COVID-19 CASE ANALYSIS

Abstract

Today world thinks about coronavirus disease that which means all even this pandemic disease is not unique. The purpose of this study is to detect the role of machine-learning applications and algorithms in investigating and various purposes that deals with COVID-19. Review of the studies that had been published during 2020 and were related to this topic by seeking in Science Direct, Springer, Hindawi, and MDPI using COVID-19, machine learning, supervised learning, and unsupervised learning as keywords. The total articles obtained were 16,306 overall but after limitation; only 14 researches of these articles were included in this study. Our findings show that machine learning can produce an important role in COVID-19 investigations, prediction, and discrimination. In conclusion, machine learning can be involved in the health provider programs and plans to assess and triage the COVID-19 cases. Supervised learning showed better results than other Unsupervised learning algorithms by having 92.9% testing accuracy. In the future recurrent supervised learning can be utilized for superior accuracy.

Introduction

Recently, the world gained rapid progression in technology and it shows an important role in the developed countries. Nowadays all daily life sectors such as education, business, marketing, militaries, and communications,

engineering, and health sectors are dependent on the new technology applications. The health care center is a crucial field that strongly needs to apply the new technologies from defining the symptoms to the accurate diagnosis and digital patient's triage. Coronavirus-2 (SARSCoV-2) causes severe respiratory infections, and respiratory disorders, which results in the novel coronavirus disease 2019 (COVID-19) in humans who had been reported as the first case in Wuhan city of China in December 2019. Later, SARS-CoV-2 was spread worldwide and transmitted to millions of people and the world health organization (WHO) have announced the outbreak as a global pandemic since the number of infected people is still increasing day by day. As of 16th December 2020, the total (global) coronavirus cases were approximately 73,806,583 with reported deaths of 1,641,635 (Pasupuleti et al. 2021). The novel coronavirus appeared in December 2019, in the Wuhan city of China and the World Health Organization (W.H.O) reported it on 31st December 2019. The virus produced a global risk and W.H.O named it COVID-19 on 11th February 2020 (Wu 2020). Up to the present time, there was no specific medication that deals directly with this new generation of COVID-19 virus, but some of the companies produced several combination drugs that basically made up from ethanol, isopropyl alcohols, and hydrogen peroxides in different combinations show a significant reaction to the novel virus and had been confirmed and accepted by WHO to be used in the world (Mahmood et al. 2020). The artificial intelligence and deep learning algorithm show the ability to diagnose COVID-19 in precise which can be regarded as a supportive factor to improve the common diagnostic methods including Immunoglobulin M (IgM), Immunoglobulin (IgG), chest x-ray, and computed tomography(CT) scan, also

reverse transcription-polymerase chain reaction (RT-PCR) and immunochroma to graphic fluorescence assay. The developments of a potential technology are one of the currently used methods to identify the infection, such as a drone with thermal screening without human intervention, which needs to be encouraged (Manigandan 2020). The assessment of the research that had been produced whether it hits the target of the existing knowledge gaps or not can be done by applying an artificial intelligence/machine learning-based approach to analyze COVID-19 literature (Doanvo et al. 2020). Thus, the acceleration of the diagnosis and treatment of COVID-19 disease is the main advantage of these AI-based platforms (Naseem et al. 2020) which finally shows a huge potential to exponentially enhance and improve health care research (Jamshidi et al. 2020). Corona Virus Disease 2019 (COVID-19), has become a matter of serious concern for each country around the world. AI applications can assist in increasing the accuracy and the speed of identification of cases through data mining to deal with the health crisis efficiently, the rapid expansion of the pandemic has created huge health care disorders which as a result encouraged the real need for immediate reactions to limit the effects. Artificial Intelligence shows great applications in dealing with the issue on many sides (Tayarani-N 2020). The COVID-19 is an epidemic disease that challenged human lives in the world. The systematic reviews showed that machine learning ML training algorithms and statistical models that are used computers to perform various tasks without explicit commands (Bishop 2006). Currently, machine learning techniques are used internationally for predictions due to their accuracy. However, machine learning (ML) techniques, have few challenges such as the new poor database that is available

online. For instance, the selection of the appropriate parameters is one of the challenges involved in training a model or the selection of the best Machine learning model for prediction. Depending on the available dataset researchers obtained predictions by using the best Machine learning model that suits the dataset (Shinde 2020). Machine learning techniques can be used to extract hidden patterns and data analytics (Khan 2020). The algorithms of Machine-learning are designed for identifying complex patterns and interfaces in the data, in the context of unknown and complicated correlation patterns among risk factors (Hossain 2019).

Related work

COVID-19

The contagion disease caused by the SARS-COV-2 virus named COVID-19 is requiring extraordinary responses of special intensity and possibility to more than 200 countries around the world, the first 4 months from its epidemic, the number of infected peoples ranged from 2 to 20 million, with at least 200,000 deaths. To manage the spread of the COVID-19 infection among people rapidly, all the governments around the world applied severe actions, such as the quarantine of hundreds of millions of citizens worldwide (Alimadadi 2020). Nevertheless, the difficulty of distinguishing between the positive and negative COVID-19 individuals depending on the various symptoms of COVID-19, all of these efforts are limited. Therefore, tests to detect the SARS–CoV-2 virus are believed to be critical to recognize the positive cases of this infection to limit the (Brinati et al. 2020). Radiology and imaging are some of the most beneficial and critical modalities used for diagnosis

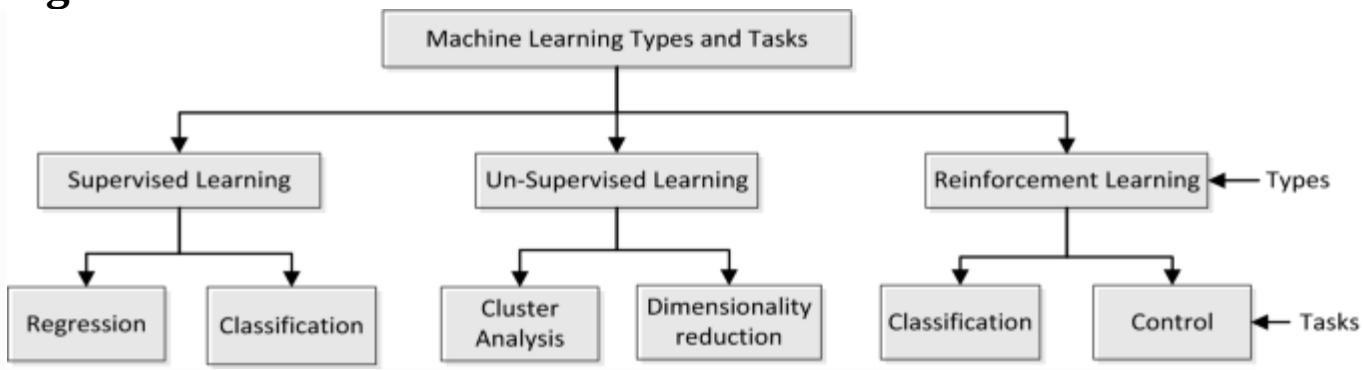
COVID-19 stage and hazards on the patient's lungs specifically by chest CT scan (Day 2020). Early diagnosis of COVID-19 is important to minimize human-to-human transmission and patient care. Recently, the separation and quarantine of healthy people from the infected or persons who suspect that they are carrying the virus is the most effective technique to avoid the spread of COVID-19 (Deng 2020). Machine-learning techniques role showed an important understandings of the COVID-19 diagnosis, such as lung computed tomography (CT) scan whether it can be regarded as the first screening or an alternative test for the real-time inverse transcriptase–polymerase chain reaction (RT–PCR), and the differences between COVID-19 pneumonia and other viral pneumonia using CT scan of the lungs(Kassani et al. (2004)).

Machine learning

Machine learning is one of the most promising tools in classification (Hossain 2019). In essence; machine learning is a model that aims to discover the unknown function, dependence, or structure between input and output variables. Usually, these relations are difficult to be existed by explicit algorithms via automated learning process (Zhang 2020a). Machine-learning methods are applied to predict possible confirmed cases and mortality numbers for the upcoming (Hastie et al. 2009). Machine learning can be divided into two parts. The first part is to define the optimal weight of data fusion of multi-node perception outcomes and eliminate unusable nodes based on the genetic algorithm, while the second part is to find fault nodes through a fault recognition neural network (Ünlü and Namlı 2020). Machine learning is a subsection of

Artificial Intelligence (AI), and it involves several learning paradigms, such as Supervised Learning (SL), Un-supervised Learning (UL), and Reinforcement Learning (RL) (Shirzadi 2018). Typical ML models consist of classification, regression, clustering, anomaly detection, dimensionality reduction, and reward maximization (Gao 2020). The ML algorithms are trained in the SL paradigm, on labeled data sets, meaning that they exist to a ground-truth output (continuous or discrete) for every input. Conversely, in UL (Bishop (2006)) there is no ground-truth output, and the algorithms normally attempt to discover patterns in the data. Reinforcement Learning aims to raise the cumulative reward so that it is more suitable for sequential decision-making tasks (Zhang 2020b). Supervised learning has regression and classification; unsupervised learning includes cluster analysis and dimensionally reduction, also Reinforcement Learning (RL) includes classification and control, as illustrated in Fig. 1.

Fig. 1



Overview of machine-learning types and tasks

COVID-19 with machine learning

Recently there are three different perspectives of work that had been done on edge computing and the detection of (COVID-19) Cases. The viewpoints are including the recognizing of (COVID-19) cases by machine-learning systems (Table 1). The algorithms for the recognition of activity from machine learning and the approaches which used in edge computing are considered the Imaging workflows that can inspire machine-learning methods that are able of supporting radiologists who search for an analysis of complex imaging and text data. For the novel COVID-19 there are models capable of analyzing medical imaging and recognizing COVID-19 (Shirzadi 2018). Artificial intelligence AI has various types, machine learning (ML), is one of these applications, it had been applied successfully to different fields of medicine for detection of new genotype–phenotype associations, diagnosis, which showed effects on assessment, prediction, diseases classification, transcriptomic, and minimizing the death ratio(Gao 2020).

Literature searching strategy and article selection

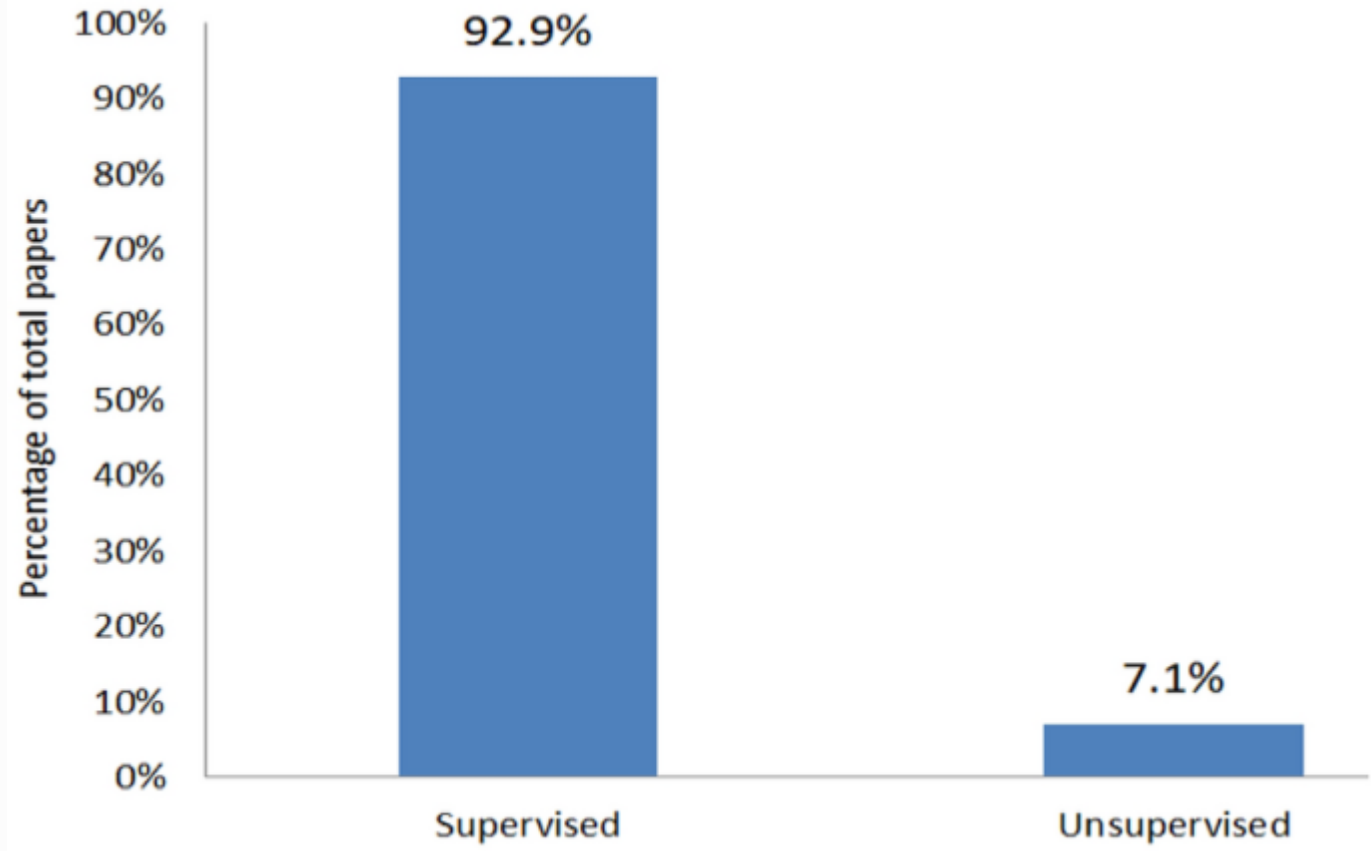
This systematic review paper used articles from online digital databases, which include Science Direct, Springer, Hindawi, and MDPI databases, two independent authors started the search strategy from October 2020 until December 2020. The used keywords were “COVID-19; Machine Learning;

Supervised Learning; Un-supervised Learning.'' They were connected to the relevant articles using "and'', or "or'' to find the studies that deals with human disease and COVID-19. The total number of the studies were (16,306) articles from all the databases, according to the inclusion and exclusion criteria this number was limited. The limitation includes selecting the publication year (2019–2021), the articles type original articles that had been published as journal articles in English language only included. This selection strategy reduced the total number to 5054 articles, then after quality assessment of these studies there was 395 articles which remained, then finally the full text article reading minimized the last included articles to 14. The included articles are presented according to the author's name, publication's year, country, the used dataset, the applied method, and finally their results in (Table 2).

Table 2 Supervised and un-supervised machine learning for analyzing the COVID-19 disease that included articles with the related details of the Dataset, author name, country of publication, year of publication, the used method in the study, and their results

Machine-learning types applied

According to Fig. 2, supervised learning is the dominant machine-learning type applied for production lines. The majority of studies used both supervised learning methods which were (92.9%), whereas unsupervised learning was (7.1%).



Distribution of machine-learning types

PROPOSED:

The aim of this project is to create an innovative application for testing Coronavirus disease which has made an economic and social disruption. The motivation behind creating this web application was the need of a platform to showcase how this pandemic can be stopped from further spread. Testing of all people for SARSCoV-2, including those who have no symptoms, who show symptoms of infection such as trouble breathing, fever, sore throat or loss of the sense of smell and taste, and who may have been exposed to the virus

will help prevent the spread of COVID-19 by identifying people who are in need of care in a timely fashion. A positive test early in the course of the illness enables individuals to isolate themselves – reducing the chances that they will infect others and allowing them to seek treatment earlier, likely reducing disease severity and the risk of long-term disability, or death. A positive test for SARS-CoV-2 alerts an individual that they have the infection. Not only can they get treated faster, but they can take steps to minimize the spread of the virus. From the academic perspective, the objective of this project to develop a scalable, reliable and secure website with an efficient storage backend mechanism and a neat frontend. To grasp the concepts of Database Management System, designing a real-time database, maintaining it and manipulating it is essential. With this project, these principles are to be learned and applied. The focus is towards designing an ER diagram that caters to our data needs, mapping this ER diagram into relational schema, creation of tables using this schema and establishment of relationships within these tables. The tables are aimed at achieving the highest Normal Forms, thus being efficient and reduce redundancy. Following are the features of Covid-19 website: • Admin's signup/login • Phlebotomist portal for testing samples • User registration • Three tests available for user to get tested • Guidelines to follow Covid-19 protocols • Website uses cookie policy to track user's behavior

EXISTING:

January 25-March 13, 2020

Health screenings at airports and border crossings

February 26-March 20, 2020

Introduction of quarantine policies: gradually for passengers coming from different countries

February 26-March 13, 2020

Visa restrictions: gradually for different countries

March 5, 2020

Limit public gatherings (closure of some selected public institutions like museums, religious places, and postponing of several local elections to stop public gatherings)

March 11, 2020

Border checks

March 13-15, 2020

Border closure

March 16, 2020

Limit public gatherings (ban on all sorts of public gatherings and meetings, and stopping people from making any congregation)

March 18, 2020

Travel restrictions

March 20, 2020

Testing for the coronavirus disease (before this point, only people who had traveled from abroad were tested; this point onwards, testing was also introduced for symptomatic contacts of laboratory-confirmed cases, symptomatic health care workers, and all hospitalized patients with severe acute respiratory illness)

March 22, 2020

Flight suspensions

March 22, 2020

Cancellation of passenger train services until March 31, 2020

March 24, 2020

Suspension of domestic airplane operations

March 25, 2020

21-day lockdown of entire country

March 25, 2020

Cancellation of passenger train services extended to April 14, 2020

March 30, 2020

Increase of quarantine/isolation facilities

April 14, 2020

Extension of lockdown until May 3, 2020

May 1, 2020

Extension of lockdown until May 17, 2020

SYSTEM DESIGN:

Coronavirus or 2019-nCoV is not, at this point, pandemic but instead endemic, with in excess of 14 million complete cases all throughout the planet getting the infection. At present, there is no particular treatment or solution for Coronavirus, and hence living with the sickness and its manifestations is unavoidable. The connection coefficient examination between different needy and free highlights was done to decide a strength connection between every reliant

element and autonomous component of the dataset before building up the models. The database is divided into two parts, 80% of the database is used for model training and the remaining 20% is used for model testing and evaluation. In 2019, early Coronavirus predictions is useful to reduce colossal weight on medical service panels through the diagnosis of coronavirus patients. In the proposed work in this paper, Naive Bayes, Decision tree, Support Vector Machine (SVM) and Artificial neural network (ANN) models are used for forecasting COVID-19 prediction and occurrences.

SYSTEM REQUIREMENT:

From the interviews, 130 requirements could be retrieved. Because of the high number of requirements, the most important ones are presented here. On the one hand, requirements are considered important when they were mentioned by at least four interviewees, called multiply mentioned requirements in the further course, and on the other hand, requirements that are part of high priority and high severity use cases. Requirements with multiple mentions are listed by name in **Table 1**.

Table 1. Overview of multiply mentioned requirements (REQ) with their respective title.

The requirements can be divided into organizational (REQ 10, REQ 23, REQ 55), artifact-related (REQ 02, REQ 05, REQ 19), where REQ 60 and REQ 81 can be interpreted as child-requirements of REQ 02, and non-functional requirements (REQ 46). Artifact-related requirements refer directly to the artifacts or their change. In the project context, artifacts include all elements that can be subject to changes: requirements, documents, products or components, (software) functions, and processes. The requirements that form the basis of the prioritized use cases can be found in **Table 2**, associated with the underlying use cases that are examined in more detail in the following section.

Table 2. Requirements assigned to the most important use cases (UC).

The prioritization of all use cases as the average interviewee assessment illustrates **Figure 3**. As can be observed, use cases with a low (high) priority were also evaluated with a low (high) severity. The prioritization serves as a basis for deciding which use cases and respective requirements will be implemented first in the implementation phase of the assistance system development. In the given prioritization, use cases UC001, UC008.a, UC012.a, UC012.b, UC012.f and UC003.b represent those with the highest priority and severity.

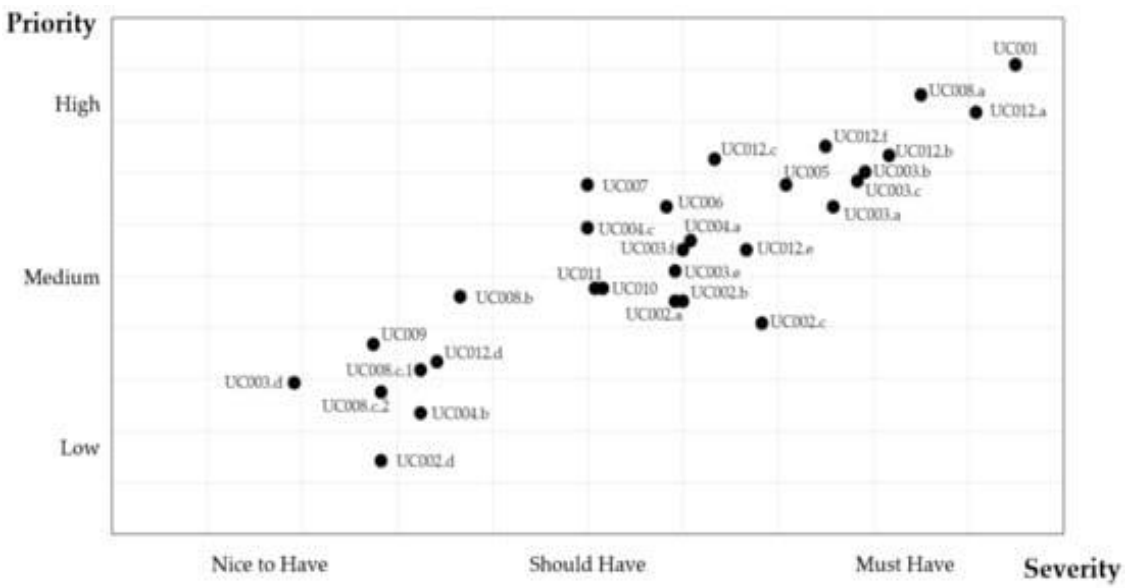


Figure 3. Use case prioritization regarding priority and severity.

Program 1:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv('case_time_series.csv')
Y = data.iloc[61:,1].values
R = data.iloc[61:,3].values
D = data.iloc[61:,5].values
X = data.iloc[61:,0]
plt.figure (fig size = ( 25, 8 ))
```

```
ax = plt.axes()
ax.grid (linewidth=0.4, color=
'#8f8f8f')
ax.set_facecolor("black")
ax.set_xlabel('\nDate'
,size=25,color=
'#4bb4f2')
ax.set_ylabel('Number of Confirmed Cases\n',
size=25,color=
'#4bb4f2')
plt.xticks(rotation=
'vertical'
, size=
'20
, color=
'white')
plt.yticks(size=20, color=
'white')
plt.tick_params(size=20,color=
'white')
for i, j in zip(X, Y):
```

```
ax.annotate (str (j), xy = (i, j+100), color=
'white',
size=
'13')ax.annotate('Second Lockdown 15th April'
, xy = (15.2,
860),
xy text = (19.9, 500),
color =
'white',
size =
'25',
arrow props = dict(color=
'white'
, linewidth=0.025))
plt.title("COVID-19 IN: Daily Confirmed\n",
size = 50, color =
'#28a9ff')
ax.plot(X, Y,color =
'#1F77B4',
marker =
'o',
```

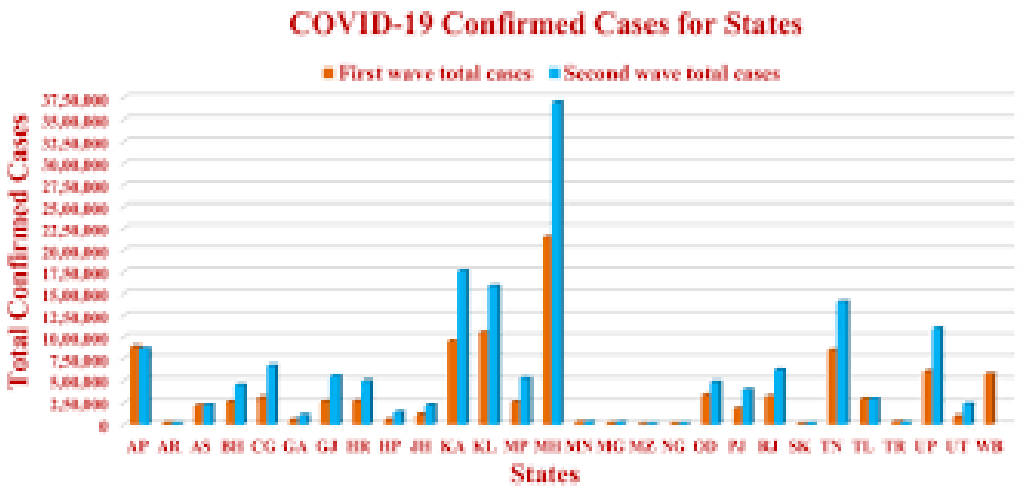
linewidth = 4,

markersize =15,

markeredgecolor=

'#035E9B')

OUTPUT:



PROGRAM:

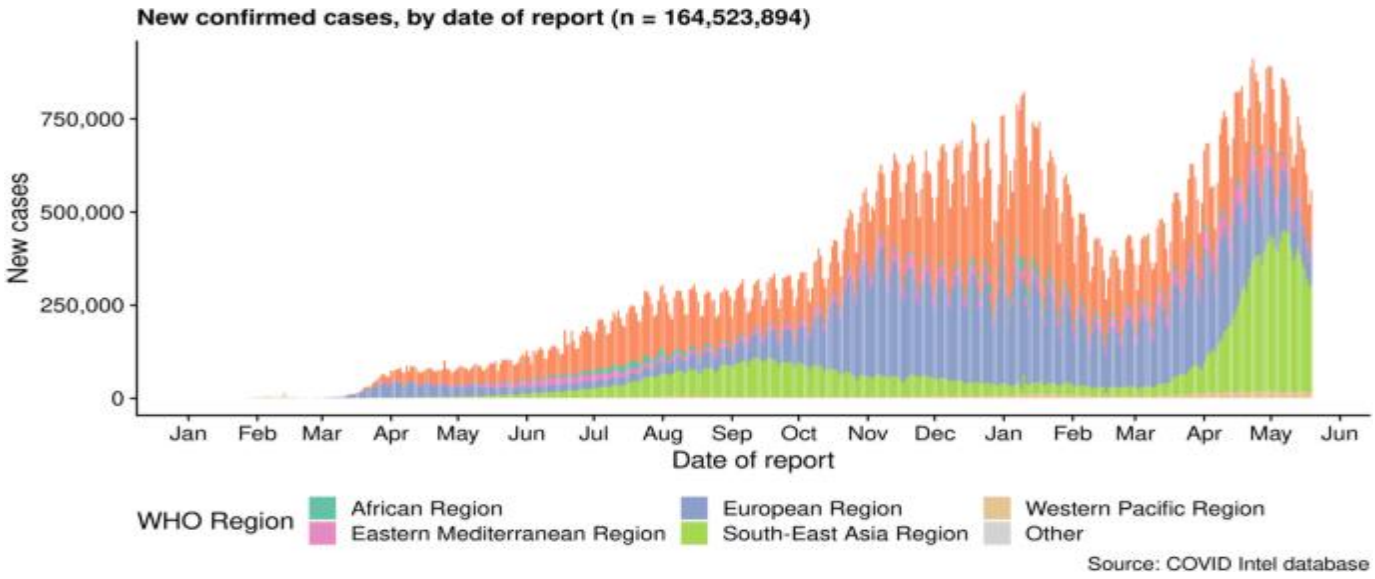
df_latest = df_latest.sort_values(by=[‘Confirmed’], ascending = False)

plt.figure(figsize=(12,4), dpi=80)

plt.bar(df_latest[‘States’][:10], df_latest[‘Confirmed’][:10],

```
align=

’center’

, color=

’blue’)

plt.ylabel(‘Number of Confirmed Cases’

, size = 12)

plt.title(“States with maximum confirmed cases till Aug’21”

, size =

16)

plt.show()
```

OUTPUT:



CONCLUSION:

The corona virus disease continues to spread across the world following a trajectory that is difficult to predict. The health, humanitarian and socio-economic policies adopted by countries will determine the speed and strength of the recovery. The ILO's four pillar policy framework presented in this brief provides guidance not only for countries as they progress through the various phases of the crisis, but also for the international community as a whole. There must be a global human-centred response which is grounded in solidarity.

International Standard Labour contain guidance for ensuring decent work that is applicable even in the

unparalleled context of the COVID 19 crisis. In particular, the _ that crisis responses need to “ensure respect for all human rights and the rule of law, including respect for fundamental principles and rights at work and for international labour standards”. The standards dealing with safety and health at work, social security, employment, non discrimination, working arrangements and the protection of specific categories of workers provide guidance on the design of rapid responses that can facilitate a stronger recovery from the crisis.

A coordinated global effort is required to support countries that currently do not have sufficient fiscal space to finance social policy, in particular universal social protection systems. Debt sustainability should be prioritized in this endeavour.

Without long-term structural changes, the deep-rooted inequalities exposed by the crisis will merely intensify. As well as tackling the immediate effects of the crisis, the international community now has a unique opportunity to adopt policies aimed at achieving social justice and a human centred future of work.