JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 40 - 45
https://publishoa.com
ISSN: 1309-3452

Prediction Of Covid-19 Infection Based on Lifestyle Habits Employing Random Forest Algorithm

Farooq Sunar Mahammad[1*], P. Bhaskar[2], A Prudvi[3], N Yugandhar Reddy[4], P Jaswanth Reddy[5], Harshavardhan[6], S Mahammad Ali[7]

[1-7]Department of Computer Science Engineering, Santhiram Engineering College, Nandyal
Email: hod.cse@srecnandyal.edu.in

ABSTRACT

A data of four factors contributing to covid-19 infection and whether the person has been infected or not is taken as input to train a Random Forest Algorithm. Factors like wearing mask, regularity of exercise, consumption of pepper and area population density are taken as input factors and the status of covid-19 infection is taken as output parameter. A random forest model trained with at least 50 data instances will become a powerful predictive model for assessing the risk of covid-19 infection.

Keywords: Random Forest, Covid-19, Prediction, Machine learning

INTRODUCTION

The covid-19 pandemic shows the huge impact on human life and the economy. Covid is a wakeup call to consider health as the highest priority to the global community. The human body is a complex system with different departments that serve various functions. The immune system is the body's defense system, designed to ward off what it sees as outside invaders. A healthy immune system consists of two components: the innate system you are born with and the adaptive immune system that develops when exposed to those outside invaders. Both works together to keep a person healthy. In fact, the immune system is the most powerful weapon against infections. While it is crucial to mention hygiene standards like washing of hands frequently, especially if the person has travelled by public transport. It is important to use an alcohol sanitizer, in case the person is travelling to disinfect their hands, and also to wear a mask and to avoid touching hands or mouth.

WORKING

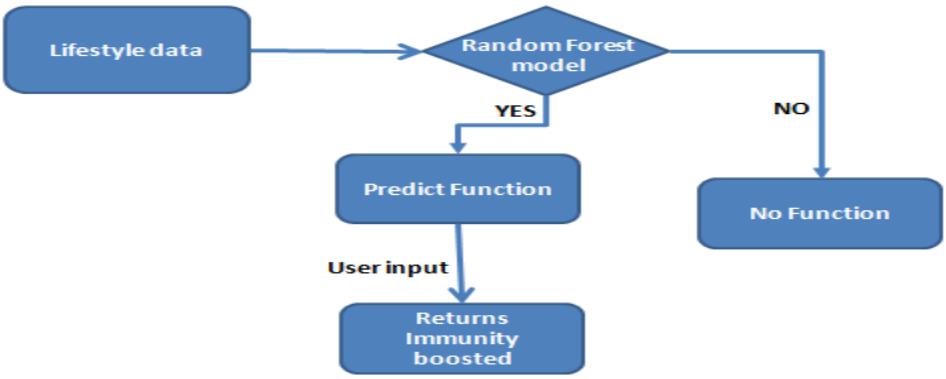

Figure 1 Block Diagram

Proposed system:

In this paper Random Forest algorithm with six input parameters and one output parameter is presented for covid-19 infection possibility. Based on binary states of 1 and 0 for the presence of positive and negative lifestyle practices respectively, the possibility of covid-19 infection is predicted. The lifestyle data is collected and given as input to Random Forest model,

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 40 - 45
https://publishoa.com
ISSN: 1309-3452

Random Forest model predict function analyze and predict the user input and returns immunity boosted.

ARTIFICIAL INTELLIGENCE

The algorithmic flow presented in this paper falls under the field of Artificial Intelligence and Machine Learning. Presented below is an outline of those technologies.

Artificial intelligence is brain power established by machines, not like the natural intelligence displayed by human beings and animals. The distinction between the former and the last groups is often exposed by the ellipsis chosen. Strong AI is generally called as artificial general intelligence (AGI) and usual intelligence is called as artificial biological intelligence (ABI). We can say that AI is the study of Intelligent Agents and any device that acts on the environment and increase the chances to reach goals. Colloquially, artificial intelligence is frequently used to describe machines that mimic cognitive functions that humans associate with the human mind, such as learning and problem solving.[3]

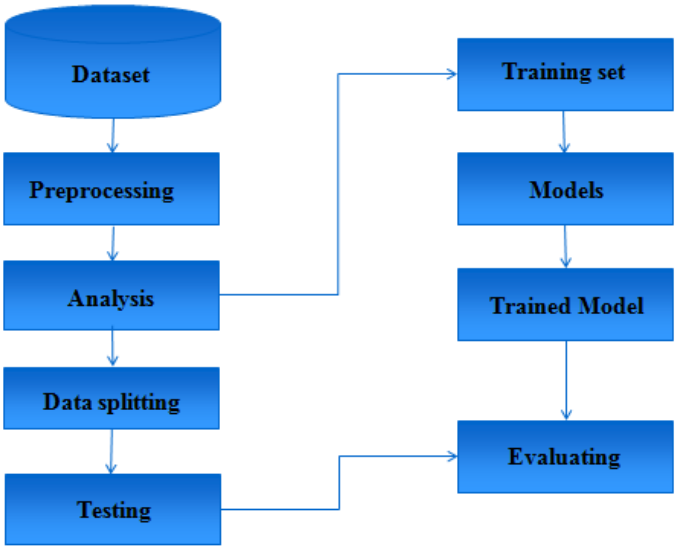

Figure 3 Flow Chart

Artificial intelligence lookup has been remoted into sub-handles that often forget about to communicate with one another. These sub-fields rely on specialized contemplations, like unique goals (for instance mechanical science or AI), the utilization of particular apparatuses (rationale or artificial brain organizations), or profound philosophical contrasts. Sub-fields have moreover been established on pleasant elements (specific agencies or crafted with the aid of precise scientists.[3]

AI techniques have experienced a huge demand in handling large amounts of data and theoretical understanding; and AI techniques have become an essential part of the technology industry, helping to solve many challenging problems in computer science, software engineering and operations research.[3]

MACHINE LEARNING

Machine Learning is substantially grounded on experience and it improves without human intervention by using the data, which may labeled or unlabeled. We can say that Machine Learning is the sub field of Artificial Intelligence. A Model can be make grounded on some sample data well known as training or labeled data which is used to read or assessment devoid of being explicitly programmed in Machine Learning. These Algorithms substantially used in different areas like Healthcare, Speech Recognition, Image processing, education, service where the use of conventional algorithm won't give stylish reports [4].

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 40 - 45
https://publishoa.com
ISSN: 1309-3452

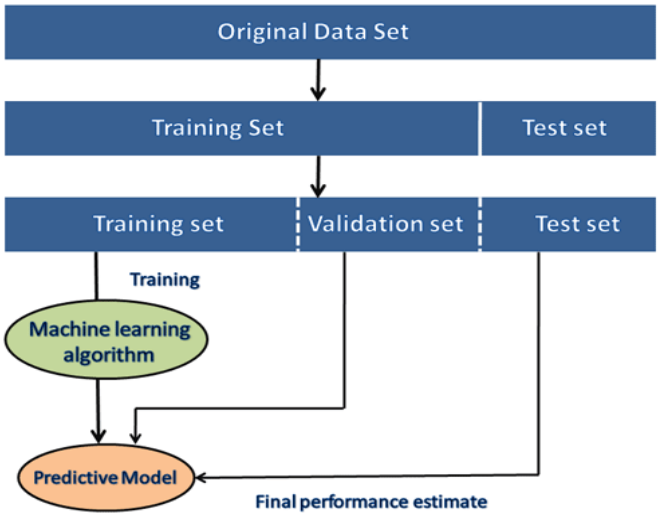

Figure 4 Machine learning dataset

A subset of computing device getting to know is cautiously related with computational statistics, which makes a speciality of making predictions the use of computers; on the other hand now no longer all system reading is statistical studying. The have a appear at of mathematical optimization offers you methods, thinking and alertness area names to the sphere of machine learning. Data mining is a related vicinity of have a seem at, specializing in exploratory statistics comparison by using unsupervised studying. In its utility at some point of organisation problems, laptop mastering is likewise known as predictive analytics. [4]

RESULTS AND DISCUSSION

Running/Push-ups	Ginger Garlic	No Stress	No Smoking	Mask Wearing	No Alcohol Consumption	Covid Infection
1	1	1	0	1	1	0
1	1	1	1	1	1	0
1	0	1	0	1	1	1
1	1	1	0	1	1	1
0	0	1	1	0	1	0
0	0	1	1	0	1	1
0	0	1	1	0	0	0
0	1	0	1	0	0	0

Figure 6 Input Sample Dataset

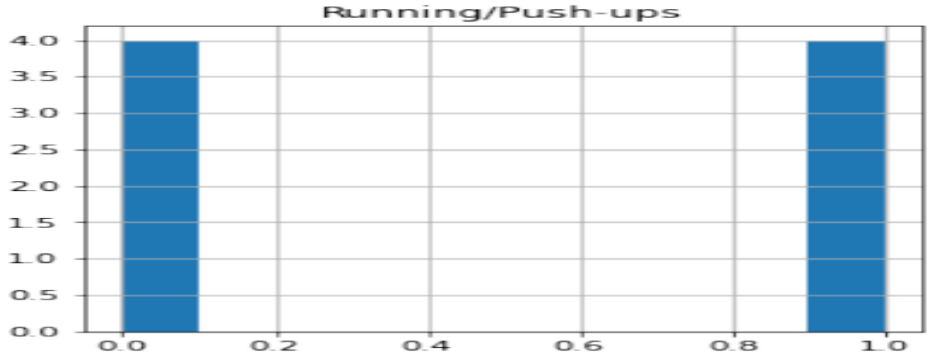

Figure 7 Running/Push ups

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 40 - 45
https://publishoa.com
ISSN: 1309-3452

Figure 8 Ginger Garlic

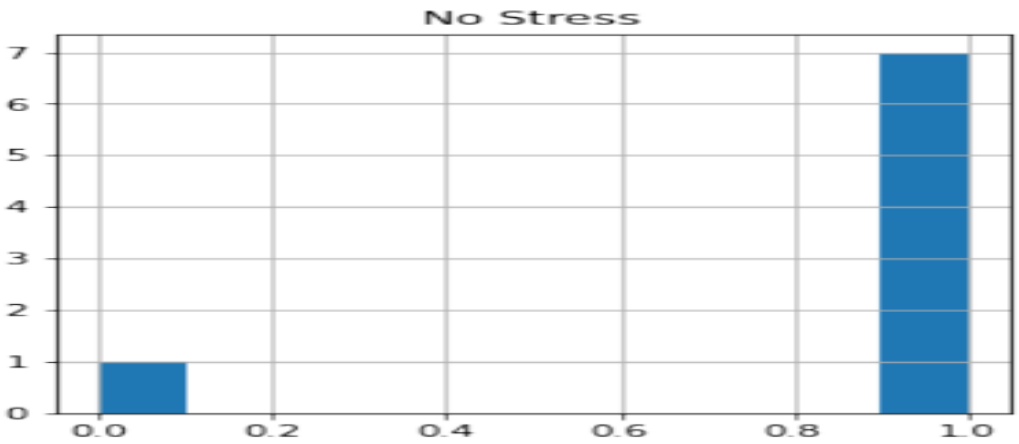

Figure 9 No stress

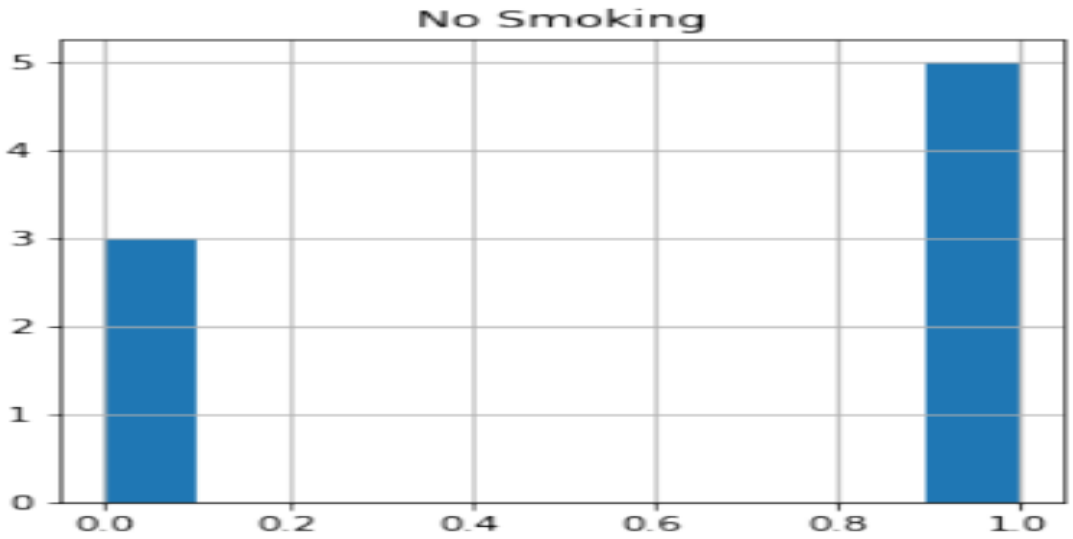

Figure 10 No smoking

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 40 - 45
https://publishoa.com
ISSN: 1309-3452

Figure 11 Mask wearing

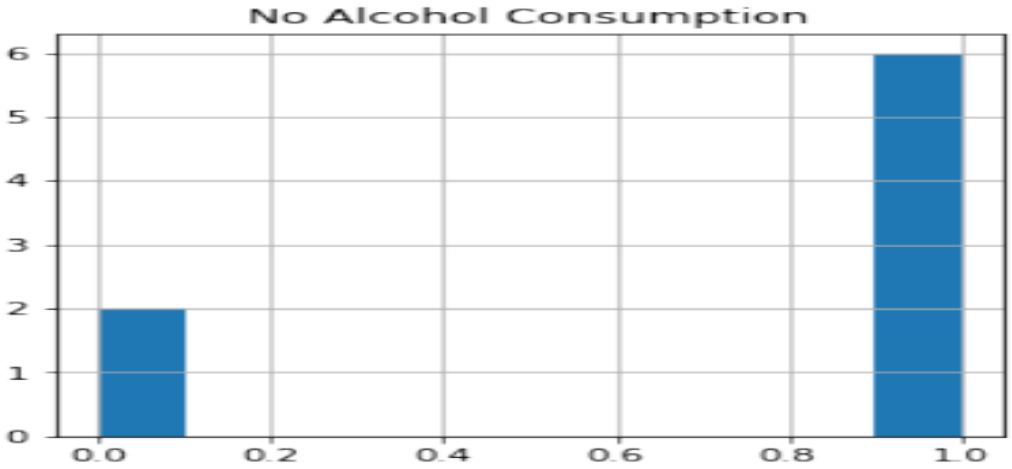

Figure 12 No alcohol consumption

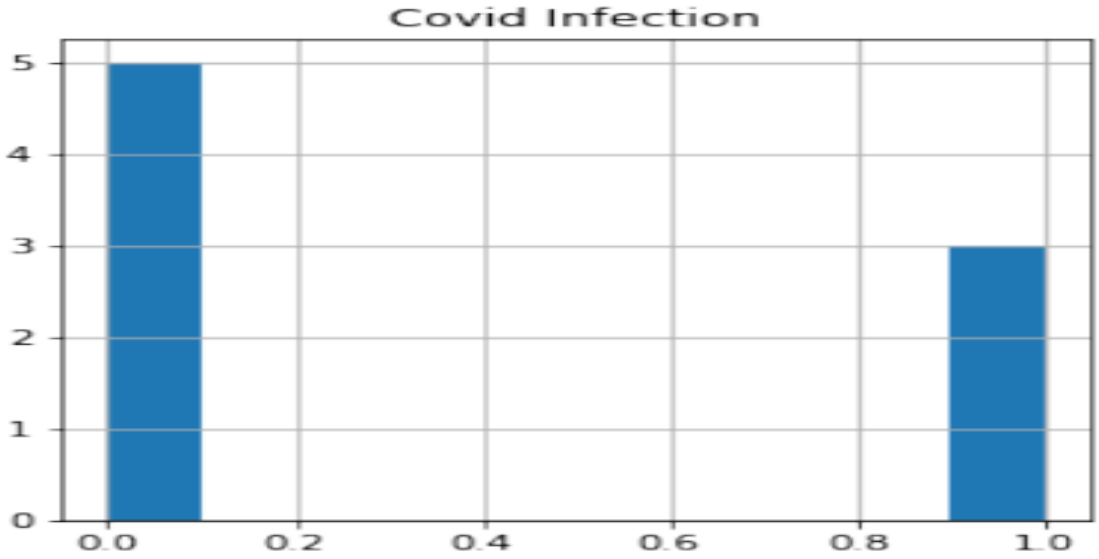

Figure 13 Covid Infection

This machine learning based predictive model is implemented employing Python programming language. The relevant library files have been included for execution of the code. The dataset was given as input and the predictive model obtained. The prediction model was tested and the results were satisfactory. The output graph of the machine learning model has been presented above. In this project, a Random Forest algorithm is employed. There are six binary state lifestyle based input parameters and one binary state output parameter that predicts the risk of covid-19 infection for a person. The random forest algorithm maps both regression and classification with its voting based decision trees algorithmic approach.

CONCLUSION AND FUTURE WORK

This machine learning based trained software automaton can be utilized globally for improving immunity against covid-19 with lifestyle practice. The automaton mentioned here can come as mobile applications or just small dialog box based immunity calculators on big websites. The model can be made better and better by adding more features and larger datasets. Automation below Industry 4.zero has a selected schema or sample at its outset. The software program automaton of the traditional automation version, that is the fame quo, changed into constructed with the aid of using a human professional or a crew of human professionals until now. With the appearance of gadgets studying technology, the software program automaton changed now no longer completely immediately designed with the aid of human professionals. The human professionals construct the gadgets studying software programs and supply the actual international records set as schooling information. The gadget studying software program identifies the sample among the center and the output parameters of the dataset in the shape of a mathematical version. This mathematical version may be downloaded as a operating software program module to different digital computing devices.

REFERENCES

[1] LixiangLia, ZihangYanga and Zhongkai, “Propagation analysis and prediction of the COVID-19”, Infectious Disease Modelling, Volume 5, 2020, Pages 282-292
[2] https://www.javatpoint.com/machine-learning-random-forest-algorithm
[3] https://en.wikipedia.org/wiki/Artificial_intelligence
[4] https://en.wikipedia.org/wiki/Machine_learning
[5] https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution