

# USE OF MACHINE LEARNING TO ASSESS THE INFLUENCE OF COMORBIDITIES ON THE RISK OF DEATH FROM COVID-19

*Taian Fonseca Feitosa*<sup>1</sup>

<sup>1</sup> Universidade Federal da Bahia

**Abstract:** The objective of this work is to use machine learning algorithms to evaluate the influence of comorbidities on the risk of death from COVID-19. The data of patients with Severe Acute Respiratory Syndrome (SARS) were extracted from OPENDATASUS, the portal of the Brazilian Ministry of Health, on October 12, 2022. Seven algorithms were used to evaluate the data. Some of them, such as the Logistic Regression algorithm, generate coefficients that have mathematical interpretation, which were used to assess and compare the degree of risk of death in patients with COVID-19, with and without comorbidities.



taianmeca@gmail.com



4th Conference on  
**Statistics and  
Data Science**  
Salvador, Brazil (online)  
December 1-3, 2022

# Problem and Data

This paper tries to classify the comorbidities according to their risk of death. The method chosen in this work is to use some machine learning algorithms and extract information from the generated models.

The data about hospitalization with SARS is available at OPENDATASUS. This data were used to train the models.

- **Postpartum**
- **Chronic Cardiovascular Disease**
- **Chronic Hematologic Disease**
- **Down's Syndrome**
- **Chronic Liver Disease**
- **Asthma**
- **Diabetes Mellitus**
- **Chronic Neurological Disease**
- **Other Chronic Pneumopathy**
- **Immunodeficiency or Immunodepression**
- **Chronic Kidney Disease**
- **Obesity**



taianmeca@gmail.com



4th Conference on  
**Statistics and  
Data Science**  
Salvador, Brazil (online)  
December 1-3, 2022

# Methods

Model	Training time (s)	Models trained	Time per model (s)	F1
Multilayer Perceptron	1759	18	97.7	0.5969435474385554
Linear Support Vector Machine	813	27	30.1	0.5964053797731378
Gradient Boosted Tree	2082	27	77.1	0.5948102650730435
Random Forest	1002	45	22.3	0.5921345034100683
Factorization Machines	1317	135	9.8	0.5828965711040142
Logistic Regression	1140	135	8.4	0.5712969341163462
Naive Bayes	106	3	35.3	0.5434226785062449

- Some models have been trained to try to predict the outcome of internalization based on comorbidities.
- Some models, such as Logistic Regression, have interpretable coefficients that allow us to analyze the impact of each comorbidity making them a good model to use, even if the model has a slightly lower F1 score.



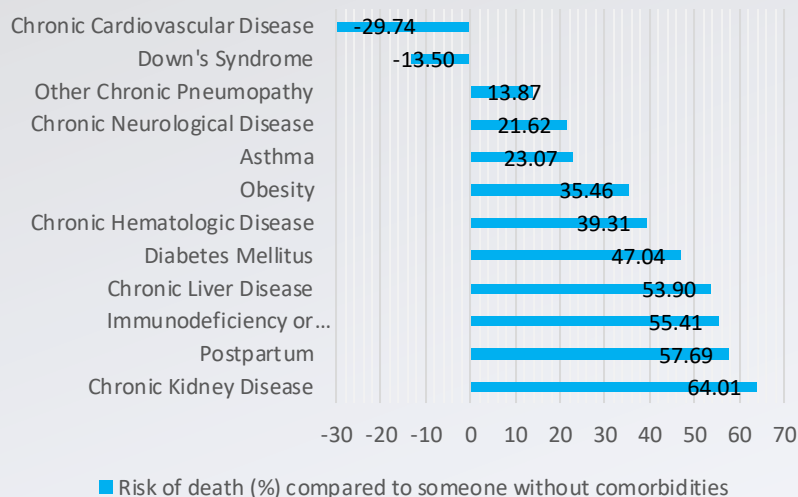
taianmeca@gmail.com



4th Conference on  
**Statistics and  
Data Science**  
Salvador, Brazil (online)  
December 1-3, 2022

# Results

Risk of death (%) compared to someone without comorbidities



Intercept	Risk of death without comorbidity (%)
-0.9615349252132588	27.65709825488868

- The risk of death by COVID for someone without any comorbidity hospitalized is around 27.66%.
- Some comorbidities (chronic cardiovascular disease and down's syndrome) have lower risk of death.

