International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Accuracy of Machine Learning Based Covid19 Immunity and Survival Analysis

[1]G. Deepthi, [2]Dr. R. Kaviarasi

[1]PG Student, Department of Computer Applications, University College of Engineering, BIT Campus-Anna University, Tiruchirappalli-620024
[2]Assistant Professor, Department of Computer Applications, University College of Engineering, BIT Campus-Anna University, Tiruchirappalli-620024

ABSTRACT:

The aim of this project was to conduct a survival analysis to establish the variability in survivorship of patients.It shows Immunity Analysis,vaccination & its details, state wise through an graphic model use the concept of supervised machine learning to analyse the data and predict the future data-based situation. Using python with Machine Learning using different modules, Packages, Libraries and different techniques related to modeling. Data mining are used for the past or present in order to forecast future trends based on the kind of data prediction and analytical techniques. Using Tableau for data visualization to predict immunity analysis, vaccination details and covid cases. To find patients immunity level, immunity used in this project are White Blood Cell (WBC), Red Blood Cells (RBC), Hemoglobin (HB), Mean Corpuscular Volume(MCV).Mean Corpuscular Hemoglobin (MCH) and Platelets. The WBC which fight infections and other diseases. There are five types of white blood cells. A CBC test measures the total number of white cells in our blood. A different test called a CBC with differential measures the number of each type of white blood cells. The RBC which carry oxygen from your lungs to the rest of your body. The Hemoglobin a protein in red blood cells that carries oxygen from your lungs to the rest of your body. The MCV stands for Mean Corpuscular Volume it use a measure of the average size of your red blood cells. The MCH stands for Mean Corpuscular Hemoglobin, It tells the average quantity of hemoglobin in one red blood cell. The Hemoglobin is present in red blood cells. It transports oxygen into the tissues of human body. Platelets are the cells that circulate our blood and bind together when they recognize the damaged blood vessels.

Keywords: Covid-19, Immunity, Algorithms, Prediction, Analysis

1. INTRODUCTION:

Coronavirus are a type of virus. There are many different kinds, and some cause disease. COVID-19 can be severe, and has caused millions of deaths around the world as well as lasting health problems in some who have survived the illness. The coronavirus can be spread from person to person. It is diagnosed with a test. The best way to protect yourself is to get vaccinated and boosted when you are eligible, follow testing guidelines, wear a mask, wash your hands and practice physical distancing.

Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus. Most people infected with the virus will experience mild to moderate respiratory illness and recover without requiring special treatment. However, some will become seriously ill and require medical attention. Older people and those with underlying medical conditions like cardiovascular disease, diabetes, chronic respiratory disease, or cancer are more likely to develop serious illness. Anyone can get sick with COVID-19 and become seriously ill or die at any age.

The best way to prevent and slow down transmission is to be well informed about the disease and how the virus spreads. Protect yourself and others from infection by staying at least 1 metre apart from others, wearing a properly fitted mask, and washing your hands or using an alcohol-based rub frequently. Get vaccinated when it's your turn and follow local guidance.

The virus can spread from an infected person's mouth or nose in small liquid particles when they cough, sneeze, speak, sing or breathe. These particles range from larger respiratory droplets to smaller

aerosols. It is important to practice respiratory etiquette, or example by coughing into a flexed elbow, and to stay home and self-isolate until you recover if you feel unwell.

Coronavirus disease 2019 (COVID-19) was first reported in Wuhan, China, in 2019 On the 7th of January, the Chinese authorities identified a novel coronavirus, temporally named 2019-nCoV, as the cause of these cases.

Weeks later, the WHO declared the rapidly spreading COVID-19 outbreak as a Public Health Emergency of International Concern on the 30th of January 2020. It wasn't until the following month, however, on the 11th of February that the novel coronavirus got its official name - COVID-

19. Nine days later, the US Centers for Disease Control and Prevention (CDC) confirmed the first person to die of COVID-19 in the country. The individual was a man in his fifties who lived in Washington state.

In this project we can use immunity to predict the patients immunity level and Analyze the conformed cases, recovered cases, vaccinated details and future prediction.. For predicting purpose 3 Algothims can be implemented that is Support Vector Algorithm, Random Forest Algorithm and Linear Regression Algorithm. The random forest algorithm can predict the accurate results.

The immunity can classify the patient's Complete Blood Count(CBC) by the following normal person immunity level is WBC is 4000 to 10000 mcl(cells/microliter),RBC is 3.8 to 5.9 mcl(cells/microliter),HB is 12 to 17 dl(grams/deciliter),MCV is 83 to 101(femtoliters/cell),MCH is 27.0 to 32.0 pg(picograms) and Platelets is 150000 to 400000(platelets/microliter).

The Complete Blood Count(CBC) by the following covid affected patients immunity level is WBC is 4000 to 8000 mcl(cells/microliter),RBC is 3.8 to 5.1 mcl(cells/microliter), HB is 9.2 to 16 dl(grams/deciliter),for men15.5 dl and for women 13.3 dl, MCV is 83 to 93(femtoliters/cell),MCH is 26 to 30 pg(picograms) and Platelets is 150000 to 300000(platelets/microliter).

2. LITERATURE SURVEY:

[1]. J E T Akinsola,June 2017 ,"Supervised machine learning algorithms: classification and comparison[34]";

This paper determines the most efficient classification algorithm based on a clinical data-set. Seven supervised machine learning algorithms were considered concluding SVM (Support Vector Machines) followed by RF (Random Forests) that were found with most prediction and accuracy [34], June 2017.

[2]. Hetal Bhavsar, 2012, "A comparative study of training algorithms for supervised machine learning";

To understand the training algorithms for supervised machine learning in International Journal of Soft Computing and Engineering (IJSCE), 2(4):2231–2307, 2012.

[3]. K. Vijiyakumar, 2019, "Random forest Algorithm for the prediction";

Random forest Algorithm used for classification and regression.The random forest accuracy level is higher,IEEE international conference,2019.

G. Kesavaraj G K,2013, "A study on classification techniques in data mining [23]"; The experimental results state that it is difficult to choose one algorithm superior to another. It summarizes that classification algorithms are strictly confined to their problem domain [23],2013

[4]. Ritika Jain,2020, "Linear Regression";

The Linear Regression Analysis of COVID- 19 using Machine Learning Algorithms (IEEE Publishers),2020 International Conference.

[5]. Vansh Jatana,June 2019, " Machine learning";

The machine learning is a branch of artificial intelligence that allows computer systems to learn directly from examples, data, and experience, June 2019.

[6]. Khushbu Kumari,2018, "Linear regression in machine learning";

Linear regression is a statistical procedure for calculating the value of a dependent variable from an independent variable. Linear regression measures the association between two variables.Journal of the practice 4(1):33, January 2018.

[7]. Priya Pedamkar,2007, "Support Vector Regression (SVR)":

Support Vector Regression attempts to minimize the generalization error bound so as to achieve generalized performance,11(10), November 2007.

3. DATA VISUALIZATION TOOLS:

1. Tableau:

Tableau is a powerful and fastest growing data visualization tool used in the Business Intelligence Industry. It helps in simplifying raw data in a very easily understandable format. Tableau helps create the data that can be understood by professionals at any level in an organization. It also allows non-technical users to create customized dashboards. Data analysis is very fast with Tableau tool and the visualizations created are in the form of dashboards and worksheets. For a clear understanding, data analytics in Tableau tool can be classified into two section.

Tableau is an online sharing tool of Tableau. Its functionalities are similar to Tableau Server, but the data is stored on servers hosted in the cloud which are maintained by the Tableau group.There is no storage limit on the data that can be published in the Tableau Online. Tableau Online creates a direct link to over 40 data sources that are hosted in the cloud such as the MySQL, Hive, Amazon Aurora, Spark SQL and many more.

To publish, both Tableau Online and Server require the workbooks created by Tableau Desktop. Data that is streamed from the web applications say Google Analytics,

Salesforce.com are also supported by Tableau Server and Tableau Online.

2. Pycharm:

PyCharm is the most popular IDE and It shows Data visualization in graphical view.It is used for Python scripting language. Lets give an introduction to PyCharm and explains its features.

PyCharm offers some of the best features to its users and developers in the following aspects,

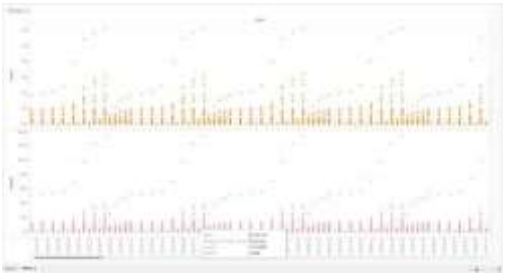
- Code completion and inspection
- Advanced debugging
- Support for web programming and frameworks such as Django and Flask

PyCharm is one of the most famous Integrated Development Environment (IDE) for Python, developed by a Czech organization called JetBrains. The integrated environment comes with the platform for Code Analysis, Graphical Debugger, Unit Tester and Version Control Systems, etc. It also supports web development with the Django framework.

4.EXPERIMENTAL RESULTS:

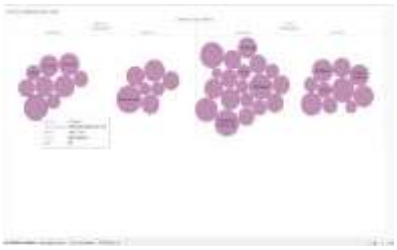
By using Pycharm and Tableau can be used for analysis purpose and graphical representation for analysis like

4.1 Cured and Death Rate



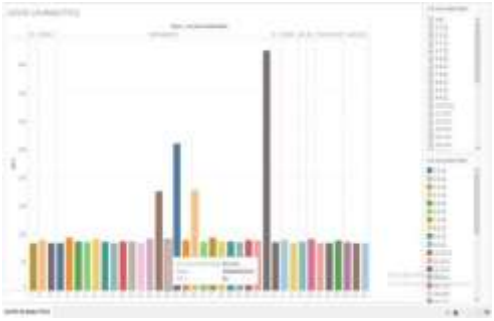
4.2 side-by-side circles

4.3 Vaccinated Date

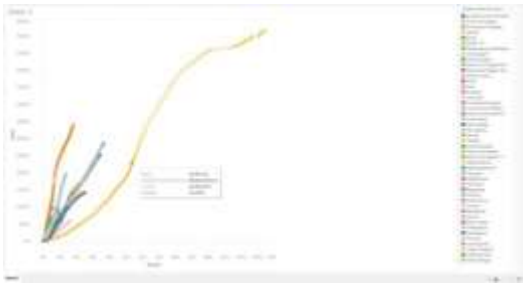


4.4 Packed Bubbles

Analyze the immunity details, conformed cases, recovered cases, affected places, vaccinated details.



4.5 Side-by-side Bars



4.6 Line Graph

4.7 Tree Maps

5. CONCLUSION:

we can use immunity analyze to predict the patients immunity level and using Tableau tool to import csv files to Analyze the conformed cases, recovered cases, vaccinated details, affected places, immunity details etc..

Thus the above project mentioned, mostly old age people can easily affected by the virus and vaccinated people can be saved from this disease and upcoming work is to identify which machinelearning algorithms can gives the accurate results.

Reference:

1. Hetal Bhavsar, Amit Ganatra .A comparative study of training algorithms for supervised machine learning [14]; International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-4, September 2012.

2. K. VijiyaKumar; B. Lavanya; I. Nirmala; S. Sofia Caroline.Random forest Algorithm for the prediction, IEEEInternational Conference,2019.

3. Linear Regression Analysis of COVID-19 using Machine Learning Algorithms (IEEE Publishers),2020 International Conference.

4. Hetal Bhavsar and Amit Ganatra. A comparative study of training algorithms for supervised machine learning. International Journal of Soft Computing and Engineering (IJSCE), 2(4):2231–2307, 2012.

5. Osisanwo F.Y, Akinsola J.E.T, Awodele O, Hinmikaiye J. O, Olakanmi O, Akinjobi J. "Supervised Machine Learning Algorithms: Classification and Comparison". International Journal of Computer Trends and Technology (IJCTT) V48(3):128-138, June 2017.

6. G. Kesavaraj, S.Sukumaran.A study on classification techniques in data mining [23],2013.

7. Sina F Ardabili, Amir Mosavi, Pedram Ghamisi, Filip Ferdinand, Annamaria R Varkonyi-Koczy, Uwe Reuter, Timon Rabczuk, and Peter M Atkinson. Covid-19 outbreak prediction with machine learning. Available at SSRN 3580188, 2020.

8. Ali Al-Hazmi. Challenges presented by MERS corona virus, and SARS corona virus to global health. Saudi journal of biological sciences, Publisher: Elsevier, 23(4):507–511, 2016.

9. Vansh Jatana, The machine learning is a branch of artificial intelligence that allows computer systems to learn directly from examples, data, and experience, June 2019.

10. H. Bansal, K. Sharmi, A Review study on various algorithms of machine learning, Journal of emerging technologies and innovative research, April 1,2020.