**PROJECT DOCUMENTATION PART I**

**TABLE OF CONTENTS Page**

[CHAPTER 1: INTRODUCTION 1](#_Toc86335576)

[1.0 Overview 1](#_Toc86335577)

[1.1 Project Description 1](#_Toc86335578)

[1.2 Problem Statement 1](#_Toc86335579)

[1.3 Research Objectives 2](#_Toc86335580)

[1.4 Research Questions 2](#_Toc86335581)

[1.5 Research Objectives Align with Research Questions 3](#_Toc86335582)

[1.6 Project Scope 4](#_Toc86335583)

[1.7 Project Limitation 4](#_Toc86335584)

[1.8 Methodology 4](#_Toc86335585)

[1.8.1 Phase 1: Fact Finding 4](#_Toc86335586)

[1.8.1.1 Qualitative Research 4](#_Toc86335587)

[1.8.1.1.1 Interview 5](#_Toc86335588)

[1.8.1.2 Quantitative Research 5](#_Toc86335589)

[1.8.1.2.1 Questionnaires 5](#_Toc86335590)

[1.8.2 Phase 2: Crisp-DM 5](#_Toc86335591)

[1.8.3 Phase 3: Web-based dashboard 6](#_Toc86335592)

[1.9 Target Audience 7](#_Toc86335593)

[1.10 Summary 7](#_Toc86335594)

[CHAPTER 2: LITERATURE REVIEW 8](#_Toc86335595)

[2.0 Overview 8](#_Toc86335596)

[2.1 Covid-19 8](#_Toc86335597)

[2.1.1 Introduction 8](#_Toc86335598)

[2.1.2 Effects of Covid-19 8](#_Toc86335599)

[2.1.3 Vaccines 9](#_Toc86335600)

[2.1.4 Impacts of Vaccines 9](#_Toc86335601)

[2.1.5 Problems facing Vaccination Rates 10](#_Toc86335602)

[2.2.1 Current Practices/ Existing System 11](#_Toc86335603)

[2.3 Techniques 12](#_Toc86335604)

[2.3.1 Classification 12](#_Toc86335605)

[2.3.1.1 Artificial Neural Networks (ANN) 12](#_Toc86335606)

[2.3.1.1.1 Applications of ANN 13](#_Toc86335607)

[2.3.1.2 Support Vector Machines (SVM) 13](#_Toc86335608)

[2.3.1.2.1 SVM Applications 13](#_Toc86335609)

[2.3.1.3 Decision Tree 14](#_Toc86335610)

[2.3.1.3.1 Advantages of Decision Trees 14](#_Toc86335611)

[2.3.1.3.2 Application of Decision Trees 14](#_Toc86335612)

[2.4 Comparative Study 1](#_Toc86335613)

[2.5 Tools 1](#_Toc86335614)

[2.5.1 RStudio 1](#_Toc86335615)

[2.5.2 Microsoft Excel 1](#_Toc86335616)

[2.6 Summary 1](#_Toc86335617)

[References 2](#_Toc86335618)

[Appendix 6](#_Toc86335619)

[Figure 1. 1 Research Methodology using CRISP-DM (Smart Vision Europe, 2021) 9](#_Toc86259684)

[Figure 2. 1 Fully and partially vaccinated people in Southeast Asia (Ritchie et al., 2021) 15](#_Toc86259685)

# CHAPTER 1: INTRODUCTION

## 1.0 Overview

Predictive analytics is a large field of using machine learning techniques, statistical algorithms and data in order to identify the possibility of future outcomes based on data collected regarding past events. It useful in a wide array of topics and has practical uses around the world for different studies. The use of predictive analytics in this study is to assist in the vaccination process by providing insights into the areas that will have higher vaccinations. This is important as it allows the hospitals and policy workers to make decisions such as stockpiling on extra vaccines that will prove useful in the future depending on the predictions.

## Project Description

According to (World Health Organization,2021) the corona virus disease is a contagious illness which is caused by the SARS-CoV-2 virus. The illness causes the majority of infected to experience mild to moderate symptoms related to the respiratory system and can recover without the need for special treatment. Some of the infected however may suffer from more severe symptoms and may require special medical treatment. Nobody is immune to the illness and can get sick from it or may even die from it. Through the use of artificial neural networks model to predict the rate of vaccinations policy makers and hospitals can be assisted in making smart and timely decisions. The predictions can help to prepare for vaccine demand, such as financially or through stockpiling vaccines, assist in running targeted awareness campaigns, help to improve the accuracy of policies and help in improving vaccine uptake behaviors through the creation of appropriate policies thus helping to better guide compliance towards taking the vaccine (Zhou and Li, 2021; Shmueli, 2021).

## 1.2 Problem Statement

Covid-19 is a very serious illness with repercussions that has impacted every part of the lives of people around the world (Chu, 2021). The Corona virus epidemic has caused more than 17 million people around the world to have been infected, causing more than 667,000 cases of death from 20 January to 30 July 2020(Niazkar and Niazkar, 2020). The loss of life around the world has been dramatic and provides a challenge to food systems, public health and jobs around the world. The pandemic has caused disruptions in both the economic and social sectors which has devastating results with poverty being a real threat to tens of millions of people. Along with that the number of malnourished people which have been estimated to be near 690 million can increase by 132 million by the end of 2020 (World Health Organization,2020).

The pandemic has impacted the world in many ways. The psyche and health of individuals has been affected all around the globe with experiences such as trauma, shock, fear, grief and existential anxiety. This can be seen with the increase in the amount of visits to mental clinics and hospitals along with reports of domestic abuse cases and other similar issues. Health services due to the pandemic came under pressure, ethical and practical dimmas arose such as which individuals need to be prioritised for treatment and which individuals to not provide treatment to (Pawar, 2020).

On a societal level the pandemic has caught the world unprepared due to societies reliance on importing products such as food, energy and medical equipment along with preparations for pandemic outbreaks being limited. Due to this many businesses around the world have been forced to close down which in turn leads to a disruption of commerce in the majority of the world’s industrial sectors (Donthu and Gustafsson, 2020).

According to an article by (Hariharan et al., 2020) the provision of accurate utilization forecasts are important in maintaining an optimal level of stock of vaccines in any health facilities. This means that machine learning forecasts are important as they allow medical facilities and policy makers to look at the predictions and use that insight to make decisions on how to increase the number of vaccinations in order to improve the pandemic situation.

## 1.3 Research Objectives

RO1: To study and understand the current Covid-19 situation using the pandemic dataset and patient information, the research articles on the Pandemic Covid-19, the challenges faced by the health department in terms of logistic or patient management and the solutions implemented by Southeast Asia to increase the vaccination rates;

RO2: To perform predictive analytics using machine learning algorithm to measure the accuracy on the prediction of the percentage of vaccination rate;

RO3: To implement the predictive model and to use the power of visualization by developing a user-friendly dashboard to visualize the rate of vaccinations.

## 1.4 Research Questions

RQ1: What are the factors that lead to the decrease in the percentage of vaccinations?

RQ2: What are the factors that lead on the logistics for the vaccinations?

RQ3: What are the factors that influence the health department on managing vaccinations?

RQ4: What are the factors that were implemented to increase the percentage of vaccinations?

RQ5: How does the health department in identifying the vaccination rate of Covid-19?

RQ6: Which solution to apply on the prediction of the vaccination rate of Covid-19?

RQ7: Which algorithm have been applied in vaccination rate of Covid-19, what are the lacking and the performance measure from the past studies, which algorithms or techniques perform better than others and why?

## 1.5 Research Objectives Align with Research Questions

|  |  |  |  |
| --- | --- | --- | --- |
| **Research Objectives** | | **Research Questions** | |
| **RO1** | To study and understand the current Covid-19 situation using the pandemic dataset and patient information, the research articles on the Pandemic Covid-19, the challenges faced by the health department in terms of logistic or patient management and the solutions implemented by Southeast Asia to increase the vaccination rates; | **RQ1** | What are the factors that lead to the decrease in the percentage of vaccinations? |
| **RQ2** | What are the factors that lead on the logistics for the vaccinations? |
| **RQ3** | What are the factors that influence the health department on managing vaccinations? |
| **RQ4** | What are the factors that were implemented to increase the percentage of vaccinations? |
| **RO2** | To perform predictive analytics using machine learning algorithm to measure the accuracy on the prediction of the percentage of vaccination rate; | **RQ5** | How does the health department in identifying the vaccination rate of Covid-19? |
| **RQ6** | Which solution to apply on the prediction of the vaccination rate of Covid-19? |
| **RO3** | To implement the predictive model and to use the power of visualization by developing a user-friendly dashboard to visualize the rate of vaccinations. | **RQ7** | Which algorithm have been applied in vaccination rate of Covid-19, what are the lacking and the performance measure from the past studies, which algorithms or techniques perform better than others and why? |

**Table 1.1 Research Objectives Align with Research Questions**

## 1.6 Project Scope

This study will look into the percentage rate of Covid-19 vaccines this will allow users to glean into the possibility of whether there will be an increase in the number of vaccines or a reduction. This information will assist the users to make decisions on the betterment of the future in regard to the illness through a number of options such as proper utilization of vaccine resources, raising awareness regarding vaccines, creating accurate policies to improve uptake of vaccines, etc (Zhou and Li, 2021; Shmueli, 2021). The use of ANN is to attempt to get a degree of certainty for the prediction on the rate of vaccinations.

## 1.7 Project Limitation

Covid-19 pandemic has intensified the challenge of performing research as a rush for data there have been multiple report and blogs that have been published but the amount of data that is scientifically robust is limited (Weiner, Balasubramaniam, Shah and Javier, 2020). The data set that has been acquired has records from 2020 to 2021. It is also difficult to acquire unequivocally accurate data without enough resources and time. The journal by (Ioannidis, Cripps and Tanner, 2020) states that epidemic prediction has an uncertain track record in that it is difficult to have a completely accurate prediction due to a number of reasons such as lack of determinacy, lack of expertise in certain required disciplines, high sensitivity of estimates, selective recording, poor past evidence on effects of available interventions, lack of transparency, errors in the data, etc.

## 1.8 Methodology

### 1.8.1 Phase 1: Fact Finding

#### 1.8.1.1 Qualitative Research

Qualitative research is used to understand thoughts, experiences or concepts which have been expressed in words. This type of research is suitable when gathering data that is in-depth regarding topics that are not understood well. Interviews are structured conversation where one participant asks the other questions to gain an in-depth understanding regarding a particular topic. Interview is the most effective method for qualitative research in that it helps to explain, understand and explore the interviewees’ opinions and experiences better according to (Research Guides: Research Methods Guide: Interview Research, 2021), which also states that due to the open-ended nature of interview questions the information collected is more in-depth.

##### 1.8.1.1.1 Interview

Interview is the process of asking questions that are open-ended in order to converse with respondents and collect data regarding a topic. The people who will be interviewed will be health officials or people who work in healthcare as they are the people most knowledgeable regarding the cases of infected patients.

#### 1.8.1.2 Quantitative Research

Quantitative research is used to confirm or test assumptions and theories through the use of numbers and graphs. This type of research allows for the establishment of generalizable facts regarding a topic. Questionnaires are a set of written or printed questions that have choices for answers and are used for the purpose of collection large amount of information from a large number of people. Questionnaires according to (McLeod, 2018) are a cheap, quick and efficient method of acquiring information from a large number of people and can be collected quickly as the researcher is not required to be present.

##### 1.8.1.2.1 Questionnaires

Questionnaires are a set of standardized questions which follow a fixed arrangement which can be used to collect data from individuals regarding specific topics. The target for the questionnaire will be ordinary people this is to understand the effects of the illness better on normal everyday people. The questionnaires will be answered on google forms as it is an easy way to create questionnaire on it. It is also easy to send out google form links to many people hence easing the process of sending out questionnaires to a large number of people.

### 1.8.2 Phase 2: Crisp-DM

Cross-Industry Process for Data Mining or CRISP-DM for short. According to (IBM Docs, 2021) CRISP-DM is an industry-proven methodology that allows for a guided method when conducting data mining projects. It is a flexible and useful method when conducting analytics. CRISP-DM has six phases as follows:

1. Business Understanding

On the first phase of CRISP-DM is necessary to understand the objectives of the project. It is the phase of the methodology where hidden risks are looked for that might have an effect on the project in the future. This then involves planning to ensure that those risks do not cause the project to fail. It also involves setting the goals for project.

1. Data Understanding

In this phase the data set is looked into in order to gain an understanding regarding what type of data will be suitable for the project whether it fits the quality, the correct values, etc. This phase is necessary to make the decision to choose the right type of data set. This in order to know what is to be expected and achieved from the data. Hence, this phase is used to set the sources of data and the data itself.

1. Data Preparation

This phase involves turning the data collected into useful information through the process of cleaning and the use of algorithms. Data is thus transformed in a way that is necessary to the project.

1. Modeling

This is the fundamental part of a machine learning project which is responsible for the outcome that should help in achieving, with satisfaction, the goals set.

1. Evaluation

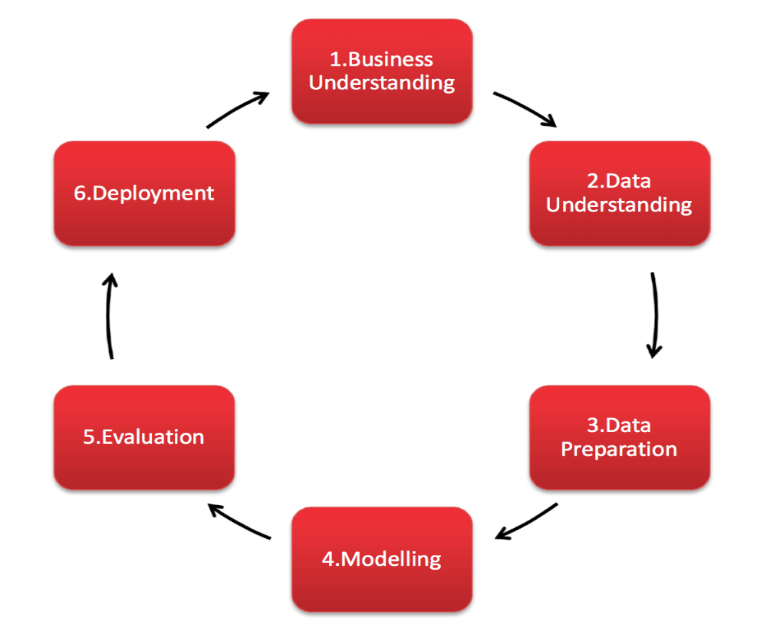
In this phase the results acquired from the modeling phase are checked and verified to be valid and correct. The methodology in the event that the results are incorrect permits that there be a review from the first phase. This is done to understand why the results are not correct.

1. Deployment

This is the final phase of the methodology. In this phase the results that have been verified to be correct are presented in such a way to be seen in an understandable and useful manner hence achieving the projects goals with satisfaction.

### 1.8.3 Phase 3: Web-based dashboard

A web-based dashboard is a page on a website or an online interface that is used to present real-time data using charts and reports. Web-based dashboards are great for visualizing data better and making data easier to understand.



**Phase 3:** Web-based dashboard

**Phase 1: Fact Finding**

* Qualitative Research
* Quantitative Research

**Phase 2:** Cross-Industry Process for Data Mining (CRISP-DM)

Figure 1.1 Research Methodology using CRISP-DM (Smart Vision Europe, 2021)

## 1.9 Target Audience

The target audience for this project would be the hospitals and policy makers. This research can benefit them by allowing them to have a degree of certainty on when and where there will be a demand for vaccines and where the number of vaccinations is lower than the required amount and thus allow them to have enough information to make smart and timely decisions to increase the number of vaccinations and improve the health of the people.

## 1.10 Summary

This study aims to assist in making the pandemic situation better through the portrayal of the predictions from the predictive model to a user-friendly web-based dashboard that will assist the target users to make timely and smart decisions which would be useful in improving the pandemic situation.

# CHAPTER 2: LITERATURE REVIEW

## 2.0 Overview

In this chapter the author will discuss more about the Covid situation, the vaccines and the different types of machine learning techniques that have been used in similar studies and their outcome as well as the machine learning techniques that the author will use in this study.

## 2.1 Covid-19

### 2.1.1 Introduction

According to an article by (Li et al., 2020; Habas et al., 2020), on December 29, 2019, the first 4 cases were reported in the city of Wuhan, a city of 11 million people. The cases that were the earliest were identified through the use of the surveillance mechanism known as the “pneumonia of unknown etiology.” The pneumonia of unknown etiology is an illness without an identified pathogen which caused it that fulfills a set of criteria such as fever above 38 degree Celsius, evidence of pneumonia through radiography, low normal white-cell or lymphocyte count and no systematic improvement after 3 to 5 days of antimicrobial treatment.

Covid-19 has had two preceding instances of emergence in the past 18 years these have been as SARS (severe acute respiratory syndrome) in 2002 and 2003, and MERS (Middle East respiratory syndrome) from 2012 to present year (Fauci, Lane and Redfield, 2020).

### 2.1.2 Effects of Covid-19

Pandemics despite being rarely occurring catastrophes have occurred in the past and will occur in the future, the current pandemic outbreak has had sever effects on the economy around the world and it is unlikely that any country will remain unaffected (Donthu and Gustafsson, 2020).

The Covid-19 infection known as SARS-CoV-2 has fast spread worldwide which has caused various levels of illness by March 11, 2020, the SARS-CoV-2 was announced as a pandemic which has persisted to present day (Lai et al., 2020). The fast spread of Covid-19 has caused a dramatic change in the world with more than 73 million individuals having contracted the illness and more than 50,000 deaths having been reported on August 15, 2020, in sub-Saharan African countries alone (Assefa et al., 2021).

Global pandemics such as Covid-19 far reaching adverse effects towards society through the high number of mortalities, the different levels of health risks and the expected loss of 2.6 trillion dollars due to long term complications (Allocati et al. 2016; Cutler and Summers, 2020; Fan, Zhao, Shi and Zhou, 2019).

### 2.1.3 Vaccines

Vaccines are an amazing example of the ability of mankind in understanding the world of biology that can be found everywhere from within us to around us and using that knowledge in the preservation of human life and health. Allowing the body to engage its adaptive immune systems to produce the specific antibodies and using immunological memory that are required to combat a future infection is the use of vaccines.

Vaccines perform this task because they comprise of a weakened or inactive version of the particular pathogen. The pathogens are rendered incapable of inducing an infection but remaining intact thus allowing the body to recognize the pathogen as foreign (Federman, 2021).

### 2.1.4 Impacts of Vaccines

There has been an explosive surge in vaccine development since the emergence of Covid-19 such that by 24 September 2020 a staggering number of more than 200 vaccines had started going through preclinical development and from those 43 had started clinical trials (Tregoning et al., 2020).

Amid the large amount of negative news regarding the pandemic, reports on the vaccines that could combat Covid-19 are rays of hope. Despite this hope the vaccines are not a cure all solution to the illness (Su et al., 2021). According to an article by (Paltiel, Schwartz, Zheng and Walensky, 2021) which states that the benefits brought about by the vaccine may be reduced significantly in the event there is an increase in epidemic severity, substantial amount of vaccine hesitancy or manufacturing delays. The article goes on to state that the vaccines can provide direct protection thus reducing the possibility of those uninfected from succumbing to the virus and indirect protection which means that it can reduce further spread of the illness to those who have been infected. The FDA established a minimum threshold for the vaccines efficacy which would in the view of the FDA would allow for a widely deployed vaccine to hold its effectiveness.

The vaccine might not be a silver bullet but as a preventive measure it can provide assistance in preventing the emergence of new variant strands of the virus. The virus has opportunities to change the more it spreads thus vaccinations to a large number of the populace can reduce the spread of the virus thus preventing new variants from emerging (CDC, 2021). It is also recommended by the CDC that everyone from age 12 and above should get vaccinated at the soonest possible time.

### 2.1.5 Problems facing Vaccination Rates

There are doubts about getting the Covid-19 vaccine and have little confidence on the safety of the vaccine and whether the distribution of the vaccine is impartial. With 26% of Americans planning on not taking the vaccine and 27% percent being unsure on whether they want to take the vaccine which leaves only 47% who plan to take the vaccine. The percentage of people who are confident in the quick and safe distribution of the vaccines are only 22%, the people confident in the fair distribution of the vaccine are 20% percent and the people confident in the effectiveness and safety of the vaccine are 28%. This data comes from an AP-NORC poll of 1117 adults on December 3-7th of 2020 (Benz et al., 2021).

Many countries in Southeast Asia are now under a new surge of infections and deaths that are being caused by the new variants of the coronavirus. The vaccine campaigns that are sluggish marked by the lack of supply and the hesitancy influenced by the distrust in the authorities and the supplier of the vaccine to this region, which is China, has caused the situation to become exacerbated (Nachemson, 2021). Despite being able to initially keep the worst of the pandemic at bay through the use of lockdowns and social distancing restrictions Southeast Asia still requires strong programs that can facilitate vaccinations to assist in moving past the acute phase of the pandemic. Till this date Southeast Asia has suffered setback due to a number of challenges one which being the shortage of sufficient vaccines. The lower income countries of Southeast Asia also require significant support in increasing the vaccination rates as well overcoming vaccination hesitancy and other challenges (Maude, 2021).

The extent of the consequences of Covid-19 have yet to be established yet the evident impact, it has and will have in the present and the future, on the global trade. The likely disruption of supply chains of global pharmaceuticals and its impact on the global access to medication in countries low- and middle-income range will have to face dramatic ramifications (Guerin, Singh-Phulgenda and Strub-Wourgaft, 2020). Due to the sudden demand in some places and decrease in demand in other places, the pandemic has caused a major disruption in the existing medical supply chain making vaccinating the global population a significant challenge (Khan, Haleem, Deshmukh and Javaid, 2021).

The rate of vaccinations around the world is slow with only 48.7% of the World’s population having gotten the first dose of the Covid-19 vaccine, with 6.88 billion doses having been administered all around the world and 24.16 million being administered daily. In low-income countries the percentage of the populace that has at least acquired one dose is a measly 3.1%. Among the countries in Southeast Asia the highest percentage of fully vaccinated country is Singapore with 80% fully vaccinated and 1% partially on the other hand the lowest is Myanmar with the fully vaccinated being 8.3% and partially being 12%. Majority of the countries in Southeast Asia are below the 50% mark for being fully vaccinated with only 5 countries having a percentage of higher than 50% of fully vaccinated people by 25th October 2021 as shown in the figure below (Ritchie et al., 2021).

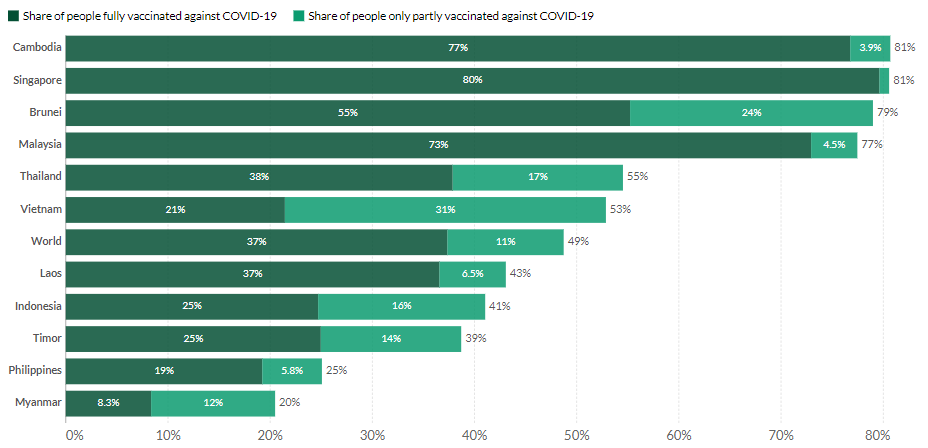


Figure 2. 1 Fully and partially vaccinated people in Southeast Asia (Ritchie et al., 2021)

**2.2 Current Problems**

There are countless number of problems that researchers face when conducting studies such as the amount of data being collected being too much to the point of being overwhelming, despite having a lot of data it is difficult to find and access what is needed most from within the data that would be useful or meaningful, data from multiple sources is difficult to analyze as it become difficult to combine the data and can limit the number of insights that would be able to be viewed, there are also problems of getting poor quality data or being denied access to the data needed, besides the problems that come with data there are also issues on maintaining the budget and finally there might be problems that arise due to the lack of the required skills needed to perform the tasks.

### 2.2.1 Current Practices/ Existing System

There are a number of practices that are conducted as the complexity and abundance of data increases causing a need to harness in order to make quick and smart decisions. In order to keep up with this need there a few current practices that have surfaced. Current practices such as the use of automation which can be useful in making the process of data discovery, preparation and blending of contrasting data to be done automatically. Democratization which is the thought that analytics and data science can be used by anyone despite not being an expert in the field this is done through the use of easy-to-use analytics automation platforms. Through use of such unified platforms, it is possible for any user to start gaining insights through simply dragging and dropping automation blocks thus reducing the amount of time needed to perform data discovery and preparation in order to gain insights. Having analytics as a core function of businesses this is when businesses embrace analytics as a critical and necessary function of a business.

## 2.3 Techniques

### 2.3.1 Classification

The process of recognizing, understanding and grouping ideas and objects is known as classification. In machine learning classification algorithms are used to input training data in order to predict the possibility that the subsequent data will fall into one of the categories that are being predetermined. Classification in summary is a type of pattern recognition using machine learning algorithms applied to training data in order to find similar patterns in future sets of data.

#### 2.3.1.1 Artificial Neural Networks (ANN)

ANN, Artificial Neural Networks are a subset of machine learning and are at the core of deep-learning algorithms. The structure of ANN as well as the name are inspired by the brains of human beings, as they copy the way biological neuron pass signals from one to another. ANNs are made up of nodes in layers such as an input layer, a single layer or multiple hidden layers and an output layer. Each node is connected to another and has an associated weight and threshold value (Education, 2021).

According to (Introduction to Neural Networks, Advantages and Applications, 2021) Artificial Neural Networks are suitable for certain types of problems and situations some of which are as follows:

Non-linear and complex relationships can be learnt and modelled by Artificial Neural Networks hence making ANNs useful in handling non-linear data. This is important because a lot of data is non-linear.

ANNs are able to generalize which means that after learning from the original inputs and their associations, it may infer from previously unknown relationships on previously unseen data thus allowing the model to be able to generalize and make prediction on the previously unseen data.

ANN, unlike many other prediction approaches doesn’t place any constraints on the input variables. ANN can also describe the data with high volatility and non-constant variance due to their capacity to gain understanding over hidden relationships in the data without establishing any preset relationships in the data. This can be extremely beneficial in situations such as financial time forecasting such as stock prices when the volatility of data is very high.

##### 2.3.1.1.1 Applications of ANN

ANNs are useful due to many of its applications such as forecasting which is important in many situations mostly ones conducted within businesses in order to assist businesses in making strategic decisions. Forecasting problems can be very complex such as predicting the prices of stocks which can have many unseen factors. Unlike traditional forecasting models which are limited in being unable to take into account non-linear and complex relationships ANNs when applied correctly can provide a robust alternative.

Image and character processing: ANNs are able to take a large number of inputs and process them which allows them to infer hidden, complex and non-linear relationships hence are able to be very useful when used to perform character recognition such as handwriting.

#### 2.3.1.2 Support Vector Machines (SVM)

This is a set of supervised learning methods that are utilized to perform classification, outlier detection as well as regression. SVM is a simple algorithm that is highly preferred by many as it provides the users with a significant amount of accuracy with little computation power. Despite its usage in both regression and classification tasks it is mostly used by many in classification objectives (Gandhi, 2021). There are multiple advantages to using Support Vector Machines such as

Being effective in high dimensional spaces such as dealing with data with many attributes. SVM is effective in the event that the number of dimensions is more than the number of samples as well as being relatively memory efficient.

##### 2.3.1.2.1 SVM Applications

Image-based analysis: SVM can be used in quite a number of tasks when it comes to working with images especially when it comes to facial recognition and extraction. When working with facial features there is a need for an algorithm that can perform classification well based on fine-tuned feature extractions.

Text-based application: SVM can be used to classify handwriting of multiple people. Support Vector Machine models train better when applied to tasks such as detection of the fine structures of handwriting such as straights and curves.

Security: SVM is very much applicable in security-based tasks this is due to the use of SVMs for basic encryption as well as the complex analytics of different materials in order to view and even break through encryptions and other measures of security.

#### 2.3.1.3 Decision Tree

This is a support tool which has a tree-like structure that models possible outcomes, cost of resources, utilities and possible consequences. Decision trees provide a means to portray algorithms with conditional control statements (Decision Tree, 2021).

##### 2.3.1.3.1 Advantages of Decision Trees

Easy to read: Decision trees are simple and easy to read and interpret without needing extra statistical knowledge to understand the outputs.

Easy to prepare: It takes little effort in order to prepare data but users do require to have ready information in order to create new variables with the power to predict the target variable.

Little data cleaning needed: A major advantage of decision trees is the amount of data cleaning needed after the creation of the variables. This is due to there being less significance on the decision tree’s data when there are missing values and outliers.

##### 2.3.1.3.2 Application of Decision Trees

Assessing possible opportunities: Decision trees can be used to evaluate possible opportunities for businesses based on historical data such as on sales which can assist businesses in making strategic decisions to aid in the growth of the business.

Finding prospective clients: Through the use of demographic data decision trees are able to be used to assist in streamlining a marketing budget and in making informed decisions on the business’s target market. This allows the business to be focused on a target demographic hence improving revenue.

Useful support tool: Decision trees are used in research operations. They can assist in figuring out the best and most appropriate strategies that will allow the company to achieve its intended goals.

## 2.4 Comparative Study

Comparative study is done to compare and analyze two or more ideas. These are used to examine how two or more studies vary from each other or how alike the studies are. This allows for one study to be compared and contrasted with multiple studies thus showing how similar or contrasted two subjects or ideas are.

**2.4.1 Existing Studies**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Author/ Date**  **Date** | **Theoretical/**  **Conceptual**  **Framework** | **Research**  **Question(s)/**  **Hypotheses** | **Methodology** | **Analysis &**  **Results** | **Conclusions** | **Implications for**  **Future research** | **Implications**  **For practice** |
| Benny, D. (2021) | Time Series Analysis: Forecast COVID-19 Vaccination Rate | Asks the questions when 100 percent vaccination rate per 100 people will be achieved and how many people will be vaccinated in the incoming 50 days. | For the time series analysis ARIMA was used which was for the purpose of forecasting large number of time series from past values of that time series. The data for the analysis was acquired from the University of Oxford’s Our World in Data project. | Different orders of differencing of time series were done to remove all the trend and seasonality in order to make the data stationary then put into the ARIMA model then AIC, Akaike’s Information Criterion, is used to observe and identify the best ARIMA model from amongst them. The best ARIMA model is the one which has second order differencing. | There are many factors affecting the vaccination rate in a country. The predictors comprised of lags of the dependent feature which is the vaccination rate here and lags of the forecasting errors. |  |  |
| Hariharan, R., Sundberg, J., Gallino, G., Schmidt, A., Arenth, D., Sra, S. and Fels, B. (2020) | An Interpretable Predictive Model of Vaccine Utilization for Tanzania | Proposes that accurately utilizing forecasts is important in maintaining stocks of vaccines in medical facilities. States that current approaches to vaccine utilization forecasts are based many a times on outdated data. | The data used was the recent vaccine utilization data from 710 health facilities in Tanzania along with the publicly available highly multivariate data from health facilities. Random Forest Regressor was trained to accurately forecast vaccine utilization. | Two different measures were used to perform an evaluation of model performance, these are Root Mean Square Error, RMSE, used to evaluate performance between predicted and biweekly utilization rates and Fractal Error, F. E, used to measure the difference between the actual value and the predicted value. Multiple common machine learning algorithms were used in order to select the best model with the results being that the models were unable to succeed in yielding comparable performance to other classic machine learning models as well as gave predictions that could not be understood. This improved the choice of RFR to be used to forecast vaccine utilization. | This is the first time an interpretable predictive model has been used to forecast the utilization of vaccines which has a broad scope and can be changed to be used in different regions and countries. | The data used is subject to a few restrictions such as requiring permission of the United Republic of Tanzania and in order to view the data permission must be acquired from the Government of Tanzania, the rightful owners of the data. | From the different RFR models used one has high predictive performance and can be utilized in places where there is availability of recent data on vaccine utilization while another model has less predictive performance but is adaptable to different regions and countries. |
| Shaharudin, S., Ismail, S., Hassan, N., Tan, M. and Sulaiman, N. (2021) | Short-Term Forecasting of Daily Confirmed COVID-19 Cases in Malaysia Using RF-SSA Model | The RF-SSA model has evidently had over-forecasted cases by 0.36% which signifies that its use in predicting the impending number of Covid-19 cases as competent. Despite the competence in its predictions an enhanced RF-SSA algorithm should be developed for higher effectiveness in capturing extreme changes in data. | Data was gathered from MOH records regarding the daily Covid-19 prevalence from 25th January to 29th April 2020. SSA model was used to decompose, reconstruct the data into several additive components which could then be defined into trends, seasonal and noise components. The forecasting algorithm of SSA, RF, is used to predict the future cases of Covid-19. The model is henceforth known as SSA-RF. | The SSA-RF model was able to predict 32 days of cases accordingly except for a drop in cases from 5th April due to the announcement of MCO. The SSA-RF model is seen to work best when data is in a consistent or stable pattern over time with low number of outliers. Sudden spikes in data lead to low forecasting performance. Mainly used for short-term forecasts. | The application of the SSA-RF model is specified to be advantageous for health authorities to use in terms of devising prompt and efficient strategies to flatten the curve of Covid-19 cases. The model is useful to health authorities in better comprehend the outbreak pattern. Overall results show that there is reasonable accuracy in forecasting this pandemic using this model. | The model failed to capture the sudden drop of cases and thus more information is required in order to improve the model to more accurately predict cases over a long period of time. For now the case definition and data collection must be maintained in real-time in order to enhance RF-SSA for further evaluation. |  |
| Shastri, S., Singh, K., Kumar, S., Kour, P. and Mansotra, V. (2020) | Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study | This study proposes the use of deep learning-based models to make predictions on the confirmed and death cases of Covid-19 for USA and India and present comparative case analysis. Through the use of RNN based LSTM, Bi-directional LSTM and Convolutional LSTM to derive forecasted results which can produce better future prediction results than traditional LSTM models. | For this study datasets of India and USA confirmed and death cases of Covid-19 were acquired from the  the Centers for Disease Control and Prevention, U.S Department of Health and Human Services and the Ministry of Health and Family Welfare, Government of India respectively. A variant of RNN is used to overcome the limitations of RNN known as LSTM. | Three variants of LSTM are used to carry out experimentation stacked LSTM which combines multiple LSTM layers to create a mode of greater complexity and increase depth of the model, Bi-directional LSTM which splits standard RNN into forward and backward states with both states not being inputs for each other and Convolutional LSTM which is able to determine the state of cell in the future which goes along to be determined by inputs and past states of its local neighbors. The best model among the three is ConvLSTM model which has predicted results that are quite similar with real case scenarios of both countries. | The study used deep learning predictions with data from both USA and India with the death and confirmed cases of Covid-19 being taken into consideration. The best model was decided by the error rate of the predictions which is calculated using MAPE. |  |  |
| Zhou, X. and Li, Y.  (2021) | Forecasting the vaccine uptake rate: An infodemiological study in the US | The CDC publishes daily statistics on vaccinations across the US but due to a 1–3-week delay for accurate data, traditional forecasting methods are lacking in the accuracy needed when using clinical data. This study thus developed a framework for predicting vaccination rate on the basis of traditional clinical data using ARIMA. | Data sources for this study were obtained from the CDC’s daily publishes from January 2nd to July 27th, 2021 which was used as the clinical data. AR/ARIMA models are used to forecast future vaccine uptakes solely with the use of the clinical data. | AR/ARIMA models require stationary time-series data which can be confirmed through the use of Dickey-Fuller and unit-root tests. To make the data stationary it needs to go through differencing. Different models use data at different levels of differencing. Models that used web data were also tested in comparison with the clinical data. Multiple models were created such as AR, ARIMA, OLS, LASSO, boost and randomforest using tested on both clinical data and web data.  The results were that models which used only clinical data fared better than those that used only web data with ARIMA performing better than all other models. | In conclusion the forecasts in this study are useful to policymakers and health officials in foreseeing the vaccination uptake rate and relevant forecast models can be created and applied to other countries and future epidemic events. |  |  |
| Cihan, P. (2021) | Forecasting fully vaccinated people against COVID-19 and examining future vaccination rate for herd immunity in the US, Asia, Europe, Africa, South America, and the World | This study aims to study the amount of people that have been fully vaccinated against Covid-19 around the world and perform an analysis on whether the number is sufficient for herd immunity. | Proposes the use of the ARIMA model to estimate the prevalence trend of the vaccinations for Covid-19. Dataset for the study was acquired from the Our World in Data. | The criteria for the most successful ARIMA model selection were through the use of AIC, RMSE and MAPE. The results of which were an ARIMA model with minimum AIC, RMSE and MAPE criteria were selected as the best models. | The models selected by the AIC, RMSE and MAPE criteria were selected with the results of the study that on 1st June 2021 in US, Asia, Europe Africa, South America and the rest of the world will have 41.8%, 2.3%, 17%, 0.6%, 8.8 % and 5.6% fully vaccinated people, respectively. |  |  |
| Toharudin, T., Pontoh, R., Caraka, R., Zahroh, S., Kendogo, P., Sijabat, N., Sari, M., Gio, P., Basyuni, M. and Pardamean, B. (2021) | National Vaccination and Local Intervention Impacts on COVID-19 Cases | The study was conducted to determine how effective the PPKM in Jakarta and its vaccination program is against the number of new cases daily. | Datasets for this study are from Jakarta, Indonesia on the historical data of daily new cases, recovered cases and death cases of Covid-19 from March 1st 2020 to March 18th 2021. The analysis is performed using NNAR, an artificial neural network method. | Proposes the use of the lag values, in the time series, as input for an artificial neural network. Trial and error method is used to determine the best model hence the best network architecture chosen will be the one with lowest error value in the validation phase. | This method is concluded to be useful in the generation of daily forecasts of Covid-19 cases. | The success of vaccination programs should be evaluated over a significant period of time to gauge their efficacy in general. | There were limitations due to the availability of data and understanding of the rates of death and transmissibility. |
| Liao, Z., Lan, P., Fan, X., Kelly, B., Innes, A. and Liao, Z. (2021) | SIRVD-DL: A COVID-19 deep learning prediction model based on time-dependent SIRVD | Proposes a SIRVD-DL Covid-19 prediction model through the use of deep learning method based on time-dependent SIRVD to create a combination of the two in order to create a model that is more explanatory and effective in prediction | The sources of data are from two places. The first time series data from JHU CSSE and the second from Our World in Data. The model to be used is mainly the SIRVD model. | The SIRVD-DL prediction model was proposed. Numerical experiments were made and analyzed to gauge the effectiveness of the model. The results of which was that the prediction model is a 50% improvement on single-day prediction in comparison to pure deep-learning methods and is adaptable in short-term and medium-term predictions. | The proposed model combines mathematical model and deep-learning model to improve upon the problem in single deep-learning methods. |  |  |
| Yadav, M., Perumal, M. and Srinivas, M. (2020) | Analysis on novel coronavirus (COVID-19) using machine learning methods | The main objectives of this study are to predict the spread of Covid-19, Analyze growth rates and methods of mitigation, predict the end of the epidemic, Analyze virus transmission rate and correlation between weather and Covid-19 | The proposed method is the use of SVR which is used for classification and regression problems. | In order to analyze the performance of the proposed method simple well known regression methods have been used for comparisons such as Simple Linear Regression, Polynomial Regression with the results being 98.8%, 96.2% and 62.1% accuracy respectively. | This study proposes that the use of SVR method for analysis by using supported vectors instead of simple regression line in order to gain a higher classification accuracy. In most of the tasks the proposed method has provided promising results when compared to the well-known regression methods. |  |  |
| Marzouk, M., Elshaboury, N., Abdel-Latif, A. and Azab, S. (2021) | Deep learning model for forecasting COVID-19 outbreak in Egypt | The aim of this study is to apply artificial intelligence-based models to be used to predict the prevalence of the coronavirus in Egypt. | Dataset used is obtained from a website(<https://flevy.com/coronavirus>) which has been acknowledged by the Egyptian ministry of health and population. The models used are LSTM, convolutional neural network and multilayer perceptron neural network. | The three chosen models are applied to forecast the total number of infections over a short period of time. Through the application of different assessment criteria in order to find out the optimal model which results in LSTM outperforming the other two models with MLP model failing to provide acceptable results. | The LSTM model outperformed the remaining machine learning and deep-learning models due to the feedback connections characterizing the LSTM network to propagate the data in the backward pass which would thus improve the accuracy of predictions. | The reliability of the results is mainly reliant on the quality of the data because under-testing could lead to a lower number of reported cases and thus a larger number of forecasting errors. It is recommended to apply moving average techniques in order to smooth out variations and separate noise from trends in the time series data. |  |

**Table 2.1 Literature Review Matrix**

## 2.5 Tools

### 2.5.1 RStudio

RStudio is an IDE, Integrated Development Environment, for the programming language R. RStudio contains features such as a console, an editor that highlights syntax and supports direct code execution, it also has tools needed for plotting, history, debugging and workspace management. RStudio will mainly be used to execute the R code which will be used to perform all the important tasks such as data cleaning, data, data preparation, creation of plots, etc.

### 2.5.2 Microsoft Excel

Microsoft office has a suite of applications that are useful in different circumstances one such program is Microsoft excel which is a spreadsheet program. Spreadsheets present values in tables arranged in rows and columns that can be manipulated in order to acquire the desired outcome. Microsoft excel will be used to store the datasets to allow for ease of importing the datasets when needed into RStudio.

## 2.6 Summary

The author looks into the effects caused by Covid-19, how vaccines can be useful in the battle against Covid-19 and why the rate of vaccination is decreasing rather than increasing. The author also looks into research articles to look at their data sources, the techniques they have used, the effectiveness of those techniques and finally the author talks about the tools and programs that will be used in the course of this study.

# References

Assefa, N., Sié, A., Wang, D., Korte, M., Hemler, E., Abdullahi, Y., Lankoande, B., Millogo, O., Chukwu, A., Workneh, F., Kanki, P., Baernighausen, T., Berhane, Y., Fawzi, W. and Oduola, A., 2021. Reported Barriers to Healthcare Access and Service Disruptions Caused by COVID-19 in Burkina Faso, Ethiopia, and Nigeria: A Telephone Survey. The American Journal of Tropical Medicine and Hygiene, 105(2), pp.323-330.

Benny, D., 2021. COVID-19 Vaccination Rate | Analysis of COVID-19 Vaccination Rate. [online] Analytics Vidhya. Available at: <https://www.analyticsvidhya.com/blog/2021/04/time-series-analysis-forecast-covid-19-vaccination-rate/> [Accessed 16 October 2021].

Benz, J., Research, P., Oppenheimer, C., Health, P. and Burke-Garcia, A., 2021. Many remain doubtful about getting COVID-19 vaccine - AP-NORC. [online] AP-NORC. Available at: <https://apnorc.org/projects/many-remain-doubtful-about-getting-covid-19-vaccine/> [Accessed 16 October 2021].

Centers for Disease Control and Prevention. 2021. COVID-19 Vaccine Facts. [online] Available at: <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/facts.html> [Accessed 16 October 2021].

Chu, J., 2021. A statistical analysis of the novel coronavirus (COVID-19) in Italy and Spain. PLOS ONE, 16(3), p.e0249037.

Cihan, P., 2021. Forecasting fully vaccinated people against COVID-19 and examining future vaccination rate for herd immunity in the US, Asia, Europe, Africa, South America, and the World. Applied Soft Computing, 111, p.107708.

Corporate Finance Institute. 2021. Decision Tree. [online] Available at: <https://corporatefinanceinstitute.com/resources/knowledge/other/decision-tree/> [Accessed 26 October 2021].

Cutler, D. and Summers, L., 2020. The COVID-19 Pandemic and the $16 Trillion Virus. JAMA, 324(15), p.1495.

Donthu, N. and Gustafsson, A., 2020. Effects of COVID-19 on business and research. Journal of Business Research, 117, pp.284-289.

Education, I., 2021. What are Neural Networks?. [online] Ibm.com. Available at: <https://www.ibm.com/cloud/learn/neural-networks> [Accessed 26 October 2021].

Fauci, A., Lane, H. and Redfield, R., 2020. Covid-19 — Navigating the Uncharted. New England Journal of Medicine, 382(13), pp.1268-1269.

Federman, R., 2021. Understanding Vaccines: A Public Imperative. [online] PubMed Central (PMC). Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4257029/> [Accessed 16 October 2021].

Gandhi, R., 2021. Support Vector Machine — Introduction to Machine Learning Algorithms. [online] Medium. Available at: <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47> [Accessed 26 October 2021].

Guerin, P., Singh-Phulgenda, S. and Strub-Wourgaft, N., 2020. The consequence of COVID-19 on the global supply of medical products: Why Indian generics matter for the world?. F1000Research, 9, p.225.

Guides.lib.vt.edu. 2021. Research Guides: Research Methods Guide: Interview Research. [online] Available at: <https://guides.lib.vt.edu/researchmethods/interviews> [Accessed 20 September 2021].

Habas, K., Nganwuchu, C., Shahzad, F., Gopalan, R., Haque, M., Rahman, S., Majumder, A. and Nasim, T., 2020. Resolution of coronavirus disease 2019 (COVID-19). Expert Review of Anti-infective Therapy, 18(12), pp.1201-1211.

Hariharan, R., Sundberg, J., Gallino, G., Schmidt, A., Arenth, D., Sra, S. and Fels, B., 2020. An Interpretable Predictive Model of Vaccine Utilization for Tanzania. Frontiers in Artificial Intelligence, 3.

Ibm.com. 2021. IBM Docs. [online] Available at: <https://www.ibm.com/docs/en/spss-modeler/SaaS?topic=dm-crisp-help-overview> [Accessed 20 September 2021].

Khan, S., Haleem, A., Deshmukh, S. and Javaid, M., 2021. Exploring the Impact of COVID-19 Pandemic on Medical Supply Chain Disruption. Journal of Industrial Integration and Management, 06(02), pp.235-255.

Liao, Z., Lan, P., Fan, X., Kelly, B., Innes, A. and Liao, Z., 2021. SIRVD-DL: A COVID-19 deep learning prediction model based on time-dependent SIRVD. Computers in Biology and Medicine, 138, p.104868.

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K., Lau, E., Wong, J., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Liu, M., Tu, W., Chen, C., Jin, L., Yang, R., Wang, Q., Zhou, S., Wang, R., Liu, H., Luo, Y., Liu, Y., Shao, G., Li, H., Tao, Z., Yang, Y., Deng, Z., Liu, B., Ma, Z., Zhang, Y., Shi, G., Lam, T., Wu, J., Gao, G., Cowling, B., Yang, B., Leung, G. and Feng, Z., 2020. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. New England Journal of Medicine, 382(13), pp.1199-1207.

Marzouk, M., Elshaboury, N., Abdel-Latif, A. and Azab, S., 2021. Deep learning model for forecasting COVID-19 outbreak in Egypt. Process Safety and Environmental Protection, 153, pp.363-375.

Maude, R., 2021. Southeast Asia and COVID-19 Vaccines Explained - Asia Society Policy Institute. [online] Southeast Asia and COVID-19 - Asia Society Policy Institute. Available at: <https://southeastasiacovid.asiasociety.org/southeast-asia-and-covid-19-vaccines-explained/> [Accessed 16 October 2021].

McLeod, S., 2018. Questionnaire: Definition, Examples, Design and Types. [online] Simplypsychology.org. Available at: <https://www.simplypsychology.org/questionnaires.html> [Accessed 20 September 2021].

Medium. 2021. Introduction to Neural Networks, Advantages and Applications. [online] Available at: <https://towardsdatascience.com/introduction-to-neural-networks-advantages-and-applications-96851bd1a207> [Accessed 26 October 2021].

Nachemson, A., 2021. As COVID surges, Southeast Asia struggles to vaccinate citizens. [online] Aljazeera.com. Available at: <https://www.aljazeera.com/news/2021/8/24/as-covid-surges-southeast-asia-struggles-to-vaccinate-citizens> [Accessed 16 October 2021].

Niazkar, H. and Niazkar, M., 2020. Application of artificial neural networks to predict the COVID-19 outbreak. Global Health Research and Policy, 5(1).

Paltiel, A., Schwartz, J., Zheng, A. and Walensky, R., 2021. Clinical Outcomes Of A COVID-19 Vaccine: Implementation Over Efficacy. Health Affairs, 40(1), pp.42-52.

Pawar, M., 2020. The Global Impact of and Responses to the COVID-19 Pandemic. The International Journal of Community and Social Development, 2(2), pp.111-120.

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D. and Roser, M., 2021. Coronavirus Pandemic (COVID-19). [online] Our World in Data. Available at: <https://ourworldindata.org/covid-vaccinations> [Accessed 27 October 2021].

Shaharudin, S., Ismail, S., Hassan, N., Tan, M. and Sulaiman, N., 2021. Short-Term Forecasting of Daily Confirmed COVID-19 Cases in Malaysia Using RF-SSA Model. Frontiers in Public Health, 9.

Shastri, S., Singh, K., Kumar, S., Kour, P. and Mansotra, V., 2020. Time series forecasting of Covid-19 using deep learning models: India-USA comparative case study. Chaos, Solitons & Fractals, 140, p.110227.

Shmueli, L., 2021. Predicting intention to receive COVID-19 vaccine among the general population using the health belief model and the theory of planned behavior model. BMC Public Health, 21(1).

Su, Z., Wen, J., McDonnell, D., Goh, E., Li, X., Šegalo, S., Ahmad, J., Cheshmehzangi, A. and Xiang, Y., 2021. Vaccines are not yet a silver bullet: The imperative of continued communication about the importance of COVID-19 safety measures. Brain, Behavior, & Immunity - Health, 12, p.100204.

Toharudin, T., Pontoh, R., Caraka, R., Zahroh, S., Kendogo, P., Sijabat, N., Sari, M., Gio, P., Basyuni, M. and Pardamean, B., 2021. National Vaccination and Local Intervention Impacts on COVID-19 Cases. Sustainability, 13(15), p.8282.

Tregoning, J., Brown, E., Cheeseman, H., Flight, K., Higham, S., Lemm, N., Pierce, B., Stirling, D., Wang, Z. and Pollock, K., 2020. Vaccines for COVID‐19. Clinical & Experimental Immunology, 202(2), pp.162-192.

Who.int. 2021. Coronavirus. [online] Available at: <https://www.who.int/health-topics/coronavirus#tab=tab\_1> [Accessed 20 September 2021].

Who.int. 2021. Immunization coverage. [online] Available at: <https://www.who.int/news-room/fact-sheets/detail/immunization-coverage> [Accessed 27 October 2021].

Who.int. 2020. Impact of COVID-19 on people's livelihoods, their health and our food systems. [online] Available at: <https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems> [Accessed 20 September 2021].

Yadav, M., Perumal, M. and Srinivas, M., 2020. Analysis on novel coronavirus (COVID-19) using machine learning methods. Chaos, Solitons & Fractals, 139, p.110050.

Zhou, X. and Li, Y., 2021. Forecasting the vaccine uptake rate: An infodemiological study in the US. [online] arXiv.org. Available at: <https://arxiv.org/abs/2109.13971> [Accessed 11 October 2021].

# Appendix