

Mosquito Borne Disease Incidence Prediction System using Fuzzy Weighted Associative Classification

M. Kalpana Devi

Research Scholar, Dept. of Computer Science,
Sri Padmavathi Mahila University, Tirupati
Andhra Pradesh, India

M. Usha Rani

Professor, Dept. of Computer Science,
Sri Padmavathi Mahila University, Tirupati
Andhra Pradesh, India

ABSTRACT

Recently, applications attracted rampant attention in Epidemiology, Medical Entomology, Bio informatics, and Bio surveillance. Data mining applications is greatly useful to all stake holders in the healthcare industry. Associative Classification (AC) is a branch of data mining, a larger area of scientific study. To build a model for the purpose of prediction, AC is a suitable prediction technique, which integrates two data mining tasks, association rule mining and classification. The main aim of classification is the prediction of class labels, while association rule discovery describes relationship between items in a transactional database. Of late, Associative Classifier is having better accuracy as compared to that of traditional classifiers. Mosquito Borne Diseases that place a heavy burden on public health system on most of the tropical countries around the world. There is a need to develop prediction methods to augment existing control strategies. In this paper we use Fuzzy Weighted Associative Classifier to build an effective prediction model to predict mosquito borne disease incidence

Keywords

Epidemiology, Medical Entomology, Bioinformatics, Bio surveillance, Associative Classification, Mosquito Borne Diseases, Fuzzy Weighted Associative Classifier.

1. INTRODUCTION

Besides causing heavy burden on public health systems, the widespread of mosquito borne diseases to humans is also causing substantial morbidity and economic loss to our nation and also most of the tropical countries around the world. The mosquito borne diseases such as Malaria, Dengue, Chikungunya, Lymphatic Filariasis and Japanese Encephalitis are major threat to human population.

Disease incidence statistics, given by Priya shetty is, in the South-East Asian Region of WHO, out of 1.4 billion people living in 11 countries (land area 8,466,600 km, i.e. 6% of global area), 1.2 billion are exposed to the risk of malaria and most of them live in India. However, the South-East Asia contributed only 2.5 million cases to the global burden of malaria. Of this, India alone contributed 76% of the total cases [1]. So, there should be a proper planning and implementation of control measures which require proper exploitation of the available data. There is a need to develop prediction methods to augment existing control strategies. Mosquito borne disease management requires a clear understanding of all concerned variables that directly or indirectly led to mosquito proliferation, pathogen sustainability and disease epidemics. Many diseases are influenced by the environmental conditions, but Mosquito borne diseases are more sensitive, since environmental conditions affect the infecting pathogens,

human hosts and the insects and other intermediate hosts that transmit them.

The spread of Mosquito borne diseases are influenced by the change in climate both directly and indirectly. Global warming and increased rainfall contribute to the abundance and distribution of vectors such as mosquitoes. Rainfall leads to stagnant water pools in which mosquitoes breed. Global warming leads to increased global temperatures allowing mosquitoes to survive in winters where they would otherwise have perished, subsequently more mosquitoes are alive to breed and transmit diseases during the summer season. Besides above, mosquito larvae develop much faster and after ingestion of the virus become more infectious at higher temperatures. Other factors such as deforestation and natural disasters, decreased hygiene and stagnant water beds increase the proliferation and survival of vectors, hence the incidence of vector-borne diseases [2] [3].

However, the quantitative relationship between meteorological variables and the spatial and temporal distributions of these infectious diseases is still not clear. Hence, in this paper we try to make an attempt to establish the association between meteorological variables and disease incidence. Associative classification is one of the best data mining techniques to identify associations between items and also is a suitable for building an effective prediction model. This paper describes the prediction model, which predicts mosquito borne disease incidence by identifying associations between climatic variables and disease incidence using Fuzzy Weighted Associative Classification technique.

2. CLIMATIC FACTORS AND DISEASE INCIDENCE

It was found by many researchers that there is a strong relationship between climatic conditions and disease epidemics, because mosquitoes can survive and also transmit diseases in certain climatic conditions. According to Tabachnik, Climate change can impact the mosquito-borne disease epidemiology and it will influence arthropod vectors, their lifecycles and the human or animal host, consequently changes both vector and pathogen distribution and changes in the ability of arthropods to transmit pathogens [4]. **Temperature**, **rainfall**, and **humidity** are especially important, but others, such as **wind velocity** and the duration of **daylight (sunshine)**, can also be significant. These factors play a vital role in the survival and transmission rate of pathogens and mainly temperature affects their rate of multiplication in the mosquito. Hence, this affects the rate at which the salivary secretions become infected, and thus the successful transmission of pathogens to another host.

As it was envisaged by Environmental Information System, Center for bioinformatics and vector control [5], rainfall can impact on the risk of transmission of vector-borne disease. Because increased surface water can provide breeding sites for vectors. Increased rain can increase vegetation and allow expansion in population of vertebrate host. Mosquitoes are highly sensitive to humidity also. Relative humidity will have an impact on the breeding and transmission characteristics of vectors.

Based on the review regarding climatic factors which influences disease incidence, we have considered six variables such as Temperature, Humidity, Rainfall, Evaporation, Wind Velocity and Sunshine. Now the question arises over the duration for which, climatic variables should be sensibly averaged. Shorter the aggregation period makes stronger the likelihood of a higher degree of serial autocorrelation in the values. So, for explaining the variation in disease prevalence, it has been decided to average the monthly climatic data over climatic seasons in order to reflect the variations in weather. In this paper the individual monthly averages of the climatic variables have been taken as predictor variables and malaria disease incidence as predictive variable.

3. RELATED REVIEW

3.1 Associative Classification

The concept of association rule was first introduced by Agrawal et al. [6]. It proposed the support-confidence measurement framework and association rule mining to the discovery of frequent item sets. The following year a fast mining algorithm, Apriori, was proposed [7]. Since then much effort has been dedicated to the classical association rule mining. Numerous algorithms have been proposed to extract the rules more efficiently [8], [9], [10]. These algorithms follow the same classical measurement framework and produce good results once the minimum support and minimum confidence are given.

Classification is an important process in many of the real time applications. Examples of Classification tasks may be found in nearly every field of endeavor such as banking, weather, education, medicine, homeland security etc. a classification method is seen as a prediction task, because it aims at predicting the value of an attribute in a data set.

Associative Classification is an integrated framework of Association Rule Mining and Classification [11]. A special subset of association rules whose right-hand-side is restricted to the classification class attribute is used for classification. Taking into account that, there is one and only one predetermined target for classification rule mining, while for association rule mining the target of discovery is not predetermined, it is necessary to constrain the rules consequent terms to encompass only one attribute. This way, the consequent term of an association rule will represent the target, or class attribute. Therefore, such rule can play a prediction role in a given system: in order to classify an item, the rule's properties are matched to every rule's antecedents and the attribute value of the consequent term will be the predicted class. Prediction is a vital and important task in Data Mining. The aim of the proposed prediction model is to predict a future value. This can be achieved by building a model that generates and evaluates a set of rules for prediction.

3.2. Classification based on fuzzy association rule mining

Numerous Fuzzy Data mining algorithms have been evolved recently especially in associative rule mining, there are many applications [12], [13] [14]. Fuzzification is an effective model in prediction and also in the construction of expert systems. The main purpose of using fuzzy logic is to overcome the sharp boundary problem. It can be used to transform quantitative data into fuzzy data through identification of suitable membership function (Trapezoidal, Triangular, Gaussian, Bell). Fig. 1 shows membership functions we considered for Temperature attribute and linguistic variables Low, Moderate and High.

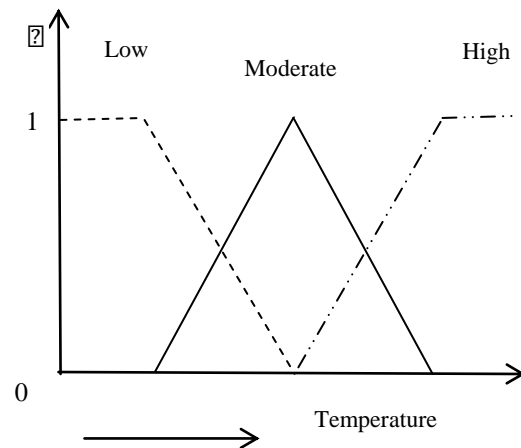


Fig. 1 Fuzzy sets Low, Moderate and High.

3.3 Weighted Association Rule Mining

Traditional association rule mining algorithm consider all items have equal importance, whereas weighted association rule mining assign weights to each item based on their significance. The concept of assigning weights to each attribute is first introduced by Ramkumar et al. [15]. He has considered cost of the item for assigning weights and a new measure was introduced, *weighted support*. Recently, many algorithms have been successfully proposed for mining association rules with weighted settings [16], [17], [18].

For example in the supermarket domain, weight is assigned to each of the items as per the profit it generates to the store, rather than simply calculating the percentage of transactions. In our proposed domain, some of the attributes have much impact to predict disease incidence. For example in the prediction of mosquito borne disease incidence, the attribute "Temperature" has more impact than the attribute "Wind velocity". So we have used the attribute weightage concept to incorporate attribute importance.

Table.4. shows fuzzy attribute weights. In our prediction model, weights assigned to each attribute were estimated based on the correlation between the attribute and the disease incidence.

4. MATERIALS AND METHODS

4.1. Study Area

Tirupati is an urbanized city situated at 13° 39' 0" N, 79° 25' 0" E. Tirupati is a major pilgrimage and cultural city in the Chittoor district of Indian state, Andhra Pradesh. Monsoon remains moderate and summer experiences temperatures ranging from 35 to 40 degrees Celsius. In winter the minimum temperatures will be between 18 and 20 degrees Celsius. Usually summer starts from March which ends in June, with the advent of rainy season in July, followed by winter which lasts till the end of February. As per provisional reports of Census India, population of Tirupati is 459,985 and is the 9th most populous urban agglomeration in Andhra Pradesh [19].

4.2. Data Collection

Malaria cases reported month wise were obtained from Medical and Health department, Chittoor district. Monthly averages of climatic data such as Temperature, Total rainfall, Relative humidity, Wind speed, Sunshine and Evaporation were collected from Regional Agricultural Research Station, Agricultural Meteorology Division, India meteorological department, Ministry of earth science, Tirupati.

4.3. Prediction Model

This prediction model was constructed using classification based on association. Given a dataset with class labels as a training dataset, classification is to build a model to predict future data object for which the class label is unknown.

4.3.1 Fuzzy Weighted Associative Classifier

FWAC is a new concept proposed by Sunita et al. [20] Fuzzy Weighted ARM uses Fuzzy Weighted Support and Fuzzy Weighted Confidence Framework to extract Association rules from data repository.

The major steps are as follows.

- 1) The monthly averages of climatic data and malaria cases of the following months were placed in a database in order to make it suitable for the mining process.
- 2) Each attribute is partitioned into fuzzy sets.
- 3) Transform original database into binary by comparing with the fuzzy ranges.
- 4) Select appropriate membership function and transform original database into fuzzy membership values.
- 5) Each fuzzy attribute is assigned a weight ranging from 0 to 1 to reflect their importance in this prediction model. Attributes that have more impact will be assigned a high weightage and attributes having less impact are assigned low weightage.
- 6) Generate Classification Association Rules using frequent itemset generation and is represented as $X \rightarrow C$ where X is subset of climatic variables and C is a class label such as malaria cases. Examples of such rules are $\{(Temperature, 'Moderate'), (Rainfall, 'High')\} \rightarrow (Malaria\ Incidence, 'High')$ and $\{(Humidity, 'Low'), (Rainfall, 'Moderate')\} \rightarrow (Malaria\ Incidence, 'Moderate')$.
- 7) Compute Fuzzy Weighted Support (FWS) and Fuzzy Weighted Confidence (FWC) for each rule.
- 8) Select rules that have a FWS and FWC larger than a given threshold values. These rules will be stored in Rule Base. Then these rules are sorted in an order of higher precedence.
- 9) Whenever a new record is provided, the rules from the rule base are used to predict the class label.

Table.1 Database D with continuous domain.

| R. No | Temperature(T) | Humidity(H) | Rainfall(RF) | Wind Velocity(WV) | Malaria Cases(MC) |
|-------|----------------|-------------|--------------|-------------------|-------------------|
| 1 | 22.9 | 71.5 | 0 | 5 | 16 |
| 2 | 19.55 | 43.8 | 15.5 | 7.3 | 7 |
| 3 | 27.3 | 54.1 | 60 | 6.3 | 22 |
| 4 | 30.965 | 75.05 | 20.5 | 3.9 | 27 |
| 5 | 32.55 | 88.1 | 117.7 | 3.6 | 42 |

Table.2 Transformed Binary Database BD from Database D.

| S. No | T | | | H | | | RF | | | WV | | | MC | | |
|-------|-----|----------|------|-----|----------|------|-----|----------|------|-----|----------|------|-----|----------|------|
| | Low | Moderate | High | Low | Moderate | High | Low | Moderate | High | Low | Moderate | High | Low | Moderate | High |
| 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 2 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 3 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| 4 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 5 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |

Table.3 Database FD with Fuzzy membership values.

| S. No | T | | | H | | | RF | | | WV | | | MC | | |
|-------|-----|----------|------|-----|----------|------|-----|----------|------|-----|----------|------|-----|----------|------|
| | Low | Moderate | High | Low | Moderate | High | Low | Moderate | High | Low | Moderate | High | Low | Moderate | High |
| 1 | 0.3 | 0.6 | 0.1 | 0.1 | 0.2 | 0.7 | 0.9 | 0.1 | 0.0 | 0.4 | 0.5 | 0.1 | 0 | 1 | 0 |
| 2 | 0.7 | 0.2 | 0.1 | 0.6 | 0.2 | 0.2 | 0.8 | 0.2 | 0.0 | 0.2 | 0.6 | 0.2 | 1 | 0 | 0 |
| 3 | 0.1 | 0.7 | 0.2 | 0.1 | 0.6 | 0.3 | 0.4 | 0.6 | 0.0 | 0.2 | 0.5 | 0.3 | 0 | 1 | 0 |
| 4 | 0.1 | 0.4 | 0.5 | 0.0 | 0.3 | 0.7 | 0.8 | 0.2 | 0.0 | 0.6 | 0.3 | 0.1 | 0 | 1 | 0 |
| 5 | 0.1 | 0.3 | 0.6 | 0.0 | 0.2 | 0.8 | 0.1 | 0.1 | 0.8 | 0.7 | 0.3 | 0.0 | 0 | 0 | 1 |

Table.4 Fuzzy Attribute Weight

| Attribute | Weight | Attribute | Weight | Attribute | Weight | Attribute | Weight |
|-------------|--------|-------------|--------|--------------|--------|--------------|--------|
| T, Low | 0.2 | H, Low | 0.2 | RF, Low | 0.2 | WV, Low | 0.6 |
| T, Moderate | 0.5 | H, Moderate | 0.4 | RF, Moderate | 0.6 | WV, Moderate | 0.2 |
| T, High | 0.7 | H, High | 0.6 | RF, High | 0.8 | WV, High | 0.2 |
| Total | 1.4 | | 1.2 | | 1.6 | | 1.0 |

Table.1 presents a part of the database (D) with its transformed binary database (BD) as shown in Table.2, in

which quantitative attributes have been partitioned into categorical attributes. For example attribute ‘Rain Fall’ in D is

replaced with three attributes such as (RF, Low), (RF, Moderate) and (RF, High) in the Binary Database BD. In the present model, class variable is also partitioned into three categorical attributes such as (MC, Low), (MC, Moderate) and (MC, High). Table.3 illustrates an example of the new database FD obtained from the original database D, given fuzzy sets {Low, Moderate, High} as characterized by membership functions. Here Fuzzy logic is incorporated to split the domain of quantitative attribute into intervals, and to define a set of meaningful linguistic labels represented by fuzzy sets and use them as a new domain.

A fuzzy dataset consists of fuzzy relational database $D = \{r_1, r_2, r_3, \dots, r_i, \dots, r_n\}$ with a set of attributes $I = \{I_1, I_2, \dots, I_k, \dots, I_m\}$, each I_k can be associated with a set of linguistic labels $L = \{l_1, l_2, \dots, l_L\}$ for example $L = \{Low, Moderate, High\}$. Let each I_k is associated with fuzzy set $F_k = \{(I_k, l_1), (I_k, l_2), \dots, (I_k, l_L)\}$. So that a new Fuzzy Database FD is defined as $\{(I_1, l_1), \dots, (I_1, l_L), \dots, (I_k, l_1), \dots, (I_k, l_L), \dots, (I_m, l_1), \dots, (I_m, l_L)\}$. Each attribute I_i in a given transaction t_k is associated (to some degree) with several fuzzy sets. The degree of association is given by a membership degree in the range $[0..1]$. $t_k[\mu(I_i, l_j)]$ will denote the degree of membership for Fuzzy Attribute I_i to fuzzy set l_j in transaction t_k .

A weight of an item is a non-negative real number that shows the importance of each item. A pair (x, w) is called a weighted item where $x \in I$ is an item and $w \in W$ is the weight associated with x . A transaction is a set of weighted items, each of which may appear in multiple transactions with different weights. Table.4 shows weights assigned to each fuzzy attributes.

4.3.2 Fuzzy Weighted Support (FWS)

In this present associative classification model, rules are of the form $X \rightarrow \text{Class label}$, Where X is a subset of fuzzy weighted attributes. FWS of a rule $X \rightarrow \text{Class label}$ is calculated as sum of weight of all transaction in which the given class label is true, divided by total number of transactions, denoted by FWS ($X \rightarrow \text{Class label}$). And is given by

$$\forall (I_i, l_j) \in X$$

'n' the number of records in D. Then

$$FWS(X \rightarrow \text{Class label}) =$$

$$\frac{\sum_{\text{For which Class label} = 1} \prod_{i=1}^{|X|} (\mu(I_i, l_j) \times W(I_i, l_j))}{n}$$

Where $\mu(I_i, l_j)$ is a degree of membership for fuzzy attribute and $W(I_i, l_j)$ is weight of a fuzzy item.

For Example: FWS of a rule $\{(T, Moderate), (H, High)\} \rightarrow (MC, Moderate)$

$$FWS =$$

$$\frac{[(0.6 \times 0.5)(0.7 \times 0.6) + (0.7 \times 0.5)(0.3 \times 0.6) + (0.4 \times 0.5)(0.7 \times 0.6)]}{5}$$

$$FWS = \frac{0.273}{5} = 0.0546$$

4.3.3 Fuzzy Weighted Confidence (FWC)

Fuzzy Weighted Confidence of a rule $X \rightarrow Y$ where Y represents the Class label can be defined as Fuzzy Weighted Support of $(X \rightarrow Y)$ divided by Fuzzy Weighted Support of (X) . And is given by

$$FWC(X \rightarrow Y) = \frac{\text{Fuzzy Weighted Support}(XUY)}{\text{Fuzzy Weighted Support}(X)}$$

For Example: FWC of a rule $\{(T, Moderate), (H, High)\} \rightarrow (MC, Moderate)$

$$FWC =$$

$$\frac{[(0.6 \times 0.5)(0.7 \times 0.6) + (0.7 \times 0.5)(0.3 \times 0.6) + (0.4 \times 0.5)(0.7 \times 0.6)]}{[(0.6 \times 0.5)(0.7 \times 0.6) + (0.2 \times 0.5)(0.1 \times 0.6) + (0.7 \times 0.5)(0.3 \times 0.6) + (0.4 \times 0.5)(0.7 \times 0.6) + (0.3 \times 0.5)(0.8 \times 0.6)]}$$

$$FWC = \frac{0.273}{0.351} = 0.78 (78\%)$$

5. EXPERIMENTAL RESULTS

To implement this work, Java has been used as front end and Oracle as back end (NetBeans IDE 6.8). In this present study, we got 17 years data ie.204 records. 120 records were used as training dataset and 84 records for testing. Accuracy of the system is 87% with 20% as minimum support and 80% as minimum confidence. Some of the screen shots have been shown in the following figures.

| YEAR | MONTH | TEMP | HUM | EP | RAINFALL | W | SS | MILARGA_CA |
|------|-------|-------|-------|------|----------|------|------|------------|
| 2011 | JAN | 22.9 | 77.5 | 4.4 | 0 | 0 | 77.5 | 17 |
| 2011 | FEB | 24.55 | 63.8 | 4.9 | 15.5 | 5.7 | 77.7 | 16 |
| 2011 | MAR | 27.3 | 54.1 | 7 | 0 | 6.3 | 8.2 | 7 |
| 2011 | APR | 30.96 | 55.05 | 6.6 | 20.5 | 5.9 | 7.8 | 22 |
| 2011 | MAY | 32.55 | 48.1 | 7.7 | 117.7 | 7.6 | 8.7 | 17 |
| 2011 | JUN | 33.55 | 70.8 | 4 | 78.8 | 5.53 | 4.06 | 21 |
| 2011 | JULY | 30.7 | 55.5 | 5.8 | 112.6 | 0 | 3.6 | 41 |
| 2011 | AUG | 28.17 | 64.65 | 4.43 | 121 | 7.83 | 3.4 | 22 |
| 2011 | SEPT | 25.39 | 59.35 | 5.46 | 119.8 | 6.22 | 6.38 | 13 |
| 2011 | OCT | 28.22 | 66.98 | 3.64 | 115 | 6 | 6.5 | 39 |
| 2011 | NOV | 24.75 | 71.70 | 4.07 | 124.4 | 5.5 | 5.12 | 31 |
| 2011 | DEC | 22.4 | 66.95 | 7 | 64.4 | 12.3 | 5.2 | 18 |
| 2010 | JAN | 27.98 | 64.99 | 4.02 | 0 | 57.5 | 6.49 | 5 |
| 2010 | FEB | 26 | 56.85 | 5.1 | 0 | 5.7 | 8.1 | 7 |
| 2010 | APR | 33.2 | 49.3 | 7.6 | 0 | 6.1 | 7.9 | 11 |
| 2010 | MAY | 32.85 | 49.5 | 8.3 | 109.8 | 0 | 6.8 | 15 |
| 2010 | JUNE | 30.35 | 60.1 | 5.5 | 119.8 | 6.9 | 5.7 | 20 |
| 2010 | JULY | 28.65 | 67 | 4.1 | 162 | 6.7 | 2.8 | 25 |
| 2010 | AUG | 29.05 | 66.5 | 4.1 | 145.7 | 5.3 | 3.9 | 31 |
| 2010 | SEPT | 28.55 | 67.7 | 5.7 | 138.6 | 4.7 | 4 | 26 |
| 2010 | OCT | 27.95 | 70.2 | 3.5 | 103.1 | 4 | 3.7 | 45 |
| 2010 | NOV | 25.8 | 75 | 2.7 | 175.7 | 3.9 | 4.1 | 45 |
| 2010 | DEC | 23.3 | 64.5 | 5.3 | 38.4 | 4.8 | 3.6 | 19 |
| 2010 | MAR | 29.95 | 68.35 | 6.7 | 0 | 5.4 | 8.1 | 6 |

Fi g.2 GUI for Input dataset display

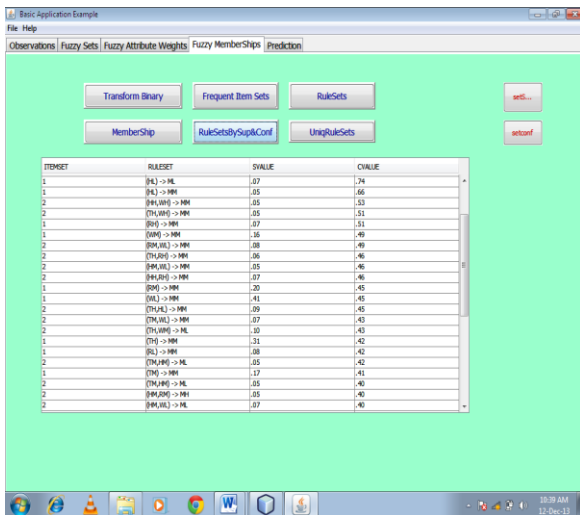


Fig. 3 GUI for Rule set display

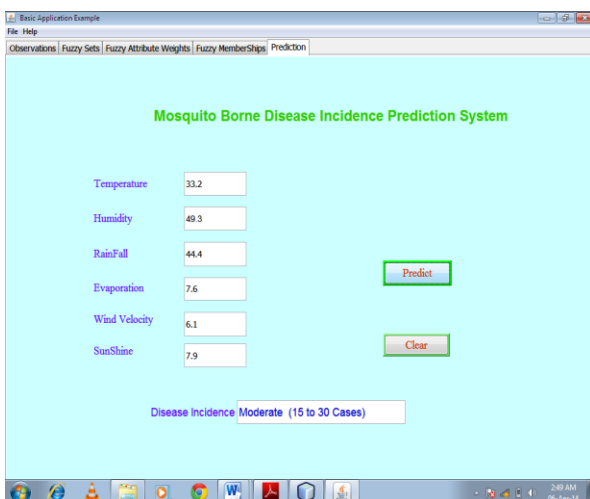


Fig.4 GUI for Disease Incidence Prediction System

6. CONCLUSION

Several studies documented relationship between climatic variables and malaria cases. Therefore it is reasonable to assume that climate would be a vital factor in mosquito borne disease epidemics. This study reveals that climatic variables could be important factors for the development of a low cost functional mosquito borne disease prediction system which is simple and precise. Our model forecasts disease incidence for the following month using monthly averages of climatic variables of present month. As climatic conditions differs from place to place, climatic data of a particular area to be considered for the prediction. Associative classification provides a fast and comprehensible learning model, when compared with that of the majority of traditional classifiers. The fuzzy approach can be used to assist in extracting knowledge from a database by transforming quantitative data (crisp data) into fuzzy data. Attribute weightage concept was also used to assign weights to each fuzzy attribute depending upon their importance in predicting the class label. In the future work, other interesting measures have to be implemented to improve the prediction accuracy and different geographical area has to be taken for prediction to identify the climatic conditions in which mosquitos can survive and transmit diseases.

7. REFERENCES

- [1] Priya Shetty, "Climate change and insect-borne disease: Facts and figures", Science and Development Network, 9 September 2009.
- [2] Zhou G, Minakawa N, Githeko AK, et al. Association between climate variability and malaria epidemics in the East African highlands. *Proc Natl Acad Sci U S A* 2004.
- [3] Hales S, de Wet N, Maindonald J, et al. Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. *Lancet* 2002.
- [4] W. J. Tabachnick, "Challenges in predicting climate and environmental effects on vector-borne disease systems in a changing world", University of Florida, *The Journal of Experimental Biology* 213, 946-954 © 2010.
- [5] Environmental Factors in the View of Vectors & Vector borne Diseases, Environmental Information System, Center for bioinformatics and vector control.
- [6] R. Agrawal, T. Imielinski, and A. Swami, 'Mining Association Rules between Sets of Items in Large Datasets' *Proc. ACM SIGMOD '93*, pp. 207-216, 1993.
- [7] R. Agrawal and R. Srikant, 'Fast Algorithms for Mining Association Rules', *Proc. 20th Int'l Conf. Very Large Data Bases (VLDB '94)*, pp. 487-499, 1994.
- [8] F. Bodon, 'A Survey on Frequent Itemset Mining', technical report, Budapest Univ. of Technology and Economics, 2006.
- [9] S. Brin, R. Motwani, J.D. Ullman, and S. Tsur, 'Dynamic Itemset Counting and Implication Rules for Market Basket Data', *Proc. ACM SIGMOD*, 1997.
- [10] J.S. Park, M. Chen, and P.S. Yu, 'An Effective Hash Based Algorithm for Mining Association Rules' *Proc. ACM SIGMOD*, 1995.
- [11] B. Liu, W. Hsu, and Y. Ma. 'Integrating classification and association rule mining', In *KDD'98*, pp. 80-86, New York, NY, Aug. 1998.
- [12] Delgado, M., Marin, N., Martin-Bautista, M. J., Sanchez, D. and Vila, M. A. (2003), 'Mining Fuzzy Association Rules: An Overview', In: *Proceedings Of the BISC International Workshop on Soft Computing for Internet and Bioinformatics*. Springer, Vol. 224, pp. 351-374.
- [13] Kuok, C.M., Fu, A., Wong, M.H.: Mining Fuzzy Association Rules in Databases. *SIGMOD Record* 27(1), 41-46 (1998)
- [14] Z. Chen and G. Chen, An approach to classification based on fuzzy association rules, in T. Li, Y. Xu, D. Ruan (Eds.), *Proc. Int. Conf. Intelligent Systems and Knowledge Engineering*, 2007.

- [15] G.D. Ramkumar, S. Ranka, and S. Tsur, “Weighted Association Rules: Model and Algorithm” Proc. ACM SIGKDD, 1998.
- [16] C.H. Cai, A.W.C. Fu, C.H. Cheng, and W.W. Kwong, “Mining Association Rules with Weighted Items” Proc. IEEE Int’l Database Eng. and Applications Symp. (IDEAS ’98), pp. 68-77, 1998.
- [17] W. Wang, J. Yang, and P.S. Yu, “Efficient Mining of Weighted Association Rules (WAR),” Proc. ACM SIGKDD ’00, pp. 270-274, 2000.
- [18] F. Tao, F. Murtagh, and M. Farid, “Weighted Association Rule Mining Using Weighted Support and Significance Framework” Proc. ACM SIGKDD ’03
- [19] [en.wikipedia.org/wiki/Tirupati_\(city\)](http://en.wikipedia.org/wiki/Tirupati_(city)).
- [20] Sunita Soni and O.P. Vyas ‘Fuzzy Weighted Associative Classifier: APredictive Technique For Health Care DataMining’, IJCSEI, Vol.2, No.1, February 2012.