



PROEXC
PRÓ-REITORIA
DE EXTENSÃO E CULTURA

Título original: SEIR and Regression Model based COVID-19 outbreak predictions in India

Título traduzido: Predições do surto de COVID-19 na Índia, baseadas em um modelo SEIR e um modelo de Regressão

Autores: Rajan Gupta¹, Gaurav Pandey², Poonam Chaudhary², Saibal Kumar Pal³

Projeto Covid-19 e a Matemática das Epidemias - Fazendo a Ponte entre Ciência e Sociedade

Tradução: Danillo Barros de Souza e Jonatas Teodomiro

Síntese: Camila Sousa e Júlia Lyra

Coordenação: Felipe Wergete Cruz

Introdução

Enquanto uma vacina eficiente e bem testada contra a Covid-19 não é lançada, uma das partes chave na administração da pandemia segue sendo o achatamento da curva de propagação. Mas, para tomar boas decisões e montar um plano de ação concreto, é necessário, antes, entender o vírus e as suas características.

Com esse intuito, cientistas e pesquisadores em mineração de dados têm utilizado diversas tecnologias a seu favor para prever a dinâmica da pandemia. Entre elas, a aprendizagem de máquina (do inglês *machine learning*) - campo de estudo que dá aos computadores a habilidade de aprender sem serem explicitamente programados - que foi destaque em uma pesquisa desenvolvida pela Organização de Pesquisa de Defesa e Desenvolvimento e pelas universidades de Delhi e NorthCap, na Índia.

¹ Faculdade Deen Dayal Upadhyaya, Universidade de Delhi, Índia

² Departamento de CSE & TI, A Universidade NorthCap, Índia

³ Organização de Pesquisa de Defesa e Desenvolvimento, Índia

Destrinchando

Os principais objetivos do trabalho foram três:

1. Achar a taxa de propagação da doença na Índia;
2. Desenvolver um modelo matemático SEIR (Suscetível, Exposto, Infectado, Recuperado) para avaliar a propagação da doença;
3. Previsão do surto de COVID-19 com modelos SEIR e de regressão.

No modelo SEIR, a parte mais importante é calcular o valor do R_0 (número básico de reprodução, que indica quão contagiosa é uma doença infecciosa). Em termos simples, o R_0 determina uma média de quantas pessoas podem ser afetadas, ao longo do tempo, por uma única pessoa infectada.

Se $R_0 < 1$, significa que a propagação irá parar;

Se $R_0 = 1$, significa que a propagação está estável

Se $R_0 > 1$, significa que a propagação crescerá na ausência de uma intervenção, como mostrado na Figura 1.

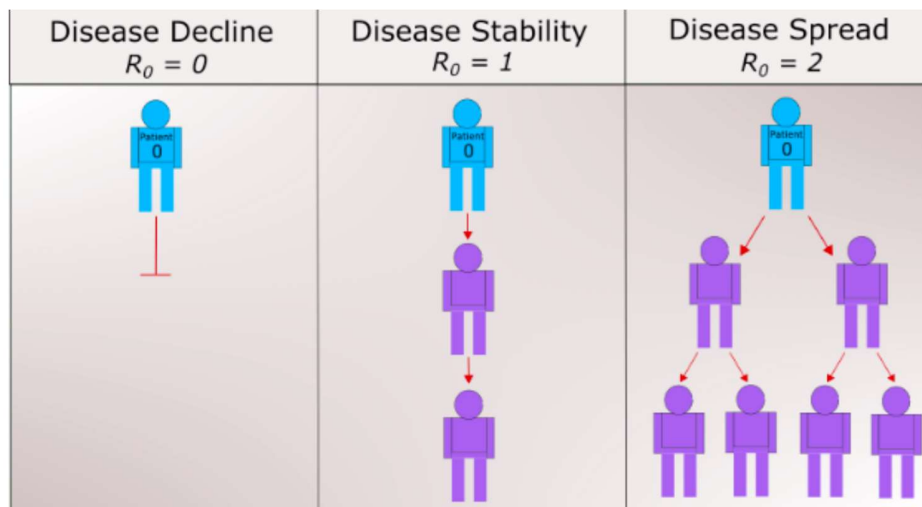


Figura 1. Demonstração da propagação da doença baseada no valor do R_0

A Equação (1) calcula a porcentagem da população que precisa ser vacinada para que a propagação da doença se estabilize:

$$\text{População precisando de vacina} = 1 - \frac{R_0 \text{ alvo}}{R_0 \text{ atual}} \times \text{População} \quad (1)$$

O modelo SEIR tem 4 componentes principais, Suscetíveis (S), Expostos (E), Infectados (I) e Recuperados (R) como mostrado na Figura 2. O número S é a parte da população que pode contrair a doença, E a parte da população que foi infectada, mas que não

infecta; I a parte de indivíduos capazes de transmitir a doença e R a parte de indivíduos que se tornaram imunes.

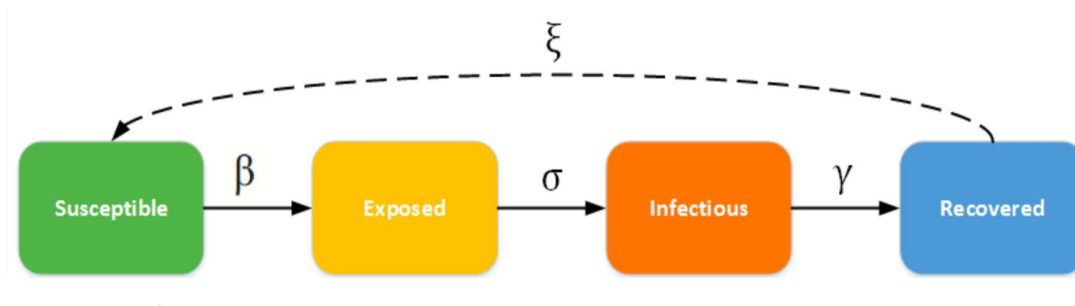


Figura 2. Ilustração do modelo SEIR e suas quatro componentes

Na Figura 2, β é a taxa de infecção que representa a probabilidade da doença se transmitir para uma pessoa suscetível; σ é a taxa de incubação, que representa a taxa na qual uma pessoa infectada se torna infecciosa; γ é a taxa de recuperação, determinada por 1D (onde D é a duração da infecção); e ξ é a taxa na qual uma pessoa recuperada se torna suscetível devido à baixa imunidade ou outros problemas de saúde relacionados.

Notas explicativas + gráficos

Para descrever a propagação da COVID-19 usando o modelo SEIR, o estudo fez algumas considerações:

1. O número de nascimentos e mortes continua o mesmo;
2. O período latente da doença é o período de infecção;
3. Pessoas que se recuperaram não ficaram doentes novamente no período calculado.

O estudo utilizou também o modelo de regressão (conjunto estatístico usado para estimar a variável alvo ou dependente se baseando nas variáveis independentes) linear e polinomial para prever os casos de Covid-19.

Vale frisar que os dados de séries temporais fornecidos pela Universidade John Hopkins, dos EUA, foram usados para análise de resultados empíricos durante um período de dois meses (30/01/2020 até 30/03/2020).

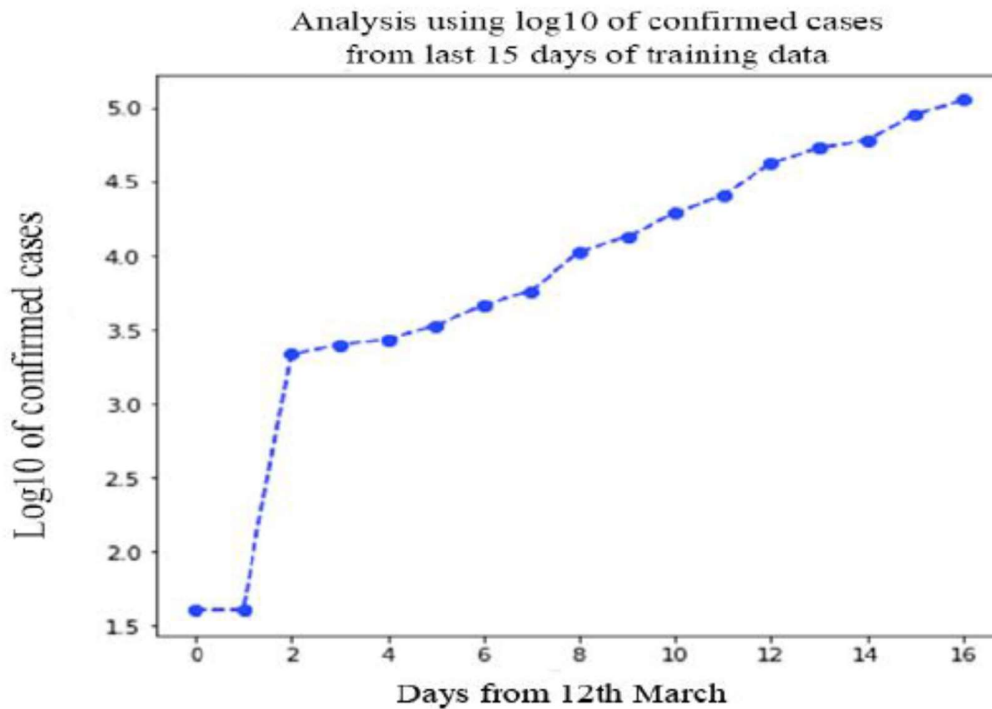


Figura 6. Mostra a tendência de crescimento linear na Índia entre os dias 49 e 65

A Figura 6 mostra o resultado de casos confirmados na escala \log_{10} versus os últimos 15 dias de dados de treinamento. A razão de serem usados os últimos 15 dias dos dados de treinamento se dá porque ele mostra o principal crescimento nos casos confirmados, como mostrado na Figura 4.

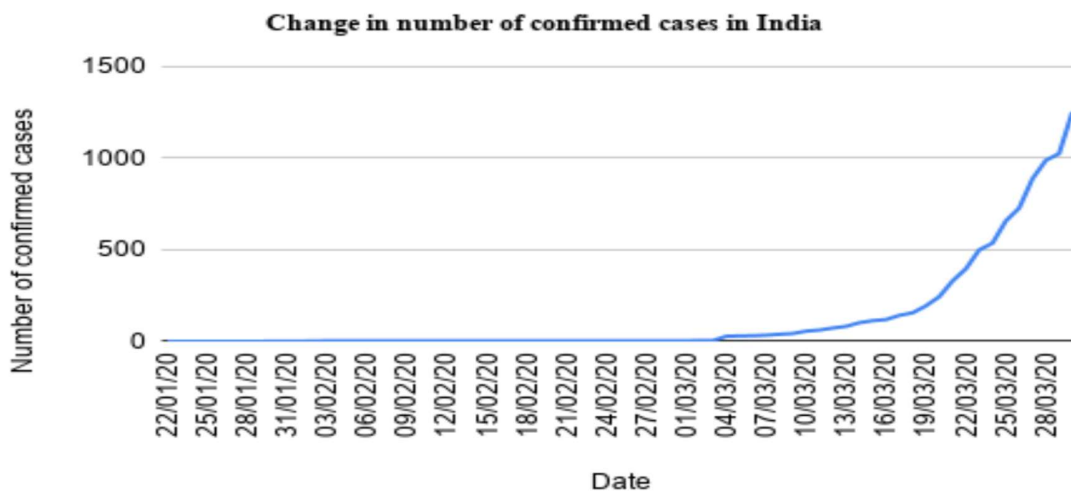


Figura 4. Mudança no número de casos confirmados de COVID-19 na Índia

Os dados foram treinados com o modelo SEIR e o resultado da curva de aproximação é mostrada na Figura 7. Enquanto o SEIR foi aplicado, foram adicionadas algumas

intervenções feitas pelo governo indiano, como quarentena e *lockdown*. Para isso, nós aplicamos uma função de decaimento que é mostrada na Equação (9), em que L é uma descrição da taxa de decaimento, T é o tempo e k é o parâmetro de forma.

$$\text{Função de decaimento} = \frac{1}{1 + (T/L)^k} \quad (9)$$

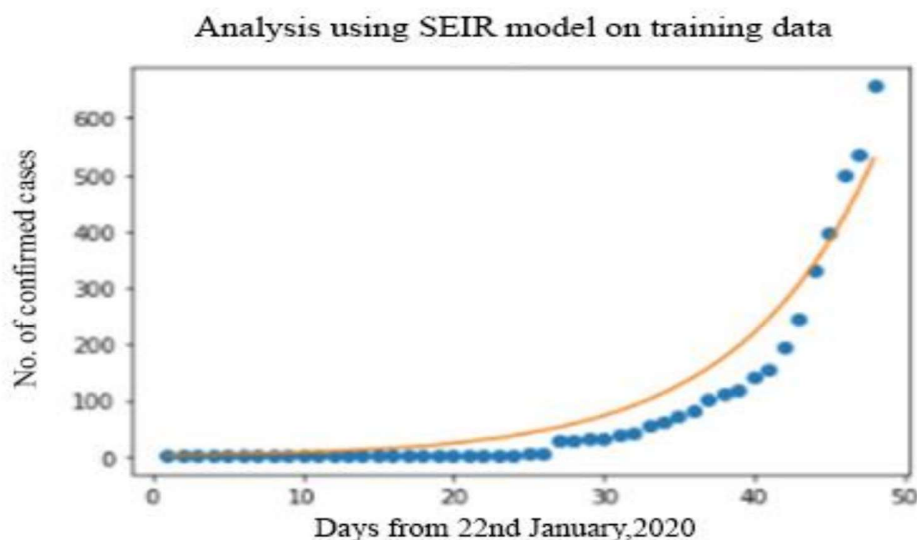


Figura 7. Análise usando o modelo SEIR, pontos azuis são os valores dos dados de treinamento e a linha amarela é a curva de aproximação do modelo.

REFERÊNCIAS

1. World Health Organization, 2020. Coronavirus disease 2019 (COVID19): situation report, 67.
2. Wu, Z. and McGoogan, J.M., 2020. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *Jama*.
3. Jiang, F., Deng, L., Zhang, L., Cai, Y., Cheung, C.W. and Xia, Z., 2020. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). *Journal of General Internal Medicine*, pp.1-5.
4. Ye, Q.H., Qin, L.X., Forgues, M., He, P., Kim, J.W., Peng, A.C., Simon, R., Li, Y., Robles, A.I., Chen, Y. and Ma, Z.C., 2003. Predicting hepatitis B virus–positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. *Nature medicine*, 9(4), pp.416-423.

5. Mai, M.V. and Krauthammer, M., 2016. Controlling testing volume for respiratory viruses using machine learning and text mining. In AMIA Annual Symposium Proceedings (Vol. 2016, p. 1910). American Medical Informatics Association.
6. Purcaro, G., Rees, C.A., Wieland-Alter, W.F., Schneider, M.J., Wang, X., Stefanuto, P.H., Wright, P.F., Enelow, R.I. and Hill, J.E., 2018. Volatile fingerprinting of human respiratory viruses from cell culture. *Journal of breath research*, 12(2), p.026015.
7. Nsoesie, E.O., Brownstein, J.S., Ramakrishnan, N. and Marathe, M.V., 2014. A systematic review of studies on forecasting the dynamics of influenza outbreaks. *Influenza and other respiratory viruses*, 8(3), pp.309-316.
8. Pirouz, B., ShaffieeHaghshenas, S., ShaffieeHaghshenas, S. and Piro, P., 2020. Investigating a Serious Challenge in the Sustainable Development Process: Analysis of Confirmed cases of COVID-19 (New Type of Coronavirus) Through a Binary Classification Using Artificial Intelligence and Regression Analysis. *Sustainability*, 12(6), p.2427.
9. More, G.D., Dunowska, M., Acke, E. and Cave, N.J., 2020. A serological survey of canine respiratory coronavirus in New Zealand. *New Zealand Veterinary Journal*, 68(1), pp.54-59.
10. Wu, C., Chen, X., Cai, Y., Zhou, X., Xu, S., Huang, H., Zhang, L., Zhou, X., Du, C., Zhang, Y. and Song, J., 2020. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA internal medicine*.
11. Mandal, S., Bhatnagar, T., Arinaminpathy, N., Agarwal, A., Chowdhury, A., Murhekar, M., Gangakhedkar, R.R. and Sarkar, S., 2020. Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical modelbased approach. *The Indian journal of medical research*.
12. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis*; published online Feb 19. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1).
13. Gupta, Rajan and Pal, Saibal K. Trend Analysis and forecasting of COVID-19 outbreak in India. Published online 30/03/2020 in preprint archive MedRxiv <https://www.medrxiv.org/content/10.1101/2020.03.26.20044511v1>