Abstract—In this paper, we present a hybrid intelligent system based on Neutrosophic Logic (NL). In conjunction with Genetic Algorithm (GA) for classification, the neutrosophic logic is adapted for representing different forms of knowledge. GA is used to refine the generated neutrosophic rules. The performance of the proposed system is tested on three real-world databases Iris, Wine, and Wisconsin Diagnostic Breast Cancer (WDBC). In a series of experiments, we compare the performance of the proposed genetic neutrosophic rule-based classification system with that of the neutrosophic rule-based classification system. The performance of both classifiers is measured for the three real-world data sets. We have reached an average accuracy 98.39% in genetic neutrosophic against 94.78% for the corresponding neutrosophic.

Keywords—Genetic Neutrosophic Rule Based System, Neutrosophic Rule Based System, Neutrosophic set, Neutrosophic Logic, Neutrosophic Classification System, Rule Based System.

I. INTRODUCTION

Neutrosophic Rule-based Classification System (NRCS) proposed in [17] as rule based classification system where neutrosophic logic (NL) is used to describe various forms of knowledge, as well as, for describing the relationships and interactions that appear between its variables. The two main tasks in the NRCS design process are: (1) the design of the inference mechanism, and (2) the generation of the rule set (Knowledge Base (KB) or Rule Base (RB)).

One of the major drawbacks of NRCS is that they are not able to learn, instead require the Knowledge Base to be derived from expert knowledge. The key point is to employ an evolutionary learning process to automate the NRCS design.

Computational Intelligence techniques such as artificial neural networks, fuzzy logic, neutrosophic logic, and genetic algorithms (GAs) deal with more complex problems which are difficult to solve with traditional methods. Each component of computational intelligence is complementary to each other. Using combinations of several technologies such as fuzzy-neural, genetic-fuzzy systems will generally get better solutions.

We will use genetic algorithms as a machine learning tool for designing rule-based classification systems. Genetic algorithms for machine learning are defined as genetics-based machine learning (GBML) algorithms. GBML algorithms are usually categorized to: Michigan approach and Pittsburgh approach [6]. In this paper we propose a new genetic neutrosophic rule-based classification system (GNRCS). The NRCS is one sort of generalization of fuzzy rule-based system. It deals with "IF-THEN" rules whose antecedents and consequents are composed of neutrosophic logic statements, instead of fuzzy logic ones.

The GNRCS will improve NRCS by using genetic algorithm to produce the best rules for classifications. The proposed system may also be successfully applied to a wide range of problems which uncertainty, incompleteness, inconsistence, and vagueness that emerge in different ways. A genetic algorithm is able to incorporate with other techniques to produce more efficient solution to real-world problems.

One of research areas is to combine genetic algorithm to build rule based systems especially to fuzzy logic since fuzzy rule-based system is the most famous type of rule based systems and most used among. Ishibashi, R. and Nascimento, C.L. [13] aim to get better classification by showing a system able to automatically generate fuzzy rule sets with less human participation.

They used genetic algorithm to trained fuzzy rule-based system to perform a classification and adapt the parameters of the membership functions after created decision tree to generate a fuzzy rule-based system.

This execution gives as output an Genetic Fuzzy Rule Based System assisted by Decision Trees able to automatically generate the fuzzy rules and calibrate the membership functions.

They applied their proposed method to an appendicitis data set with 106 instances, 7 normalized real-valued inputs and one binary output.

Jose Antonio Sanz et al [11] present IVTURS, which is a new linguistic fuzzy rule-based classification method based on a new completely interval-valued fuzzy reasoning method. This inference process uses interval-valued restricted equivalence functions to increase the relevance of the rules. They using the recent fuzzy rule learning algorithm known as FARC-HD [9] to learn the rule base in order to generation of an initial Interval-Valued Fuzzy Rule-Based Classification System (IV-FRBCS).

After that, for each variable of the problem They model its
linguistic labels with interval-valued fuzzy sets (IVFSs) and initialize the interval-valued restricted equivalence functions (IV-REF).

Using the proposed synergy between the tuning of the equivalence and rule selection in the optimization step. They used genetic algorithm in their proposed method to tune the values of the parameters and to perform a rule selection process in which we obtain a compact and cooperative fuzzy rule set.

J Casillas et al [10] proposed a method to deal with the difficulty comes from the exponential growth of the fuzzy rule search space with the increase in the number of features considered in the learning process. They presented a genetic feature selection process that can be integrated in a multistage genetic learning method to obtain, in a more efficient way, Fuzzy Rule based Classification Systems (FRBCSs) composed of a set of comprehensible fuzzy rules with high-classification ability.

The proposed process fixes, a priori, the number of selected features, and therefore, the size of the search space of candidate fuzzy rules.

O. Cordon et al in [12] proposed a new method by using a Genetic Algorithm and considering a simple generation method to derive the Rule Base to automatically learn the Knowledge Base of a Fuzzy Rule-Based System by finding an appropriate Data Base.

Their genetic process using a non-linear scaling function to adapt the fuzzy partition contexts to learns all the components of the Data Base (number of labels, working ranges and membership function shapes for each linguistic variable).

Genetic Neutrosophic Rule-Based System (GNRCS) aims to the genetic learning process at designing or optimizing the KB.

Consequently, the proposed method is a design method for NRCSs which incorporates evolutionary techniques to achieve the automatic generation or modification of the entire or part of the KB.

The rest of this paper is organized as follows. Section II presents basics technology used in our system. In section III the proposed genetic neutrosophic classifier is presented. Section 4 shows the experimental results obtained. Finally, section 5 presents conclusions and discusses future work.

II. BASICS

In this section we present a basics technologies used in our system such as Neutrosophic Logic and Genetic Algorithm.

A. Neutrosophic Logic

Neutrosophic, introduced by Smarandache [19], can handle incomplete information as well as inconsistent information without danger of trivialization [5]. Neutrosophy is a relatively new branch of philosophy, with ancient roots, handling neutralities, as well as their interactions with different ideational spectra. This theory considers every notion or idea <X> together with its opposite or negation <AntiX> and the spectrum of neutralities <NeutX> which neither <X> nor <AntiX>. The <NeutX> and <AntiX> can considered together as <Non X>. According to this theory every idea <X> tends to be balanced and neutralized by <Anti X> and <Non X> ideas - as a state of equilibrium [5].

Neutrosophic logic was developed to represent mathematical model of ambiguity, vagueness, uncertainty, imprecision, inconsistency, incompleteness, contradiction and redundancy [19].

Neutrosophic logic is a logic in which each proposition is estimated to have the percentage of truth in a subset T, the percentage of indeterminacy in a subset I, and the percentage of falsity in a subset F, where T, I, and F are standard or non-standard real subsets of \([-0,1]\) where \([-0,1]\) is non-standard unit interval [1] [15].

with
\[
sup T = t_{sup}, \quad inf T = t_{inf}
\]
\[
sup I = i_{sup}, \quad inf I = i_{inf}
\]
\[
sup F = f_{sup}, \quad inf F = f_{inf}
\]

T, I, and F are called neutrosophic components, representing respectively the truth, indeterminacy, and falsehood values referring to neutrosophy, neutrosophic set, neutrosophic logic, neutrosophic statistics, neutrosophic probability [3].

In real world applications, it is easier to use standard real interval [0,1] for T, I and F instead of non-standard unit interval \([-0,1]\) [1]. All the factors stated in neutrosophic logic are very integral to human thinking. It is very rare that we tend to conclude/judge in definite environments. Impression of human systems could be due to the imperfection of knowledge that human receives (observation) from the external world [1].

The fundamental concepts of neutrosophic set, introduced by Smarandache in [19] [18], Salama et al. in [16] [7] [8] [2], provides a natural foundation for treating mathematically the neutrosophic phenomena which exists pervasively in our real world and for building new branches of neutrosophic mathematics.

Now we explain a formal Definition, Neutrosophic set [5]. Let X be a space of objects, with a generic element x in X. A neutrosophic set A in X is characterized by three membership functions \(T_A, I_A, F_A\) which are represent respectively a truth-membership function, an indeterminacy-membership function, and a falsity membership function. Where \(T_A(x), I_A(x)\) and \(F_A(x)\) are real standard or non-standard subsets of non-standard interval \([-0,1]\).

\[
T_A : X \rightarrow [-0,1]
\]
\[
I_A : X \rightarrow [-0,1]
\]
\[
F_A : X \rightarrow [-0,1]
\]

There is no any restriction on the sum of superior \(T_A(x), I_A(x)\) and \(F_A(x)\), so \(0 \leq sup T_A(x) + sup I_A(x) + sup F_A(x) \leq 3\).

B. Genetic Algorithm

Genetic Algorithms (GAs) are general-purpose search algorithms that use principles inspired by natural population genetics to evolve solutions to problems [4].

In GA, First generate an initial population of solutions usually random. In order to achieve the best solution GA will operates on these population by applying genetic operators such as, crossover and mutation.
The algorithm is stopped when the population converges toward the optimal solution. The simple genetic algorithm is as follows:

- **[start]** Genetic generate random population of n chromosomes.
- **[Fitness]** Evaluate the fitness function \( f(x) \) of each chromosome \( x \) in the current population.
- **[New population]** Create a new population by repeating the following operations until the new population is complete.
  - **[selection]** Select two parent from a current population according to their fitness value, the better the fitness is, the bigger chance to get selected.
  - **[crossover]** Crossover the parents to form new offspring. If no crossover was performed, offspring is the exact copy of parents.
  - **[Mutation]** Mutate new offspring at each position in chromosome.
  - **[Accepting]** Place new offspring in the new population.
- **[Replace]** Use new generated population for a further sum of the algorithm.
- **[Test]** If the end condition is satisfied: stop and return the best solution in the current population.
- **[Loop]** Go to step 2 for fitness evaluation.

### III. THE PROPOSED APPROACH FOR GENETIC NEUTROSOFPIC CLASSIFICATION SYSTEM

One of successfully systems used to model human problem-solving activity and adaptive behavior is rule-based systems which use IF-THEN rules to represent human knowledge.[4].

Neutrosophic Rule-based Classification System (NRCS) which is a rule based system where neutrosophic logic is used as a tool for representing different forms of knowledge about the problem at hand, as well as for modeling the interactions and relationships that exist between its variables[17].

The generic structure of a NRCS shown in Fig. 1[17]. We build a hybrid system between Genetic Algorithm and Neutrosophic Rule Based Classification System that target the genetic learning process at designing or optimizing the knowledge base as Fig. 2.

Consequently, the proposed Genetic neutrosophic Rule Based Classification System (GNRCS) is a design method for NRCSs which incorporates evolutionary techniques to achieve the automatic generation or modification of the entire or part of the knowledge base.

The automatic definition of NRCS can be seen as an optimization or search problem. GA is a well known and widely used global search technique with the ability to explore large search space for suitable solutions.

In our approach, neutrosophic if-then rules are generated from data set as initial population, Each rule is represented by a string (each string divided to parts each of them contains a triple values) and as in the Michigan approach each string (i.e. each rule) handled as an individual Thus a fitness value is calculated and assigned to a single rule. A population of individuals corresponds to a single rule set. The rule sets are not directly optimized therefore the performance of a rule set is not utilized in the evolution of rules[6].

GNRCS approach contains six phases: Extract Information phase, Neutrosop hic phase, Generate set of rules phase, Genetic Based Machine Learning phase, Classification phase, and Accuracy Phase. These six phases are described in detail in the rest of this section.

#### A. Extract Information phase

In this phase, as in NRCS we extract some significant features from the data set, such as number of attributes, number of classes, and the maximum, the minimum for each attribute which will be needed in section B.

#### B. Neutrosop hic phase

In this phase, as in NRCS we apply three membership functions \( T_A, I_A, F_A \) on the data set to represent each value in the three neutrosophic components \(< T,I,F >\). Where \( T, I, \text{and } F \) are restricted to be subsets of standard real interval \([0,1]\) instead of non-standard unit interval \([-0,1+\]).

#### C. Generate set of rule phase

In this phase, for initial population we generate the training rules, and for testing generate both testing rules, and the exact rules. In neutrosophic rule based classification system, each variable in each rule has three components which represent respectively degree of truth, degree of indeterminacy, and degree of falsity as example in Fig. 4[17]. Then, after checking for redundancy the rules are added to the training rules. Therefore, in most cases the number of training rules generated...
is less than the number of training instances and moreover, number of generated training rules in GNRCs is less than that in NRCS as in Table I. In case of generating the testing rules we remove the class labels instead of checking for redundancy [17].

<table>
<thead>
<tr>
<th>Rule</th>
<th>Iris Data Set</th>
<th>Wine Data Set</th>
<th>Wdbc Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRCS</td>
<td>51</td>
<td>89</td>
<td>269</td>
</tr>
<tr>
<td>GNRC</td>
<td>45</td>
<td>85</td>
<td>264</td>
</tr>
</tbody>
</table>

D. Genetic Based Machine Learning phase

Genetics-based machine learning (GBML) algorithms are genetic algorithms for machine learning which are usually classified into two categories: Michigan approach and Pittsburgh approach[6]. The comparison between these approaches are summarized in Table II.

<table>
<thead>
<tr>
<th>Rule</th>
<th>Michigan approach</th>
<th>Pittsburgh approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>A single rule</td>
<td>A single rule set</td>
</tr>
<tr>
<td>Population</td>
<td>Yes</td>
<td>Multiple rule set</td>
</tr>
<tr>
<td>Evaluation of each rule</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Fitness calculation</td>
<td>For each rule</td>
<td>Good rules sets</td>
</tr>
<tr>
<td>Selection</td>
<td>Good rules</td>
<td>For each rule set</td>
</tr>
<tr>
<td>Crossover</td>
<td>Between rules</td>
<td>Between rules sets</td>
</tr>
</tbody>
</table>

In our proposed GNRCs we apply Michigan approach to refine a linguistic rules. Michigan-Style Neutrosophic GBML Algorithm Step 1: Parameter Specification. Specify the number of linguistic rules N rule, the number of replaced rules N replace, the crossover probability P_c, the mutation probability P_m, and the stopping condition. Step 2: Initialization. generates a pre-specified number of linguistic rules (say, N rule linguistic rules) from data set as an initial population.

E. Classification Phase

For Initial Population as in NRCS, After generating training and testing rules, we construct the testing rule matrix after removing class labels. Then, to get the class label we compare each rule in the testing set with all rules in the training set. If we do not find any one of the training rules intersected with the current testing rule -at least in half of the attributes-, from the exact rule set we get this rule completed and add it to the training rules set, then for each class calculate the total accuracy, precision, recall, and specificity [17], then we begin with these training rules as initial population for genetic algorithm and begin evaluate fitness function then select two parent for crossover operation to generate new linguistic rules from these rule. Then, replace the worst two rules from our population and check classification accuracy again and so on until get the best.

F. Accuracy Phase

In this phase, in each generation as in NRCS we compare testing matrix, obtained from phase E, with the exact matrix, obtained from phase C. Then, we compute the confusion matrix; which is a matrix that is often used to describe the performance of a classifier on a set of testing data for which the true values are known[17]. Finally, we calculate the total accuracy, precision, recall, and specificity for each class. AS in Fig. 5., Fig. 6., Fig. 7., and Fig. 8., we compare our results with the results obtained from the corresponding neutrosophic rule based system in [17].

IV. EXPERIMENTAL RESULTS

A. Dataset

Since we have limited accessibility to the private datasets, we made our test on public open datasets until we are able to
Fig. 4: Accuracy of classification for Iris, Wine, and Wdbc data sets in Genetic Neutrosophic and Neutrosophic

Fig. 5: Precision, recall, and specificity for Iris data set in Genetic Neutrosophic and Neutrosophic

Fig. 6: Precision, recall, and specificity for Wine data set in Genetic Neutrosophic and Neutrosophic

Fig. 7: Precision, recall, and specificity for Wdbc data set in Genetic Neutrosophic and Neutrosophic

B. Evaluation Results

The proposed Genetic neutrosophic rule based system improves the neutrosophic rule based system and gives better results. A comparison, in Fig. 5, shows that for all the data sets used our proposed GNRCS is more accurate in classification. Also it reduces the computational and complexity by removing the overlaps between sets and classes. Which was a result of introducing the indeterminacy term in neutrosophic logic, then by hybrid genetic algorithm get the best rules for classification from the initial set of rules comes from data sets.
V. CONCLUSIONS AND FUTURE WORK

The proposed Genetic neutrosophic rule based classification system improves the neutrosophic rule based system and gives better results. Better rules were obtained in shorter time. Moreover, the results shows that GNRC is more robust in classification.

In future work, we aim to build a hybrid system between a neutrosophic rule based system and rough set which not only applies to analyze the attributes of quality but also to study the attributes of quantity.

REFERENCES


