rho_d)}{\tau} - \log(\sum_{d=1}^{D}(\frac{x_{id}}{\rho_d})^{\frac{1}{\tau}})\Big)\end{aligned} \tag{8}$$

while Ψ is the digamma function.

$$\rho_d = \frac{\sum_{i=1}^{N} g_2'(\mu_i) \frac{\alpha_+ x_{id}}{\tau \rho_d \sum_{d=1}^{D} \frac{x_{id}}{\rho_d}} - \frac{\alpha_d}{\tau}}{\sum_{i=1}^{N} g_2'(\mu_i) \sum_{d=1}^{D} \frac{\alpha_+ x_{id}}{\tau \rho_d \sum_{d=1}^{D} \frac{x_{id}}{\rho_d}} - \sum_{d=1}^{D} \frac{\alpha_d}{\tau}} \tag{9}$$

$$\begin{aligned}\tau = \sum_{i=1}^{N} g_3'(\mu_i) \frac{\sum_{d=1}^{D} \alpha_d \Big(\log(\rho_d) - \log(x_{id})\Big)}{D-1} \\ + \sum_{i=1}^{N} g_3'(\mu_i) \frac{\alpha_+ \log \sum_{d=1}^{D} \frac{x_{id}}{\rho_d}}{D-1}\end{aligned} \tag{10}$$

The reader may realize that Eq.(8) for α parameter is a non-linear function. Hence, there is no closed-form solution for it. This issue leads to the necessity of an optimization technique where we deploy the Newton Raphson method that allows fast convergence [23]. The Newton-Raphson method for $\boldsymbol{\alpha}$ parameter is an iterative procedure expressed as:

$$\alpha^{new} = \alpha^{old} - H(\alpha)^{-1} G \tag{11}$$

where $\boldsymbol{H}$ is the Hessian matrix that includes the second derivatives vectors, and G is the gradient that holds the first derivatives vectors. Note that, the Hessian matrix should be transformed to its inverse before it can be calculated in the Newton-Raphson maximization step. By calculating the second and mixed derivatives with respect to α_d, we obtain,

$$\frac{\partial^2 \mathcal{L}(\mathcal{X}|\Theta)}{\partial^2 \alpha_d^2} = \sum_{i=1}^{N} g_1''(\mu_i)\Big(\Psi'(\alpha_+) - \Psi'(\alpha_d)\Big) \tag{12}$$

$$\frac{\partial^2 \mathcal{L}(\mathcal{X}|\Theta)}{\partial^2 \alpha_{d_1} \alpha_{d_2}} = \sum_{i=1}^{N} g_1''(\mu_i)\Psi'(\alpha_+) \tag{13}$$

where Ψ' is the trigamma function.

The online phase: After building the model and estimated its parameter based on the available set of images, we need to update the model by treating each newly coming x-ray or CT-image as soon as it becomes available. Let us assume that at time $t+1$ a new image $\mathbf{X}_{N+1}$ is presented to the model and the estimated parameters, thus, should be updated incrementally considering the new data. For updating the parameters, following the practice in [42], we utilize gradient ascent [39] where the parameters are derived iteratively. For each iteration, we first derive one parameter and fixing the other two. The iteration can be represented as follows:

$$\alpha_d^{(t+1)} = \alpha_d^{(t)} + \delta_N \frac{\partial \mathcal{L}(\mathcal{X}|\Theta)}{\partial \alpha_d} \tag{14}$$

$$\rho_d^{(t+1)} = \rho_d^{(t)} + \delta_N \frac{\partial \mathcal{L}(\mathcal{X}|\Theta)}{\partial \rho_d} \tag{15}$$

$$\tau^{(t+1)} = \tau^{(t)} + \delta_N \frac{\partial \mathcal{L}(\mathcal{X}|\Theta)}{\partial \tau} \tag{16}$$

where δ_N is a sequence of positive number that decreases to zero chosen to be $\delta_N = 1/(N+1)$ [39, 11].

3.3. Initial Parameter Settings

To boost the computation for deriving the parameters in our shifted-scaled Dirichlet-based regression model, we try to make initial parameters close to the value of convergence, and then the number of iterations can be reduced to achieve a faster converges. For obtaining the initial values of the shape parameter $\alpha^{(0)}$, we apply the method of moment [27]. To initialize the location parameter vector $\rho^{(0)}$, we propose to use a constant proportion vector where the dimensions summation for the parameter vector equals to one, *i.e.* $\rho_d = 1/D$ for $d = 1, \ldots, D$. Finally, to initialize the scale parameter vector $\tau^{(0)}$, we used a vector of ones.

4. Performance Evaluation

In this section, we measure the decency of our model performance by testing it on novel Covid-19 datasets and compare it with ordinary linear regression, logistic regression, and Dirichlet regression models. We apply the following criteria for this evaluation:

- **Mean square error (MSE)**: This measure is the deviation of predicted value from true value of Y and defined as follows [14]:

$$MSE = \frac{1}{N}\sum_{n=1}^{N} \mid \hat{Y} - Y \mid^2 \tag{17}$$

where N is the number of samples, and $\hat{Y}$ is the predicted value. In this criterion, larger-value deviations will be imprisoned as all errors are squared; therefore the smaller values indicate better performance.

- **Akaike information criterion (AIC)**: The AIC [13] is a sample prediction error estimator and therefore the relative accuracy of statistical models for a given data set. Given a collection of models for the data, AIC estimates the quality of each model. Thus, AIC provides a means for model selection. A lower AIC score is better [12]:

$$AIC = -2\mathcal{L}_n + 2N \tag{18}$$

where N is the number of data points, and $\mathcal{L}_n$ is the log-likelihood of the complete data.

- **Bayesian information criterion (BIC)**: The low BIC[12] value means low test error. When we try to train our model it has very high chance that it has some underfitted over overfitted models [38] and these methods will induce some noise and bias to counter that and due to such methods we can able to reduce test error [12], where N is the number of data points and D is the dimension of the data:

$$BIC = -2\mathcal{L}_n + N\log(D) \tag{19}$$

- **Accuracy**: Accuracy is the last criterion which is the average of difference between expected variables and real dependent variables. Here the dependent variable is the type of infection that the patient has. As compared to three other metrics, greater values of accuracy are better:

$$Accuracy = \Big(1 - \frac{\mu\big(\mid \hat{Y} - Y \mid}{\mu \quad \mid Y \mid}\Big) \times 100 \tag{20}$$

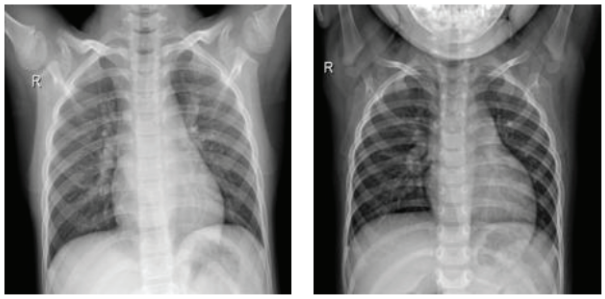

Figure 2. Representative normal chest X-ray images.

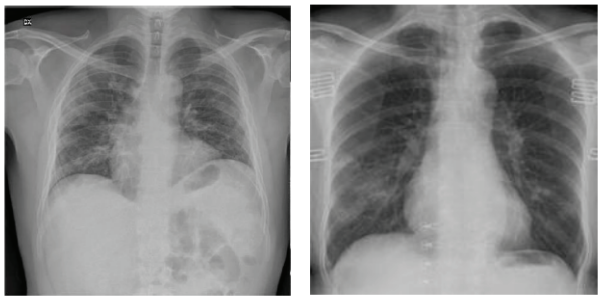

Figure 3. Representative chest X-ray images of COVID-19 patients.

4.1. Dataset Description

In this work, we have used a public database of pneumonia cases with chest X-ray or CT images, specifically COVID-19 cases, as well as the Middle East respiratory syndrome (MERS), severe acute respiratory syndrome (SARS), and acute respiratory distress syndrome (ARDS). The dataset has been obtained from the open-source GitHub repository shared by Dr. Joseph Cohen [15]. All images in this dataset were resized to 64x64 pixel size. In Figure (2) and Figure (3), representative chest X-ray images of normal and COVID-19 patients are given, respectively. With this dataset we want to predict the type of infection, allowing a physician to plan ahead the appropriate treatment and/or advice the immediate isolation of the suspicious COVID-19 cases to control its spread.

We use Histogram of Oriented Gradients (HOG) [16] as a feature descriptor technique. HOG decomposes an image into windows, each window into small squared 8×8 cells, where in each cell, a HOG that contains two values (magnitude and direction) is computed at each pixel, and represented using a 9-bin histogram. It can be stored as an array of nine numbers where the histogram ranges from 0

Table 1. Models performance comparison for COVID-19 detection in the considered dataset ($D = 36$).

Model	Performance metrics			Accuracy
	MSE	AIC	BIC	
Linear Regression	1.830	-51777894.165	-51776731.77	87.09%
Logistic Regression	1.343	-93888945.425	-98537360.245	90.06%
Dirichlet Regression	2.37	-97772.07	-97647.84	81.14%
Shifted-scaled Dirichlet Regression	1.55e+40	-76549959.1921	-76549834.95	95.40%

Table 2. Models performance comparison for COVID-19 detection in the considered dataset ($D = 324$).

Model	Performance metrics			Accuracy
	MSE	AIC	BIC	
Linear Regression	1.196e+5	-10250398.02	-10276845.41	89.00%
Logistic Regression	1.187e+7	-61704206.78	-61039604.19	91.30%
Dirichlet Regression	2.77	-1329578.08	-1328459.95	87.14%
Shifted-scaled Dirichlet Regression	1.24e+17	-100151479.62	-100152597.762	96.20%

to 180, so there are 20 per bin. The result is then normalized by using a block-wise pattern, so they are not affected by any variations (*e.g.*, lighting), and finally, a descriptor for that cell is returned [28]. The number of dimensions for the HOG descriptors depends on the selected cell size. In our experiments, we have chosen 2 different cell sizes such that we represent the images with 36 and 324-dimensional vectors.

4.2. Evaluation of Offline Training and Real-time Prediction Efficiency

The prediction results for the considered dataset using the four tested regression models: Linear regression, logistic regression, Dirichlet regression and shifted-scaled Dirichlet regression which are given in Table 1 and Table 2 reported using different performance metrics namely MSE, AIC, BIC and accuracy. According to the results, one may notice that the Dirichlet regression model has the lowest accuracy for both datasets comparing to the other tested models. On the other hand, the proposed shifted-scaled Dirichlet based regression has achieved the best performance according to the prediction accuracy.

To evaluate the proposed online training, we randomly selected 60% observations to initialize, where we updated the regression parameters using the Maximum Likelihood Estimate (MLE) algorithm, and later we used the proposed online algorithm, where we insert a new data each time until the end. Table 3 illustrates robustness of our proposed model where the accuracy of prediction is reported each time we reach 20% of remaining data. Figure 4 presents the accuracy (%) of the model each time a new feature vector is inserted during the algorithms running, where we update the model according to the proposed online algorithm and re-evaluate the accuracy of the whole model. We can see from the Figure that as more images are inserted, the models performance has been improved. This result confirms that the proposed approach supports life-long learning (*i.e.*, the model improves each time new data are added). It is worth mentioning that the proposed model is efficient also in terms of time, where the time required for each update is 0.15 and 1.79 seconds for the dataset with 36 and 324 features, respectively.

5. Conclusion

In this paper, we focus on detecting the COVID-19 cases from chest x-ray and CT images. A shifted-scaled Dirichlet-based censored linear regression model is proposed, where the response variables are compositional and are associated with the shape, scale, and location parameters of the corresponding shifted-scaled Dirichlet distribution. To derive the parameters in the proposed model, we use the maximum likelihood estimation approach, and the parameters are then updated using the gradient ascent during the real-time prediction. In the experiments, the results show superior performance with a prediction accuracy of up to 97% accuracy and an update time as low as 0.15 seconds. This indicates that our proposed regression model is more effective and robust than comparable state-of-the-art approaches.

Table 3. Performance of the proposed online model for COVID-19 detection in the considered dataset.

Proportion of the data in the stream	Accuracy ($D = 36$)	Accuracy ($D = 324$)
1/5	95.87%	95.90%
2/5	94.78%	96.30%
3/5	96.50%	96.10%
4/5	96.30%	96.20%
5/5	97.10 %	96.67%

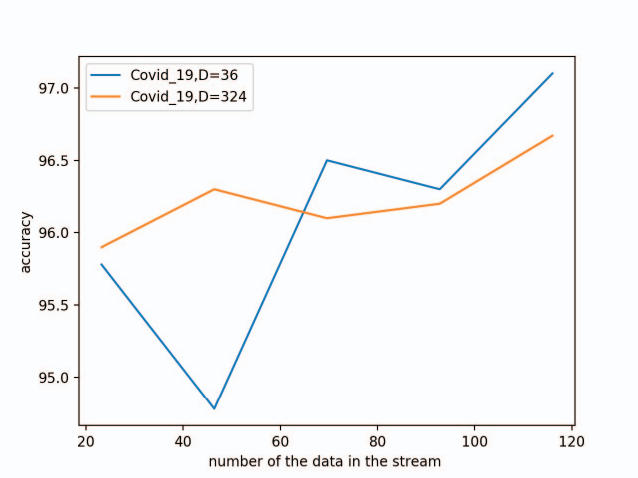

Figure 4. Accuracy of the proposed model each time new images are inserted.

References

[1] COVID-19 situation reports-world health organization. *https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports*, May 20, 2020.

[2] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia. Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: a report of 1014 cases. *Radiology*, page 200642, 2020.

[3] R. Alsuroji, N. Bouguila, and N. Zamzami. Predicting defect-prone software modules using shifted-scaled dirichlet distribution. In *2018 First International Conference on Artificial Intelligence for Industries (AI4I)*, pages 15–18. IEEE, 2018.

[4] R. Alsuroji, N. Zamzami, and N. Bouguila. Model selection and estimation of a finite shifted-scaled dirichlet mixture model. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 707–713. IEEE, 2018.

[5] D. Ankam and N. Bouguila. Generalized dirichlet regression and other compositional models with application to market-share data mining of information technology companies. In *21st International Conference on Enterprise Information Systems (ICEIS)*, 2019.

[6] D. Ankam, N. Bouguila, and M. Amayri. Beta-liouville regression and applications. In *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, pages 1740–1745. IEEE, 2019.

[7] I. D. Apostolopoulos and T. A. Mpesiana. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, page 1, 2020.

[8] C. L. Bayes, J. L. Bazán, C. García, et al. A new robust regression model for proportions. *Bayesian Analysis*, 7(4):841–866, 2012.

[9] N. Bouguila and D. Ziou. On fitting finite dirichlet mixture using ECM and MML. In P. Wang, M. Singh, C. Apté, and P. Perner, editors, *Pattern Recognition and Data Mining, Third International Conference on Advances in Pattern Recognition, ICAPR 2005, Bath, UK, August 22-25, 2005, Proceedings, Part I*, volume 3686 of *Lecture Notes in Computer Science*, pages 172–182. Springer, 2005.

[10] N. Bouguila and D. Ziou. Using unsupervised learning of a finite dirichlet mixture model to improve pattern recognition applications. *Pattern Recognit. Lett.*, 26(12):1916–1925, 2005.

[11] N. Bouguila and D. Ziou. Online clustering via finite mixtures of dirichlet and minimum message length. *Engineering Applications of Artificial Intelligence*, 19(4):371–379, 2006.

[12] K. P. Burnham and D. R. Anderson. Multimodel inference: understanding aic and bic in model selection. *Sociological methods & research*, 33(2):261–304, 2004.

[13] K. P. Burnham, D. R. Anderson, and K. P. Huyvaert. Aic model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral ecology and sociobiology*, 65(1):23–35, 2011.

[14] T. Chai and R. R. Draxler. Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. *Geoscientific model development*, 7(3):1247–1250, 2014.

[15] J. P. Cohen, P. Morrison, and L. Dao. Covid-19 image data collection. *arXiv preprint arXiv:2003.11597*, 2020.

[16] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893. IEEE, 2005.

[17] T. Elguebaly and N. Bouguila. Bayesian learning of generalized gaussian mixture models on biomedical images. In F. Schwenker and N. E. Gayar, editors, *Artificial Neural Networks in Pattern Recognition, 4th IAPR TC3 Workshop, ANNPR 2010, Cairo, Egypt, April 11-13, 2010. Proceedings*, volume 5998 of *Lecture Notes in Computer Science*, pages 207–218. Springer, 2010.

[18] W. Fan, N. Bouguila, and D. Ziou. Variational learning for finite dirichlet mixture models and applications. *IEEE Trans. Neural Networks Learn. Syst.*, 23(5):762–774, 2012.

[19] R. Gueorguieva, R. Rosenheck, and D. Zelterman. Dirichlet component regression and its applications to psychiatric data. *Computational statistics & data analysis*, 52(12):5344–5355, 2008.

[20] R. K. Hankin et al. A generalization of the dirichlet distribution. *Journal of Statistical Software*, 33(11):1–18, 2010.

[21] R. H. Hijazi and R. W. Jernigan. Modelling compositional data using dirichlet regression models. *Journal of Applied Probability & Statistics*, 4(1):77–91, 2009.

[22] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, et al. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *The lancet*, 395(10223):497–506, 2020.

[23] J. Huang. Maximum likelihood estimation of dirichlet distribution parameters. *CMU Technique Report*, 2005.

[24] P. Koochemeshkian, N. Zamzami, and N. Bouguila. Flexible distribution-based regression models for count data: Application to medical diagnosis. *Cybernetics and Systems*, pages 1–25, 2020.

[25] H. Liang, B. Y. Tsui, H. Ni, C. C. Valentim, S. L. Baxter, G. Liu, W. Cai, D. S. Kermany, X. Sun, J. Chen, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. *Nature medicine*, 25(3):433–438, 2019.

[26] E. Mahase. Coronavirus: covid-19 has killed more people than sars and mers combined, despite lower case fatality rate, 2020.

[27] T. Minka. Estimating a dirichlet distribution, 2000.

[28] K. Mizuno, Y. Terachi, K. Takagi, S. Izumi, H. Kawaguchi, and M. Yoshimoto. Architectural study of hog feature extraction processor for real-time object detection. In *2012 IEEE Workshop on Signal Processing Systems*, pages 197–202. IEEE, 2012.

[29] G. Monti, G. Mateu-Figueras, V. Pawlowsky-Glahn, and J. Egozcue. A regression model for compositional data based on the shifted-dirichlet distribution. In *International Workshop on Compositional Data Analysis*, pages 127–143. Springer, 2015.

[30] G. Monti, G. Mateu i Figueras, V. Pawlowsky-Glahn, J. J. Egozcue, et al. The shifted-scaled dirichlet distribution in the simplex. 2011.

[31] A. Narin, C. Kaya, and Z. Pamuk. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *arXiv preprint arXiv:2003.10849*, 2020.

[32] K. W. Ng, G.-L. Tian, and M.-L. Tang. *Dirichlet and related distributions: Theory, methods and applications*, volume 888. John Wiley & Sons, 2011.

[33] B. S. Oboh and N. Bouguila. Unsupervised learning of finite mixtures using scaled dirichlet distribution and its application to software modules categorization. In *IEEE International Conference on Industrial Technology, ICIT 2017, Toronto, ON, Canada, March 22-25, 2017*, pages 1085–1090. IEEE, 2017.

[34] R. Rajalakshmi, R. Subashini, R. M. Anjana, and V. Mohan. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye*, 32(6):1138–1144, 2018.

[35] A. S. S. Rao and M. P. Diamond. Deep learning of markov model-based machines for determination of better treatment option decisions for infertile women. *Reproductive Sciences*, 27(2):763–770, 2020.

[36] A. S. S. Rao and J. A. Vazquez. Identification of covid-19 can be quicker through artificial intelligence framework using a mobile phone–based survey when cities and towns are under quarantine. *Infection Control & Hospital Epidemiology*, pages 1–5, 2020.

[37] A. Sefidpour and N. Bouguila. Spatial color image segmentation based on finite non-gaussian mixture models. *Expert Syst. Appl.*, 39(10):8993–9001, 2012.

[38] S. I. Vrieze. Model selection and psychological theory: a discussion of the differences between the akaike information criterion (aic) and the bayesian information criterion (bic). *Psychological methods*, 17(2):228, 2012.

[39] J.-F. Yao. On recursive estimation in incomplete data models. *Statistics: A Journal of Theoretical and Applied Statistics*, 34(1):27–51, 2000.

[40] N. Zamzami, M. Amayri, N. Bouguila, and S. Ploix. Online clustering for estimating occupancy in an office setting. In *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, pages 2195–2200. IEEE, 2019.

[41] Y. Zhang, H. Zhou, J. Zhou, and W. Sun. Regression models for multivariate count data. *Journal of Computational and Graphical Statistics*, 26(1):1–13, 2017.

[42] W.-Y. Zhu, W.-Y. Shih, Y.-H. Lee, W.-C. Peng, and J.-L. Huang. A gamma-based regression for winning price estimation in real-time bidding advertising. In *2017 IEEE International Conference on Big Data (Big Data)*, pages 1610–1619. IEEE, 2017.