# Trend Analysis of Infected Cases with Corona Virus in Nepal by Using Polynomial Regression Model

Madhav Prasad Poudel
School of Engineering, Pokhara University, Kaski, Nepal.

**Abstract**

Among the viral infections, the infection of the Novel Coronavirus (COVID-19) is the most recent one in the world Firstly; it was detected in Wuhan city of China in December, 2019 and became pandemic in the short span of time. Since the pandemic of COVID-19 is alarming the globe, the number of infected cases around the world is increasing day by day. Several mathematical models studies are conducted based on the infected, death and recovered cases It is still unknown whether this epidemic will be under control or not. In context to Nepal, the infected cases at present scenario are uncontrolled. By the first week of September, the number of positive cases recorded was 44236 with recovery case 25561. The death rate is increasing day by day and till now it is 300 In spite of Nepal government's effort to minimize the pandemic, the cases of infected rates have been increasing linearly. In this present study, the main goal is to find out the trend of spread of COVID-19 in Nepal. For the analysis of the trend, the polynomial regression model is fitted in the curve obtained in the timeline to the number of infectious cases. Through the polynomial regression, the infectious rate and inflection points are studied. By the time of the study, the highest rate was measured on 30 August 2020. The following two remedies are recommended to minimize the spread of virus which are as follows: keeping the infected person in insolation and increase of the immune power of the individual by taking medicine.

**Keywords**

Pandemic, COVID-19, coronavirus, mathematical modeling, Polynomial regression model.

## 1. Introduction

The history of human corona virus started in 1965 with the first coronavirus B814. It was discovered by Tyrell and Byoe. It was detected in the trachea of the respiratory tract in an adult who was suffering from common cold[1]. Since then different forms of coronavirus exit in earth. Recently , the world is under the epidemic of new type of coronavirus (COVID-19) which is believed to be worldwide through its origination in Wuhan, China. Till now it has reached to across the globe. It is also a virus that is transmitted through contact with an infected person[1]. This contact can be through respiratory droplets or salivary droplets during the coughing or sneezing of the infected person[2]. It has led to high mobility rate and mortality rate in China, Europe, United States and most of the other countries in the world[3]. It has lead to unprecedented health crises throughout the world. Many people are dying because of disease and hunger in the world due to coronavirus. So on March 11, 2020 the world health organization (WHO) declared COVID-19 as a global pandemic. COVID 19 is caused by a novel corona virus which is also known as severe acute respiratory syndrome coronavirus (SARS-COV-2)[4, 5].

Although the large numbers of individuals infected with COVID-19 globally, it is bitter fact that there is no medicine invented by now. Yet international organizations are playing a vital role in reducing the panic and for the removal of the terror [6].

With this global expansion of the Covid-19, Nepal cannot be isolated from the pandemic. In Nepal, the first case of Covid-19 was confirmed on 23-January 2020 who returned from Wuhan of China, a thirty-two years old Nepalese male. It was tested in a reference laboratory in Hong Kong [7, 8]. Similarly, the second case was reported on 23rd March, 2020 in a 19 years old female, returned from France and Doha Qatar [8, 9]. On 24th

March, 2020 the government of Nepal implemented the nationwide lockdown so that the spread of the diseases through close contact can be minimized.

In this hard time of pandemic, researches from various fields and areas are conducting the research to analyse it and to curb it's ill effects. Apart from the nature of virus and experimental studies statistical analysis are also carried out. These studies will definitely help the policy makers to evaluate the current situation and to take necessary action as a remedy in future. Furthermore, the models developed can also provide an estimate of the ending time and tentative numbers of people that may be infected by this pandemic[10].

Likewise, in other countries, the study on the spread of corona virus is felt to be necessary in Nepal. In this context, the spread and impact of the public health interventions were mathematically analyses and modeled by Bhandary et al. [7]. Most often in pandemic of the disease the reproduction rate in the susceptible population is studied and hence the same is done in Covid-19. High reproduction number in short interval of time generates a rapid growth. In this phase, it is obvious that the number of cases exhibits an exponential growth. The basic reproduction number is a function of exponential growth rate with the time [11].

Some works are recently published on the COVID-19 pandemic. They are focused in and control of covid-19. Some models based on practices to minimize the disease transmission in some States of India. These studies are useful to forecast and to control the transmission in susceptible humans [11]. The Mathematicians are using different types of models to study the final size of COVID-19 epidemic in the world. Applying numerical methods to calculate some approximate solutions could improve the predictions and estimations of spread of COVID-19in the country. Batista M [12] studied on estimation of final size of the coronavirus epidemic by logistic model for China, South Korea and the rest of the world.

Niazkar, H. K. and Niazkar, M. has studied the spread of COVID-19 through a second order polynomial equation with constant coefficients based on based on infected cases. The model seems to be inadequate and therefore concluded to be insufficient in studying the spread of coronavirus[13]. The study seems to be inadequate to fit the trend of coronavirus cases on higher order polynomial with the constant coefficients.

The aim of this study is to fit the polynomial regression model of the Covid-19 ongoing phase of pandemic in Nepal. Nevertheless, as discussed earlier, the lockdown issued by the government, has somehow helped to retard the growth rate in Nepal. So instead of using the exponential growth, the polynomial equation can be used to study the trend of spread of corona virus. This study will be useful to inform the official and the public in preparedness spread, forecasting the trend and highlighting the importance of implement strict measures against Covid-19 to resist the increase of the virus.

## 2. Materials and methods

### 2.1 Polynomial Regression model

The Mathematical models are beneficial, convenient and efficient tools to understand the behavior of an infection when it enters a community and to investigate under which conditions or when it will be under control. Some diseases are not totally eliminated but continue in a normal form[14, 15]. Thus the models are essential to make a therapeutic choice when surge capacity has been exceeded or without ready access to laboratory testing. The models are necessary for the policymaker to acquire medical supplies, to allocate human resources and to plan for hospital beds and ensure the sustainability of the health system throughout the peak and duration of the epidemic [15]. Currently, COVID-19 is of great concern to researcher, governments, and all people around the globe due to the high transmission rate and the significant number of deaths that occurred [14]

Many scientific and technological problems are modeled mathematically by systems of polynomial regression, for example, chemical reactions, ecological interactions, biological process. Normally in a pandemic of the certain diseases, the exponential or the logistic models are used. In pandemic of an disease the SIR, SEIR, SARS and MERS models are also very common [4, 9,13, 16, 17]. In exponential model, the growth has a uniform pattern and in logistic growth, the rate of increase ceases to certain finite value for $t>\infty$, where t is the time frame. These types of graphs are obtained if the physical conditions remain constant. But the study of COVID-19, the physical constraint such as lockdown, mobility, medical tests etc. were not remaining constant as the number of cases increased. No medicines are made yet and there is no issue of medicines for the effect of the curve. So at certain points some changes in trend were observed in the curve. Therefore a polynomial curve was attempted to be fitted in the infectious function for the COVID-19 study.

Several studies are done to study the behavior of COVID-19 by polynomial regression method [18-21]. This method is also regarded as one of the best tool to analyze and predict the growth of the pandemic. In other study, Yadav [19] has mentioned that the accuracy of the epidemic of the disease by this method is 99.85%.

Polynomial regression models are special types of regression models. The simple linear regression is a type of polynomial regression. A polynomial regression model is hierarchical if all the terms of the independent variable raised to a power are present. In this case, the observation, C(t) is dependent on t and higher powers of *t*. It is a diffusion model because the rate of spread is time is dependent. As time *t* increases, the spread measured in number of persons per time unit also increases until it reaches to a saturation. Any further increase in time will decrease in the number of cases. Since, it is not a theoretical model; there could be increase and decrease between small change in t. But in the long run, it shows a pattern that can be modelled by the so called diffusion model[20]. The coefficients of $t^n$ in the polynomial equation indicates the behavior, relation with the environment and action on the host [18].

Polynomial regression is a special case of multiple regression. The polynomial Regression model of degree n is given by[20]

$$C(t)= a_n t^n+a_{n-1}t^{n-1}+a_{n-2}t^{n-2}+\text{--------------}+a_1t+a_0+e_t \qquad (1)$$

where $e_t$ is the error term. This is also time dependent and follows the probability distribution of C(t). t is a independent variable (time in day) and C(t) is the number of positive cases of the coronavirus.

The equation (1) can be written as $C(t)= AT + e_t$ (2)

Where $AT= a_n t^n+a_{n-1}t^{n-1}+a_{n-2}t^{n-2}+\text{--------------}+a_1t+a_0$

The equation (1) can have a matrix representation for t denoted by i=1, 2, 3,-------------------------

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ - \\ - \\ C_n \end{pmatrix} = \begin{pmatrix} 1 & t_1 & t_1^{\,2} & - & - & t_1^{\,m} \\ 1 & t_2 & t_2^{\,2} & - & - & t_2^{\,m} \\ 1 & t_3 & t_3^{\,2} & - & - & t_3^{\,m} \\ - & - & - & - & - & - \\ - & - & - & - & - & - \\ 1 & t_n & t_n^{\,2} & - & - & t_n^{\,m} \end{pmatrix} . \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ - \\ - \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ - \\ - \\ \varepsilon_n \end{pmatrix}$$

Where $\beta$ is the parameter vector and a vector $\mathcal{E}$ is of random errors

Then the vector of estimated polynomial regression coefficient is

$$\beta=(T^TT)^{-1}T^TC \qquad (3)$$

The instantaneous rate is calculated as

$$C'(t)=n\, a_n t^{n-1}+(n-1)a_{n-1}t^{n-2}+(n-2)a_{n-2}t^{n-3}+\text{--------------}a_1 \quad (4)$$

And the inflection occurs at $C''(t)=0$

i.e $$n(n-1)\, a_n t^{n-2}+(n-1)(n-2)a_{n-1}t^{n-3}+(n-2)(n-3)a_{n-2}t^{n-4}+\text{--------------}a_2 \quad (5)$$

By the solution of equation (2) , the rate of increase or decrease of reproduction can be found out and the inflection point helps to predict the nature of pandemic.

## 3. **Results and Discussion**

### 3.1 Mathematical Modeling of a curve

For the study, the data of daily Covid-19 cases were obtained from the Nepal government situation reports [22] from 23, January to 4 September, 2020. Although Nepal had a confirmed case in 23, January to 12 March 2020, this duration was not considered being a single case in that time frame. The study is conducted from the number of positive cases starting from 12 March 2020 predicting that the growth follows the polynomial regression model. The modeling of epidemic curve was plotted by using the polynomial regression function. The slope at the certain points were found out to predict the growth rates.

The entries were placed in the excel spreadsheet The situation of total number of cases, Recovery cases and the number of death cases from 12 March 12, 2020 to 4 September 20202 are as follows;

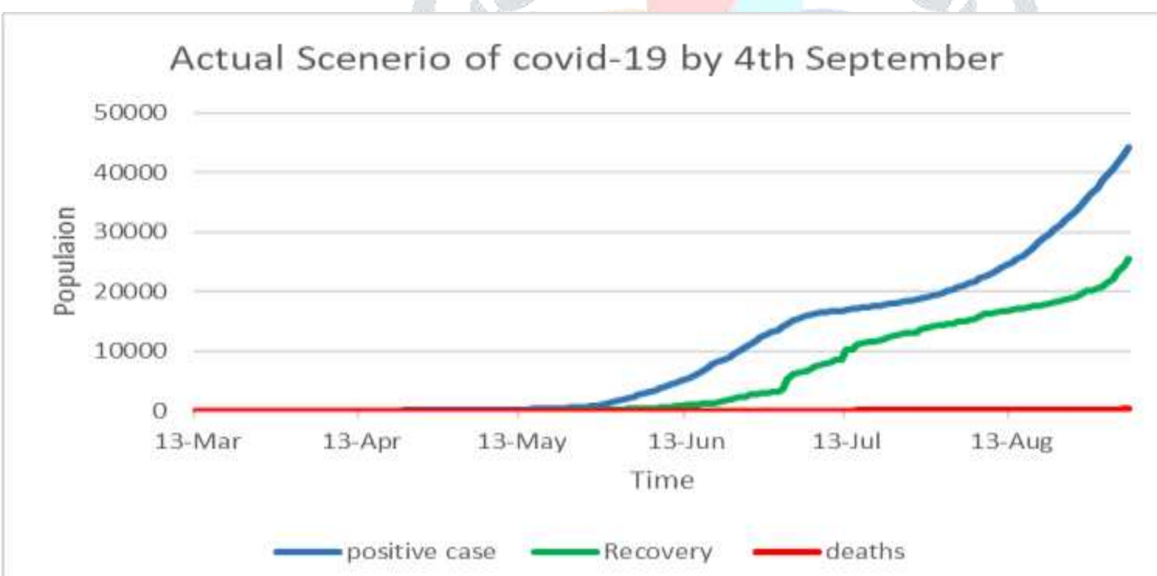


Figure 1 . Total number of coronavirus positive, Recovery and Death cases.

The general polynomial equation of the number of infected cases was modelled through the trend fitting in excel and the equation of the trend line is given below.

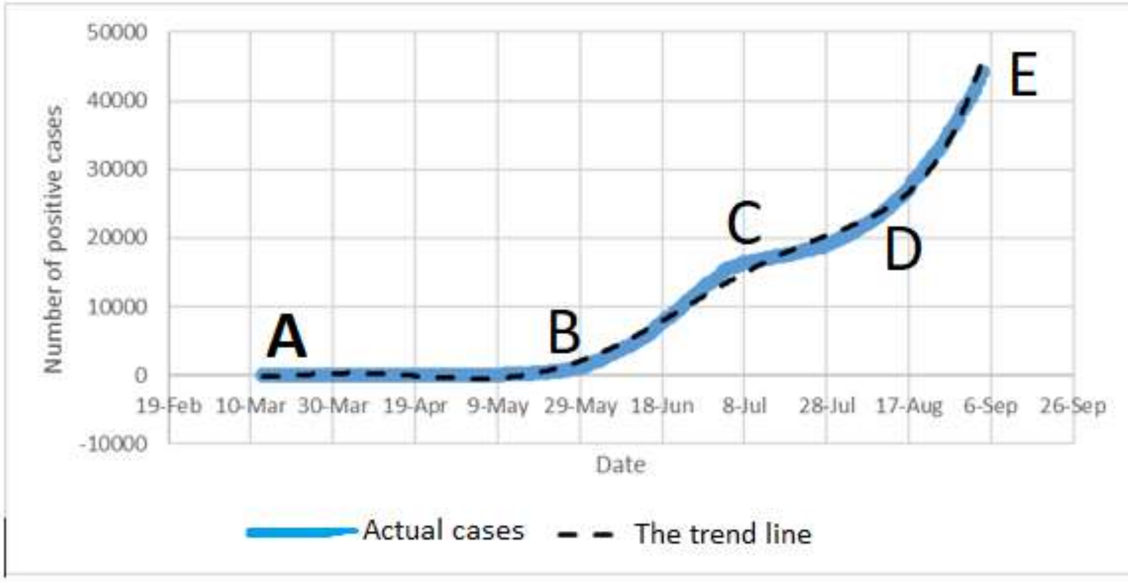


Fig 2 Number of Actual case with the trend line of the equation

In figure 2 the polynomial function is fitted with the curve to the number of infections timeline. The equation of the polynomial is of sixth degree (sextic equation).The equation so obtained is

$$C=7*10^{-8}t^6-3*10^{-5}t^5+0.0036t^4-0.1581t^3+0.3365t^2+77.544t+548.68 \quad (5)$$

Here t is a number of days from 12th March 2020

The value of coefficient of determination $R^2=0.9957$

The curve seems to rising up from B to C and then retards from C to D. It is again rising from D onwards to E. The explanations behind this nature of the curve are mentioned in the finding section. However for the perfect fit of the regression curve, it is divided into two parts. A to C is considered as the first part and the portion of the curve from C to E is fitted with second part of the curve. They are given by figures 3 and 4 respectively.

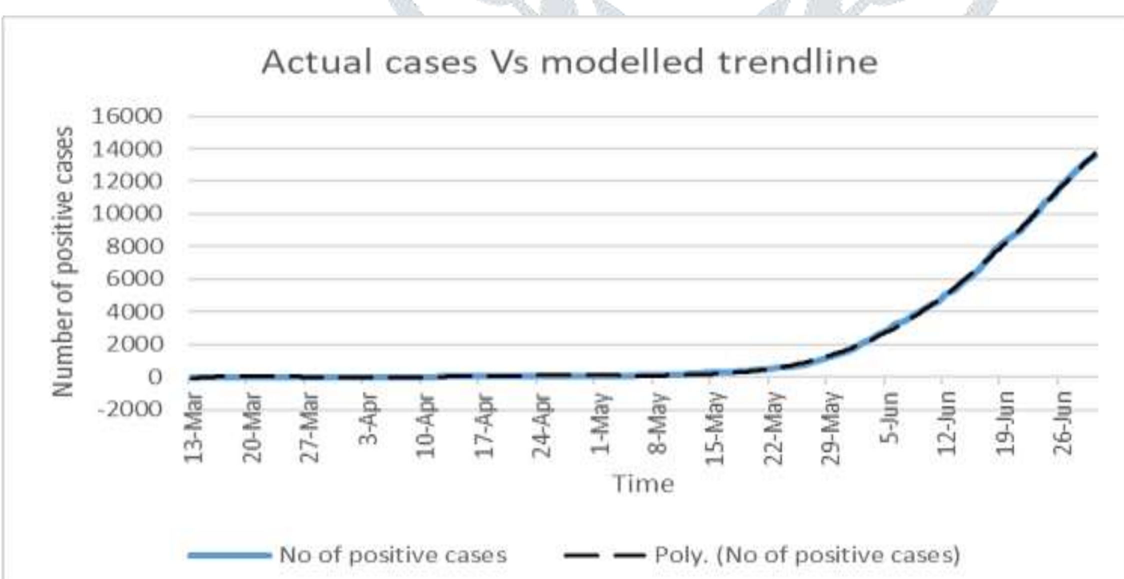


Figure. 3, First phase of rise

The equation fitted to the curve in figure 3 is given by

$$C(t)=-2*10^{-7}t^6+7*10^{-5}t^5-0.0073t^4-0.3567t^3-0.0468t^2+72.7249t-164.77 \quad (6)$$

With coefficient of determination $R^2$= 0.9996

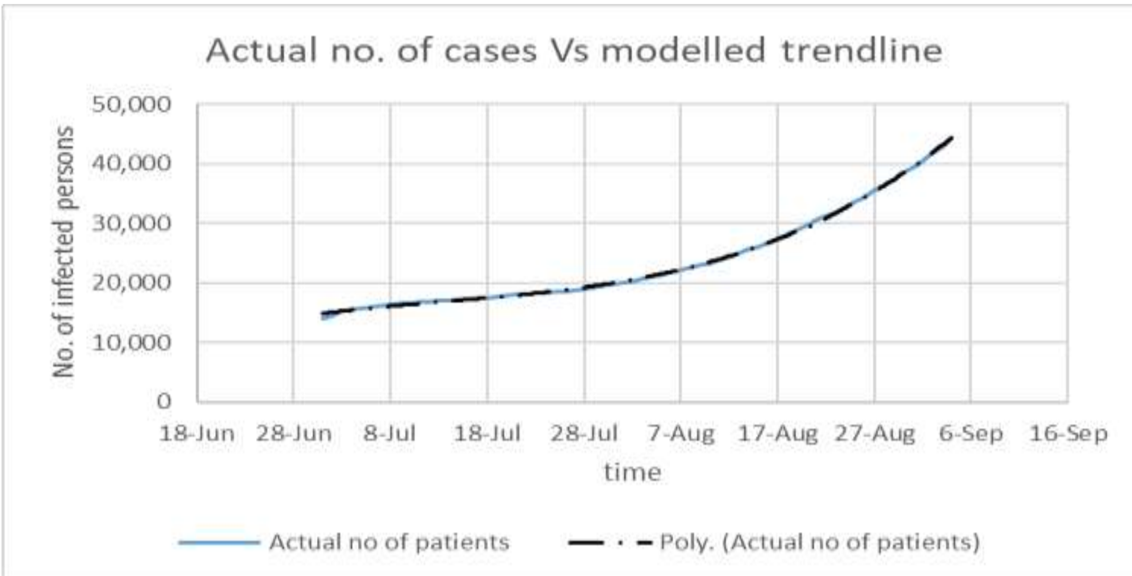


Figure 4.Second part of the rise

Similarly the equation fitted with the curve in figure 4 is

$C(t)=0.1429t^3+58699t^2-216.16t+14657$ with $R^2=0.9995$ (7)

**3.2 Results from the curve modelling**

Through the analysis of the curves obtained in figures 2, 3, and figure 4 following results are observed.

i) The value of $R^2$ for the curves modelled in figure 1, 2, and 3 indicate the good fit for the given curve of the number of positive test for coronavirus (COVID-19) in Nepal. Since the $R^2$ value in figure 2 is less than in figures 3 and 4, the polynomial function is better fitted in figure 4. The function obtained from figure 4 will give the approximate values for the prediction in coming days if the physical conditions remains constant.

ii) The overall number of cases can be expressed in terms of two conditional functions.

$$C(t)=\begin{cases} -2*10^{-7}t^6+7*10^{-5}t^5-0.0073t^4-0.3567t^3-0.0468t^2+72.7249t-164.77 & \text{for } t<111 \\ 0.1429t^3+58699t^2-216.16t+14657 & \text{for } t\geq 111 \end{cases}$$

iii) The slopes at the point A, B, C and D if the figure was calculated. The slopes and y intercepts were as follows. The slope (rate of increase) was compared with the actual value.

| S. n. | Position in figure 2 | Days acquired | Equation of the tangent | Slope (Rate of infection) |
|---|---|---|---|---|
| 1 | A | 72 | 102t+576 | 102 |
| 2 | B | 111 | 240t+15024 | 240 |
| 3 | C | 150 | 526.1t+23403 | 526 |
| 4 | D | 170 | 1234t+39236 | 1234 |

iv) While observing the latest trend of the regression model (fig 4) It is seen that if t = 0, C(0) =14657. This number refers to new coronavirus cases will be equivalent to 14657 at the steady state of inflection.

v) On solving equation (7), t = -136923.2days. The negative time indicates that it is nearly impossible to bring the infection to zero. Changing habits and weakening of virus can be a remedy, which is not a part of this study.

v) The recovery function of novel corona virus (COVD-19) in Nepal is fitted as follows.

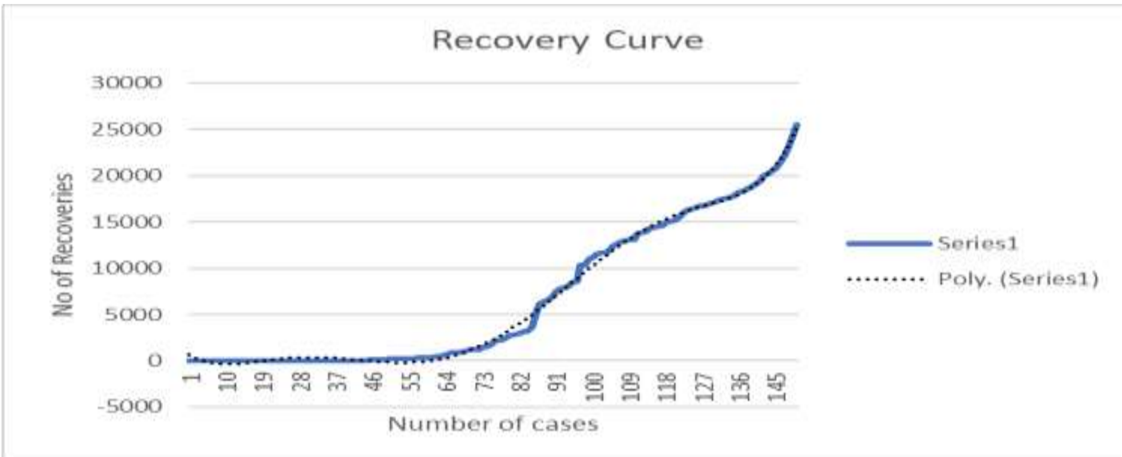


Fig 4. Recovery curve

The recovery curve is plotted from 2 April 2020 and the recovery (R) equation is

$R(t)= 2*10^{-7}t^6-8*10^{-5}t^5+0.01293t^4-0.8842t^3+27.26t^2-328.11t+966.89$ coefficient of determination with $R^2=0.9972$

vi) The maximum no of deaths in 30 August 2020 with the total number of 14 until 4 September 2020[12].

## 4. Conclusions

The study from the fitting of the model, it is seen that sixth order polynomial (sexist equation) is suitable to fit the data. Polynomial regression is adequate to reveal some intensive characteristics of coronavirus (COVID-19) in Nepal.

The data and the models show the following conclusions.

a) The inflection rate experiences a slow increase for first 73 days and after initial report identify the virus.
b) Then the virus started to rise to towards the peak for the rest 39 days till 12th may 2020
c) The coefficients of $t^n$ suggested the following facts
   i) $t^6$ and $t^5$: The positive coefficients for $t^6$ and $t^5$ indicates the favorable condition of the virus while negative coefficients indicates harshening of environment conditions and reducing virus availability. The actual values of these coefficients are supposed to be dependent on population density and mobility. In the equation of the figure 4 (equation 7) shows the absence of terms of $t^6$ and $t^5$ indicated that there is no relation for the virus multiplication. Virus multiplication with assault host is neither negative nor positive.
   ii) $t^4$and $t^3$: The positive coefficient of $t^4$ and $t^3$ measures the virus success in using the host RNA to reproduce. The negative coefficient measures host success in reducing the virus reproduction using their immunological and inflammatory systems. The coefficient of t4 in equation 3 is absent. However it is the coefficient of $t^3$ is present. It means the virus are using the host RNA to reproduce. The coefficient is 0.1429. i.e the reproduction rate is slower.
   iii) $t^2$ and t : The polarity of the coefficients of $t^2$ and t is inversely correlated with the quality of medical assistance to the infected host. High quality services promote negative coefficients and vice versa. The availability and quality of primary health care and hospital services are directly related to the available resources of each country health systems. The negative coefficient of $t^2$ in equation 7 gives the positive vibes that health awareness and medical efforts are helping to reduce the inflammation.
d) In figure 2, the curve of the actual statistics was seen to be flattened because the local government issued the lockdown in major cities and hotspots of corona virus as Kathmandu and Rautahat from 17

August 2020. Further restrictions were imposed to reduce the spread of the coronavirus (COVID-19)[23]

e) The rate of increase of the infection of corona virus at 4 September 2020 from the fitted polynomial curve is 1234. Since the rate of increase is very high. It indicates that the coronavirus has gone into the community.

f) From the above study, the isolation of corona positive patients is only an ultimate option to minimize the spread of coronavirus. Keeping distance and scheduling the times for works with sanitizers can minimize the transmission of the virus.

**Author Contributions**

Madhav Prasad Poudel is the sole author of this research work. The conceptualization, methodology, analysis and preparation of the manuscript have been done by myself.

**Funding**

There is not any specific funding available for this research work.

**Conflicts of Interest**

The author declares no conflict of interest.

**References**

1. Tyrrell, D. A. J., and M. L. Bynoe. "Cultivation of viruses from a high proportion of patients with colds." *Lancet* (1966): 76-7.
2. He, Sha, Sanyi Tang, and Libin Rong. "A discrete stochastic model of the COVID-19 outbreak: Forecast and control." *Math. Biosci. Eng* 17 (2020): 2792-2804.
3. Kahn, Jeffrey S., and Kenneth McIntosh. "History and recent advances in coronavirus discovery." *The Pediatric infectious disease journal* 24.11 (2005): S223-S227.
4. Wang, Jin. "Mathematical models for COVID-19: Applications, limitations, and potentials." *Journal of public health and emergency* 4 (2020).
5. Gilbert, Marius, et al. "Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study." *The Lancet* 395.10227 (2020): 871-877.
6. Ahmed, Ayub, et al. "Analysis coronavirus disease (COVID-19) model using numerical approaches and logistic model." *AIMS Bioengineering* 7.3 (2020): 130.
7. Bhandary, Shital, et al. "Trend analysis, modelling and impact assessment of COVID-19 in Nepal." *medRxiv* (2020).
8. MoHP. Coronoavirus disease (COVID-19) outbreak updates and resource material. Health Emergency and Disaster Management Unit, Health Emergency Operation Center 2020. Available from: https://heoc.mohp.gov.np/update-on-novel-corona-virus-COVID437 19
9. Bastola, Anup, et al. "The first 2019 novel coronavirus case in Nepal." *The Lancet Infectious Diseases* 20.3 (2020): 279-280.
10. Ahmed, Ayub, et al. "Analysis coronavirus disease (COVID-19) model using numerical approaches and logistic model." *AIMS Bioengineering* 7.3 (2020): 130.
11. Ma, Junling. "Estimating epidemic exponential growth rate and basic reproduction number." *Infectious Disease Modelling* 5 (2020): 129-141.
12. Batista, Milan. "Estimation of the final size of the second phase of coronavirus epidemic by the logistic model." *MedRxiv.[Internet]* (2020).
13. Niazkar, Hamid Reza, and Majid Niazkar. "COVID-19 international outbreak and the need for a suitable estimation model: A second-order polynomial equation with constant coefficients based on imported infected cases seems inadequate." *Asian Pacific Journal of Tropical Medicine* 13.4 (2020): 185.
14. Koltsova, E. M., E. S. Kurkina, and A. M. Vasetsky. "Mathematical Modeling of the Spread of COVID-19 in Moscow and Russian Regions." *arXiv preprint arXiv:2004.10118* (2020).

15. Wang, Ning, et al. "An evaluation of mathematical models for the outbreak of COVID-19." *Precision Clinical Medicine* (2020).

16. Chowdhury, Rajiv, et al. "Dynamic interventions to control COVID-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries." *European journal of epidemiology* 35.5 (2020): 389-399.

17. Liang, Kaihao. "Mathematical model of infection kinetics and its analysis for COVID-19, SARS and MERS." *Infection, Genetics and Evolution* (2020): 104306.

18. da Rocha, Armando. "Regression Polynomial Analysis of the COVID 19 Epidemics: Progress Report 1." *Available at SSRN 3636824* (2020).

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

20. Ekum, M., and A. Ogunsanya. "Application of hierarchical polynomial regression models to predict transmission of COVID-19 at global level." *Int J Clin Biostat Biom* 6.1 (2020): 027.

21. Ostertagová, Eva. "Modelling using polynomial regression." *Procedia Engineering* 48 (2012): 500-506.

22. MOHP. A report on Ministry of Health and Population, Nepal. Accessed on 5th September, 2020. https://covid19.mohp.gov.np/situation-report

23. Online news. Accessed on 4th September, 2020. https://reliefweb.int/report/nepal/covid19-nepal-covid-19-and-flood-response-situation-report-noxx-17-august-2020