# Covid-19 Hospitalization and ICU Prediction

## An Interactive Visualization Based on Patients' Preconditions


Jae Hyen Park
jpark7ca@gatech.edu

Salim Khan
skhan1975@gatech.edu

Salman A Azmi
sazmi3@gatech.edu

Simon Yee
syee30@gatech.edu

Soniya Goyal
soniyagoyal@gatech.edu

Yunlin Qi
yqi82@gatech.edu


## 1 ABSTRACT


COVID-19 infections have been on the rise since the beginning of this year, and there has been massive pressure on the hospitals and medical facilities to provide appropriate treatment for all patients[3, 5, 12, 16, 18, 25]. Since Intensive Care Unit (ICU) beds and ventilators aren't sufficient for all those who require it, it has become imperative to predict a patient's need for such medical facilities early on based on their preconditions and make them available to as many patients as possible. In this study, we have used some commonly known models[3, 9, 21] and combined them with 2D visualization of the data[19] to understand how patients' preconditions can help us forecast the need for a hospital bed or ICU for them.


## 2 PROBLEM DEFINITION

In this project, we have predicted and visualized hospitalizations and ICU admissions for COVID-19 infected patients, utilizing a preprocessed multi feature COVID-19 patient precondition dataset provided by the Mexican Government[17]. The dataset does not contain any Protected Health Information (PHI) of the patient and a unique random ID is assigned to every row. Our study focuses on feature selection using stepwise regression and Boruta, measures the prediction performance and discriminatory power of logistic regression (LR) and standard neural network (SNN)[3, 9], investigates the impact and correlations of key predictors[12, 18, 25] via 2D interactive data visualizations[4, 15, 19, 23], and finally provides a prediction of hospital beds and ICU capacity required.

## 3 CURRENT PRACTICES AND LIMITATIONS

Current practices are typically carried out in 3 steps - feature selection, model training and visualization. On the feature selection side, it involves determining the presence of categorical, non-categorical information, patient's preconditions, and temporal information[3, 24]. Limitations include imbalanced classes and missing information in the dataset, different range of data values, and omissions of significant predictors[24]. For example, since the number of patients in ICU is far less than the number of patients who are not, we can anticipate skewness in class distribution (ICU vs non-ICU). In fact, non-ICU class makes up a large proportion of our dataset. Because of class imbalance issues, researchers need to carefully account for it and ensure that the minority class is properly handled during model training. On the model side, various classification and probability models[3, 5, 9, 13, 16, 21, 25] were built with frequently-used predictors[12, 14, 18] to predict diagnostic and prognostic risks among various targeted populations[26]. Limitations include insufficiently preprocessed, limited, or outdated dataset, lack of model validation, arbitrary data selection, low accuracy predictions, and arbitrary thresholds in probability models which resulted in misconception and bias risks in the outcome[4–6, 11, 14, 18, 24, 26]. On the data visualization side, techniques such as 2D graphs, matrices for variable correlation, and charts[4, 15, 23] were adopted for visualizing COVID-19 data. A more complex 3D visualization was developed for viewing the WHO data[19]. Limitations include lack of interactive visual techniques for COVID-19 data and some innovative yet vaguely illustrated procedures[4, 15, 22, 23].

## 4 PROPOSED METHOD : INTUITION

### 4.1 Feature Selection

Our feature selection process aims to clean, process, and select the appropriate independent variables to ensure that our models measure real relationships and ensure the highest accuracy possible. **Our innovative process includes traditional data cleaning and preprocessing methods, and more importantly, incorporating medical research and CDC guidelines into our consideration.**

## 4.2 Predictive Model Analysis

We carefully selected two machine learning algorithms for our prediction analysis: a logistic regression model (LR) employed as a baseline and a multilayer standard neural network model (SNN) employed to better capture the nonlinear and complex relationship between various patient's preconditions (features) and outcomes (ICU or hospitalization). Our intuition is that the LR model will act as a single layer neural network with sigmoid nonlinearity that gives us a performance baseline, which can be then compared with the ideally suited SNN model through its multiple layers of interconnected neurons for performance improvement. We expect to better understand the hidden relationships and patterns in the dataset and make more accurate predictions.

To potentially outperform the state of the art, we decided to innovatively implement the algorithms incorporating miscellaneous precautions and improvements to tackle the limitations in current research work. **Our innovative process includes carrying out more comprehensive model validations across different analyzing tools like TensorFlow [2], Scikit-learn [20], Weka [10] and Azure ML Studio [1] to make sure our models are working properly before diving into hyperparameter tuning and metrics evaluation, as well as investigating and minimizing the impact of imbalanced dataset with downsampling, oversampling and initial weights exploration.** Besides, we evaluated our model performances with comprehensive metrics such as area under the receiver operating characteristic curve (AUC)[7, 8] and confusion matrix, in addition to accuracy and precision. Different classification discriminant thresholds were evaluated to determine their influence on prediction quality, thus accurately and reasonably identifying patients' needs.

## 4.3 Data and Model Visualization

We believe the results from our study will help hospital staff and government officials predict hospitalization needs in the right manner. That's why we want to visualize the distribution of data and results from our model both on the training data and test data, as well as relate them back to the preconditions that played a major role in determining the results for a patient. **Through these procedures, we can provide great insights to our stakeholders and help them make the right decisions around the needs of their patients.**

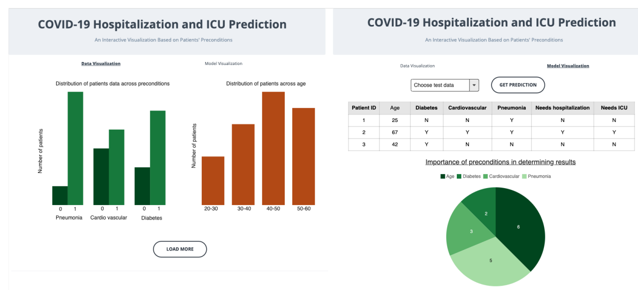


**Figure 1: Visualization prototype**

# 5 PROPOSED METHOD : APPROACH

## 5.1 Feature Selection

The first step in our feature selection process is to ensure that there is no bad data. Examples include checking for instances where the patient is male and is pregnant, or if expired dates (date of death) are incorrectly recorded and are prior to hospitalization dates, as this would imply that an individual died and then got checked into the hospital. The next step is to ensure that our data makes sense for determining whether an individual will be hospitalized and in ICU or not. For example, to ensure we are not measuring randomness, we only kept records of patients who are hospitalized or in the ICU if their COVID results are known and positive.

## 5.2 Predictive Model Analysis

We built our primary model prediction platform and pipelines in a Docker image, which was equipped with multiple data analysis and visualization packages. We used Jupyter Notebooks to carry out various predictive model experiments, where the datasets were processed with Numpy and Pandas, the LR and SNN algorithms were implemented with Tensorflow/Keras and Scikit-learn, and the performance results were visualized with Tensorboard, Matplotlib and Seaborn.

We conducted ICU prediction first using LR and SNN, validated the models in other software, then predicted hospitalization in Jupyter similarly with another dataset. To implement our algorithms, we first imported the preprocessed dataset with selected features, identified features and labels in the dataframe, further processed features into categorical or normalized format, and then split the dataset for training, validation, and testing. Next, we defined our LR and SNN classifiers with pre-configured parameters and hyperparameters. The LR has a single output layer with sigmoid nonlinearity and stochastic gradient descent (SGD) optimizer[2], and the SNN has one dense layer with 32 neurons, one dropout layer to avoid overfitting during training, and one output layer with sigmoid nonlinearity and Adam optimizer[2]. Both models use binary cross-entropy loss to train labeled datasets. Then, we started training the model with training and validation datasets, plotted metric results including loss, AUC, precision, and recall for model complexity analysis. To validate the ICU models, LR and SNN models were built in Weka[10] and Azure ML Studio[1] with the same configurations for performance comparison, where we also used 10-fold cross validation. During the validation phase, exploration of hyperparameter tuning was used to improve prediction performance. After validating our Jupyter Notebook models and finalizing hyperparameters, we evaluated our trained models with the test dataset and interpreted in plots various metrics including precision, AUC[7, 8], and confusion matrix. To investigate the impact of our imbalanced dataset, we also tried down-sampling, oversampling, and initial SNN weights for any potential improvements. The result metrics were then used for the interactive visualization.

## 5.3 Data and Model Visualization

We have built three forms of visualization - first, we visualize the dataset itself to show its distribution across various features as well as derive some intuitions from there. We also display the correlation between these features through graphs to facilitate removal of highly correlated features for the model training. Next, we visualize the machine learning models and their respective results with ROC curves to better understand their performance. Finally, we have built a predictor which, given a test dataset, runs the model and predicts the number of cases for hospitalization and ICU. This also shows the accuracy of our model and how much each of the preconditions contributed to the results.

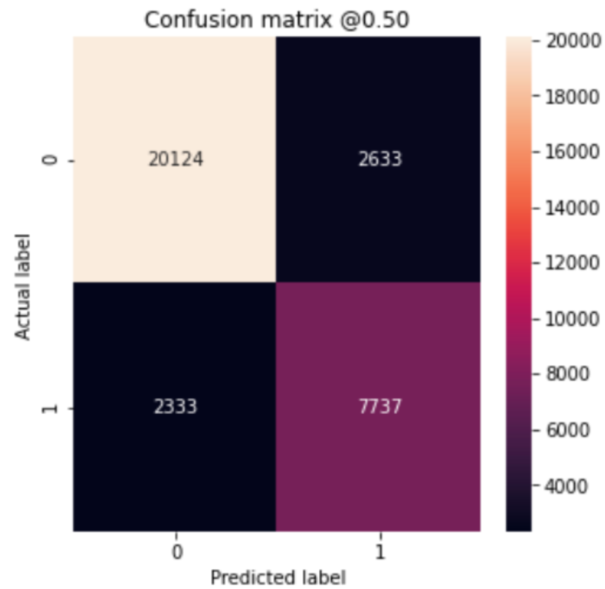


**Figure 2: Confusion Matrix**

We wanted to take these simple visualizations and make them interactive - for example, for the dataset, we have added the capability to view the distribution of the data across any feature for both hospitalization and ICU. For the machine learning models, we have provided various forms of comparison like ROC, loss, precision and recall numbers for both SNN and LR over both ICU and Hospitalization predictions. With the logistic regression model, we have even provided tools to visualize the coefficients of the features and understand which user health condition contributes how much in determining if a patient needs hospitalization or ICU. Finally, our stakeholders will be able to also visualize, given a test set of data, how many hospital and ICU beds are needed for that list of patients.

# 6 EXPERIMENTS AND EVALUATION

## 6.1 Feature Selection

Prior to the model experiment, we excluded irrelevant features in the dataset such as potential contact with COVID-positive individuals, date of hospital admission, date of first symptoms and date of death. Our thought process is that whether someone died or not, as an example, is not predictive of the need for ICU or hospitalization, because it is a result of contracting COVID-19 and might be better used as a response feature rather than predictor.

In our experimentation process, a logistic regression base model was built with a 70/30 split and cross validated with all the features selected. Then we chose the model with the lowest variance inflation factor (VIF) score from the base model, and looked for statistical significance of 95% and multicollinearity when considering a feature for the final model. To diversify our options and validate our base model, we also ran a step regression model to compare the selected features. After consolidating our selected features from each model, the final process was to cross-reference medical research and CDC guidelines to check how accurate our predictors were. Experiments were also carried out using Boruta charts which displayed the importance of each predictor and correlation matrices.

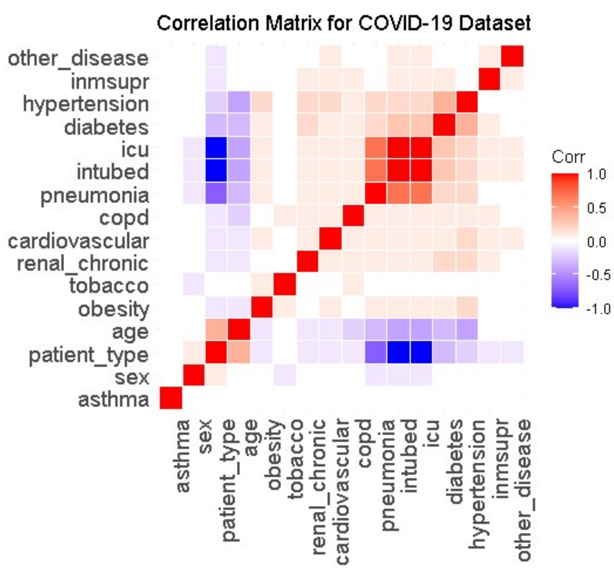


**Figure 3: Correlation Matrix**

After the experimentation process, we started our evaluation by building a simple table that displayed the selected predictors from the aforementioned models (see figure 4), and cross referencing each predictor with medical research and CDC guidelines to ensure that our variable selection process made sense from a practical perspective. Immunocompromised (inmsupr) and obesity features were not selected to be evaluated with the ICU model despite CDC's guideline. It was due to the fact that only about 3.88% of individuals who are in ICU for COVID-19 in our dataset are immunocompromised. Similarly, pregnancy was also not selected due to skewness in data and collinearity with sex. Cardiovascular was correlated with hypertension, therefore, it was dropped from the Hospitalization model. Renal chronic was selected in the ICU models, although it appeared only to have an effect on individuals with preconditions already included in the ICU models. Lastly, since intubation was an intuitive equivalence of ICU admission, it was also excluded from the ICU model. While we were able to determine which features suggested likely relationships, we were unable to quantify the relative importance of those selected features in this process.

| | ICU - CV (95% stat signif.) | ICU - Step Regression | Hospitalization - CV (95% stat signif.) | Hospitalization - Step Regression | Potential Objective Evidence |
|---|---|---|---|---|---|
| sex | | ✓ | ✓ | ✓ | Article |
| pneumonia | ✓ | ✓ | ✓ | ✓ | See Illness Severity section and linked studies |
| age | ✓ | ✓ | ✓ | ✓ | CDC article |
| pregnancy | | | | | An MMWR study suggests that pregnant women with COVID-19 are more likely to be hospitalized and are at increased risk for intensive care unit (ICU) admission and receipt of mechanical ventilation than nonpregnant women. Risk of death is similar for both groups. But much remains unknown |
| diabetes | ✓ | ✓ | ✓ | ✓ | Having type 2 diabetes increases your risk of severe illness from COVID-19. Based on what we know at this time, having type 1 or gestational diabetes may increase your risk of severe illness from COVID-19 |
| copd | | | ✓ | ✓ | See COPD - Having COPD (including emphysema and chronic bronchitis) is known to increase your risk of severe illness from COVID-19. Other chronic lung diseases, such as idiopathic pulmonary fibrosis and cystic fibrosis, may increase your risk of severe illness from COVID-19. |
| asthma | ✓ | | ✓ | ✓ | See Asthma section |
| inmsupr | | | ✓ | ✓ | CDC guidelines and recommendations for the immunocopromised |
| hypertension | | | ✓ | ✓ | CDC - Having any of the following serious heart conditions increases your risk of severe illness from COVID-19: heart failure, coronary artery disease, cardiomyopathies, pulmonary hypertension |
| other_disease | | | ✓ | ✓ | |
| cardiovascular | | ✓ | | | See "Table 7" for takeaways |
| obesity | | | ✓ | ✓ | "Obesity worsens outcomes from COVID" |
| renal_chronic | ✓ | ✓ | ✓ | ✓ | See Chronic kidney disease section |
| tobacco | | | ✓ | ✓ | what is unchallengeable is that cigarette smoke is detrimental for the lungs in several ways, and further studies are needed to clarify the reasons behind the reported low prevalence of current smokers among hospitalized patients with COVID-19 |

**Figure 4: Feature selection**

## 6.2 Predictive Model Analysis

Our primary testbeds include LR and SNN models operating on the preprocessed ICU and hospitalization datasets. Both datasets after data cleaning and feature selection were split into three (3) parts: train (70%), validation (15%) and test (15%). Both datasets were imbalanced, where we had 8.42% (5,667) positive labels for ICU and 30.79% (218,842) positive labels for hospitalization. Therefore, our experiments first tried to address the impact of the imbalanced datasets on the prediction results. To address this, we make the model pay more attention to samples from an under-represented label by implementing both oversampling and undersampling on train dataset in addition to custom class weights for positive labels. Secondly, our experiments aimed to address the impact of various hyperparameters (batch size, epochs, learning rate and momentum) on the result metrics such as AUC. Thirdly, our experiments tried to demonstrate how certain feature combinations affect the prediction results. Lastly, our experiments would also investigate how classification threshold affects the prediction results and what is the reasonable threshold for our predictions.

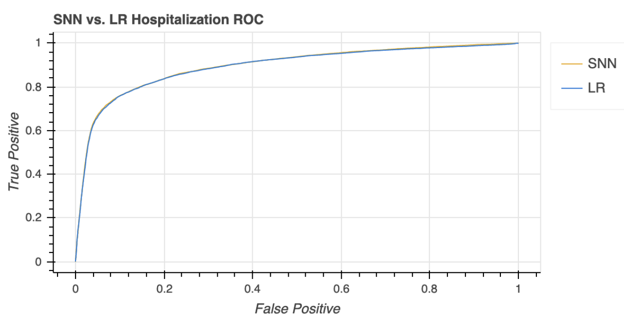


**Figure 5: Hospitalization ROC for SNN vs LR**

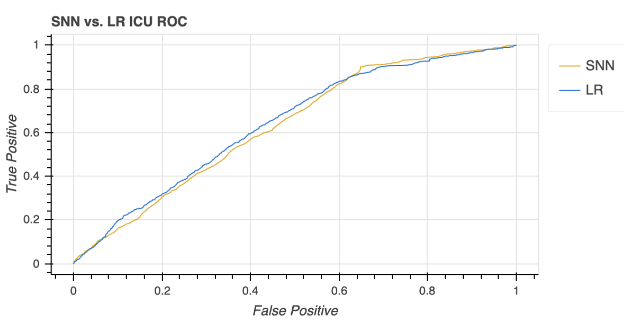


**Figure 6: ICU ROC for SNN vs LR**

Prior to the model experiments, for both ICU and hospitalization predictions, we created two separate train datasets that were randomly resampled by duplicating samples from the minority class (oversampling) and removing samples from the majority class (undersampling). In addition, we carefully adjusted the class weights for positive labels when training the model on the original train dataset. Our experiment to predict a need for ICU using LR and SNN models produced low accuracy results on undersampled and oversampled datasets. We had an average AUC score of 0.65 compared to our expected performance benchmark 0.80. However, we noticed that after adding the previously excluded intubation feature, both models had improved AUC at 0.75. This suggested that chosen combinations of selected features except intubation were not strong indicators of the ICU admission. We were not able to increase the performance of the ICU models with different combinations of feature sets along with hyperparameter adjustment. For the final prediction model, we justified lowering the classification threshold to capture more false positives for ICU needs.

On the other hand, our experiment to predict hospitalization using LR and SNN models produced a 0.89 AUC value with a 85% classification accuracy on the test data at 0.5 classification threshold. We observed that large batch sizes helped models converge faster, which might be attributed to the imbalanced dataset. In addition, SNN with Adam optimizer performed better compared to LR with SGD with fewer iterations (epochs). We had to increase the number of epochs for LR to produce a similar result. The effect of oversampling was basically identical to that of adjusting under-represented labels using custom weights (class weighting). We did not observe any significant improvement with an oversampled dataset compared to a carefully weighted dataset when training the model. During the LR experiments, we have also analyzed the odds ratio to understand the effect of a predictor on the model outcomes. We identified 5 features that were important in predicting hospitalization: pneumonia (OR = 29.93), chronic renal (OR=2.34), age (OR=1.9 for all age groups), diabetes (OR=1.75) and immunosuppression (OR=1.74). For example, the odds of being hospitalized were 30 times higher for patients with pneumonia than those without if all other conditions were considered constant. Here we used the mean value 45 and scale value 16 for the non-categorical age feature just to give some guidance. We also noticed that the first four identified features in the hospitalization prediction were also the strongest ICU predictors from our feature selection process.

## 6.3 Data and Model Visualization

We started with a prototype design of what we wanted to achieve in visualization. Our goal from the outset was to design an understandable and interactive visualization, not just a static set of charts and data. We first designed our prototype online, and in return it guided us throughout the implementation process. Once we started on the actual implementation, we analyzed multiple Python libraries and decided on the one best suited to our needs. It was important to us that the libraries and language chosen supported the ability to run, visualize and compare models. We chose Bokeh due to its native support for interactive and customizable visualizations and its interoperability with other Machine learning based Python packages.

Next, we had to determine our hosting infrastructure. Settling for a local setup is easy, but we determined a website would improve accessibility to the public. We explored options like GitHub pages, Tableau and Heroku and decided to use Heroku because of its simplicity in deploying a Python based website with minimum fees and setup.

We were finally able to build our website which has **3 interactive tabs** as discussed in our original intuition section above. The first tab "Show data distribution" allows for interactive feature visualization across the dataset along with a correlation plot of the preconditions. The second tab, "Show model results" focuses on model output analysis with AUC, precision, loss, recall, and ROC. Finally, the last tab, "Show model predictor" includes prediction of number of hospitalizations given a test data and also focuses on key features contributing towards the estimated counts of Hospitalizations. Our website can be accessed at this url: https://tinyurl.com/covid-visualisation.

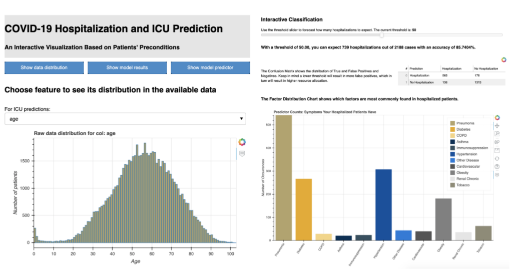


**Figure 7: Visualization website**

## 7 CONCLUSION AND DISCUSSION

Based on the results of this study, we found out that our selected patients' precondition features led to better predictions for hospitalization (AUC = 0.89) than for ICU (AUC = 0.65). No previously selected features (age, pneumonia, diabetes, renal chronic) or additional pre-excluded features (except intubation) in the feature selection process appeared to perform strongly in predicting ICU need for a COVID-19 positive individual, and neither did the sampling strategies discussed in our approach and analysis help. In contrast, five strong indicators/features were identified for hospitalization prediction, four of which coincided with the previously selected ICU predictors, followed by the additional immunosuppression. We concluded a few possibilities for ICU prediction outcome; ranging from having insufficient data, lack of real relationships between the predicted and predictors in the data gathering process, and skewed class labels. In contrast, while hospitalization has a significant performance over ICU, we are cautiously optimistic because explaining all predictors or having all the information may be difficult in practice. Specific predictors, like pneumonia and age appear to play a significant role in hospitalization and we recommend medical professionals/practitioners to be extra cautious when working with older and pneumonia positive patients when no other patient/medical information is immediately available.

We hope our study motivates other researchers to work with new data findings about COVID-19 and further improve the prediction for ICU needs. In addition, the web application we developed for interactive visualization provided us with instant visual feedback on feature distribution, model outcomes, and adjustable classification thresholds. It played a significant role in multiple aspects of our study such as model validation, hyperparameter tuning, and result interpretation. We hope our visualization tools could help hospitals and public policy makers make informed decisions on effectively distributing critical resources during the pandemic. We included the web application in the distribution package to facilitate the use of our visualization tools.

## 8 DISTRIBUTION OF TEAM MEMBER EFFORT

All team members have contributed similar amount of effort.

# REFERENCES


[1] 2010. Microsoft Azure Machine Learning Studio (classic). https://studio.azureml.net/

[2] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. https://www.tensorflow.org/ Software available from tensorflow.org.

[3] Ahmed Allam, Mate Nagy, George Thoma, and Michael Krauthammer. 2019. Neural networks versus Logistic regression for 30 days all-cause readmission prediction. *Scientific Reports* 9 (06 2019), 9277. https://doi.org/10.1038/s41598-019-45685-z

[4] Pradipta Biswas, Kamalpreet Singh Saluja, Somnath Arjun, Lrd Murthy, Gowdham Prabhakar, Vinay Krishna Sharma, and Jeevitha Shree Dv. 2020. COVID-19 Data Visualization through Automatic Phase Detection. *Digit. Gov.: Res. Pract.* 1, 4, Article 25 (July 2020), 8 pages. https://doi.org/10.1145/3411756

[5] Theodora S Brisimi, Tingting Xu, Taiyao Wang, Wuyang Dai, and Ioannis Ch Paschalidis. 2019. Predicting diabetes-related hospitalizations based on electronic health records. *Statistical Methods in Medical Research* 28, 12 (2019), 3667–3682. https://doi.org/10.1177/0962280218810911 PMID: 30474497.

[6] Yongli Cai, Jianjun Jiao, Zhanji Gui, Yuting Liu, and Weiming Wang. 2018. Environmental variability in a stochastic epidemic model. *Appl. Math. Comput.* 329 (2018), 210 – 226. https://doi.org/10.1016/j.amc.2018.02.009

[7] Xue-wen Chen and Michael Wasikowski. 2008. FAST: A Roc-Based Feature Selection Metric for Small Samples and Imbalanced Data Classification Problems. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '08)*. Association for Computing Machinery, New York, NY, USA, 124–132. https://doi.org/10.1145/1401890.1401910

[8] Olga V. Demler, Michael J. Pencina, and Ralph B. D'Agostino Sr. 2011. Equivalence of improvement in area under ROC curve and linear discriminant analysis coefficient under assumption of normality. *Statistics in Medicine* 30, 12 (2011), 1410–1418. https://doi.org/10.1002/sim.4196

[9] Stephan Dreiseitl and Lucila Ohno-Machado. 2002. Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics* 35, 5 (2002), 352 – 359. https://doi.org/10.1016/S1532-0464(03)00034-0

[10] Eibe Frank, Mark Hall, Geoffrey Holmes, Richard Kirkby, Bernhard Pfahringer, Ian H. Witten, and Len Trigg. 2010. *Weka-A Machine Learning Workbench for Data Mining*. Springer US, 1269–1277. https://doi.org/10.1007/978-0-387-09823-4_66

[11] Raphaël Gazzotti, Catherine Faron-Zucker, Fabien Gandon, Virginie Lacroix-Hugues, and David Darmon. 2019. Injecting Domain Knowledge in Electronic Medical Records to Improve Hospitalization Prediction. In *The Semantic Web*, Pascal Hitzler, Miriam Fernández, Krzysztof Janowicz, Amrapali Zaveri, Alasdair J.G. Gray, Vanessa Lopez, Armin Haller, and Karl Hammar (Eds.). Springer International Publishing, Cham, 116–130. https://doi.org/10.1007/978-3-030-21348-0_8

[12] Wei-jie Guan, Zheng-yi Ni, Yu Hu, Wen-hua Liang, Chun-quan Ou, Jian-xing He, Lei Liu, Hong Shan, Chun-liang Lei, David S.C. Hui, Bin Du, Lan-juan Li, Guang Zeng, Kwok-Yung Yuen, Ru-chong Chen, Chun-li Tang, Tao Wang, Ping-yan Chen, Jie Xiang, Shi-yue Li, Jin-lin Wang, Zi-jing Liang, Yi-xiang Peng, Li Wei, Yong Liu, Ya-hua Hu, Peng Peng, Jian-ming Wang, Ji-yang Liu, Zhong Chen, Gang Li, Zhi-jian Zheng, Shao-qin Qiu, Jie Luo, Chang-jiang Ye, Shao-yong Zhu, and Nan-shan Zhong. 2020. Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine* 382, 18 (2020), 1708–1720. https://doi.org/10.1056/NEJMoa2002032

[13] Julien I.E. Hoffman. 2015. Chapter 33 - Logistic Regression. In *Biostatistics for Medical and Biomedical Practitioners*, Julien I.E. Hoffman (Ed.). Academic Press, 601 – 611. https://doi.org/10.1016/B978-0-12-802387-7.00033-0

[14] Dong Ji, Dawei Zhang, Jing Xu, Zhu Chen, Tieniu Yang, Peng Zhao, Guofeng Chen, Gregory Cheng, Yudong Wang, Jingfeng Bi, Lin Tan, George Lau, and Enqiang Qin. 2020. Prediction for Progression Risk in Patients With COVID-19 Pneumonia: The CALL Score. *Clinical Infectious Diseases* 71, 6 (04 2020), 1393–1399. https://doi.org/10.1093/cid/ciaa414

[15] B. Liu, Y. Peng, Y. Zou, J. Wang, and T. Jiang. 2015. Web-Weka Meets D3.js in Web Based Medical Data Mining. In *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, Vol. 2. 180–183.

[16] Seyed M. Moghadas, Affan Shoukat, Meagan C. Fitzpatrick, Chad R. Wells, Pratha Sah, Abhishek Pandey, Jeffrey D. Sachs, Zheng Wang, Lauren A. Meyers, Burton H. Singer, and Alison P. Galvani. 2020. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proceedings of the National Academy of Sciences* 117, 16 (2020), 9122–9126. https://doi.org/10.1073/pnas.2004064117

[17] Tanmoy Mukherjee. 2019. COVID-19 patient pre-condition dataset. https://www.kaggle.com/tanmoyx/covid19-patient-precondition-dataset?select=covid.csv

[18] Paulo Nogueira, Miguel De Araújo Nobre, Andreia Costa, Ruy Ribeiro, Cristina Furtado, Leonor Bacelar-Nicolau, Catarina Camarinha, Márcia Luís, Ricardo Silva Macedo Abrantes, and António Carneiro. 2020. The Role of Health Preconditions on COVID-19 Deaths in Portugal: Evidence from Surveillance Data of the First 20293 Infection Cases. *Journal of Clinical Medicine* 9 (07 2020), 2368. https://doi.org/10.3390/jcm9082368

[19] Dale Patterson. 2016. Interactive 3D Web Applications for Visualization of World Health Organization Data. In *Proceedings of the Australasian Computer Science Week Multiconference (ACSW '16)*. Association for Computing Machinery, New York, NY, USA, Article 76, 8 pages. https://doi.org/10.1145/2843043.2843477

[20] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 85 (2011), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

[21] Benjamin Quost, Thierry Denoeux, and Shoumei Li. 2017. Parametric classification with soft labels using the Evidential EM algorithm . *Advances in Data Analysis and Classification* 11, 4 (2017), 659–690. https://doi.org/10.1007/s11634-017-0301-2

[22] P. Rheingans and M. DesJardins. 2000. Visualizing high-dimensional predictive model quality. In *Proceedings Visualization 2000. VIS 2000 (Cat. No.00CH37145)*. 493–496.

[23] Michael Schroeder, Abel Gonzalez-Perez, and Nuria López-Bigas. 2013. Visualizing multidimensional cancer genomics data. *Genome Medicine* 5 (01 2013), 9. https://doi.org/10.1186/gm413

[24] Yinxiaohe Sun, Vanessa Koh, Kalisvar Marimuthu, Oon Tek Ng, Barnaby Young, Shawn Vasoo, Monica Chan, Vernon J M Lee, Partha P De, Timothy Barkham, Raymond T P Lin, Alex R Cook, Yee Sin Leo, and National Centre for Infectious Diseases COVID-19 Outbreak Research Team. 2020. Epidemiological and Clinical Predictors of COVID-19. *Clinical Infectious Diseases* 71, 15 (03 2020), 786–792. https://doi.org/10.1093/cid/ciaa322

[25] Salomón Wollenstein-Betech, Christos G. Cassandras, and Ioannis Ch. Paschalidis. 2020. Personalized predictive models for symptomatic COVID-19 patients using basic preconditions: Hospitalizations, mortality, and the need for an ICU or ventilator. *International Journal of Medical Informatics* 142 (2020), 104258. https://doi.org/10.1016/j.ijmedinf.2020.104258

[26] Laure Wynants, Ben Van Calster, Gary S Collins, Richard D Riley, Georg Heinze, Ewoud Schuit, Marc M J Bonten, Darren L Dahly, Johanna A A Damen, Thomas P A Debray, Valentijn M T de Jong, Maarten De Vos, Paula Dhiman, Maria C Haller, Michael O Harhay, Liesbet Henckaerts, Pauline Heus, Nina Kreuzberger, Anna Lohmann, Kim Luijken, Jie Ma, Glen P Martin, Constanza L Andaur Navarro, Johannes B Reitsma, Jamie C Sergeant, Chunhu Shi, Nicole Skoetz, Luc J M Smits, Kym I E Snell, Matthew Sperrin, René Spijker, Ewout W Steyerberg, Toshihiko Takada, Ioanna Tzoulaki, Sander M J van Kuijk, Florien S van Royen, Jan Y Verbakel, Christine Wallisch, Jack Wilkinson, Robert Wolff, Lotty Hooft, Karel G M Moons, and Maarten van Smeden. 2020. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* 369 (2020). https://doi.org/10.1136/bmj.m1328