Impact of Covid Vaccination on the Globe using data analytics

[1]Arun Velu, [2]Pawan whig
Director Equifax USA,
Dean Research VIPS
* **arun.velu@equifax.com**

* corresponding author

ARTICLE INFO

Article History:
Received Nov 11, 2020
Revised Jan 31, 2021
Accepted April 30, 2021

Keywords:
World Economic Forum , covid , vaccination

Correspondence:
E-mail:
arun.velu@equifax.com

ABSTRACT

The new World Economic Forum-Ipsos vaccine trust study reveals that clear intent to have a vaccine for COVID-19 has improved in the United Kingdom and the United States, two countries where vaccines have begun to be delivered. In most other countries surveyed, however, vaccine trust has fallen. The primary reason people claim they won't get a vaccine for COVID-19 is that they fear side effects. In this research study, the impact of covid vaccination on the globe is studied using machine learning, and the trends found to be very interesting and presented in the result section.

Introduction

A wide family of distinct viruses are coronaviruses. Any of them trigger the common cold in persons. Others, including bats, camels, and horses, infect wildlife. But how did SARS-CoV-2, the new COVID-19-causing coronavirus, come into being? SARS-CoV-2 emerged in bats, researchers say. That's also how the Middle East respiratory syndrome (MERS) and extreme acute respiratory syndrome (SARS) coronaviruses began [1]. At one of Wuhan's open-air "wet markets," SARS-CoV-2 made the transition to humans. It's where consumers buy fresh meat and fish, even animals that are slaughtered on the spot. Any wet markets, including cobras, wild boars, and raccoon dogs, sell wild or banned animals [2]. Crowded environments can cause genes to be swapped by viruses from various species [3]. The virus often changes so much that it may begin to affect and propagate among individuals [4]. It affected people who had no close contact with wildlife, as SARS-CoV-2 spread both within and outside China. This meant that the virus was spread from one person to the next. In the U.S. and across the world, it is already spreading, suggesting people are inadvertently receiving and sharing the coronavirus. This rising dissemination worldwide is what is becoming a pandemic [5].
In 1965, scientists first observed a human coronavirus [6]. This produced a widespread cold. Researchers identified a group of related human and animal viruses later that decade and named them for their crown-like shape. The Origin of is how it can be transmitted is shown in Fig. 1

They will infect humans with seven coronaviruses [7]. The one that causes SARS established in 2002 in southern China and spread rapidly to 28 other countries. By July 2003, more than 8,000 people were sick, and 774 died. In 2004, a minor epidemic only affected four additional cases [8]. This

coronavirus causes problems with fever, headache, and coughing, including cough and shortness of breath.

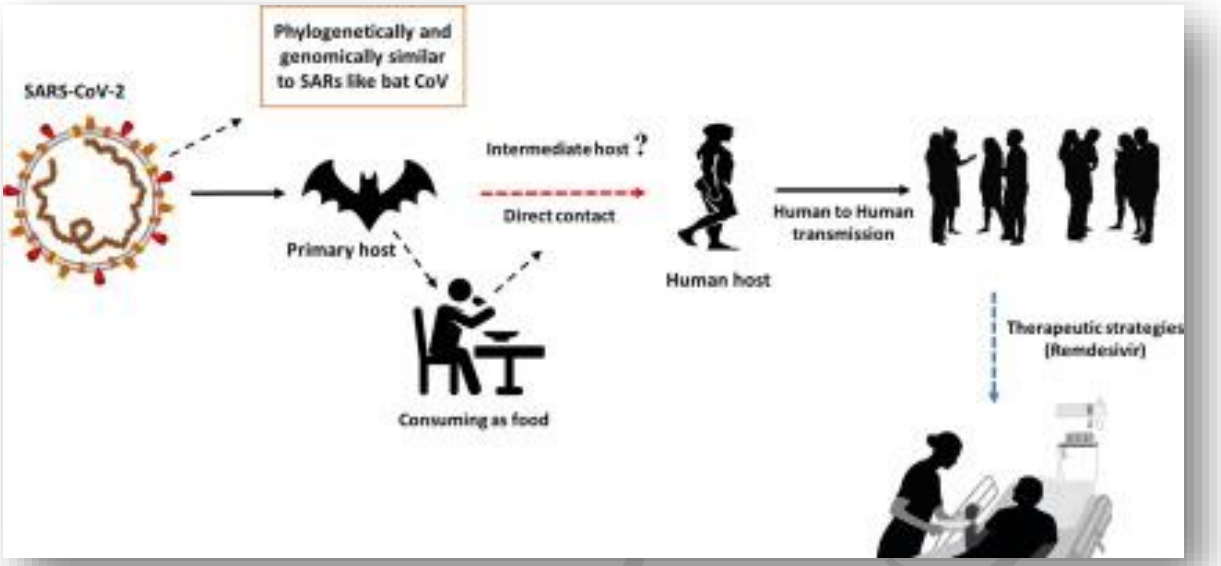

Figure 1 Proposed Origin of Corona Virus

As per date 31 jan 2021 the total number of covid cases were 101561219 with confirmed deaths of 2196944 with total 223 countries are infected as reported by who. The stats of covid is shown in the Fig.2.

Figure 2 Stats of Covid till Jan 2021

Data analysis using machine learning

Machine learning is a data processing tool that automates the creation of analytical models. It is a subset of artificial intelligence focused on the premise that, with minimal human interaction,

computers can learn from data, recognize trends and make decisions. Machine learning today is not like machine learning in the past, thanks to emerging computing developments [8-12]. It was born from the identification of patterns and the idea that computers would learn to do complex tasks without being programmed; academics interested in artificial intelligence decided to see how computers could learn from results[13-16]. The iterative nature of machine learning is important because they are able to adapt independently when models are introduced to new data [16-20]. To generate accurate, repeatable decisions and outcomes, they learn from previous computations. It is a science that is not new, but has obtained new traction. In this research studies we used the various machine learning libraries like Pandas ,NumPy etc for obtaining the impact of covid vaccination on the globe. The block diagram of the process involved is shown in Fig. 3.

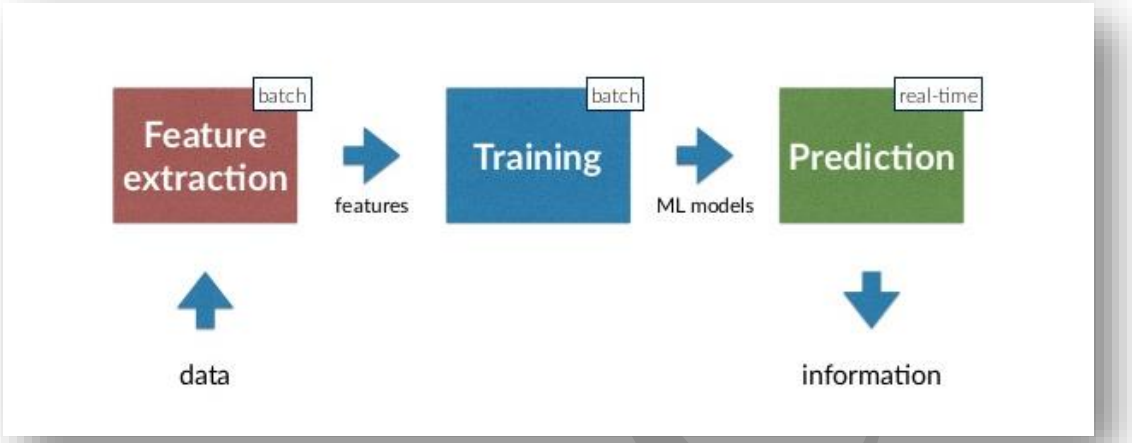

Figure 3 Block diagram of process involved

Result and Process Step Wise

For the data analysis google colab is used , "Colab" is a Google Analysis product, for short. Colab enables anybody through the browser to write and execute arbitrary python code, and is specifically well suited to computer learning, data processing and education. The detailed analysis is presented stepwise to understand the process in easy ways.

Step 1
Importing Packages

```
import pandas as pd
import numpy as np
import plotly.express as px
import plotly.graph_objects as go
from plotly.subplots import make_subplots
```

Step 2
Importing Dataset

```
data=pd.read_csv('country_vaccinations.csv')
data2=pd.read_csv('countries-aggregated.csv')
```

Step 3
Calulate the

```
data.dropna(subset=['daily_vaccinations'],inplace=True)
s=data['date'].str.split('-',expand=True)
data['Year']=s[0]
data['Month']=s[1]
data['Date']=s[2]
fig1=px.scatter_geo(data,color='vaccines',locationmode="ISO-
3",locations="iso_code",opacity=0.6,
```

```
                hover_name="iso_code", size="daily_vaccinations",projection='conic equal
area',animation_group ="iso_code",color_continuous_scale='blackbody',

animation_frame="Date",scope='world',symbol='vaccines',template="plotly_dark",title='Vaccinati
on Count ' )
fig1.layout.updatemenus[0].buttons[0].args[1]["frame"]["duration"] = 400

fig1.update_geos(
   landcolor="red",
   oceancolor="#006994",
   showocean=True,
   lakecolor="Blue"
)

fig1.update_traces(
   marker_coloraxis=None
)
fig1.show()
```

Number of vaccination available in the world the result obtained is shown in Fig 4

Figure 4 Various Vaccination count available roud the world

It is easily seen that the influence of Pfizer/Biotech in the USA is immense and it tends to grow by the date rises. We can also see CNBG/strong Cinovac's influence beginning in China. Sinovac takes a sudden roll on the 14th in Turkey and starts to develop afterwards.

Step 4

Analysis of Vaccine distributed the most

```
s=data.drop_duplicates(subset=['iso_code'])['vaccines'].apply(lambda x: x.split(','))
dic={}
for i in s:
   for j in i :
      if j[0]==' ':
         k= j[1:]
      elif j[-1]==' ':
         k=j[:-1]
      else:
         k=j
      if k not in dic :
         dic[k]=1
      else:
         dic[k]+=1
px.bar(x=list(dic.keys()),y=list(dic.values()),color=list(dic.keys()),template='plotly_white',labels={'
x':'Vaccine Name','y':'Total Count'})
```

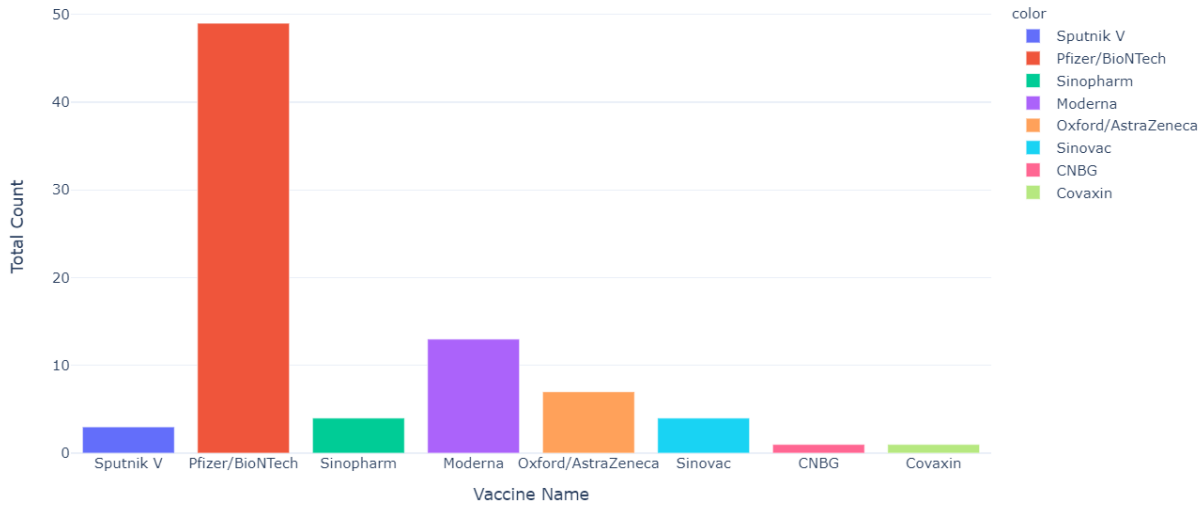

Figure 5 Result obtain in Step 4

It is clearly observed that Pfizer in January 2021, followed by Moderna and Sinovacac, became the most common vaccine in the world.

Step 5
Observe the deathrate after vaccine

```
arr=[]
index1=data.groupby(['country','date']).count().index
index2=data2.groupby(['Country','Date']).count().index
for i in index2:
    if i in index1:
        arr.append(1)
    else:
        arr.append(0)
data2['Vaccine_is_there']=arr
data2.head()
```

Out[14]:

	Confirmed	Recovered	Deaths	Vaccine_is_there
count	7.180800e+04	7.180800e+04	71808.000000	71808.000000
mean	1.440877e+05	8.692296e+04	3921.234305	0.020221
std	8.780563e+05	5.053113e+05	19220.500408	0.140755
min	0.000000e+00	0.000000e+00	0.000000	0.000000
25%	6.800000e+01	1.600000e+01	0.000000	0.000000
50%	2.645500e+03	1.290000e+03	53.000000	0.000000
75%	3.695150e+04	1.803075e+04	638.000000	0.000000
max	2.592928e+07	1.040915e+07	436678.000000	1.000000

Figure 6 Result obtain in step 5

Figure 7 Results of different countries showing effect on deathrate after vaccine
After a line indicated in the figures the vaccine was introduce in the country. The Country names are indicated on the figure for better understanding of trends.

Conclusion

"Vaccination intent is highest in China, according to the survey, where 80 percent of respondents strongly or somewhat agreed with the statement "If a COVID-19 vaccine were available, I would get it. Countries with a reasonably high degree of intent include Brazil (78%), the United Kingdom (77%), Mexico (77%), Australia (75%) and South Korea (75 percent). Among the countries surveyed, South Africa (53 percent), Russia (43 percent) and France were those whose populations recorded the lowest intentions (40 percent). It is clearly observe from the Fig 7 that the deathrate either stagnant or go low in almost all countries but it is still sever in US and UK where a new mutant of Covid 19 has been found.

Future Scope

This research study will be very helpful for the researchers working in the same field

[1] Singh, R., Singh, R., Bhatia, A. (2018). Sentiment analysis using Machine Learning technique to predict outbreaks and epidemics. Int.J.Adv.Sci.Res,3(2),19-24

[2] Ting, Daniel Shu Wei, Lawrence Carin, Victor Dzau, and Tien Y. Wong. "Digital technology and COVID-19." Nature Medicine (2020): 1-3.

[3] Benvenuto, Domenico, Marta Giovanetti, Lazzaro Vassallo, Silvia Angeletti, and Massimo Ciccozzi."Application the ARIMA model on the COVID-2019 epidemic dataset." Data in brief (2020):105340.

[4] Deb, Soudeep, and Manidipa Majumdar. "A time series method to analyze incidence pattern and estimate reproduction number of COVID-19." arXiv preprint arXiv:2003.10655(2020).

[5] Kucharski, Adam J., Timothy W. Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, Rosalind M. Eggo et al. "Early dynamics of transmission and control of COVID-19: a mathematical modelling study." The lancet infectious diseases(2020).

[6] Dey, Samrat Kumar, Md Mahbubur Rahman, Umme Raihan Siddiqi, and Arpita Howlader. "Analyzing the Epidemiological Outbreak of COVID-19: A Visual Exploratory Data Analysis(EDA)Approach. Journal of Medical Virology(2020).

[7] R. Meredith, Andrew S. Azman, Nicholas G. Reich, and Justin Lessler. " The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application." Annals of internal medicine (2020).

[8] Singer, H. M. "Short-term predictions of country-specific Covid-19 infection rates based on power law scaling exponents."arXivpreprintarXiv:2003.11997(2020).

[9] Awad, Mariette, and Rahul Khanna. "Support vector regression." In Efficient Learning Machines, pp. 67-80. Apress, Berkeley, CA,2015.

[10] De Castro, Yohann, Fabrice Gamboa, Didier Henrion, Roxana Hess, and Jean-Bernard Lasserre."Approximate optimal designs for multivariate polynomial regression."The Annals of Statistics 47, no. 1 (2019):127-155.

[11] Greff,Klaus,RupeshK.Srivastava,JanKoutn´ık,BasR.Steunebrink,andJu¨rgen Schmidhuber. "LSTM: A search space odyssey." IEEE transactions on neural networks and learning systems 28, no. 10 (2016):2222-2232.

[12] U. Ojha and S. Goel, "A study on prediction of breast cancer recurrence using data mining techniques," *Proc. 7th Int. Conf. Conflu. 2017 Cloud Comput. Data Sci. Eng.*, pp. 527–530, 2017.

[13] S. Ghosh, S. Mondal, and B. Ghosh, "A comparative study of breast cancer detection based on SVM and MLP BPN classifier," *1st Int. Conf. Autom. Control. Energy Syst. - 2014, ACES 2014*, pp. 1–4, 2014.

[14] A. Osareh and B. Shadgar, "Machine learning techniques to diagnose breast cancer," *2010 5th Int. Symp. Heal. Informatics Bioinformatics, HIBIT 2010*, pp. 114–120, 2010.

[15] D. Bazazeh and R. Shubair, "Comparative study of machine learning algorithms for breast cancer detection and diagnosis," *Int. Conf. Electron. Devices, Syst. Appl.*, pp. 2–5, 2017.

[16] M. S. B. M. Azmi and Z. C. Cob, "Breast cancer prediction based on backpropagation algorithm," *Proceeding, 2010 IEEE Student Conf. Res. Dev. - Eng. Innov. Beyond, SCOReD 2010*, no. SCOReD, pp. 164–168, 2010.

[17] B. M. Gayathri and C. P. Sumathi, "Comparative study of relevance vector machine with various machine learning techniques used for detecting breast cancer," *2016 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2016*, pp. 0–4, 2017.

[18] Purva Agarwal and Pawan Whig, "Low Delay Based 4 Bit QSD Adder/Subtraction Number System by Reversible Logic Gate", 2016 8th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE Xplore: 26 October 2017.

[19] Jacob B. Chacko ; Pawan Whig, " Low Delay Based Full Adder/Subtractor by MIG and COG Reversible Logic Gate", 2016 8th International Conference on Computational Intelligence and Communication Networks (CICN), IEEE Xplore: 26 October 2017

[20] Ajay Rupani , Pawan Whig , Gajendra Sujediya and Piyush Vyas ,"A robust technique for image processing based on interfacing of Raspberry-Pi and FPGA using IoT," International Conference on Computer, Communications and Electronics (Comptelix), IEEE Xplore: 18 August 2017

References

[1] Robert Driskill and John Riedl, "Recommender Systems for E-Commerce: Challenges and Opportunities", AAAI Technical Report WS-99-01, 2011, pp. 73-76

[2] Mr.Avadhut D.Wagavkar et. al., "Weighted Hybrid Approach in Recommendation Method", (IJCST) – Volume 5 Issue 2, Mar – Apr 2017, pp. 346-350

[3] Prof. Mehul Barot et. al., "Fuzzy Logic Based Recommender System", (IJRSI) Volume IV, Issue VIS, June 2017 | ISSN 2321–2705, pp. 71-80

[4] Rupali Hande et. al., "MOVIEMENDER- A Movie Recommender System", International Journal Of Engineering Sciences & Researchtechnology ISSN : 2277-9655, pp. 469-73

[5] Pawan Whig and S. N. Ahmad, A CMOS Integrated CC-ISFET Device for Water Quality Monitoring, International Journal of Computer Science Issues ,Volume 9, Issue 4, July 2012, ISSN (online): 1694-0814 pp: 365-371.

[6] Pawan Whig and S. N. Ahmad, Performance Analysis of Various Readout Circuits for Monitoring Quality of Water Using Analog Integrated Circuits, International Journal of Intelligent Systems and Applications (IJISA) ISSN: 2074-904X (Print), ISSN: 2074-9058 (Online) Volume 4, No.11, October 2012 pp:91-98.

[7] Pawan Whig and S. N. Ahmad, A Novel Pseudo PMOS Integrated CC-ISFET device for water quality monitoring, Journal of integrated circuit and system published 2013 Volume 8, No.2, October 2013 pp:1-6. ISSN, 1807-1953 (Scopus).

[8] Pawan Whig and S. N. Ahmad, "Simulation of Linear Dynamic Macro Model of Photo Catalytic Sensor in SPICE" Compel, the international journal of computation and mathematics in electrical and electronic engineering, Vol. 33 No. 1/2, 2014. ISSN: 0332-1649 (SCI, ISI index)

[9] Vaibhav Bhatia and Pawan Whig" A secured dual tune multi frequency based smart elevator control system," International journal of research in engineering and advanced technology",Vol. 4 Issue 1 , 2013. ISSN (Online): 2319-1163

[10] Pawan Whig and S. N. Ahmad, A Novel Pseudo NMOS Integrated ISFET device for water quality monitoring, Active and Passive Components Hindawi article i.d 258970. Vol. 1 Issue 1, 2013(Scopus). ISSN 0882-7516

[11] Vaibhav Bhatia and Pawan Whig, "Modeling and Simulation of Electrical Load Control System Using RF Technology, International Journal of multidisplinary science and engineering",2013,Vol. 4 No.2, pp 44-47 ISSN 2045-7057.