Epidemic Analysis and Prediction of Novel Coronavirus based on XGBoost Algorithm

Jiaye Xie

Michigan State University, East Lansing, MI 48825, USA

Abstract. The sudden outbreak of COVID-19 poses a great threat to the health and safety of people all over the world. Since the outbreak of the epidemic, the number of suspected and confirmed infections in the world has continued to rise. During the period of epidemic prevention and control, it is very important to carry out scientific research quickly, such as finding the source of the virus, controlling the spread of the virus, studying the pathogenic mechanism, collecting data, looking for scientific treatment and prevention and control programs, screening and developing effective drugs, and so on. At the same time, according to the data collected and published by the World Health Organization, it can reflect the response measures in different regions and the number of people found to be affected by COVID. The epidemic situation in COVID-19 is a critical period for major scientific research projects to tackle key problems, and an important stage to win time and save lives. Today, Covid is still rampant, and there are still different varieties. Based on the existing data, the COVID infection model can be calculated to predict the infectivity of new varieties, and different measures can be taken for countries with high mortality or infection rates. The experimental results show that, compared with three traditional machine learning algorithms, the model built based on the integrated algorithm XGBoost has the highest prediction accuracy rate for whether human beings are infected with novel coronavirus, and the accuracy rate reaches 95.6%, providing a non-destructive auxiliary method for biomedical detection of COVID-19.

Keywords: Covid-19; XG-Boost Algorithm; Evaluation of Epidemic Prevention Measures.

1. Introduction

In 2019, the pneumonia epidemic in novel coronavirus broke out all over the world and quickly spread around the world, which made coronavirus known to the world again and had a serious impact on the world [1]. Coronavirus is a sense RNA virus with envelope. Coronavirus consists of spike protein, lipid membrane, envelope, nucleoprotein shell and sense single stranded RNA. Its characteristics are stick-like spikes protruding from its surface, huge RNA genome and its unique replication strategy [2]. Novel coronavirus is also known as COVID-19. This disease spreads rapidly all over the world. Patients infected with COVID-19 have many symptoms, but they usually include fever, cough, fatigue, dyspnea and loss of smell and taste. The virus that causes COVID-19 mainly spreads when infected people are in close contact with others. Small droplets and aerosols containing viruses can spread in the air. If the virus enters people's mouth, nose or eyes, others will be infected [3]. The virus may also spread through contaminated surfaces, although this is not considered to be the main route of transmission. The exact route of transmission is rarely proved, but the infection mainly occurs when people are close to each other for long enough. Infected people can transmit the virus to others two days before they appear symptoms, and people without symptoms can also [4].

After the full outbreak of the epidemic, the whole country invested a lot of human and financial resources to fight the epidemic. Since the treatment of COVID-19 is mainly based on supportive treatment and mainly depends on the patient's own immunity, it is extremely important to judge the effect of current epidemic prevention policies, the ability of virus infection, and predict the trend of the epidemic in the fight against COVID-19 [5]. It is of great significance to quantitatively predict the trend of the epidemic through modeling for implementing corresponding means to fight against the epidemic and actively resume production. The prevention and control of infectious diseases has always been a major issue of common concern to all people in the world. Generally speaking, infectious diseases can be studied from many aspects, such as their transmission rules, pathogenesis of patients, transmission tendency of infectious sources, etc. [6]. For biological problems that can be practically studied, they can be carried out in laboratories and other areas, but we can't carry out experiments on real people to repeat the whole process of infectious diseases, so it is particularly important to carry out theoretical derivation research such as mathematical modeling [7]. Mathematical modeling is to use mathematical theory to summarize objective problems, grasp the main contradiction and ignore the secondary contradiction. Find out the pathogenesis of infectious diseases and the transmission mechanism in the population, and establish an appropriate model to restore the transmission law of infectious diseases [8].

Therefore, a rapid identification method of COVID-19 infection by Raman spectroscopy combined with XG-Boost is proposed. XGBoost ensemble learning algorithm is used to model the classification and prediction of Covid-19 infected and healthy people. Finally, compared with the classification results of K nearest neighbor, support vector machine (SVM) and Logistic Regression, it is concluded that Raman spectroscopy combined with XGBoost ensemble learning algorithm can accurately classify and identify the sample data of Covid-19 infected and Covid-19 uninfected, and it also provides an auxiliary nondestructive testing method for biomedical detection of other diseases.

2. Theoretical Basis and Methods

2.1 Overall Software Architecture Design

Nucleic acid detection software contains many complex and trivial functions. The previous generation of software has not been carefully designed, the realization of various functions has not been fully modularized, and the coupling between functions is too strong. As a result, it is very difficult to modify and reconstruct the program in the process of demand changes, and subsequent maintenance cannot be carried out, so it has to be rewritten based on new needs. Therefore, maintainability becomes the primary design requirement of new software [9]. In order to increase the adjustability of modules and reduce coupling, the program uses dependency injection to turn the direction of dependencies into abstract rather than concrete implementation. Based on this method, although the process of program execution at runtime remains unchanged, the introduction of interfaces means that different implementations of these interfaces can be easily inserted, and the implementation of modules is more independent [10].

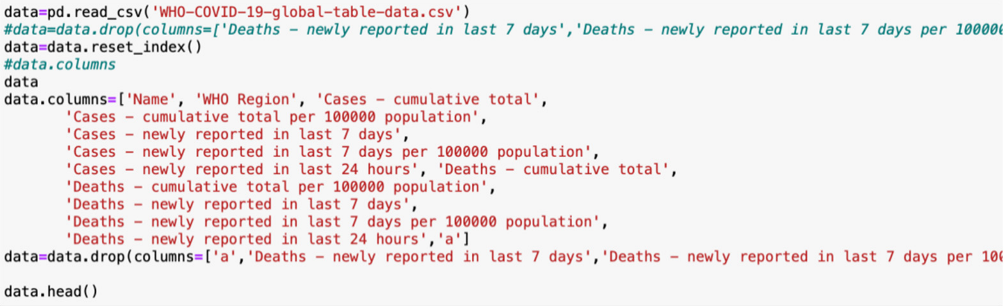
Service is the foundation of software dependency injection, and it is a part of the core library. Various components in the program create access through the service host. In the initialization stage, the program registers the used service objects through the host builder, configures plug-ins and settings, and creates the service host through the construction method. The provider is located at the lowest level of the setting module, and provides services such as saving and loading different setting sources. Each setting key in the settings will be bound to a setting provider for, so as to read and write the stored contents of the settings. Item is the direct operation layer of setting, including setting root and setting section, and the corresponding setting value can be read or modified by setting key. Among them, the setting section is only mapped to part of the setting keys, which separates the concerns, which is beneficial to avoid misoperation.

Since the description of patients contacting vulnerable populations in classical infectious disease models is mostly simply set as bilinear contact rate (proportional to the number of vulnerable populations in the environment and the number of patients) or standard contact rate (patients contacting vulnerable populations in a fixed ratio), this setting is more appropriate for closed systems with small populations, But no matter in the system with large population or the system with artificial information intervention, this setting obviously does not conform to the actual situation. So on this basis, for the system with large population, the epidemic model with saturated contact rate and the epidemic model with non-monotonic incidence rate are extended. The above are all continuous models based on differential equations. However, in practice, both the number of patients with disease and the actual release of epidemic data are often discrete, so it is more convenient to use discrete infectious disease models to model, and it is more convenient to identify parameters. There are two main methods to construct the discrete model. One is to directly establish the discrete model on the basis of the continuous model assumption; The other is to obtain the discrete model by using the Euler method to discrete the equations of the continuous model.

2.2 Theoretical Basis of Hydrophobic Action

Due to the strong polarity of water molecules, hydrogen bond competition always occurs with water molecules in extracellular medium. Before protein docking, water molecules should be able to saturate the hydrogen bond formation on the hydrophilic surface of protein. The interaction between the surface of protein and the surrounding water is usually called protein hydration shell (sometimes also called hydration layer), which is the foundation of the structural stability of protein, because non-aqueous solvents usually denature protein. The change caused by the exposure of hydrophilic or hydrophobic groups on the surface of protein will affect the collective vibration of water molecules in hydration layer. Therefore, the hydrophilic side chain of protein usually bonds with surrounding water molecules in water environment, thus preventing the hydrophilic side chain of protein from randomly bonding together. This is the reason why protein usually does not aggregate and crystallize in unsaturated aqueous solution. The change of protein and ligand binding to hydration layer can be generally divided into the following six processes: (1) the contact formation of protein and ligand; (2) water rearranges, initially dissolving protein; (3) rearranging the water to dissolve the ligand at first; (4) forming a hydrated structure around the protein-ligand complex; (5) the change of protein conformation; (6) Changes of ligand conformation. The hydrophobic groups, polar functional groups or shape matching of the binding region of protein only consider the interaction of protein interface (that is, the way of thinking of functional sites is used to consider the binding of protein). When the hydrophobic interaction occurs, that is, the water molecules around the hydrophobic substances will make the hydrophobic substances close to each other and concentrate, which also greatly reduces the influence of hydrophobic groups on the hydrogen bond system between water molecules or between water molecules and hydrophilic substances.

According to COVID data and vaccination data from WHO, Pandas, matplotlib and Python based numpy software package are used for multiple visualization and mapping to predict the new variation transmission rate of COVID. First, I download the data from the official WHO website, then use the panda’s package to extract the data and rename the columns to fit the data. Deleted columns. As shown in Figure 1.

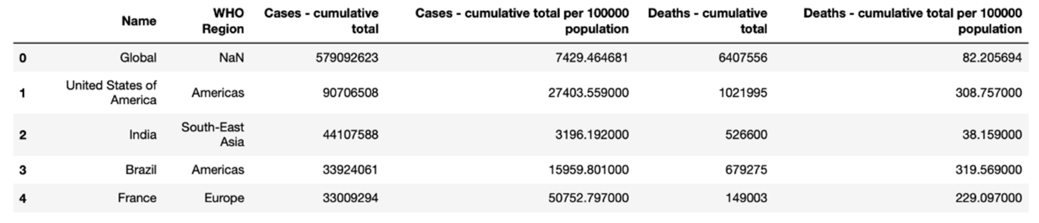


```
data=pd.read_csv('WHO-COVID-19-global-table-data.csv')
#data=data.drop(columns=['Deaths - newly reported in last 7 days','Deaths - newly reported in last 7 days per 10000
data=data.reset_index()
#data.columns
data
data.columns=['Name', 'WHO Region', 'Cases - cumulative total',
       'Cases - cumulative total per 100000 population',
       'Cases - newly reported in last 7 days',
       'Cases - newly reported in last 7 days per 100000 population',
       'Cases - newly reported in last 24 hours', 'Deaths - cumulative total',
       'Deaths - cumulative total per 100000 population',
       'Deaths - newly reported in last 7 days',
       'Deaths - newly reported in last 7 days per 100000 population',
       'Deaths - newly reported in last 24 hours','a']
data=data.drop(columns=['a','Deaths - newly reported in last 7 days','Deaths - newly reported in last 7 days per 10

data.head()
```

Figure 1. Use the panda’s package to extract the data

My first dataset, as shown in Figure 2.



	Name	WHO Region	Cases - cumulative total	Cases - cumulative total per 100000 population	Deaths - cumulative total	Deaths - cumulative total per 100000 population
0	Global	NaN	579092623	7429.464681	6407556	82.205694
1	United States of America	Americas	90706508	27403.559000	1021995	308.757000
2	India	South-East Asia	44107588	3196.192000	526600	38.159000
3	Brazil	Americas	33924061	15959.801000	679275	319.569000
4	France	Europe	33009294	50752.797000	149003	229.097000

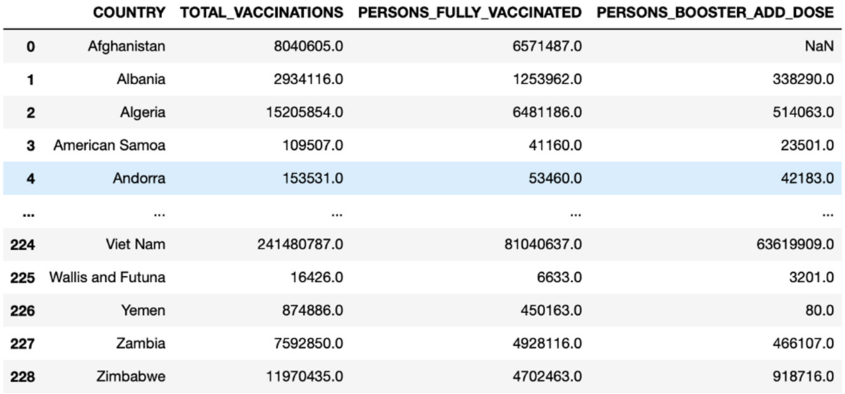
Figure 2. My first dataset

Separate data sets by different regions, as shown in Figure 3.

```
amdf=data.loc[data['WHO Region'].isin(['Americas'])]
seadf=data.loc[data['WHO Region'].isin(['South-East Asia'])]
edf=data.loc[data['WHO Region'].isin(['Europe'])]
wpdf=data.loc[data['WHO Region'].isin(['Western Pacific'])]
emdf=data.loc[data['WHO Region'].isin(['Eastern Mediterranean'])]
afdf=data.loc[data['WHO Region'].isin(['Africa'])]
```

Figure 3. Separate data sets by different regions

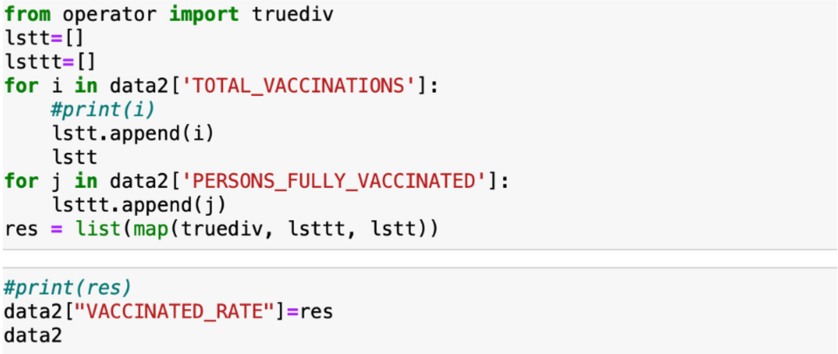
My second dataset:
The vaccination data set is shown in Figure 4.



	COUNTRY	TOTAL_VACCINATIONS	PERSONS_FULLY_VACCINATED	PERSONS_BOOSTER_ADD_DOSE
0	Afghanistan	8040605.0	6571487.0	NaN
1	Albania	2934116.0	1253962.0	338290.0
2	Algeria	15205854.0	6481186.0	514063.0
3	American Samoa	109507.0	41160.0	23501.0
4	Andorra	153531.0	53460.0	42183.0
...	...	...	...	...
224	Viet Nam	241480787.0	81040637.0	63619909.0
225	Wallis and Futuna	16426.0	6633.0	3201.0
226	Yemen	874886.0	450163.0	80.0
227	Zambia	7592850.0	4928116.0	466107.0
228	Zimbabwe	11970435.0	4702463.0	918716.0

Figure 4. The vaccination data set

Add the vaccination rate column to the data frame, as shown in Figure 5.

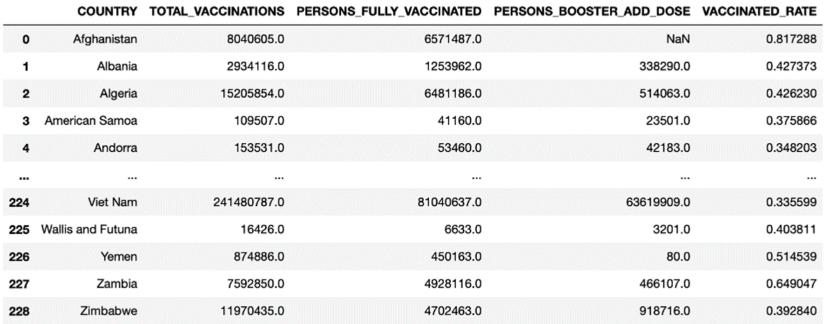


```
from operator import truediv
lstt=[]
lsttt=[]
for i in data2['TOTAL_VACCINATIONS']:
    #print(i)
    lstt.append(i)
    lstt
for j in data2['PERSONS_FULLY_VACCINATED']:
    lsttt.append(j)
res = list(map(truediv, lsttt, lstt))
```

```
#print(res)
data2["VACCINATED_RATE"]=res
data2
```

Figure 5. Add the vaccination rate column to the data frame

The new dataset is shown in Figure 6.



	COUNTRY	TOTAL_VACCINATIONS	PERSONS_FULLY_VACCINATED	PERSONS_BOOSTER_ADD_DOSE	VACCINATED_RATE
0	Afghanistan	8040605.0	6571487.0	NaN	0.817288
1	Albania	2934116.0	1253962.0	338290.0	0.427373
2	Algeria	15205854.0	6481186.0	514063.0	0.426230
3	American Samoa	109507.0	41160.0	23501.0	0.375866
4	Andorra	153531.0	53460.0	42183.0	0.348203
...	...	...	...	...	...
224	Viet Nam	241480787.0	81040637.0	63619909.0	0.335599
225	Wallis and Futuna	16426.0	6633.0	3201.0	0.403811
226	Yemen	874886.0	450163.0	80.0	0.514539
227	Zambia	7592850.0	4928116.0	466107.0	0.649047
228	Zimbabwe	11970435.0	4702463.0	918716.0	0.392840

Figure 6. The new dataset

3. Empirical Analysis of covid-19

3.1 Principle and Implementation of XGBoost Algorithm

XGBoost integration algorithm mainly uses tree structure to continuously split features to grow into a new tree. That is, XGBoost continuously generates new decision trees A, B, C, D... The algorithm of the final decision tree is the decision tree of the sum of A+B+C+D+. Each time a function is added to fit the error predicted by the previous layer. To prevent over fitting, XGBoost introduces the L2 regular term to smooth the predicted value of the leaf node. The objective function consists of loss function and regularization term. Its objective function is:

$$L(\varphi) = \sum_{i} l(\hat{y}_i, y_i) + \sum_{k} \Omega(f_k) \tag{1}$$

The regularization items are:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|\omega\|^2 \tag{2}$$

Where T represents the number of leaf nodes, and ω represents the score of leaf nodes. The regularization term represents the function of the complexity of the tree. The smaller the value, the lower the complexity and the stronger the generalization ability. XGBoost allows cross-validation to be used in each boosting iteration. Therefore, the optimal boosting iteration times can be conveniently obtained.

Classify the unlabeled data samples one by one. First, calculate the average value of network metrics formed by inserting all the data in X_{items} into each network, which has been stored in the two-dimensional array in the first step; Then, a new data sample from X_{test} is inserted into the two networks that have been built so far. Then, the same metric $G_{after}(class), i = 1,2$ of each network after insertion is calculated. Finally, the influence of inserting new samples into each class is ΔG by Formula 3:

$$\Delta G(class_t) = \left\| G_{after}(class_i) - G_{before}(class_i) \right\|, i = 1,2 \tag{3}$$

Finally, this new sample data is classified into class j, where:

$$G(class_i) = \min\{\Delta G(class_i)\}, i = 1,2 \tag{4}$$

In other words, the new sample conforms to the pattern formed by the network j, and will not cause great interference to the network j. Note that the new sample can even be far away from the elements of the class. At this point, classification is complete.

The box number method is used to calculate the fractal dimension. The method is applied to binary image. Therefore, given a human chest X-ray image, first use some thresholds to generate several binary images. For each binary image, use a grid to cover the image, and then calculate how many boxes in the grid cover the pattern in the image. Then use the small box to repeat the process. By repeatedly reducing the size of the box, you can accurately capture the structure of the pattern. In a word, XGBoost model is easier to understand and many XGBoost models that are difficult to solve often need to be converted to XGBoost model for solution. XGBoost model is easier to conduct qualitative analysis, whereas discrete model has more complex dynamic characteristics. In practical application, the discrete model has better effect and is easy to solve, but when analyzing the properties of the system, the continuous form is more convenient to study.

3.2 Resumption and Analysis of Epidemic Situation

We can see that both random XGBoost model and deterministic XGBoost model have played a good role in fitting the epidemic data in Wuhan, especially after the 19th. But there is a problem here. The number of people who have been removed, that is, those who have been infected, is a part more than the official statistics. However, in the actual fitting process, the number of infected people is basically in line with objective changes. After January 23rd, the difference between the fitted value and the actual value can be understood as the gap between the medical testing resources and the real situation at this time, and this gap is gradually smoothed out with the establishment of the shelter hospital and the improvement of testing instruments and equipment. It is reflected in the data that the error between statistical data and fitted data is getting smaller and smaller after February 18th. Therefore, according to our model, there are some asymptomatic infected people who have not been detected in Wuhan, and some infected people who have not been to the hospital for systematic treatment in the early stage of self-healing, which constitute a lack of statistical data in Wuhan. The XGBoost classifier model constructed from the infected and uninfected COVID-19 data selected in this experiment is evaluated. For this two-classification model, ROC curve is selected for evaluation. Taking KNN classifier and XG Boost classifier model as examples, the AUC area in their ROC curve shows that the classification model constructed by XGBoost algorithm selected in this experiment is closer to the ideal effect. Its ROC curve is shown in Figure 7 and Figure 8.

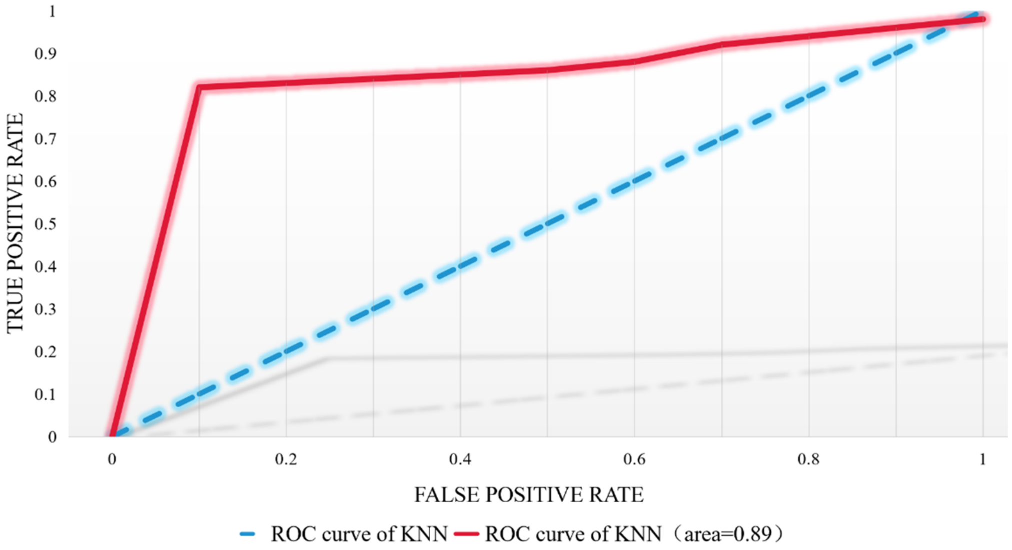


Figure 7. ROC curve of KNN algorithm

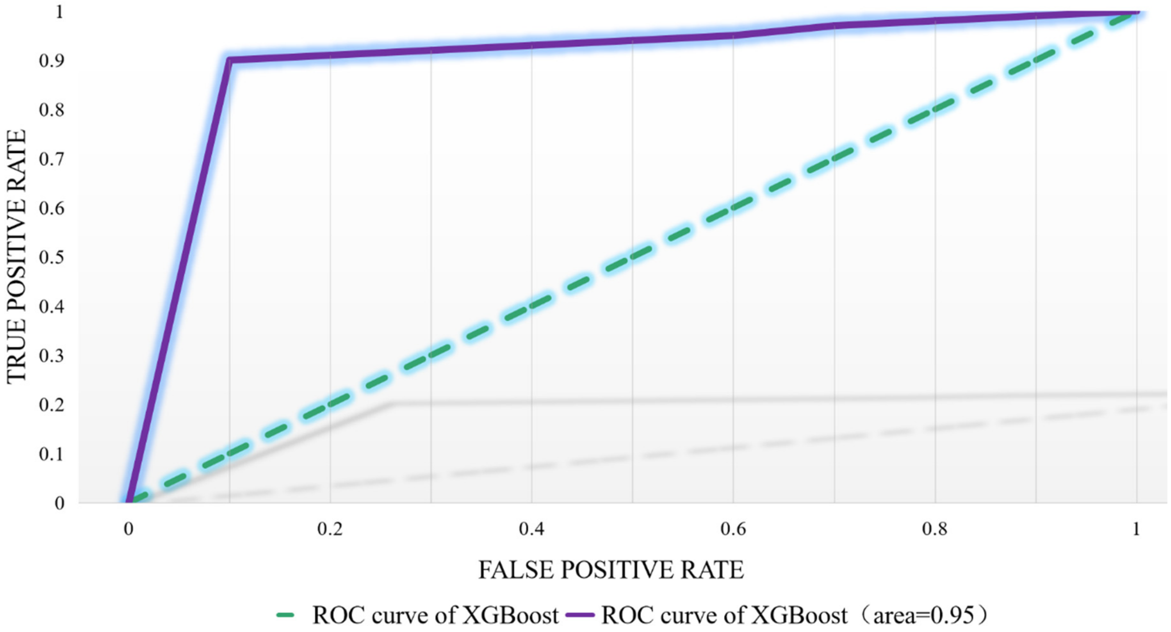


Figure 8. ROC Curve of XGBoost Algorithm

To evaluate the classification performance of XGBoost model, this experiment finally selects K nearest neighbor, support vector machine, logistic regression model and XGBoost model for comparison. The comparison results are shown in Table 1.

Table 1. Comparison of diagnostic results

Classification model	Accuracy rate
LDA+KNN	89.8
LDA+SVM	85.3
LDA+XGBoost	95.6
LDA+Logistic Regression	87.1

It can be seen from Table 1 that this experiment selects two types of spectral data of COVID-19 infected people and healthy people, and uses several classification models to predict the corresponding results. Compared with the traditional algorithm, XGBoost splits the decision tree for many times, iteratively generates the base learner, adds and updates the base learner, and finally gets a powerful learner. On the basis of the traditional boost integration algorithm, XGBoost uses CPU multithreading to introduce regularization terms into its objective function to prevent data over fitting. In addition, pruning can effectively control the complexity of the model and achieve better classification results.

Use the European mobility index and COVID-19 infection growth rate data to investigate the relationship between population mobility and the spread of new coronavirus. The mobility index data are the European driving travel index, public transport travel index and walking travel index from January 22, 2020 to August 16, 2020 (Figure 1), which are all from Apple's website in the United States (https://www.apple.com/covid-19/mobility). This data set is a relative number based on January 13, 2020 (100). Among them, the data of population mobility from May 11 to May 12, 2020 is missing, and this paper uses linear interpolation to make up. As shown in Figure 9.

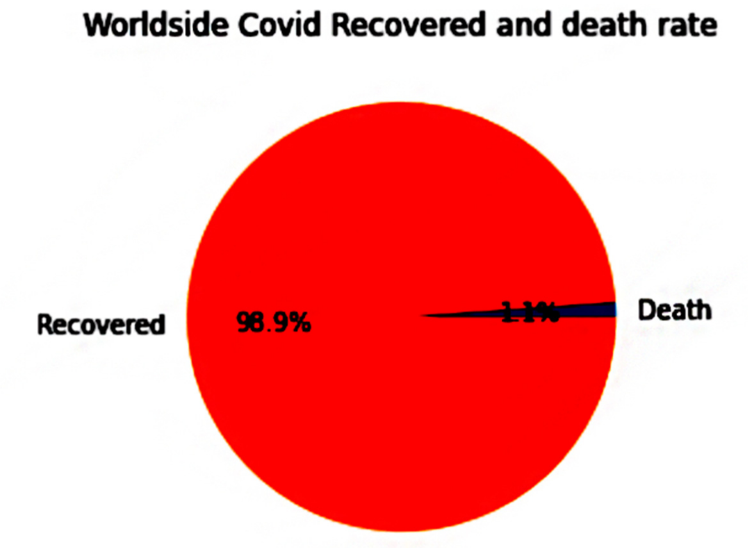


Figure 9. Covid Recovery and Rating Pie Chart

According to my data set, most worldside Covid has been recovered, but the mortality rate is only 1.1%. As shown in Figure 10.

As I can see, Europe has the largest number of Covid patients in the world, accounting for 41.9% before the latest updated data, while Africa has the smallest number, only 1.6%.

But one interesting thing is that Europe has the largest number of patients, but the mortality rate is lower than the average. The western Pacific region has the lowest mortality rate, only 0.3%. Africa has the lowest number of patients, but the mortality rate is the highest. As shown in Figures 11 and 12.

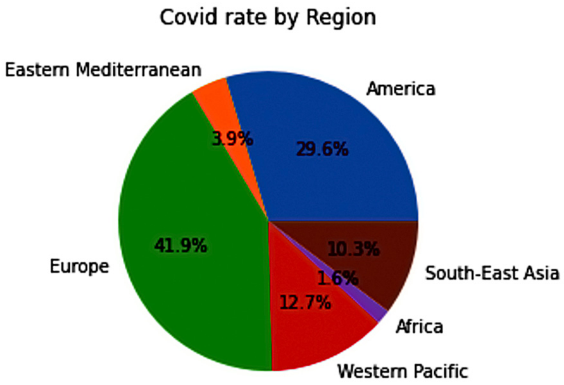


Figure 10. Comparison of COVID-19 infection rates by region

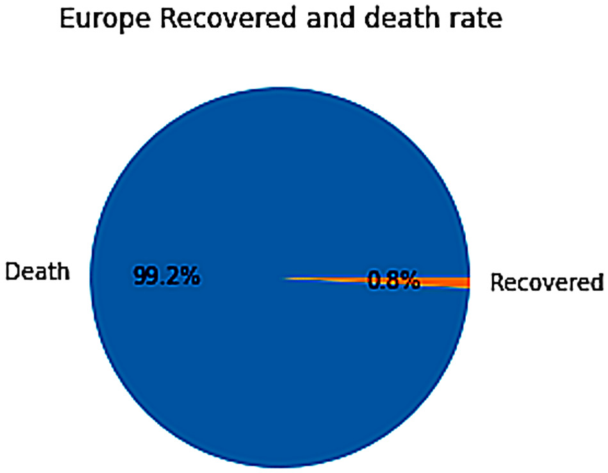


Figure 11. Comparison of recovery and mortality in Europe

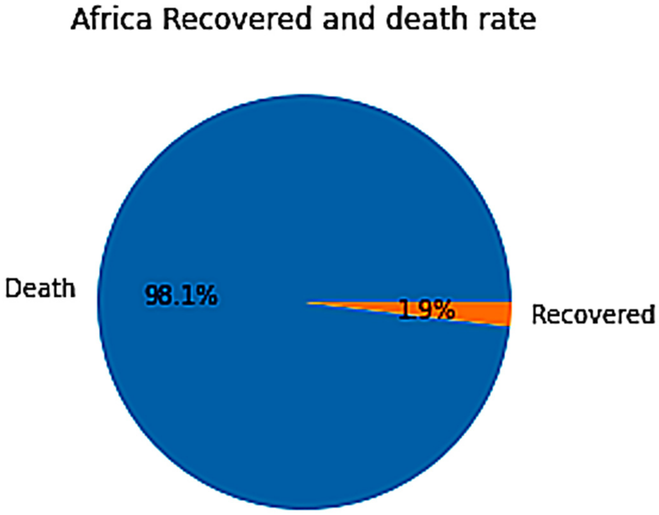
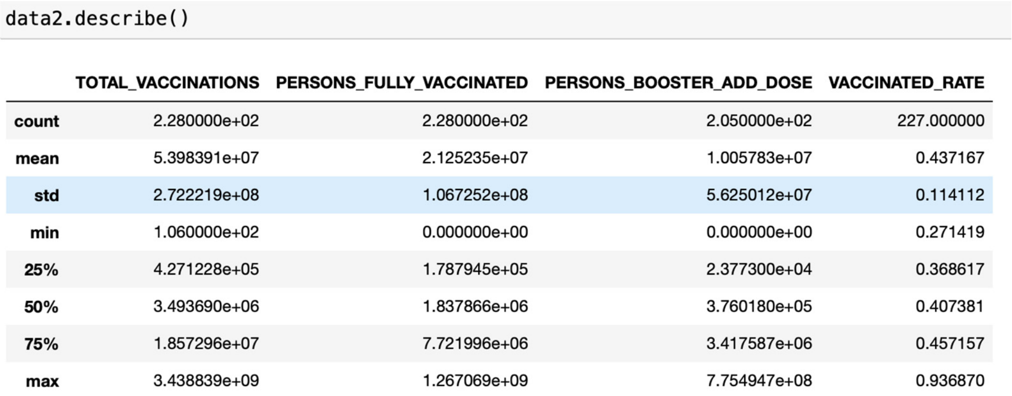


Figure 12. Comparison of recovery and mortality in Africa

As shown in the figure, the recovery and mortality in Africa and Europe are in sharp contrast. Firstly, we fitted the VAR model (model 1) constructed by the growth rate of COVID-19 infected people and the driving trip index. The fitting results with lninf as dependent variable and lndriv as independent variable showed that the goodness of fit was 0.7234, and the F statistical value was 34.56, which was significant at the level of 1%, and the fitting effect of the model was good. The fitting results with lndriv as dependent variable and lninf as independent variable are as follows: goodness of fit is 0.9503, and the model fits very well. The statistical value of f is 252.4, and it is significant at 1% level, which indicates that the linear relationship among variables is significant. Then, the VAR model (model 2) constructed by COVID-19 infected population growth rate and public transport travel index was fitted. The fitting results with lninf as dependent variable and lntran as independent variable showed that the goodness of fit was 0.7335, and the statistical value of F was 36.37, which was significant at the level of 1%, and the fitting effect of the model was also good. As we can see from the data set of vaccination, the average of receiving full vaccination is about 43.71%. The highest is 93.69%. This means the popularization of vaccines in COVID-19. As shown in Figure 13.



```
data2.describe()
```

	TOTAL_VACCINATIONS	PERSONS_FULLY_VACCINATED	PERSONS_BOOSTER_ADD_DOSE	VACCINATED_RATE
count	2.280000e+02	2.280000e+02	2.050000e+02	227.000000
mean	5.398391e+07	2.125235e+07	1.005783e+07	0.437167
std	2.722219e+08	1.067252e+08	5.625012e+07	0.114112
min	1.060000e+02	0.000000e+00	0.000000e+00	0.271419
25%	4.271228e+05	1.787945e+05	2.377300e+04	0.368617
50%	3.493690e+06	1.837866e+06	3.760180e+05	0.407381
75%	1.857296e+07	7.721996e+06	3.417587e+06	0.457157
max	3.438839e+09	1.267069e+09	7.754947e+08	0.936870

Figure 13. The prevalence of COVID-19 vaccines

This chapter puts forward the dispersion mode of COVID-19 through visual observation and data level calculation and analysis of the lung X-ray images of normal people and people infected with new coronavirus. Then the core/edge network is used to characterize this dispersion pattern. finally, the core/edge network is used to classify the lung infection images of new coronavirus, and good classification results are obtained. Through the change of the hydrophobic surface area of the new coronavirus docking with ACE2 protein in the original state in the temperature range of 300K to 320K with the increase of temperature, it can be clearly found that the hydrophobic surface area also increases with the increase of temperature, which also shows that the hydrophobic force between the binding conformation gradually increases with the increase of temperature; Later, the binding conformation of the RBD region and ACE2 protein after the mutation of the new coronavirus was analyzed and calculated by the same method. The calculation results showed that the hydrophobic area of N501Y mutation and E484K mutation increased significantly, indicating that their hydrophobic force increased significantly, which was also consistent with the reality. At the same time, the correctness of entropy enthalpy compensation mechanics theory was confirmed.

This chapter puts forward the dispersion mode of COVID-19 through visual observation and data level calculation and analysis of the lung X-ray images of normal people and people infected with COVID-19. Then the core/edge network is used to characterize this dispersion pattern. Finally, the core/edge network is used to classify the lung infection image of COVID-19, and good classification results are obtained. According to the change of the hydrophobic surface area between the original state of Covid-19 and ACE2 protein in the temperature range of 300K to 320K with the increase of temperature, it can be clearly found that with the increase of temperature, the hydrophobic surface area also increases, which also indicates that the hydrophobic force between binding conformations gradually increases with the increase of temperature. After the same method, the binding conformation between RBD region and ACE2 protein after Covid-19 mutation was analyzed and calculated. Through the calculation results, it can be found that the hydrophobic area of N501Y mutation and E484K mutation increased greatly, indicating that their hydrophobic force increased greatly, which is also consistent with the reality, and the correctness of entropy enthalpy compensation mechanics theory was confirmed.

4. Conclusion

The results showed that the areas with high mortality of NCP were lack of medical resources. In the two years that the world has experienced AIDS, the disease has had a serious impact on the lives of all people. At the same time, however, scientists and researchers have also conducted in-depth research on the virus and made much progress, which has brought the epidemic situation under control, reduced the damage to human lives and saved countless lives. Medical staff have also made great contributions to protecting people's safety.

In the third year of the epidemic, WHO hopes Covid can be truly controlled. This hope comes first from vaccination and continuous research and development. Experts once again stressed that at present, the vast majority of hospitalizations and deaths occur in unvaccinated people, rather than vaccinated people. The given vaccine is still effective for delta and omicron varieties. The popularization of Covid vaccine has brought great improvement to areas lacking medical resources. Secondly, XGBoost, an integrated algorithm and traditional machine learning algorithm are selected to construct the prediction model of the sample data. Later, deep learning algorithms such as convolutional neural network and recurrent neural network can be used to construct the classification model of the sample data. The algorithm is further improved and optimized, and the prediction accuracy of the classification model is continuously improved by comparing the results.

References

[1] https://covid19.who.int/who-data/vaccination-data.csv.

[2] Liu L, Chen X, Li X. Online Teaching Practice under the Epidemic Situation of Novel Coronavirus Pneumonia. Asian Agricultural Research, vol.12, no.8, pp.5,2020.

[3] Yang T, Yu X, He X. Early clinical manifestations and pulmonary imaging analysis of patients with Novel coronavirus pneumonia. Chinese Journal of Emergency Medicine, vol.29, no.1, pp.5,2020.

[4] Wang Ming. Analysis of 8274 cases of novel coronavirus nucleic acid detection and co-infection in Wuhan. China Medical Abstracts (Internal Medicine), vol.37, no.3, pp.8-8,2020.

[5] Babu R, Shebana M, Harish R M. Data science: a survey on the statistical analysis of the latest outbreak of the 2019 pandemic novel coronavirus disease (COVID-19) using ANOVA - ScienceDirect. Data Science for COVID-19, vol.57, no.13, pp.6,2022.

[6] Fan Y J, Chan K H, Hung F N. Safety and Efficacy of COVID-19 Vaccines: A Systematic Review and Meta-Analysis of Different Vaccines at Phase.vo.33, no.7, pp.10, 2021.

[7] https://covid19.who.int/WHO-COVID-19-global-table-data.csv.

[8] Naveed M, Rafiq M, Raza A. Mathematical Analysis of Novel Coronavirus (2019-nCov) Delay Pandemic Model. Computer, materials and continuum (English), vo.57, no.9, pp.14-14,2020.

[9] Liu C X. Pay attention to situation of SARS-CoV-2 and TCM advantages in treatment of novel coronavirus infection. Chinese herbal medicine: English version, vol.78, no.2, pp.7,2020.

[10] Yousef A, Bozkurt F, Abdeljawad T. Qualitative Analysis of a Fractional Pandemic Spread Model of the Novel Coronavirus (COVID-19). Computer, materials and continuum (English), vol.66, no.1, pp. 000, 2021.