Epidemic Outbreak Prediction Using Machine Learning Model

Soham Shinde
Information Technology
K.J Somaiya Institute of Engineering and Information Technology
Mumbai, India
soham.ns@somaiya.edu

Seema Yadav
Information Technology
K.J Somaiya Institute of Engineering and Information Technology
Mumbai, India
syadav@somaiya.edu

Ashelesha Somvanshi
Information Technology
K.J Somaiya Institute of Engineering and Information Technology
Mumbai, India
a.somavanshi@somaiya.edu

2022 5th International Conference on Advances in Science and Technology (ICAST) | 978-1-6654-9263-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAST55766.2022.10039594

Abstract— **The intelligent models is used for prediction of diseases as well as creation of model that helps doctor to prevent spreading of disease globally is increased day by day. When a disease spreads rapidly in a short period of time in a specific area, it is called an epidemic outbreak. An outbreak might start in a single community or spread across multiple countries. It might last anywhere from a few days to several years. PHO (Public health organizations) are taking preventative efforts to stop the disease from spreading besides that they are highly benefited from accurate prediction of infectious disease. The emergence of big data in the sectors of health and biomedicine, precise data analysis aids early disease identification and better patient treatment. It is now increasingly viable to use massive computing power to predict and manage outbreaks. Our goal is to investigate and determine how outbreaks spread in villages and suburbs where medical care may be limited. A machine learning model is required to forecast epidemic dynamics and identify where the next outbreak is most likely to occur. Because these are important features that contribute subtly to the dynamics of the disease epidemic, our method considers the climate, geography, and distribution of population in impacted region. Our approach will assist health authorities in taking the necessary steps to guarantee that there are sufficient resources to fulfil demand and, if feasible, to prevent epidemics from arising.**

Keywords— Epidemic, Zika Virus, Outbreak, Cases, Classi fication, Regression, Weather, Population Density, Prediction

I. Introduction

The globe is currently dealing with an evolving spectrum of infectious diseases, which has been influenced by organismal evolution, human demographics, global travel, environmental modifications, and closure of institutions for public health. To meet growing needs for prevention, detection, reporting, and response, current infectious disease surveillance and response procedures are insufficient. The ability to predict diseases will equip governments and healthcare practitioners with a way to promptly respond to epidemics., reducing the damage and conserving limited resources. Many infectious diseases, especially those conveyed by arthropod vectors, possess the capacity to anticipate the potential for epidemics in time and space using sophisticated monitoring and modelling techniques that combine environmental data. These tactics can offer useful, timely, and cost-effective tools when paired with communication technologies. The focus of this study work is on the Zika virus outbreak, and we took into account multiple disease dynamics to enable us make accurate predictions.

II. Literature Survey

Effective outbreak prediction models are needed to learn more about the anticipated spread and infectious disease effects, insights from other legislative bodies and government. The prediction models is for suggestion of new strategies and assess effectiveness of those that have already been put in place. (A & G., 2020) [1].

Model uncertainty was greatly enhanced by the complexity of community activities in different geographic regions and changes in control techniques, in addition to the many unknown and known underlying factors in the transmission. (Darwish, Rahhal, & Jafar, 2020)[2]. Traditional epidemiological models are therefore faced with new challenges in providing more reliable data. Many new models have emerged to address this problem, each of which adds a set of assumptions to the modelling process. (Scarpino & Petri, 2019) [3].

SEIR models demonstrated improved accuracy of model for the Zika and Varicella outbreaks by accounting for the lengthy incubation period that infected patients experience. (Pan, Huang, & Chen, 2012)[4]. A random variable that affects the incubation duration, and a disease-free equilibrium is assumed in SEIR models, just like the SIR model [5].

Due to the complexity and magnitude of the challenge of developing health-related models, recently machine learning has received importance for producing epidemic prediction models. Machine learning methods intent to make models having improved generalisation capability and reliable prediction with increased lead periods. (Yin, Tran, Zhou, Zheng, & Kwoh, 2018) [6].

When using textblob in Python to perform sentiment analysis, polarity and subjectivity are two important factors to consider. As proposed by (Alka, 2018), it focuses on some common areas such as parts of speech, nouns and phrases from text, text classification, sentiment analysis, and so on. In Python, tokens supplied to textblob can be processed as strings for natural language processing. The sentiment analyzer provides a tuple of the type sentiment (polarity, subjectivity), with polarity ranging between [-1.0, 1.0] and subjectivity ranging between [0.0, 1.0].

The outcome of data analysis is greatly improved when using an interactive approach. It also significantly enhances comparative analysis [8]. (Buja, Cook, & Swayne., 1996) suggested an interactive data visualization approach focused on specific analytic tasks like comparisons. Plotly, an open source and interactive python graphing package, is used in this study. Statistical charts, financial charts, scientific charts, and other sorts of charts can all be plotted.

Traditional prediction methodologies make it difficult to manage time components, however time series forecasting techniques take time components into consideration with

978-1-6654-9263-8/22/$31.00 ©2022 IEEE

other factors. Time series forecasting produces far more accurate results than traditional prediction approaches. According to the findings of (Taylor, 2008) study, there is a lot of potential for using time series forecasting to anticipate future outcomes.[9]

Artificial Intelligence model as a Dengue Outbreak Predictor: This model was turned into a Graphical User Interface, which was intended to help and educate the general population in areas at risk of a dengue outbreak. It employs a Bayesian network machine learning technique with a 79-84 percent accuracy (Chenar & Deng, 2018)[10]. A fusion of meteorological data and random forest algorithms was used to forecast global African swine disease outbreaks [11].

The 12 features used for modelling were subjected to a logistic regression analysis after that Receiver Operating Curves were produced, and it was assumed that precipitation had a substantial impact on the pandemic's onset. It decided to use random forest algorithms in conjunction with the subset Evaluator-Best First feature selection method using ASF outbreak data and weather data from the world climb database [12].

III. Problem Definition

COVID19 has shown the true state of our healthcare infrastructure, planning, and preparedness to deal with a pandemic with limited resources, unprepared personnel, and shattered supply chains. If government agencies are given this information ahead of time, they may plan and execute projects in a well-organized manner, maximizing the utilization of employees and resources. The objective of this project is to investigate and develop a multimodal model that may foretell the likelihood of an outbreak in a certain place.

Motivation

The sole motivation of this problem statement is the current pandemic situation that has occurred abruptly. The viral condition known as "coronavirus disease" is caused by a coronavirus that has just been discovered (COVID-19). The COVID-19 situation could have long-term and severe impacts on countries, humanities, and cooperation between nations as well as posing complications in management of sickness and crisis management. There are growing signs that the world after the crisis will change and that globalisation will be called into question in a number of situations.

Scope

The scope of our project is broad and widespread, globally. We will be including numerous ML algorithms for predicting the outbreak of an epidemic disease in a certain region. Because these factors are pertinent and slightly influence the dynamics of the disease outbreak, our methodology considers climate, distribution of population in the afflicted area and geography.

1. Reducing avoidable pain from sickness Reduce the expenditure load on government and healthcare systems by giving them first-hand knowledge of outbreak hotspots and the agents that cause epidemics to spread.
2. The machine learning model must identify the subsequent outbreak-prone locations and attributes that greatly aid in the propagation of the outbreak given a region where an epidemic outbreak has already occurred.

A. Project Architecture

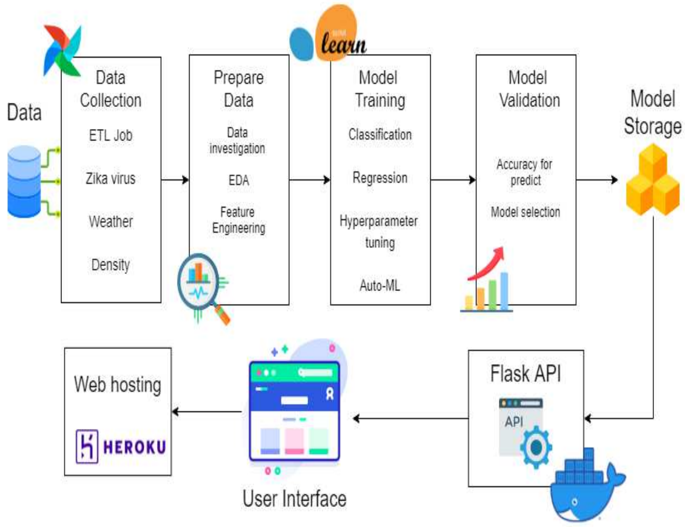


Fig.1. Project Architecture

We begin our project by collecting data from several sources and parameters, such as the Zika Virus dataset, weather dataset, and population density dataset of virus-affected areas. Following that, we do exploratory data analysis in order to extract useful information based on the dataset and the data suitably prepared. In addition, we generate two datasets for classification and regression models, respectively, and perform hyperparameter tuning to pick the best functioning model.

Finally, we design a user interface to display the predictions and results.

IV. Implementation and Results

A. Technology Stack

For all machine learning activities, we primarily employed the Python programming language. Along with Pandas, Numpy, and Seaborn, the Scikit-Learn package is used to build models. In Tableau, we've also generated a dashboard. To make the front end of our website responsive, we used HTML, CSS, JavaScript, and Bootstrap. The Flask API is utilised in the backend for Heroku deployment.

B. Data Collection

We collected three major datasets to help us do the prediction for the outbreak of Zika virus-

- The Zika Data Repository from the CDC and Prevention makes information on the Zika virus accessible to the general public. It gave us enough information to build and test the model.
- Historical weather dataset collected from world weather online through API.
- Population density dataset of the virus struck zones where we are doing the prediction.
- Also, latitude and longitude information of these areas for plotting purpose.

C. *Exploratory Data Analysis*

We study the data gathered in this process to see if there are any relevant insights that can help us make sense of it. We analyse two major elements that may have a significant impact on the frequency of Zika virus cases: The population density of a region, as well as the weather in that region.

Effect of Population density on the number of cases

In any epidemic disease, the population plays a critical role. As a result, we took into account the population density per sq.km of the specific place where instances were discovered. We saw that for a country with the least number of cases, the population density is high whereas for a country with the most number of cases, the population density is low.

TABLE I. POPULATION DENSITY OF COUNTRIES WITH MOST AND LEAST NUMBER OF CASES

Country	No. of cases	Population Density
Panama Metro Las_Garzas	9389432	1966.26
Virgin Islands (US)	19	24970.13

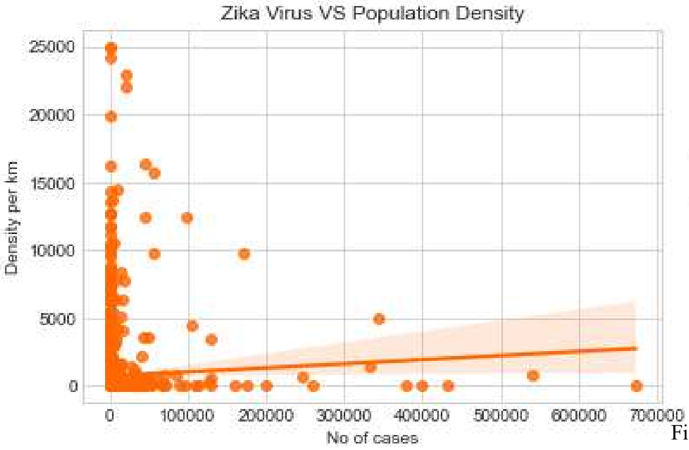


Fig.2. Scatter plot for number for cases vs population density

Effect of Weather on the number of cases:

We noticed that certain countries are tropical, while others are subtropical, in our list of countries. As a result, we decided to track the weather trends in these areas and assess their impact on the number of instances.

Tropical locations are those where the months of a specific season are seen in a consistent manner. The term "subtropical" refers to areas where there is no definite season period.

Countries belonging to tropical regions- Colombia, Brazil, Ecuador, Dominican Republic, El Salvador, Haiti Guatemala, Panama, Nicaragua.

Countries belonging to subtropical regions- United States, Argentina, Mexico.

As we know, Zika Virus is a mosquito-borne disease and mosquitoes are most active in the rainy season. So, for weather analysis, we considered three main factors that will help us to understand if there was any rain- Precipitation, Humidity, and maximum temperature of the area.

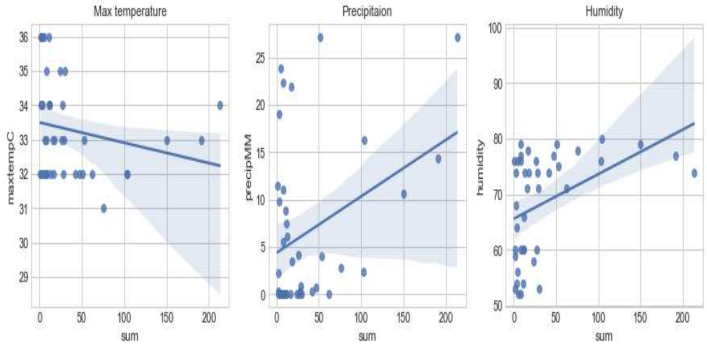


Fig.3. Weather effect on number of cases in Nicaragua

In case of Nicaragua, which is a tropical region, we can see that; Precipitation and Humidity have a positive correlation with the quantity of cases. Hence indicating that the quantity of cases in this country are likely to increase with the commencement of rainy season.

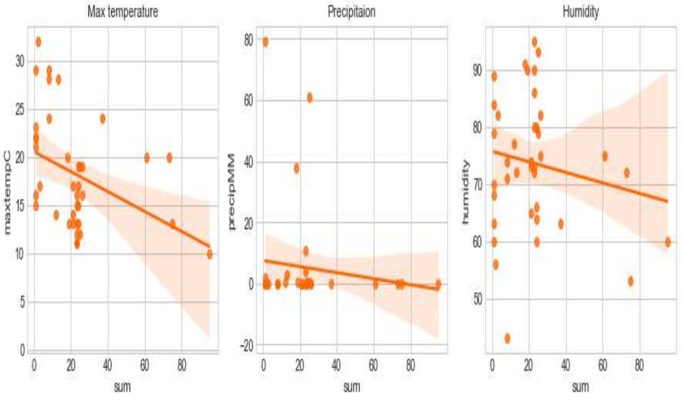


Fig.4. Weather effect on number of cases in Argentina

In case of Argentina, we can see that Precipitation, Humidity, and temperature have a negative correlation with the number of cases. Hence indicating that in such regions weather has minimal effect on number of cases.

Analysis of weather with incubation period of Zika virus-

For any case observed, the process of getting infected happens 7 to 14 days prior. So, we have to observe the trend of weather in those 7 days. The incubation period of Zika virus is generally 3-14 days. In our analysis, we've considered the incubation period of 7 days.

Case 1: When number of cases are 0, the weather conditions 7 days prior are:

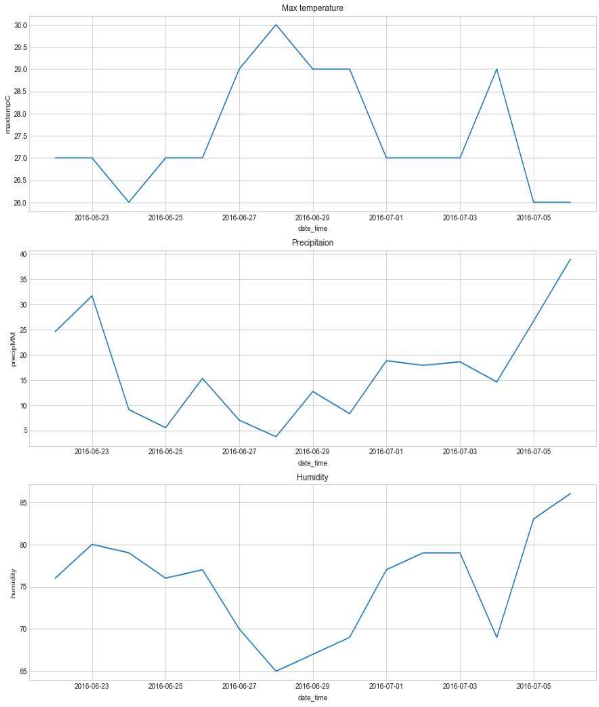


Fig.5. Previous 7 days weather condition for case 0

We can observe that temperature, precipitation and humidity levels were quite normal, indicating ordinary days.

Case 2: When number of cases are 214, the weather conditions 7 days prior are:

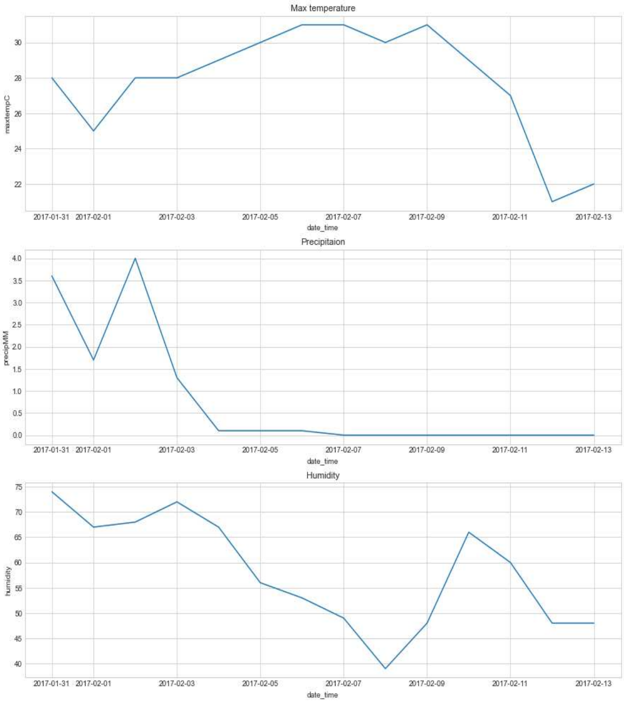


Fig.6. Previous 7 days weather condition for cases 214

Here, we can observe an uneven trend in weather conditions 7 days prior to when maximum number of cases were observed on a single day.

D. *Model Building*

For prediction, we built two kinds of model-Classification model for probability prediction of any cases and Regression model to predict the probable number of cases.

For classification following baseline models were built and tested.

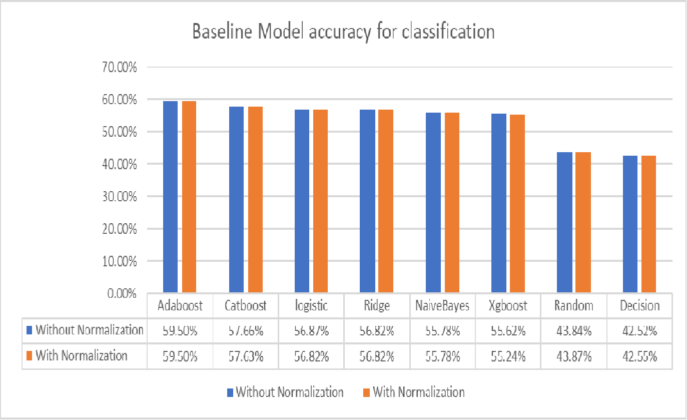


Fig.7. Performance evaluation of classification models

We can see that the ADABoost model performed pretty well as compared to other classification models.

After hyperparameter tuning we can observe that Catboost classification model gave better accuracy as compared to other models which is 60.40%.

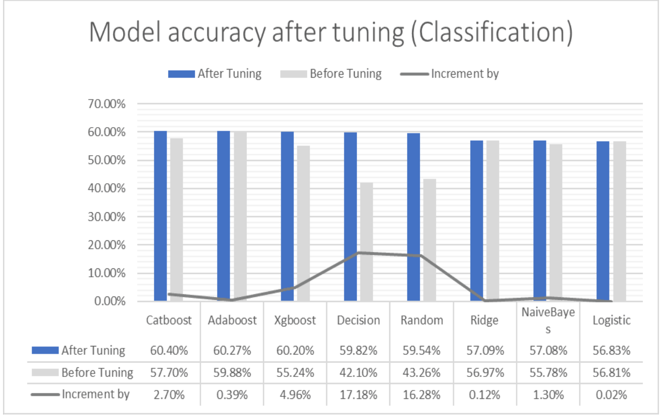


Fig.8. Performance evaluation after tuning

Regression models:

For regression following baseline models were built and tested.

TABLE II. PERFORMANCE EVALUATION OF REGRESSION MODELS

Model	MSE	R2 score
Linear Regression	482209.82	0.00673
Lasso	482521.69	0.00624
Ridge	482442.85	0.00673
Decision Tree	902975.40	-0.859
Random Forest	669564.42	-0.37
XGBoost	512600.67	-0.005

Here, we can see that 7 models were tested for prediction out of which Linear Regression model gave quite standard results.

After hyperparameter tuning of these models:

TABLE III. PERFORMANCE EVALUATION AFTER TUNING

Model	MSE	R2 score	% increase
Lasso	546999.53	0.0067	7.46
Ridge	514096.20	0.0066	1.49
Decision Tree	481223.73	0.0060	101.02
Random Forest	441754.63	0.090	124.32

It has been observed that after hyperparameter tuning the performance of Random Forest Regressor is better compared to other models.

Model Selection

We have used Auto ML to select the best performing model for classification and regression in order to gain better conclusions and results

For Classification:

Auto ML method used: TPOT

Best Model: XGBoost Classifier

TABLE IV. BEST PARAMETERS AFTER HYPERPARAMETER TUNING

Parameter	Parameter
base_score	max_depth
booster	min_child_weight
bylevel	missing
colsample_l	monotone_constraints
colsample_bynode	n_estimators
colsample_bytree	n_jobs
gamma	num_parallel_tree
gpu_id	random_state
importance_type	reg_alpha
interaction_constraints	reg_lambda, validate_parameters
learning_rate	scale_pos_weigh
max_delta_step	tree_method , verbosity

Performance Evaluation: Accuracy = 60.373%

For Regression:

AutoML method used: Auto Sklearn

Best Model: Random Forest Regressor

Performance Evaluation:

Mean Squared Error = 441754.63

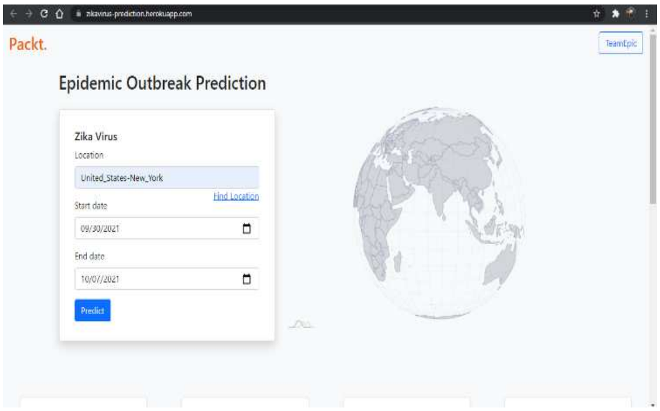
Coefficient of determination (R2 score) = 0.09

These two models were selected in case of classification and regression for prediction.

E. Model Deployment

We used the Flask API to deploy the model on Heroku, and the front end was created entirely with HTML, CSS, and Bootstrap. The website's functionality is as follows: For prediction on the webpage, users will need to provide a country name and a start and finish date. The model will forecast the possibility of an outbreak based on the number of cases.

We may also analyze the pattern of new instances to see if they will rise or fall.



zikavirus-prediction.herokuapp.com

Packt.

TeamEpic

Epidemic Outbreak Prediction

Zika Virus

Location: United_States-New_York

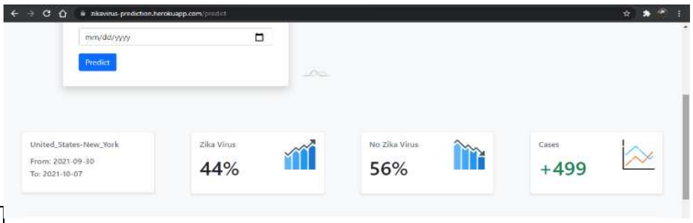
Find Location

Start date: 09/30/2021

End date: 10/07/2021

Predict

Fig.9. Enter location and duration for prediction



zikavirus-prediction.herokuapp.com/predict

mm/dd/yyyy

Predict

United_States-New_York

From: 2021-09-30

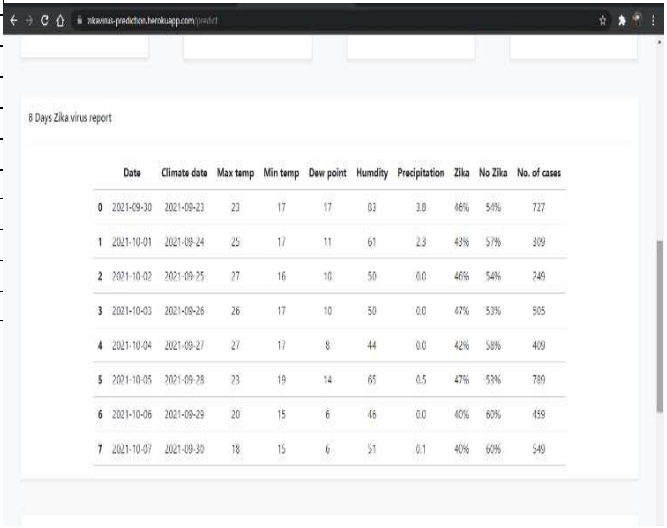
To: 2021-10-07

Zika Virus 44%

No Zika Virus 56%

Cases +499

Fig.10. Predicted results



zikavirus-prediction.herokuapp.com/predict

8 Days Zika virus report

	Date	Climate date	Max temp	Min temp	Dew point	Humdity	Precipitation	Zika	No Zika	No. of cases
0	2021-09-30	2021-09-23	23	17	17	83	3.8	46%	54%	727
1	2021-10-01	2021-09-24	25	17	11	61	2.3	43%	57%	309
2	2021-10-02	2021-09-25	27	16	10	50	0.0	46%	54%	249
3	2021-10-03	2021-09-26	26	17	10	50	0.0	47%	53%	505
4	2021-10-04	2021-09-27	27	17	8	44	0.0	42%	58%	409
5	2021-10-05	2021-09-28	23	19	14	65	0.5	47%	53%	789
6	2021-10-06	2021-09-29	20	15	6	46	0.0	40%	60%	459
7	2021-10-07	2021-09-30	18	15	6	51	0.1	40%	60%	549

Fig.11. Detailed report of next 7 days prediction

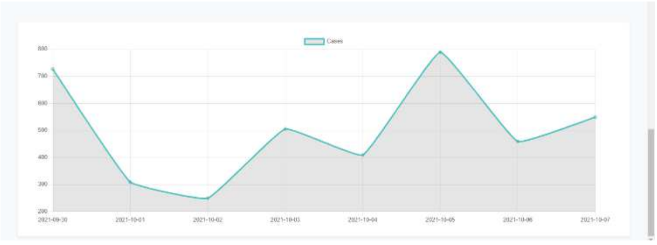


Fig.12. Line graph to observe the trend of cases

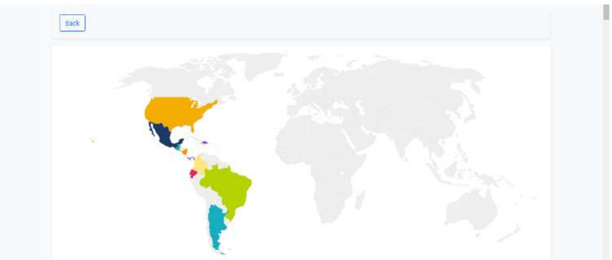
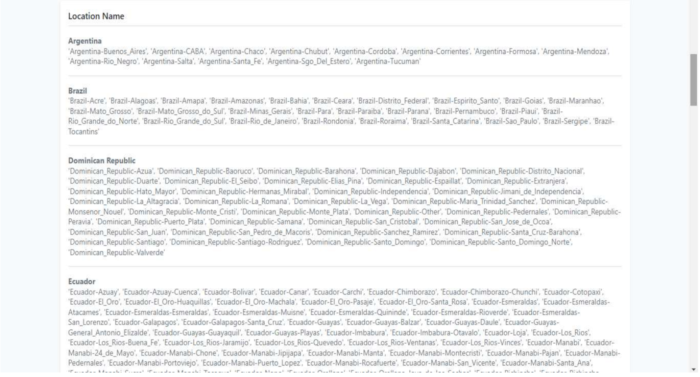


Fig.13. Map to know the countries which are observed for prediction.



Location Name

Argentina
'Argentina-Buenos_Aires', 'Argentina-CABA', 'Argentina-Chaco', 'Argentina-Chubut', 'Argentina-Cordoba', 'Argentina-Corrientes', 'Argentina-Formosa', 'Argentina-Mendoza', 'Argentina-Rio_Negro', 'Argentina-Salta', 'Argentina-Santa_Fe', 'Argentina-Sgo_Del_Estero', 'Argentina-Tucuman'

Brazil
'Brazil-Acre', 'Brazil-Alagoas', 'Brazil-Amapa', 'Brazil-Amazonas', 'Brazil-Bahia', 'Brazil-Ceara', 'Brazil-Distrito_Federal', 'Brazil-Espirito_Santo', 'Brazil-Goias', 'Brazil-Maranhao', 'Brazil-Mato_Grosso', 'Brazil-Mato_Grosso_do_Sul', 'Brazil-Minas_Gerais', 'Brazil-Para', 'Brazil-Paraiba', 'Brazil-Parana', 'Brazil-Pernambuco', 'Brazil-Piaui', 'Brazil-Rio_Grande_do_Norte', 'Brazil-Rio_Grande_do_Sul', 'Brazil-Rio_de_Janeiro', 'Brazil-Rondonia', 'Brazil-Roraima', 'Brazil-Santa_Catarina', 'Brazil-Sao_Paulo', 'Brazil-Sergipe', 'Brazil-Tocantins'

Dominican Republic
'Dominican_Republic-Azua', 'Dominican_Republic-Baoruco', 'Dominican_Republic-Barahona', 'Dominican_Republic-Dajabon', 'Dominican_Republic-Distrito_Nacional', 'Dominican_Republic-Duarte', 'Dominican_Republic-El_Seibo', 'Dominican_Republic-Elias_Pina', 'Dominican_Republic-Espaillat', 'Dominican_Republic-Extranjera', 'Dominican_Republic-Hato_Mayor', 'Dominican_Republic-Hermanas_Mirabal', 'Dominican_Republic-Independencia', 'Dominican_Republic-Jimani_de_Independencia', 'Dominican_Republic-La_Altagracia', 'Dominican_Republic-La_Romana', 'Dominican_Republic-La_Vega', 'Dominican_Republic-Maria_Trinidad_Sanchez', 'Dominican_Republic-Monsenor_Nouel', 'Dominican_Republic-Monte_Cristi', 'Dominican_Republic-Monte_Plata', 'Dominican_Republic-Other', 'Dominican_Republic-Pedernales', 'Dominican_Republic-Peravia', 'Dominican_Republic-Puerto_Plata', 'Dominican_Republic-Samana', 'Dominican_Republic-San_Cristobal', 'Dominican_Republic-San_Jose_de_Ocoa', 'Dominican_Republic-San_Juan', 'Dominican_Republic-San_Pedro_de_Macoris', 'Dominican_Republic-Sanchez_Ramirez', 'Dominican_Republic-Santa_Cruz-Barahona', 'Dominican_Republic-Santiago', 'Dominican_Republic-Santiago-Rodriguez', 'Dominican_Republic-Santo_Domingo', 'Dominican_Republic-Santo_Domingo_Norte', 'Dominican_Republic-Valverde'

Ecuador
'Ecuador-Azuay', 'Ecuador-Azuay-Cuenca', 'Ecuador-Bolivar', 'Ecuador-Canar', 'Ecuador-Carchi', 'Ecuador-Chimborazo', 'Ecuador-Chimborazo-Chunchi', 'Ecuador-Cotopaxi', 'Ecuador-El_Oro', 'Ecuador-El_Oro-Huaquillas', 'Ecuador-El_Oro-Machala', 'Ecuador-El_Oro-Pasaje', 'Ecuador-El_Oro-Santa_Rosa', 'Ecuador-Esmeraldas', 'Ecuador-Esmeraldas-Atacames', 'Ecuador-Esmeraldas-Esmeraldas', 'Ecuador-Esmeraldas-Muisne', 'Ecuador-Esmeraldas-Quininde', 'Ecuador-Esmeraldas-Rioverde', 'Ecuador-Esmeraldas-San_Lorenzo', 'Ecuador-Galapagos', 'Ecuador-Galapagos-Santa_Cruz', 'Ecuador-Guayas', 'Ecuador-Guayas-Balzar', 'Ecuador-Guayas-Daule', 'Ecuador-Guayas-General_Antonio_Elizalde', 'Ecuador-Guayas-Guayaquil', 'Ecuador-Guayas-Playas', 'Ecuador-Imbabura', 'Ecuador-Imbabura-Otavalo', 'Ecuador-Loja', 'Ecuador-Los_Rios', 'Ecuador-Los_Rios-Buena_Fe', 'Ecuador-Los_Rios-Jaramijo', 'Ecuador-Los_Rios-Quevedo', 'Ecuador-Los_Rios-Ventanas', 'Ecuador-Los_Rios-Vinces', 'Ecuador-Manabi', 'Ecuador-Manabi-24_de_Mayo', 'Ecuador-Manabi-Chone', 'Ecuador-Manabi-Jipijapa', 'Ecuador-Manabi-Manta', 'Ecuador-Manabi-Montecristi', 'Ecuador-Manabi-Pajan', 'Ecuador-Manabi-Pedernales', 'Ecuador-Manabi-Portoviejo', 'Ecuador-Manabi-Puerto_Lopez', 'Ecuador-Manabi-Rocafuerte', 'Ecuador-Manabi-San_Vicente', 'Ecuador-Manabi-Santa_Ana',

Fig.14. List of locations for prediction

We have to select the location from this list and thus make the prediction for the number of cases in that location.

V. CONCLUSION

The main aspect of this paper is to make people aware of any epidemic diseases that prevail in their surroundings. Our key emphasis was on studying and analyzing the data we gathered before moving on to building prediction models. With Epidemic Outbreak Prediction, not only the general public, but also the government, can take the required procedures that limit epidemic diseases and prevent their spread. This project is very adaptable, and we can use it to anticipate various diseases. Because the UI is simple to use, individuals of all ages may access and use it.

VI. FUTURE SCOPE

We can expand this project to predict zika virus cases all across the world. We can predict the probability of an epidemic for a few more days now that weather forecasting data is available. Given the success of our proof of concept, we may improve it by adding more criteria such as social media symptomatic data, lifestyle, demographic dynamics, and so on. Using the same method, we can investigate and predict a variety of additional epidemics. We can also seek additional data from government and healthcare institutions that isn't currently available in the public domain

REFERENCES

[1] Remuzzi A, Remuzzi G., "COVID-19 and Italy: what next ? ", Lancet. 2020;395(10231):pp.1225-1228.

[2] Darwish, A.; Rahhal, Y.; Jafar, A., "A comparative study on predicting influenza outbreaks using different feature spaces: Application of influenza-like illness data from Early Warning Alert and Response System in Syria". BMC Res. Notes 2020, 13, pp. 33.

[3] Scarpino, S.V., Petri, G., "On the predictability of infectious disease outbreaks". Nat Communication, 10, 2019, pp. 898 .

[4] Pan, J.-R.; Huang, Z.-Q.; Chen, K., "Evaluation of the effect of varicella outbreak control measures through a discrete time delay SEIR model". Zhonghua Yu Fang Yi Xue Za Zhi2012, 46, pp. 343–347

[5] Dantas, E.; Tosin, M.; Cunha, A., Jr., "Calibration of a SEIR–SEI epidemic model to describe the Zika virus outbreak in Brazil", Applied. Math. Computation. 2018, 338, pp. 249–259.

[6] Yin, R.; Tran, V.H.; Zhou, X.; Zheng, J.; Kwoh, C.K., "Predicting antigenic variants of H1N1 influenza virus based on epidemics and pandemics using a stacking model. PLoS ONE", 2018,pp. 13.

[7] Das Adhikari, Nimai & Alka, Arpana & Kushwaha, Jitendra & Nayak, Ashish, "Sentiment Classifier and Analysis for Epidemic Prediction"., 10.5121/csit.2018.81004.

[8] Andreas Buja, Dianne Cook & Deborah F. Swayne Research Scientist, "Interactive High-Dimensional Data Visualization", Journal of Computational and Graphical Statistics , 1996, 5:1, pp. 78-99.

[9] Taylor, James., "A Comparison of UnivariateTime Series Methods for Forecasting Intraday Arrivals at a Call Center", ManagementScience,2008,54.253-265.0.1287/mnsc.1070.0786.

[10] Chenar, S.S.; Deng, Z, "Development of genetic programming-based models for predicting oyster norovirus outbreak risks.", Water Res., 2018, 128, pp. 20–37.

[11] Liang, R.; Lu, Y.; Qu, X.; Su, Q.; Li, C.; Xia, S.; Liu, Y.; Zhang, Q.; Cao, X.; Chen, Q.; et al. "Prediction for global African swine fever outbreaks based on a combination of random forest algorithms and meteorological data." Transbound. Emerg. Dis., 2020, 67, pp. 935–946.

[12] Raja, D.B.; Mallol, R.; Ting, C.Y.; Kamaludin, F.; Ahmad, R.; Ismail, S.; Jayaraj, V.J.; Sundram, B.M., "Artificial Intelligence Model as Predictor for Dengue Outbreaks". Malays. J. Public Health Med. ,2019, 19, pp. 103–108.