International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

# Using Fuzzy Classifier for Cholera Analysis

**Okpor. M. D**

Department of Computer Science, Delta State, Polytechnic, Ozoro, Nigeria

**Abstract:** ***Cholera is a water borne disease which is epidemic and life-threatening in nature. It is usually characterized by numerous, voluminous watery stools, often accompanied by vomiting, and resulting in hypovolemic shock and acidosis. It is caused by a bacteria species Vibrio cholera which can also cause mild or in apparent infections. In this research paper fuzzy classification (classifier) a branch of soft-computing and subsequently a branch of Artificial intelligence is applied to objective recognition of cholera which invariably eliminates the uncertainty associated with conventional approach for cholera diagnosis.***

**Keywords:** Cholera, Fuzzy classifier, Fuzzy set, Fuzzy logic, Expert system

## 1. Introduction

Vibrio cholera is a Gram-negative, comma-shaped bacterium. Some strains of V. cholerae cause the disease cholera. V. cholerae is afacultative anaerobic organism (Davis, 2003) and has a flagellum at one cell pole.

This bacteria is residence in most contaminated water consumed by the host do not survive the highly acidic conditions of the human stomach. The few bacteria that do survive conserve their energy and stored nutrients during the passage through the stomach by shutting down much protein production. When the surviving bacteria exit the stomach and reach the small intestine, they need to propel themselves through the thick mucus that lines the small intestine to get to the intestinal wall where they can thrive (Who, 2009). The main reservoirs of *V. cholerae* are people and aquatic sources such as brackish water and estuaries, often in association with copepods or other zooplankton, shellfish, and aquatic plants.Cholera infections are most commonly acquired from drinking water in which *V. cholerae* is found naturally or into which it has been introduced from the feces of an infected person. Other common vehicles include contaminated fish and shellfish, produce, or leftover cooked grains that have not been properly reheated. Transmission from person to person, even to health care workers during epidemics, is rarely documented. *V. cholerae* thrives in water ecology, particularly surface water. The primary connection between humans and pathogenic strains is through water, particularly in economically reduced areas that don't have good water purification systems (Faruque et al., 1998)

Non-pathogenic strains are also present in water ecologies. It is thought that it is the wide variety of strains of pathogenic and non-pathogenic strains that co-exist in aquatic environments that allow for so many genetic varieties. Gene transfer is fairly common amongst bacteria and recombination of different *V. cholerae* genes can lead to new virulent strains (Faruque and Nair, 2002).

Transmission is primarily by the acquisition of the pathogen through contaminated drinking water or infected food. It is a fatal intestinal disease that produces severe gastrointestinal symptoms such as vomiting, abdominal pain and severe diarrhea (Abramowicz, 2008). An estimated 3.5 million cases and over 100,000 death cases occur each yeardue to cholera around the world. The infection is often mild without symptoms but sometimes is severe. Approximately one in 20 (5%)infected persons will have severe disease characterised with watery diarrhoea, vomiting, leg cramps. In these people, rapid loss of body fluid leads to dehydration and shock (WHO, 2009).

Existing methods of medical diagnosis employed by physicians for the analysis of cholera uses subjective approach characterized by the inability to handle uncertain or vague data existing between intervals. These papers have chosen to solve this problem by employing the rich facilities of fuzzy classification embedded in union approach which is more objective in analyzing cholera.

## 2. Review of Related Literature

The approach on ground for the medical diagnosis of cholera is subjective based on the will, exposure and experience of the physician administrating the diagnosis. Therefore an objective approach is relevant in complement the effort of theses physicians.

A **Fuzzy *classifier*** is an algorithm that assigns a class label to an object, based on the object description. It is also said that the classifier *predicts* the class label (Angelov and Zhou, 2008). The object description comes in the form of a vector containing values of the features (attributes) deemed to be relevant for the classification task (Ishibuchi et al., 1995). Typically, the classifier learns to predict class labels using a training algorithm and a training data set. When a training data set is not available, a classifier can be designed from prior knowledge and expertise. Once trained, the classifier is ready for operation on unseen objects (Cordon et al, 1999).

Classification belongs to the general area of pattern recognition and machine learning (Babuska, 1998)

a. *Soft labelling.* The standard assumption in pattern recognition is that the classes are mutually exclusive. A standard classifier will assign a single *crisp* label (rain). A fuzzy classifier can assign degrees of membership (*soft* labels) in all four classes {rain, clouds, wind, sunshine}, accounting for the possibility of winds and cloudy weather throughout the day. A standard classifier can output posterior probabilities, and offer soft-labellingtoo. However, a probability of, say, 0.2 for

cloudy weather means that there is 20% chance that tomorrow will be cloudy. A probabilistic model would also assume that the four classes form a full group, i.e., snow; blizzards or thunderstorms must be subsumed by one of the existing four classes. Soft labelling is free from this assumption. A fuzzy classifier, *D* ,producing soft labels can be perceived as a function approximator $D:F \rightarrow [0,1]c$, where *F* is the feature space where the object descriptions live, and *c* is the number of classes. While tuning such a function approximator outside the classification scenario would be very difficult, fuzzy classifiers may provide a solution that is both intuitive and useful(Kuncheva, 1998)

b. *Interpretability.* Automatic classification in most challenging applications such as medical diagnosis has been sidelined due to ethical, political or legal reasons, and mostly due to the *black box* philosophy underpinning classical pattern recognition. Fuzzy classifiers are often designed to be *transparent*, i.e., steps and logic statements leading to the class prediction are traceable and comprehensible (Kuncheva 1998).
c. *Limited data, available expertise.* Examples include predicting and classification of rare diseases, oil depositions, terrorist activities, natural disasters. Fuzzy classifiers can be built using expert opinion, data or both.

***Fuzzy Rule-Based Classifiers***

The simplest fuzzy rule-based classifier is a fuzzy if-then system, similar to that used in fuzzy control. Consider a 2D example with 3 classes. A fuzzy classifier can be constructed by specifying classification rules, e.g.

IF X1 is medium and X2 is small Then Class is 1
IF X1 is Medium and X2 is large Then Class is 2
IF X1 is large and X2 is small Then Class is 2
IF X1 is Large and X2 is small Then class is 3
If X1 is small and X2 is large Then Class is 3

The two features $x1$ and $x2$ are numerical but the rules use *linguistic values*. If there are *M* possible linguistic values for each feature, and *n* features in the problem, the number of possible different if-then rules of this conjunction type (AND) is *Mn*. If the fuzzy classifier comprises of all such rules, then it turns into a simple look-up table. Unlike look-up tables, however, fuzzy classifiers can provide outputs for combinations of linguistic values that are not included as one of the rules. Each linguistic value is represented by a membership function.

## 3. Methodology

The methodology adopted in this research paper is geared toward specifying fuzzy rules utilizing fuzzy set theory application.

We utilize several symptoms for diagnosis cholera which includes: (S: Watery-diarrhea, Muscle cramps, Vomiting, Thirst, Dehydration, Cold skin, sunkeneyes, Weak pulse). Each of these criteria fall into a rule (P01-P09) and the fuzzy rules thus specifies:

a. If the Patient is exhibitingS $\leq$ 3 and each S $\geq$ (50%)THENNot Cholera Diagnosed
b. If the Patient is exhibitingS=4and each S $\geq$ (50%)THENMight be Cholera Diagnosed
c. If the Patient is exhibitingS $\geq$ 4and S $\geq$ (50%)THENCholera Diagnosed

In set theory, the union (denoted by ∪) of a collection of sets; is the set of all distinct elements in the collection. It is one of the fundamental operations through which sets can be combined and related to each other. The initial ∪ is initialized as P ∪∅ = P, for the set P. Therefore the fuzzy set rules are thus:

**P0:** P ∪∅
**P01:** {∅∪Watery-diarrhea} = Not Cholera Diagnosed.
**P02:** {∅∪Watery-diarrhea} ∪Muscle cramp = Not Cholera Diagnosed.
**P03:** {∅∪Watery-diarrhea∪Muscle cramp}∪Vomiting = Not Cholera Diagnosed.
**P04:** {∅∪Watery-diarrhea∪Muscle cramp∪ Vomiting}∪Thirst = Might be Cholera diagnosed.
**P05:** {∅∪Watery-diarrhea∪Muscle cramp∪ Vomiting∪ Thirst}∪Dehydration = Cholera Diagnosed.
**P06:** {∅∪Watery-diarrhea∪Muscle cramp∪ Vomiting∪ Thirst∪Dehydration}∪Cold skin = Cholera Diagnosed.
**P07** {∅∪Watery-diarrhea∪Muscle cramp∪ Vomiting∪ Thirst∪Dehydration∪Cold Skin}∪Sunken Eyes=Cholera Diagnosed.
**P08** {∅∪Watery-diarrhea∪Muscle cramp∪ Vomiting∪ Thirst∪Dehydration ∪Cold skin∪Sunken Eyes}∪Weak pulse= Cholera Diagnosed.

## 4. Simulation Result

The simulation Results were based on the dataset derived online pertaining to cholera patient in Nigeria.

**Table1:** Dataset for Cholera Patient

| Fuzzy Codes | Fuzzy Set | Online Values Pertaining to Fuzzy Set for cholera patient (%) | Membership Function (Scale Value 0.00 -1.0) |
|---|---|---|---|
| P01 | Watery-diarrhea | 70 | 0.70 |
| P02 | Muscle cramps | 80 | 0.80 |
| P03 | Vomiting | 90 | 0.90 |
| P04 | Thirst | 40 | 0.40 |
| P05 | Dehydration | 20 | 0.20 |
| P06 | Cold skin | 50 | 0.50 |
| P07 | Sunken eyes | 10 | 0.10 |
| P08 | Weak pulse | 50 | 0.50 |

Table 1, represents the degree of membership function for a patient with cholera symptoms, for instance, P05 in Column 4, we notice it has 0.20. In percentage, it can be represented as 20% that is 20% of the symptoms of dehydration are present in such patient. This means that the degree of membership function of P05 matches **0.20 of the fuzzy scaled ranged values for that patient.** The Fuzzy clustering graphical distribution shown Figure 1, depicts five symptoms with high degree of membership function for "Cholera Diagnosed", one criterion for "Might Cholera

Diagnosed“ and two criteria igniting “ Not Cholera Diagnosed”.

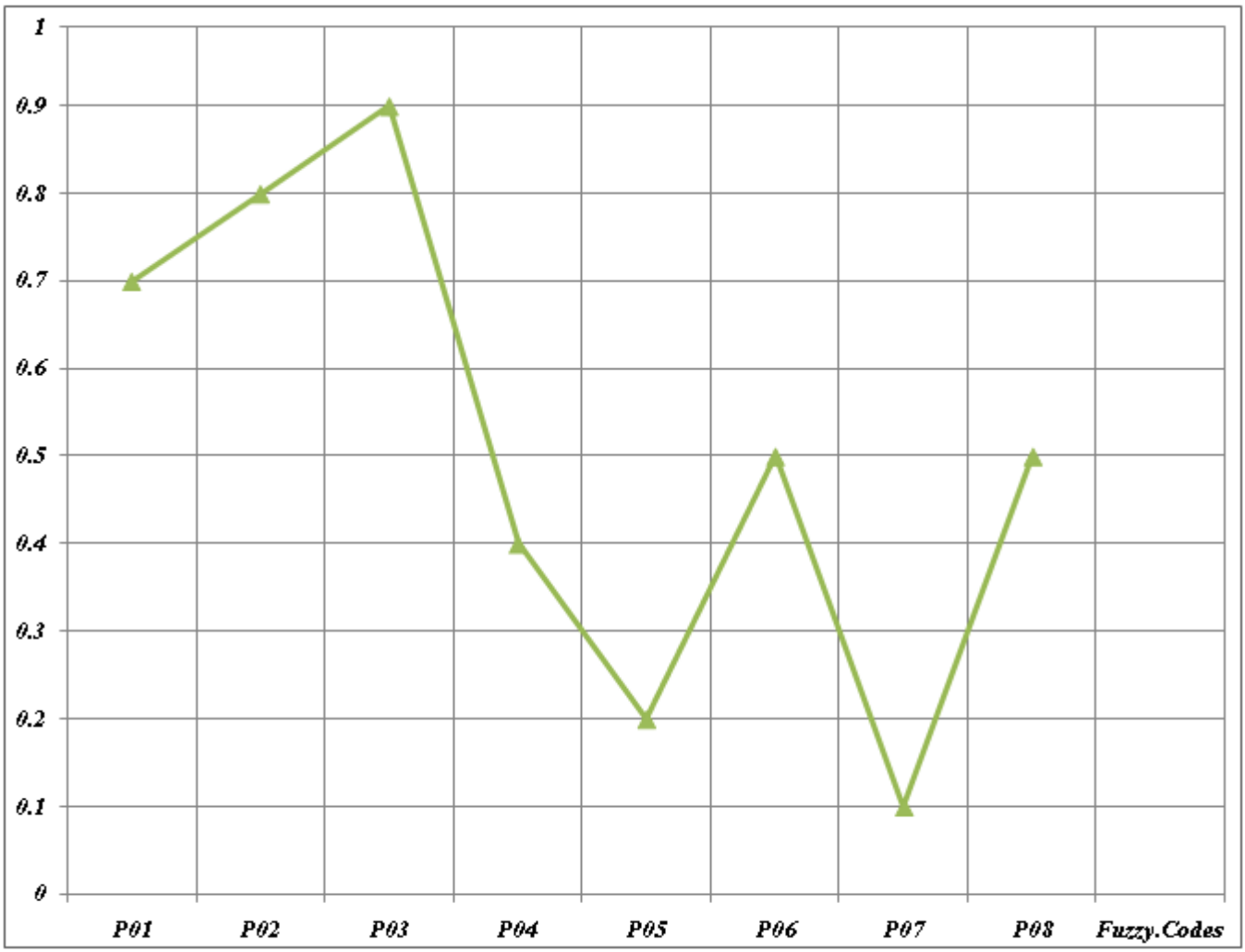


**Figure 1:** Graphical Fuzzy Distribution for Cholera

## 5. Discussion

The main focus of our approach is geared toward recognizing or diagnosing cholera in varied patient using the rich facilities of fuzzy set theory application which is more pivotal in nature, flexible and robust. Unlike previous approaches which are times consuming and quite expensive because of subject approach this approach diagnoses based on the tunes up decision variables exhibited within each region in achieving objectives results.

## 6. Conclusion

This work demonstrates the application of fuzzy set theory in the diagnosis of cholera utilizing the symptoms presented by the patient at a particular point in time. Using fuzzy classifier methodology, differential diagnosis of cholera into three major classes “Cholera diagnosed, Might be cholera diagnosed, Not cholera diagnosed” was presented. The simulated result was satisfactory having been able to tune-up our membership function in achieving our desired results.

## 7. Scope for Future Study

The diagnosis of cholera can be done using hybrid soft computing approaches to achieve a more optimized and robust result. Also, the system can be expanded to be used in the diagnosis of other tropical infections.

## References

[1] Abramowicz M. (2008), Cholera, Microsoft ® Encarta® 2009. [DVD]. Redmond, WA: Microsoft corporation, 2008.
[2] Angelov P., Zhou X. (2008), Evolving Fuzzy-Rule-based Classifiers from Data Streams, IEEE Transactions on Fuzzy Systems, ISSN 1063-6706, special issue on Evolving Fuzzy Systems, December 2008, vol. 16, No6, pp.1462-1475.
[3] Babuska R. (1998), Fuzzy Modeling for Control, Kluwer Academic Publishers, Boston, USA.
[4] Cordon O., Jesus M. J., Herrera F. (1999), A proposal on reasoning methods in fuzzy rule-based classification systems, International Journal of Approximate Reasoning, 20 (1), 1999, pp.22-45, 1999.

[5] Davis, B (2003), Filamentous phages linked to virulence of Vibrio cholera, Current Opinion in Microbiology, 6 (1): 35–42. doi: 10.1016/S1369-5274(02)00005-X.
[6] Faruque, S. M., Albert M. J.;Mekalanos, J.J. (1998), Epidemiology, genetics, and ecology of toxigenic Vibrio cholera, Microbiology and molecular biology reviews: MMBR 62 (4): 1301–14, PMID 9841673.
[7] FaruqueS.M.; Nair G.B. (2002).,Molecular ecology of toxigenic Vibrio cholera, Microbiology and immunology 46 (2): 59–66.,
[8] Ishibuchi H., Nozaki K., Yamamoto N., Tanaka H. (1995), Selecting fuzzy if-then rules for classification problems using genetic algorithms, IEEE Trans. on Fuzzy Systems, 3(3), 1995, pp.260-270.
[9] Kuncheva L.I. (2003)Fuzzy vs Non-fuzzy in combining classifiers designed by boosting, IEEE Transactions on Fuzzy Systems, 11 (6), 2003, pp. 729-741.
[10] World Health Organization (2009), Cholera in Zimbabwe: Epidemiological Bulletin Number 16 Week 13 (22-28 March 2009). March 31, 2009; WHO Zimbabwe Daily Cholera Update, 16 April 2009)

## Authors Profile

**Mrs. Margaret Dumebi Okpor** has B.Sc and M.Sc degrees in Computer Science from University of Benin, Edo State, Nigeria. She has been involved in teaching Computer science courses for over 10 years in the Delta State Polytechnic, Ozoro, Delta State, Nigeria. She was Head of Department in the Department of Computer Science in the same institution. Her research interests include machine learning, soft computing, computational linguistics and evolutional computing. She has published several papers in local and international journals. She loves Reading Novels, Singing and doing research.