Journal of Theoretical and Applied Information Technology
30th November 2022. Vol.100. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SWINE FLU HOTSPOT PREDICTION IN REGIONS BASED ON DYNAMIC HOTSPOT DETECTION ALGORITHM

P.NAGARAJ[1], DR.A.V.KRISHNA PRASAD[2], Dr.M.VENKATA DASS[3], KALLEPALLI ROHIT KUMAR[4]

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

[2] Associate Professor in IT Department, Maturi Venkata Subba Rao Engineering College, Hyderabad, India,

[3]Assistant Professor in CSE, University College of Engineering (A), Osmania University, Hyderabad, India,

[4]Assistant Professor in CSE, Koneru Lakshmaiah Education Foundation,Vijayawada, India.

E-mail: [1] nagarajcsedept@gmail.com, [2] kpvambati_it@mvsrec.edu.in, [3]m.venkatadass@osmania.ac.in, [4]krk542@gmail.com,

ABSTRACT

The purpose of this work is to provide a strategy for locating outbreaks of swine flu by making use of geographical data. Following their generation by the Dynamic Boundary Location Algorithm (DBLA), the localization coordinates are then sent on to the suggested KDH approach. Identifying swine flu cases, locating clusters of cases, and pinpointing their geographical locations using geographic information system coordinates are all accomplished with the help of the Dynamic Boundary Localization approach. A severity index is assigned to each illness, and data on its location is also recorded. The technique makes use of a Dynamic Hotspot detection algorithm (DHDA) which gives max number of patients death i.e., hotspot and Gaussian mixture model in conjunction with the Severity index. Data collected globally using Moran's I are what's utilized to calculate the severity index. The Gi instrument is what we use to determine how accurate the predicted hotspots are. This technique makes use of statistical analysis to provide a more accurate estimation of the locations in which swine flu is most common.

Keywords: *Swine Flu, Gaussian Mixture Model, Machine Learning, Global Moron, Gi*, Statistical Evaluation*

1. INTRODUCTION

The swine flu is highly infectious and has the potential to spread rapidly. There is a distinct possibility that both air and water might play a role in the spread of the disease. This particular strain of the influenza virus is sometimes referred to by the label H1N1. It makes no difference if the viral epidemic is \it will always result in the loss of human life. There is evidence that the influenza virus has caused human illnesses in a number of different countries. This chemical has an effect on the respiratory system of human beings.

According to information kept by the World Health Organization (WHO), between 5 and 15 percent of the world's population gets the influenza virus each year (WHO). It is believed that between 250,000 and 500,000 people have passed away as a result of respiratory illnesses brought on by this virus. [1] The process of identifying regions with abnormally high incidence rates for a certain event, often known as "spatial hotspots," is a frequent research technique and an important endeavour in a variety of domains. If you are seeking for locations in which a certain kind of event is more likely to occur, you may consider these regions to be "hotspots." For instance, epidemiologists utilize geographical hotspots to detect outbreaks of illness and deploy efficient treatments in order to limit the risks to public health. The field of public health encompasses the application of this practice. In the same way as "hotspots" in the food business identify locations with disproportionately high crime rates, "hotspots" in public safety identify places with disproportionately high crime rates, which can help investigations focus on "repeat offenders" or other core causes. On the subject of hotspot detection in clusters, several research have been carried out, and among these studies, comparisons have been made between approaches

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

a, b, c, and d and e. The kdh algorithm that was suggested possesses a number of significant benefits. First the hazard level of the spread of swine flu is defined by the algorithm as a region close to the epicenter that has a high chance of generating Swine Flu positive cases based on a geographical link. This is what the algorithm considers to be the most dangerous area. The decade algorithm that is based on spatial relationships first conducts an analysis and comparison of performance, and then it classifies the data. In the end, brand names are assigned to a raster layer and its individual pixels in a manner that is proportional to the degree of the condition. As a consequence of this, the kdh approach is particularly useful for visualization since it enables the simple identification of classes that generate high-density hotspots. This is an important benefit. Nevertheless, the dynamic boundary technique does not evaluate the statistical significance of the observed hotspots. [2] [3][4]. Instead, then using it to pinpoint specific hotspot locations, the kdh is more useful for visualizing patterns and trends. At this point in time, the provision of trustworthy techniques for evaluating hotspot detection using kdh data is really essential. Nevertheless, the statistical importance of the hotspots that were found was ignored in a great number of the earlier investigations.

2. LITERATURE REVIEW

Manu Jiang1 et al. 2019[5] conducted a study to locate and identify HIV-1 epidemic hotspots. The authors recognized pathways and patterns of transmission of HIV in countryside areas. The prevention strategies of mature people were developed. Li et al. 2020[6] Suggested a method for performing cluster analysis using Gaussian Mixture Model (GMM). A compact index (CI) is anticipated to authenticate the clustering outcomes rationally. An adjacent association was detected among the GMM clustering mechanism and the CI evaluation index. Ghan et al. 2009 [7] suggested techniques for hotspot detection implementing clustering and pattern matching. The hotspot fragments have been organized into manageable groups. A rapid incremental clustering method is used to do this task with the same huge datasets in a time-efficient way. Each cluster is analyzed to provide a classification of a certain type of hotspots, and then a pattern matcher is developed utilizing the hotspot class reports as a basis to locate hotspots in freshly conceived scheme diagrams. In 2020, Khan et al. [8] proposed using deep hybrid neural networks for hotspot prediction. They used Convolution Neural Networks (CNN) with Long Short-Term Memories to get their results (LSTM). The technology makes use of convolution neural networks. As a next step, the characteristics are fed into the LSTM model. Short-term, medium-term, and long- term dependencies are all separately accounted for in the LSTM model. It was created by Sun et al. 2019 [9] to use the FCM clustering technique to examine hot spot zones using time variable characteristics. Damage indices are first extracted from the collected signals and used to build both the baseline model and the pre-existing FCM model, as proposed. Bandyopadhyaya, in 2014 [10], created the fuzzy centred cluster technique as a way to find hotspots with limited information. The technique of Hot-Spot Identification is used here. Using severity models and disaggregated crash account data, this technique calculates the probability of various accident severity outcomes depending on the major underwriting criteria. Numerous estimated amounts of severe and fatal crashes are based on these probabilities. The places are divided into two vague categories based on these totals. Guleria et al. 2022 [6] accomplished the timely treatment prediction for the covid-19 infection thanks to their empirical inquiry on classifiers. In order to predict mortality and response rates among people with infectious illnesses, the suggested method of Collective Classification employs Machine Learning techniques. The classifiers predicted the severity of the patients' illnesses across multiple areas, which improved the capacity to organize resources and response care systems. The strategy competes with support vector machines, naive bayes, and decision trees. There has been no consideration given to the swine flu severity index in any of the research (SI). The Swine Flu Severity Index should be included in research results, according to the findings of several different investigations. The impacts of the disease, such as those of permanent disability, are very helpful to research and knowledge of the present situation of swine flu in the cluster. The swine flu is a highly contagious illness. The primary objective of the study is to evaluate the significance of the hotspots that were discovered using the kdh method in combination with SI. It focuses in on Hyderabad, which is the capital city of the Indian state of Telangana. The on-and-off patterns of taxi clients were discovered to have a correlation with the social functions of urban centers by researchers Qi and colleagues [14], which used data mining to make this discovery. This demonstrates that one way to determine the level of sociability present at a site is to measure the number of times passengers enter and exit taxis. It is essential to reference this term in your paper. Zheng et al. [15] conducted an investigation of the efficiency of urban planning by using GPS data collected from city taxis. They were looking for a way to strike a healthy balance between the city of Taipei, China's supply of taxis and the city's demand for them, and they found that urban planning was an efficient way to accomplish so. Chang et al. [16] employed geographical statistical analysis, data mining, and clustering to historical data. They then used the results to anticipate probable centers of taxi request activity based on the data. These various methods were utilized in the processing of the data Lee et al. [17] conducted research on taxi pickup patterns by analyzing historical trajectory data from the system. Phithakkitnukoon et al. provide a model in [18] that can be used to make predictions about the number of taxis that are wheelchair accessible. The model takes into account a variety of elements, including the day, the time of day, and the weather, among other things. Li and Kautz [19] developed a strategy for discovering passengers that is dependable as well as effective by doing research on a sizable urban taxi GPS dataset. As a result, this made the design of the system simpler. Liao et al. [20] investigated the activity level of users at each place in order to make an educated guess as to how drivers will conduct in the future. Zheng et al. [21-23] performed an analysis on the GPS data that was collected by individuals. Recent research carried out by Zhu and colleagues [24] provides an in-depth analysis of activity detection using trajectory data. Zheng and Zhou presented a computer analysis method in their paper [25] that makes use of spatial trajectories as one of its key components. A recommendation system was built by Yuan and his colleagues [26] in order to assist travelers in locating taxis that are conveniently located. In addition to this, Yuan et al. [27, 28] devised a technique for learning taxi routes. The spatio-temporal characteristics of cab GPS trajectories were utilized in the investigation of the anomalies by both Liu et al. [29] and Pang et al. [30]. The ever-changing nature of the taxi business in New York City was a major factor in Qian and his colleagues' [31] decision to do research on the sector. Zhu et al.[32] made the suggestion that a two-stage method should be utilized in order to establish whether or not a taxi cab was operating within the bounds of the law (occupied, non-occupied or parked). Taxi tracks were used as a stand-in for a variety of other urban land uses, and Pan et al. [33] described a novel way to categories them as such. The forecast for the current state of the traffic situation Wenet al. [34] conducted an analysis of the data collected from taxis equipped with GPS in order to determine the amount of traffic congestion that existed in Beijing during the Olympic Games. You may examine the outcomes of their work here. Pele and Norency [35] looked into the ways in which individuals in Montreal use taxis as part of their daily commutes by analyzing the GPS data collected by taxis. The objective of the research was to determine the degree to which taxis are responsible for the aforementioned patterns of behavior. The article that detailed their findings was published in the journal Transportation Research Part D: The findings of their investigation point to the possibility that taxis might play an important part in this subject. In order to analyze the current state of traffic in a number of European cities and to gather data on traffic patterns, the authors Schafer et al. [36] utilized automobiles that were outfitted with GPS technology. Because of this, they were able to carry out their research in a manner that was more effective. The authors Castro et al. [37] developed an approach that may be used to generate a model of traffic density. The data source for their research is comprehensive taxi traces. This methodology might be implemented. It was Castro and the others involved. This method could have some useful uses in the real world. Lu and Li [38] made their predictions about the impending OD using data obtained from the GPS devices installed in Singapore taxis. Retaining complete dominance and authority over the administration of the cab firm. Using a model developed by Chang et al. [39], one is able to make accurate predictions regarding the demand distribution for a taxi management system. The four components that make up this process are called data filtering, clustering, semantic annotation and hotness calculation respectively. This model was first conceived of with the intention of being utilized for the purpose of anticipating shifts in client interest. [40] In order to get insight into the routines of taxi drivers, Liu and his colleagues come up with a novel approach. A system that provides users with recommendations based on their previous travel itineraries was developed by Ge et al. [41] with the intention of assisting taxi drivers in reducing their costs for both money and petrol. The ever-increasing cost of gasoline was a driving factor in the government deciding to take this action. A thorough taxi assessment index was created by Zhang et al. [42] in order to evaluate the level of service provided by Shenzhen's taxi companies. Using GPS data, Zhang et al. [43] devised a method that they referred to as isolation-based anomalous trajectory (iBAT). The objective of this method was to identify taxi drivers that engage in driving behaviors that are not typical.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

KPIs were utilized by Amat et al. [44] in order to evaluate the level of productivity and profitability within the taxi sector.

3. METHODOLOGY

The credit-recommended strategy was initially used in the search for the hotspot, and it was successful. Within the framework of the Gaussian mixture model, the kdh approach may be uncovered at some point (GMM). The second phase consisted of establishing the statistical significance of the hotspots that were created by kdh by using both the local and global Moran's I approach. There was a natural hierarchy present that was determined by the seriousness of the threats. The Gi* data were utilized in order to conduct the investigation into the Hotspots [46]. The following is a rundown of how the remaining sections of the article are organized: In Section 2, we discuss the methodology of the study. In Section 3, we show instances of the technique in operation together with an algorithmic analysis. Section 1 introduces the problem. Section 2 discusses the methodology of the research. In the next section, "Section 4, We Provide a Brief Overview of the Research Results," and in the following section, "Section 5, We Present Our Findings and Conclusions Based on the Findings,"

3.1 KDH algorithm

Researchers were able to locate the precise location of the cluster of confirmed swine flu infections by employing a process known as dynamic boundary mapping. The approach, on the other hand, does not have the potential to do a statistical analysis of the significance of the hotspots that are detected. As a consequence of this, the authors present an approach that possesses the capability of enhancing the performance of the kdh algorithm [49]. Figure 1 demonstrates how we used the KDH approach in conjunction with an analysis of the statistical significance of the clusters and hotspots that were formed. A technique known as dynamic boundary analysis was utilized in the first round of efforts to determine the locations of people who had tested positive for swine flu. A family member or friend of the patient who is staying close may see an improvement in the patient's condition. As a result, the gathered events are categorized before being put to use to conduct out integration and get information regarding events that have place in the same areas. It is possible to piece together the events in this manner into a chronological narrative [52]. In addition, a substantial amount of weight was given to this SI in the process of finding the locations in where the Swine Flu was most common. Utilizing the kdh method allowed for the creation of the swine flu density map as well as its subsequent calculation. The following is a description of the procedures that need to be followed in order to evaluate whether or not the findings of an evaluation are statistically significant. The operation will be halted if there is sufficient evidence to show that the spread of Swine Flu infections follows a random pattern [53]. This will be determined by first examining the random nature of the distribution of Swine Flu diseases within a region. If there is sufficient evidence to show that the spread of Swine Flu infections follows a random pattern, then the operation will be halted. However, it is abundantly evident that the distribution does not follow a normal distribution; hence, the challenge now is how to compute the bandwidth in such a way that it maximizes clustering or autocorrelation. This was accomplished in particular through the utilization of the rising Global morons I. However, in order to go on to the next phase, it is necessary to compute the beginning distance at which a single coordinate has at least one matching neighbor. Only then can this phase be completed [50].

After determining an acceptable bandwidth in the first step, the second step is to. Within the context of the boundary estimate technique, the term "broadband," which is symbolized by the letter d, refers to the threshold distance input. Through the employment of the regional Moran's I, this results in the appearance of enormous clusters on the map. A density map and a map depicting major clusters have been combined to create this Swine Flu priority index. In order to assess the Quality of the newly created hotspot, the Gi* tool was utilized. After this, you will see some specifics that go into further depth on the KDH approach, as well as information regarding local and global moron I and Gi* [51].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

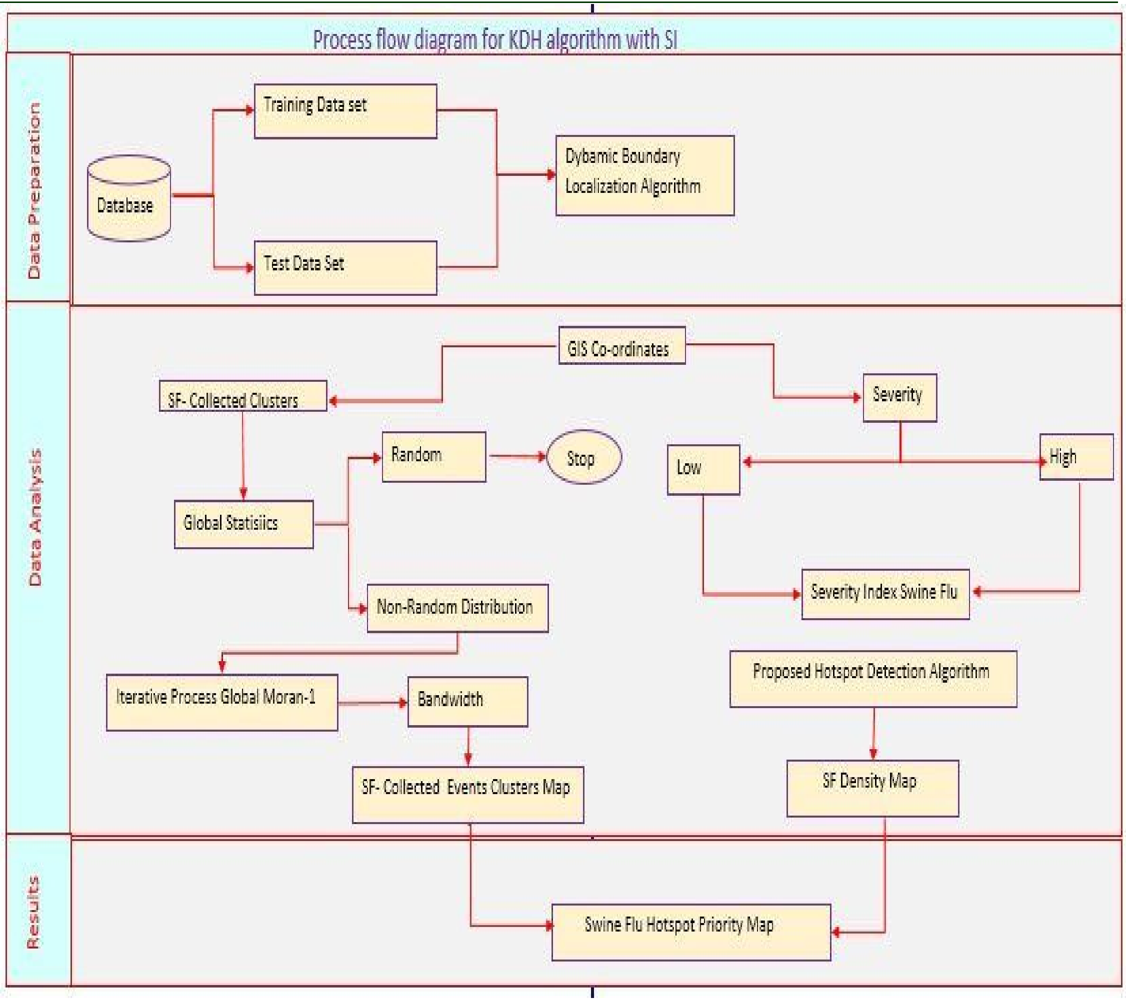

Figure 1 Flow of KDH algorithm

3.2 Dataset

The ICMR generously provided the dataset that was utilized in the investigation. The database has a total of 1264443 entries, which includes 122751 confirmed cases of swine flu as well as 115517 survivors. Between the years of 2015 and 2020, a total of 7,336 persons will have lost their lives within the borders of our country. Table 1 contains the most important aspects of the dataset that have been collected [55]

Table 1: Dataset

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

Journal of Theoretical and Applied Information Technology
30th November 2022. Vol.100. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

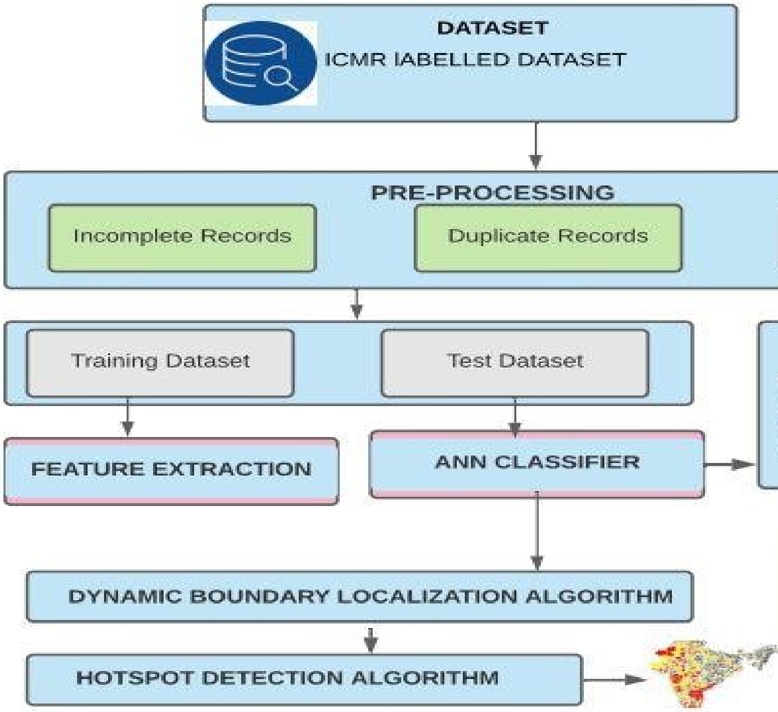

Figure 2 Proposed System Architecture

Proposed Dynamic Hotspot Detection Algorithm
Algorithm Hotspot DHDA
L1 Begin
L2 Read -> no of centroids //calculate no. of nodes in each centroids
L3 For i = 1 to n /*Compute distance Assign the points to the nearest function Colour points according to the function*/
L4 noc=count (cluster)
L5 nopi =noc
L6 print (noc) End for //Find max value of each cluster
L7 Initialize Max
L8 for i = 1 to n
L9 If nopi is greater than Max
L10 Max= nopi
L11 End for
L12 print Max
L13 END

- Read the number of centroids in location and calculate the number of nodes in each centroids.
- For each of the coordinates in the location we try to estimate the distance of the centroids.
- In each centroids we calculate the distance between nodes.
- Assign the points to the nearest functions color points according to the function.
- Count the number of centroids and print it.
- Count the number of points in each centroids
- Initially assign the nopi is max
- If any changes have taken place repeat the process until maximum value in each centroid
- Now in the cluster that is achieved the hotspot is identified based on the population of the cluster region.
- Hotspot color is filled with red color.

The figure shows the frequency of swine flu attack in India from 2015 to 2020[45].

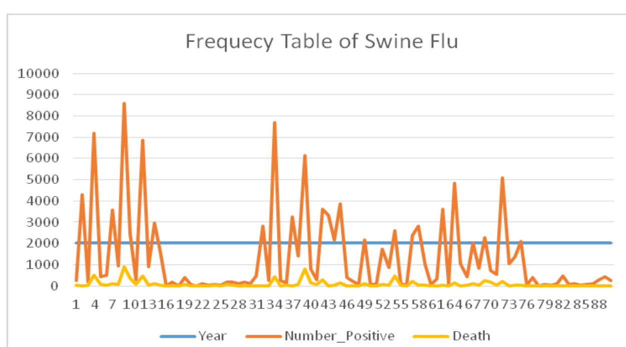

Figure 3 Cluster and hotspot formation using the proposed algorithm

Journal of Theoretical and Applied Information Technology
30th November 2022. Vol.100. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. RESULTS

By using Dynamic Boundary Location algorithm, we can find the location. In each location swine flu effected patients are find out. Below figure 4 and figure 5 locations of swine flu effected locations namely Warangal, Mahaboobnagar, Nalgonda, Rangareddy, Hyderabad.

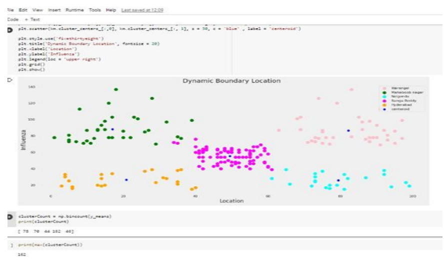

Figure4 Cluster and hotspot locations of Rangareddy

The above diagram contains swine flu affected locations in that maximum effected Region is Ranga Reddy districts. In that we are applying Dynamic Hotspot Detection Algorithm (DHDA), then we can get maximum number of dead ratio is called Hotspots. So finally the location marked as a RED color below shows result of DHDA.

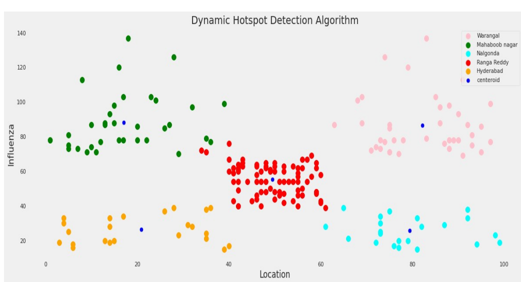

Figure 5 Clustering hotspots locations using DHDA

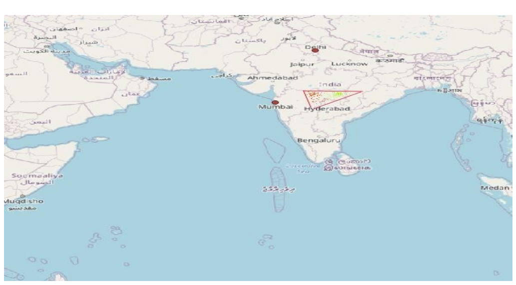

Figure 6 Clustering hotspots location using QGIS

Table 2 Evaluation results

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°

The researchers were able to make use of a comprehensive dataset in their inquiry because to the generosity of the ICMR, which made it possible for them to do so. The database has a total of 1264443 records; of them, 122751 are confirmed cases of swine flu, while 115517 are recorded as having survived the virus. Within our boundaries, a total of seven thousand three hundred thirty-six lives will have been lost by the year 2020, having occurred between the years of 2015 and 2020 [47]. The most notable features of the assembled dataset are summarized in Table 1, which contains a summary of those features.

3. CONCLUSION

The research presents a summary of the attempts that have been undertaken up to this point to anticipate swine flu hotspots using geographical data. The DHDA and kdh approach that was described here is used to make the forecast, which was done with its assistance. The classifier is able to forecast the location of the next epidemic depending on the severity of the ill individual by using geographical data, Moran I statistics, and the Gi* tool as inputs in the prediction process [54]. The severity of swine flu symptoms, as well as the mortality rate associated with such symptoms, is both taken into account by the algorithm before it makes its prediction. The authors of the earlier study focused on two different methods: one for detecting swine flu, and the other for determining a person's geographic location. The first approach utilizes an artificial neural network-based back propagation classifier as its foundation. The dataset compiled by the ICMR documenting cases of swine flu in India is utilized as an input for the algorithm.

The procedure of the classifier may be broken down into two distinct stages. The first step is to identify individual cases of swine flu, and then we apply the dynamic boundary approach to narrow down the specific locations of those individual cases. It is feasible to identify particular localities that are undergoing an epidemic of this highly contagious illness [48]. In hotspots—areas where the disease is quickly spreading—containment zones can be formed to limit the disease's ability to spread across a wider geographic area. Eighty percent of the data set is utilized in the training process for the classifier. During the phase of testing, the classifier is able to pick up on and learn from any and all possible permutations that have led to a positive or negative result in previous tests. During this phase, the classifier is also able to learn from any and all possible permutations that have led to an incorrect result. The detection rate of the program was analyzed with a variety of algorithms, some of which are standard in the field of machine learning. These algorithms include logistic regression, KNN, and SVM. The recommended classifier has a higher detection rate than other classifiers that were considered to be state-of-the-art. The algorithm obtained 96% accuracy. The DHDA, kdh method and Gi* work together to locate the hotspots by making use of the localization data that is provided by the dynamic boundary technique that is recommended.

REFERENCES:

[1] R. Krueger, P. Bansal, and P. Buddhavarapu, “A new spatial count data model with Bayesian additive regression trees for accident hot spot identification,” Accident Analysis & Prevention, vol. 144, p. 105623, Sep. 2020, doi: 10.1016/j.aap.2020.105623.

[2] H. Yu, P. Liu, J. Chen, and H. Wang, “Comparative analysis of the spatial analysis methods for hotspot identification,” Accident Analysis & Prevention, vol. 66, pp. 80–88, May 2014, doi: 10.1016/j.aap.2014.01.017.

[3] T. Yabe, K. Tsubouchi, and S. V. Ukkusuri, “Early Warning of COVID-19 Hotspots using Mobility of High Risk Users from Web Search Queries,” arXiv:2010.13254 [physics], Oct. 2020, Accessed: May 04, 2022. [Online]. Available: http://arxiv.org/abs/2010.13254

[4] Z. Yao, J. Ge, Y. Wu, X. Lin, R. He, and Y. Ma, “Encrypted traffic classification based on Gaussian mixture models and Hidden Markov Models,” Journal of Network and Computer Applications, vol. 166, p. 102711, Sep. 2020, doi: 10.1016/j.jnca.2020.102711.

[5] J. Jiang et al., “A Geographic Hotspot and Emerging Transmission Cluster of the HIV-1 Epidemic Among Older Adults in a Rural Area of Eastern China,” AIDS Research and Human Retroviruses, vol. 36, no. 9, pp. 712–720, Sep. 2020, doi: 10.1089/aid.2019.0293.

[6] Y. Li, J. Zhang, Z. Ma, and Y. Zhang, “Clustering Analysis in the Wireless Propagation Channel with a Variational Gaussian Mixture Model,” IEEE Trans. Big Data, vol. 6, no. 2, pp. 223–232, Jun. 2020, doi: 10.1109/TBDATA.2018.2840696.

[7] J. Ghan et al., “Clustering and pattern matching for an automatic hotspot classification and detection system,” San Jose, California, USA, Mar. 2009, p. 727516. doi: 10.1117/12.814328.

[8] S. D. Khan, L. Alarabi, and S. Basalamah, “Toward Smart Lockdown: A Novel Approach for COVID-19 Hotspots Prediction Using a Deep Hybrid Neural Network,” Computers, vol. 9, no. 4, p. 99, Dec. 2020, doi: 10.3390/computers9040099.

[9] H. Sun, J. Yi, Y. Xu, Y. Wang, and X. Qing, “Crack Monitoring for Hot-Spot Areas Under Time-Varying Load Condition Based on FCM Clustering Algorithm,” IEEE Access, vol. 7, pp. 118850–118856, 2019, doi: 10.1109/ACCESS.2019.2936554.

[10] R.Bandyopadhyaya and S. Mitra, “Fuzzy Cluster–Based Method of Hotspot Detection with Limited Information,” Journal of Transportation Safety & Security, vol. 7, no. 4, pp. 307–323, Oct. 2015, doi: 10.1080/19439962.2014.959583.

[11] P. Guleria, S. Ahmed, A. Alhumam, and P. N. Srinivasu, “Empirical Study on Classifiers for Earlier Prediction of COVID-19 Infection Cure and Death Rate in the Indian States,” Healthcare, vol. 10, no. 1, p. 85, Jan. 2022, doi: 10.3390/healthcare10010085.

[12] J. Tang, J. Hu, Y. Wang, H. Huang, and Y. Wang, “Estimating hotspots using a Gaussian mixture model from large-scale taxi GPS trace data,” Transportation Safety and Environment, vol. 1, no. 2, pp. 145–153, Nov. 2019, doi: 10.1093/tse/tdz006.

[13] J. Shi, Q. He, and Z. Wang, “GMM Clustering-Based Decision Trees Considering Fault Rate and Cluster Validity for Analog Circuit Fault Diagnosis,” IEEE Access, vol. 7, pp. 140637– 140650, 2019, doi: 10.1109/ACCESS.2019.2943380.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[14] Qi G, Li X, Li S, et al. Measuring social functions of city regions from large-scale taxi behaviors. In: IEEE International Conference on Pervasive Computing and Communications Workshops, Seattle, WA, USA, 2011, 384–8.

[15] Zheng Y, Liu Y, Yuan J, et al. Urban computing with taxicabs. In: 13th ACM International Conference on Ubiquitous Computing, Beijing, China, 2011, 89–98.

[16] Chang H, Tai Y, Chen H, et al. iTaxi: context-aware taxi demand hotspots prediction using ontology and data mining approaches. In: Proceedings of the 13th Conference on Artificial Intelligence and Applications (TAAI), Taipei, Taiwan, 2008.

[17] Lee J, Shin I, Park G. Analysis of the passenger pick-up pattern for taxi location recommendation. In: Fourth International Conference on Networked Computing and Advanced Information Management, Gyeongju, South Korea, 2008, 199–204.

[18] Phithakkitnukoon S, Veloso M, Biderman A, et al. Taxi-aware map: identifying and predicting vacant taxis in the city. In: International Joint Conference on Ambient Intelligence, Malaga, Spain, 2010, 86–95.

[19] Li B, Zhang D, Sun L, et al. Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset. In: IEEE International Conference on Pervasive Computing and Communications Workshops, Seattle,WA, USA, 2011, 63–8.

[20] Liao L FD, Kautz H. Extracting places and activities from GPA traces using hierarchical conditional random fields. Int J Robot Res 2007; 26:119–34.

[21] Zheng Y, Liu L, Wang L, et al. Learning transportation mode from raw GPS data for geographic applications on the web. In: Proceedings of the 17th International Conference on the World Wide Web, Beijing, China, 2008, 247–56.

[22] Y Z, X X, Ma W. Understanding mobility based on GPS data. In: Proceedings of the 10th ACM International Conference on Ubiquitous Computing. Seoul, South Korea 2008; 312–21.

[23] Zheng Y, Y C, Q L, et al. Understanding transportation modes based on GPS data for web applications. ACM T Web 2010; 4:1.

[24] Zhu Y, Zheng VW, Yang Q. Activity recognition from trajectory data. In: Zheng Y, Zhou X (eds). Computing with Spatial Trajectories. New York: Springer, 2011,179–212.

[25] Zheng Y and Zhou X. Computing with Spatial Trajectories. New York: Springer, 2011

[26] J Y, Y Z, L Z, et al. Where to find my next passenger? In: Proceedings of the 13th ACM International Conference on Ubiquitous Computing. Beijing, China 2011; 109–18.

[27] Yuan J ZY, C Z, et al. T-Drive: driving directions based on taxi trajectories. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems. San Jose, CA, USA, 2010, 99–108

[28] J Y, Y Z, X X, et al. Driving with knowledge from the physical world. In: Proceedings of the 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. San Diego, CA, USA 2011; 316–24.

[29] Liu W, Zheng Y, Chawla S, et al. Discovering spatio-temporal causal interactions in traffic data streams. In: Proceedings of the 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 2011, 1010–18.

[30] Pang XL, Chawla S, Liu W, et al. On mining anomalous patterns in road traffic streams. In: Proceedings of the 7th International Conference on Advanced Data Mining and Applications. Beijing, China, 2011, 237–51.

[31] Qian X, Zhan X, Ukkusuri SV. Characterizing urban dynamics using large scale taxicab data. In: Lagaros ND, Papadrakakis M (eds). Engineering and Applied Sciences Optimization. Cham: Springer, 2015.

[32] Zhu Y, Zheng Y, Zhang L, et al. Inferring taxi status using GPS trajectories. arXiv:1205.4378 [cs.AI], 2012.

[33] G P, G Q, Z W, et al. Land-use classification using taxi GPS traces. IEEE T Intell Transp 2013; 14:113–23.

[34] H W, Z H, J G, et al. Operational analysis on Beijing road network during the Olympic games. J Transp Syst Eng Inf Technol 2008; 8:32–7.

[35] N P, C M. When, where and how taxis are used in Montreal. Presented at the 93rd Annual Meeting of the Transportation Research Board. Washington, DC, USA, 2014

[36] Schäfer R, Thiessenhusen K, Wagner P. A traffic information system by means of real-time floating-car data. In: Proceedings of the 9th World Congress on Intelligent Transport Systems, Chicago, USA, 2002, 2.

[37] Castro PS, Zhang D, Li S. Urban traffic modeling and prediction using large scale taxi

GPS traces. In: Proceedings of the 10th International Conference on Pervasive Computing, Newcastle, UK, 2012, 57–72.

[38] Lu Y, Li S. An empirical study of with-in day OD prediction using taxi GPS data in Singapore. Presented at the 93rd Annual Meeting of the Transportation Research Board, Washington, DC, USA, 2014.

[39] H C, Y T. Hsu YJ. Context-aware taxi demand hotspots prediction. Int J Bus Intell Data Min 2009; 5:3–18.

[40] Liu L RC, Uncovering cabdrivers AC. behavior patterns from their digital traces. Comput Environ Urban 2010; 34:541–8.

[41] Ge Y, Xiong H, Tuzhilin A, et al. An energy-efficient mobile recommender system. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Washington, DC, USA, 2010, 899–908.

[42] D Z, Z P. Sun DJ. A comprehensive taxi assessment index using floating car data. J Harbin Inst Technol 2014; 21:7–16.

[43] D Z, N L, Z Z, et al. iBAT: detecting anomalous taxi trajectories from GPS traces. In: Proceedings of the 13th International Conference on Ubiquitous Computing. Beijing, China, 2011, 99–108

[44] Bertran C A i, Marin JO, Romeu M E. Presented at the 93rd Annual Meeting of the Transportation Research Board. In: Assessment of the taxi sector efficiency and profitability based on continuous monitoring and methodology to review fares. Washington, DC, USA, 2014

[45] 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.

[46] 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

[47] P. Nagaraj, Dr A. V. Krishna Prasad, “A Novel Technique to Predict the Hotspots Swine Flu Effected Regions”, THINK INDIA JOURNAL, ISSN: 0971-1260, Vol-22-Issue-41- December-2019

[48] P. Nagaraj and Dr A. V. Krishna Prasad, “A Novel Technique to Detect the Hotspots Swine Flu Effected Regions”, Published in: 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 15 November 2021 DOI:10.1109/ICRITO51393.2021.9596422, Electronic ISBN: 978-1-6654- 1703-7 CD: 978-1-6654-1702-0

[49] 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 Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Special Issue 2, December 2017)

[50] Nagaraj P,Rohit Kumar K, Rajesh Banala b,” Energy efficient 2 tier data aggregation scheme in-Sensor networks”, Accepted 7 March 2021,https://doi.org/10.1016/j.matpr.2021.03.140 2214-7853/ 2021 Elsevier Ltd.scientific committee of the Emerging Trends in Materials Science, Technology and Engineering.

[51] 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 Conference on Image Information Processing (ICIIP),26-28 Nov. 2021, 10 February 2022 DOI: 10.1109/ICIIP53038.2021.9702607 Publisher: IEEE Conference Location: Shimla, India.

[52] P.Nagaraj Rajesh Banala b, Vicky Nair “Performance 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.

[53] P. Nagaraj, ChennaSriya, SarikondaMahidhar Raju, SabaNaazneenKauser,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

[54] 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.

[55]P.Nagaraj,Dr.A.V.KrishnaPrasad,Dr.V.B.Narsimha, Dr.B.Sujatha “Swine flu Detection and Location using Machine Learning Techniques and GIS” in International Journal of Advanced Computer Science and Applications (IJACSA) , Vol. 13, No. 9, 2022,page 901 to 909, www.ijacsa.thesai.org.