Journal of Theoretical and Applied Information Technology
29th February 2024. Vol.102. No 4
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A NOVEL TECHNIQUE TO DETECT THE HOTSPOTS IN INFLUENZA EFFECTED REGIONS

P. NAGARAJ[1], KALLEPALLI ROHIT KUMAR[2], V.BIKSHAM[3],CHUNDURI KIRAN KUMAR[4], MANDADI VASAVI[5]

[1, 3] Assistant Professor, Anurag University, Department of CSE, India.

[2] Associate Professor, Sreyas Institute of Engineering and Technology, Department of CSE, India.

[4] Assistant Professor, Koneru Lakshmaiah Education Foundation, Department of CSE, India.

[5] Assistant Professor, R.V.R and J.C. College of Engineering, Department of CSE, India.

Email: [1]nagaraj.cse@anurag.edu.in, [2]krk542@gmail.com,[3] vbm2k2@gmail.com,

[4]kirancvl17@gmail.com, [5]vasavilahari@gmail.com,

ABSTRACT

In modern days, healthcare is very important in human life. The World Health Organization (WHO) publishes the latest articles on the humans suffering from different types of flues. Machine Learning (ML) plays a major role in predicting the diseases. The report of patients sometimes fails to provide the needed information if the data is not recorded properly. So recently we have made an endeavour to find the most globally spreading influenza virus in all over the world called as Swine flu or the Influenza (H1N1) flu. Firstly we filterthe number of Patients died by swine flu using Artificial Neural Network (ANN) and detected the location using Dynamic Boundary Location Algorithm (DBLA) and geographic information system (GIS). The next finding is location where more number of patients died termed as hotspots, applying Dynamic Hotspots Detection Algorithm (DHDA), then using High Ranking Frequency Prioritization (HRFP) algorithm high risk hotspots known as prioritized hotspots are detected.

Keywords: *Machine Learning, Ann, Gis, Dbla, Dhda, Hrfp.*

1. INTRODUCTION

Influenza viruses, which spread globally, are the cause of seasonal flu.The flu (Influenza) tends to spread very rapidly. Influenza is an infectious disease caused by an influenza virus. This is a worldwide infection and causes local or widespread epidemics and pandemics. People who cough or sneeze with influenza might easily spread the illness to others. It is best to get vaccinated against the sickness. Flu symptoms include fever with an abrupt onset, cough, sore throat, body aches, and fatigue. Influenza viruses are categorized as types A, B, C, and D. Influenza viruses are further classified into subtypes according to the combinations of the haemagglutinin (HA) and the neuraminidase (NA), the proteins on the surface of the virus. In influenza family there are many types of viruses like Ebola, dengue, malaria, swine flu, bird flu, covid-19, omicron etc., here my research on Swine flu, recently we made an endeavour to find the globally spreading influenza virus all over the world, known as Swine flu, Swine influenza have been causing mortality in recent past, Swine flu is a respiratory disease caused by influenza virus infected by the respiratory tract of the pigs, one kind of influenza A virus is the H1N1 flu, also referred to as the swine flu. A whole new strain of the H1N1 virus started sickening people in the 2009–10 flu seasons. It was a novel influenza virus combination that can infect humans, pigs, and birds. It was frequently referred to as the “swine flu”.

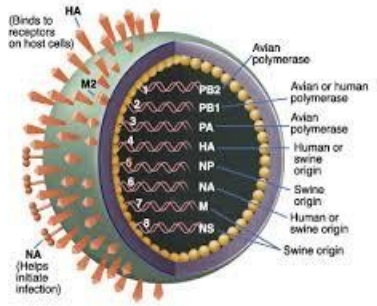


Figure 1.1 Structure of swine flu

2. LITERATRE SURVEY

The H1N1 viral strains occurring in 2009 flu pandemic. The objective of this research is to characterize the viral strains and identify their evolutionary origin phylogenetic analysis, Decision Tree and SVM analysis to find out the swine origin. They find latest pandemic viral of swine origin and they develop web tool and Hidden Markov Model (HMM) for accurate results [1]. They used Naive bayes classifier for classifying the patients of swine flu into 3 categories. The significant predictors are identified by Naive bayes classifier and ISWPS (Intelligent swine flu prediction software). They achieved 63.33% accuracy [2]. The main motive of D.M is to predicting diseases by K-NN Algorithm, patient suffering from which type of swine flu virus by using k-NN algorithm [3]. It recognizes user location by using k nearest neighbor and decision trees, and predicts user destination using SVM. In mobile devices, we achieved the average prediction accuracy higher than 90% through the experiments [4]. This research is prediction techniques to apply to healthcare and finding the application, tools used in health care. In this SMARTDIAB more Tech. has been applied to extract relevant features required for prediction [5].Prediction of Swine Flu using Dynamic Node Creation, feed forward neural network and cascade correlation algorithms. We have observed neural network works better to identify disease compare to naive based, SVM etc [6]. Measurements of location and acceleration information obtained with the neck tag proved to be acceptable for the conditions of this study when cows were housed in the barn. However, data packet dropout and unexpected network delay were observed. The results of the proposed estrus detector based on machine learning techniques showed improved performance, an enhanced number of successful alerts, and a reduced number of false positives compared to statistical analysis methods. The BPNN algorithm with a 0.5-h time window achieved the ideal identification performance. Our results suggest that the combination of the data acquisition system and machine learning methods Animals 2020, 10, 1160 16 of 17 is an alternative to visual observations for indoor-housed cows. Furthermore, the use of the PCA in dimension reduction of correlated behavioral metrics should be advised for the determination of estrus indicators. The estrus indicators originated from the location and acceleration data and the appropriate time window were verified for the positive effects on detection rate of estrus [7].]. It recognizes user location by combining KNN and DT, predicts user destination using hidden Markov models, we achieved the average prediction accuracy is 80% through the experiments [8].We designed a web-GIS system HoLSAT to automatically calculate and visualize the location on different Ml algorithms. We did not consider the dynamic nature of hotel performance and the change of location attributes [9]. The proposed method based on SVM acts like a inputs features which are calculated from processing voltage and current sampled data (in MATLAB software) that were obtained from output of MATLAB. All methods was obtained the right location prediction accuracy is 88% [10]. We have proposed three layered model for flu prediction using support vector-based for predicting trends of influenza spread. The proposed work can be extended in future by hooking a GIS in parallel with the three layers [11]. We concentrate on the prediction of user home locations, tweet locations, and mentioned locations, we make a conclusion of the survey and list future research directions [12]. We have presented a machine learning based intelligent and automated intrusion detection and location solution using PLC modems installed in the grid for communication purposes. Identify potential energy theft without requiring any additional data acquisition devices or the decoding of energy consumption data [13]. We collecting data from Twitter and compared with real time data and approach is Markov Model. Utilized dataset is very small. Result is70% accuracy at detection [14].Machine-Learning-Based Hotspot Detection Using Topological Classification and Critical Feature Extraction, in this Uses Machine learning techniques to find hotspot in fabrication technology. A combination of classification and feature extraction process are used. SVM kernel was used with an accuracy of 78% [15]. We Present SpotOn, a web server to identify and classify interfacial residues as Hot-Spots (HS) and Null-Spots (NS). Spot ON implements a robust algorithm with a

demonstrated accuracy of 0.95 and sensitivity of 0.98 on an independent test set. The predictor was developed using an ensemble machine learning approach with up sampling of the minor class. It was trained on 53 complexes using various features, based on both protein 3D structure and sequence [16]. Hot spots are the subset of interface residues that account for most of the binding free energy, and they play essential roles in the stability of protein binding. We describe the basic concepts and recent advances of machine learning applications in inferring the protein–protein interaction hot spots, and assess the performance of widely used features, machine learning algorithms, and existing state-of-the-art approaches. We also discuss the challenges and future directions in the prediction of hot spots [17]. Used ML based hotspot detection that uses lithography information to build SVM during its learning process, did not detect all hotspots together, no comparisons made [18]. Novel methodology for machine learning (ML) based hotspot detection that uses lithography information to build SVM (Support Vector Machine) during its learning process. Unlike previous studies that use only geometric information or require a post- OPC (Optical Proximity Correction) mask, this proposed method utilizes detailed optical information but bypasses post-OPC mask by sampling latent image intensity and uses those points to train an SVM model. The results suggest high accuracy and low false alarm, and faster runtime compared with methods that require a post-OPC mask [19]. Dataset from 2012 to 2018 and ML based hotspot detection using SVM Limitation: we critically analyze the ML-based hotspot detection literature and we highlight common misconceptions which are found therein [20]. Hotspot analysis is a technique through which potential diseases outbreak areas can be identified based on responsible factors. Looking towards the nature of the COVID-19, district-wise total population, population density, reported confirmed COVID-19 cases, and foreign tourist arrivals in India were taken as a responsible factor for detecting hotspot areas. The geospatial analysis and spatial statistics suggested that these all factors are spatially clustered (p value\0.01). The hotspot generated using Getis-Ord Gi* statistic also revealed that the total population and reported COVID-19 cases produced more accurate hotspot as they produced higher z-score value (z-score [3and [0.7295 respectively). Thus, the resultant hotspot generated using these two factors will be considered as the COVID-19 hotspot for India. The spatial locations of the hotspot in our study will help the Government, district administrative, local authorities, as well as residents of hotspot location for taking actions and precautions against this novel infectious disease [21].

3. PROBLEM STATEMENT

To prevent the swine flu there are several methodologies currently available, but not by detecting the hotspot regions. So we have planned to apply machine learning algorithms (1) Dynamic Boundary Location Algorithm (DBLA), to predict the influenza effected zones (2) Dynamic Hotspot Detection Algorithm(DHDA), to Detect the hotspots in Swine Flu effected Regions. (3) High Ranking Frequency Prioritization (HRFP), to prioritize the detected Hotspots.

4. RESEARCH GAP

Research gap as a topic or area for which missing or inadequate information limits the ability of reviewers to reach a conclusion for a given question. Swine flu detection and prevention are the major issues to the Government Agencies. There are number of patients who are suffering from H1N1 influenza virus in Telangana but mainly two test centres i.e., (Institute of Preventive Medicine (IPM) in Narayanaguda and Govt. Fever Hospital).

https://ncdc.gov.in/showfile.php?lid=325

The H1N1 and H3N2 are seasonal viruses. Till now we have used different type of techniques and algorithms but could not get expected results.

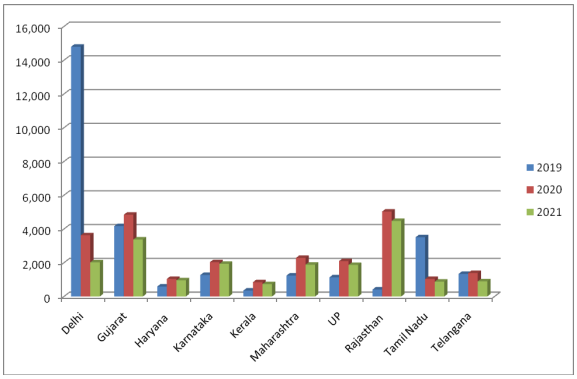


Figure 4.1 Swine flu effected states in last three years

5. SYSTEM ARCHITECTURE

The data is taken from the WHO, NCDC, ICMR data set and this data set we need to do pre-processing. In Preprocessing we delete duplication

records, incomplete data and Normalization. The next stage it will go two phases i.e., Training and Testing. In a Training data(80%) set we are using Artificial Neural Network (ANN), in this finally we find Number of swine flu effected persons and Testing phase (20%) we use confusion matrix and classifiers the values. Finally we get 96% accuracy. On this dataset, now apply our first algorithm Dynamic Boundary Location algorithm (DBLA) to find the Location of swine flu effected persons, Second we will apply Dynamic Hotspot Detection Algorithm (DHDA) then we get more number of people effected died i.e., Hotspots, Finally Cumulative Rank Frequency (CRF) gives more death ratio but if you want prioritize the death ration we have High Rank Frequency Prioritization (HRFP) to rank detected hotspots.

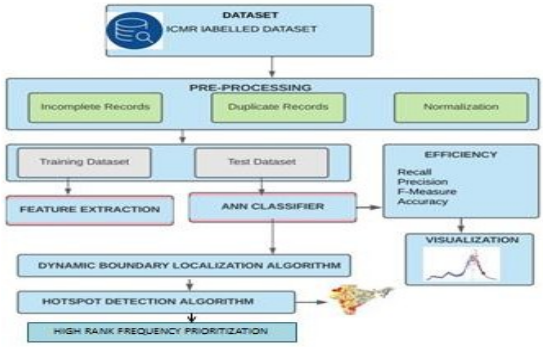


Figure 5.1 Proposed Hotspots in swine flu regions

Table 5.1 The Different State of data from NCDC

S. No	Name of State	Period
1	AndhraPradesh	2015 2016 2017 2018 2019 2020 2021 2022
2	Delhi	
3	Gujarat	
4	Goa	
5	Haryana	
6	Jammu& Kashmir	
7	Karnataka	
8	Kerala	
9	Madhya Pradesh	
10	Maharashtra	
11	Punjab	
12	Rajasthan	
13	Tamil Nadu	
14	Telangana	
15	Uttar Pradesh	

5.1 Data Collection:

In the table 5.1.1, we took only Telangana state swine flu affected data. The data set consists of 1264443 records out of which 1252751 are positive cases of swine flu and 1115517 records all of people you have recovered from this deadly disease. During the period of 2015 to 2022 the total numbers of 137234 of deaths have taken place in Telangana state.

Table 5.1.1 Swine flu Dataset

No of Records	1264443
Positive Cases	1252751
Recovered	1115517
Deaths	137234
No of Training Records	1002200
No of Test Records	250551

In Pre processing we delete duplication records, incomplete data and Normalization. The next stage it will go two phases i.e., Training and Testing. In a Training data(80%) set we are using Artificial Neural Network (ANN), in this finally we find Number of swine flu effected persons and Testing phase (20%) we use confusion matrix and classifiers the values.

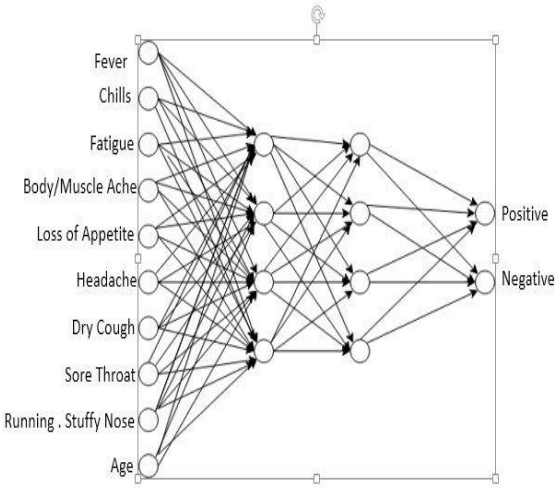


Figure 5.1.1Working of ANN

Training method: In ANN classifier first we will give the 10 inputs to hidden layer1, each and every node is connect to hidden layer1, in the hidden layer one contains 4 nodes, all 10 inputs connect to every node. First feed forward process will do, input is connecting to hidden

layer with random weights and adds bias, the final result will go second hidden layer, finally get output. If the output is not match with desired value it called as Error. So will solve the error by using Output Error <- 1/n sigma i=1 to n (Expected - Output) ^{2}. The activation functions sigmoid or "S"-shaped curve and binary classification (output gives either 0 or 1) are used.

$$F(x) = 1/1 + e^{-x}$$

Testing method: In a Testing phase 20% of data will send to confusion matrix, it performs the classification model and find the classifier values (Recall, precision, f-measure, Accuracy).

5.2 Proposed Predicting ANN Algorithm

Algorithm Prediction Detecting Diseases
Formation
Input
Weights wi
node
Threshold

L1 Begin
L2 {
L3 Initialize the network
L4 WI<-Random
L5 Node<-Threshold

L6 f=Output Expected - Output Observed
L7 f1= Set of f<- minimize error
L8 Repeat L6 to L7 until optimization criteria
L9 End
L10 Output Error <- 1/n sigma i=1 to n (Expected - Output)2
L11 Hidden error <- Back Propagation method
L12 If Error criteria checking -> False Go to Line 10
L13 Repeat Until Expected = Output
L14}
L15 End

The above pseudo code first read the 10 inputs given by patients; it is initialize the network with random weights. Each node is connecting to hidden layer1 and finds the threshold value. If the threshold value is between 0 to 1. Here we calculate f value using expected- observed. The f value minimizes the error and assign to variable f, this step repeat until data can be optimized. The back propagation method can done using f(x) = 1/1 + e-x is the sigmoid function. Now, we need to verify error if it false to line1 to line 10, this step will repeat until we will get expected output.

Table 5.2 System Accuracy Rate

Iteration	Upper value for -ve case	Lower valuefor +ve case	Required Threshold value	Number of Records per iteration	No. of correctly classified patterns	Accuracy for each partition
1	0.287656	0.565139	0.4263975	202310	197252	78.00 %
2	0.3315276	0.524858	0.4281928	404621	192194	76.00 %
3	0.79043	0.379523	0.5849765	606932	212425	84.00 %
4	0.878156	0.335547	0.6068515	809243	232656	96.00 %

Confusion Matrix:

A table known as a confusion matrix is intended to describe how a classification models, moreover known as a "classifier," a set of test data have been run since the true values have been received.

Correct P	Incorrect N	Result
39450	11127	positive
16184	186127	negative

Figure 5.2 Confusion Matrix

Classifier Values:

TP: True Positive

TN: True Negative

FP: False Positive

FN: False Negative

Table 5.3 Classifier Values

Recall	TP/(TP+TN)	0.17488

Journal of Theoretical and Applied Information Technology
29th February 2024. Vol.102. No 4
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Precision	TP/(TP+FP)	0.79481
Frequency measure	(2*Recall*Precision) / (Precision-Recall)	0.448440
Accuracy	(TN+TP)/(FP+FN+TP+TN)	0.96690

Performance of Classifier:

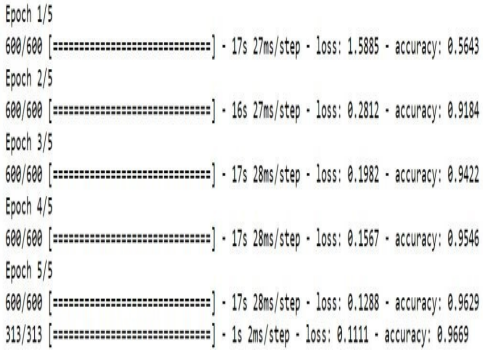


Figure 5.3 Output of Classifier

The above figure show the code written in Python programming language, it contain the five epochs. Each epoch have the accuracy value.

Epoch 1: 56%

Epoch 2: 92%

Epoch 3: 94%

Epoch 4: 95%

Epoch 5: 96%

Performance of Training and Testing:

Figure 5.4 Output of Training and Testing

6. WORKING PROCEDURE OF DBLA ALGORITHM

Proposed Dynamic Boundary Location Algorithm (DBLA)

INPUT ICMR Dataset, Distance D=50, GIS Co-ordinatesBegin
Read-> Training Data, Distance, GIS Co-ordinates
For
Each training data record
Do
{
Mark first data point ->X
Identify points B1,B2,B3 -> D/2 from X
Plot point-> GIS
}
End Do
If region of B1 B2 B3 locate-> X1
X1<- Shortest Distance point-> X
Else
X1<- Outside region B1 B2 B3
Locate nearest Boundary point <- X1 Repeat L5 to L10
until END of training data record

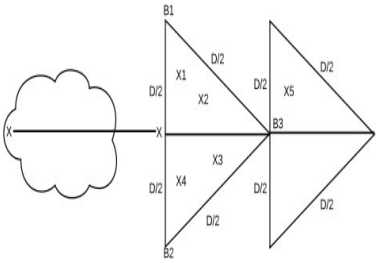
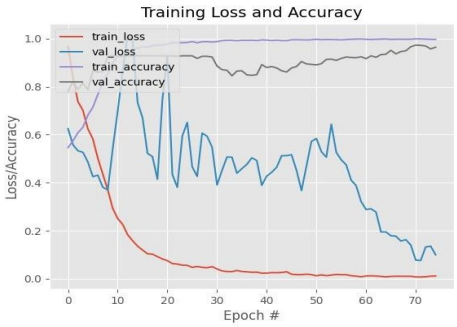


Figure 6.1 working of proposed DBLA algorithm

The explanation of the Dynamic edge boundary algorithm: - -> Consider X as the point where the first instance of influenza/swine flu is detected. The point X (Co-ordinates are input to the Google map) acts as the reference and based on the country/region/area the value of D=Distance is selected. For this example, say D=50 KM. From point X at a distance of D/2 in North, South and East direction mark points B1, B2, and B3 respectively. This region is considered as the outer cluster. The max spread of the disease. For any point say (x1,x2….xn) occurring inside the triangulated region find the nearest data point to x and keep forming small or joint clusters. The condition when a data point is found outside the triangulated area marks the spread of the disease



outside the area. Now dynamically the algorithm has to expand the area. This is performed by searching for the Boundary point nearest to the outside data point (X3 in this example). The nearest boundary point is B3. Now from B3 marks B4, B5 and B6 and repeat the steps as before. The algorithm generates triangular area and keeps

spreading to the three sides North, South and East. The globe/Earth being a circle any expansion of three sides will automatically cover the four i.e.; west.

Results:

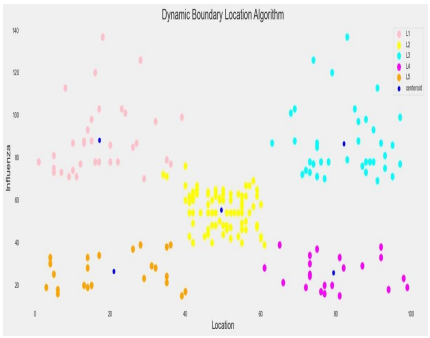


Figure 6.2 Results of locality Prediction using DBLA

Here the output show the Location of five regions, each region can be represented different colour. We take X axis Location and Y axis Influenza, middle have five centroids.

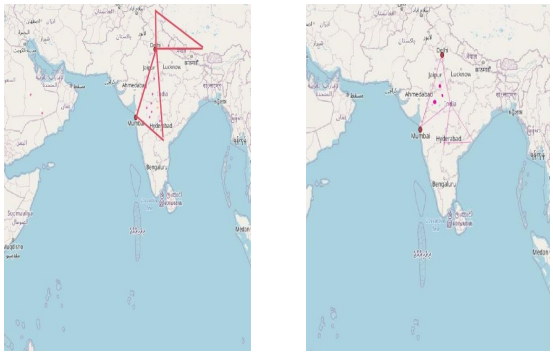


Figure 6.3 Results of locality Prediction using QGIS

Finally we get 96% accuracy. On this dataset, now

apply our first algorithm Dynamic Boundary Location algorithm (DBLA) to find the Location of swine flu effected persons. Second we will apply Dynamic Hotspot Detection Algorithm (DHDA) then we get more number of people effected died i.e., Hotspots.

7. WORKING OF DHDA ALOGIRTHM

Proposed Dynamic Hotspot Detection Algorithm

```
DHDA (cluster)
Begin
//calculate number of nodes in each centroid
Read-> number of centroid
/* Compute distance assign the points to the nearest
function colour points according to the function*/
For i = 1 to n
 noc = count (cluster)
 nopi = noc
  print (noc)
End for
//Find max value of each cluster Initialize max
 for i= 1 to n
     if nopi is greater than max
           max = nopi
 End for
 print Max
 End
```

The Dynamic Hotspot Detection Algorithm (DHDA) is detecting hotpots in dynamically. "Hotspot" is frequently used in infectious disease epidemiology to reveal areas with a high disease burden or high transmission efficiency. The Dynamic Boundary Location Algorithm (DBLA) predicting the location of swine flu effected people, in that location it forms clustering. In each Region form the number centroid, the Dynamic Hotspot Detection Algorithm read the centroid number and calculates each node's number centroid. A loop will continue and finally it prints the number of node in each centroid (Compute distance assign the points to the nearest function colour points according to the function). Now we got total number of centroids and each centroids holds number of nodes. Find max value of each cluster, initially maximum values is assign to max, loop will start if number of points is greater than max, number of point's value is assign to max, and finally print the maximum number of points in each region is called Hotspot.

Results:

Figure 7.1 Clustering results of Location using DBLA

.

In the above diagram we use dynamic boundary location algorithm to detect the location in Telangana state. Here there are five different regions are detected namely Warangal, Mahaboob nagar, nalgonda, Hyderabad, Secundrabad. Take x axis is location and Y axis is influenza virus effected ratio. If we want find maximum number of effected nodes in each cluster, we use logic is bin count in python. The bin count displays the maximum number of count in each region. Each location contain different type of colour like Warangal cyan, Mahaboobnagar green, Nalgonda blue, Ranga reddy pink, Hyderabad yellow. Now we need to find maximum number of swine flu affected people use logic is Max (cluster count). Finally it will print maximum number; this count is known as Hotspots. The fourth location contain maximum value i.e., Rangareddy region.

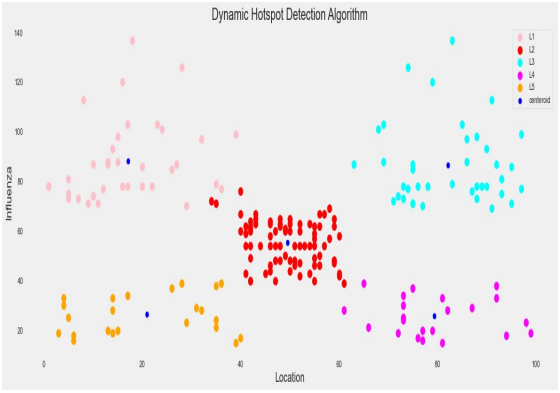


Figure 7.2 Clustering results of locations using DHDA

The above diagram contains the five centroids (color is dark blue) and five regions (color is cyan, blue, yellow, red, pink). The fourth region have maximum number of death ratio in Ranga reddy region, this is color changed pink to Red. So red color denotes the Hotspot.

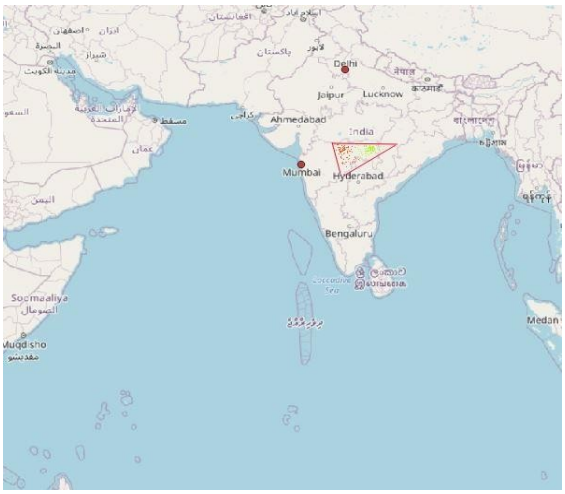


Figure 7.3 Results of Hotspot detection

The same data set can be imported into QGIS software and gives the latitude and longitude it will display the maximum number of patient dead is call Hotspot.

Cluster Point No	Gi*	Latitude(in degrees)	Longitude (in degrees)
1	4.35	17.3850°	78.4867°
2	2.13	19.0760°	72.8777°
3	2.1	28.7041°	77.1025°

Table 7.1 Evaluation results

Finally, prioritization uses High Ranking Frequency Prioritization (HRFP) to rank detected hotspots. We have Cumulative Rank Frequency (CRF) gives more death ratio but if you want prioritize the death ration we have High Rank Frequency Prioritization (HRFP), this will work based on the Prioritization matrix and proposed algorithm will find out the region and performer cluster whether person is infected with Swine Flu or not based on the GIS coordinates the hotspot is detected. Cumulative frequency has two varieties: lesser than type and larger than type. The number of observations in a given data collection that fall above (or below) a specific frequency is calculated using cumulative frequency. In this which has the heights rank death ration is called cumulative rank frequency (CRF).

Journal of Theoretical and Applied Information Technology
29th February 2024. Vol.102. No 4
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 7.2 Data set of Cumulative Rank Frequency

s.no	Location	Districts	Latitude	Longitude	Population	Tested	Infected	Deaths
1	Balanagar	Rangareddy	17.465	78.450994	567996	220170	17057	584
2	Saroornagar	Rangareddy	17.354	78.531467	349004	256420	194892	12973
3	Qutubullapur	Rangareddy	17.4917	78.436458	495683	236342	113641	11985
4	Malkajgiri	Rangareddy	17.4478	78.52633	413571	292602	248139	15982
5	Uppal	Rangareddy	17.4018	78.560188	384835	289125	230628	12634
6	Serilingampally	Rangareddy	17.4795	78.319902	309320	203479	109910	32105
7	Rajendranagar	Rangareddy	17.2367	78.429882	307175	136799	73858	54627
8	Hayathnagar	Rangareddy	17.3271	78.591241	227195	208174	17438	8647
9	Ghatkesar	Rangareddy	17.4468	78.67672	188380	56598	25311	8710
10	Keesara	Rangareddy	17.4777	78.577859	177288	130126	101371	67043

In a Rangareddy Region total 37 Districts are there, in that region we taken most affected 10 districts. These districts contain Latitude, longitudes, Tested, Population, Tested, and Deaths. The above table contains most effected deaths in Keesara, Rajendranagar so on. If you are using Cumulative Rank-Frequency (CRF), we can give first rank goes to Keesara because it has more number of Deaths. Second Rank goes to Rajendranagar; Third Rank goes to serilingpally so on.

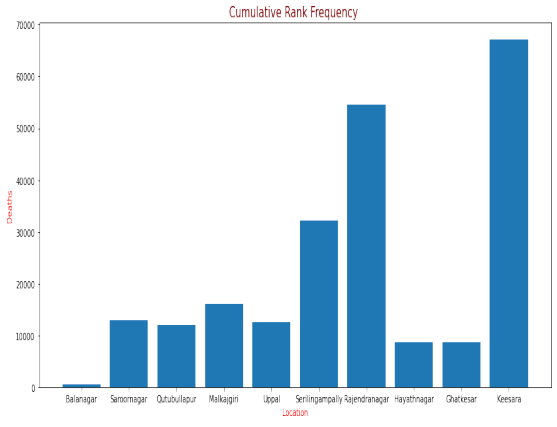


Figure 7.4 Bar chart of Cumulative Rank Frequency

If we are using Prioritization matrix we need consider total population, infected, death. The best common methods for priority is a prioritization matrix, when evaluating health concerns in criterion or when a particular activity is only controlled to concentrate on single high-precedence healthiness problem. Decision matrices provide visual methods for priority and criteria with degrees of importance, more than measures.

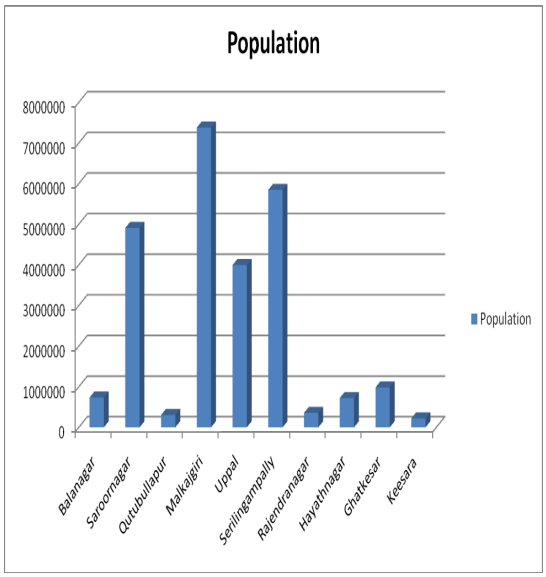


Figure 7.5 Total Population of Ranga reddy Region

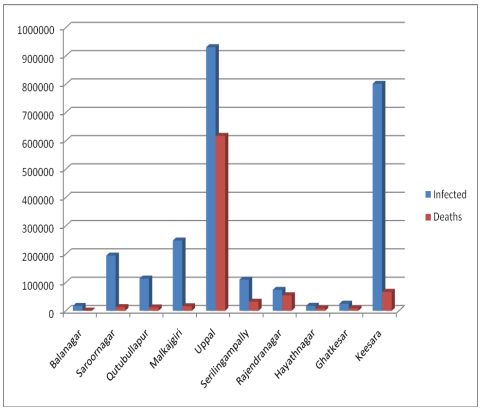


Figure 7.6 Total Infected, Deaths of Ranga Reddy Region

In the above figure 5.1 we consider the total population of Rangareddy district. Here Malkajgiri, serilingampally, saroornagar, uppal have more populations. In the above figure5.2, we consider amount of infected public by pig infection and calculate the amount of demise public by swine flu. The uppal have most effected by swine flu and death ratio also more, second place is keesara have most effected and death by swine flu.

8. WORKING OF HRFP ALOGIRTHM

Algorithm: High Ranking Frequency Prioritization

L1 Read Data set
L2 Initialize xi, yi
L3 yi = regionsi
L4 xi = deathsi, infectedi, populationi
L5 Read values from Data set and assign matrix xi, yi
L6 Give each parameter a weight in order to take the differences into consideration.
For e.g., if "Deaths" was twice as essential as "Infected" and "Population," "Deaths" may have a weight of 0.5 and "Infected" and "Population" could have a weight of 0.25.
HRFP = death*0.5+Infected*0.25+Population*0.25
L7 Assign waits 0.50 to deathsi, 0.25 to infectedi, 0.25 to populationi
L8 Compute infectedi * 0.25, deathsi * 0.5, populationi * 0.25
L9 Store sum of each matrix row into Priority_Scorei
L10 Give each region a rank, with 1 being the highest, according to Priority_Scorei using rank method
L11 Highest ranki specifies most affected regions.

Journal of Theoretical and Applied Information Technology
29th February 2024. Vol.102. No 4
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 8.1 Evaluation of HRFP Method

S.no	Location	Districts	Latitude	Longitude	Tested	Population	Infected	Deaths	Total	Rank
1	Balanagar	Rangareddy	17.46501	78.450994	220170	567996*0.25=141999	17057*0.25=4264	584*0.50=292	146555	4
2	Saroornagar	Rangareddy	17.35402	78.5314672	256420	349004*0.25=87251	194892*0.2 5=48723	12973*0.5=6486	142460	5
3	Qutubullapur	Rangareddy	17.49173	78.4364582	136342	495683*0.25=123920	113641*0.2 5=28410	11985*0.5=5992	158322	3
4	Malkajgiri	Rangareddy	17.44781	78.52633	292602	413571*0.25=103392	248139*0.2 5=62034	15982*0.5=7991	173417	1
5	Uppal	Rangareddy	17.40181	78.560188	289125	384835*0.25=96208	230628*0.2 5=57657	12634*0.5=6317	160182	2
6	Serilingampally	Rangareddy	17.47946	78.3199015	203479	309320*0.25=77330	109910*0.2 5=27477	32105*0.5=1605	120859	7
7	Rajendranagar	Rangareddy	17.23667	78.4298816	136799	307175*0.25=76793	73858*0.25 =18464	54627*0.5=2731	122570	6
8	Hayathnagar	Rangareddy	17.32711	78.5912409	208174	227195*0.25=56798	17438*0.25 =4359	8647*0.5=4323	65480	9
9	Ghatkesar	Rangareddy	17.44682	78.6767195	56598	188380*0.25=47095	25311*0.25 =6327	8710*0.5=4355	57777	10
10	Keesara	Rangareddy	17.47772	78.5778592	130126	177288*0.25=44322	101371*0.2 5=25342	67043*0.5=3352	103185	8

The above table contain 10 different regions of Ranga reddy district and the attributes are Location, district, latitude, longitude, number of tested, population, infected , death, total and finally we give the ranks. The weight of "Deaths" is 0.5 and the weight of "Infected" and "Population" could be.25 if "Deaths" is twice as critical as "Infected" and "Population." Assign waits 0.50 to deathsi, 0.25 to infectedi, and 0.25 to populationi and calculate the total, finally assign to priotization score variable. Based on total which have highest total marked as a Rank 1, second height total marked as Rank 2, so on. This processing knows as High Rank Frequency Prioritization.

Results Of Hrfp

The below results show the Line chart, Bar chart, Dark grid, Face grid of death ratio, the ten different regions in Rangareddy districts. In this distinct the death ratio calculated based on HRFP (High Ranking Frequency Prioritization). The ranking is done based on death, here malkajgiri is first rank on swine flu death, second is serilingampally, third is uppal so on.

Journal of Theoretical and Applied Information Technology
29th February 2024. Vol.102. No 4
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 8.1 Results of HRFP

To overcome this problem personally we developed one website **www.Pranasetu.com.** In this website we can read any type of virus data.

9. CONCLUSION

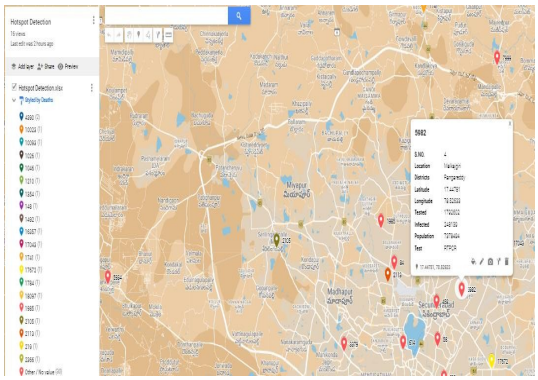
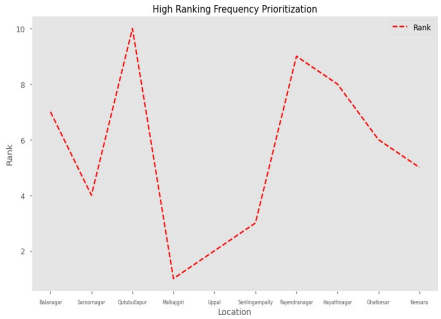
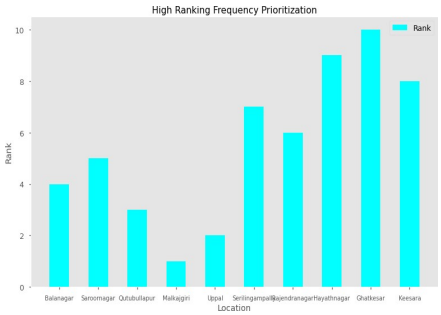
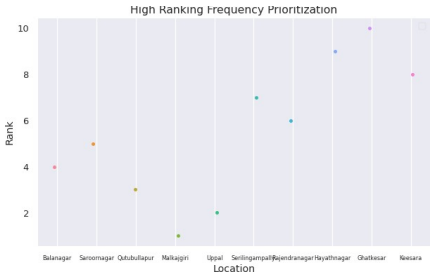


Figure 9.1 Working of Pranasetu

The DBLA algorithm used to find zones based on number deaths occurred in a location gives accuracy of 96% than any other method used for this purpose. The number of infections and population also may be taken into consideration to broaden the scope of this study. The hotspot regions are detected based on severity by applying DHDA algorithm, but density also may be taken into consideration. The priority to the hotspot given based on factors population, infected and death ratio with different cost factors using HRFP algorithm. The matrix has to be found with varying weights to the factors so that best possible combination of factors can be chosen.







10. FUTURE SCOPE

The proposed research provides the advantages of Influenza security and minimum time complexity. The future work will focus on the possible all virus attacks and Location prediction of Influenza virus, Geographic Information System and measure its strength. Another future direction is to enhance the flexibility of allowing dissimilar influenza datasets to be uploaded into the Pranasethu and compute for as long as accuracy medicine and thus improving the health and welfare of patients and recommended in more critical to identify virus attacked in humans. Also, this research will continue to possess adequate opportunities in achieving Hotspot detection in influenza effected regions and Prioritization hotspot defines the High rank frequency prioritization. Last but not least in the future, implement the any type of virus can be prevented.

REFERENCES

[1]Pavan K Attaluri,Ximeng Zheng,Zheng xin, Guoqing Lu," Applying machine learning techniqu es to classify H1N1 viral strains occurring in 2009 flu pandemic.

[2] Binal A. Thakkar, Mosin I . Hasan, Mansi A. Desai "Health care decision support system for swine flu prediction using Naive bayes classifier" in 2010.

[3] Ms.Ankita,R.Borkar,Dr.Prashant R.Deshmukh "K-NN Classifier for prediction of swine influenza virus" in 2015

[4] Sung-Bae Cho "Exploiting machine learning techniques for location recognition and prediction with smart phone logs" in 2016

[5] Sujata joshi, k.Nair "Survey of Classification based prediction techniques in healthcare" in 2018

[6] Dvijesh Bhatt, Daiwat vyas, Malaram kumhar and Ajay patel "Swine flu predication using Machine learning" in 2019

[7] Jun Wang, Matt Bell , Xiaohang Liu and GangLiu" Machine-Learning Techniques Can Enhance Dairy Cow Estrus Detection Using Location and Acceleration Data" in 2020

[8] Sung-Bae Cho "Exploiting machine learning techniques for location recognition and prediction with smart phone logs" in 2015.

[9] Yang Yanga, JingyinTang ,Hao Luoc, Rob Lawd "Hotel location evaluation: A combination of machine learning tools and web GIS" in 2015

[10] Younes Mohammadi,Mohammad H.Moradi ,Roberto Chouhy Leborgne "A novel method for voltage-sag source location using a robust machine learning approach" in 2016

[11] Sandeep Singh, Harjot Kaur "Influenza prediction from social media texts using machine learning" in 2018

[12] Xin Zheng , Jialong Han , Aixin Sun "A Survey of Location Prediction on Twitter" in 2018

[13] Gautham Prasad, Yinjia Huo, Lutz Lampe, and Victor C. M. Leung "Machine Learning Based Physical-Layer Intrusion Detection and Location for the Smart Grid" in 2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm).

[14] Thanh Voa, Rohit Sharmab, Raghvendra Kumar, Le Hoang Sond, Binh Thai Phame , Dieu Tien Buif,g,Ishaani "Crime rate detection using social media of different crime locations and Twitter part- Of-speech tagger with Brown clustering" in 2020

[15] Y.T.Yu,G.H.Lin,I.H.R. Jiang, and C. Chiang, "Machine-Learning-Based Hotspot Detection Using Topological Classification and Critical Feature Extraction," IEEE Trans. Comput Aided Des.Integr.CircuitsSyst.,vol.34,no.3,pp.460–470, Mar. 2015, doi: 10.1109/TCAD.2014.2387858.

[16] Sumanta Das,Malini Roy Choudhury "A Geo-Statistical Approach for Crime hot spot Prediction "International Journal of Criminology and Sociological Theory, Vol. 9, No. 1, August 2016, 1-11.

[17] Siyu Liu, Chuyao Liu and Lei Deng "Machine Learning Approaches for Protein–Protein Interaction Hot Spot Prediction: Progress and Comparative Assessment", 4 October 2018 Litho-aware Machine Learning for Hotspot Detection Jea Woo Park , Andres Torres, Xiaoyu Song, IEEE Transactions On Computer-Aided Design Of Integrated Circuits And Systems, Vol. , No. , September 2017.

[18] Siyu Liu, Chuyao Liu and Lei Deng "Machine Learning Approaches for Protein–Protein Interaction Hot Spot Prediction: Progress and Comparative Assessment" in 2018.

[19] Gaurav Rajavendra Reddy, Kareem Madkour and Yiorgos Makris" Machine Learning-Based Hotspot Detection: Fallacies, Pitfalls and Marching Orders" in June 28, 2019 IEEE xplore.

[20] Mohsen Shariati, Tahoora Mesgari Mahboobeh Kasraee Mahsa Jahangiri-rad "Spatiotemporal analysis and hotspots detection of COVID-19 using GIS" April, 2020.

[21] Farhana Parvin, Sk Ajim Ali1 • S. Najmul Islam Hashmi , Ateeque Ahmad "Spatial prediction and mapping of the COVID-19 hotspot in India using geo statistical technique" in Korean Spatial Information Society 2021, https://doi.org/10.1007/s41324-020-00375-1

[22] Swine Flu Hotspot Prediction In Regions Based On Dynamic Hotspot Detection Algorithm"Nagaraj, P., Krishna Prasad, A.V., Venkata Dass, M. , Kumar, K.R." Journal of Theoretical and Applied Information Technology this link is disabled, 2022, 100(22), pp. 6535–6544

[23] Swine flu Detection and Location using Machine Learning Techniques and GIS Nagaraj,P., Narsimha, V.B. , Prasad, A.V.K., Sujatha, B. International Journal of Advanced

Computer Science and Applications, 2022, 13(9), pp. 901–909

[24] Breast Cancer Risk Detection using XGB Classification Machine Learning Technique Nagaraj, P., Dass, M.V. , Mahender, E. , Kumar, K.R. Proceedings of 2022 IEEE International Conference on Current Development in Engineering and Technology, CCET 2022, 2022

[25]Performance of Secure Data Deduplication Framework in Cloud Services Banala, R., Nair,V., Nagaraj, P.Communications in Computer and Information Science, 2022, 1673 CCIS, pp.392–403

[26]Real Time Face Recognition using Effective Supervised Machine Learning Algorithms Nagaraj, P., Banala, R. , Krishna Prasad, A.V. Journal of Physics: Conference Series, 2021, 1998(1), 012007

[27]A Novel Technique to Classify Face Mask for Human Safety Nagaraj, P., Phebe,G.S., Singh, A. Proceedings of the IEEE International Conference Image Information Processing, 2021, 2021-November, pp. 235–239

[28]A Novel Technique to Detect the Hotspots in Swine Flu Effected Regions Nagaraj, P., Prasad, A.V.K.2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), ICRITO 2021, 2021