Publisher
ADSRS EDUCATION AND RESEARCH
www.adsrs.net

E-ISSN:2583-4487

Global Journal of Novel Research in Applied Sciences

2022 Volume 1 Number 2 July-Dec

Research Article

Covid-19

Prediction of the evolution of corona-virus using Machine Learning Technique

Kossi T.[1], Sahana S.[2*]

DOI: https://doi.org/10.58260/j.nras.2202.0106

[1] Tokpe Kossi, Dept. of Computer Science & Engineering, School of Engineering &Technology, Sharda University, Greater Noida, Uttar Pradesh, India.

[2*] Subrata Sahana, Dept. of Computer Science & Engineering, School of Engineering & Technology, Sharda University, Greater Noida, Uttar Pradesh, India.

The whole world has seen a change in habits in the past two years due to the discovery of the new corona family virus. The new corona virus is classified by The World Health Organization as Covid-19. Everywhere on the planet we are witnessing confirmed decision-making, curfew, restrictions on people, vaccinations, wearing of masks, etc. This paper aims to show how advancement of Machine Learning have provided excellent tools capable of reducing the number of infections while helping medicine in making decisions about testing infected people, reliability, accuracy of tests. We have improved the existing algorithms to produce software that can be able to do the prediction of evolution of corona virus. The implementation is done through Python language. We ensure that the results produced will be reliable and with fewer errors.

Keywords: Curfew, Python, Algorithms, Covid-19, Machine Learning, World Health Organization

Corresponding Author

Subrata Sahana, , Dept. of Computer Science & Engineering, School of Engineering & Technology, Sharda University, Greater Noida, Uttar Pradesh, India.
Email: tokpe@pg.sharda.ac.in

How to Cite this Article

Tokpe Kossi, Subrata Sahana, Prediction of the evolution of corona-virus using Machine Learning Technique. Glo.Jou.Nov.Res.App.Sci. 2022;1(2):1-6. Available From http://nras.adsrs.net/index.php/nras/article/view/6

To Browse

Manuscript Received	Review Round 1	Review Round 2	Review Round 3	Accepted
2022-11-03	2022-11-08	2022-11-10	2022-12-16	2022-12-30

Conflict of Interest	Funding	Ethical Approval	Plagiarism X-checker	Note
Nil	Nil	Yes	16%	

Introduction

The last decade has had a lot of revolutions in new technologies of computer science areas which keep revolutionizing the technological world thus involving the improvement of human existence.

Technology professions are not following the way of this evolution and the professionals must know that their job will be automated by Artificial Intelligence tools one day. Our research will clarify Machine Learning changes bring in terms of opportunities. In this research we will be focused on some applications of Machine Learning. By studying the existing algorithms models, we proposed a prediction model used in Machine Learning in order to see in real time the evolution of this pandemic of corona virus in the world. Dummy data and local file content is used for testing purpose to produce software as outcome of the work [1].

Brief introduction of Artificial Intelligence and machine learning discussed in the following section.

A. Artificial Intelligence and Machine Learning

1. Artificial Intelligence: Tasks usually done by humans require discernment and intelligence and Artificial Intelligence provide to computer the way to automate these tasks. With AI a combination with huge amounts of data and required algorithms is made in the goal to provide features or patterns to the software to be capable to learn automatically from them.AI includes many concepts based on technologies, methodologies, theories and methods [2].

2. Machine Learning: Machine learning by definition is a way to give computers programs the capability to learn without automated explicitly even if changes arrive in existing programs [3].

Machine Learning has following fields:

A. A process of decision: Process of prediction is in general used with Machine Learning algorithms. An estimation of a pattern of data is processed through inputs of the data and can be categorized as labelled or unlabelled [4].

B. A function of error: The prediction of the model is evaluated by error function. The validity of the model can be made by comparing existing examples [5].

The process of error function is shown in Fig.1.

Figure 1: Process error function

C. Process of Optimization of the model: Adjustment of the weights between the existing example and the estimated model is made. Evaluation and optimization in the same way the updates of the process are made for the sole purpose of finding an acceptable threshold of precision [6].

Figure 2: Adjustment of the weights

Literature Survey: In this section, we have shown what researchers about Machine Learning have done and not done.

A. What they have done?

1. Representation of the source of knowledge

Difference with the traditional machine learning was done by bringing the blocks into the new concept called informed machine learning.

This approach is made by considering the knowledge source, the representation of the knowledge and the integration into the tube of machine learning. Researchers describe the entire schema by using learning systems tools like results of simulation, logic rules, algebraic equations, theories [7].

2. Modelling and classification approaches of human machine dialogs systems

Researchers focused on studying human models of dialogs by providing a survey. Firstly, the establish the fundamentals of a dialog model.

Secondly, they make examination of the components of the system dialog model. They compared dialog models' based on their drawbacks and advantages. They made evaluation in details and provide system dialog metrics using databases. Existing issues of the dialog systems are also analysed and propositions for new directions of human dialogs systems are mentioned [8].

3. Application of data mining in the CRM

This method means Customer Relationship and is the knowledge used to maximize profits based from the learning or statistical method to instruct the strategic behaviour.

Support Vector Machine (SVM) has been used in recent years as an important tool in data mining and machine learning

Thus, SVM can be considered as solution for CRM problem [9].

4. Boosting semi-supervised learning

Limited number of labeled and unlabeled examples is used to show the importance of the problem.

Some advantages of this model are essentially [10:

1) Multiple unlabelled data can be used to improve this algorithm of supervised learning.

2) Iterative way to compute efficiently the learning algorithm

3) In classification training models Assumption of cluster and manifold are exploited.

A large number of unlabeled data are used as an empirical study that is made on different data sets and text categorization.

B. What they have not done?

These points follow-up is not considered by the researchers [11]:

1. Evaluation of the success of a model

The model used can be difficult for human to interpret it. Deep study can help in this situation to understand properly the model [20]

2. Accuracy and interpretability importance

The different interactions between models can be difficult to interpret, the accuracy of result can be difficult to evaluate. It can be difficult to get errors of accuracy.

3. Time and resource allocation

The resources that will be allocated must be studied and the time that the resources will consume must be studied also [18].

4. Quality of labeled data

The quality of labeled data must be studied. Researchers must check if they have quality data [19].

Flow Information in Informed Machine Learning.

The different ways of elements in Machine Learning are shown Fig.3. Here is the proposed architecture of our model [12]:

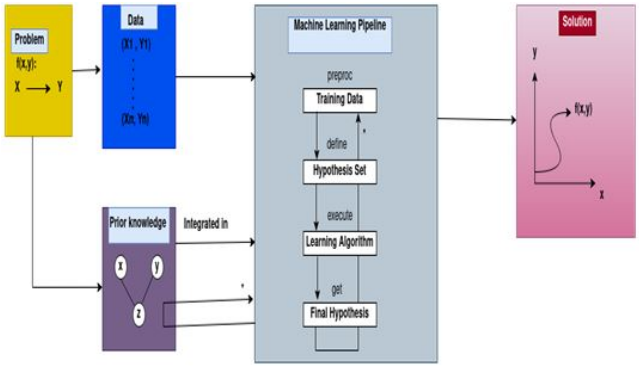

Figure 3: Technical architecture of Informed Machine Learning

1. Step 1: Problem

The problem is defined:

What do we want to do?

What information do we have?

What Statistics information are provided?

By analysing supplies X and Y and applying X to outputs Y, the problem is interpreted as a regression model [13].

Fig 4: Problem

2. Step 2: Information

Information is generated or collected, and algorithms are now used to try to approximate the unknown mapping.

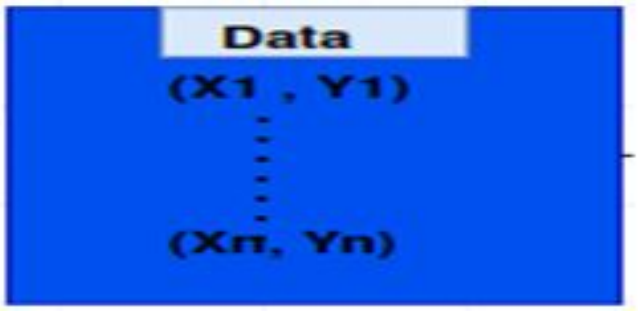

Figure 5: Data

3. Step 3: Prior knowledge

It appears as relevant information that is then verified.

The more formalised knowledge is, the easier it is to incorporate into machine learning [14].

These are the kinds of inquiries that need to be asked [17].

1. Which sources of information are incorporated?

2. Representation: What is the method for representing knowledge?

3. Integration: Where does knowledge fit into the learning process?

4. Step 4: Machine Learning Pipeline

Training Data: Extra independent variables are synthesized and used to create model parameters when matched with real data.

Hypothesis Set: A linkage from reduced numerical simulations to a few real-world observations or slightly elevated simulations can be performed through components of the system [15].

Protocol for learning: It is easy to embed a simulation straight into iterations. For example, by incorporating input from a dynamic simulation into learning, a corona prediction system can be enhanced [16].

Final Hypothesis: It combines the outcomes of simulations. Simulations, in particular, can be used to verify the findings of a trained model.

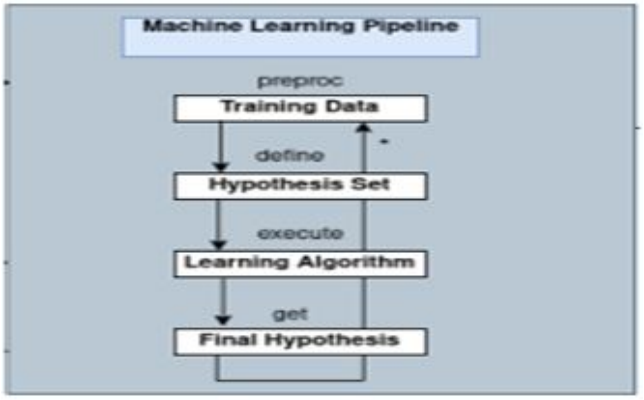

Figure 6: Final Hypothesis trained model.

5. Step 5: Solution

The expected result is provided. It can be a software or website or excel file containing data.

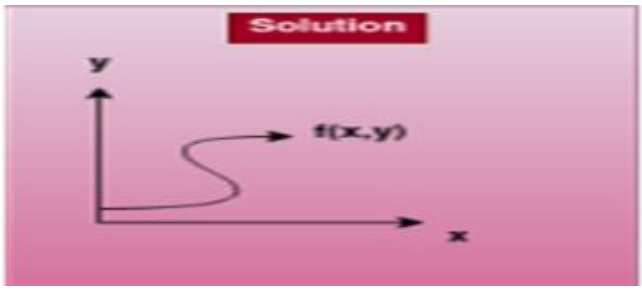

Figure 7: Solution

Corona prediction software using Python and Machine Learning

We use Python applying it with machine learning which goal is to provide prediction of corona evolution around the world.

The form of the equation used is in linear form: y=ax+b with variables a and b inputs.

We consider two files in Python language coronaCases.csv which has fields id and cases. Dummy data are used to fill this file as it is shown as below:

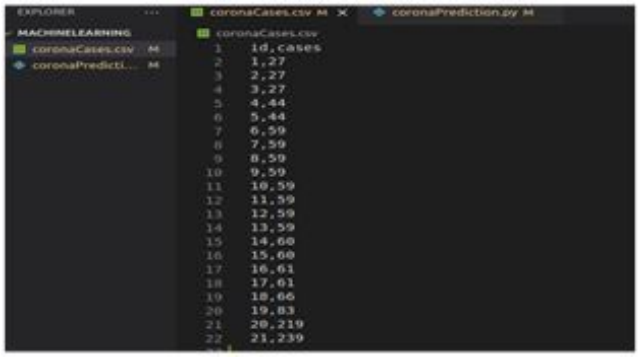

Figure 8: Cases of corona file

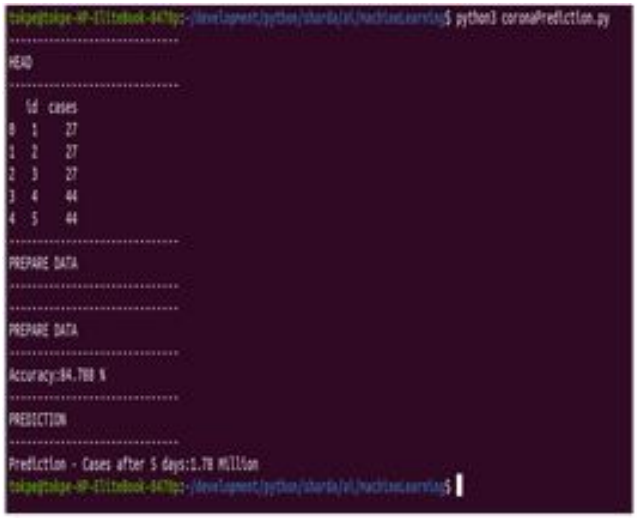

Figure 9: Results of execution

In our implementation we want to know the evolution of corona-virus in 5 next days. The execution is made as shown by the below screenshot: The head of the file is read by the program and Linear Regression model is used and provides an accuracy result of 84.788%. Our output matches with the prediction made on the code source. So, 1.78 Million cases will be the number of infected persons in next 5 days. The graph bellows shows the evolution in 5 next days the variation around the world of corona virus.

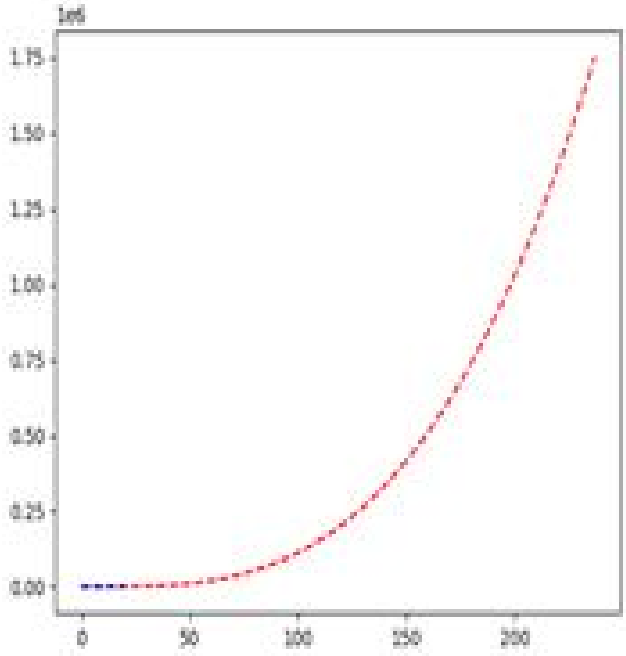

Figure 10: Five next day's corona-virus evolution graph

We can see in this graph that the number of cases of corona infection is less important at the start, then quickly grows and continues to grow,

Which represents the line of the actuality of the evolution of this pandemic. The prediction algorithm confirms perfectly the expected result.

Conclusion

Machine Learning brings enormous promise for making the world a better place to live. Humans' most crucial responsibility will be to guarantee that the emergence of Machine Learning does not go out of control. Machine Learning has an undeniable impact on the worldwide industry, despite its arguable benefits and drawbacks. Every day, it grows, ensuring the long-term viability of businesses. To succeed in many new age jobs, AI literacy and up skilling are unquestionably required. Throughout the paper demonstrated the importance of the Machine Learning in general in the fight against covid-19. We apply Machine Learning concepts to build a software which allows to predict the evolution of corona-virus. We achieved the goals fixed at the beginning and the software model developed is working correctly.

Reference

1. von Rueden, Laura, et al. "Informed Machine Learning--A Taxonomy and Survey of Integrating Knowledge into Learning Systems." arXiv preprint arXiv:1903.12394 (2019).

2. Liu, Bo, et al. "Apply support vector machine for CRM problem." 2007 International Conference on Machine Learning and Cybernetics. Vol. 6. IEEE, 2007.

3. Lu, Mei, and Fanzhang Li. "Survey on lie group machine learning." Big Data Mining and Analytics 3.4 (2020): 235-258.

4. Mallapragada, Pavan Kumar, et al. "Semiboost: Boosting for semi-supervised learning." IEEE transactions on pattern analysis and machine intelligence 31.11 (2008): 2000-2014.

5. Mallapragada, Pavan Kumar, Rong Jin, Anil K. Jain, and Yi Liu. "Semiboost: Boosting for semi-supervised learning." IEEE transactions on pattern analysis and machine intelligence 31, no. 11 (2008): 2000-2014.

6. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 2015 Jul 17;349(6245):255-60.

7. Kourou, Konstantina, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis, and Dimitrios I. Fotiadis. "Machine learning applications in cancer prognosis and prediction." Computational and structural biotechnology journal 13 (2015): 8-17.

8. Seko, Atsuto, Hiroyuki Hayashi, Keita Nakayama, Akira Takahashi, and Isao Tanaka. "Representation of compounds for machine-learning prediction of physical properties." Physical Review B 95, no. 14 (2017): 144110.

9. Yarkoni, Tal, and Jacob Westfall. "Choosing prediction over explanation in psychology: Lessons from machine learning." Perspectives on Psychological Science 12, no. 6 (2017): 1100-1122.

10. Menden, M. P., Iorio, F., Garnett, M., McDermott, U., Benes, C. H., Ballester, P. J., & Saez-Rodriguez, J. (2013). Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. PLoS one, 8(4), e61318.

11. Krittanawong, C., Virk, H. U. H., Bangalore, S., Wang, Z., Johnson, K. W., Pinotti, R., ... & Tang, W. H. (2020). Machine learning prediction in cardiovascular diseases: a meta-analysis. Scientific reports, 10(1), 1-11.

12. Rottondi, C., Barletta, L., Giusti, A., & Tornatore, M. (2018). Machine-learning method for quality of transmission prediction of unestablished lightpaths. Journal of Optical Communications and Networking, 10(2), A286-A297.

13. Chibani, Siwar, and François-Xavier Coudert. "Machine learning approaches for the prediction of materials properties." APL Materials 8, no. 8 (2020): 080701.

14. Uddin, Shahadat, Arif Khan, Md Ekramul Hossain, and Mohammad Ali Moni. "Comparing different supervised machine learning algorithms for disease prediction." BMC medical informatics and decision making 19, no. 1 (2019): 1-16.

15. Lin, W.Y., Hu, Y.H. and Tsai, C.F., 2011. Machine learning in financial crisis prediction: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(4), pp.421-436.

16. Zoabi, Y., Deri-Rozov, S., & Shomron, N. (2021). Machine learning-based prediction of COVID-19 diagnosis based on symptoms. npj digital medicine, 4(1), 1-5.

17. Lykourentzou, Ioanna, Ioannis Giannoukos, Vassilis Nikolopoulos, George Mpardis, and Vassili Loumos. "Dropout prediction in e-learning courses through the combination of machine learning techniques." Computers & Education 53, no. 3 (2009): 950-965.

18. Rose, S. (2018). Machine learning for prediction in electronic health data. JAMA network open, 1(4), e181404-e181404.

19. Korup, O., & Stolle, A. (2014). Landslide prediction from machine learning. Geology today, 30(1), 26-33.

20. Hindman, Matthew. "Building better models: Prediction, replication, and machine learning in the social sciences." The ANNALS of the American Academy of Political and Social Science 659.1 (2015): 48-62.