Abstract

By the emergence and rapid spread of 2009 pandemic influenza A (H1N1) virus through the world, several methods have been developed to predict and prevent this lethal disease. Although many efforts have been made by statistical and traditional intelligent methods to anticipate this disease, but none of them could satisfy the expectations of specialists. This paper aims to present an efficient hybrid method to predict H1N1 with a reliable sensitivity. In this way, three methods including Gaussian mixture model (GMM), neural network (NN), and fuzzy rule-based system (FRBS) have been fused in order to provide an accurate and reliable prediction scheme to anticipate presence of H1N1 influenza. In this study, 220 individuals were participated and their clinical data were collected. The proposed hybrid scheme was implicated and the results showed to be superior to using each of the decision components containing NN, FRBS, and GMM classifiers. The achieved results produced 95.83% sensitivity and 80.95% specificity on unseen data which support the effectiveness of the hybrid method to predict the influenza in its golden time.

Keywords: GMM; Neural network; FRBC; H1N1 influenza; Prediction

Introduction

Few months after announcement of world health organization (WHO), regarding the start of phase 6 pandemic of swine flu on 11 June 2009, we have witnessed the rapid progression of the first pandemic of twenty first century [1]. Since introduction of swine flu (H1N1) infection different reports have been released indicating that H1N1 behavior has some unique features including: high rate of hospitalization and mortality, involvement of young age group rather than extreme of ages and higher risk of influenza complications among patients with morbid obesity, that differentiate it from seasonal flu [2]. Furthermore, the magnitude of hospitalization and mortality rate and to top it off imposed burden on health care system and affected patients showed a substantial difference in comparison with seasonal flu [3]. In this setting the adverse impact of this pandemic have not limited to the health care systems while a great social fear in the large gatherings and travelers (e.g. among pilgrims) was among consequences of this event [4].

Due to unusual and atypical clinical and laboratory presentations of H1N1 influenza with its consequent difficulties in differentiation of swine flu from seasonal one and even common cold, late or misdiagnosing have led to some complications for the affected patients [5]. To avoid delaying in proper decision making and rational prescription of antiviral medication, oseltamivir, this study aims to develop a hybrid model based on fuzzy, statistical and neural learning theories in order to predict presence of this infection in its vital period. The proposed hybrid scheme benefits all strong points of human-like decision making (fuzzy method), generalization and robustness of statistical models and adopt the ability of human-like neural learning to gather all these capabilities in one hybrid model. The rest of this paper is structured as follows. First, the related works are explained and their positive and negative points are discussed. Next, the methodology has been used in this article is presented. Afterward, experimental results are brought and analyzed in the results and discussions section. Finally, the paper is concluded by a conclusion and future works part.

Related Works

Prediction of lethal diseases has attracted many attentions and the number of research teams who are working on this issue is significantly increasing [6]. Several the state-of-art methods have been employed in this field such as polynomial predictors [7], support vector machines (SVMs) [8], fuzzy predictors such as adaptive neuro-fuzzy interference system (ANFIS) [9], as well as neural approaches including multi-layer perceptron (MLP), radial basis functions (RBFs) and modular neural networks [10].

A variety of medical tasks has been successfully performed using such methods including detection of electrocardiogram (ECG) arrhythmia using fuzzy classifier [11], decision-making in pathology by neural networks [12,13], prediction of influenza vaccination [14], presenting a stochastic model to predict pandemic influenza [15], prediction of gastro-intestinal absorption using adaptive splines [16], prediction of pulmonary embolism [17] by neural network, and early detection in mammographic images by SVMs and fuzzy SVMs [18,19]. Moreover, the mentioned prediction methods were used commonly in other fields such as predicting of stock market indexes [20] by SVM, forecasting of weather status by neural networks [21], predicting on sea waves behaviors to anticipate Tsunamis using neural network [22], prediction of Alzheimer prognosis based on estimating the electroencephalogram (EEG) behaviors [23], forecasting the seizure attack time for epileptic patients [24], and prediction the cancer growth [25]. Although lots of research has been done for preparation of societies to inhibit the spreading of influenza A (H1N1) but no reliable prediction method has been proposed which is verified FDA institute. Therefore, this paper is aimed at predicting H1N1 influenza...
using quantitative clinical signs and symptoms in its golden time. To achieve this aim, this study presents a hybrid model based on statistical methods as well as neural network and fuzzy logic to predict the influenza A (H1N1) in its golden time.

**Materials and Methods**

**Subjects and source of information**

Based on the national protocol in Islamic Republic of Iran [26,27], a suspected case of influenza A (H1N1) virus infection was defined as one who has high grade fever (>38°C) or has at least two acute respiratory symptoms including: nasal obstruction/rhino rhea, sore throat, cough, fever (feverishness) and met at least one of the following criteria:

1) Within the last seven days returned from a country or region with an epidemic of influenza A (H1N1).

2) Being in close contact (within two meters) with a confirmed case of influenza A (H1N1) within the past seven days.

3) Patients with moderate to severe respiratory illness requiring hospitalization, or unexplained or unusual clinical patterns associated with serious or fatal cases.

On the other hand confirmed case of influenza A (H1N1) was defined as one who has high grade fever (>38°C) or has at least two acute respiratory symptoms including: nasal obstruction/rhino rhea, sore throat, cough, fever (feverishness) and influenza A (H1N1) virus infection that has been confirmed by Reverse Transcriptase PCR (RT-PCR). RT-PCR is the gold standard test to rule in or rule out the presence of virus and for this reason we provided the information for training of the software from confirmed cases and suspected cases for those the result of RT-PCR became negative.

**Artificial neural network**

Artificial neural networks [12] have been developed for a wide variety of problems such as classification, function approximation, and prediction. A neural network is structured by a parallel full-connection computation units arranged in layers mimicking the physiologic structure of the brain. There are multiple connections within and between the layers which indicate the strengths or weights between neurons that are learned under an optimization criterion. All Information learnt by the network is stored in the interconnection weights between each two successive layers. In medical research, the most commonly used artificial neural networks (ANN) are multi-layer perceptrons (MLP) which schematically depicted in Figure 1.

The implemented neural network uses back propagation in order to learn the parameters (weights) of the model, usually squared error or maximum likelihood, using a gradient optimization method [12]. In MLP networks, the error (the difference between the predicted and the target values) is propagated back from the output to the connection weights and updates the weights to minimize the prediction error [12]. The employed MLP is constructed by three layers with back propagation learning algorithm. Quantitative clinical symptoms of are used as input features to the network and the training algorithm iteratively modified the numeric values of these weights to decrease the training error of the network. In order to avoid over-fitting, 10-times 10-folds cross validation is utilized to determine the optimum parameters of the MLP in terms of number of neurons in the hidden layer (here is selected 6) and the function type of neurons (which is selected as sigmoid). After training of the MLP, the weights are fixed for the test phase. When a set of input feature values are presented to the trained network, an output is generated to predict the label of the input vector. By applying a sign function the predicted value is classified to the positive or negative classes which show that the patient is with or without the H1N1, respectively. It should be noted that there is no overlap between the train and test data and in other words, test set is considered as blind or unseen data.

**Gaussian mixture model (GMM)**

Statistical models have always been applied to prediction of different diseases. One of the models which are repeatedly being utilized for influenza prediction is Gaussian mixture model (GMM) [28,29]. GMM models the probability density function of the observed data using a multivariate Gaussian mixture density. This mixture is formulated in terms of Gaussian functions and their importance. The implemented components are expressed as follows.

\[
F(x; \theta) = \sum_{i=1}^{K} p_i \cdot g(x; m_i, \sigma_i)
\]

\[
g(x; m_i, \sigma_i) = \frac{1}{(\sqrt{2\pi} \sigma_i)^D} e^{-\frac{1}{2} \frac{\|x - m_i\|^2}{\sigma_i}}
\]

where \(p_i\) is the weight of \(i\)th Gaussian function. A graphic representation of GMM is depicted in Figure 2. In the current problem, we are faced with a two-class problem; therefore, a GMM should be trained to recognize the patterns of each class separately. First we apply K-means to clustering input data of each class separately. K-means tries to find those clusters in which data are distributed. The point is...
that K-means is not a stable algorithm due to this fact that the formed clusters depend seriously on the initial centers considered in each run. To improve stability of K-means, we have applied a method to initialize the cluster centers such that the outcome will be more stable [30]. In this approach the cluster centers have the maximum distance from each other.

Another problem in K-means is the number of clusters which should be known beforehand. There are several methods for finding the number of clusters among which Gap-Statistic has proven to be quite applicable [31]. In this algorithm the optimum number of clusters is found. Therefore, it is expected that the formed clusters would be denser and are placed far away from each other.

After finding the initial clusters, we apply expectation maximization (EM) algorithm to find the Gaussian components for each of the two classes.

EM is a soft version of K-means algorithm in which the data membership in each cluster is a binary value (0 or 1) while in EM, an observation can have a membership degree ranged from 0 to 1. It means data can belong to more than one Gaussian component with different probabilities. The clusters’ centers and number of instances in each cluster are used as Gaussian mixture model parameters.

Finally the probability of data is calculated in each class and the division is compared to a threshold; consequently, the class label can be determined. Afterward, the model is trained and assessed by 10-times 10-folds cross validation in order to avoid over-fitting. Regarding the training results, a threshold is determined for each GMM model. Next, the trained GMM is applied to the test data and the weighted summation of Gaussian outputs for each input pattern determines its probability given the model. Finally, to take the last decision, each pattern is applied simultaneously to the trained models, and a simple maximizing operand determines the predicted label. In order to use GMM as a predictor instead of a classifier, the maximum GMM real value is considered as the predicted value with regard to that input sample.

Fuzzy rule-based classifier

One of the efficient learning methods which mimic human-like decision, especially in disease diagnosis, is fuzzy rule-based system (FRBS) which is a special case of fuzzy modeling where the output of the system is crisp and discrete. The main advantage of this classifier is the interpretability and human understandability of the model [6]. Basically, to train a FRBS, a compact set of fuzzy If–Then rules should be found in order to model the input–output behavior of the system [32].

Rules are generated from a set of labeled data from all classes [33]. After the training phase, we can test the classifier with unseen patterns. Many approaches have been proposed to construct a rule engine from input numerical data. These include heuristic approaches [34,13], neuro-fuzzy techniques [12], and clustering methods [35]. Here, we use fuzzy rules of the following type for this problem:

Rule Rj: If \(x_1 \) is \(A_{j1} \), and \( \ldots \) and \( x_n \) is \(A_{jn} \), then C with \( CF=\sigma_j \), \( j = 1, 2, \ldots, N \), where \( X = [x_1, x_2, \ldots, x_n] \) is the input feature vector, \( A_{ji} \) is the fuzzy set associated to \( x_i \), C is the class label, \( CF \) is the certainty degree of rule \( Rj \) and \( N \) is the number of fuzzy rules in the rule-base system. In order to classify an input pattern \( X = [x_1, x_2, \ldots, x_n] \), the degree of compatibility of that pattern with each rule should be computed (i.e., using a T-norm to model the “and” connectives in the rule antecedent).

In case of using product as T-norm, the compatibility grade of rule \( Rj \) with the input pattern \( X \) can be calculated as:

\[
\mu j(X t)=\prod_{i=1}^{n} \mu_i (A_j(x_i)).
\]

Where \( \mu(X t) \) is the fuzzy membership value of its input pattern \( X \) in the \( j \)th membership function. In case of using weighted vote (10,20,21) as the reasoning mechanism, each fuzzy rule gives a vote for its consequent class.

\[
\mu+=\sum_{j=1}^{N} \mu_j(X t)>CFj \ \mu_j(X t) \text{and } j=1,...,N
\]

\[
\mu-=\sum_{j=1}^{N} \mu_j(X t)<CFj \ \mu_j(X t) \text{and } j=1,...,N
\]

1. If \( \mu+>\mu- \), then \( X \) is classified as positive.
2. If \( \mu+<\mu- \), then \( X \) is classified as negative.
3. If \( \mu+=\mu- \), then don’t classify \( X \).

Two stages for generating rules are used in this study as shown in Figure 3. In the first stage, rules are generated for each training data. Triangular fuzzy sets are used and number of fuzzy sets and parameters of membership functions (MFs) on attributes are dependent to values of that attribute. We also determine degree of each rule and select rules with maximum degrees. Degrees are determined based on compatibility of training data with fuzzy sets defined on each attribute as described in [29,36]. In the second stage, the objective function is constructed in terms of sensitivity and specificity (i.e., simulated annealing) to improve the FRBS performance. Finally, a majority vote mechanism is used to determine the class label.

The proposed hybrid method

We have proposed a method composed of three components in terms of artificial Neural Network (ANN), fuzzy rule-based classifier (FRBS), and Gaussian mixture model (GMM). In the proposed method, the decision is made based on the result of each of the three methods. A weighted voting mechanism is used and a class with the major vote is selected to determine the label of the input data. We first trained each of the above methods independently using the training data to find their proper parameters. Finally a hybrid method is constructed using the three methods and weights of each method are adjusted using the training data as shown in Figure 4.


![Figure 3: Stages of constructing of an FRBS.](image)

![Figure 4: Structure of the proposed hybrid model using MLP, FRBS and GMM modalities are shown.](image)
Results and Discussions

In this section, the achieved results by each of the modalities along with the hybrid method are presented. In this study, 230 cases were selected randomly from the examination results of patients at Namazi and Aliashgar hospitals which were under supervision of Shiraz University of Medical Sciences. The diagnoses were performed by professional specialists from 15 September to 10 December 2009. A list of clinical attributes and their corresponding positive percent used in this experience is shown in the Table 1.

We used polymerase chain reaction (PCR) test results as the label for each case. PCR tests showed that 37% of patients were infected by H1N1. We trained the proposed method using 10-times 10-folds cross validation to evaluate the effectness of our method. The results achieved by applying each decision making component was separately determined. Then, the hybrid method was employed and the results are compared to those achieved by each modality. A comparison of the performance of each method is shown in Table 2. Different terminology is used to show the performance of each method including sensitivity or true positive rate (TPR) which measures the proportion of actual positives which are correctly identified as such; specificity or true negative rate (TNR) which measures the proportion of actual negatives which are correctly identified. Moreover, other indexes such as positive predictive value (PPV) which measures the proportion of the identified positives that are actually positive, negative predictive value (NPV) which measures the proportion of the identified negatives which are actually negative and finally the accuracy which measures the proportion of all cases which are identified correctly are calculated using applying the data to the models. Another significant criterion, especially in the medical applications, is area under curve (AUC) which is a nonparametric measure of discrimination. The receiver operating characteristic (ROC) area measures the relative goodness of the predictions entirely by comparing the predicted probability of each patient with that of all the other patients. Receiver operating characteristic area is independent of both the prior probability of each prediction and the threshold cut-off for classification. In addition, the mean and standard deviation for each terminology is also reported in Table 2. It can be seen that the hybrid method outperformed the other modalities in terms of TPR, NPV, accuracy, and AUC indexes.

To test for significant difference between the proposed hybrid method and its decision making components, the statistical T-test is utilized to reveal the significance of our results. The statistics t-test allows us to determine p-value that indicates how likely we could have gotten these results by chance. If the p-value is less than 0.05, we can conclude that there is a statistically reliable difference between the mean results of the two methods. The p-value result of T-test for each terminology between the hybrid method and ANN, FRBS, and GMM methods is reported in Table 3. We are able to test whether the significant improvements of the hybrid method in prediction has enough generalization across the unseen data (test set) or not. T-test reveals that the hybrid method predictions of unseen patients were significantly more accurate than each of the FRBC, MLP, and GMM methods separately.

Our hybrid method exploits its components and compensate for each method’s disadvantages. Neural network derives its computing power through its massively parallel distributed structure and its ability to learn and generalize. The nonlinearity, adaptability, confidence in decision making and neurobiological analogy of neural network makes it a powerful tool for a reliable prediction. Fuzzy inference systems can incorporate and handle uncertainty; therefore, it provides a systematic calculus to deal with it linguistically, and it performs numerical computation by using linguistic labels specified by membership functions. A selection of fuzzy if-then rule can effectively model human expertise but it suffers adequate adaptability to deal with changing environment. Gaussian mixture model is useful in case of multi modal distributions but is an unstable method due to its high sensitivity to its initialization values. In order to have a method that benefits all advantages of each described modality, a hybrid scheme is proposed...
which fuse the results of different modality and takes its decision according to the majority vote scheme.

In order to make our work applicable in hospitals, we designed a user-friendly graphic user interface (GUI) software to predict H1N1 influenza by each modality or using hybrid method. The GUI of this software allows the user to select one of the four models which is shown in Figure 5. This software is now being used in many hospitals which are all under supervision of Shiraz University of Medical Sciences to assist low-experienced graduated doctors to detect H1N1 influenza.

Conclusion and Future Works

In this paper we proposed a hybrid method to predict H1N1 influenza with high sensitivity and specificity that can satisfy the specialist expectations. The proposed method is able to learn well, takes a robust decision and models the statistics of input data. The reason of inheriting such strong capability is to use three prediction modalities like the chest status, for each suspicious subject, the hybrid method is expected to result in a higher accuracy, robustness, and generalization properties.

References
