REVISTA MEXICANA DE ECONOMÍA Y FINANZAS
Nueva Época
REMEF
(THE MEXICAN JOURNAL OF ECONOMICS AND FINANCE)

Revista Mexicana de Economía y Finanzas, Nueva Época

Volumen 20 Número 1, Enero – Marzo 2025, pp. 1-33, e911

DOI: https://doi.org/10.21919/remef.v20i1.911

imef fundación de investigación

(Received: August 7, 2023, Accepted: July 25, 2024, Published: December 11, 2024)

Modeling economic growth in pandemic times with machine learning regression algorithms

J. Alejandro Navarro-Acosta - Universidad Autónoma de Coahuila, México
Valeria Soto-Mendoza[1] - Universidad Autónoma de Coahuila, México
Laura Policardo - Customs and Monopolies Agency, Italy
Edgar J. Sánchez-Carrera - University of Florence, Italy & Universidad Autónoma de Coahuila, México

Abstract

Objective: The aim of this paper is to analyze the contrast of policies to face the Covid-19 pandemic in the socioeconomic performance of three representative economies: Italy, Mexico, and United States. Methodology: Machine learning (ML) techniques are applied to analyze the socioeconomic effects of the pandemic (containment measures, infection rates, total deaths, vaccination, etc.) on GDP growth in those countries. The experiment is that New Zealand's reference stringency index replaces each of those countries' own stringency index and the forecasts for GDP growth, Covid-19-induced deaths, and the Covid-19 reproduction rate. Thus, we show that ML techniques are robust tools for multiple outcome regressions and for experimental scenarios on the socioeconomic impact of the Covid-19 pandemic. Results: The experimental results revealed that the Regression Tree and Random Forest techniques successfully estimate and predict the cases of Italy, Mexico, and the United States. Conclusions: The proposal is that stringency measures and vaccination policies are undoubtedly successful in the fight against a pandemic, in addition to measuring the effects of such policies when data is available through the use of novel techniques such as ML.
JEL Classification: C8, I1, O1, O5, Y1.

Keywords: Economic Growth, Data Driven Analysis, Machine Learning, Stringency Index, Pandemic COVID-19.

Modelado del crecimiento económico durante la pandemia con algoritmos de regresión de aprendizaje automático

Resumen

Objetivo: El objetivo es analizar el contraste de políticas para enfrentar la pandemia de Covid-19 en el desempeño socioeconómico de: Italia, México y Estados Unidos. Metodología: Aplicando técnicas de aprendizaje automático (machine learning, ML) para analizar los efectos socioeconómicos de la pandemia (medidas de contención, tasas de infección, muertes totales, vacunación, etc.) sobre el crecimiento del PIB en esos países. El experimento es que el índice de contingencia referencial de Nueva Zelanda reemplaza el propio índice referencial de cada uno de los países para predecir el PIB, muertes inducidas por Covid-19 y tasa de reproducción de Covid-19. Se muestra que las técnicas de ML son herramientas sólidas para regresiones de resultados múltiples y para escenarios experimentales sobre el impacto socioeconómico de la pandemia de Covid-19. Resultados: Los resultados experimentales revelaron que las técnicas de Árbol de Regresión y Bosque Aleatorio estiman y predicen con éxito los casos de Italia, México y Estados Unidos. Conclusiones: La propuesta es contingencia y vacunación son sin duda exitosas en la lucha contra una pandemia, además de medir los efectos de dichas políticas con el uso de técnicas novedosas como el ML.
Clasificación JEL: C8, I1, O1, O5, Y1.

Palabras clave: Crecimiento Económico, Análisis Basado en Datos, Aprendizaje Automático, Indice de Rigurosidad, Pandemia COVID-19.

[1] Corresponding author: Saltillo, Coahuila de Zaragoza, México. E-mail: vsoto@uadec.edu.mx
* No source of funding for research development

1. Introduction

Machine Learning (ML) algorithms have attracted increasing attention during the last decade due to its power to perform out-of-sample forecasts and by discovering potentially very complex structures in the data that were not specified in advance. Their potentiality was first developed using neural network models to exploit the increasing amount of big data collected by the several sources of information like Google, Facebook, and other internet-based services providers[2].

The use of ML models in macroeconomic forecasting is quite a recent practice. Until recently, “traditional” macro-economists have tended to prefer standard econometric techniques to more sophisticated models like those involving ML. One reason could be that they often work with datasets that fit in a spreadsheet, so there was no need to use algorithms that find most of their applications to big data. Moreover, political economists have always been interested in investigating the effects of one variable (covariate) on a target variable (Y or dependent variable), focusing mainly on its causal effect. Models of causal effects developed by economists focus their attention in estimating unbiased values of the coefficients attached to the covariates of interests, namely, the $\hat{\beta}$, but often have the disadvantage of suffering of poor predictive power.

In our paper, the goal is not to compare the accuracy of predictions attained with various ML models with respect to standard regression techniques - fact that is widely, even not unanimously, documented in the literature - but to estimate possible impacts of different Covid-19 containment policies on various economic and epidemiological indicators for three countries, namely, Italy, Mexico, and United States. We decided to use a ML approach because we want to handle potential various sort of non-linearities in the data that standard econometric techniques find hard to do, without making any assumption and being constrained to assume a functional form for this non-linearities, as it happens when one uses a standard regression technique.

So, in this paper, we use four ML algorithms to forecast three economic and epidemiological indexes, namely, GDP growth, Covid-19 reproduction rate and induced deaths during the period 01-01-2020 - 02-09-2021 in Italy, Mexico and USA, as a function, among the others, of their respective containment policies. Since these three indicators depend on the mitigation policies undertaken by each country, we use as a ``control'', among others, the Covid-19 containment index in each country. Once the models were estimated, in order to forecast the three outcome variables (GDP growth, Covid-19 reproduction rate and Covid-19 induced deaths), we substituted, for all the three countries, the containment index of New Zealand. Subsequently, for Mexico and USA only, we repeated the same experiment substituting the containment index series of Italy.

Our goal is to verify whether a different policy would have generated higher or lower economic and epidemiological performances for those countries. In order to do that, we use - as a benchmark - the containment indexes of New Zealand and Italy. The New Zealand's index was then substituted to each of the three country's own stringency index and forecasts on GDP growth, deaths induced to Covid-19 and Covid-19 reproduction rate performed for all the three countries analyzed.

[2] Think for example to the internet searches done by all the users, that all ow to profile them and submit proper advertisements, as suggested by Einav & Levin (2014), or to the purchase habits of the customers of a supermarket, or by Facebook, though which an analysis of the photos in the network may help the Police, for example, to detect nets of terrorist organizations as suggested by Macdonald et al. (2019).

Revista Mexicana de Economía y Finanzas, Nueva Época, Vol. 20 No. 1, pp. 1-33, e911
DOI: https://doi.org/10.21919/remef.v20i1.911

Moreover, the Italian Stringency index was substituted to the Mexican and U.S. true index as a potential counterfactual to forecast GDP growth and the two epidemiological indexes in these two countries as if that policy was implemented in the place of their own.

2. Literature review

In recent years, however, ML methods have improved substantially their forecasting capacity through the implementation of new algorithms and have acquired a new centrality in the economic literature (Athey, 2018; Einav & Levin, 2014; Gogas & Papadimitriou, 2021; Monroe et al., 2015; Varian, 2014). In the realm of macroeconomics, machine learning algorithms were pitted against traditional econometric models. The comparison revealed that ML methods outperform conventional time series analysis in estimating the US Gross Domestic Product (GDP) (Maccarrone et al., 2021). Other scholars have forecast several macroeconomic variables using ML like inflation (Liao, 2017), unemployment rate (Katris, 2020) and other indicators (Maehashi & Shintani, 2020), finding better forecasts than standard regression models.

In addition, recent and important research papers have developed on the application of ML techniques to measure and construct predictions about the effect of Covid-19 in specific country-cases (see, for instance, Celestine Iwendi & Mohan (2024), Tiwari et al. (2022), Kumar et al. (2022)), and in general cases for the global economy (see, for instance, Sengupta et al. (2020); Tiwari et al. (2022)). In those research studies, the application of ML is assessed for mining time series data during the covid-19 pandemic and to make predictions about the socioeconomic costs of such a shock. Recent studies tackle the socioeconomic costs of the pandemic on economic growth (see for instance, Bischi et al. (2022); Gubar et al. (2023)). Nevertheless, our research proposal differs from those researches in that, considering the most important measure of economic growth, that is GDP, with the application of ML we can measure and predict lockdowns or quarantine scenarios during the pandemic to know the effect of this situation on the GDP of the particular economies.

It is worth to mention that the higher predictive capacity of ML algorithms has not always confirmed in the literature: Katris (2020) conducted a comparison between the forecasts of unemployment rates generated by classical time series models and machine learning techniques and Fischer et al. (2018) made a similar exercise under different scenarios, finding standard time series econometric techniques more appropriate than ML algorithms (and viceversa) depending on the underlying data generating process. Bajari et al. (2015) use a panel dataset to estimate demand with ML algorithms and found better results than standard regression estimations forecasts.

In this paper, we use four ML algorithms to forecast three economic and epidemiological indexes, namely, GDP growth, Covid-19 reproduction rate and induced deaths during the period 01-01-2020 - 02-09-2021 in Italy, Mexico, and US, as a function, among the others, of their respective containment policies. Our goal is to verify whether a different policy would have generated higher or lower economic and epidemiological performances for those countries. In order to do that, we use - as a benchmark - the containment indexes of New Zealand and Italy. The New Zealand's index was then substituted to each of the three country's own stringency index and forecasts on GDP growth, deaths induced to Covid-19 and Covid-19 reproduction rate performed for all the three countries analyzed. Moreover, the Italian Stringency index was substituted to the Mexican and U.S. true index as a potential counterfactual to forecast GDP growth and the two epidemiological indexes in these two countries as if that policy was implemented in the place of their own.

3. Machine learning techniques

Data is used as an input for machine learning techniques. Firstly, a procedure called *k-fold cross validation* was employed to divide the dataset into k equal-sized sub-samples (Jung & Hu, 2015). Then a single sub-sample is retained as the validation data for testing the regression model, and the remaining $k-1$ sub-samples are used as training data. This approach is repeated k times and the k results can then be averaged to produce a single estimation (Arlot & Celisse, 2010). This approach presents some advantages such as every data points get to be tested exactly once and is used in training $k-1$ times. In addition, the variance (overfitting) of the resulting estimate is reduced as k increases. Moreover, this method reduces the bias (underfitting) in the machine learning techniques (Browne, 2000).

Secondly, the *Grid Search* procedure was applied. In the Machine Learning field, hyper-parameter tuning is a crucial task to enhance the model learning capabilities, thus in this research, a Grid Search procedure is applied to find the hyper-parameters that allow us to obtain a better fit for the datasets. Grid search is a versatile method for exploring various model configurations. It involves discretizing a target range of values for each hyper-parameter of interest, then training and testing models across all hyper-parameters for every combination of values (Mesafint & Huchaiah, 2021). This procedure is depicted in Figure 1.

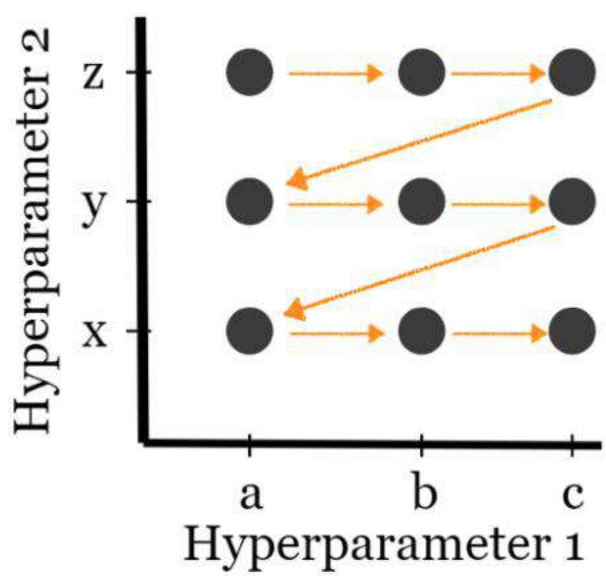

Figure 1. Two hyper-parameters grid search example.

3.1 Regression Tree

Regression trees (RT) generate regression models represented by a tree arrangement. A decision tree maps each vector of feature variables to a predicted value of the target variable, and it is composed of a set of "questions" organized hierarchically with the shape of a tree. Usually, a "question" is represented by a condition or split. The terminal nodes are called "leafs" and contain predictions. All non-leaf nodes are conditions. The initial node is the "root" node, that is, the first condition from which the tree grows. Each path starting from the root and ending in a leaf is called "inference path"[3]. The fundamental algorithm for constructing decision trees, as introduced by Quinlan (1986), is known as ID3. It employs a top-down, greedy search to explore the potential branches without backtracking. While typically used for classification tasks, ID3 can also be adapted for regression by substituting Information Gain with Standard Deviation Reduction. Beginning from a root node, the decision tree is constructed in a top-down manner, involving the division of data into subsets with similar instances. Standard deviation serves to measure the homogeneity of numerical samples, where a standard deviation of zero indicates complete homogeneity. For more details consult De Ville (2013); Loh (2014).

3.2 Random Forest

The Random Forest (RF) algorithm stands as a widely embraced supervised learning method. RF is an ensemble learning method that combines a series of k decision trees, $T_1, T_2, \ldots, T_k$, usually trained with the "bagging" method. This approach strives to generate an enhanced composite classification or prediction model, T^*. Using a given dataset D, RF constructs multiple decision trees and combines them to yield a prediction that is more accurate and stable. When presented with a new sample, each model within the ensemble generates a predicted outcome. The random forest then aggregates these individual predictions, ultimately returning a final predicted value based on the mean of the predicted results from each tree (Genuer & Poggi, 2020). This method exhibits robustness as it is applicable to both classification and regression tasks. Furthermore, its efficacy can be enhanced by augmenting the diversity among the constituent decision tree models. Conversely, while the generalization error tends to diminish with a larger number of trees in the forest, mitigating the risk of overfitting (Izquierdo-Verdiguier & Zurita-Milla, 2020). In this algorithm, two parameters must be considered to achieve a good performance in regression tasks. The first is the number of trees that make up the forest N_{trees}. The second N_i, is the number of split variables of the tree node preselected. In practice, these parameters are tuned empirically, for this reason, is difficult to ensure the best performance of Random Forests. Nevertheless, alternative methodologies like Grid search have been implemented to address this concern.

[3] In general, the dataset is broken into smaller subsets in the meantime a decision tree is created. At the end, a tree with decision nodes and leaf nodes is obtained. A decision node has branches, each branch is a value for the attribute proved. A leaf node is a decision on the target. The root node is the topmost decision node of the tree which coincides with the best predictor.

3.3 Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is a model that has undergone continuous optimization and enhancement through the collective efforts of numerous researchers. In contrast to Random Forest (Bagging method), XGBoost operates as a learning framework centered around Boosting Tree models (Le Nguyen, 2022). This approach performs a second-order Taylor expansion on the loss function. Given a training set $\{x_i, y_i\}_{i=1}^{N}$, a differentiable loss function $L(y, F(x))$, a number of weak learners M and a learning rate α. A generic XGBoost algorithm first initializes a model with constant value:

$$\hat{f}_{(0)}(x) = arg_{\theta} \min \sum_{i=1}^{N} L(y_i, \theta)$$

Then for each model $m \in M$, compute the gradient and Hessian:

$$\hat{g}_m(x_i) = \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]$$

$$\hat{h}_m(x_i) = \left[\frac{\partial^2 L(y_i, f(x_i))}{\partial f(x_i)^2}\right]$$

now fit the weak learner using the training set $\left\{x_i - \frac{\hat{g}_m(x_i)}{\hat{h}_m(x_i)}\right\}_{i=1}^{N}$ by solving the optimization problem:

$$\hat{\phi}_m = arg \min_{\phi \in \Phi} \sum_{i=1}^{N} \frac{1}{2} \hat{h}_m(x_i) \left[\phi(x_i) - \frac{\hat{g}_m(x_i)}{\hat{h}_m(x_i)}\right]^2$$

$$\hat{f}_m(x) = \alpha \hat{\phi}_m(x)$$

for update the model compute $\hat{f}_m(x) = \hat{f}_{m-1}(x) + \hat{f}_m(x)$, and finally the output is: $\hat{f}(x) = \hat{f}_M(x) = \sum_{m=0}^{M} \hat{f}_m(x)$.

3.4 K-Nearest Neighbor Regression

K-nearest neighbors (KNN) (Meade, 2002) is a simple algorithm that stores all available cases and predicts the numerical target based on a similarity measure i.e. distance functions such as the Euclidean distance

$$\sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}$$

or the Minkowski distance

$$\left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q}$$

A straightforward implementation of KNN regression involves computing the average of the numerical target values of the K nearest neighbors. KNN regression utilizes the same distance functions as KNN classification and finds applications in statistical estimation and pattern recognition. For details see Hu et al. (2022).

4. Data

The analysis begins with the collection of all possible relevant data and available features that affect the three key targets of interest (GDP, Covid-19 Reproduction Rate, and Total Deaths). Our main sources are:

- OECD weekly tracker of GDP[4]. This source provides data for several economic indicators like production, consumption, labor markets, housing, trade, industrial activity, and economic uncertainty. It also provides, as the name suggests, weekly observations for GDP growth (as percent change with respect to the same week in the previous year and in the previous two years). Despite these were not official OECD forecasts, we claim that they could be successfully used for our purpose since this indicator is one that feeds the OECD forecasting process. From this source, we used the variable indicating GDP growth observed weekly with respect to the same indicator two years earlier. We decided to use the variables "% change in GDP in the same week over two years earlier" instead of the same variable recorded one year earlier because the pandemic has lasted more than a year, and therefore, the use of this indicator could be better for estimating the effect of the Covid-19 on the economy than the same indicator expressed one year earlier.
- Our World in Data (OWID)[5]. The full dataset on Covid-19 from OWID was downloaded. This dataset includes daily observations and metrics on vaccinations, hospitals, and intensive care units (ICU), tests and positivities, confirmed cases and deaths, reproduction rates and policy responses, in addition to other potentially relevant variables. Some variables are collected by Our World in Data's team, other features are collected by other scholars or institutions and are disseminated by OWID. Since this data, contrary to the previous one, are recorded daily, the dataset was adjusted taking weekly means of all the series.

The full list of variables and their detailed descriptions can be found in Table 1.

[4] The OECD Weekly Tracker of GDP growth (https://www.oecd.org/economy/weekly-tracker-of-gdp-growth/ is a real-time high- frequency indicator of economic activity using machine learning and Google Trends data (Woloszko, 2020).
[5] Source: (Mathieu et al., 2020). These data have been downloaded from https://github.com/owid/Covid-19-data/tree/master/public/data.

Table 1. Complete list of variables.

Variable name	Description	Source
OECD weekly tracker of GDP		
gdp	% change of GDP recorded each week over two years earlier.	OECD**
Confirmed cases		
total_cases	Total confirmed cases of Covid-19. Counts can include probable cases, where reported.	OWID
new_cases	New confirmed cases of Covid-19. Counts can include probable cases, where reported. In rare cases where our source reports a negative daily change due to a data correction, we set this metric to NA.	OWID
new_cases_smoothed	New confirmed cases of Covid-19 (7-day smoothed). Counts can include probable cases, where reported.	OWID
total_cases_per_million	Total confirmed cases of Covid-19 per 1,000,000 people. Counts can include probable cases, where reported.	OWID
new_cases_per_million	New confirmed cases of Covid-19 per 1,000,000 people. Counts can include probable cases, where reported.	OWID
new_cases_smoothed_per_million	New confirmed cases of Covid-19 (7-day smoothed) per 1,000,000 people. Counts can include probable cases, where reported.	OWID
Confirmed deaths		
total_deaths	Total deaths attributed to Covid-19. Counts can include probable deaths, where reported.	OWID**
new_deaths	New deaths attributed to Covid-19. Counts can include probable deaths, where reported. In rare cases where our source reports a negative daily change due to a data correction, we set this metric to NA.	OWID
new_deaths_smoothed	New deaths attributed to Covid-19 (7-day smoothed). Counts can include probable deaths, where reported.	OWID
total_deaths_per_million	Total deaths attributed to Covid-19 per 1,000,000 people. Counts can include probable deaths, where reported.	OWID
new_deaths_per_million	New deaths attributed to Covid-19 per 1,000,000 people. Counts can include probable deaths, where reported.	OWID
new_deaths_smoothed_per_million	New deaths attributed to Covid-19 (7-day smoothed) per 1,000,000 people. Counts can include probable deaths, where reported.	OWID
Reproduction rate		
reproduction_rate	Real-time estimate of the effective reproduction rate (R) of Covid-19. See https://github.	OWID**

	com/crondonm/TrackingR/tree/main/ Estimates-Database.	
Excess of mortality[6]		
excess_mortality	Percentage difference between the reported number of weekly or monthly deaths in 2020–2021 and the projected number of deaths for the same period based on previous years.	OWID *
excess_mortality_cumulative	Percentage difference between the cumulative number of deaths since 1 January 2020 and the cumulative projected deaths for the same period based on previous years.	OWID
excess_mortality_cumulative_absolute	Cumulative difference between the reported number of deaths since 1 January 2020 and the projected number of deaths for the same period based on previous years.	OWID
excess_mortality_cumulative_per_million	Cumulative difference between the reported number of deaths since 1 January 2020 and the projected number of deaths for the same period based on previous years, per million people.	OWID
Hospital & ICU		
icu_patients	Number of Covid-19 patients in intensive care units (ICUs) on a given day.	OWID
icu_patients_per_million	Number of Covid-19 patients in intensive care units (ICUs) on a given day per 1,000,000 people.	OWID
hosp_patients	Number of Covid-19 patients in hospital on a given day.	OWID
hosp_patients_per_million	Number of Covid-19 patients in hospital on a given day per 1,000,000 people.	OWID
weekly_icu_admissions	Number of Covid-19 patients newly admitted to intensive care units (ICUs) in a given week (reporting date and the preceeding 6 days).	OWID
weekly_icu_admissions_per_million	Number of Covid-19 patients newly admitted to intensive care units (ICUs) in a given week per 1,000,000 people (reporting date and the preceeding 6 days).	OWID
weekly_hosp_admissions	Number of Covid-19 patients newly admitted to hospitals in a given week (reporting date and the preceeding 6 days).	OWID
weekly_hosp_admissions_per_million	Number of Covid-19 patients newly admitted to hospitals in a given week per 1,000,000 people (reporting date and the preceeding 6 days).	OWID
Policy responses		
stringency_index	Government Response Stringency Index: composite measure based on 9 response indicators including school closures, workplace closures, and travel bans,	OWID*

[6] See https://github.com/owid/Covid-19-data/tree/master/public/data/excess_mortality.

	rescaled to a value from 0 to 100 (100 = strictest response).	

Source: Own elaboration.

Table 1 (continued). Complete list of variables.

Variable name	Description	Source
Policy responses		
containment_index[7]	The index on any given day is calculated as the mean score of thirteen metrics, each taking a value between 0 and 100. A higher score indicates a stricter response (i.e. 100 = strictest response).	OWID*
Test and positivity		
total_tests	Total tests for Covid-19.	OWID
new_tests	New tests for Covid-19 (only calculated for consecutive days).	OWID
total_tests_per_thousand	Total tests for Covid-19 per 1,000 people.	OWID
new_tests_per_thousand	New tests for Covid-19 per 1,000 people.	OWID
new_tests_smoothed	New tests for Covid-19 (7-day smoothed). For countries that don't report testing data on a daily basis, we assume that testing changed equally on a daily basis over any periods in which no data was reported. This produces a complete series of daily figures, which is then averaged over a rolling 7-day window.	OWID
new_tests_smoothed_per_thousand	New tests for Covid-19 (7-day smoothed) per 1,000 people.	OWID
positive_rate	The share of Covid-19 tests that are positive, given as a rolling 7-day average (this is the inverse of tests_per_case).	OWID*
tests_per_case	Tests conducted per new confirmed case of Covid-19, given as a rolling 7-day average (this is the inverse of positive_rate).	OWID*
tests_units	Units used by the location to report its testing data. A country file can't contain mixed units. All metrics concerning testing data use the specified test unit. Valid units are 'people tested' (number of people tested), 'tests performed' (number of tests performed. a single person can be tested more than once in a given day) and 'samples tested' (number of samples tested. In some cases, more than one	OWID

[7] This indicator is available for download at https://ourworldindata.org/policy-responses-covid.

	sample may be required to perform a given test).	
Vaccinations		
total_vaccinations	Total number of Covid-19 vaccination doses administered.	OWID
people_vaccinated	Total number of people who received at least one vaccine dose.	OWID
people_fully_vaccinated	Total number of people who received all doses prescribed by the initial vaccination protocol.	OWID*
total_boosters	Total number of Covid-19 vaccination booster doses administered (doses administered beyond the number prescribed by the vaccination protocol).	OWID
new_vaccinations	New Covid-19 vaccination doses administered (only calculated for consecutive days).	OWID*
new_vaccinations_smoothed	New Covid-19 vaccination doses administered (7-day smoothed). For countries that don't report vaccination data on a daily basis, we assume that vaccination changed equally on a daily basis over any periods in which no data was reported. This produces a complete series of daily figures, which is then averaged over a rolling 7-day window.	OWID
total_vaccinations_per_hundred	Total number of Covid-19 vaccination doses administered per 100 people in the total population.	OWID
people_vaccinated_per_hundred	Total number of people who received at least one vaccine dose per 100 people in the total population.	OWID
people_fully_vaccinated_per_hundred	Total number of people who received all doses prescribed by the initial vaccination protocol per 100 people in the total population.	OWID
total_boosters_per_hundred	Total number of Covid-19 vaccination booster doses administered per 100 people in the total population.	OWID
new_vaccinations_smoothed_per_million	New Covid-19 vaccination doses administered (7-day smoothed) per 1,000,000 people in the total population.	OWID
new_people_vaccinated_smoothed	Daily number of people receiving their first vaccine dose (7-day smoothed).	OWID
new_people_vaccinated_smoothed_per_hundred	Daily number of people receiving their first vaccine dose (7-day smoothed) per 100 people in the total population.	OWID
Others		
iso_code	ISO 3166-1 alpha-3 – three-letter country codes. Note that OWID-defined regions (e.g.	OWID

	continents like ’Eu- rope’) contain prefix ’OWID_’.	
continent	Continent of the geographical location	OWID
location	Geographical location. Location ’International’ considers special regions ("Diamond Princess" and "MS Zaandam" cruises).	OWID
date	Date of observation.	OWID
population	Population (latest available values). See https://github.com/owid/Covid-19-data/blob/master/scripts/input/un/population_latest.csv for full list of sources.	OWID
population_density	Number of people divided by land area, measured in square kilometers, most recent year available.	OWID
median_age	Median age of the population, UN projection for 2020.	OWID
aged_65_older	Share of the population that is 65 years and older, most recent year available.	OWID

Source: Own elaboration.

Table 1 (continued). Complete list of variables.

Variable name	Description	Source
Others		
aged_70_older	Share of the population that is 70 years and older in 2015.	OWID
gdp_per_capita	Gross domestic product at purchasing power parity (constant 2011 international dollars), most recent year available.	OWID
extreme_poverty	Share of the population living in extreme poverty, most recent year available since 2010.	OWID
cardiovasc_death_rate	Death rate from cardiovascular disease in 2017 (annual number of deaths per 100,000 people).	OWID
diabetes_prevalence	Diabetes prevalence (% of population aged 20 to 79) in 2017.	OWID
female_smokers	Share of women who smoke, most recent year available.	OWID
male_smokers	Share of men who smoke, most recent year available.	OWID
handwashing_facilities	Share of the population with basic handwashing facilities on premises, most recent year available.	OWID
hospital_beds_per_thousand	Hospital beds per 1,000 people, most recent year available since 2010.	OWID
life_expectancy	Life expectancy at birth in 2019.	OWID

human_development_index	A composite index measuring average achievement in three basic dimensions of human development—a long and healthy life, knowledge and a decent standard of living. Values for 2019, imported from http://hdr.undp.org/en/indicators/137506.	OWID

Source: Own elaboration.

4.1 Country selection

As a policy counterfactual, we used the Covid-19 containment indexes of Italy and New Zealand[8]. In this context, being ML algorithms more effective in producing good forecasts of a variable of interest than any other traditional econometric method, we believe that they could be suitable for our purpose, that is, find forecasts of three variables of interest in a given country once a counterfactual policy is implemented.

The Italian and New Zealand experiences in the global panorama of non-medical containment policies are two important examples. Italy bore the brunt of the virus in its most aggressive manifestation, being among the initial western nations to face its impact. On March 9th, Prime Minister Giuseppe Conte announced a total closure of economic activities and personal movements leaving opened only essential activities. After the decline of the contagion rate the strictness of the containment policy was gradually relaxed, and selective/differentiated policies based on municipalities implemented depending on the number of Covid-19 cases recorded in each municipality. New Zealand's elimination (and not mitigation) strategy implying total closure of their boundaries at the beginning of the first outbreak (shifted to a mitigation policy later on, with the outbreak of the less severe Omicron) has let New Zealand to record the smaller number of deaths and the best economic performance at the same time, according to The Guardian[9].

For what concerns Mexico and the United States (US), their containment policies have been characterized by less strict and poorly coordinated rules. In the United States, the lack of a unified national strategy and contradictory messaging regarding social distancing measures was particularly evident, notably amidst the U.S. election campaign (Yamamoto et al., 2021). Knaul et al. (2021a) illustrated in Mexico the lack of a consistent national approach and the considerable disparity in the stringency of responses across different states. Differently from New Zealand, where the policies have been explicitly advised by public health experts (Gauld, 2023). In Mexico the responses were not grounded in testing and did not accurately mirror the local disease burden (Knaul et al., 2021).

[8] For Italy, the alternative policy was New Zealand's one, and for Mexico and USA, the alternative policies were the Italy and New Zealand's ones.
[9] https://www.theguardian.com/world/commentisfree/2022/apr/05/new-zealands-covid-strategy-was-one-of-the-worlds-most-successful-what-can-it-learn-from-it

5. Methodology

All the selected variables from the dataset presented in Section 4 are numeric (float) with the only exceptions of the variables describing the country code, the continent, and the location (*iso_code, continent, location*). Three are target variables: *reproduction_rate*, *total_deaths* and *gdp*, denoted by two asterisks in the "Source" column; the remaining are feature variables. The feature variables denoted by one asterisk (which can be found in the "source" column) are those that remained after the five phases of the pre-processing process and that were used to infer the target variables in the machine learning models presented.

5.1 Data pre-processing

As we have already mentioned, our three target variables are: annual GDP growth (recorded every week during the analysis period with respect to the same week two years earlier), virus reproduction rate, and total deaths due to Covid-19. The data were analyzed and pre-processed as a first step. The entire data pre-processing procedure consisted of five steps:

1. Check for missing data and replace them with a mean of the nearest values.
2. Average daily records in weekly records. The averaging methods consisted in the division of the year in weeks (for example, the first seven days of the year were classified as "week 1", the eighth to the 14th day of the year as "week 2", etc.) and a computation of the mean for each week of the daily observations of the variables, so to have weekly observations of those variables.
3. Order all the observations according to the date (in ascending order) and merge the several variables coming from the different sources according to the date.
4. Drop all the variables having zero variance.
5. Each remaining variable was paired with all the other remaining variables two-by-two and the correlation was computed. If the correlation between the two variables paired was greater than 0.7, one variable was dropped, and the other one was kept for the analysis. This operation allowed us to reduce the number of highly correlated features, to avoid potential problems such as multicollinearity and endogeneity between variables. Since the nonlinear machine learning models used in this work are free of most statistical assumptions compared to methods such as linear regression, the possible effects of the issues mentioned earlier are often ignored. This happens when the purpose of the ML model is solely to be accurate in predictions without the need to be explanatory (a black box). In the work presented by De Veaux & Ungar (1994), it was demonstrated how a neural network achieved good predictive performance in contrast with a spline-based regression when multicollinearity is present, however neural networks like other ML approaches such as KNN and SVM do not provide interpretability. On the other hand, regression tree-based models such as Random Forest provide some interpretability through an interesting property known as feature importance (described in Section 6.2). Although the way of estimating this importance differs completely from linear regression where the coefficients are directly affected by multicollinearity and

endogeneity, these issues could produce undesirable effects in ML algorithms, such as dealing with unnecessary high dimensionality as well as the loss of generalization in prediction (overfitting).

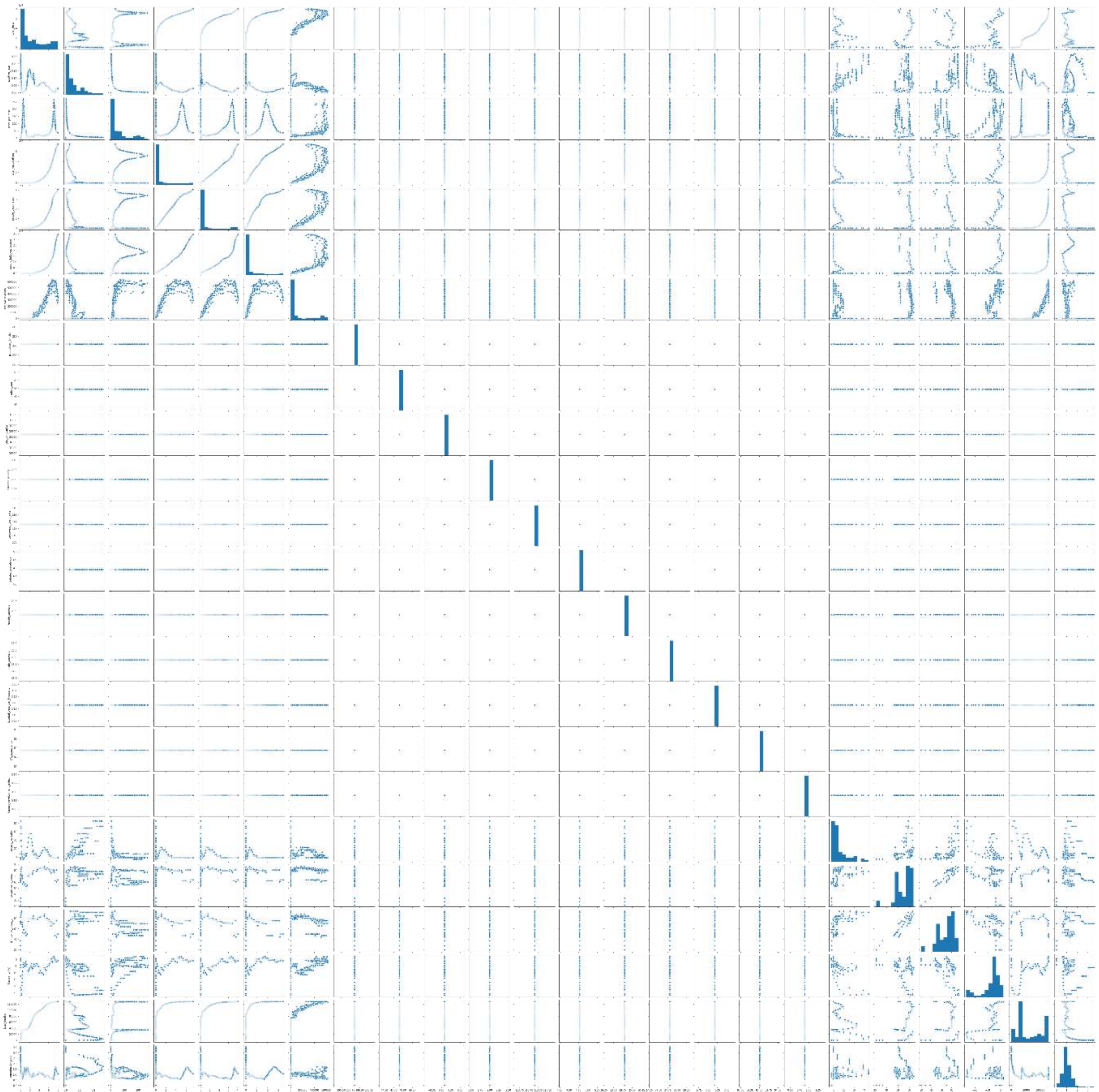

Figure 2. Pair chart of all variables for Italy.
Source: Own elaboration.

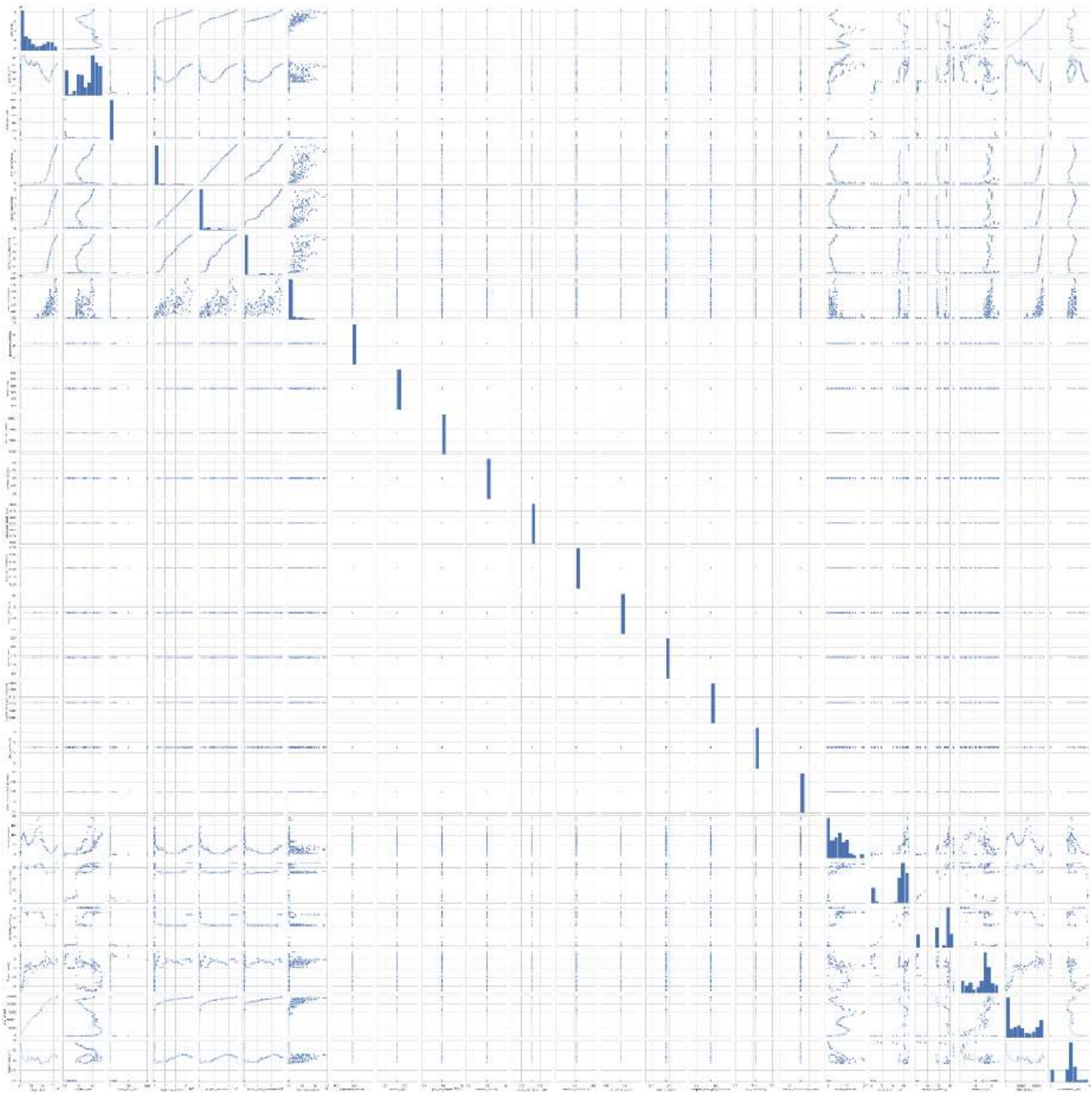

Figure 3. Pair chart of all variables for Mexico.
Source: Own elaboration.

Revista Mexicana de Economía y Finanzas, Nueva Época, Vol. 20 No. 1, pp. 1-33, e911
DOI: https://doi.org/10.21919/remef.v20i1.911

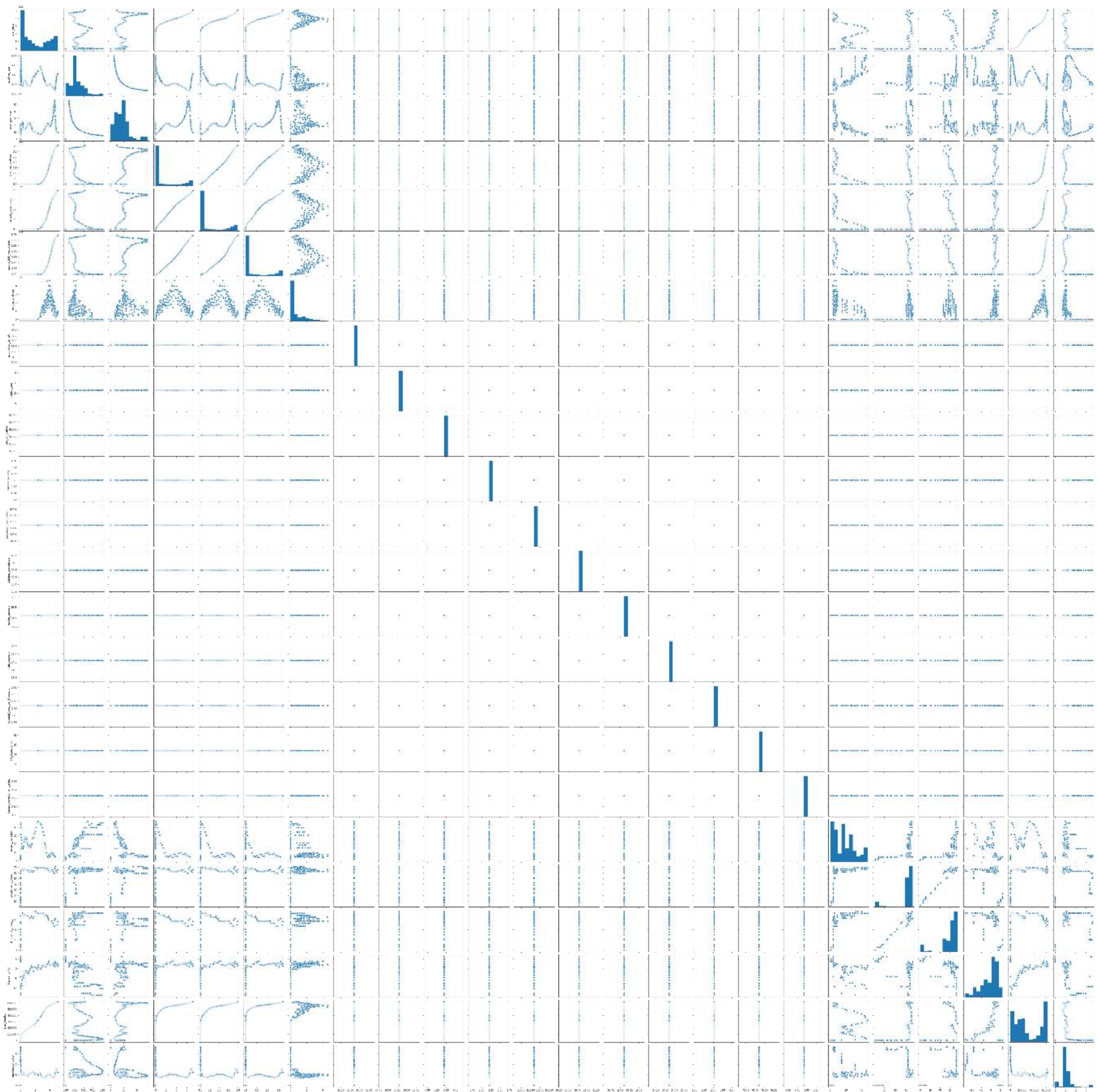

Figure 4. Pair chart of all variables for United States (USA).
Source: Own elaboration.

From the whole dataset whose variables are listed in Table 1, we performed all the pre-processing operations, and we ended up with the dataset characterized by the variables indicated with a star (one star in the "source" column denotes a feature, two stars denote a target. In total, we ended up with seven features. Targets are three by hypothesis).

5.2 Exploratory analysis

First, a pair plot was made to observe the behavior of the variables. Figure 2, Figure 3, and Figure 4 show the graphs corresponding to Italy, Mexico and the United States respectively. As can be noticed, the behavior of the variables within the central square of each graph is constant over time because they are ratios or averages. For this reason, those variables were temporarily eliminated to perform a correlation analysis.

Then, the distribution of the variables is studied to make inferences and implement the most appropriate correlation test. For example, Figure 5 shows the distribution of the variable containment index for the case of Italy, Mexico and the US. As can be seen, the data does not follow a normal distribution.

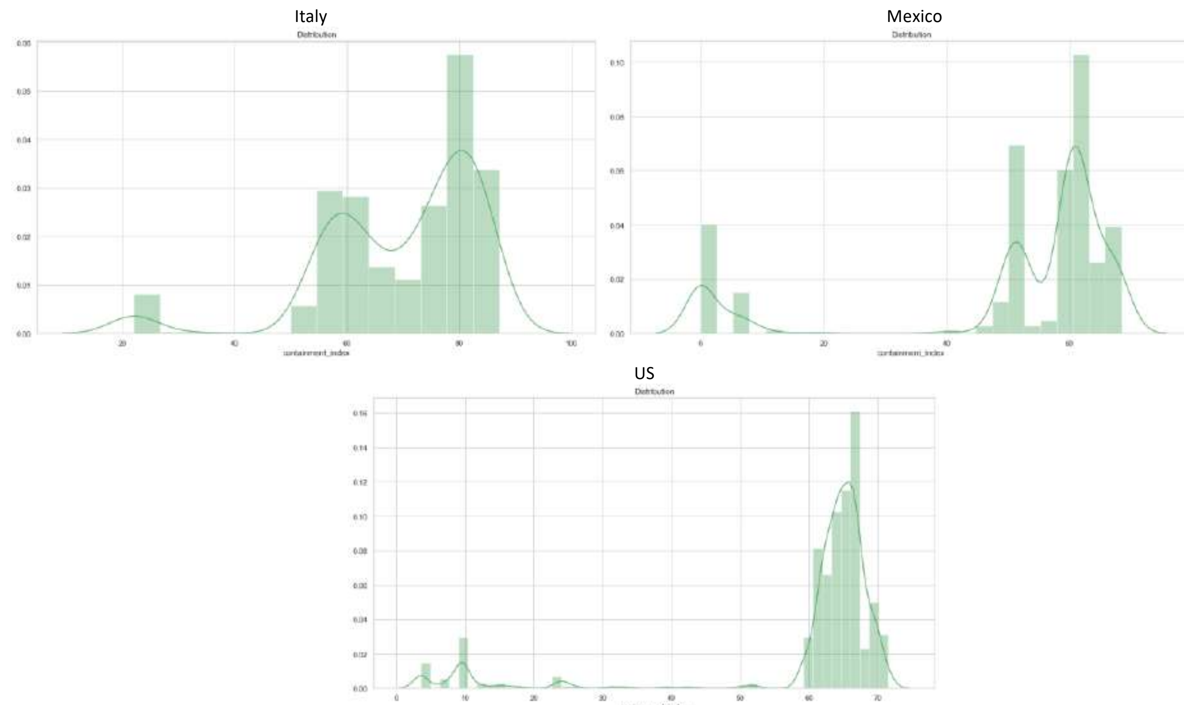

Figure 5. Distribution of containment index data for Italy, Mexico, and United States (US).
Source: Own elaboration.

The rest of the variables present the same characteristics for the countries analyzed. Hence, a Kendall correlation analysis was conducted to assess the extent of dependence among all the variables. Kendall's tau, introduced by Maurice Kendall in 1938, serves as a non-parametric gauge of correlation strength. Specifically tailored for ordinal level variables, Kendall's Tau quantifies the relationship's intensity between two such variables. Alongside Spearman's rank correlation coefficient, Kendall's Tau stands as a widely embraced metric for rank correlations, representing popular statistics in this domain (Puka, 2011). Figure 6 displays the correlation coefficients over a heat map for Italy, Mexico and the United States. A high correlation between variables is considered if Kendall's coefficient is greater than 0.7.

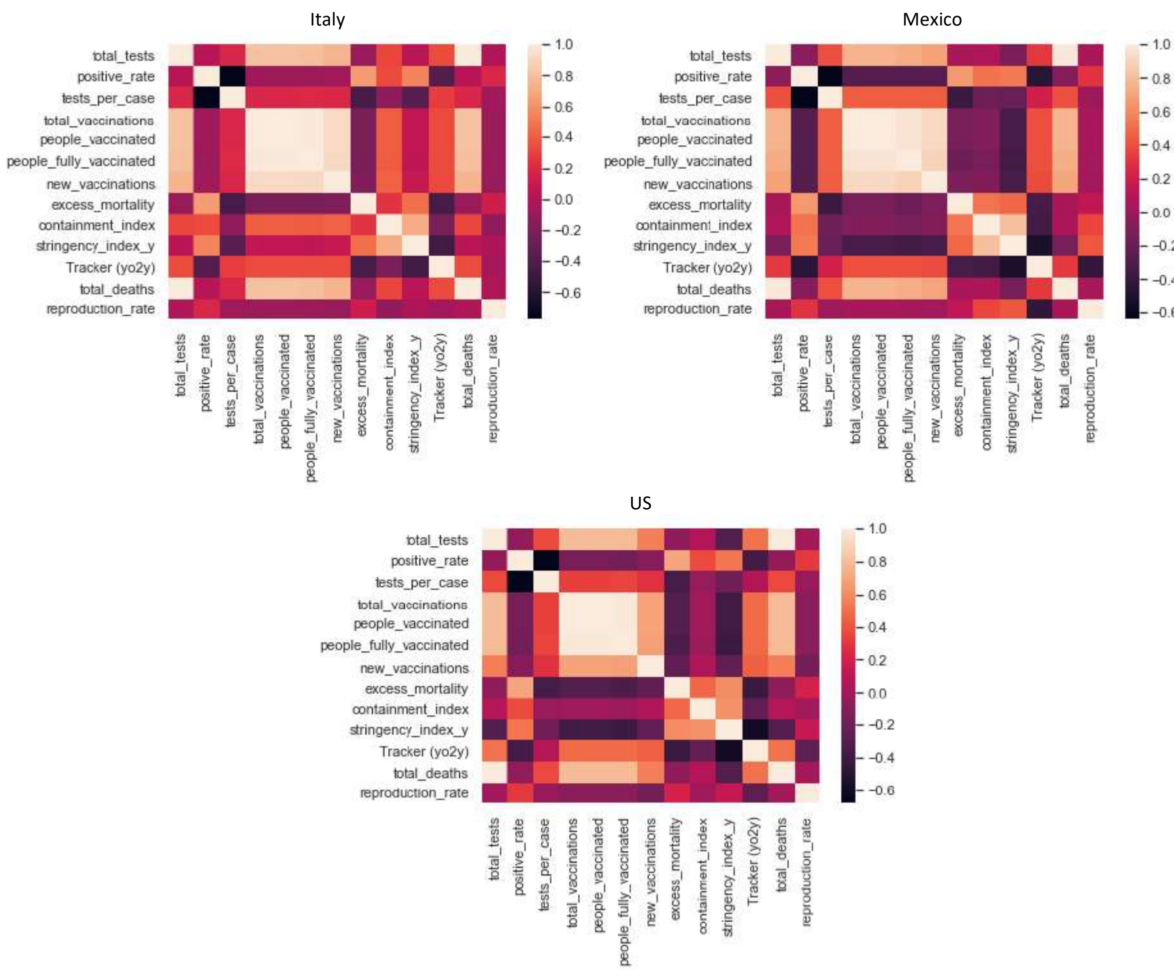

Figure 6. Heat map of Kendall's correlation index for Italy, Mexico, and United States (US).
Source: Own elaboration.

The total size of the different datasets after the pre-processing process are depicted in Table *2*.

Table 2. Original and pre-processed data distribution.

Country	Continent	Initial date	Original data		Pre-processed data	
			End date	Number of rows	End date	Number of rows
MEX	North America	01-01-2020	02-09-2021	611	06-08-2021	547
US		22-01-2020	02-09-2021	590	16-08-2021	545
ITA	Europe	31-01-2020	02-09-2021	581	13-08-2021	561

Source: Own elaboration.

6. Results

The problem can be set as a multi-output regression (Borchani et al., 2015) (also known as multi-variate (Breiman & Friedman, 1997), multi-response (Similä & Tikka, 2007), multi-target (Tsoumakas et al., 2014), or multi-label (Madjarov et al., 2012)). It is the simultaneous prediction of multiple output variables given an input set of variables (Xu et al., 2019a). The multi-output problem (MOP) is associated with multi-output learning (MOL) which "maps each input instance to multiple outputs. Let's assume that $X = \mathbb{R}^d$ is a d-dimensional input space, and $X = \mathbb{R}^m$ is an m −dimensional output label space. The objective of multi-output learning is to learn a function $f : X \to Y$ from the training set $D = \{(x_i, y_i) \mid 1 \leq i \leq n\}$" (Xu et al., 2019b). The general definition of MOL is "finding a function $F: X \times Y \to \mathbb{R}$ based on the training sample of input-output pairs, where $F(x, y)$ is a compatibility function that evaluates how compatible the input x and the output y are. Then, given an unseen instance x at the test state, the output is predicted to be the one with the largest compatibility score, namely, $f(x) = \tilde{y} = argmax_{y\in Y}\, F(x, y)$" (Tsochantaridis et al., 2005). For the multi-output regression specific case, the learned multi-target regression function $f(\cdot)$ predicts a real-valued vector $f(x) \in Y$ as the output (Borchani et al., 2015).

6.1 Fitting the regression models

Each of the datasets considered in this research was divided into two groups: 80% of the data is put through the 10-fold cross-validation procedure to fitting the ML techniques. This means that 10 different models are constructed for each ML algorithm in order to observe their performances. Then the other 20% of the data is used to perform another extra validation procedure, i.e. the models are tested whit data completely unknown to them. This is to reduce as possible the over-fitting and under-fitting. The tuning of the hyper-parameters was made by a Grid Search method. Since it is known that many regression techniques are affected by the different scales of the variables, a statistical standardization has been carried out, i.e. $z = \frac{x - \bar{a}}{s}$ where x is an observed value from a sample with mean $\bar{a}$ and standard deviation s. Table 3, Table 4 and Table 5 show the values of the evaluation metrics achieved by the four proposed approaches. These are the results for the three analyzed countries: Mexico (MEX), the United States (US), and Italy (ITA).

As can be seen, the R^2 achieves good values in the total of cases for the proposed ML approaches, whereas the Linear Regression model (LR) presents the worst performance. However, as previously noted in Figure 2, the datasets present a non-linear behavior and non-normal distribution, which makes the metric R^2 unreliable for choosing the best models. While analyzing the MAE and RMSE scores, it is observed that the RF, XG, and RT algorithms show the best performance when fitting the data, since they present the lowest values both in the training and validation phase. Furthermore, in some cases, the values reached in the validation phase are lower than those obtained in the training phase, which means that the models do not present over-fitting.

DOI: https://doi.org/10.21919/remef.v20i1.911

Table 3. Regression models metrics for Mexico dataset.

MEX						
	Training			Validation		
	R^2	MAE	RMSE	R^2	MAE	RMSE
RF	0.9858	348.77	1369.18	0.9906	371.49	1661.78
XG	0.9638	311.14	1015.66	0.9110	416.48	1434.90
KNN	0.9499	644.29	2209.40	0.9611	592.20	1818.02
RT	0.9752	250.74	783.35	0.9866	301.47	1076.74
LR	0.7055	5695.22	13598.34	0.7417	5608.32	13549.16

Source: Own elaboration.

Table 4. Regression models metrics for United States dataset.

US						
	Training			Validation		
	R^2	MAE	RMSE	R^2	MAE	RMSE
RF	0.9852	755.05	2986.80	0.9874	408.94	1312.11
XG	0.9851	817.01	2926.24	0.9819	610.11	1963.52
KNN	0.9440	2258.80	9305.52	0.9772	1863.55	8404.61
RT	0.9758	846.83	3689.35	0.8110	492.09	1254.50
LR	0.6523	16904.46	35671.24	0.7609	14741.61	31329.33

Source: Own elaboration.

Table 5. Regression models metrics for Italy dataset.

ITA						
	Training			Validation		
	R^2	MAE	RMSE	R^2	MAE	RMSE
RF	0.9447	298.97	1784.99	0.9627	410.56	2387.32
XG	0.9485	256.78	1324.67	0.9788	213.59	913.47
KNN	0.9324	403.86	1884.52	0.9739	324.37	1927.23
RT	0.9310	229.05	1311.07	0.9918	140.22	481.40
LR	0.5225	3477.38	7238.37	0.6085	3293.31	6986.4143

Source: Own elaboration.

6.2 Feature importance

Once the models have been fitted, an analysis of the importance of the variables from the pre-processed dataset is carried out by using an inherent property of the construction of the RF algorithm called feature importance. This property measures how variations in the input variables affect the response. In this way, it is concluded that those features that most influence the model in order to make better predictions of the output values are more important.

For an input dataset X with $x_1, x_2, \dots, x_j, \dots x_p$ variables, the Feature Importance (FI) for the variable x_j is denoted as follows:

$$FI_j = \frac{1}{N_{trees}} \sum_{v \in S} G(x_j, v)$$

Where S is the set of nodes where x_j is considered to partition the samples and $G(x_j, v)$ is known as the Random Forest gain due to this particular x_j feature. Therefore, the gain is based on the impurity measure when the samples are split at each node. In this research, the impurity Gini index was considered, for details see Wang & Luo (2016). Table 6, Table 7 and Table 8 show the feature importance for the three countries considered. As can be seen, the feature *new_vaccinations* is the most important for the RF algorithm in all cases, i.e, this feature significantly contributes to the algorithm achieving better predictions.

Table 6. Feature importance for Mexico.

Feature	Predictor	Importance
3	*New vaccinations*	0.7989
4	*Excess mortality*	0.0921
0	*Positive rate*	0.0372
2	*People fully vaccinated*	0.0366
6	*Stringency index*	0.0150
5	*Containment index*	0.0146
1	*Test per cases*	0.0053

Source: Own elaboration.

Table 7. Feature importance for the US.

Feature	Predictor	Importance
3	*New vaccinations*	0.8309
2	*People fully vaccinated*	0.0459
1	*Test per cases*	0.0370
4	*Excess mortality*	0.0366
6	*Stringency index*	0.0253
5	*Containment index*	0.0195
0	*Positive rate*	0.0044

Source: Own elaboration.

Table 8. Feature importance for Italy.

Feature	Predictor	Importance
3	*New vaccinations*	0.8364
1	*Test per cases*	0.0683
2	*People fully vaccinated*	0.0626
0	*Positive rate*	0.0112
4	*Excess mortality*	0.0082
5	*Containment index*	0.0068
6	*Stringency index*	0.0062

Source: Own elaboration.

DOI: https://doi.org/10.21919/remef.v20i1.911

Nevertheless, for the purpose of this study it is necessary to experiment with the predictions given by the levels of response by governments. Given the imperative for governments to enact effective measures in controlling the pandemic, the term "efficient" in this context pertains to evaluating the efficacy of interventions based on a defined objective function. Such functions may encompass various considerations, including economic metrics such as GDP per capita lost, alongside health-related aspects (for example, the value of lives lost)[10]. In this sense, Mahmoudi J (2022) and Santini et al. (2022) argue that lower rates of Covid-19 infections and mortality are correlated with the implementation of stricter enforcement policies and the imposition of more severe sanctions for non-compliance with control measures. Likewise, Qiu Z (2022) highlights that travel bans, school closures, economic activity shutdowns, and other restrictions have proven to be highly effective in curbing the transmission of Covid-19. However, a lingering question remains regarding the most effective containment policies to address the Covid-19 crisis: namely, whether the optimal strategy involves implementing public policies with either high or low levels of societal restrictions (Qiu Z, 2022).

6.3 What if scenarios

Estimates of scenarios and comparative predictions on the evaluation of the effectiveness of response policies (containment measures) to face the crisis of the Covid-19 pandemic are presented. These scenarios encompass the reduction of infections and fatalities, as well as the impact on the dynamics of economic growth (GDP). The predictions are driven by the Oxford Containment and Health Index, processed by the Oxford Coronavirus Government Response Tracker project (Hale et al., 2021)[11]. This index is computed by considering all ordinal containment and closure policy indicators, along with health system policy indicators. These indicators include a range of metrics such as school closures, vaccine distribution, testing policies, movement restrictions, cancellation of public events, implementation of face coverings, and others. The index is then normalized to a scale ranging from 0 to 100, with a score of 100 indicating the strictest level of policies.

The countries of New Zealand and Italy were selected as benchmarks in order to estimate the predictions. These countries were chosen because they imposed strict measures to face the Covid-19 pandemic. Hence, both Italy and New Zealand are important reference points (see Figure *7*). The regression models with the smallest MAE for training and validation were used (see Table 3, Table 4 and Table 5) to test these hypothetical scenarios where some features (columns) of the original dataset were modified by the values of the reference countries. That is, the values in the *containment_index* column in the dataset for Mexico, Italy, and the US have been replaced by the New Zealand and Italy values.

[10] Due to the dynamic nature of a pandemic, the application of dynamic optimization and optimal control theory may be appropriate (see, for example, Buratto et al. (2022)).

[11] The Containment and Health Index measures a mean value of the scope of 13 policy responses from governments around the world and covers 185 countries, regions, and territories. See: https://ourworldindata.org/grapher/covid-containment-and-health-index.

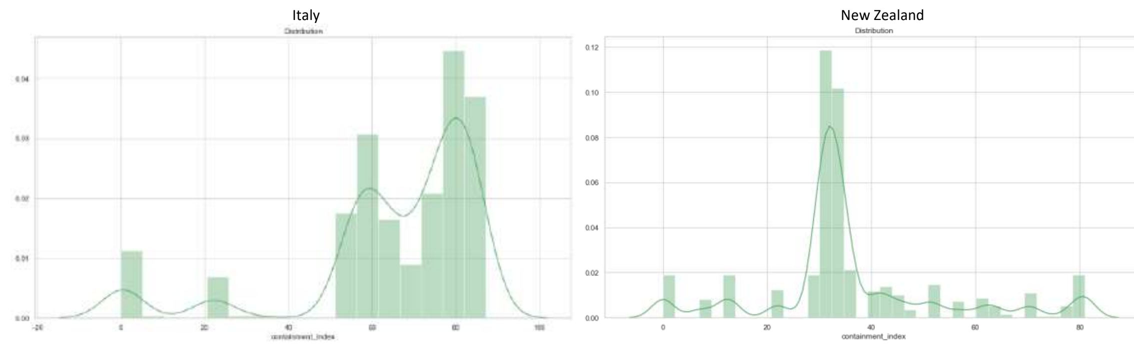

Figure 7. Containment health index distribution.
Source: Own elaboration.

Italy became the initial European nation to confront a significant outbreak of the Covid-19 pandemic and subsequently introduced various policies aimed at mitigating the virus's spread. In particular, lockdown policies, ergo, widespread mobility restrictions, have been used to combat the pandemic waves (Capano, 2020; Galeazzi et al., 2021). The New Zealand government response turned out to be one of the strictest in terms of lockdowns, resulting in low levels of disease disparities in the population and the initial achievement of elimination of Covid-19 (Gauld, 2022; Wilson, 2020). New Zealand moved from a national planning- led response to an influenza pandemic to a tailored approach to Covid-19 focused on suppression (stopping community spread of SARS-CoV-2) over mitigation (slowing transmission), with the goal of eliminating Covid-19. The New Zealand experience outlines the multifaceted components of a national response as a feasible route to the elimination of Covid-19. New Zealand was characterized by decisive and timely government leadership, combining rigorous measures of case detection, isolation, contact tracing and quarantine, driven by education and population participation (Baker et al., 2020; Jefferies et al., 2020).

The experimental scenarios are shown below. Each figure (Figure 8, Figure 9, Figure 10, Figure 11, and Figure 12) presents the three target variables (GDP, Total deaths, Reproduction rate) predicted using the best regression model for each country in the training phase. These are: Regression Tree for Italy and Mexico, and Random Forest for the US.

First, the Italian scenario considering the application of the New Zealand policy. Figure 8 shows the Italian scenario driven by the confinement measures in New Zealand, as can be seen the predictions fit quite well to the reality experienced in Italy. This indicates that both Italy and New Zealand implemented very similar measures with effects on GDP, the total number of deaths and the reproduction rate of Covid-19, under a singularly already given dynamic.

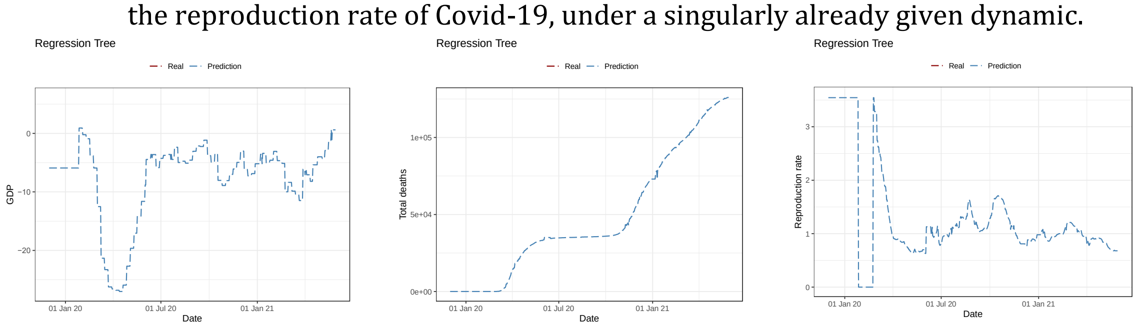

Figure 8. Italian scenario under the New Zealand containment index.
Source: Own elaboration.

DOI: https://doi.org/10.21919/remef.v20i1.911

One thing to keep in mind is that the Italian containment policy implemented by the central government was applied in such a way that the confinement policy was divided into regions and municipalities and each municipality was given an emergency color that was reviewed very frequently (sometimes weekly or every two weeks and even more often in some circumstances, such as during the Christmas holidays, where tougher policies were implemented due to the increased incentive to disobey). Each color represented a different degree of confinement (red = hard confinement; orange = medium confinement; yellow = soft confinement). During that period, all regions and municipalities alternated short periods of hard policies with periods of soft policies. This dynamic was broken in Italy only after a massive public vaccination policy that began at the end of February 2021 (Caselli et al., 2022; Merkaj E, 2022).

On the other hand, the scenarios for Mexico are presented in Figure 9 and Figure 10. Let us briefly explain the response of the Mexican government to the pandemic. Mexico has not pursued a complete lockdown strategy, opting instead for various partial lockdown measures. However, this approach to containment has faced criticism regarding its alignment with federal recommendations and mandates[12]. In a broader context, particularly in a deeply decentralized nation, it seems plausible to admit that modulating or clearly distinguishing the restrictions and recommendations is a reasonable strategy[13]. Regarding the economic, social and cultural aspects, it's crucial to acknowledge that nearly half of Mexico's employment falls under the "informal" sector[14]. Consequently, the implications are evident: limited savings capacity and challenges in accessing healthcare. These factors, coupled with low incomes, contribute to profound vulnerabilities in the region. Therefore, enforcing strict and prolonged lockdowns in a regulatory manner becomes nearly impractical.

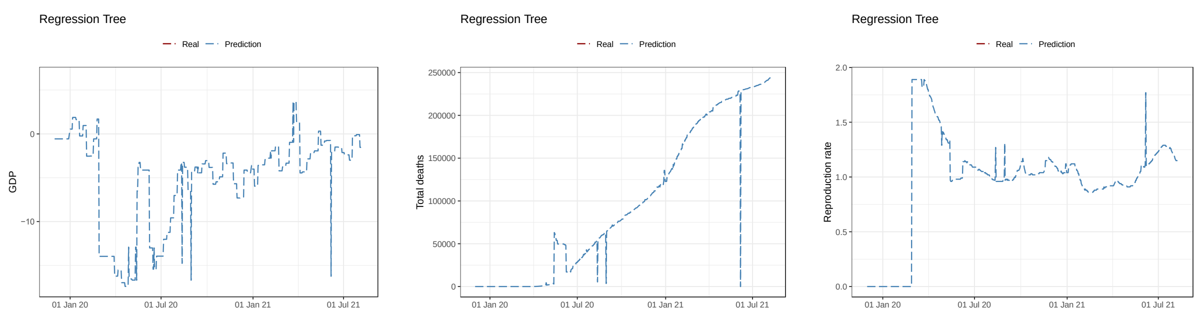

Figure 9. Mexican scenario under the New Zealand containment index.
Source: Own elaboration.

12 Mexico is made up of 32 entities that make up the country officially called: United Mexican States (*Estados Unidos Mexicanos*). Each entity has independent sovereignty. The heterogeneity of Mexico's cultural, political, and economic conditions shapes the government response to the pandemic (Campos-Vazquez & Esquivel, 2021).

13 For Mexico, a strict and prolonged confinement would have been possible only if the Federation had implemented a wide range of subsidiary resources, which was not the case. The only solution given, then, was the recommendation to stay at home (where it was possible to implement homeworking) and the healthy distance between people. In this context, it could be understandable that many of the policies remain as recommendations instead of requirements or regulations, and they are addressed precisely in the moral and subjective dimensions of citizens (de Anda-Jáuregui et al., 2022; Perea Tinajero & Bąk, 2021).

14 This is consistent with the general trend in the LA region where the rate of informal economic activities is around 62% (Perea-Tinajero & Bąk, 2021).

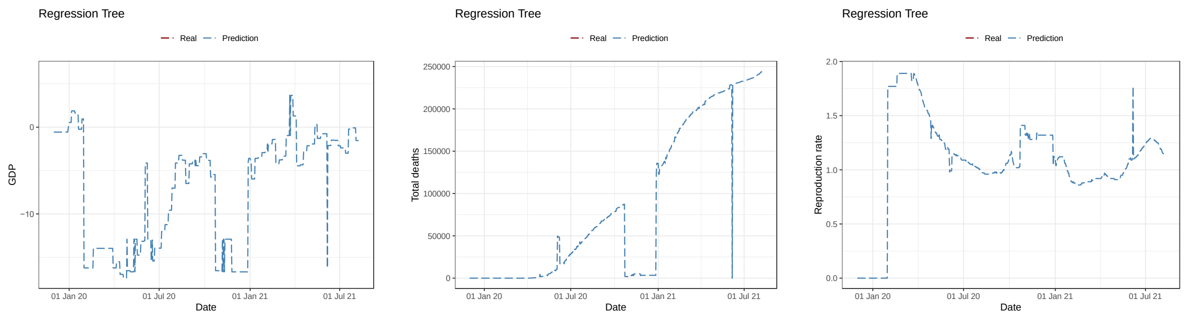

Figure 10. Mexican scenario under the Italian containment index.
Source: Own elaboration.

The scenario in which Mexico had followed the containment policy as applied in New Zealand (see Figure 9), then the total number of deaths from Covid-19 would have been more or less the same, with very marked jumps in more weekly deaths before July 2020, drops in the number of deaths between August and September 2020, and a significant drop in the number of deaths before the first week of July 2021. But outside of those weeks the trend in the number of deaths is practically the same as the real value. Similarly, the reproduction rate of Covid-19 in Mexico would have been forecast weekly for a few weeks of maximums (before July 2021, which is consistent with an increase in deaths) and minimums (before June 2020). In the same way, the expected Mexican GDP would have presented very contrasting ups and downs with respect to real GDP, but only for brief weekly periods. Overall, there is consistency in the trend of maximums and minimums between the actual and predicted values. In other words, Mexico was not a country that applied inefficient confinement policies, since if it had applied policies like those of New Zealand, in general terms it would have given the same results in terms of GDP, Reproduction Rate and Total Deaths.

If Mexico had applied a containment index like Italy's (shown in Figure 10), then the real and predicted results, although they fit well, it can be observed abrupt weekly changes of the predicted values with respect to the real ones. For example, it is interesting to note how the total weekly deaths prior to January 2021 (at least three consecutive weeks prior to that date) would have been considerably less than the real value, as well as the reproduction rate with high weekly levels prior to January 2021. But obviously reproduction and deaths do not occur in the same periods, since we well know that it takes weeks for an infected person to have severe symptoms and death or none at all. In fact, these drops in deaths can be explained by the slight drops in reproduction that would have occurred under the Italian containment policy that are observed around May 2020. It is difficult to explain the ups and downs that are observed in the predicted Mexican GDP. But as we have pointed out, it is interesting to observe how in the expected period of the drop in deaths applying the Italian policy also produces a considerable weekly drop in Mexican GDP (before January 2021). In other words, the application of the Italian policy seems to be efficient in terms of reducing the number of deaths but very costly in economic aspects. However, generally speaking, the forecast data fits the real data well. This confirms once again that the Mexican containment policy was not at all inefficient.

Revista Mexicana de Economía y Finanzas, Nueva Época, Vol. 20 No. 1, pp. 1-33, e911
DOI: https://doi.org/10.21919/remef.v20i1.911

The scenarios for the United States applying the measures of New Zealand (Figure 11) and Italy (Figure 12) are very interesting[15]. The predictions show periods of high economic costs (sharp falls in GDP in the weeks before January 2021), but also falls in the total number of deaths from Covid-19 (in the weeks before January 2021). Although reproduction rates, applying either the New Zealand or Italian containment policy, would have always been higher in the period from May 2020 to the end of that year.

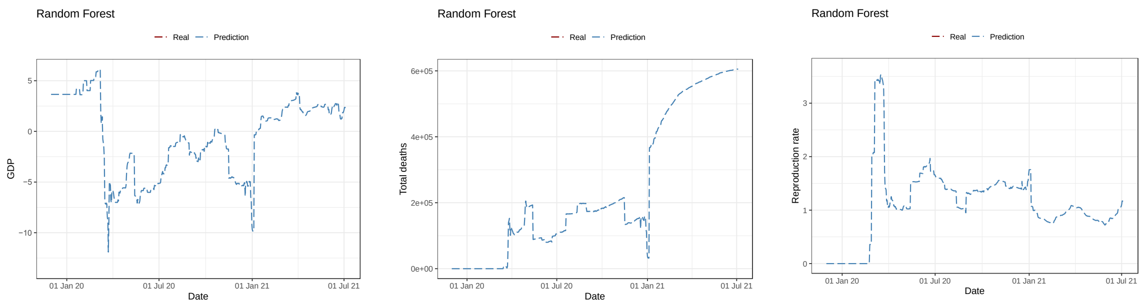

Figure 11. United States scenario under the New Zealand containment index.
Source: Own elaboration.

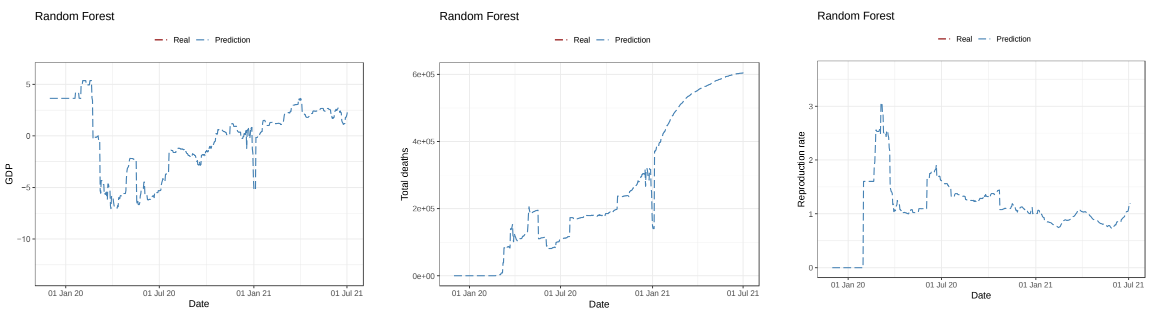

Figure 12. United States scenario under the Italian containment index.
Source: Own elaboration.

During the peak of restrictions in late March and early April 2020, over 310 million Americans were subject to directives spanning from "shelter in place" to "stay at home" (Abouk & Heydari, 2021; Walmsley et al., 2021). In May 2021, following the announcement that vaccinated individuals were no longer required to wear masks in public, the Centers for Disease Control and Prevention (CDC) reversed its stance amidst the delta wave. They recommended indoor mask-wearing for those residing in areas with high transmission rates. Subsequently, local governments reintroduced their own mask mandates. Later, during the omicron wave, several major cities enforced new vaccination proof requirements. Meanwhile, the CDC shortened the recommended isolation period for asymptomatic individuals who tested positive for the virus[16]. Benefitting from an early and swift

[15] Throughout the pandemic, United States officials have implemented a range of restrictions on social distancing, mask wearing, and other aspects of social life. The orders vary by the American decentralization: according to the American state, the county and even the city. See: https://eu.usatoday.com/storytelling/coronavirus-reopening-america-map/
[16] This latest move was aimed, at least in part, at addressing widespread worker shortages, including at struggling airlines during the height of the holiday travel season. See: https://www.pewresearch.org/2022/03/03/ two-years-into-the-pandemic-americans-inch-closer-to-a-new-normal/

vaccine rollout, along with robust fiscal support, the United States experienced a robust recovery. By the second quarter of 2021, US real GDP had surpassed its pre-pandemic levels, outpacing most other major economies[17].

7. Conclusions

In this paper, we used four ML algorithms (namely, Regression tree, random forest, Extreme gradient boosting, and K-nearest neighbor regression) on a dataset of weekly observations of several covid-related variables to predict GDP, Covid-19 reproduction rate and total deaths for four countries, namely Italy, New Zealand, US and Mexico. We first used these data to find the best model and specification with the highest predictive capacity for each country, then we simulated the outcome variables using, for three out of the four countries, a different policy, in order to predict what could have happened in terms of GDP, total deaths and reproduction rate. We predicted the “what if” scenario using the regression tree model for Italy and Mexico, and Random Forest for the US.

For what concerns Italy, we used as a counterfactual the containment policy implemented by New Zealand, and for US and Mexico, we used both policies implemented by Italy and New Zealand. Italy and New Zealand's policies were chosen as a “benchmark” because they have been recognized as the most restrictive during the first months of pandemic, which have no similar in other countries.

The prediction of GDP, total deaths and Covid reproduction rate in Italy if the New Zealand's containment policy was implemented does not produce substantially different results from what Italy recorded in reality, as a consequence of implementing its own policy. Italy and New Zealand, after all, are well known cases of very hard lockdown implemented by their governments, and it is reasonable to think that similar policies could produce similar outcomes.

Mexico, contrary to Italy and New Zealand, has not sought a total lockdown, although different States have implemented different partial lockdown policies (often disattending federal recommendations and guidelines). One important characteristic of the Mexican economy is, contrary to the other three countries, the preponderance of the ``informal sector'' which is difficult to regulate, and which causes low saving capacity for the population and a low access to healthcare services and facilities. However, the ``what if'' Mexican scenario if the New Zealand policy was implemented did not produce substantially different results from what we observed in reality after September 2020. Substantial differences are represented by marked positive jumps in deaths before July 2020, which are counterbalanced by drops between August and September 2020, and the first week of July 2021. The estimation of the covid reproduction rate follows substantially the pattern observed for deaths, but for what concerns GDP in the “what if” scenario presents contrasting ups and downs, but only for few weeks in the whole period, leaving substantially unchanged the estimated values of GDP with respect to the realized values.

If Mexico, instead, had implemented Italy's containment policy, our estimations predict that the number of total deaths prior January 2021 would have been reduced with respect the real value,

[17] Output growth picked up in the euro area and Canada in the third quarter of 2021; but by the end of 2021, output in most major US trading partners had just reached its pre-pandemic level, while US output was three percent higher than before the pandemic. See: https://www.whitehouse.gov/wp-content/uploads/2022/04/Chapter-3-new.pdf

which can be explained by the reduction of the covid reproduction rate which we observe in the "what if" scenario during May 2020. The greatest differences however are observed for the predictions of GDP with respect to the observed real value. Under the Italian containment policy, Mexico would have suffered big losses in GDP. So, if the Italian containment policy would probably have been more effective in reducing the number of deaths and the spread of the virus, it would have been highly depressing for GDP.

For US instead, the "what if" scenarios under the containment policies of both Italy and New Zealand would have predicted a higher GDP prior June 2020, but also a higher number of deaths. The reproduction rate of the virus in this period is substantially the same under the New Zealand containment policy, but it is higher under the Italian policy prior February 2020, and it is lower between February and March 2020.

Both Italian and New Zealand's policy would have predicted a lower GDP and number of deaths after June 2020 with respect to the real value, but a higher reproduction rate. So, for US, both "what if" scenarios would probably have generated less deaths, but also less income, at least after June 2020. In this case, the result might be due to a different vaccination policy (especially for what concerns the timing of mass vaccination), which, in Mexico, occurred later and that can justify the contemporaneous presence of a higher reproduction rate of the virus with a substantial drop in the number of deaths.

So, what can be concluded from this paper - beyond what we have learned from an economic and political perspective - is that ML is a powerful tool which can successfully guide data-driven decision making. Future works may consist of implementing more advanced hyper-parameter tuning techniques such as Bayesian optimization, Genetic Algorithms etc., developing a metaheuristic hybrid model in order to perform an optimal selection of feature, which we made by simply excluding, from all the couples, some variables if they are correlated more than 0.7, which we deliberately chose as a threshold in order to avoid potential overfitting problems.

Supplementary Materials: There is no Supplementary Material citation.

Declarations: Opinions expressed in this research paper are those of the authors and do not necessarily reflect the official opinion of the Italian Customs and Monopolies Agency. The authors did not receive support from any organization for the submitted work. The authors have no relevant financial or non-financial interests to disclose.

Author Contributions: J. Alejandro Navarro-Acosta: Conceptualization, simulations, numerical test, methodology, investigation, writing- original draft preparation, review and editing. **Valeria Soto-Mendoza:** Data gathering, simulations, methodology, investigation, writing- original draft preparation, review and editing. **Laura Policardo:** Formal analysis, investigation, writing- original draft preparation, review and editing. **Edgar J. Sánchez-Carrera:** Data gathering, methodology, formal analysis, investigation, writing- original draft preparation, review and editing. The authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The reader can find the input data on public websites.

Conflicts of Interest: The authors declare no conflict of interest in the tests and results presented herein.

References

[1] Abouk, R., & Heydari, B. (2021). The Immediate Effect of COVID-19 Policies on Social-Distancing Behavior in the United States. *Public Health Reports*, *136*(2), 245-252. https://doi.org/10.1177/0033354920976575
[2] Arlot, S., & Celisse, A. (2010). *A survey of cross-validation procedures for model selection*. https://doi.org/10.1080/10.1214/09-SS054
[3] Athey, S. (2018). The impact of machine learning on economics. En *The economics of artificial intelligence: An agenda* (pp. 507-547). University of Chicago Press.
[4] Bajari, P., Nekipelov, D., Ryan, S. P., & Yang, M. (2015). Machine learning methods for demand estimation. *American Economic Review*, *105*(5), 481-485.
[5] Baker, M. G., Wilson, N., & Anglemyer, A. (2020). Successful Elimination of Covid-19 Transmission in New Zealand. *The new england journal of medicine*, *338*(8). https://doi.org/10.1056/NEJMc2025203
[6] Bischi, G. I., Grassetti, F., & Carrera, E. J. S. (2022). On the economic growth equilibria during the Covid-19 pandemic. *Communications in Nonlinear Science and Numerical Simulation*, *112*, 106573. https://doi.org/10.1016/j.cnsns.2022.106573
[7] Borchani, H., Varando, G., Bielza, C., & Larranaga, P. (2015). A survey on multi-output regression. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *5*(5), 216-233.
[8] Breiman, L., & Friedman, J. H. (1997). Predicting multivariate responses in multiple linear regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *59*(1), 3-54.
[9] Browne, M. W. (2000). Cross-validation methods. *Journal of mathematical psychology*, *44*(1), 108-132. https://doi.org/10.1006/jmps.1999.1279
[10] Buratto, A., Muttoni, M., Wrzaczek, S., & Freiberger, M. (2022). Should the COVID-19 lockdown be relaxed or intensified in case a vaccine becomes available? *PLoS One*, *17*(9). https://doi.org/10.1371/journal.pone.0273557
[11] Campos-Vazquez, R. M., & Esquivel, G. (2021). Consumption and geographic mobility in pandemic times. Evidence from Mexico. *Review of Economics of the Household*, *19*, 353-371. https://doi.org/10.1007/s11150-020-09539-2
[12] Capano, G. (2020). Policy design and state capacity in the COVID-19 emergency in Italy: If you are not prepared for the (un)expected, you can be only what you already are. *Policy and Society*, *39*(3), 326-344. https://doi.org/10.1080/14494035.2020.1783790
[13] Caselli, M., Fracasso, A., & Scicchitano, S. (2022). From the lockdown to the new normal: Individual mobility and local labor market characteristics following the COVID-19 pandemic in Italy. *J Popul Econ*, *35*. https://doi.org/10.1007/s00148-022-00891-4
[14] Celestine Iwendi, C. C., C. G. Y. Huescas, & Mohan, S. (2024). COVID-19 health analysis and prediction using machine learning algorithms for Mexico and Brazil patients. *Journal of Experimental & Theoretical Artificial Intelligence*, *36*(3), 315-335. https://doi.org/10.1080/0952813X.2022.2058097
[15] de Anda-Jáuregui, G., García-García, L., & Hernández-Lemus, E. (2022). Modular reactivation of Mexico City after COVID-19 lockdown. *BMC Public Health*, *22*(1), 961. https://doi.org/10.1186/s12889-022-

13183-z

[16] De Veaux, R. D., & Ungar, L. H. (1994). Multicollinearity: A tale of two nonparametric regressions. *Selecting Models from Data*, 393-402. https://doi.org/10.1007/978-1-4612-2660-4_40

[17] De Ville, B. (2013). Decision trees. *Wiley Interdisciplinary Reviews: Computational Statistics*, *5*(6), 448-455.

[18] Einav, L., & Levin, J. (2014). Economics in the age of big data. *Science*, *346*(6210), 1243089.

[19] Fischer, T., Krauss, C., & Treichel, A. (2018). *Machine learning for time series forecasting-a simulation study*. FAU Discussion Papers in Economics.

[20] Galeazzi, A., Cinelli, M., Bonaccorsi, G., Pierri, F., Schmidt, A.-L., Scala, A., Pammolli, F., & Quattrociocchi, W. (2021). Human mobility in response to COVID-19 in France, Italy and UK. *Scientific Reports*, *11*, 1-10. https://doi.org/10.1038/s41598-021-92399-2

[21] Gauld, R. (2022). A review of public policies on Covid-19: The New Zealand experience. *Public Administration and Policy: An Asia-Pacific Journal*. https://doi.org/10.1108/PAP-04-2022-0028

[22] Gauld, R. (2023). A review of public policies on Covid-19: The NewZealand experience. *Public Administration and Policy*, *26*(1), 10-20.

[23] Genuer, R., & Poggi, J.-M. (2020). Random Forests. En *Random Forests with R* (pp. 33-55). Springer International Publishing. https://doi.org/10.1007/978-3-030-56485-8_3

[24] Gogas, P., & Papadimitriou, T. (2021). Machine Learning in Economics and Finance. *Computational Economics*, *57(1)*, 1-4.

[25] Gubar, E., Policardo, L., Sánchez Carrera, E. J., & Taynitskiy, V. (2023). On optimal lockdown policies while facing socioeconomic costs. *Annals of Operations Research*, 1-34. https://doi.org/10.1007/s10479-023-05454-8

[26] Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & others. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature human behaviour*, *5*(4), 529-538.

[27] Hu, X., Wang, J., Wang, L., & Yu, K. (2022). K-Nearest Neighbor Estimation of Functional Nonparametric Regression Model under NA Samples. *Axioms*, *11*(3).

[28] Izquierdo-Verdiguier, E., & Zurita-Milla, R. (2020). An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, *88*, 102051.

[29] Jefferies, S., French, N., Gilkison, C., Graham, G., Hope, V., Marshall, J., McElnay, C., McNeill, A., Muellner, P., Paine, S., Prasad, N., Scott, J., Sherwood, J., Yang, L., & Priest, P. (2020). COVID-19 in New Zealand and the impact of the national response: A descriptive epidemiological study. *The Lancet Public Health*, *5*(11), e612-e623. https://doi.org/10.1016/S2468-2667(20)30225-5

[30] Jung, Y., & Hu, J. (2015). AK-fold averaging cross-validation procedure. *Journal of nonparametric statistics*, *27*(2), 167-179. https://doi.org/10.1080/10485252.2015.1010532

[31] Katris, C. (2020). Prediction of unemployment rates with time series and machine learning techniques. *Computational Economics*, *55*(2), 673-706.

[32] Knaul, F., Arreola-Ornelas, H., Porteny, T., Touchton, M., Sánchez-Talanquer, M., Méndez, Ó., Chertorivski, S., Ortega, S., Chudnovsky, M., Kuri, P., & Americas, group from the O. for the C. of C.-19 in the. (2021). Not far enough: Public health policies to combat COVID-19 in Mexico's states. *PLoS ONE*, *16*(6). https://doi.org/10.1371/journal.pone.0251722

[33] Kumar, Y., Koul, A., Kaur, S., & Hu, Y.-C. (2022). Machine learning and deep learning based time series prediction and forecasting of ten nations' COVID-19 pandemic. *SN Computer Science*, *4*(1), 91. https://doi.org/10.1007/s42979-022-01493-3

[34] Liao, Y. (2017). Machine learning in macro-economic series forecasting. *International Journal of Economics and Finance*, *9*(12), 71-76.

[35] Loh, W.-Y. (2014). Fifty years of classification and regression trees. *International Statistical Review*, *82*(3), 329-348.

[36] Maccarrone, G., Morelli, G., & Spadaccini, S. (2021). GDP Forecasting: Machine Learning, Linear or

Autoregression? *Frontiers in Artificial Intelligence*, *4*.

[37] Macdonald, S., Correia, S. G., & Watkin, A.-L. (2019). Regulating terrorist content on social media: Automation and the rule of law. *International Journal of Law in Context*, *15*(2), 183-197.

[38] Madjarov, G., Kocev, D., Gjorgjevikj, D., & Džeroski, S. (2012). An extensive experimental comparison of methods for multi-label learning. *Pattern recognition*, *45*(9), 3084-3104.

[39] Maehashi, K., & Shintani, M. (2020). Macroeconomic forecasting using factor models and machine learning: An application to Japan. *Journal of the Japanese and International Economies*, *58*, 101104. https://doi.org/10.1016/j.jjie.2020.101104

[40] Mahmoudi J, X. C. (2022). How social distancing, mobility, and preventive policies affect COVID-19 outcomes: Big data-driven evidence from the District of Columbia-Maryland-Virginia (DMV) megaregion. *PLoS One*, *17*(2). https://doi.org/10.1371/journal.pone.0263820

[41] Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., Ortiz-Ospina, E., & Roser, M. (2020). Coronavirus Pandemic (COVID-19). *Our World in Data*.

[42] Meade, N. (2002). A comparison of the accuracy of short term foreign exchange forecasting methods. *International Journal of Forecasting*, *18*(1), 67-83.

[43] Merkaj E, S. R. (2022). Italian national policies in response to the COVID-19 pandemic: The case of the Friuli-Venezia-Giulia and Umbria Regions. *Health Policy*, *126*(4), 287-293. https://doi.org/10.1016/j.healthpol.2022.02.0042

[44] Mesafint, D., & Huchaiah, M. D. (2021). Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. *International Journal of Computers and Applications*, *44*, 1-12. https://doi.org/10.1080/1206212X.2021.1974663

[45] Monroe, B. L., Pan, J., Roberts, M. E., Sen, M., & Sinclair, B. (2015). No! Formal theory, causal inference, and big data are not contradictory trends in political science. *PS: Political Science & Politics*, *48*(1), 71-74.

[46] Perea Tinajero, G., & Bąk, A. (2021). To lock or not to lock? Mexico case. *History and Philosophy of the Life Sciences*, *43*(4), 117. https://doi.org/10.1007/s40656-021-00462-y

[47] Perea-Tinajero, G., & Bąk, A. (2021). To lock or not to lock? Mexico case. *History and Philosophy of the Life Sciences*, *43*(117). https://doi.org/10.1007/s40656-021-00462-y

[48] Puka, L. (2011). Kendall's Tau. En M. Lovric (Ed.), *International Encyclopedia of Statistical Science* (pp. 713-715). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-04898-2_324

[49] Qiu Z, D. X., Cao Z, Zou M, (. .. ..), Wang D. (2022). The effectiveness of governmental nonpharmaceutical interventions against COVID-19 at controlling seasonal influenza transmission: An ecological study. *BMC Infect Dis*, *22*(1). https://doi.org/10.1186/s12879-022-07317-2

[50] Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, *1*(1), 81-106.

[51] Santini, G., Fordellone, M., Boffo, S., Signoriello, S., De Vito, D., & Chiodini, P. (2022). Modeling for the Stringency of Lock-Down Policies: Effects of Macroeconomic and Healthcare Variables in Response to the COVID-19 Pandemic. *Frontiers in Public Health*, *10*. https://doi.org/10.3389/fpubh.2022.872704

[52] Sengupta, S., Mugde, S., & Sharma, G. (2020). Covid-19 pandemic data analysis and forecasting using machine learning algorithms. *MedRxiv*, 2020-06. https://doi.org/10.1101/2020.06.25.20140004

[53] Similä, T., & Tikka, J. (2007). Input selection and shrinkage in multiresponse linear regression. *Computational Statistics & Data Analysis*, *52*(1), 406-422.

[54] Tiwari, D., Bhati, B. S., Al-Turjman, F., & Nagpal, B. (2022). Pandemic coronavirus disease (Covid-19): World effects analysis and prediction using machine-learning techniques. *Expert Systems*, *39*(3), e12714. https://doi.org/10.1111/exsy.12714

[55] Tsochantaridis, I., Joachims, T., Hofmann, T., Altun, Y., & Singer, Y. (2005). Large margin methods for structured and interdependent output variables. *Journal of machine learning research*, *6*(9).

[56] Tsoumakas, G., Spyromitros-Xioufis, E., Vrekou, A., & Vlahavas, I. (2014). Multi-target regression via random linear target combinations. *Joint european conference on machine learning and knowledge discovery in databases*, 225-240.

[57] Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), 3-28.

[58] Walmsley, T., Rose, A., & Wei, D. (2021). The Impacts of the Coronavirus on the Economy of the United States. *Economics of Disasters and Climate Change*, *5*, 1-52. https://doi.org/10.1007/s41885-020-00080-1

[59] Wang, F., H. ,. Yang, & Luo, Z. (2016). An experimental study of the intrinsic stability of random forest variable importance measures. *BMC Bioinformatics*, *17*.

[60] Wilson, S. (2020). Pandemic leadership: Lessons from New Zealand's approach to COVID-19. *Leadership*, *16*(3), 279-293. https://doi.org/10.1177/1742715020929151

[61] Woloszko, N. (2020). *Tracking activity in real time with Google Trends*.

[62] Xu, D., Shi, Y., Tsang, I. W., Ong, Y.-S., Gong, C., & Shen, X. (2019a). Survey on multi-output learning. *IEEE transactions on neural networks and learning systems*, *31*(7), 2409-2429.

[63] Xu, D., Shi, Y., Tsang, I. W., Ong, Y.-S., Gong, C., & Shen, X. (2019b). Survey on multi-output learning. *IEEE transactions on neural networks and learning systems*, *31*(7), 2409-2429. https://doi.org/10.1109/TNNLS.2019.2945133

[64] Yamamoto, N., Jiang, B., & Wang, H. (2021). Quantifying compliance with COVID-19 mitigation policies in the US: A mathematical modeling study. *Infectious Disease Modelling*, *6*, 503-513. https://doi.org/10.1016/j.idm.2021.02.004