Swine Flu Detection and Location using Machine Learning Techniques and GIS

P. Nagaraj[1]
Research Scholar of Osmania University and
Associate Professor in Sreyas Institute of Engineering, and Technology, CSE
Hyderabad, India

Dr. V. B. Narsimha[3], Dr. B. Sujatha[4]
Assistant Professor
Department of Computer Science & Engineering
University College of Engineering (A), Osmania University,
Hyderabad, India

Dr. A. V. Krishna Prasad[2]
Associate Professor in IT Department
Maturi Venkata Subba Rao Engineering College
Hyderabad, India

***Abstract*—The H1N1 virus, more commonly referred to as swine flu, is an illness that is extremely infectious and can in some cases be fatal. Because of this, the lives of many individuals have been taken. The disease can be transmitted from pigs to people. This research presents an artificial neural network (ANN) classifier for disease forecasting, as well as a technique for detecting people who are sick based on the geographic region in which they are found. The source codes for these two algorithms are provided below. These coordinates serve as the foundation for the GIS coordinates that are utilized in the method for assessing the extent to which the illness has spread. The ICMR and NCDC datasets were utilized in the study. They used Dynamic Boundary Location algorithm to detect swine flu affected person's location, the researchers discovered that the accuracy of the proposed classifier was 96 standard classifiers.**

Keywords*—Swine Flu; influenza; machine learning; GIS; classifiers; ANN; virus; algorithm

I. Introduction

The swine flu is very contagious and spreads quickly. Potential vectors for the propagation of the disease include air and water. The influenza virus strain is also known as H1N1, which is another name for the H1N1 virus. Virus outbreaks might be global or local in scope, but they invariably end in the loss of human life. More than one country has reported human cases of influenza virus infection. It affects the respiratory system in humans. The WHO estimates that the influenza virus kills between 250,000 and 500,000 people each year, infecting 5 to 15% of the world's population and causing respiratory illnesses and other complications. Due to influenza-related infections, the United States of America loses between $70 billion and $170 billion. This has a huge impact on the economics of the rest of the globe [1] [2].

An artificial neural network (ANN)-based system for swine flu categorization and forecasting is suggested in this paper. GIS coordinates of the person who is infected are acquired, and the second approach is used to locate the individual's position, as outlined in the article. The algorithm determines how far the sickness may spread, and it is constantly being updated as more people become ill. Confinement zones are defined by a set of boundaries generated by an algorithm. Machine learning may help in the analysis and exploration of patterns in a dataset. Machine learning-based classifiers take in data from training datasets and then use that data to make predictions.

Contribution from us: We present a cutting-edge technique that use machine learning to reliably forecast the incidence of swine flu in real time. Sore throats, chills, weariness, nausea, runny noses, body pains, coughs, and fevers are all common symptoms of swine flu infections. In certain cases, a cough may accompany these symptoms. ANN classifiers are able to predict whether or not a person has swine flu with a "Yes" or "No" answer, despite the difficulty of the task. We've come up with a supervised classifier as a solution to this problem. The classifier's output is sent into the proposed boundary algorithm, which generates the disease's sphere of influence's outer border.

II. Literature Review

The Table I shows the Literature review of swine flu affected areas and the limitation of each paper. In this, we explain the main research gap of the latest significant state of the art approaches. Based on Table I we found research gaps in Machine learning for predicting location. We have seen various kinds of research papers and we conclude that most of the research papers are used Support vector machine, Random forest, Artificial Neural Network, NavieBayes, Ad boost and KNN. The main gap is quality of dataset is not used, accuracy and no comparisons are not made.

TABLE I. LITERATURE REVIEW

Ref No	Approach	Limitations	Gaps
[3]	Describe the use of social network plat forms to track people infected with swine flu based on their post on socialmedia like twitter	Social media data like twitter is not trustworthy	No detection method is discussed
[4]	Discusses the symptoms of the disease along with the hotspot detection	No comparisons made	No proper dataset And detection mechanism
[5]	Discusses 7 ML Classifiers for influenza detection	Dataset was randomly selected and consisted of 31268 records over a period from 2008-11	The number of records and time period both are less
[6]	Study and review of the Existing ML techniques that could detectother diseases were performed. The techniques like SVM, Random Forest, ANN, Naïve Bayes and Adaboost are used.	No classifier was proposed. Only standard base line algorithm were used	A quality dataset was not used for thestudy
[7]	Use deep learning neural network architecture for predicting thoracic disease using x-ray images. 50 layersResnet architecture is used. Dynamic routing approach isused in between convolution layers	No classifier was Proposed. Data set is specific to application	Semi supervised approach not used inlocation information
[8]	Predicts and localization of disease donesimultane ously. The use of deep earningclassifier is used	Classifier is nor explained	A quality dataset was not used for thestudy
[9]	Larger dataset arerequired to get better results	Gaussian distribution notused	No comparisons made
[10]	Images of pneumonia may be identified and pinpointed thanks to a model constructed using deep learning. Utilizes a segmentation-based approach.	Dataset was small	Average segmentation approach is used
[11]	It has been hypothesised that there exists a system that is capable of concurrently learning discriminative brain-region localization andsickness detection. The datacame from ADNI, which was used as the source. The algorithm that will be used to categorise the data has been suggested.	The model works well for small patches	Classifier for prediction is not available
[12]	Proposed a classifier based on Random forest algorithmon individual symptoms of swine flu. Probability [reduction of eachsymptomswas detected leading to the disease ofswine flu	Dataset of GCI used is small	Did not perform proper clustering to detect hotspot of the disease
[13]	Create a concept for an artificial neural network (ANN) that loops back on itself. In comparative testing, the performance ofthe suggested method is Superior than that of SVMand Naive Bayes.	Algorithm and dataset not described properly	No comparisons made
[14]	Discusses the methods like density estimation, Model based approaches to clustering	Detection process is not mentioned	A quality dataset was not used for the study
[15]	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%	Accuracy is low
[16]	Detects hotspot in online Forums. The proposed algorithm works in collaboration with K Meansand SVM. The centre of thecluster is the hotspot	Apart from Detection. No proper algorithm is used to identify the hotspots	No comparisons made
[17]	Uses KNN algorithm to find hotspot with ML algorithm and	Aproper dataset is not used	Segmentation of cluster region notspecified
[18]	Applied a ML-based hotspot identification approach thatincludes lithography data into the development of theSVMduring the learning phase. True alarms are few and far between.	Did not detect all hotspots together	No comparisonsmade
[19]	Used feature extraction with tensor generation in CNN that has spatial relationships.The learning process is further extended to batch learning. Used gradient decent approach with lowfalse alarm	Comparison Between different methods learning'snot specified	A quality dataset was not used for thestudy

III. Methodology

Dataset

The ICMR supplied the data utilized in the study. 122751 cases of swine flu have been verified and 115517 people have recovered from the infection, making up 1264443 entries in the database. Between 2015 and 2021, 7336 Americans lost their lives as a result of various causes around the country. Shown in Table II are the dataset's most essential features.

TABLE II. DATASET

No of Records	1264443
Positive Cases	122751
Recovered	115517
Deaths	137234
No of Training Records	1002200
No of Test Records	250551

The Table III shows the attributes that was selected for the study after performing a feature selection. 10 attributes where selected for the study.

TABLE III. ATTRIBUTES SELECTED FOR THE STUDY

Sl. No.	Symptoms/Attributes
1	Fever
2	Chills
3	Fatigue
4	Body/Muscle ache
5	Loss of Appetite
6	Headache
7	Dry Cough
8	Sore throat
9	Running/stuffy nose
10	Age

The Proposed System architecture of the system is displayed in Fig. 1. The ICMR and NCDC data sets supply the information that is used in the recommended technique. The recommended approach can determine whether or not a person is sick, and hence at danger, when GIS coordinates are used to find a likely hotspot for the Swine Flu. The purpose of this research is to identify cases of swine flu using the application of machine learning. This goal has been accomplished as a direct consequence of implementing the classifier that was recommended [30].

For the purpose of anticipating Swine Flu symptoms, this system makes use of an implementation that is based on artificial neural networks. Even though there are 10 separate tones being sent into the network in Fig. 2, it only produces two distinct tones. The missing layer, which is composed of two hidden levels, has had a total of eight nodes removed from it. The artificial neural network (ANN) is a type of classifier that sorts incoming data by employing a multilayer feed forward back propagation method. The technique of modeling makes use of the activation function that is associated with the sigmoid function. Using this procedure, swine flu can be identified in one of two different ways: Training and tests are included in this package. Each individual neural network must be trained on all ten characteristics that comprise a record in order to function properly. Please be aware that this is an activity that will be logged.

The Table IV show the parameters of Artifical Nural Network(ANN), it takes 10 input nodes, two output nodes , two hidden layer(each hidden layer contain four nodes)and used backpropagation method. The Activation funcation is Sigmoid funcion.

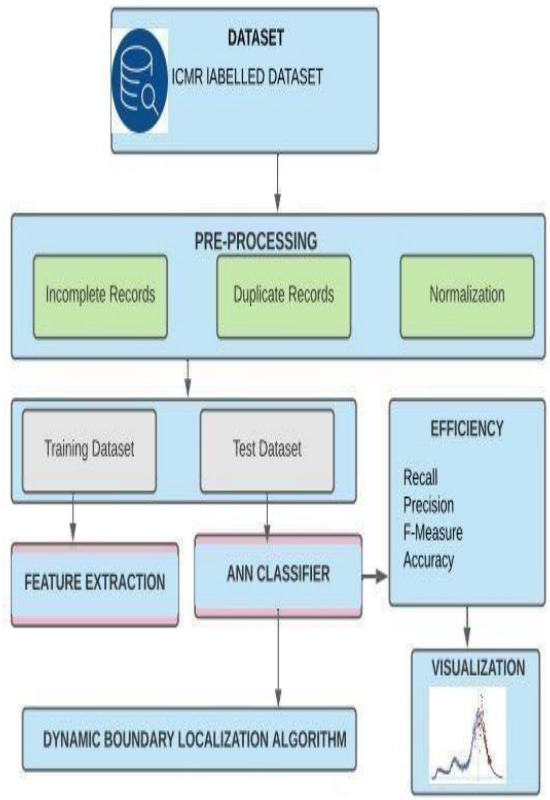

Fig. 1. Proposed System Architecture.

TABLE IV. PARAMETERS OF THE ANN

Number of input nodes	10
Number of output nodes	2
Number of hidden layers	2
Number of nodes in hidden layer	8
Paradigm	Multilayer Feed-Forward Network (Backpropagation)
Weight updating rue	Delta weight
Activation function	Sigmoid function

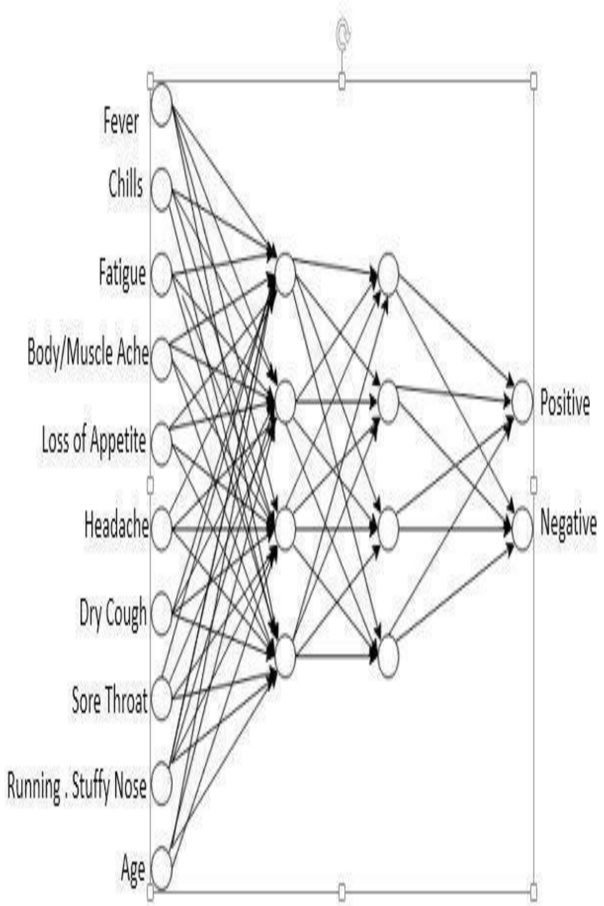

Fig. 2. Flow of ANN using Selected Features.

The Fig. 3 describes the how Artifical Neural Netwok(ANN) works.

If you go through the data set, you will learn about the classes. To categorise all records, this process is repeated as many times as necessary. Errors in training will be propagated backward and new penalties or delta error values will be applied if necessary. The figure shows how the algorithm for making predictions works in a visual way. Next, a random weight is supplied to each node in the artificial neural network to "initialize" it. Every node has been given a certain threshold. Calculating how much of a miscalculation occurred begins by determining the difference between projected and actual data. Efforts are taken to minimize the chance of making a mistake. An optimization condition must be met in order for checks to be performed. The output error can also be affected by errors that occur in the hidden layers. At every step of this procedure, it is necessary to use the gradient Descent function. It is important to meet the erroneous criteria in order to generate an accurate forecast for swine flu. A total of four data sets are available for training [31].

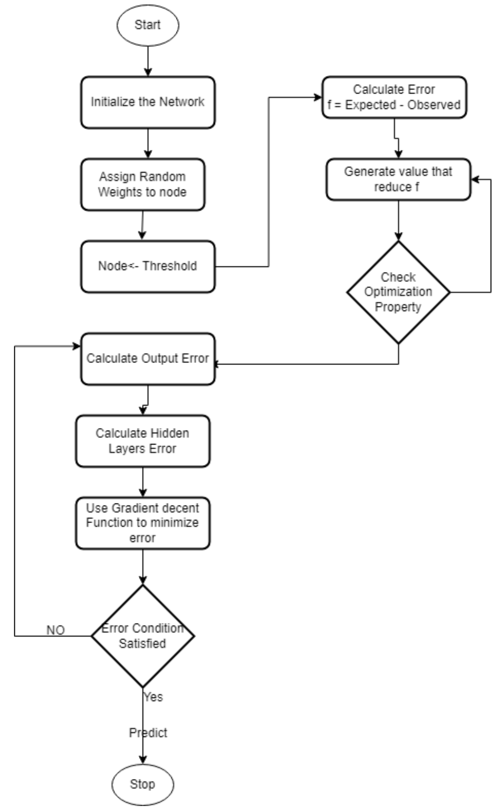

Fig. 3. Flowchart of the Proposed Algorithm.

Proposed Predicting 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

For the suggested classifier, a random seed is used to initialize the ANN network's weights. Each node has its unique threshold. The node is activated when the reset threshold is met. The total number of mistakes in each node's output is calculated. By using it, we may get an accurate estimate of the variance between what was expected and what was really achieved, which is known as the error value. Each node in the network calculates the ANN's error. By utilizing the new value, we may achieve a reduced level of node-level inaccuracy. Errors should be kept to a minimum as the primary goal of the project. To determine hidden layer errors, the output error is first decreased.

The gradient descent function has been used to lower the total amount of error in our calculations. It is still possible for the gradient descent process to reach a point of convergence even while the data cannot be separated linearly. Each time an incorrect number is rectified, a little amount is deducted from the overall sum. When an error number rises, the error value that was there just before the rise is displayed on the screen as well. Because of this, the function is referred to as "decent." The gradient descent technique can be used to break out of a loop if an error continues to occur [32].

For starters, only 20% of the data is really used for training. Our second iteration uses the first 20% of data, as well as the remaining 20% of data. More than half of the records are still in use in the third iteration. We were able to use about 80% of the data in the final dataset. Approximately 20% of the entries in the database are taken from which a sample of the full database may be utilized for testing. At the conclusion of each cycle, the accuracy or detection rate is shown in the Table V. With a 96 percent accuracy rate, the classifier is clearly doing its job. For your perusal, Table 5 displays the matrix of perplexity. There were 39450 positive test findings, and 1861 27 negative test results were uncovered in this investigation. The values are shown in their proper

Training Data Accuracy Rate

places on the table. Classifier's findings from past study are summarized in the table that follows. 96 percent of the time, the proposed neural network-based classifier works as expected. The Fig. 4 shows the accuracy and detection graph [33].

Testing data Confusion Matrix:

It is standard practice to use Table VI a confusion matrix to explain the performance of a classification model (also known as a "classifier") on a set of test data for which the actual values are known and finally Table VII shows the classifier values accuracy is 0.96[34].

Classifier Values:

TP: True Positive

TN: True Negative

FP: False Positive

FN: False Negative

Performance of Training and Testing

The Fig. 6 shows the performance of training and testing i.e., Training loss and epochs,

The Accuracy rate of Training and Testing, here we execute code in python programming the Epoch are from one to five. Here the final loss 0.1111 and accuracy rate is 96% in Fig. 5.

TABLE V. SYSTEM ACCURACY RATE

Ieration	Upper value for -ve case	Lower valuefor +ve case	Required Threshold value	Number of Records per iteration	No. of correctly classifiedpatterns	Accuracy foreach partition
1	0.287656	0.565139	0.4263975	202310	192196	78.00%
2	0.3315276	0.524858	0.4281928	404621	192194	76.00%
3	0.7904	0.379523	0.584976	606932	212425	84.00%
4	0.878156	0.335547	0.6068515	809243	232656	96.00%

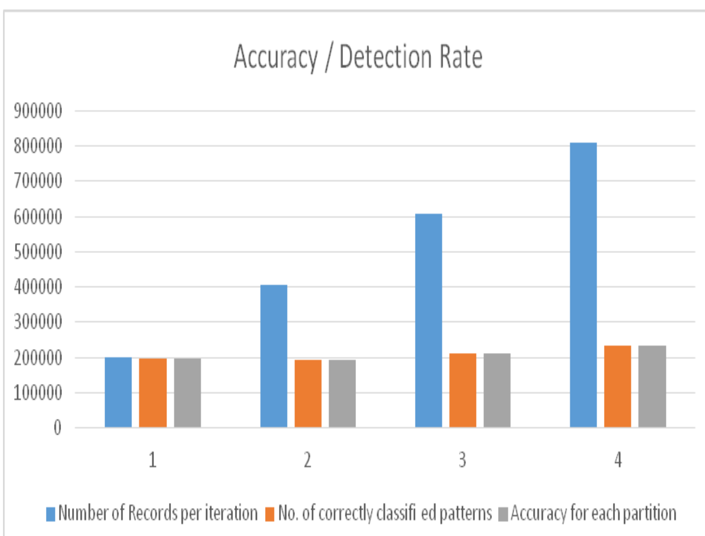

Fig. 4. Graph of Accuracy/Detection Rate.

TABLE VI. CONFUSION MATRIX

Correct P	Incorrect N	Result
39450	9217	positive
10184	186127	negative

TABLE VII. CLASSIFIER VALUES

Recall	TP/(TP+TN)	0.17488
Precision	TP/(TP+FP)	0.79481
F-measure	(2*Precision*Recall)/(Precision-Recall)	0.448440
Accuracy	(TP+TN)/(TP+FN+FP+TN)	0.96690

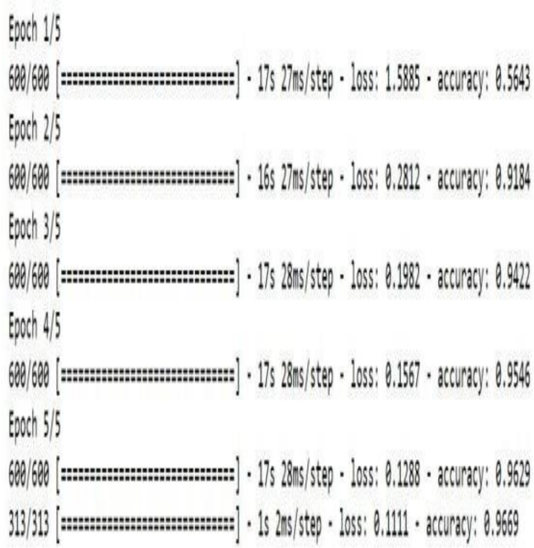

Fig. 5. Results of Training and Testing.

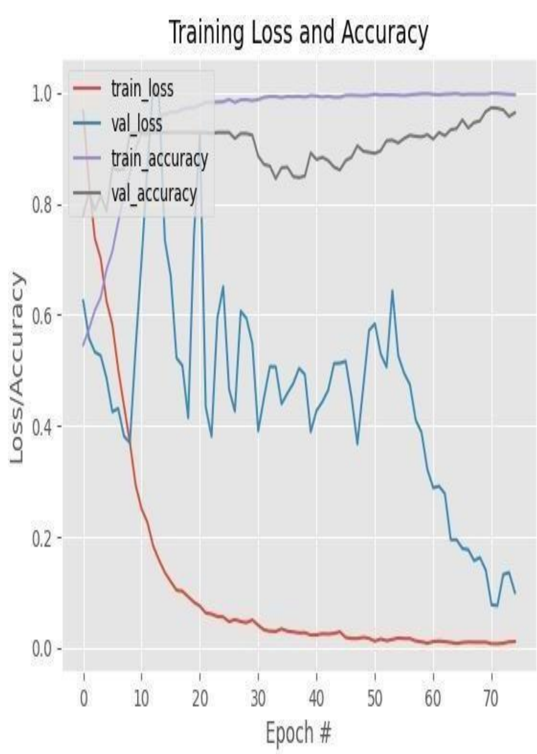

Fig. 6. Performance of Training and Testing.

IV. PROPOSED DYNAMIC BOUNDARY LOCATION ALGORITHM

Pseudo code

Algorithm Dynamic Boundary Location Algorithm INPUT

ICMR Dataset, Distance D=50, GIS Co-ordinates
L1 Begin
L2 Read-> Training Data, Distance, GIS Co-ordinates L3 For
L4 Each training data record L5 Do
L6 {
L7 Mark first data point ->X
L8 Identify points B1, B2, and B3 -> D/2 from X L9 Plot point ->GIS
L10}
L11 End Do
L12 if region of B1 B2 B3 locate ->X1 L13 X1<- Shortest Distance point-> X L14 Else
L15 X1<- Outside region B1 B2 B3
L16 Locate nearest Boundary point <- X1 L17 Repeat L5 to L10
L18 until END of training data record

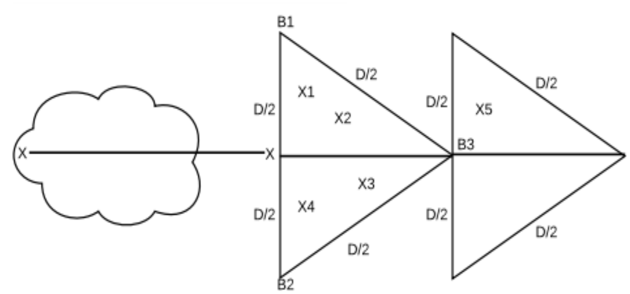

Fig. 7. Working of Proposed Dynamic Boundary Localization Algorithm.

This Fig. 7 article provides an explanation of the Dynamic Boundary Localization technique. To phrase it in a different way: Imagine that X is the location where the first case of the flu or swine flu was found. As a point of reference, Google Maps is utilized, and the value of the country/region/Distance area is selected based on the location of the point X. Let's say D represents a distance of 50 kilometers in this scenario. From point X, mark the locations of points B1, B2, and B3 at a distance of D/2 in the north, south, and east directions, respectively. This location houses the outermost cluster. The maximal possible spread of the illness. Find the data point that is geographically closest to each point (x1, x2, etc.) that makes up the triangulated region so that you may create tiny clusters or joint clusters. If a data point is located outside of the region that was triangulated, this indicates that the disease has spread to a location that is not included in the triangulation. In this stage, the algorithm will need to dynamically expand the area. This is accomplished by looking for the boundary point that is geographically closest to the data point that is located outside the boundary. The boundary points X3 and B3 are the ones that are located the most closely to one another. It is necessary to repeat the stages from B3 to B4 to B5 to B6 once more. The algorithm first creates a zone in the shape of a triangle, and then it expands to the north, south, and east sides of the triangle. If it grows on just three of the globes or Earth's sides, then it will encompass all of the other four sides [35].

V. Results And Discussion

The suggested classifier has a consistent upward trend over the length of the learning process. According to Table IV, there was a success rate of 78% in the first 20% of data that were looked at. The right classification was applied to a total of 197252 items. The accuracy of the classifier increases from 80 percent after the third iteration to 84 percent after the fourth iteration, and then to 92 percent after the fifth iteration. We can ensure that our classifiers continue to improve as we increase the number of iterations that we feed into them as well as the quantity of training data that we feed into them by utilizing the techniques of machine learning [36]. This is an easy assignment to do because the ANN classifier that was explained has an accuracy rate of 96 percent. Fig. 4 provides a visual representation of the ANN (artificial neural network) prediction classifier detection rate. Table V presents the matrix of ambiguity for your review and consideration. In the database, there are a total of 39450 positive entries in addition to 186127 negative records [37].

The Table VIII show the comparison of results with existing literature i.e., it will be compared with Machine learning algorithms and get the accuracy results. The Dynamic Boundary Location Algorithm (DBLA) and finally we get the 96 accuracy.

TABLE VIII. Comparison of Results with Existing Literature

Literature	Approach	Accuracy %
[12] Kakulapati et al. 2020	ANN and Random Forest algorithm	87
Proposed D B L A classification method		**96**
[8] Li et al. 2013	Bayesian and Markov network	83
[20] Xue et al. 2018	Regression and ANN	81
[21] Biswas et al. 2015	ANN classifier	78
[22] Srinivas et al. 2018	Naive Bayesian Classifier	84
[3] Kostkova et al. 2014	Cross correlation	76
[23] Volkova et al. 2017	Neural networks	87
[24] Raval et al. 2016	feed-forward neural network construction	73
[25] Singh and Kaur et al	support-vector-regression	91
[26] Tate et al. 2017	Random forest	78
[27] Byrd et al. 2016	Web based sentimental analysis Twitter	86
[28] Xue et al. 2019	support-vector-regression	89
[29] Rao et al. 2021	Hybrid voting algorithm	73

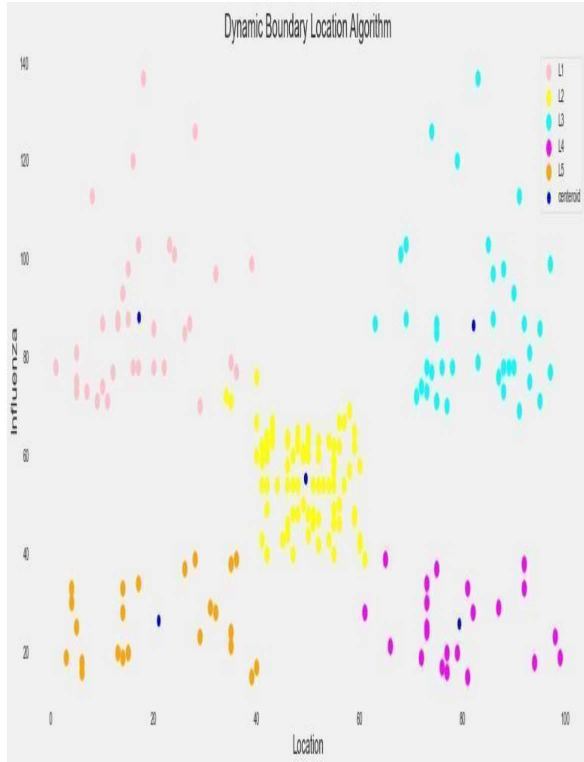

Fig. 8. Results of the Second Technique was Provided for Dynamic Border Localization.

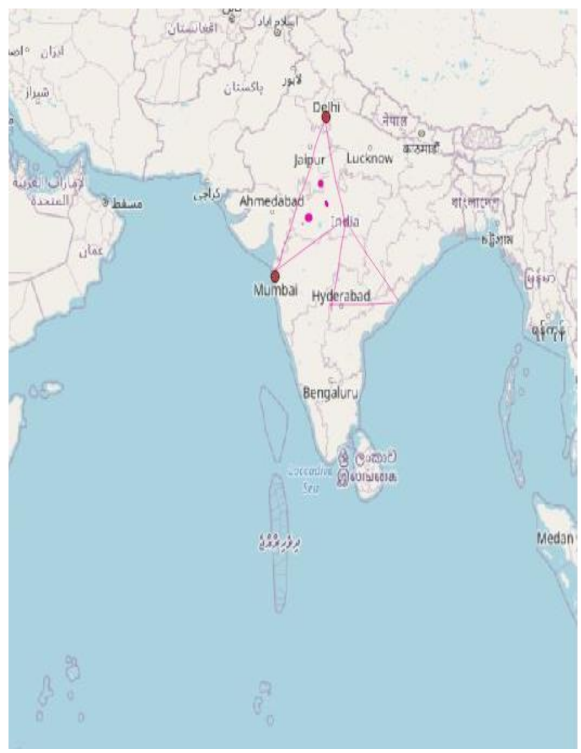

Fig. 9. The Output from Both the Existing Standard Literature and the GIS Tools.

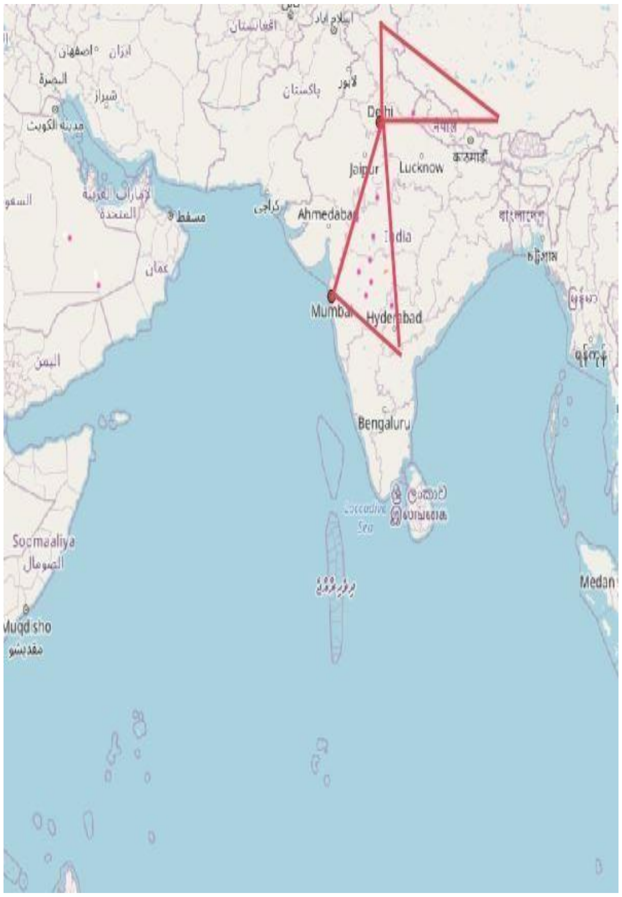

Fig. 10. Localization of Confirmed Swine Flu Cases using Proposed Dynamic Boundary Location Algorithm.

The total number of genuine positives was 9217, whereas the number of erroneous negatives was 10184. The method for predicting the spread of swine flu has an accuracy rate of 96 percent overall. According to the findings, the accuracy of the classifier is measured at 17.4 percent recall. According to the findings of the test, the accuracy was found to be 79.48 percent, and the F-measure was found to be 44.84 percent. The results of the experiments are presented in Table VII, which also serves as a comparison of the experimental outcomes of the recommended classifier with previously published research. The ANN classification algorithm that has been presented in this article cannot be compared in any way to the approaches that have been discussed in the previous paragraphs. Fig. 8 illustrates the results of the second technique that was provided for dynamic border localization. This method was proposed before [38]. The Fig. 9 and Fig. 10 show the output from both the existing standard literature and the GIS tools that are now available thanks to the method's superior area localization accuracy and GIS co-ordinate precision.

VI. Conclusions

The use of localization for predictive purposes about swine flu is covered in great detail in the study. Both prediction and location may be accomplished with the use of algorithms. A back propagation classifier based on an artificial neural network should be considered as a first step. The Indian Council of Medical Research (ICMR) is the source of the data that feeds the algorithm. The classification process is broken up into two distinct steps by the classifiers. In the first step, instances of swine flu are identified, and in the second stage, positive cases are localized with the use of an algorithm known as the dynamic boundary. The locations in the world where this extremely contagious virus is spreading may be easily identified, and containment zones may be set up in regions where the prevalence of the disease is high. Eighty percent of the data in the dataset is utilized in the training of the classifier. During the testing phase, there are an infinite number of potential combinations and permutations that might result in a record testing positive or negative. In order to assess the level of accuracy achieved by the methodology, a wide variety of well-established machine learning approaches were used as benchmarks. The recommended classifier has a detection rate that was significantly higher than the detection rates of the typical classifiers that are currently being utilized. The likelihood that the algorithm has gotten the answer correct is 96 percent. In order to evaluate the precision of the dynamic boundary method, many different GIS tools were utilized. It has been demonstrated that the localization accuracy achieved using the dynamic boundary technique is on par with that achieved using more conventional GIStools.

Reference

[1] J. Li and C. Cardie, " Early Stage In fluenza Detection from Twitter," *arXiv:1309.7340 [cs]*, Nov. 2013, Accessed: Nov. 25, 2021. [Online]. Available: http://arxiv.org/abs/1309.7340.

[2] R. Singh and R. Singh, " Applications of sentiment analysis and machine learning techniques in disease outbreak prediction – A review," Materials Today: Proceedings, p. S2214785321032764, May 2021, doi: 10.1016/j.matpr.2021.04.356.

[3] P. Kostkova, M. Szomszor, and C. St. Louis, "#swineflu: The Use of Twitter as an Early Warning and Risk Communication Tool in the 2009 Swine Flu Pandemic," ACM Trans. Manage. Inf. Syst., vol. 5, no. 2, pp. 1–25, Jul. 2014, doi: 10.1145/2597892.

[4] R. Chen, W. Zhong, H. Yang, H. Geng, X. Zeng, and B. Yu, "Faster Region-based Hotspot Detection," in Proceedings of the 56th Annual Design Automation Conferen ce 2019, Las Vegas NV USA, Jun. 2019, pp. 1–6. doi: 10.1145/3316781.3317824.

[5] A. López Pineda, Y. Ye, S. Visweswaran, G. F. Cooper, M. M. Wagner, and F. (Rich) Tsui, "Comparison of machine learning classi fiers for in fluenza detection from emergency department free-text reports," Journal of Biomedical In fo rmatics, vol. 58, pp. 60–69, Dec. 2015, doi: 10.1016/j.jbi.2015.08.019.

[6] Laura K. Borkenhagen, Martin W. Allen & Jonathan A. Runstadler (2021) In fluenza virus genotype to phenotype predictions through machine learning: a systematic review, Emerging Microbes & Infections, 10:1, 1896-1907, DOI: 10.1080/22221751.2021.1978824.

[7] Y. Shen and M. Gao, " Dynamic Routing on Deep Neural Network for Thoracic Disease Classification and Sensitive Area Localization," arXiv:1808.05744 [cs], Aug. 2018, Accessed: Dec.07,2021.[Online]. Available: http://arxiv.org/abs/1808.05744.

[8] Z. Li et al., "Thoracic Disease Identi fication and Localization With Limited Supervision," p. 10. 2013.

[9] W. Huang, C. Yang, and T. Hou, " Spine Landmark Localization with combining of Heatmap Regression and Direct Coordinate Regression," p. 6.

[10] R. Amer, M. Frid-Adar, O. Gozes, J. Nassar, and H. Greenspan, " COVID-19 in CXR: from Detection and Severity Scoring to Patient Disease Monitoring," IEEE J. Biomed. Health Inform., vol. 25, no. 6, pp. 1892–1903, Jun. 2021, doi: 10.1109/JBHI.2021.3069169.

[11] C. Park and H.-I. Suk, " Deep Joint Learning of Pathological Region Localization and Alzheimer's Disease Diagnosis," arXiv:2108.04555 [cs], Aug. 2021, Accessed: Dec. 07, 2021. [Online]. Available:

http://arxiv.org/abs/2108.04555.

[12] V. Kakulapati, V. K. Kumar, V. S. Srikar, Y. S. Rao, and T. O. Edoh, " An intelligent fram ework of Swine flu st atus prediction byrain forest algorithm," p. 8.

[13] D. Bhatt, D. Vyas, M. Kumhar, and A. Patel, "Swine Flu Predication Using Machine Learning," in In fo rmation and Communication Technology for Intelligent Systems, vol. 107, S. C. Satapathy and A. Joshi, Eds. Singapore: Springer Singapore, 2019, pp. 611–617. doi: 10.1007/978-981-13-1747- 7_60.

[14] A. B. Lawson, " Hotspot detection and clustering: ways and means," Environ Ecol Stat, vol. 17, no. 2, pp. 231–245, Jun. 2010, doi: 10.1007/s10651-010-0142-z.

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

[16] N. Li and D. D. Wu, " Using text mining and sentiment analysis for online fo rums hotspot detection and forecast," Decision Support Systems, vol. 48, no. 2, pp. 354–368, Jan. 2010, doi: 10.1016/j.dss.2009.09.003.

[17] M. U. Ali et al., "Early hotspot detection in photovoltaic modules using color image descriptors: An infrared thermography study," Int J Energy Res, p. er.7201, Aug. 2021, doi: 10.1002/er.7201.

[18] J. W. Park, A. Torres, and X. Song, "Litho-Aware Machine Learning fo r Hotspot Detection," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst., vol. 37, no. 7, pp. 1510–1514, Jul. 2018, doi: 10.1109/TCAD.2017.2750068.

[19] H. Yang, J. Su, Y. Zou, Y. Ma, B. Yu, and E. F. Y. Young, "Layout Hotspot Detection With Feature Tensor Generation and Deep Biased Learning," IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst., vol. 38, no. 6, pp. 1175–1187, Jun. 2019, doi: 10.1109/TCAD.2018.2837078.

[20] H. Xue, Y. Bai, H. Hu, and H. Liang, " Influ enza Activity Surveillance Based on Multiple Regression Model and Artifi cial Neural Network," IEEE Access, vol. 6, pp. 563–575, 2018, doi: 10.1109/ACCESS.2017.2771798.

[21] S. Kr. Biswas, B. Baruah, B. Purkayastha, and M. Chakraborty, " An ANN based Classification Algorithm for Swine Flu Diagnosis,"IJKBCS, vol. 3, no. 1, 2015, doi: 10.21863/ijkbcs/2015.3.1.005.

[22] P. Srinivas, D. Bhattacharyya, and D. M. Chakkaravarthy, " An Artificial Intelligent based System for Efficient Swine Flu Prediction using Naive Bayesian Classifier," IJCRR, vol. 12, no. 15, pp. 134–139, 2020, doi: 10.31782/IJCRR.2020.121519.

[23] S. Volkova, E. Ayton, K. Porterfi eld, and C. D. Corley, " Forecasting influen za-like illness dynamics for military populations using neural networks and social media," PLoS ONE, vol. 12, no. 12, p. e0188941, Dec. 2017, doi: 10.1371/journal.pone.0188941.

[24] D. Raval, D. Bhatt, M. K. Kumhar, V. Parikh, and D. Vyas, "Medical Diagnosis System Using Machine Learning," vol. 7, no. 1, p. 6,2015.

[25] S. Singh and H. Kaur, " Influenza prediction from social media texts using machine learning," J. Phys.: Conf. Ser., vol. 1950, no. 1,p. 012018, Aug. 2021, doi: 10.1088/1742-6596/1950/1/012018.

[26] A. Tate, U. Gavhane, J. Pawar, B. Rajpurohit, and G. B. Deshmukh, "Prediction of Dengue, Diabetes and Swine Flu Using Random Forest Classification Algorithm," vol. 04, no. 06, p. 7.

[27] K. Byrd, A. Mansurov, and O. Baysal, " Mining Twitter data fo r influ enza detection and surveillance," in Proceedings of the International Workshop on Software Engineering in Healthcare Systems, Austin Texas, May 2016, pp. 43–49. doi: 10.1145/2897683.2897693.

[28] H. Xue, Y. Bai, H. Hu, and H. Liang, " Regional level influenza study based on Twitter and machine learning method," PLoS ONE, vol. 14, no. 4, p. e0215600, Apr. 2019, doi: 10.1371/journal.pone.0215600.

[29] N. T. Rao, D. Bhattacharyya, E. S. N. Joshua, and C. V. Satyanarayana, "Prediction of Swine Fluusing a Hybrid Voting Algorithm," p. 9, 2021.

[30] P.NAGARAJ, Dr.A.V.Krishna Prasad, "Survey on Swine flu Prediction", International Journal of Management, Technology and Engineering, Volume IX, Issue V,May/ 2019 , ISSN No:2249-7455,Page no: 937-941.

[31] P.Nagaraj , Rajesh Banala and A.V.Krishna Prasad, " Real Time Face Recognition using Effective Supervised Machine Learning Algorithms", Journal of Physics: Conference Series 1998 (2021) 012007 IOP Publishing doi:10.1088/1742-6596/1998/1/012007.

[32] P. Nagaraj and Dr A. V. Krishna Prasad, "A Novel Technique to Detect the HotspotsSwine Flu Effected Regions", Published in: 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO),15November2021DOI:10.1109/ICRITO513 93.2021.9 596422, Electronic ISBN:978-1-6654- 1703-7 CD:978-1-6654-1702-0.

[33] Nagaraj P, Dr A.V. Krishna Prasad, " A Cloud Computing Emerging Security Threats and Its Novel Trends in Knowledge Management Perception", International Journal of Emerging Technologyand Advanced Engineering Website: www.ijetae.com (ISSN 2250- 2459, ISO 9001:2008 Certified Journal, Volume 7, Special Issue 2, December 2017).

[34] Nagaraj P,Rohit Kumar K, Rajesh Banala b," Energy effici ent 2 tier data aggregation scheme in- Sensor networks", Accepted 7 March 2021,https://doi.org/10.1016/j.matpr.2021.03.1402214-7853/ 2021 Elsevier Ltd.scientific committee of the Emerging Trends in Materials Science, Technology and Engineering.

[35] P. Nagaraj, Gunta Sherly Phebe, Anupam Singh, "A Novel Technique to Classify Face Mask for Human Safety ", 2021 Sixth ICIIP Published in: 2021 Sixth International Conferen ce on ImageIn formation Processing (ICIIP),26-28 Nov. 2021, 10 February 2022.DOI:10.1109/ICIIP53038.2021.9702607Publisher: IEEE Conference Location: Shimla, India.

[36] P. Nagaraj Rajesh B anala b, Vicky Nair " Perform ance of Secure Data Deduplication Framework in Cloud"(ICAIDS-2022) organized by Sreyas Institute of Engineering and Technology, Hyderabad, Telengana, India during March 11-12, 2022.

[37] P. Nagaraj, ChennaSriya, SarikondaMahidhar Raju, SabaNaazneen Kauser,Chintala Rakesh " SMART Street – An Artificial Intelligence(AI) Powered Street Garbage Detection & Alert System",Design Engineering ISSN: 0011-9342 | Year 2021, Issue: 9, Pages: 18131-18141.

[38] P. Nagaraj, P. Sahith Krishna, P.Shiva Sai " Forcasting Cyber Attacks Using Machine Learning" Journal of Opto Electronics Laser, ISSN: 1005-0086, DOI: 10050086.2002.07.66, Volume 41, Issue 7, 2022, Pages: 550-556.