

# Estimation and Prediction of Swine Flu Information using Speech Based Chatbot Model

<sup>1</sup>I. Surya Prabha, <sup>2</sup>Dr. M. Sriram

Submitted: 02/10/2023

Revised: 21/11/2023

Accepted: 30/11/2023

**Abstract:** In this research, we apply a complex chatbot model to determine the origin of the swine flu pandemic and develop an effective treatment strategy. The helpful patient's dataset collects a wealth of information. The lack of clarity in these depictions is often associated with the discovery of new diseases. The use of historical patient data to predict heart, lung, and other tumor infections is a promising new direction in medicine. Our research is based on a comprehensive strategy for gathering and using information about Swine Flu. As a result of this research, forecasting models and trouble spots have been identified. There was use of split classification that uses dynamic learning to make decisions. Data mining has a significant obstacle ahead of it in the shape of disease prediction. Here, we use a database of treatment patients to try to identify swine flu, the most rapidly spreading disease on the planet. Since swine flu is a respiratory illness, a comprehensive battery of diagnostic procedures must be performed on the patient. Advanced information mining methods could be used to fix this problem. The DLSC method has been used to the speech samples obtained from the patients via the cloud network. At long last, we're diagnosing utilizing data from speech samples gleaned from a training set. Chatbot technology has been used to discuss the swine flu and its treatment. Accuracy of 93.23%, sensitivity of 94.23%, recall of 92.12%, F1 score of 95.21%, and throughput of 90.34% are all achieved by this model, making it competitive with existing models.

**Keywords:** Naive Bayes Classifier, DLSC classifier, Data mining, Swine flu Disease

## 1. Introduction

The swine flu has a high incidence and transmission rate in the United States, placing it among the most widespread infectious diseases. To save time and money in the doctor's office, A Naive Bayes classification will be used to estimate the severity of this illness for the time being. How can we change accommodating data so that medical professionals may make informed clinical decisions? This idea serves as the inspiration for this writing. A. the global spread of bird flu Swine flu damages pigs' lungs and causes symptoms like coughing, wheezing, nasal discharges, and lethargy due to exhaustion. Humans could become infected as a result. Swine flu pandemics may evolve to become more contagious among humans. Infectious diseases spread among pigs can lead to a variety of symptoms, some of which are similar to those of human diseases. Social pollution from infected pigs this season refers to any harm that has arisen as a result of the spread of the current infectious disease season. The Influenza A(H1N1) and Influenza A(H2N2) Viruses In 2009, it became clear that influenza A subtypes included H3N1, H3N2, and H2N3. Swine flu begins its global journey in

pig herds. However, antibodies in human blood are occasionally sufficient to prevent the spread of a contagious contamination from pigs to humans. When the swine flu virus escapes from infected pigs and infects humans, this is known as zoonotic swine flu. Regular contact with pigs increases the likelihood of catching swine flu in humans. In the middle of the twentieth century, testing for infection subtypes became feasible, allowing for the effective stopping of human transmission. As quickly as possible, we have to validate the top 50 of these shows. It is not possible to contract swine flu or swine pneumonia from consuming cooked pork; the infection must be contracted through inhalation or ingestion of infected material. Even the most innovative methods of preventing swine flu infections cannot protect against H3N2v infections. Symptoms of swine flu include high body temperature, sore throat, runny nose, headache, fatigue, & nausea are all cold signs.

### *Statistics of Swine flu:*

The deliberate production and dissemination of an infectious disease was a contributing factor to the 2009 H1N1 swine flu pandemic. Infection A (H3N2) v, a previously unknown strain of the swine flu, was discovered in 2011. The specialists at the United States' Although just a limited number of people (mainly children) were initially infected with the 2012-2013 flu season, hotspots for disease management & avoidance (SMA) indicated that millions of persons were affected.

<sup>1</sup>Research Scholar, Computer Science and engineering Bharath institute of higher education and research, Selaiyur, Tambaram.

ipsurya17@gmail.com

<sup>2</sup>Associate professor, Bharath institute of higher education and research, Selaiyur, Tambaram, OS, Data Mining, Computer Networks, 9952124912.

Email: msr1sriram@gmail.com, sriram.cse@bharathuniv.ac.in

Approximately 1500 people in India have lost loved ones because of swine flu. Because of this, H3N2v is not anticipated to affect clearing amounts for relatives. A contaminant designated H3N2 (note the absence of the letter "v" in its name) has also been found. This type of influenza, however, has the potential to be quite widespread. The virus's H3N2v name has been used in the past. When all the numbers are tabulated, the bulk of the population will not have reported becoming sick from the current cold wave. The H1N1 virus is structurally identical to diseases, with the exception that each distinct disease requires a unique H.

Master's students need help with computer skills during the computational evaluation of disease transmission. The increasing accusations of pandemic contamination have been untangled by some. Certain trademark states may see an outbreak of infectious diseases such as swine flu, cholera, or jaundice. This research uses a variety of data mining methods to give the ongoing support necessary for addressing swine flu and its contributing variables. At An urgent check was performed to demonstrate what the future holds for k-means clustering & the Google graph. These clustering estimates not only aid in a Google visualization, but also exert dominance over the field of intelligent swine flu study. Google has a blueprint of the genuine idea behind the study of swine rotting. These bundles depict, in a nutshell, the many spots where detective work has been done. The new guidelines for data collecting are based on preliminary calculations derived from the existing data archives.

Much of the information details the devastation brought on by swine infection in a particular area, together with the natural domain variables that preceded and followed it. It's important to keep in mind not just the potential for swine contamination, but also the normal measurements for heat, stickiness, and elasticity. The usage of archived records of past data is essential for imparting lessons on the significance of such recognitions. The C4.5 scale and the believable bayes classification are highly correlated. The reasonable result indicates a real blue weight inconsistency of investigative findings with respect to the basically observed investigations. It keeps an eye on potential new domains affected by the zone to make sure they don't start claiming an excessive amount of investigations. Furthermore, this conclusion informs those who may be considering tainting claims in a new region of the challenging path that lies ahead of them. Standard basic characteristics required to image the swine flu sea can be attained through web-based collaborations amongst national atmosphere figures. As part of a comprehensive review, decision trees are factored in. Nave Bayes classifiers need agitated subject Like establishing a profit as well as accurate count for

investigations from those crucial systems, it is necessary for.

## 2. Literature survey

**Thakkar, B. A., et al [2010]** Data mountains are being produced by the healthcare sector, but they are not being mined or used effectively. In many cases, we fail to recognize latent connections & patterns. However, patient information is still being used in diagnostic medical studies to predict outcomes for diseases like heart disease and lung cancer. Our research utilizes this aspect of medical diagnosis with respect to Swine Flu data. An intelligent swine flu prediction system (ISWPS) prototype was developed as a result of this investigation. Patients with swine flu were classified using a Naive Bayes classifier into the least likely, most likely, and most likely categories, respectively. We used 17 of the swine flu symptoms recognized by medical professionals and assembled 110 symptom sets from various hospitals and clinics. We were able to improve our accuracy to 63.33 percent by employing ISWPS. Its implementation makes advantage of the JAVA platform [1].

**Baker, Q. B., et al [2020]** Opinion mining shows promise as a useful tool for gaining insight from health data. Twitter may be used to track the spread of diseases, according to numerous research. Many studies have excluded tweets written in languages other than English. Influenza and other infectious diseases are a worldwide epidemic. This study presents a novel approach to monitoring Arabic tweets from Arab countries for signs of influenza transmission using machine learning techniques. We think this is the first study to analyze disease transmission using Arabic Twitter [2].

**Alsmadi, T., et al [2021]** Covid-19 is a newly discovered coronavirus that has been linked to an infection. Wuhan, Hubei Province was the first location where this pneumonia-like illness was identified. Since there is currently no vaccine and the pandemic is rapidly expanding, medical professionals want assistance in diagnosing people. Finding a solution to this problem requires data mining techniques. Predicting recoveries is crucial for countries working to contain the virus because it helps public health specialists follow COVID-19-positive citizens, increases clinicians' capacity to forecast the general perception of occurrences over time, and lets them to assess early-risk patients using fresh data. The 3,799-person COVID-19 Corona Virus India dataset was used to split cases into survivors and non-survivors. In this article, this dataset illustrates supervised learning. On the patient dataset, decision trees (DT), support vector machines (SVM), logistic regression (LR), random forests (RF), k-nearest neighbours (KNN), naïve bayes (NB), and artificial neural networks (ANN) were used [3, 4]

**Nouira, K., et al [2020]** Since a growing percentage of the population has access to the internet, we can now quantify a variety of modern social phenomena. Monitoring the spread of the current outbreak presents a significant problem for health care practitioners. Our unique digital surveillance technique, FluSpider, measures the outbreak by monitoring internet activity. Large-scale data mining and Big Data technologies make our process most effective and exact. Our findings suggest FluSpider could follow influenza-like infections in real time two weeks before traditional centres can respond. [4].

**Haque, T. H., et al [2018]** The impacts of the swine flu on tourism in Brunei have been disentangled from those of the ongoing global financial crisis through the use of auto regressive integrated moving average and intervention time series analysis, as demonstrated in this paper. Analysis of the effects of the swine flu in the first year after the pandemic used an auto regressive integrated moving average model for the virus, models for the global financial crisis and its impact on the number of tourist arrivals, and models for intervention time series analysis. Examining the time series intervention model coefficients, we may estimate the extent to which the swine flu and global financial crisis have impacted the tourism sector. The swine flu pandemic cost Brunei an estimated B\$15 million and 30,000 visitors in the first year after the epidemic. [5].

**Signorini, A., et al [2011]** You can send each other short messages (called "tweets") on Twitter, an open-source platform for microblogging and social networking. There is a 140-character limit per tweet. Daily, the service handles more than 55 million tweets from its user base of over 190 million. Everyone's thoughts are represented in the Twitter stream to tweets, is rich with data on current affairs and international politics. The Twitter stream is rich in data that might be useful for tracking and even predicting user behavior if collected effectively. In this study, we examine how Twitter data can be used to (1) track how the public feels about H1N1 (swine flu) over time, secondly, monitor the real spread of the disease and report on it. We also show that Twitter is a valid gauge of public concern during health emergencies. Tweet volume can be used to estimate the spread of influenza-like infections, and the results are consistent with those reported by official agencies. [6].

**Peng, Y., et al [2020]** A correlation between search engine data, such as that produced by Google Trends, and the spread of COVID-19 was observed but only in a subset of nations. To anticipate the global pandemic warning level, they plan to use a model based on data from a subset of countries. Google Trends was used to look at how people were searching for things like

Searches for "coronavirus," "pneumonia," and "six COVID symptoms" ranged widely across countries, but each had distinctive characteristics. Daily incidence figures for COVID-19 in 202 countries were requested from the World Health Organisation between January 10, 2020, and April 23, 2020. There are currently three distinct "alarm states." Data from 20 nations were used to train machine learning algorithms for 10 weeks. Selected characteristics for their significance and connections to one another. After that [7] was used to assess the model, data was collected from 2,830 individuals across 202 countries.

**Biswas, S. K., et al [2014]** Applying the problem-solving method of Case Based Reasoning (CBR) to the medical sciences could have significant positive effects. One of the most important tools for achieving CBR, however, is the reliable evaluation of case similarity. Inductive learning techniques, such as weighted nearest neighbour (w-NN), approaches that employ feature weighting to account for the significance of similarity when making predictions regarding the spread of swine flu. A unique strategy of classifying instances has been developed using w-NN and extensibility in similarity, which can be utilised to foresee the spread of swine flu. The healthcare business is sitting on a data goldmine, but it is not being mined properly or exploited to its full potential, which is where the novel case classification method comes in. The empirical findings validate the tremendous potential of our method by yielding an answer that is 85.6% accurate, which is more precise than any prior model, and the two models are compared to one another. [8].

**Boulos, M. N. K., et al [2010]** Incorporating methods from the Social Web (or "Web 2.0"), such as weblogs, This research delves into the realm of Web "data mining," specifically the application of Techno social Predictive Analytics (TPA) to social media platforms, wherein users' postings and inquiries are aggregated to form unified models of real-time health occurrences. This article gives readers a crash course in the history of mashups and aggregators as well as a rundown of their meteoric rise as a people-centric, open-architecture component of the Social Web. In-depth discussions and demonstrations of health-related tools are provided, such as the use of Google Maps to the display of epidemiological data, such as the location of recent flu outbreaks. [9].

**Ong, S. Q., et al [2022]** Mosquito-borne diseases have originated and re-emerged in various places of the world since The COVID19 pandemic has hit. Text mining for infectious diseases has recently shown promising results, opening the door to instantaneous access to implicit and explicit linkages between textual material. Due to the

wealth of unstructured or semi-structured text data available online with rich content in this domain, many studies have focused on knowledge discovery about diseases, disease surveillance, early detection systems, etc. in recent years. Unfortunately, we were unable to locate any in-depth research on the topic of Diseases spread by mosquitoes and text mining. This article compiles the most recent findings on the effectiveness of text mining strategies in preventing mosquito-borne diseases. In order to better understand the recent developments and potential future directions in this domain, we assess corpus sources, technology, applications, and obstacles. In this research, we present the results of a bibliometric analysis of 294 scholarly articles from Scopus and PubMed covering the subject of text mining in mosquito-borne diseases and published between 2016 and 2021. After weeding out non-applicable studies, the remaining papers were assessed for the methodology they used to collect data on mosquito-borne illnesses. Only 27 of the 158 included studies directly addressed the issue of edition withdrawal in relation to mosquito-borne diseases. While Zika, Dengue, and Malaria received the most media coverage, other infectious diseases such as Chikungunya, Yellow Fever, and West Nile were also discussed. Usually, the fever was quite low. [10].

**aNagaraj, P., et al [2019]** In terms of data, the charity sector collects an enormous amount that is never used. Oftentimes, it serves no purpose to reveal such obscured symbols and associations. Research is ongoing, however, with the hope of developing a therapy that can analyze patient records and predict infections in the heart, lungs, and other tumors. Our study is based on a knowledge map of the current swine flu data, which we used to conclude that no further medical progress can be made. This study has uncovered uncharted territory for the development of predictive models for intelligent swine infections. We built a split classifier using Dynamic Learning and put it into action as the DLSC Classifier. Health care disease prediction relies heavily on data mining. Given the lack of improvement in the patient database, we have initiated a program to detect Swine flu is the most contagious illness in the world. As a respiratory virus, swine flu mostly affects the lungs and nose, a wide range of tests is required to confirm a diagnosis. Powerful data mining algorithms are assisting in the resolution of this issue. [11].

**Ali, A. A., et al [2020]** Given that heart disease is a due to the fact that it is a top killer and the heart is an extremely important organ, it stands to reason that doctors would aim to provide the most plausible diagnosis possible. Many of the unique hybrid approaches established by researchers have the potential to improve a variety of machine learning techniques,

which may be of benefit to those who make a living anticipating the onset of heart problems. The Gate Recurrent Unit of a Convolutional Network (CNN GRU) is a method introduced in this study. One of the primary motivations behind this strategy is to provide a more robust machine learning strategy for cardiac illness forecasting. The most important characteristics of a database can be extracted with the use of Techniques for selecting features, such as principal component analysis (PCA) and linear discriminant analysis (LDA). The proposed strategy was evaluated alongside a number of alternatives that rely on machine learning in a similar way. Using "K-fold" cross-validation, we were able to achieve better results. The (CNN GRU) technique was the most accurate, with a 94.5 percent success rate. [12].

**Olson, D. R., et al [2013]** Clinical episodes of influenza-like illness (ILI) can be tracked to determine the timing, geography, and intensity of outbreaks. Due to advancements in processing power and data storage, enormous volumes of electronic data may now be automatically collected, allowing for instantaneous evaluations. One such cutting-edge tracking system is Google Flu Trends (GFT), which analyses search query data to track the spread of influenza-like illness (ILI). The initial wave of the 2009 influenza A/H1N1 pandemic was not predicted at the regional, national, or worldwide levels by either the original or improved GFT models, and the GFT models grossly underestimated the severity of the A/H3N2 epidemic during the 2012/2013 season. The results from both the 2008 and 2009 versions of GFT were the same. Both models underperformed the gold standard, possibly because of shifts in internet search behaviour [13] and differences in the seasonality, regional heterogeneity, and age distribution of the epidemics between the timeframes of GFT model-fitting and prospective usage.

**Culotta, A. [2010]** Over the course of eight months, we analyzed over 500 million tweets and discovered that by keeping an eye on a handful of flu-related phrases, we could accurately predict flu incidence in the future, with a 95% association to data on the health of the nation as a whole. Then, we assess how resistant our method is to keyword false positives and propose including a document categorization phase to remove spurious alerts. We find that our document classifier can reduce mistake rates by more than half in simulated false alarm testing, but additional work is needed to develop methods that are robust in the face of very high noise. [14].

**Lamos, V., et al [2010]** Monitoring the transmission of an infectious disease like the flu can help prevent it from becoming widespread and give public health officials time to prepare a response. Particularly, this surveillance aids in the detection and location of an outbreak. Various

methods are used to maintain tabs on patients, such as recording how often they go to the doctor. We provide a technique that uses content analysis on Twitter and other social media to calculate disease rates in a given area. Hundreds of tweets are analysed daily for symptom references, and this data is then converted into a numerical index that could be used to monitor the spread of influenza [15].

**Achrekar, H., et al [2011]** Reducing the severity of pandemics like H1N1 and the annual flu is a major focus of public health efforts. Research has shown that early detection is key in stopping epidemics in their tracks. The CDC routinely collects data on the distribution of ILIs from a representative sample of "sentinel" healthcare facilities. Illness outbreak reports will not include individual patient diagnoses for at least two weeks. This study presents the Social Network Enabled Flu Trends (SNEFT) architecture, which monitors tweets for mentions of flu indicators to trace the emergence and spread of a pandemic. Based on our analysis of data from 2009 and 2010, we find that the volume of tweets regarding the virus is positively correlated with the number of CDC-reported cases of influenza or influenza-like illness (ILI). To further estimate the ILI gearbox, we additionally construct auto-regression models. The percentage of upcoming "sentinel" medical visits that can be linked to ILI is predicted using CDC data and algorithms. Thus, Twitter data allows for real-time tracking of disease activity [16].

**Venna, S. R., et al [2018]** Here, we present a cutting-edge LSTM-based data-driven machine learning approach to forecasting influenza seasons. Some distinguishing features of the approach are: Two significant advances have been made recently: The suggested model is put to the test against two state-of-the-art methods using two publicly available datasets. Our strategy outperforms both the current standard and the most widely used approach to influenza forecasting. The findings indicate that data-driven forecasting methods that take into account spatial, temporal, and environmental aspects can lead to more accurate predictions of influenza outbreaks. [17].

**AlAmoodi, A. H., et al [2020]** Medical photos from There is a comparison and analysis of various AI algorithms for identifying and classifying the 2019 coronavirus outbreak (COVID-19). To locate relevant research, we looked through Databases such as Scopus, IEEE Xplore, PubMed, Science Direct, and the Web of

Science. For the first 36 research projects, only 11 were determined to meet the inclusion/exclusion criteria after several rounds of screening [18]. The eleven papers were classified as either reviews or original research using taxonomies. Then, the relevant academic literature was thoroughly researched and analyzed to identify the gaps and challenges. There were no articles located that compared the efficacy of different AI algorithms for With the COVID-19, you may do a wide variety of binary, multi-class, multi-labelled, and hierarchical classifications are all examples of classification problems. picture collection as a benchmark. Future difficulties for evaluation and benchmarking may arise from the usage of many assessment criteria within each tagging assignment, the need to make trade-offs, and the criteria's relevancy [19].

### 3. Existing Methods

In this tabular display, We highlight the most critical swine flu investigations and provide a link to a Google graphic [20]. Checks for swine flu at different meetings are verified with the use of the figure 1 tally. The data and the various results are described using set theory [21].

$S = \{I, O, F\}$

Data sources, restrictions, and the results they produce are presented in a set-theoretic fashion below as Set S [22].

Inputs are 1)  $I = D, P, S, y$ .

Set of data =  $d_i$ , where  $d_i$  is a scheme for all medical history

A method for calculating Swine Flu complications using  $P = \text{Weather Parameter} = S_i$

To address the  $s_y = s_y$  conundrum: Strategy Based on the  $P_i$  of Climate Variables

2)  $F = \{K, D, B\}$

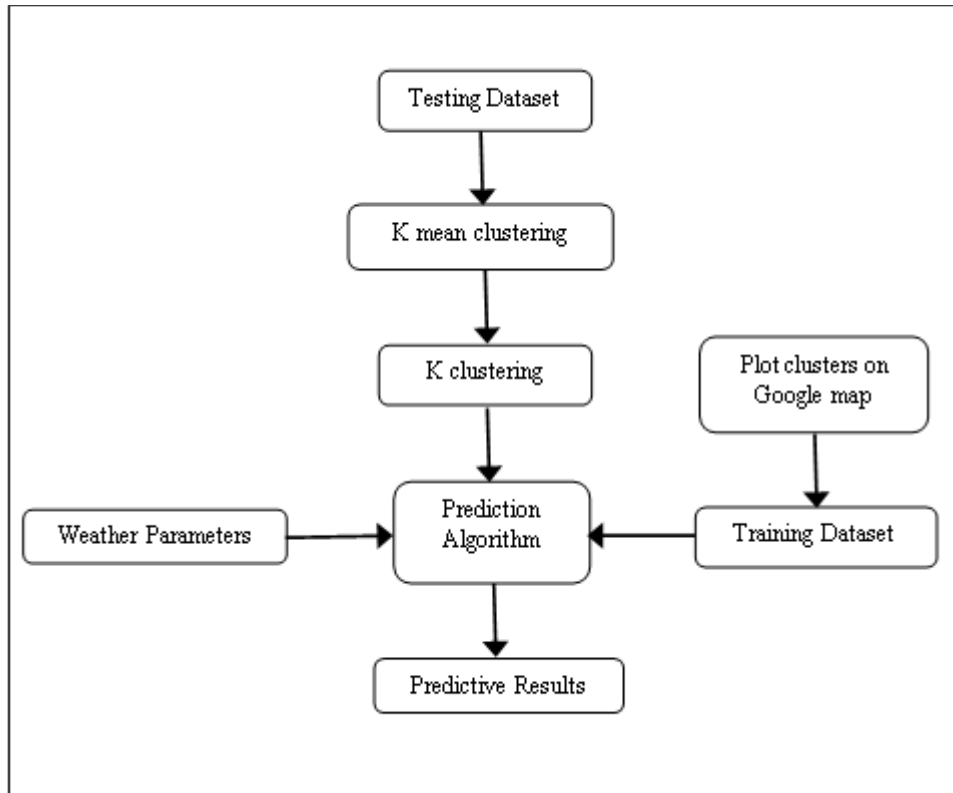
$K = \text{Mean}$

$D = \text{Tree of Choices}$

$B = \text{Bayesian Theorem}$

#### *Existing Architecture:*

When patient records have been examined in detail, they will provide the primary results of the swine flu investigation [23]. Indicators may be used to evaluate the patient's development. The exact clusters of crucial studies are outlined and displayed on Google Maps [24] [25].



**Fig.1.** System Architecture

The final tally now incorporates the findings of the probes. The following is an illuminating subset of the structure's functioning: [26][27].

**Testing Data set:**

The testing dataset [28] would be an introduction data record for people who need assistance managing the side effects of swine flu infection. Any data sets that k determines are crucial must be collected [29]. K anticipates discovering a K anticipates a high degree of association between the data records gathered during swine flu investigations and the social event estimates he plans to produce [30].

**Planning Data set:**

**Proposed Algorithm**

**Algorithm: 1** feature selection procedure

[In this definition, Pattern problem specifications have a total number of components, N, and a corresponding attribute, A.]

**Step: 1** KThe classification is performed using K-Nearest Neighbor (K-NN) [35].

1) The standards for determining the classifier's performance are fixed. A fair measure of its efficacy is the number of false positives it produces relative to the total number of classifications. Highest recognition rate characteristic is used as criterion [36].

2) Many iterations of training to train a classification algorithm on a dataset a single feature at a

The "K" Clusters feature on Google Maps [31] allows users to hone in on a specific social problem. we may visualize potential hotspots for Swine Flu surveillance. In order to track whether or not investigations are being thoroughly checked off based on the most up-to-date evidence, it is necessary to compile a dataset of accounts from the runner's previous [32].

**Perceptive Investigation:**

Based on the current patient outline, we employ dataset with real-world inspection data using a decision tree technique and Naive Bayes classifier [33]. Parts of the present setup are analysed to determine the probable impact of Swine Flu on the area [34].

time are performed using the five-folded cross-validation method [37]. As such, the classifier will utilise attribute A1 initially, then A2 subsequently, and so on up to attribute AN. We keep track of the recognition rate and the amount of false positives for each individual characteristic. Attribute Ai, with possible values between 1 and N, is chosen according to the standards set forth in Step 2. This maximises the classifier's efficacy [38].

3) Cross-validation using two features simultaneously uses the classifier:

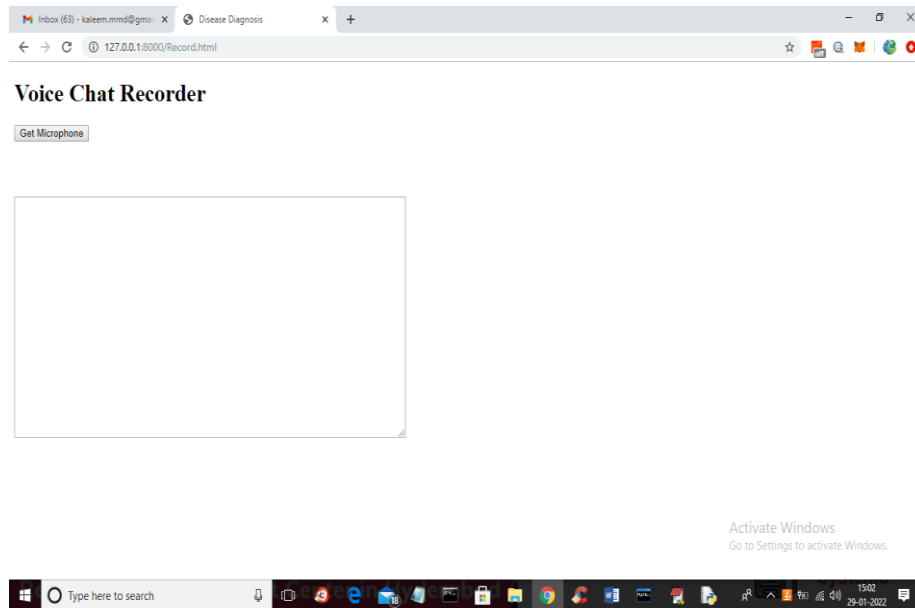
$$A_i A_1, A_i A_2, \dots, A_{i-1} A_{i+1}, \dots, A_i A_n$$

The classifier selects the optimal pairing. Let's pretend the best possible pair is  $A_i A_j$  for  $1 \leq i \leq N$  and  $1 \leq j \leq N$ .

4) Thirdly, three attributes are utilized simultaneously with the classifier in cross-validation:  $A_1A_2A_3$ ,  $A_1A_2A_4$ , and so on [39].

5) Consequently, step by step, the list of preferred characteristics grows by one feature. This process is done until an optimal classifier is found. [40].

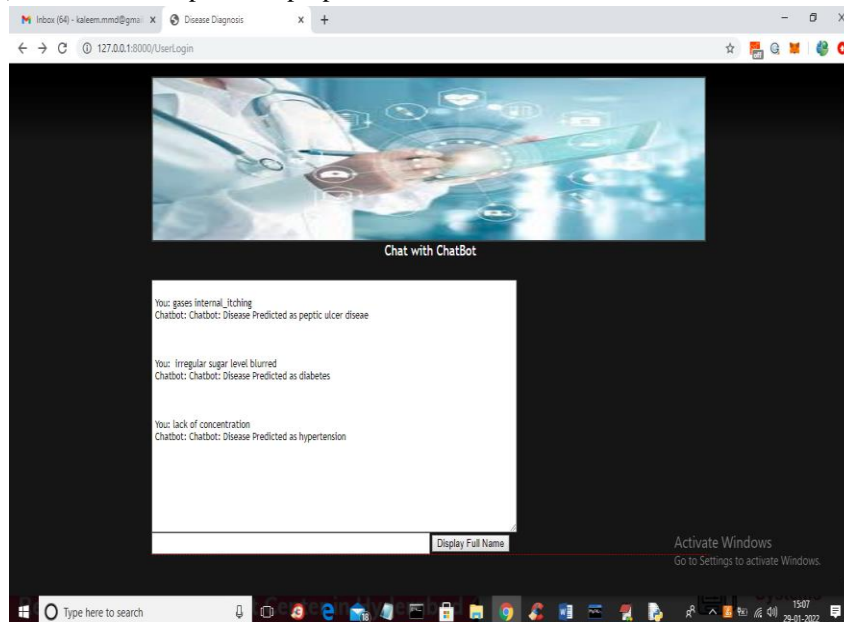
In my study, I used the DLSC Classifier (Dynamic Learning and Managed Classifier) to identify people afflicted by seasonal infections in problem areas by expanding this expectation models. Using a dynamic learning model, I was able to predict who among my contacts was most at risk of contracting the flu this year.



**Fig :2** recording voice through recorder

As far as we're aware, this approach is the only one that takes into account how to isolate situations with potential clinical significance. Straight hotspot disclosure is now scalable thanks to new algorithmic adjustments (i.e., fragmenting information that differentiates across persons and places). To back up our proposed

algorithmic improvements, we also Specify a Cost Model for the Calculations You Suggest. Particular activities that need investigation are: In accordance with the suggested strategy, we employ a dynamic division to address the sub-edge hotspots in the dataset.



**Fig :3** chat bot diagnosis for swine flu

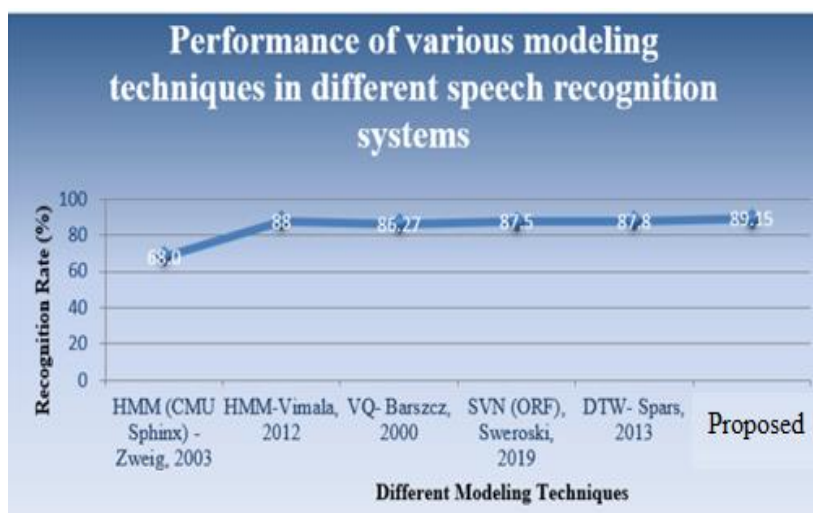
To extend the dynamic sub-division allows for direct hotspot detection's adaptability, new algorithmic additions are introduced, such as an area channel and pruning computation. We provide an expense analysis

and empirical evidence supporting the master's calculations. Two contextual investigations, depicted in Figures 2 and 3, will be provided that compare the results

of finding when compared to the results of comparable studies, such as area recognition, under strong division.

**Table :1** patient voice for swine flu diagnosis

Patient voice	Gender	Accuracy
Patient 1	F	90.23%
Patient 2	F	88.61%
Patient 3	F	87.01%
Patient 4	M	89.12%
Patient 5	F	83.01%
Patient 6	F	88.61%
Patient 7	M	87.01%
Patient 8	F	89.12%
Patient 9	M	83.01%
Patient 10	F	88.61%
Patient 11	F	87.01%
Overall accuracy of system	-	89.154%

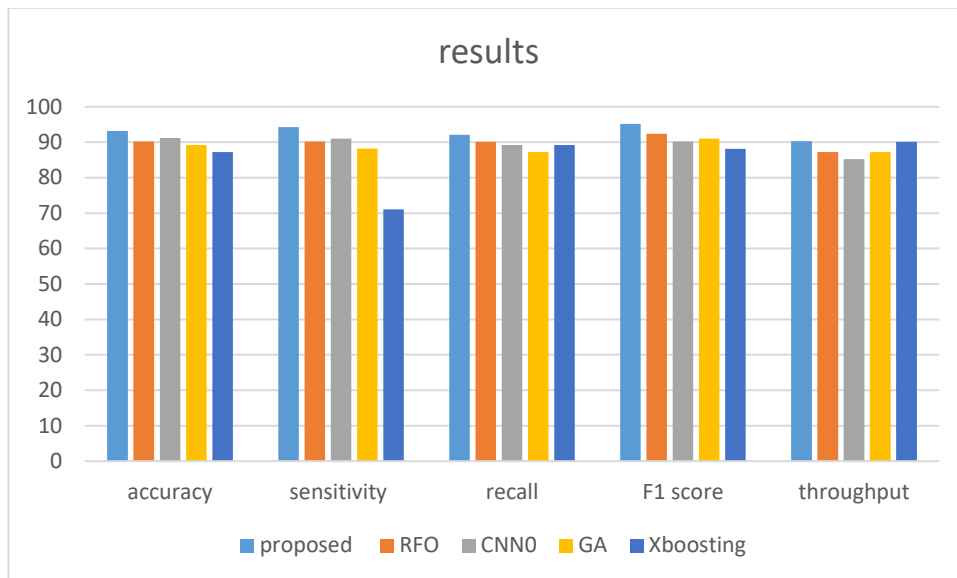


**Fig :4** comparison of results

**Table: 2** comparative analyses

	accuracy	sensitivity	recall	F1 score	throughput
proposed	93.23	94.23	92.12	95.21	90.34
RFO	90.21	90.21	90.12	92.34	87.21
CNN0	91.23	91.02	89.23	90.23	85.23
GA	89.23	88.21	87.27	91.01	87.23
X-boosting	87.21	71.02	89.23	88.12	90.12





**Fig :5** comparatives analysis

Figure 5 and Table 2 show clearly how the findings of comparisons were arrived at, demonstrating that the suggested model improves upon the state of the art.

#### 4. Conclusion

This body of work demonstrates how a data mining approach can be utilized successfully in the study of disease transmission. Patients are likely to be designed using a batching count of  $k$ , as shown by the Swine Flu symbol on Google's layout. It is feasible to put people's fantasies and potential threats to the test with the help of the desire estimates. For example, C4.5 has strict rules for collecting and acquiring data, yet it settles on a difficult decision tree structure after being subjected to the various competing atmospheric elements. When there is a change in an The Bayesian Theorem, an innovative probabilistic calculation, excels at predicting atmospheric parameters all other assessment tree estimations. Through a cloud network, voice samples from the patients were obtained and analysed using the DLSC method. Ultimately, swine flu diagnosis and therapy have been implemented using voice samples and data based on training set. This model competes with existing models and achieves accuracy of 93.23%, sensitivity of 94.23%, recall of 92.12%, F1 score of 95.21%, and throughput.

#### References

- [1] Thakkar, B. A., Hasan, M. I., & Desai, M. A. (2010, October). Health care decision support system for swine flu prediction using naïve bayes classifier. In 2010 International Conference on Advances in Recent Technologies in Communication and Computing (pp. 101-105). IEEE.
- [2] Baker, Q. B., Shatnawi, F., Rawashdeh, S., Al-Smadi, M., & Jararweh, Y. (2020). Detecting epidemic diseases using sentiment analysis of arabic tweets. *J. Univers. Comput. Sci.*, 26(1), 50-70.
- [3] Alsmadi, T., Alqudah, N., & Najadat, H. (2021, July). Prediction of Covid-19 patients states using Data mining techniques. In 2021 International Conference on Information Technology (ICIT) (pp. 251-256). IEEE.
- [4] Noura, K., & Njima, N. B. (2020, February). FluSpider as a new vision of digital influenza surveillance system: based on Big Data technologies and Massive Data Mining techniques. In 2020 International Multi-Conference on: "Organization of Knowledge and Advanced Technologies" (OCTA) (pp. 1-10). IEEE.
- [5] Haque, T. H., & Haque, M. O. (2018). The swine flu and its impacts on tourism in Brunei. *Journal of Hospitality and Tourism Management*, 36, 92-101.
- [6] Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PLoS one*, 6(5), e19467.
- [7] Peng, Y., Li, C., Rong, Y., Chen, X., & Chen, H. (2020). Retrospective analysis of the accuracy of predicting the alert level of COVID-19 in 202 countries using Google Trends and machine learning. *Journal of global health*, 10(2).
- [8] Biswas, S. K., Sinha, N., Baruah, B., & Purkayastha, B. (2014). Intelligent decision support system of swine flu prediction using novel case classification algorithm. *International Journal of Knowledge Engineering and Data Mining*, 3(1), 1-19.

- [9] Boulos, M. N. K., Sanfilippo, A. P., Corley, C. D., & Wheeler, S. (2010). Social Web mining and exploitation for serious applications: Technosocial Predictive Analytics and related technologies for public health, environmental and national security surveillance. *Computer methods and programs in biomedicine*, 100(1), 16-23.
- [10] Ong, S. Q., Pauzi, M. B. M., & Gan, K. H. (2022). Text Mining in Mosquito-Borne Disease: A Systematic Review. *Acta Tropica*, 106447.
- [11] Nagaraj, P., & Prasad, A. K. (2019). Survey on Swine flu Prediction. *Int. J. Manag. Technol. Eng.*, 9, 937-941.
- [12] Ali, A. A., Hassan, H. S., & Anwar, E. M. (2020, July). Heart diseases diagnosis based on a novel convolution neural network and gate recurrent unit technique. In *2020 12th International Conference on Electrical Engineering (ICEENG)* (pp. 145-150). IEEE.
- [13] Olson, D. R., Konty, K. J., Paladini, M., Viboud, C., & Simonsen, L. (2013). Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS computational biology*, 9(10), e1003256.
- [14] Culotta, A. (2010). Detecting influenza outbreaks by analyzing Twitter messages. arXiv preprint arXiv:1007.4748.
- [15] Lampos, V., & Cristianini, N. (2010, June). Tracking the flu pandemic by monitoring the social web. In *2010 2nd international workshop on cognitive information processing* (pp. 411-416). IEEE.
- [16] Achrekar, H., Gandhe, A., Lazarus, R., Yu, S. H., & Liu, B. (2011, April). Predicting flu trends using twitter data. In *2011 IEEE conference on computer communications workshops (INFOCOM WKSHPs)* (pp. 702-707). IEEE.
- [17] Venna, S. R., Tavanaei, A., Gottumukkala, R. N., Raghavan, V. V., Maida, A. S., & Nichols, S. (2018). A novel data-driven model for real-time influenza forecasting. *Ieee Access*, 7, 7691-7701.
- [18] Albahri, A. S., Hamid, R. A., Al-qays, Z. T., Zaidan, A. A., Zaidan, B. B., Albahri, A. O., ... & Madhloom, H. T. (2020). Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review. *Journal of medical systems*, 44(7), 1-11.
- [19] Vemuri, D. K., Akshinthala, P., Konduru, N., Kowtharapu, L. P., Katari, N. K., Jonnalagadda, S. B., & Gundla, R. (2022). Unique quality by design approach for developing HPLC and LC-MS method for estimation of process and degradation impurities in pibrentasvir, antiviral agent for hepatitis C. *ACS Omega*, 7(51), 47650-47661. doi:10.1021/acsomega.2c04617
- [20] Prudhvidhar Reddy, A., Saride, R. B., Peyyala, P. K., & Mallepally, V. R. (2022, December 23). Experimental investigation of mechanical characterization of sugarcane bagasse & Sapodilla seed shell powder reinforced polymer composite. *SAE Technical Paper Series*. Presented at the International Conference on Advances in Design, Materials, Manufacturing and Surface Engineering for Mobility. doi:10.4271/2022-28-0587
- [21] Chinnasamy, P., Sathya, K. B. S., Jebamani, B. J. A., Nithyasri, A., & Fowjiya, S. (2022). Deep learning. In *Deep Learning Research Applications for Natural Language Processing* (pp. 1-17). doi:10.4018/978-1-6684-6001-6.ch001
- [22] Comparative study of flow stress modeling of zircaloy-4 sheets manufactured from different routes. (2022). *Indian Journal of Engineering & Materials Sciences*, 29(06). doi:10.56042/ijems.v29i6.70299
- [23] Patil, V., Madgi, M., & Kiran, A. (2022). Early prediction of Alzheimer's disease using convolutional neural network: a review. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery*, 58(1). doi:10.1186/s41983-022-00571-w
- [24] Ramanjini, V., Gopi Krishna, G., Mishra, S. R., Kumari, S. V. S., & Sree, H. K. (2022). An unsteady axisymmetric Williamson nanofluid flow over a radially stretching Riga plate for the inclusion of mixed convection and thermal radiation. *Partial Differential Equations in Applied Mathematics: A Spin-off of Applied Mathematics Letters*, 6(100456), 100456. doi:10.1016/j.padiff.2022.100456
- [25] Deepanraj, B., Senthilkumar, N., Jarin, T., Gurel, A. E., Sundar, L. S., & Anand, A. V. (2022). Intelligent wild geese algorithm with deep learning driven short term load forecasting for sustainable energy management in microgrids. *Sustainable Computing Informatics and Systems*, 36(100813), 100813. doi:10.1016/j.suscom.2022.100813
- [26] Kavitha, R. J., Thiagarajan, C., Priya, P. I., Anand, A. V., Al-Ammar, E. A., Santhamoorthy, M., & Chandramohan, P. (2022). Improved Harris Hawks Optimization with Hybrid Deep Learning Based Heating and Cooling Load Prediction on residential buildings. *Chemosphere*, 309(Pt 1), 136525. doi:10.1016/j.chemosphere.2022.136525
- [27] Palle, K., Vunguturi, S., Subba Rao, K., Naga Gayatri, S., Ramesh Babu, P., Mustaq Ali, M., &

- Kola, R. (2022b). Correction to: Comparative study of adsorption isotherms on activated carbons synthesized from rice husk towards carbon dioxide adsorption. *Chemical Papers*, 76(12), 7865–7866. doi:10.1007/s11696-022-02461-0
- [28] Rao, G. P., Singh, R., Lenka, T. R., Boukourt, N. E. I., & Nguyen, H. P. T. (2022). Simulation modelling of III-Nitride / $\beta$ -Ga<sub>2</sub>O<sub>3</sub> Nano-HEMT for microwave and millimetre wave applications. *International Journal of RF and Microwave Computer-Aided Engineering*, 32(12). doi:10.1002/mmce.23416
- [29] Vasanthkumar, P., Senthilkumar, N., Rao, K. S., Metwally, A. S. M., Fattah, I. M., Shaafi, T., & Murugan, V. S. (2022). Improving energy consumption prediction for residential buildings using Modified Wild Horse Optimization with Deep Learning model. *Chemosphere*, 308(Pt 1), 136277. doi:10.1016/j.chemosphere.2022.136277
- [30] Neelakandan, S., Prakash, M., Geetha, B. T., Nanda, A. K., Metwally, A. M., Santhamoorthy, M., & Gupta, M. S. (2022). Metaheuristics with Deep Transfer Learning Enabled Detection and classification model for industrial waste management. *Chemosphere*, 308(Pt 1), 136046. doi:10.1016/j.chemosphere.2022.136046
- [31] Prusty, S., Somu, P., Sahoo, J. K., Panda, D., Sahoo, S. K., Sahoo, S. K., ... Rao, K. S. (2022). Adsorptive sequestration of noxious uranium (VI) from water resources: A comprehensive review. *Chemosphere*, 308(Pt 1), 136278. doi:10.1016/j.chemosphere.2022.136278
- [32] Palle, K., Vunguturi, S., Subba Rao, K., Naga Gayatri, S., Ramesh Babu, P., Mustaq Ali, M., & Kola, R. (2022a). Comparative study of adsorption isotherms on activated carbons synthesized from rice husk towards carbon dioxide adsorption. *Chemical Papers*, 76(12), 7525–7534. doi:10.1007/s11696-022-02371-1
- [33] Ravindranath Reddy, G., & Ramakrishna Reddy, K. (2022). Casson nanofluid performance on MHD convective flow past a semi-infinite inclined permeable moving plate in presence of heat and mass transfer. *Heat Transfer*, 51(8), 7307–7327. doi:10.1002/htj.22645
- [34] Anand, A. V., Sivaraj, G., Velmurugan, P., Rao, K. S., Khan, M. A., & Kumar, P. S. P. (2022). Interpretive structural modelling approach: Implementation of sustainability concept and lean in aerospace sectors. *THE 8TH ANNUAL INTERNATIONAL SEMINAR ON TRENDS IN SCIENCE AND SCIENCE EDUCATION (AISTSSE) 2021*. Presented at the THE 8TH ANNUAL INTERNATIONAL SEMINAR ON TRENDS IN SCIENCE AND SCIENCE EDUCATION (AISTSSE) 2021, Medan, Indonesia. doi:10.1063/5.0109432
- [35] Bala, M. N. V. S. S., Goud, N. U. R., Gara, D. K., & Kannan, R. (2022). Determination of stacking fault energy for a Nitinol based shape memory alloys by molecular dynamics simulation. *THE 8TH ANNUAL INTERNATIONAL SEMINAR ON TRENDS IN SCIENCE AND SCIENCE EDUCATION (AISTSSE) 2021*. Presented at the THE 8TH ANNUAL INTERNATIONAL SEMINAR ON TRENDS IN SCIENCE AND SCIENCE EDUCATION (AISTSSE) 2021, Medan, Indonesia. doi:10.1063/5.0109530
- [36] Mishra, N. K., Naveed, M. B., Shivam, D., Chandrashekhar, S., Hrithik, P., & Kumar, P. S. P. (2022). Numerical investigation of supersonic flow past circular cross-section bodies. *THE 8TH ANNUAL INTERNATIONAL SEMINAR ON TRENDS IN SCIENCE AND SCIENCE EDUCATION (AISTSSE) 2021*. Presented at the THE 8TH ANNUAL INTERNATIONAL SEMINAR ON TRENDS IN SCIENCE AND SCIENCE EDUCATION (AISTSSE) 2021, Medan, Indonesia. doi:10.1063/5.0108443
- [37] Anand, A. V., Velmurugan, P., Singh, R. A., Jayalakshmi, S., Rao, K. S., & Kumar, P. S. P. (2022). Macro mechanical and water absorption properties of composite laminates with novolac resin and e-glass/sisal fibers. *THE 8TH ANNUAL INTERNATIONAL SEMINAR ON TRENDS IN SCIENCE AND SCIENCE EDUCATION (AISTSSE) 2021*. Presented at the THE 8TH ANNUAL INTERNATIONAL SEMINAR ON TRENDS IN SCIENCE AND SCIENCE EDUCATION (AISTSSE) 2021, Medan, Indonesia. doi:10.1063/5.0109433
- [38] Kailasam, S., Achanta, S.D.M., Rama Koteswara Rao, P., Vatambeti, R., Kayam, S. (2022). An IoT-based agriculture maintenance using pervasive computing with machine learning technique. *International Journal of Intelligent Computing and Cybernetics*, 15(2), pp. 184–197.
- [39] Saikumar, K., Rajesh, V. A machine intelligence technique for predicting cardiovascular disease (CVD) using Radiology Dataset. *Int J Syst Assur Eng Manag* (2022). <https://doi.org/10.1007/s13198-022-01681-7>.
- [40] Saikumar, K., Rajesh, V., & Rahman, M. Z. U. (2022). Pretrained DcAlexnet Cardiac Diseases Classification on Cognitive Multi-Lead Ultrasound Dataset. *International Journal of Integrated Engineering*, 14(7), 146-161.