Research on the Machine Learning Algorithms in the Prevention Model of New Coronary Pneumonia Epidemic Prevention and Control

Yukai Hong*
Zhejiang Zhejiang University, Hangzhou 310000, China
*Corresponding author: hongyukai@zju.edu.cn

Abstract. In response to the practical application of new COVID-19 transmission dynamics models such as new coronary pneumonia, the number of contacts associated with real-time data of confirmed cases is proposed, and the new COVID-19 transmission dynamics model based on machine learning algorithms is optimized. Enhance the practical value of the model. This type of exposure can be extended to other new cases of COVID-19, revealing the risk of transmission at different times during the outbreak.

Keywords: Machine learning algorithm, new crown pneumonia epidemic prevention and control, infection model, dynamic model.

1. Introduction
The 2019 novel coronavirus pneumonia (COVID-19) and other epidemics have had a great impact on humans. As of March 8, 2020, COVID-19 has caused more than 3,000 deaths in my country, which also poses a huge threat to global public health. The current situation for the prevention and control of the COVID-19 epidemic is severe, and the spread of the epidemic in my country is well under control, and it is in a critical period of epidemic prevention and control. In the process of epidemic prevention and control, mathematical models have played an important role in epidemic prediction, early warning and risk analysis.

COVID-19 is a process evolving over time. Based on the SIR model of differential equations, this paper analyses the law of the spread of COVID-19 and establishes a dynamic model. Apply the COVID-19 dynamic model to describe the process of disease development and change and the law of spread, predict the state of disease occurrence, evaluate the effect of various control measures, provide the best decision basis for disease prevention and control, and maintain human health and social and economic development. Through the planning of the number of people, a differential equation model of COVID-19 was established, and MATLAB software was used to fit the relationship curve of the number of patients over time [1]. The method of controlling variables was used to control certain variables unchanged and change one of them. Through comparison, find out the main factors that cause the infection of COVID-19 in order to make corresponding measures.

The key to this model is to divide the confirmed patients, suspected patients, cured patients, dead and normal people into two categories: infectious and non-infectious, with some special parameters, such as infection rate, cure rate, etc. The differential equation group finds out the change in the
number of normal people per unit time, the change in the number of confirmed patients, the change in
the number of suspected patients, and the change in the number of dead or cured (i.e., those who
withdraw from the system), thereby establishing a differential equation model.

2. Problem analysis
It is known that the incubation period of a certain incompletely known infectious virus is up to and the
cure time for the patient is days. The virus can be spread and spread through direct contact or oral
droplets. The average number of contacts in this population per day is r. In order to control the spread
and spread of the virus, the population is divided into five categories: confirmed patients, suspected
patients, cured, dead and normal people. The controllable parameter is the intensity of isolation
measures p (the percentage of patients who are isolated during the incubation period). Establish a
mathematical model of the spread of COVID-19 through reasonable assumptions.

Because the population is divided into five categories in the topic: confirmed patients, suspected
patients, cured, dead and normal people, we use the SIR model. In the model, we find out the changes
in the number of these five groups of people per unit time to establish a differential equation and get
the model. Then use MATLAB to draw graphics and analyse them to achieve the purpose of obtaining
countermeasures. The people within the scope of the investigation are divided into the following
categories: healthy people, that is, people who are susceptible to infection [2]. Record the number as
S(t), which means the number of people who are not infected with the disease at time t but may be
infected with the disease; the incubation period is the population who is infected with the disease, and
the number is I(t), which means the disease may be infected at time t The number of suspected patients,
but not the number of suspected patients; E(t) means the number of suspected patients who are
infected with the disease at time t; For suspected patients, record the number as Q(t) Represents the
number of people who were infected with the disease at t and were diagnosed as patients; the number
of recovered people is R(t), which represents the number of people who have been removed from the
infected at time t. Based on the above assumptions, the process diagram of healthy people from the
incubation period to removal from the infectious system is shown in Figure 1:
The meanings of the symbols are as follows. The incubation period of the virus (days) is $d_1$-$d_2$; the healing time of patients (days) is $d_3$; the number of patients in contact per day is $r$; the intensity of isolation measures is $p$; the healthy population at time $t$ is $S(t)$; the population in the latency period at time $t$ is $I(t)$; the suspected population of the disease at time $t$ is $E(t)$; the population who has been ill at time $t$ is $Q(t)$; the cured or dead population at time $t$ is $R(t)$; the COVID-19 infection rate is $\lambda$.

3. **Build the accuracy analysis of the mean square error evaluation model**

Mean square error:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (S_{\text{pred}} - S_i)^2$$

(1)

Root mean square error:

Both the mean square error and root mean square error will be affected by outliers, which will affect the final virus model evaluation

Average absolute percentage error:

$$RMSE = \sqrt{MSE}$$

(2)

The average absolute percentage error improves the infection characteristics of the virus outlier, which is equivalent to normalizing the error of each point, thereby reducing the impact of the absolute error caused by a single new coronavirus outlier

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(S_{\text{pred}} - S_i)}{S_i} \right| \times \frac{100}{n}$$

(3)

According to the hypothesis, the changes in the healthy population at time $\Delta t$ are:

$$S(t + \Delta t) - S(t) = -\lambda Q(t)(1 - p)S(t)\Delta t$$

(4)

The changes in the cured or dead population at time $\Delta t$ are

$$R(t + \Delta t) - R(t) = \frac{1}{d_3} I(t)\Delta t$$

(5)

Among them, $\frac{1}{d_3}$ is the patient's recovery rate per unit time. The changes in people with suspected symptoms at time $\Delta t$ include:

$$E(t + \Delta t) - E(t) = -\lambda Q(t)(1 - p)[E(t)(1 - p) + E(t)p\frac{1}{d_3}]\Delta t$$

(6)

The changes in the population with the disease at time $\Delta t$ are as follows (the population with the disease is equal to the incubation period patients converted to infected people minus the number of people removed):
\[ Q(t + \Delta t) - Q(t) = \frac{2}{d_1 + d_2} I(t) - \frac{1}{d_3} Q(t) \] 

(7)

The changes in the population of the latent group at time \( \Delta t \) are:

\[ I(t + \Delta t) - I(t) = \{ \lambda Q(t)(1 - p)[S(t) + E(t)(1 - p) + E(t)p \frac{1}{d_3}] - \frac{2}{d_1 + d_2} I(t) \} \Delta t \] 

(8)

\( \frac{2}{d_1 + d_2} \) is the constant of proportionality for the incubation period of patients to become infected per unit time. According to the above changes

\[
\begin{align*}
\frac{dS}{dt} &= -\lambda Q(t)(1 - p)S(t) \\
\frac{dR}{dt} &= \frac{1}{d_3} I(t) \\
\frac{dE}{dt} &= -\lambda Q(t)(1 - p)[E(t)(1 - p) + E(t)p \frac{1}{d_3}] \\
\frac{dQ}{dt} &= \frac{2}{d_1 + d_2} I(t) - \frac{1}{d_3} Q(t) \\
\frac{dI}{dt} &= \lambda Q(t)(1 - p)[S(t) + E(t)(1 - p) + E(t)p \frac{1}{d_3}] - \frac{2}{d_1 + d_2} I(t)
\end{align*}
\]

(9)

When \( d_1 = 1, d_2 = 14, d_3 = 30, r = 20, p = 40\% \), the patient was admitted to the hospital 2 days later, and the suspected patient was isolated 2 days later [3]. The number of patients in the initial state is: \( Q = Q(0) \ast \left( \frac{N - Q(0) - E(0)}{N} \ast r \right) \ast 2 \), then the number of patients changes over time as shown in Figure 2:
From the above figure, it can be seen that when $d_1=1, d_2=14, d_3=30, r=20, p=40\%$, the patient is admitted to the hospital 2 days later, and the suspected patient is isolated 2 days later. When $t \in [0,14.324]$, the number of patients rose sharply, reaching the maximum at $t=14.324$, and the number of patients at this time was $y=228870.1396$. The number of patients has increased over time after taking medical measures, such as hospital admission and isolation of suspected patients. It showed a downward trend, and the number of patients was 126.5086 after 250 days.

**Table 1. The corresponding value of each parameter**

| Problem          | Maximum time | Maximum number of patients | Strength of isolation measures P | Number of days before patient admission n | Number of contacts per person per day r | Number of patients after 250 days |
|------------------|--------------|----------------------------|----------------------------------|------------------------------------------|----------------------------------------|----------------------------------|
| Second question  | 13.0971      | 228669.8722                | 0.6                              | 2                                        | 20                                     | 115.5535                         |
| Third question   | 13.3715      | 52851.3637                 | 0.6                              | 1.5                                      | 20                                     | 26.5958                          |
| Fourth question  | 14.324       | 228870.1396                | 0.4                              | 2                                        | 20                                     | 126.5086                         |

It can be seen from the above table that 1. When the isolation intensity is the same, the start time of patients' admission will affect the number of patients in a certain period of time. Obviously, it can be seen that the treatment time of patients is prolonged and the number of patients is increased compared with admission after 1.5 days. Therefore, relevant departments should isolate and treat the patient in time. 2. When patients are admitted at the same time, the intensity of isolation measures will affect the length of time when the number of patients reaches the maximum [4]. It can be seen that after the intensity of isolation measures is reduced, the number of patients is relatively high. Therefore, relevant departments should strengthen the intensity of isolation measures and increase vigilance. From the analysis of the changes in the values of the above two parameters, it can be seen that the "time to hospital after illness" and the "strength of isolation measures" have great sensitivity and relevance to
the development of the COVID-19 epidemic. When will the patient go to the hospital for treatment after the illness, is of greater significance to the control of the epidemic. Therefore, "early detection, early isolation, and early treatment" can help us effectively and quickly control the spread and spread of COVID-19.

4. Simulation results
According to the population status of the spread of COVID-19, the cell evolution rule is applied to Moore neighbours with a radius of 1 to simulate the spread of COVID-19. Set the cell space as a grid of $L \times L = 100 \times 100$, and the initial probability of a cell being infected with the virus is 0.0018. Assuming $k=0.5$, $l=0.2$, the duration of onset is 7 days, and the duration of immunity of the recovered person is 365 days [5]. Figure 3 shows the simulation data within 200 days, and the simulation results are basically consistent with the actual data.

As can be seen from Figure 3, in the early stage of the virus outbreak, the spread of COVID-19 was relatively fast and the number of patients increased rapidly. In the later stage, the spread of COVID-19 decreased and the number of patients began to slowly decline. The reason for the decline in the number of patients is that with the progress of medical treatment, patients are immune and no longer infected after being cured; first, some patients have died in the later stage of the virus transmission and will no longer be transmitted to other people [6]. In real life, once everyone finds that they are unwell, they will go to the hospital and ask the doctor for help, who will treat us. It is assumed here that the doctor has a special medicine, as long as the doctor asks for help, the patient will be cured, thus reducing the duration of the disease. Keep the initial parameters unchanged, and study the impact of the duration of onset on the spread of COVID-19. As shown in Figure 3.

It can be seen from Figure 3 that when the duration of the onset is $T=5$, the number of infected patients increases at a slower rate than the actual data, and the number of infected patients is also less than the actual data; when the duration of onset is $T=9$, the number of infected patients the rate of increase is greater than the actual data, and the number of infected patients is also greater than the actual data. The longer the duration of the patient's illness, the more susceptible people the patient will come into contact with during the activity, resulting in an increase in the cumulative number of patients [7]. Therefore, in the early stage of the infection outbreak, effective methods should be actively adopted to rescue patients, or to effectively isolate patients, reduce the contact between patients and other people, and inhibit the spread of COVID-19.
5. Conclusion
The differential equation model is used in this model to make reasonable assumptions about the spread of COVID-19 and over-fit it to obtain the development trend of COVID-19, which can effectively predict the arrival of the COVID-19 climax and be accepted by the masses. COVID-19 prevention knowledge plays a good warning role. However, the parameters in the model are all random, so there will still be errors in the results obtained. It is impossible to accurately predict the exact time of each COVID-19 climax, and can only be limited within a certain time. The COVID-19 model established by this model can effectively predict the climax of COVID-19, so it can be further applied in real life, especially for the outbreak and epidemic of COVID-19 in today’s society.

References
[1] Weusten, J., Van Drimmelen, H., Vermeulen, M., & Lelie, N. A mathematical model for estimating residual transmission risk of occult hepatitis b virus infection with different blood safety scenarios. Transfusion, 57(3) (2017) 841-849.
[2] Kim, J. A., Park, S., Kumar, M., Lee, C., & Shin, O. S. Insights into the role of immunosenescence during varicella zoster (shingles) virus infection in aging cell model. Journal of Dermatological Science, 86(2) (2017) 72-79.
[3] Manda, E. C., & Chirove, F. Acute hepatitis b virus infection model within the host incorporating immune cells and cytokine responses. Theory in Biosciences, 139(5258) (2019) 1-17.
[4] Zhang, Z., Dong, Z., Wei, Q., Carr, M. J., Li, J., & Ding, S., et al. A neonatal murine model of coxsackievirus a6 infection for evaluation of antiviral and vaccine efficacy. Journal of Virology, 91(9) (2017) e02450-16.
[5] Manda, E. C., & Chirove, F. Acute hepatitis b virus infection model within the host incorporating immune cells and cytokine responses. Theory in Biosciences, 139(2) (2020)153-169.
[6] Tang, S., Teng, Z., & Miao, H. Global dynamics of a reaction–diffusion virus infection model with humoral immunity and nonlinear incidence. Computers & Mathematics with Applications, 78(3) (2019) 786-806.
[7] Li, X., Wu, S., Du, Y., Yang, L., & Hong, B. Entecavir therapy reverses gut microbiota dysbiosis induced by hepatitis b virus infection in a mouse model. International Journal of Antimicrobial Agents, 56(1) (2020) 106000-106009.