

# Contact Tracing Using ML

Mrs. M. S. Sandhya Rani<sup>1</sup>, D Sanjana<sup>2</sup>, N Varshitha<sup>2</sup>, M Shemith<sup>2</sup>

<sup>1</sup>Assistant Professor, Computer Science & Engineering, Anurag University, Telangana, India

<sup>2</sup>Student, Computer Science & Engineering, Anurag University, Telangana, India

## ARTICLE INFO

### Article History:

Accepted: 10 Oct 2023

Published: 30 Oct 2023

### Publication Issue

Volume 10, Issue 5

September-October-2023

### Page Number

624-628

## ABSTRACT

At the time of Corona, we are undergoing a dangerous situation due to wide spread of "Corona Virus". If a person is found to be infected with the Corona Virus, then It is critical to find the persons got contacted with the infected person. This can be done by Manual methods or numerical methods. Technology can be used for Contact tracing is a process that produces more efficient and accurate results than when the procedure is done manually.

Contact Tracing is onetype of the processes to identify the infected persons using machine learning. It can help to reduce the infected persons and can fasten the process of treating people who have been infected with a disease. We can save a lot of lives this way.

To find the movement of infected person, we use GPS data set that contains time and location of a person. The location data contains longitude and latitude coordinates.

**Keywords:** ALU, Adders, Subtractors, Borrow

## I. INTRODUCTION

There are lot of infectious viruses which can be spread from one person to another person easily. CoronaVirus is one of the viruses among them. If a person tests positive for Covid-19, it's critical to track down the other persons who were exposed to the virus through the sick person.

We track the activities of patients who have received treatment in the last 14 days in order to detect contaminated people. Contact tracking is the term used to describe this process. This can be

accomplished using either manual or numerical methods.

In the process of contact tracing, technology can be employed, and it is more efficient and accurate than manual techniques.

The Technology that is used in the process was Clustering, which is a subclass of Machine Learning. It is used to stop the spread of viruses and even save more lives. This can be done by collecting data set of different persons at different locations with their timestamp.

Our data set contains 4columns such as:

- 1: Name of a person
- 2: Timestamp
- 3: Latitude.
- 4: Longitude

Here, We use Jupyter Notebook for analyzing and running the code .The installation of Jupyter

Notebook was simple and easily downloaded from the Internet.

Applying DBSCAN Algorithm results in the formation of clusters. We analyze the data using the help of graph. The piece of python code which returns the contacted persons.

Jupyter comes with the anaconda package. We can easily install different libraries in the Anaconda command prompt and this makes our project simple.

## II. LITERATURE SURVEY

Author(s)	Method	Advantages	Disadvantages
Tao Feng, Sirui Song	Deep Reinforcement Learning algorithm	Proposed model can estimate the spread of the virus through contacts between individuals.	Absence of imposition of constraint for control- action selection.
	SVM classifier	It can be applied anywhere, without prior training or calibration	Cannot estimate the distance between the patient and possible contacts
	Four classification methods: decision tree (DT), linear discriminant analysis (LDA), naive Bayes (NB), and k nearest neighbors (KNN).	Estimate the spread of the virus through contacts between individuals.	Cannot reframe the risk classification problem as a risk regression problem

## III. METHODOLOGY

### MACHINE LEARNING

Machine learning is the study of algorithms that improve over time as they gain experience from data. It has become a common technique in practically every operation that demands information extraction from massive data sets over the previous few decades. We are surrounded by machine learning technology: search engines learn how to deliver the best results to

us (while placing pro table advertisements), and anti-spam software is always improving. Our email messages are filtered by a machine-learning algorithm, and credit card transactions are safeguarded by software that learns how to detect fraud. Face detection in digital cameras is assisted by machine learning, and intelligent personal assistant apps on smartphones learn to identify spoken instructions. Accident-prevention systems in vehicles are focused on machine learning algorithms and artificial intelligence.

In addition to bioinformatics, daily uses medicine, and astronomy, machine learning is widely employed in scientific applications. In contrast to more traditional uses of computers, all of these applications have one thing in common, A human programmer cannot provide an explicit, precise detailed specification of how such tasks should be accomplished in these instances due to the complex of the patterns that need to be recognised. Many of our skills are obtained through learning from our experiences, as evidenced by intelligent human beings (rather than following instructions given to us). The goal of machine learning tools is to provide programs/algorithms the ability to learn and adapt.

Time (hour, day, month, year) and location (latitudes and longitudes).are some of inputs to our algorithms . The type of crime that is most likely to have occurred is the output. We use a variety of classification algorithms on Our Dataset, including DBSCAN algorithm.

### 3.1.1 Machine Learning

ML is the study of computer algorithms that improve automatically through experience and no of times it runs. It is seen as a part of artificial intelligence (AI). Machine learning algorithms build a model based on the sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed by a programmer. Machine learning algorithms are used in many applications, such as email filtering and computer vision, where it is difficult to develop conventional algorithms to perform the needed tasks.

#### DBSCAN ALGORITHM

DBSCAN is a density-based spatial clustering of applications with noise. This can detect arbitrary-shaped clusters as well as clusters with noise (i.e. outliers).

The DBSCAN algorithm works on the principle that a point belongs to a cluster whether it is close to many other points from such a cluster.

DBSCAN has two important variables:

eps: The distance between neighbourhoods that defines them. If the distance between two points is below or similar to eps, they are termed neighbours.

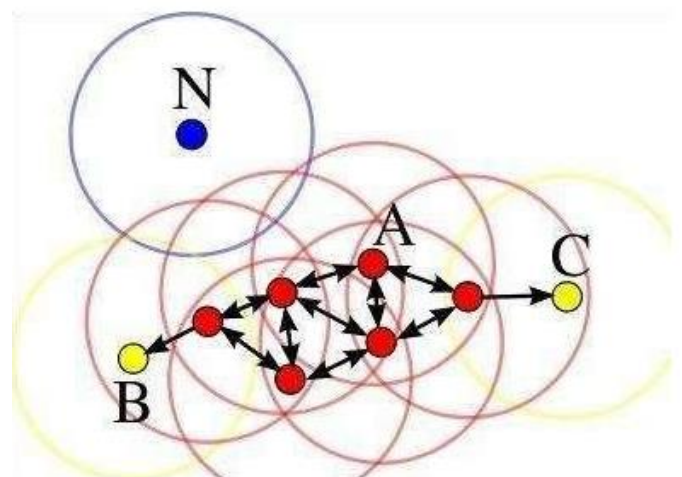
minPts: The minimal number of the data points required to define a cluster is minPts

Points are classified as core, boundary, or outlier two of the following two parameters:

- Core point: If there are minimum minPts number of points (along with the point itself) in its neighbourhood region with radius eps, it is a core point.
- Border point: A point would be a border point if it can be reached from a core point and also the number of points in its immediate vicinity is fewer than minPts.

Outlier: If a point isn't a core point and can't be reached from any other core points, it's called an outlier. The following illustrations may help to clarify these topics.

The image below is sourced from Wikipedia.



The value of minPts is 4 in this example. Because there are at least four points of radius eps within its

surrounding(neighborhood) region, red points are considered core points. The circles there in figure depict this area. Because they may be reached from a core point and have fewer than four points in their neighbourhood, the yellow points are considered very border points. The term "reachable" refers to points that are within a certain radius of a central point. Within the neighbourhood of points B and C, there are two points (along with the point including itself) ( the surrounding area with which a radius of eps). Finally, since N is not a core point and cannot be visited from one, it is an outlier.

How the algorithm works :

- Initially, minPts and eps are determined.
  - A initial point will be chosen at random from its surrounding area using radius eps. If there should be at least minPts number of points in the neighbourhood, that point is designated as a core point, and a cluster formation begins with it. If this is not the case, the point is labelled as noise. When the development of clusters (let's call it cluster A) begins, every point in the vicinity of the starting point becomes a member of cluster A. If the newly added points are already core points, the points in their immediate vicinity are likewise added to the cluster A. A point that has been labelled as noise can be revisited and become part of a cluster.
  - The next step in this process is to select another point at random among those that were not visited there in previous steps. Then follow the same instructions as before
  - When all of the points have been covered, the process is finished.
  - A distance measuring approach, similar to that used in the k-means algorithm, is used to determine the distance between each pair of points. For this, the Euclidean distance approach is often utilised.
- DBSCAN method is able to obtain high density areas and differentiate them over low density regions utilising these steps.

#### 4. RESULTS

This system can predict the infected persons close to the person with disease

```
#will return the infected name back based on the inputed id get_infected_names('Alice')
```

output:

```
['Judy', 'Carol', 'David', 'Frank', 'Erin', 'Bob', 'Grace', 'Ivan', 'Heidi']
```

#### IV. CONCLUSION

So, finally I conclude by describing that, this project Contact tracing is one of the techniques we can utilise technology save the lives and also provide people the treatment they need as soon as possible. Government and medical personnel frequently have accessibility to the GPS coordinates of some patients. The approach we went through in this study is almost the same as the method officials go through to collect probable diseases. Fortunately, packages like sklearn allow us to apply existing models on our datasets to generate results.

The creation of a contact tracking system is among the most crucial steps to do in order to restore the global economy as safely as feasible. Contact tracing is our sole option for isolating the viruses when it occurs when social distancing is loosened, and this is especially useful in the case of coronavirus.

#### V. FUTURE SCOPE

The goal of our project is to reduce manual work by adding new technologies like machine learning/artificial intelligence. This is the starting stage of project applications. As we move further one can explore and add more number of technologies to the manual projects which reduce human stress leading to the development. The Contact Tracing is one of such example which adds machine learning for finding the contacts easily and this helps the humans in reduce the work .

## VI. Predicting Future Infectious Diseases

Like Corona Virus, more number of contagious diseases can be easily identified by new technologies. Our model is very helpful in the case of finding the contacts easily.

This reduces the human work and can be updated with adding more technology for further projects.

Also, It plays an important role in the health care system. Not only in the healthcare system, the technologies can be added in any field, where the human work is more

## VII. REFERENCES

- [1]. <https://machinelearningmastery.com/clustering-algorithms-with-python>
- [2]. <https://www.kdnuggets.com/2020/04/dbscan-clustering-algorithm-machine-learning.html>
- [3]. <https://thecleverprogrammer.com/2021/02/03/dbscan-clustering-in-machine-learning>

### Cite this article as :

Mrs. M. S. Sandhya Rani, D Sanjana, N Varshitha, M Shemith, "Contact Tracing Using ML", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 10 Issue 5, pp. 624-628, September-October 2023. Available at doi : <https://doi.org/10.32628/IJSRST52310569>  
Journal URL : <https://ijsrst.com/IJSRST52310569>