

Technology Management for Accelerated Recovery during COVID-19: A Data-Driven Machine Learning Approach

Swapnil Morande *¹ , Dr. Veena Tewari ²

¹ University of Naples, Italy

² Majan University College, Oman

* Corresponding author: swapnil.morande@unina.it

Article History

Received 2020-08-23

Accepted 2020-08-31

Published 2020-09-06

Keywords

Machine Learning

Artificial Intelligence

COVID

Healthcare

Predictive Analytics

How to cite?

Morande, S., & Tewari, D. V. (2020).
Technology Management for Accelerated Recovery during COVID: A Data-Driven Machine Learning Approach.
SEISENSE Journal of Management,
3(5), 00-00. doi:
[10.33215/sjom.v3i5.445](https://doi.org/10.33215/sjom.v3i5.445)

Copyright © 2020 The Author(s)



Objective- The research looks forward to extracting strategies for accelerated recovery during the ongoing Covid-19 pandemic.

Design - Research design considers quantitative methodology and evaluates significant factors from 170 countries to deploy supervised and unsupervised Machine Learning techniques to generate non-trivial predictions.

Findings - Findings presented by the research reflect on data-driven observation applicable at the macro level and provide healthcare-oriented insights for governing authorities.

Policy Implications - Research provides interpretability of Machine Learning models regarding several aspects of the pandemic that can be leveraged for optimizing treatment protocols.

Originality - Research makes use of curated near-time data to identify significant correlations keeping emerging economies at the center stage. Considering the current state of clinical trial research reflects on parallel non-clinical strategies to co-exist with the Coronavirus.

Introduction

The discovery of Acute Respiratory Syndrome Coronavirus-2 (also known as Covid-19) as a deadly virus has put the globe on high alert. With a high infection rate and rising death toll, it has changed the way we live in society. It is beyond doubt that Coronavirus has caused massive economic distress around the world, leading to business interruptions and regional shutdowns (Martin et al., 2020a). The negative impact of Covid-19 on the economy is visible on the disruption of the supply chain and extends to stumbling socio-economic activities (Mhalla, 2020; Phillipson et al., 2020). The associated uncertainty of the pandemic is not only limited to the economic aspect but also increasing the psychological distress of people (Silva Junior et al., 2020).

As scientists are striving to contain the virus, it has become essential to adopt an approach that can take calculated risk to bring the situation to normalcy. While health professionals denounced the certification of the Russian Covid-19 vaccine as premature, as of now, social distancing appears to be the most effective measure for combating the pandemic (Mahase, 2020). Social distancing may work in high-income countries; however, low-income countries may not enjoy such benefits. On the contrary, the limitation posed due to lockdown in underdeveloped nations may lead to deprivation of people and a steep decline in the economy. This fact can make social distancing reflect poorly and make emerging countries vulnerable, and hence the exploration of alternative strategies appears to be a need of time (Barnett-Howell & Mobarak, 2020).

These alternative strategies can be drawn with the help of a technology called Artificial Intelligence (A.I.). Artificial Intelligence is viewed as a revolution in the current era, and it possesses an immense application in several areas, including healthcare. Machine Learning (ML) is the branch of Artificial Intelligence that can solve real-life problems by learning from data rather than being explicitly programmed (Saeb et al., 2016). Such digital transformation presents an opportunity to address issues across multiple facets of healthcare and administration using Machine Learning models (Feldman et al., 2017). While we are exposed to the Covid-19 pandemic, it presents an ideal scenario where we can exploit a digital transformation for the immediate benefits of society.

Considering the current state of a pandemic, it has become necessary to seek support from areas such as statistics and computer science. Accordingly, research makes use of state-of-the-art technology to extract an accelerated recovery plan with a data-driven approach. Researchers can deploy Machine Learning in the healthcare sector to predict future possibilities and propose effective strategies to deal with the Covid-19 outbreak. Based on the predictions from Machine Learning models, administrators can focus on planning and decision making during the pandemic (Madurai Elavarasan & Pugazhendhi, 2020).

The proposed research looks forward to fulfilling the same using data aggregated from the nations that have unfortunately been subjected to the ongoing pandemic. Research prominently demonstrates that Machine Learning can be used to discover interesting patterns and support sound decision-making. With efficient technology management, the study draws parallel plans to co-exist with Covid-19 while clinical trials await approval.

Literature Review

Impact of Coronavirus

Since it first appeared, Covid-19 has turned into a life-threatening pandemic that has pushed global health and socio-economic activities to the breaking point. A model developed to evaluate the socio-economic impact of pandemic estimates the negative effect of social-distancing on household income, savings, consumption, and poverty (Martin et al., 2020b). It is expected that the emerging world, including low and middle-income countries, will undergo more significant concerns during an ongoing pandemic (Klinger et al., 2020).

Countries at present are trying to inhibit the spread of the Coronavirus by adopting various strategies. While a lockdown can be an effective strategy of social distancing and successful in tackling the rapid spread of the infectious Covid-19 virus, it can also have a negative impact on society (Zhang et al., 2020). This fact is in agreement with research manifested in China by Zhu et al. (2020), where people experienced enormous psychological impact due to the pandemic.

Associated Factors

Prior research has established that the odds of death due to the Covid-19 infection increases along with age (Talukder et al., 2019). The study further elaborated on the types of diseases like hypertension, diabetes, and cardiovascular (referred to as comorbidities) that have a significant direct impact on infected patients. Apart from age and existing health conditions, there are a number of factors that can affect the recovery of Covid-19 patients. According to Livadiotis (2020), the statistical analysis of the impact of environmental temperature on Covid-19 patients suggests a negative correlation on the growth rate of infected cases. It has been suggested that Bacillus Calmette-Guérin (BCG) vaccination may also have a protective effect on the Covid-19 mortality rate. The data from publicly available resources also indicate that both Covid-19 incidences and deaths are associated with regional BCG vaccination policies. However, as per Miyasaka (2020), if BCG vaccination does contribute to lower Covid-19 mortality, it is not the only factor. The presented study looks for such associations in a systematic manner.

Another independent research by Klinger et al. (2020) concluded that the inverse correlation with BCG vaccine administration and the validated role of the young population in the spread of Covid-19 calls for revisiting the BCG immunization policies. It must be noted that BCG vaccination is routine and near-universal in many low and middle-income countries. Data suggests that many countries in Asia, the Gulf, and Africa have maintained a flat mortality rate that can be correlated with the BCG vaccination policies (Debnath et al., 2020; Gursel & Gursel, 2020).

Recent studies have observed the dietary patterns where the ability of curcumin (a type of spice) has been found to have an inhibitory potential against different types of viral infections, making it a fit for resisting the coronavirus infection (Zahedipour et al., 2020). As per the guidance provided by Rozenfeld et al. (2020), the research has also identified additional factors associated with Covid-19. Considering lifestyle choices, the current study further explores associations with smoking as well as drinking habits (Chodkiewicz et al., 2020; Reddy et al., 2020; Sidor & Rzymiski, 2020). Further, building on observations made by other studies, the proposed research attempts to re-confirm relationships among factors such as population density and pollution index (Contini & Costabile, 2020; Rashed et al., 2020; Rocklöv & Sjödin, 2020).

Treatment of Coronavirus

According to Jean & Hsueh (2020), several countries are working on medical trials that may deliver potentially applicable solutions to handle patients with various levels of Covid-19 infections. Although the Russian government has endorsed a Covid-19 vaccine - Sputnik V - it followed only a limited trial with no published results (Mahase, 2020).

Researchers believe that many of the reviewed trials lack the robust methodology required to make a sound conclusion (Sethi & Bach, 2020). Their study also suggests, these therapeutics still need a considerable level of investigation to establish efficacy and determine adverse effect profiles. Arguably, there is no valid confirmation of well-designed randomized and completed controlled studies for Covid-19 therapy.

Presented Challenges

At the same time, the existing set of drugs have been deemed as promising candidates for controlling Covid-19. These drugs have specific safety profiles. Although drug re-purposing is an essential step against the fight with Coronavirus, it may require caution. As a number of clinical trials that are underway, it is safe to assume that their results might help us defeat Covid-19, but it is likely to take a considerable amount of time (Scavone et al., 2020).

The evidence further indicates that more data is required to determine whether any therapeutic agent has strong efficacy in the treatment of Covid-19. While society is being subjected to Coronavirus's fury and working to develop an effective cure, we need to be aware that human-to-human virus spread is skyrocketing. It may result in an exponential rise of Covid-19 cases, and a sheer number of infected people can be overwhelming for populous nations. This presents a challenge to humanity in terms of limited information, limited time, and increased infection rate.

Application of Machine Learning

The above considerations in dealing with Coronavirus pave a way to address concerns using Artificial Intelligence as a method of choice. The transformative potential of Machine Learning in healthcare supports the prediction made by Alan Turing (1950) that 'machine intelligence' will have a pervasive role within our society (Ashrafian et al., 2015).

As per Tan et al. (2020), the evolution of deep learning mechanism as a sub-discipline of Machine Learning requires minimal user input and demonstrates the tremendous potential to identify patterns. The patterns retrieved using a computing algorithm can facilitate healthcare interventions as a powerful approach. (Ashrafian & Darzi, 2018; Saria et al., 2018) Machine Learning guided outcomes can offer new inroads through enhanced health-related screening (Goshen et al., 2018). Such digital transformation can help governing bodies to determine the applicability of the Machine Learning models and discover trends that are related to Covid-19. The most prominent contribution of A.I. to health policy knowledge currently resides within the application of ML to large, population-level datasets, as presented in current research.

Recommended Strategies

The pandemic originated due to Coronavirus is going to have a long-lasting global impact. It is expected that the effect of this pandemic would harshly reflect on our lives and reverberate for some time to come. Based on the review of literature, the presented research works towards a new line of research, keeping in mind that society may not be able to sustain the restriction posed by Covid-19 for too long. Without effective treatment protocol, the placement of cities in 'lockdown' can affect economies on a multi-lateral level, including both social and economic standpoints. Every country cannot afford extended lockdowns as it may affect people with limited financial means, and hence, the best approach includes a transition to live with Coronavirus.

Considering the relationships between the factors that caused the virulence of the respiratory disease, the measures are needed to be placed to control the pandemic. Also, the impact associated with the disease is necessary to be understood scientifically. Thus, the detection and management of the Covid-19 can become increasingly dependent upon the predictive capabilities of the technological backbone (Allam & Jones, 2020).

In the absence of vaccine and treatment to cure Covid-19, it is suggested that a data-driven reflection needs to be carried out. Proposed research makes the use of Machine Learning to fulfill the same for accelerated recovery. In the process, it looks forward to identifying significant correlations that are associated with Coronavirus. The research includes both supervised learning & unsupervised learning for greater accuracy.

Supervised Machine Learning identifies ‘field Importance’ and ‘partial dependence’ to identify correlations across variables at the same time, unsupervised Machine Learning enlightens ‘associations’ as well as ‘component weightage’ using the curated data. Using predictive analytics given study evaluates several permutations and combinations to offer recommendations. Such interaction and integration with near-time dataset using the Machine Learning may yield an effective containment of Coronavirus outbursts. In the current state of Covid-19, the data-driven prediction will help the society to find means to survive during the pandemic, while it lasts.

Research Objective

Research Gap

Bluhm et al. (2020) believe that the ongoing pandemic may last considerably longer, and there may be a second wave of infection. Presuming the vaccine and its mass administration is still out of reach, existing measures of social distancing do not appear to be sustainable.

As established earlier, emerging economies may suffer substantial losses by the time clinical trials develop control over the pandemic. This fact brings us to the development of parallel strategies to deal with Covid-19.

Arguably the least explored area of Artificial Intelligence in healthcare is its role in governance, and researchers have also observed the limited impact A.I. has had in the management of Covid-19 (Ashrafian & Darzi, 2018; Hu et al., 2020). It reflects the need for translating insights from Machine Learning models to Covid-19 affected environments.

Research Questions

In consideration of insights retrieved via Machine Learning, it is possible to open the global economies and establish the normalcy in economic activities.

To achieve the same, the strategic review of Covid-19 related factors needs to be made by governing authorities.

Hence proposed research aims -

- To extract a data-driven solution to contain the current virus outbreak
 - That is sustainable in emerging economies
 - That can be applied to communities
- To identify strategies to avoid the future virus outbreak
 - That exist without strict social distancing norms
 - That manages the coronavirus outbreak at a macro level

Research Methodology

Data Collection

The research followed the ‘quantitative research’ methodology to be able to draw insights from curated datasets. A total of 170 countries was selected as a sample based on the gravity of impact experienced during the Covid-19 event. The sample size represented Africa, Asia & Pacific, Europe, Oceania, Middle Eastern, and the Arab States as well as Southern & Northern American regions to achieve accurate representation. The identification of direct and indirect measures was made through a comprehensive review of the literature. In line with the assessment of the literature, eleven measures were identified for data collection.

The range of the collected dataset included the very first reported instance of Coronavirus until July 2020. Further scientific evidence was compiled from various sources and consolidated for predictive modeling. The

dataset related to food consumption patterns was collected from Helgi Library (2020), and global vaccination policies were retrieved from BCG World Atlas (2020) to identify relevant associations.

Data Modelling

The data staging was done with the Covid-19' Recovery Index' as the target value. Lower target value represented better outcomes (or improved recovery rate) arising from the pandemic. As per Ooms & Spruit (2020), there have been many instances where data science has leveraged the use of Machine Learning techniques in the healthcare sector, and considering the same, a set of models were created and evaluated to generate Covid-19 related predictions. Machine Learning offered the ability to program real-world problems, explicitly using computer algorithms and statistical techniques. To accomplish the research outcome, a Machine Learning model was trained to identify the probable effect of Covid-19.

In general, Machine Learning is categorized in supervised (i.e., consists of output variables that are predicted from input variables) or unsupervised (i.e., deals with clustering of different groups for an intervention) learning process that simulates complex real-life scenarios. (Battineni et al., 2020) In the given study, supervised Machine Learning was achieved using Ensemble Modelling, while unsupervised learning was carried using Associations Modelling. The default ML model was built using Decision Tree (DT) algorithms, where the datasets were configured for identifying the 'Recovery Index' of Covid-19 as the predictive outcome. During model building, a supervised Machine Learning algorithm required a dataset that is split into a 'training' data set and a 'test' or 'validation' data set. A Machine Learning model was then trained using 80% of the data to predict a Recovery Index accurately. As the number of measures reflected on the Covid-19 Recovery Index, the accuracy of predicting the same improved significantly over the period. Post model training, the 'trained' Machine Learning model was validated – post evaluation of its predictive performance. The evaluation was performed using 20% of the test data.

The performance of the algorithm was tested by splitting the initial data through the random sampling, into training and testing sets being mutually exclusive subsets. This event enabled validation of the Recovery Index in predicted versus actual scenarios to meet a predetermined threshold for model evaluations. This process reflected positively on both the reliability and validity of the model. The 'validated' Machine Learning model was then applied to a new dataset for gauging target value. Model deployment supported the identification of factors and observed significant correlations that predicted the Covid-19 Recovery Index. This approach of model deployment to predict the Covid-19 Recovery Index was a key component of technology management that generated recommendations for accelerated recovery.

Data Analysis

Machine Learning Model

According to Caballé et al. (2020) in Artificial Intelligence, Machine Learning stands out as a method for providing tools for intelligent data analysis. There are specific algorithms that are used to develop models with predictive capabilities, and the selection of such algorithms depends on the research objective. The algorithm learns from the existing data, can be used to build complex models and predicts the target values (Arpitha et al., 2018). However, as per Feldman et al. (2017), the ability to derive informative insights requires more than the processing and execution of Machine Learning models. Rather, it calls for a deeper understanding of the data on which the models are executed.

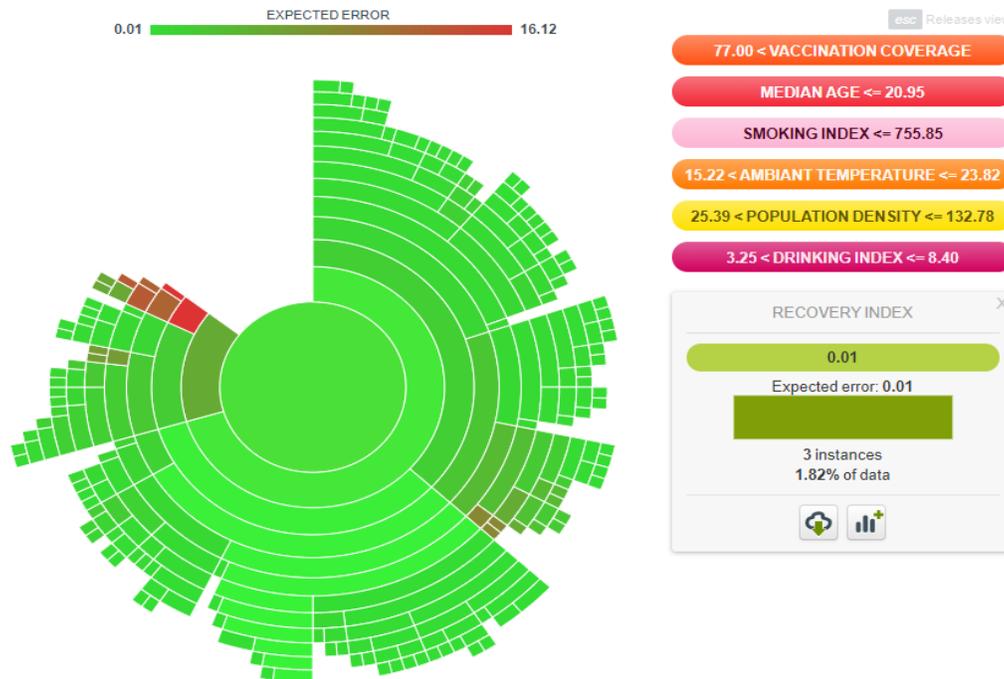


Figure 1 - Sunburst analytics chart of an instance for Covid-19 recovery

For the current study, a dataset was created with relevant fields where both supervised and unsupervised learning modes were carried out for the development of multiple models. The default model utilized Decision Trees (DT), and one of its instances as Sunburst analytics chart is as shown in Figure 1.

A Machine Learning model was trained with an algorithm that can be used to reason over and learn from previous outcomes to recognize patterns. The areas in green color represent the accuracy level of the 'Sunburst chart,' as shown in Figure 1. The accuracy is derived from associated factors predicting the Recovery Index using Decisions Trees. As can be seen from Figure 1, Machine Learning algorithms require the fitting of data to statistical models. Unlike traditional model fitting, the goal of a Machine Learning algorithm is to make accurate predictions using the input data. According to Talevi et al. (2020), the parameter values returned by the model are generally of secondary interest; however, to interprets parameter values in a meaningful way, another set of supervised learning (Ensemble modeling) and unsupervised learning (Associations modeling) was carried out during the presented research.

Ensemble Modelling

Ensemble learning is commonly applied to Machine Learning models to improve efficiency and reduce the risk in the decision-making process. In this approach, predictions from a diverse set of models are combined to yield the best fitting model (Battineni et al., 2020). In supervised learning, aggregated predictions from an ensemble model of diverse classifiers consistently outperform individual methods (Ahsen et al., 2019). However, the difficulty of the optimization problem during data fitting depends on the nature and complexity of the real-world situation, and also the determination of metrics (that are an optimal fit for precise decision making) remains a challenging task (Chui et al., 2017; Tan et al., 2020).

There were several factors being considered while building the Ensemble model. However, the Median Age showed the greatest importance (33.58 %), as per Figure 2.

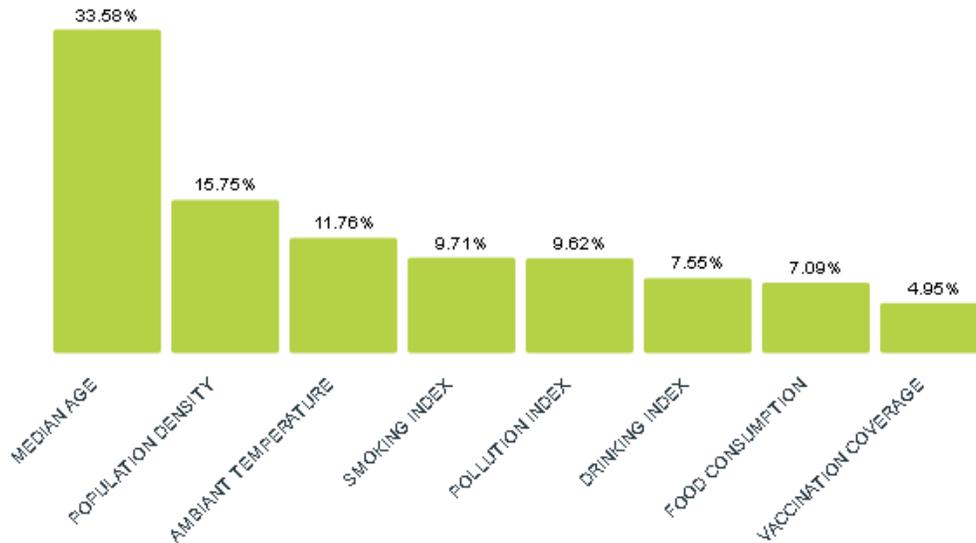


Figure 2 - Fields of Importance affecting Covid-19 recovery

Figure 2 signifies the importance of Ambient Temperature on the exponential growth rate of Covid-19 infection. (Livadiotis, 2020) Interestingly, Smoking Index has also shown considerable importance. In line with research conducted by Hayden & Parkin (2020), countries with a higher smoking index have observed an unsatisfactory Covid-19 recovery rate. The impact of the food consumption pattern can also be seen above in Figure 2. As food consumption includes uses of ‘edible spices’ as corroborated by Boukhatem & Setzer (2020), it would seem reasonable to assume that these products contain antiviral compounds. Figure 2 also shows the importance of population density that suggests the necessity of avoiding situations in highly populated areas to limit the spread of Covid-19. A similar observation was made by Rocklöv & Sjödin (2020); however, such observation may not always be universally accurate and warrants further investigation when using Machine Learning.

Although it may appear that some factors have low relevance to Covid-19, the complexity of the study involves finding that indirect associations that can be achieved using a separate Machine Learning model. Hence to re-confirm the output of the Machine Learning model, associations were identified through an unsupervised modeling technique called ‘Associations modeling.’

Associations Modelling

Associations modeling is an unsupervised Machine Learning technique that explores unique relationships among variables in a large dataset. Unlike conventional modeling techniques, associations algorithms measure degrees of similarity to identify hidden correlations to generate an exploratory graph, as displayed in Figure 3.

Associations are the product of the intersection between the antecedent variable and the consequent variable. The association is determined using the number of instances, probability, confidence level, and occurrence in the intersected dataset. The values of an association and leverage can be set by researchers to reduce the complexity of analysis.

The associations (K=5) and leverage (6.18%) based on curated data have been represented using Figure 3.

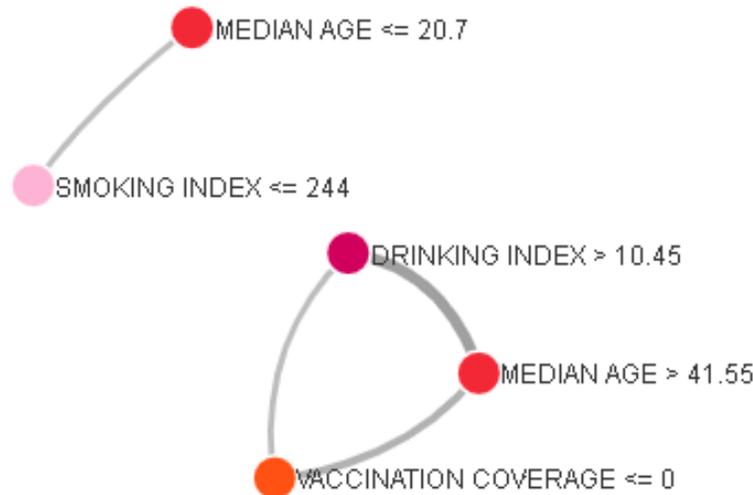


Figure 3 - Significant associations derived from the dataset

As per the model displayed in Figure 3, there exists a clear association between Median Age with Drinking Index as well as with Vaccination Coverage. The model shows that people of median age (>41.55), with no BCG vaccination, shall avoid consuming alcohol after a specific limit (Index 10.2) as it may negatively affect the Covid-19 recovery rate. Another study by Urashima et al. (2020) presented a similar hypothesis where BCG vaccination appears to be associated with reduced mortality of Covid-19.

The presented model further implies that populations with a specific age group ($20.7 <$) may get adversely affected in case the Smoking Index exceeds 244. Although this is an association based analysis, it does elaborate on the impact of smoking on the recovery rate of Coronavirus infected persons (Berlin et al., 2020; Reddy et al., 2020; Saadat et al., 2020). The act of smoking has several consequences as active smoking affects not only an individual but also affects people that are around. Further, in a social setting, smoking may also lead to minimal social distancing, which inevitably reflects on the Covid-19 Recovery Index.

Reliability

The current adoption of Machine Learning technology in healthcare emphasizes the need for characterizing model reliability. Considering the purpose as well as the application of Machine Learning in healthcare, it is imperative to conduct a thorough verification to support robust decision making (Rodr et al., 2006).

Reliability is a measure of consistency, and it focuses on the re-production of results when research is repeated under the same set of conditions. The reliability of the ML training process in the current study was achieved during model training automation process using the following steps -

1. Automated data extraction
2. Feature extraction
3. Model training and testing
4. Model selection and
5. Model evaluation

Reliability was further tested by checking the consistency of results across parts of the test itself. The approach involved the creation of ML models and the comparative analysis of the performances between the test set and a training set of data. The data reliability was identified upon training the dataset (using 80% data) and testing the model with the remaining training dataset (using 20% data).

To avoid the inefficient models in Machine Learning, quality checks were conducted on the following aspects:

- ✓ Data Quality
- ✓ Features Importance
- ✓ Model Matrices

Thus, as recommended by Cabitza et al. (2019), the presented study minimized the uncertainty of the Machine Learning model.

Validity

Validity is about the accuracy of a measure that confirms the extent to which the results measure; what they are supposed to measure. Model validation is performed to assess the model's predictive performance using a dataset that was not a part of the original build of the model.

As suggested by Huber-Carol et al. (2008), in the given study, validity was achieved by checking how well the results correspond to Goodness-Of-Fit. The Goodness-Of-Fit indicators summarize the disparity between observed values and the model's anticipated values. As far as a Machine Learning algorithm is concerned, a good fit is when both the training data error and the test data are minimal. The same is evident in the presented research based on the following characteristics -

A) MEAN ABSOLUTE ERROR - 0.11

Statistically, Mean Absolute Error refers to the results of measuring the difference between two continuous variables. It is the average of the absolute values of the differences between the target predicted by the model and the true target.

B) R-SQUARED VALUE - 0.95

R-squared is a statistical measurement of the vicinity of the data -points to the fitted regression line. It is also called the coefficient of determination for multiple regression. It is a measure of how a model performs than predicting the mean of the test set.

The Goodness-Of-Fit of a Machine Learning model explains how well it matches a set of observations and based on the ML Model output (from Table 1) following facts were validated –

More Significant Factors:

- AMBIENT TEMPERATURE, VACCINATION COVERAGE, MEDIAN AGE, SMOKING INDEX

Less Significant Factors:

- POPULATION DENSITY, POLLUTION INDEX, DIET CONSUMPTION, DRINKING INDEX

In accordance with existing literature, similar observations were reflected in the 'Discussion' sections.

Data-driven characteristics represent the level of 'Significance' as well as 'Confidence' of factors as mentioned below -

Table 1 - Significant factors and Confidence levels of test data

CONFIDENCE	POPULATION DENSITY	POLLUTION INDEX	AMBIENT TEMPERATURE	VACCINATION COVERAGE	MEDIAN AGE	DIET CONSUMPTION	SMOKING INDEX	DRINKING INDEX
3.31861	0	0	0.00209	0.71612	0.12009	0	0.16171	0
0.15803	0	0.00152	0.00175	0.71013	6.80E-04	0	0.28365	0.00228
6.21581	0	0	0	0	0.991	0	0	0.009
1.5993	0	0	0	0	0	0	0	0
0.47409	0	0	0.00664	0.74639	0.24648	0	4.90E-04	0
8.11216	0	0	0	0.95251	0.04749	0	0	0
0.77364	0.23137	0	0.01746	0.74232	0	0.00384	0.005	0
0.10535	0.23008	0	0.01737	0.74013	0.00362	0.00382	0.00498	0
0.79015	0.02908	0	0.02152	0	0.90103	0	0.04019	0.00818
0.10258	0	0	0	0.74986	0.24482	0	0.00531	0
0.05268	0.23058	0	0.0174	0.73818	0	0.00886	0.00498	0
0.33955	0	0	0.00175	0.71078	0	0	0.28389	0.00358
6.21581	0	0	0	0	0.991	0	0	0.009
0.26338	0	0	2.10E-04	0.75006	0.24442	0	0.00531	0
1.5993	0	0	0	0	0	0	0	0
0.05268	0	0	7.70E-04	0.74927	0.24463	0	0.00533	0
0.42141	0	0	0	0.74809	0.23872	0	0.01319	0
0.36873	0.03749	0	0.00977	0	0.89934	0.00512	0.04012	0.00817
0.36873	0.2302	0	0.01738	0.73858	0	0.00887	0.00498	0
0.50055	0	0	2.10E-04	0.7497	0.24477	0	0.00531	0
0.53061	0.23053	0	0.0174	0.73963	0.00362	0.00383	0.00499	0
0.03841	3.00E-05	0	7.40E-04	0.74928	0.24463	0	0.00531	0
0.15803	0.23001	0	0.01736	0.73798	0.00383	0	0.01082	0
1.5993	0	0	0	0	0	0	0	0
0.02934	0.23056	0	0.0175	0.73818	0.00294	0	0.01082	0
0.93905	0	0	0	1	0	0	0	0
0.15803	0	0.00152	0.00175	0.71013	6.80E-04	0	0.28365	0.00228
0.36873	0.03749	0	0.00977	0	0.89934	0.00512	0.04012	0.00817
0.47409	0	0	0.00175	0.71077	0.00133	0	0.28388	0.00228
0.31606	0	0.0012	0.00175	0.71037	6.80E-04	0	0.28372	0.00228
0.01174	5.00E-05	0	7.40E-04	0.74927	0.24463	0	0.00531	0
11.4336	0	0	0	0	1	0	0	0
0.02075	0.23056	0	0.01739	0.73817	0.00302	0	0.01087	0
0.08337	0.23058	0	0.01739	0.73826	0.00295	0	0.01082	0

Table 1 infers from the 20% of test data (output from ML Model, including 34 countries) that four factors have high significance as well as confidence than other factors included in the Covid-19 dataset.

Evaluation

Decision Tree

Reliability and validity have been established in given research to evaluate the quality of the proposed research. It indicates how effectively the current built of the Machine Learning model can make predictions. There are several ways to use data and reflect the trends and make predictions. Each one uses a technique that can be applied to infer from output structure (Supervised Learning) or represents specific data clusters (unsupervised

learning). The default model for given research is based on Decision Trees (DT) that are predictive representations of datasets. Decision trees are a hierarchical way of partitioning the space, where the goal is to develop a model that predicts the value of a target variable based on several input variables (Caballé et al., 2020).

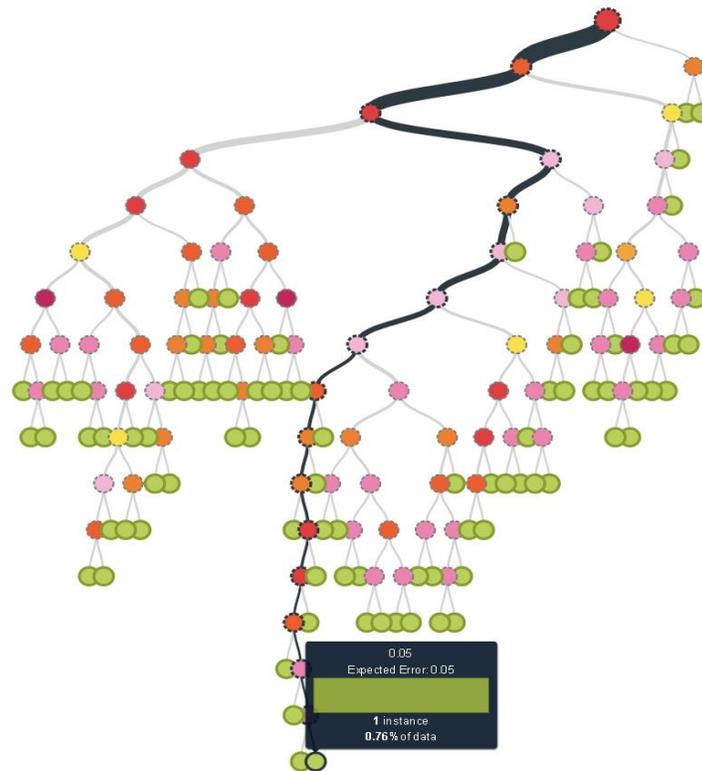


Figure 4 - Decision Trees Model developed for Covid-19 Recovery Index

As shown in Figure 4, the Decision Tree evaluated multiple fields from the given dataset and achieved Recovery Indices based on the field of importance. A Decision Tree learns by splitting the source dataset into subsets that are based on an attribute value test. When a DT is used for classification activities, it is more appropriately referred to as a classification tree. On the other hand, it is called a regression tree when used for the regression tasks. Such a model is expected to make sufficiently accurate predictions for the intended use, which in this case, is the Covid-19 Recovery Index.

As Figure 4 displayed a regression tree with a complexity, Principal Component Analysis (PCA) was conducted to be able to maintain accuracy during predictions.

Principal Component Analysis

Principal Component Analysis (PCA) is a statistical technique that is used for dimensionality reduction to improve model performance. Because of unsupervised learning, PCA seeks to cluster the data without prior training. During the presented study, the dataset was evaluated using Principal Component Analysis to overcome feature redundancy. It also reduced the complexity of the model and provided greater computation efficiency.

Figure 5 below shows the Scree plot of Principal Component Analysis with variances -

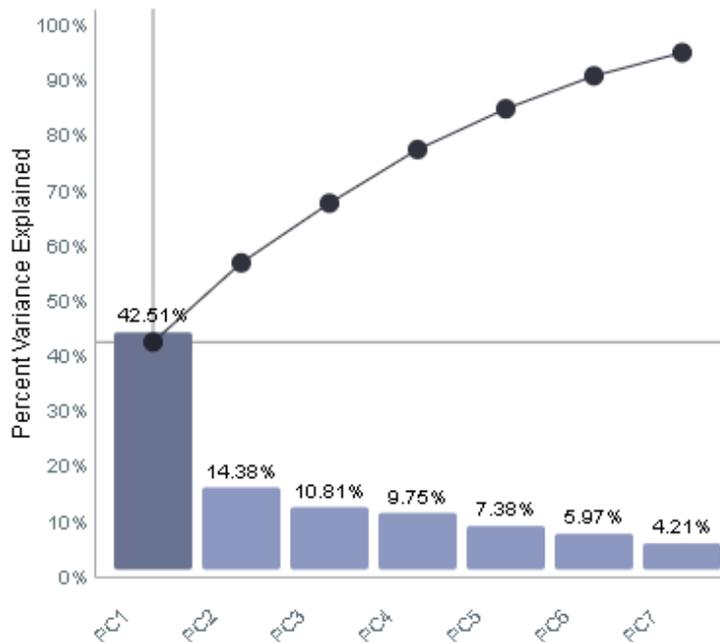


Figure 5 - SCREE plot showing PCA variances

A Scree Plot is a graphical utility used in the selection of the number of factors to be considered in a Principal Components' Analysis (or factor analysis.) Conceptually, the Scree Plot supports visualizing the magnitude of the variability associated with every component extracted through Principal Component Analysis (PCA).

In the given study, the Scree plot facilitated the examination of the pattern of decreasing variability attributable to each successive component that may be used for component selection and feature extraction.

As Figure 5 re-confirms, the selected Principal Components (Set 1) used in current research possess the highest variance. It translates into the fact that model selection by the researchers was robust. Also, the variables identified for the study were unique and statistically independent.

Discussion

Observations

Machine Learning, a subfield of Artificial Intelligence, leverages numerical techniques derived from computer science, mathematics, and statistics to automatically 'learn' by processing massive datasets. Machine Learning can provide several indispensable tools for intelligent data analysis. According to Ashrafian & Darzi (2018), Machine Learning has the potential to address complex real-world problems, and for current research, it played a significant role by supporting predictive modeling.

Predictive models in Machine Learning can effectively support decision-making activities (Battineni et al., 2020).

Additionally, Caballé et al. (2020) believe that technology is currently well suited for analyzing medical data and presents a wide range of possible applications. Unlike early A.I. systems, where it relied heavily on expert-derived rules to perform analytical processing; Machine Learning can fulfill tasks such as sensing, learning, reasoning, to make predictions and support serendipitous discovery (Morande & Pietronudo, 2020).

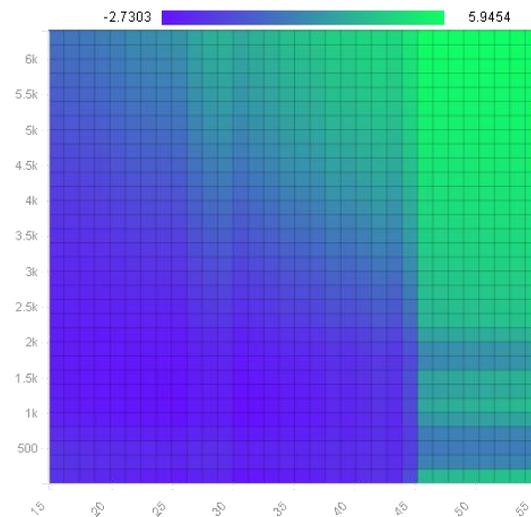


Figure 6 - PDP against Smoking Index (X-axis) & Median Age (Y-axis)

One of the outputs of Machine Learning models is in the form of Partial Dependence Plots (PDP). Partial Dependence Plots show the dependence between the target response and a set of target features, marginalizing over the values of complement features. As it can be observed from Figure 6, Partial Dependence Plots with two target features show the interactions among the two features (Molnar, 2019). Using the same, researchers can interpret the relationship between the expected target response as a function of the target features.

Using PDP, it is possible to predict and analyze the impact of different factors on the Covid-19 Recovery Index. In line with the inference drawn by Wyper et al. (2020), Figure 6 displays population vulnerability to Covid-19 concerning the Median Age. It also signifies one of the factors (Smoking Index) affecting the recovery of the patient. Partial Dependence Plots further presents a greater depth of insights, as described in the next section.

Insights

The use of real-life data identified important correlations, discovered unseen trends and suggested best practices regarding Covid-19. Following are the salient points with respect to various factors -

[A] Median Age

The age of an individual appears to be the most significant factor during the Covid-19 Pandemic (Karadag, 2020).

The Machine Learning model displayed that individuals above the age of forty-five fall under the 'High Risk' category. Other studies carried out by Solomou & Constantinidou (2020) also affirmed that age plays an important role when dealing with the Covid-19 pandemic.

[B] Smoking Index

Although the current study by Cai (2020) does not support smoking as a predisposing factor for Covid-19 infection, the data-driven approach tells a different story. It suggests that Smoking Index is one of the significant factors in the Covid-19 pandemic.

The Machine Learning model shows a significant correlation between Age and Smoking Index as well. It further suggests that smoking deteriorates Recovery Index and adversely affects an individual beyond the age of 25.

Covid-19 being a respiratory condition, supporting argument can be made, stating smoking increases the risk of viral infection (Berlin et al., 2020).

[C] Vaccination Coverage

Gursel & Gursel (2020) have hypothesized that a limited number of Covid-19 cases in Asia and Africa might reflect from the 'BCG immunization induced heterologous protective activity' of the vaccine.

According to the Machine Learning model, after the age of 45, if not vaccinated - an individual can experience a weaker Recovery Index. This observation further aligns with the claim made by Iwasaki & Grubaugh (2020), where the BCG vaccine may induce certain immune stimuli to gain resistance against pathogens. That said, the ML model showed a positive but limited impact of BCG vaccination on the recovery rate, which may not be very conclusive and may call for additional data required for future studies.

[D] Ambient Temperature

The Machine Learning model displayed that temperature range from 7-15 °C (degree Celsius) presents the lower Recovery Index, and the age group beyond the age of twenty-seven might get exposed to the adverse effect of Covid-19 in cold regions.

Research conducted by scientists on similar grounds affirms the impact of temperature on the number of Covid-19 patients (Bukhari et al., 2020; Jiang et al., 2020).

[F] Drinking Index

The Machine Learning model also inferred that after the age of forty, the Recovery Index suffers if the Drinking Index exceeds 3.0. That re-confirms, drinking alcohol does have some impact on the Covid-19 Recovery Index.

Szabo & Saha (2015) also believe that exposure to alcohol may harm host response and may impair immune function that can enhance susceptibility to viral infection. Although Drinking Index is one of the factors identified in the given study, the significance is relatively low, as shown in Table 1.

[H] Food Habits

The data modeling suggests that the inclusion of spices in the diet does positively affect the Covid-19 Recovery Index. This may also translate into a flat Covid-19 mortality rate observed in Asian and African regions based on respective diet patterns (Debnath et al., 2020).

As per Boukhatem & Setzer (2020), it would be reasonable to assume that herbs or spices that contain antiviral compounds can provide health benefits against viral infections.

[E] Population Index

Ahmed et al. (2020) suggest that population density plays an essential role in the spreading of Covid-19 & higher population density poses an extreme risk.

Surprisingly, the ML model shows that the Recovery Index improves as population density becomes higher.

This observation is inclined towards 'herd immunity,' which may not be a practical approach as far as Covid-19 is concerned. (Iwasaki & Grubaugh, 2020) It may also be an indirect effect observed by the administration of the BCG vaccine that is more common in countries with higher population density (Zwerling et al., 2011).

Such circumstances call for additional studies for the reasoning of factors in question.

[E] Pollution Level

One of the preliminary studies by Han et al. (2020) suggests that Pollution Levels are associated with Covid-19 infection. However, according to the Machine Learning model, the Pollution and Recovery Index are not highly correlated.

(In the future studies using A.I. driven fusion modeling may reflect some co-relation but is unlikely to be highly significant.)

It must be noted that the above interpretation drawn from Machine Learning modeling applies only to the macro level. This is because collected data is based on population-level, and hence insight derived for the dataset can only be transferred to masses.

Implications

Over the last few years, the immense increase in computational power and data storage has enabled A.I. (more specifically, Machine Learning) to provide almost unbelievable classification and prediction performances compared to well-trained humans (Clifford, 2020). According to Stiglic et al. (2020), emerging Machine Learning algorithms are beginning to transform the healthcare sector, where it is possible to classify interpretability approaches into two types. The first type focuses on personalized interpretation (local interpretability) while the second summarizes prediction models on a population level (global interpretability).

With a focus on the developing world, the research presents a data-driven but non-clinical approach that assumes the time barrier for the availability of Covid-19 treatment. The study attempts to provide insights that can be adopted during policy-making exercise on a population level for effective governance. Research provides a set of practical aspects related to Covid-19, including causal effect leading to the prediction of health-related outcomes. The same can be applied to the communities for a healthier society.

Limitations and Future possibilities

Every Machine Learning experiment begins with data curation required to build a model, and a model will be only as good as the data from which it is derived. Therefore, data quality is often highly dependent on data aggregators. As data curation and preparation are a cornerstone to develop a reliable model, the results presented in this study are dependent on the collected dataset. While a significant effort has been undertaken for accurate data modeling, the overfitting of data in the Machine Learning may bring in less clarity in the output. Such conditions may make it challenging to identify associations in the dataset that are not genuinely intrinsic to the predictions made. Furthermore, drawing on data from an increasingly diverse set of sources, it presents an incredibly complex set of attributes that must be accounted for throughout the Machine Learning modeling (Feldman et al., 2017).

It must also be noted that the Machine Learning outcomes are derived from data available until July 2020. Any mutation of the Covid-19 virus may differ from current observation and reduce the value presented by the existing set of data. As per the suggestions by Urashima et al. (2020), the observation made in the study needs to be further examined and proved by randomized clinical trials. The predictive inferences in the given research are broad in nature and do reflect on the macro level only, and to make these insights applicable to micro-level data needs further augmentation.

In the near future, researchers plan to engage in personalized interpretation using a longitudinal approach. Beyond that, future research would segregate and reflect on the data by classifying it as per the regions to provide detailed insights.

Conclusion

The Covid-19 pandemic has changed the people's lifestyle, caused extensive losses, and threatened the sustenance of millions of people. Overall, the economic activities have been suspended, and commercial activities have been hibernated to control the spread of the virus.

As advancements in healthcare fields are expanding the access to electronic data, the utilized approach emphasizes on accelerated Covid-19 recovery through technology management. Using substantial computation power, the healthcare field can apply Artificial Intelligence to address challenges in healthcare, especially during a current pandemic. Although many of these Machine Learning systems remain experimental, it ultimately presents the most substantial transformative role in healthcare governance. With the proposed research, Machine Learning modeling has evaluated multiple scenarios to focus on the Covid-19 Recovery Index. The research presents a strong case where Machine Learning models can be used to identify specific traits and rescue masses from the impending outcome of Covid-19.

As the given study feeds on near-time data and comprehensive academic underpinning, the generalization of developed Machine Learning models is possible. The insights derived from this research are applicable to regional and national levels and can be used by both developed and developing countries to develop strategies. Like other revolutionary technologies, Machine Learning should consider the limitations on algorithm development and understanding its appropriateness to apply.

While researchers are progressively attracted by predictive modeling techniques, Machine Learning is likely to play a vital role in the advancement of healthcare and enhancement of societal health. The data-driven insights from the presented research could offer advice to accelerate the Covid-19 recovery and make policy recommendations to help administrators develop well-informed health policies.

Funding: This research received no external funding.

Acknowledgments: Machine Learning platform/APIs for an Ambassador - BigML

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ahmed, J., Jaman, H., Saha, G., Ghosh, P., Hasnat, J., Saha, G., & Ghosh, P. (2020). Effect of Temperatures, Humidity and Population Density on the Spreading of Covid-19 at 70 Cities/Provinces. Preprints, June, 1–11.
- Ahsen, M. E., Vogel, R. M., & Stolovitzky, G. A. (2019). Unsupervised evaluation and weighted aggregation of ranked classification predictions. *Journal of Machine Learning Research*, 20, 2018–2020.
- Allam, Z., & Jones, D. S. (2020). On the Coronavirus (COVID-19) Outbreak and the Smart City Network: Universal Data Sharing Standards Coupled with Artificial Intelligence (AI) to Benefit Urban Health Monitoring and Management. In *Healthcare* (Vol. 8, Issue 1). <https://doi.org/10.3390/healthcare8010046>
- Arpitha, M. S., Mithun, K. A., Rakesh, S., Singh, A., & Yadav, A. (2018). Better Healthcare using Machine Learning. *International Journal of Advanced Research in Computer Science*, 9(3), 10–14. <https://doi.org/10.1109/ABLAZE.2015.7154917>
- Ashrafian, H., & Darzi, A. (2018). Transforming health policy through machine learning. *PLoS Medicine*, 15(11), 10–13. <https://doi.org/10.1371/journal.pmed.1002692>
- Ashrafian, H., Darzi, A., & Athanasiou, T. (2015). A novel modification of the Turing test for artificial intelligence and robotics in healthcare. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 11(1), 38–43. <https://doi.org/10.1002/rcs.1570>



- Barnett-Howell, Z., & Mobarak, A. M. (2020). The Benefits and Costs of Social Distancing in Rich and Poor Countries. 1–3. <http://arxiv.org/abs/2004.04867>
- Battineni, G., Sagaro, G. G., Chinatalapudi, N., & Amenta, F. (2020). Applications of machine learning predictive models in the chronic disease diagnosis. *Journal of Personalized Medicine*, 10(2). <https://doi.org/10.3390/jpm10020021>
- BCG World Atlas. (2020). A database of global BCG vaccination policies and practices. <http://bcgatlas.org/>
- Berlin, I., Thomas, D., Le Faou, A. L., & Cornuz, J. (2020). COVID-19 and smoking. *Nicotine & Tobacco Research : Official Journal of the Society for Research on Nicotine and Tobacco*, 1–3. <https://doi.org/10.1093/ntr/ntaa059>
- Bluhm, A., Christandl, M., Gesmundo, F., Ravn Klausen, F., Mancinska, L., Steffan, V., Stilck Franca, D., & Werner, A. (2020). SARS-CoV-2 transmission chains from genetic data: a Danish case study. *BioRxiv*, December 2019, 2020.05.29.123612. <https://doi.org/10.1101/2020.05.29.123612>
- Boukhatem, M. N., & Setzer, W. N. (2020). Aromatic herbs, medicinal plant-derived essential oils, and phytochemical extracts as potential therapies for coronaviruses: Future perspectives. *Plants*, 9(6), 1–23. <https://doi.org/10.3390/PLANTS9060800>
- Bukhari, Q., Massaro, J. M., D'Agostino, R. B., & Khan, S. (2020). Effects of Weather on Coronavirus Pandemic. *International Journal of Environmental Research and Public Health*, 17(15), 5399. <https://doi.org/10.3390/ijerph17155399>
- Caballé, N. C., Castillo-Sequera, J. L., Gómez-Pulido, J. A., Gómez-Pulido, J. M., & Polo-Luque, M. L. (2020). Machine learning applied to diagnosis of human diseases: A systematic review. *Applied Sciences (Switzerland)*, 10(15), 1–28. <https://doi.org/10.3390/app10155135>
- Cabitza, F., Ciucci, D., & Rasoini, R. (2019). A giant with feet of clay: On the validity of the data that feed machine learning in medicine. *Lecture Notes in Information Systems and Organisation*, 28, 121–136. https://doi.org/10.1007/978-3-319-90503-7_10
- Cai, H. (2020). Sex difference and smoking predisposition in patients with COVID-19. *The Lancet Respiratory Medicine*, 8(4), e20. [https://doi.org/10.1016/S2213-2600\(20\)30117-X](https://doi.org/10.1016/S2213-2600(20)30117-X)
- Chodkiewicz, J., Talarowska, M., Miniszewska, J., Nawrocka, N., & Bilinski, P. (2020). Alcohol consumption reported during the COVID-19 pandemic: The initial stage. *International Journal of Environmental Research and Public Health*, 17(13), 1–11. <https://doi.org/10.3390/ijerph17134677>
- Chui, K. T., Alhalabi, W., Pang, S. S. H., de Pablos, P. O., Liu, R. W., & Zhao, M. (2017). Disease diagnosis in smart healthcare: Innovation, technologies and applications. *Sustainability (Switzerland)*, 9(12), 1–24. <https://doi.org/10.3390/su9122309>
- Clifford, G. D. (2020). The Future AI in Healthcare: A Tsunami of False Alarms or a Product of Experts? 404, 1–49. <http://arxiv.org/abs/2007.10502>
- Contini, D., & Costabile, F. (2020). Does air pollution influence COVID-19 outbreaks? *Atmosphere*, 11(4), 377. <https://doi.org/10.3390/ATMOS11040377>
- Debnath, M., Banerjee, M., & Berk, M. (2020). Genetic gateways to COVID-19 infection: Implications for risk, severity, and outcomes. *FASEB Journal*, 34(7), 8787–8795. <https://doi.org/10.1096/fj.202001115R>
- Feldman, K., Faust, L., Wu, X., Huang, C., & Chawla, N. V. (2017). Beyond volume: The impact of complex healthcare data on the machine learning pipeline. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10344 LNAI, 150–169. https://doi.org/10.1007/978-3-319-69775-8_9
- Goshen, R., Choman, E., Ran, A., Muller, E., Kariv, R., Chodick, G., Ash, N., Narod, S., & Shalev, V. (2018). Computer-Assisted Flagging of Individuals at High Risk of Colorectal Cancer in a Large Health Maintenance Organization Using the ColonFlag Test. *JCO Clinical Cancer Informatics*, 2, 1–8. <https://doi.org/10.1200/cci.17.00130>



- Gursel, M., & Gursel, I. (2020). Is global BCG vaccination-induced trained immunity relevant to the progression of SARS-CoV-2 pandemic? *Allergy: European Journal of Allergy and Clinical Immunology*, 75(7), 1815–1819. <https://doi.org/10.1111/all.14345>
- Han, Y., Lam, J. C. K., Li, V. O. K., Guo, P., & Zhang, Q. (2020). The effects of outdoor air pollution concentrations and lockdowns on Covid-19 infections in Wuhan and other provincial capitals in China. *Preprints.Org*, March, 1–7. <https://doi.org/10.20944/preprints202003.0364.v1>
- Hayden, J. C., & Parkin, R. (2020). The challenges of COVID-19 for community pharmacists and opportunities for the future. *Irish Journal of Psychological Medicine*. <https://doi.org/10.1017/ipm.2020.52>
- Helgi Library. (2020). Spice Consumption Per Capita. <https://www.helgilibrary.com/indicators/spice-consumption-per-capita/>
- Hu, Y., Jacob, J., Parker, G. J. M., Hawkes, D. J., Hurst, J. R., & Stoyanov, D. (2020). The challenges of deploying artificial intelligence models in a rapidly evolving pandemic. *Nature Machine Intelligence*, 2(6), 298–300. <https://doi.org/10.1038/s42256-020-0185-2>
- Huber-Carol, C., Balakrishnan, N., M.S. Nikulin, & M. Mesbah. (2008). Statistics for Industry and Technology. In *Statistics for Industry and Technology*. https://doi.org/10.1007/978-0-8176-4619-6_22
- Iwasaki, A., & Grubaugh, N. D. (2020). Why does Japan have so few cases of COVID-19? *EMBO Molecular Medicine*, 12(5), 10–13. <https://doi.org/10.15252/emmm.202012481>
- Jean, S. S., & Hsueh, P. R. (2020). Old and re-purposed drugs for the treatment of COVID-19. *Expert Review of Anti-Infective Therapy*, 1–3. <https://doi.org/10.1080/14787210.2020.1771181>
- Jiang, Y., Wu, X. J., & Guan, Y. J. (2020). Effect of ambient air pollutants and meteorological variables on COVID-19 incidence. *Infection Control and Hospital Epidemiology*, 1–5. <https://doi.org/10.1017/ice.2020.222>
- Karadag, E. (2020). Increase in COVID-19 cases and case-fatality and case-recovery rates in Europe: A cross-temporal meta-analysis. In *Journal of Medical Virology* (Issue December 2019). <https://doi.org/10.1002/jmv.26035>
- Klinger, D., Blass, I., Rappoport, N., & Linal, M. (2020). Significantly improved COVID-19 outcomes in countries with higher bcg vaccination coverage: A multivariable analysis. *Vaccines*, 8(3), 1–14. <https://doi.org/10.3390/vaccines8030378>
- Livadiotis, G. (2020). Statistical analysis of the impact of environmental temperature on the exponential growth rate of cases infected by COVID-19. *PLoS ONE*, 15(5), 1–22. <https://doi.org/10.1371/journal.pone.0233875>
- Madurai Elavarasan, R., & Pugazhendhi, R. (2020). Restructured society and environment: A review on potential technological strategies to control the COVID-19 pandemic. *The Science of the Total Environment*, 725, 138858. <https://doi.org/10.1016/j.scitotenv.2020.138858>
- Mahase, E. (2020). Covid-19 : Russia approves vaccine without large scale testing or published results. August, 1–3. <https://doi.org/10.1136/bmj.m3205>
- Martin, A., Markhvida, M., Hallegatte, S., & Walsh, B. (2020a). Socio-Economic Impacts of COVID-19 on Household Consumption and Poverty. *Economics of Disasters and Climate Change*. <https://doi.org/10.1007/s41885-020-00070-3>
- Martin, A., Markhvida, M., Hallegatte, S., & Walsh, B. (2020b). Socio-Economic Impacts of COVID-19 on Household Consumption and Poverty. *Economics of Disasters and Climate Change*, 1–3. <https://doi.org/10.1007/s41885-020-00070-3>
- Mhalla, M. (2020). The Impact of Novel Coronavirus (COVID-19) on the Global Oil and Aviation Markets. *Journal of Asian Scientific Research*, 10(2), 96–104. <https://doi.org/10.18488/journal.2.2020.102.96.104>



- Miyasaka, M. (2020). Is BCG vaccination causally related to reduced COVID-19 mortality? *EMBO Molecular Medicine*, 12(6), 10–13. <https://doi.org/10.15252/emmm.202012661>
- Molnar, C. (2019). *Interpretable Machine Learning*. Book, 247. <https://christophm.github.io/interpretable-ml-book>
- Morande, S., & Pietronudo, M. C. (2020). Pervasive Health Systems: Convergence through Artificial Intelligence and Blockchain Technologies. *Journal of Commerce and Management Thought*, 11(2), 155. <https://doi.org/10.5958/0976-478x.2020.00010.5>
- Ooms, R., & Spruit, M. (2020). Self-service data science in healthcare with automated machine learning. *Applied Sciences (Switzerland)*, 10(9). <https://doi.org/10.3390/app10092992>
- Phillipson, J., Gorton, M., Turner, R., Shucksmith, M., Aitken-McDermott, K., Areal, F., Cowie, P., Hubbard, C., Maioli, S., McAreevey, R., Souza-Monteiro, D., Newbery, R., Panzone, L., Rowe, F., & Shortall, S. (2020). The COVID-19 pandemic and its implications for rural economies. *Sustainability (Switzerland)*, 12(10), 1–10. <https://doi.org/10.3390/SU12103973>
- Rashed, E. A., Kodera, S., Gomez-Tames, J., & Hirata, A. (2020). Influence of absolute humidity, temperature and population density on COVID-19 spread and decay durations: Multi-prefecture study in Japan. *International Journal of Environmental Research and Public Health*, 17(15), 1–14. <https://doi.org/10.3390/ijerph17155354>
- Reddy, R. K., Charles, W. N., Sklavounos, A., Dutt, A., Seed, P. T., & Khajuria, A. (2020). The effect of smoking on COVID-19 severity: a systematic review and meta-analysis. *Journal of Medical Virology*, 0–2. <https://doi.org/10.1002/jmv.26389>
- Rocklöv, J., & Sjödin, H. (2020). High population densities catalyse the spread of COVID-19. *Journal of Travel Medicine*, 27(3), 1–2. <https://doi.org/10.1093/jtm/taaa038>
- Rodr, E., Kypson, A. P., Moten, S. C., Nifong, L. W., & Jr, W. R. C. (2006). Robotic mitral surgery at East Carolina University : *International Journal*, April, 211–215. <https://doi.org/10.1002/rcs>
- Rozenfeld, Y., Beam, J., Maier, H., Haggerson, W., Boudreau, K., Carlson, J., & Medows, R. (2020). A model of disparities: risk factors associated with COVID-19 infection. *International Journal for Equity in Health*, 19(1), 126. <https://doi.org/10.1186/s12939-020-01242-z>
- Saadat, S., Rawtani, D., & Hussain, C. M. (2020). Environmental perspective of COVID-19. *Science of the Total Environment*, 728, 138870. <https://doi.org/10.1016/j.scitotenv.2020.138870>
- Saeb, S., Lonini, L., Jayaraman, A., Mohr, D., & Kording, K. (2016). Voodoo Machine Learning for Clinical Predictions. 059774. <https://doi.org/10.1101/059774>
- Saria, S., Butte, A., & Sheikh, A. (2018). Better medicine through machine learning: What’s real, and what’s artificial? *PLoS Medicine*, 15(12), 1–6. <https://doi.org/10.1371/journal.pmed.1002721>
- Scavone, C., Brusco, S., Bertini, M., Sportiello, L., Rafaniello, C., Zoccoli, A., Berrino, L., Racagni, G., Rossi, F., & Capuano, A. (2020). Current pharmacological treatments for COVID-19: what’s next? *British Journal of Pharmacology*. <https://doi.org/10.1111/bph.15072>
- Sethi, A., & Bach, H. (2020). Evaluation of current therapies for COVID-19 treatment. *Microorganisms*, 8(8), 1–17. <https://doi.org/10.3390/microorganisms8081097>
- Sidor, A., & Rzymiski, P. (2020). Dietary choices and habits during COVID-19 lockdown: Experience from Poland. *Nutrients*, 12(6), 1–14. <https://doi.org/10.3390/nu12061657>
- Silva Junior, F. J. G. Da, Sales, J. C. E. S., Monteiro, C. F. D. S., Costa, A. P. C., Campos, L. R. B., Miranda, P. I. G., Monteiro, T. A. D. S., Lima, R. A. G., & Lopes-Junior, L. C. (2020). Impact of COVID-19 pandemic on mental health of young people and adults: A systematic review protocol of observational studies. *BMJ Open*, 10(7). <https://doi.org/10.1136/bmjopen-2020-039426>
- Solomou, I., & Constantinidou, F. (2020). Prevalence and predictors of anxiety and depression symptoms during the COVID-19 pandemic and compliance with precautionary measures: Age and sex matter. *International Journal of Environmental Research and Public Health*, 17(14), 1–19.



- <https://doi.org/10.3390/ijerph17144924>
- Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., & Cilar, L. (2020). Interpretability of machine learning based prediction models in healthcare. In arXiv.org. Cornell University Library, arXiv.org. <https://search.proquest.com/docview/2359831537?accountid=12669>
- Szabo, G., & Saha, B. (2015). Alcohol's effect on host defense. *Alcohol Research: Current Reviews*, 37(2), 159–170.
- Talevi, A., Morales, J. F., Hather, G., Podichetty, J. T., Kim, S., Bloomingdale, P. C., Kim, S., Burton, J., Brown, J. D., Winterstein, A. G., Schmidt, S., White, J. K., & Conrado, D. J. (2020). Machine Learning in Drug Discovery and Development Part 1: A Primer. *CPT: Pharmacometrics and Systems Pharmacology*, 9(3), 129–142. <https://doi.org/10.1002/psp4.12491>
- Talukder, A., Author, C., Address, D., & Author, C. (2019). Title Page Effect of Age on Death Due to Coronavirus Disease 2019 (COVID-19): Application of Poisson Regression Model Running Head: Effect of Age on Death due to Coronavirus Disease 2019 (COVID-. 0–2. <https://doi.org/10.1111/ijcp.13649>
- Tan, Y., Jin, B., Yue, X., Chen, Y., & Sangiovanni-Vincentelli, A. (2020). Exploiting Uncertainties from Ensemble Learners to Improve Decision-Making in Healthcare AI. 1–3.
- Urashima, M., Otani, K., Hasegawa, Y., & Akutsu, T. (2020). BCG Vaccination and Mortality of COVID-19 across 173 Countries: An Ecological Study. *International Journal of Environmental Research and Public Health*, 17(15), 1–21. <https://doi.org/10.3390/ijerph17155589>
- Wyper, G. M. A., Assunção, R., Cuschieri, S., Devleeschauwer, B., Fletcher, E., Haagsma, J. A., Hilderink, H. B. M., Idavain, J., Lesnik, T., Von Der Lippe, E., Majdan, M., Milicevic, M. S., Pallari, E., Peñalvo, J. L., Pires, S. M., Plaß, D., Santos, J. V., Stockton, D. L., Thomsen, S. T., & Grant, I. (2020). Population vulnerability to COVID-19 in Europe: A burden of disease analysis. *Archives of Public Health*, 78(1), 1–9. <https://doi.org/10.1186/s13690-020-00433-y>
- Zahedipour, F., Hosseini, S. A., Sathyapalan, T., Majeed, M., Jamialahmadi, T., Al-Rasadi, K., Banach, M., & Sahebkar, A. (2020). Potential effects of curcumin in the treatment of COVID-19 infection. *Phytotherapy Research*. <https://doi.org/10.1002/ptr.6738>
- Zhang, J., Wu, W., Zhao, X., & Zhang, W. (2020). Recommended psychological crisis intervention response to the 2019 novel coronavirus pneumonia outbreak in China: a model of West China Hospital. *Precision Clinical Medicine*, 3(1), 3–8. <https://doi.org/10.1093/pcmedi/pbaa006>
- Zhu, Z., Xu, S., Wang, H., Liu, Z., Wu, J., Li, G., Miao, J., Zhang, C., Yang, Y., Sun, W., Zhu, S., Fan, Y., Hu, J., Liu, J., & Wang, W. (2020). COVID-19 in Wuhan: Immediate Psychological Impact on 5062 Health Workers. 1095. <https://doi.org/10.1101/2020.02.20.20025338>
- Zwerling, A., Behr, M. A., Verma, A., Brewer, T. F., Menzies, D., & Pai, M. (2011). The BCG world atlas: A database of global BCG vaccination policies and practices. *PLoS Medicine*, 8(3). <https://doi.org/10.1371/journal.pmed.1001012>