Artificial Immune System for Attribute Weighted Naive Bayes Classification

Jia Wu, Zhihua Cai, Sanyou Zeng, and Xingquan Zhu

Abstract—Naive Bayes (NB) is a popularly used classification method. One potential weakness of NB is the strong conditional independence assumption between attributes, which may deteriorate the classification accuracy. In this paper, we propose a new Artificial Immune System based Weighted Naive Bayes (AISWNB) classifier. AISWNB uses immunity theory in artificial immune systems to find optimal weight values for each attribute. The adjusted weight values will alleviate the conditional independence assumption and help calculate the conditional probability in an accurate way. Because AISWNB uses artificial immune system search mechanism to find optimal weights, it does not need to know the importance of individual attributes nor the relevance among attributes. As a result, it can obtain optimal weight value for each attribute during the learning process. Experiments and comparisons on 36 benchmark data sets demonstrate that AISWNB outperforms other state-of-the-art attribute weighted NB algorithms.

I. INTRODUCTION

NAIVE BAYES (NB) classification is a Bayes Theorem oriented learning model particularly useful for handling high dimensional data [1]. Different from typical Bayesian models which consider the dependency (or conditional dependence) between some random variables for classification, NB relaxes the restriction of the dependency structures between attributes by simply assuming that attributes are conditionally independent, given the class label. In reality, the attributes in real-world applications often correlate to each other, so the conditional independence assumption between attributes made by Naive Bayes [2] may impair its classification performance [3]. In order to relax NB’s attribute independence assumption and simultaneously retain its simplicity and efficiency, researchers have proposed some approaches to further improve its performance [4]. One major approach to mitigate NB’s primary weakness (the attributes independence assumption) is to assign a weight value to each individual attribute. Because weight values enforce attributes to play different roles in classification, weighted naive Bayes (WNB) helps relax the conditional independence assumption and makes NB more efficient for real-world applications [5].

In order to discover proper weight values for WNB to improve its classification accuracy, researchers have proposed many useful methods to evaluate the importance of attributes. Examples include Gain Ratio [6], CFS (Correlation-based Feature Selection) attribute selection algorithm [7], Mutual Information [8], and ReliefF attribute ranking algorithm [9]. Zhang & Sheng [6] investigated the gain ratio-based weighting scheme and several wrapper-based methods for finding attribute weights in order to improve the Area Under Curve (AUC) performance¹ for Naive Bayes. Hall [5] proposed a new attribute weighting method, based on the degree of dependence on other attributes, to improve the AUC value.

The above attribute weight setting methods for WNB have achieved good performance to solve domain specific problems. However, for all these methods, they employ some external criteria, such as Gain Ratio, to determine the weight values of the attributes. So the attribute weights are determined without taking the NB objective function into consideration. In order to address the problem, we propose in this paper a new method, which automatically calculates the optimal attribute weight values for WNB, by directly working on NB’s objective function. A new method in computational intelligence namely artificial immune system (AIS), which is inspired by the immune systems and was developed to accomplish a wide range of learning tasks [10], is used to assign proper weight values for NB classification.

In this paper, we first analyze the influence of existing weighted approaches for WNB. Then we propose to use Artificial Immune System (AIS) mechanism to adaptively determine attribute weight for Naive Bayes classification. Our method uses AIS principles to design an automated search strategy to find optimal attribute weight for each data set. The unique immune system computation processes, including initialization, clone, mutation, and selection, ensure that our method can adjust itself to the data without any explicit specification of functional or distributional form for the underlying model. Experiments and comparisons, on 36 UCI benchmark data sets [11] demonstrate that the proposed Artificial Immune System based Weighted Naive Bayes (AISWNB) classifier can successfully find optimal weight combinations for the underlying learning tasks, and AISWNB consistently outperforms other state-of-the-art NB algorithms.

1The area under the ROC curve (AUC) is a common metric used to compare algorithms by taking their accuracy with respect to different classes into consideration.
In contrast to the conventional statistical probabilistic evaluation in NB, AISWNB is a self-learning algorithm by utilizing the immunological properties, such as memory property and clonal selection. The niche and advantages of AISWNB can be understood from the following three aspects: AISWNB is a data-driven self-adaptive method because it does not require explicit specification of functional or distributional form for the underlying model. AISWNB is a nonlinear model and is flexible in modeling complex real-world relationships. AISWNB inherits the memory property of human immune systems and can recognize the same or similar antigen quickly at different times.

The rest of the paper is organized as follows. In Section II, we review related work on attribute weighted Naive Bayes classification and artificial immune systems. In Section III, we propose our new artificial immune systems based weighting scheme for Naive Bayes classification (AISWNB). In Section IV, we describe the experimental conditions, methods, and results in details. Section V concludes the paper and outlines several directions for future study.

II. RELATED WORK

A. Attribute Weighted Methods

In real-world applications, attributes may play different roles for classification. Therefore, assigning different weight values to attributes can potentially help reach higher classification accuracies. During the whole process, the way to learn the attribute weights is the most important part.

Mutual Information: According to the probability theory and information theory, the mutual information of two random variables provides a quantity measure to evaluate the mutual dependence of two variables. Mutual information has been widely used for measuring the importance between attributes and the class variable in classification. Jiang & Zhang [8] applied this method to improve the accuracy performance for AODE (Averaged One-Dependence Estimators). Zhang, Jiang & Su [1] proposed a Hidden Naive Bayes (HNB) classifier, which uses the mutual information attribute weighted method to weight one-dependence estimators. Han [12] proposed a new weighted method to propose a new algorithm called AWKNN (Attribute Weighted K-Nearest-Neighbor). In this paper, we propose to apply mutual information to calculate the weight between each attribute and the class attribute for WNB.

Gain Ratio: Zhang & Sheng [6] argued that an attribute with a higher gain ratio value [13] deserves higher weight in WNB. In their studies, they proposed a gain ratio weighted method that calculates the weight of an attribute from a data set.

Correlation-based Feature Selection: Correlation-based Feature Selection (CFS) uses a correlation-based heuristic evaluation function as the attribute quality measure [7]. The core of CFS algorithm is the heuristic process that evaluates the worth or “merit” of a subset of features. Hall [5] employed this method to evaluate the importance of attributes according to the heuristic “merit” value.

Relief-F: Relief is a feature selection method based on attribute estimation [14]. Relief assigns a grade of relevance to each feature by examining the change of the feature values with respect to instances within the same class (i.e. the nearest hit) and instances between classes (i.e. the nearest miss). If a feature’s values remain relatively stable for instances within the same class, the feature will receive a higher weight value. The original Relief only handles Binary classification problems. Its extension, Relief-F, can be applied for multi-class classifications [15]. Besides, Tucker [16] applied Relief-F attribute weighted approach to deal with top-down product, which is an engineering optimization problem.

Decision Tree-Based Attribute Weighting: Hall [5] proposed a new weighting method that assigns small weight values to the attributes, which have strong dependencies on other attributes. In order to estimate each attribute’s dependence on other attributes, an unpruned decision tree is constructed from the training instances with a minimum depth indicating the depth for testing the tree. Attributes that do not appear in the tree receive a weight of zero.

B. Artificial Immune Systems

The Artificial Immune System (AIS) theory consists of three major components, including representation, recognition, and clone selection. The representation, known as shape-shape problem, focuses on how to model antibodies and antigens. When the immune system is attacked by antigen, antibodies try to neutralize the infection by binding to the antigen through the recognition process. Binding strength, also regarded as affinity, is used as a threshold for the immune system to respond to the antigen. The clone selection is corresponding to an affinity maturation process, which means immune individuals with low affinity will gradually increase during clone and mutation process. At the same time, some immune individuals will polarize into memory individuals.

AIS has been used in various areas of research including pattern recognition [17], clustering [18], optimization [19] and Remote Sensing [20]. However, few applications have been reported in Bayesian classification. In this paper, we propose a new attribute weighted method based on AIS, which has high classification accuracy performance for weighted NB (a benchmark classifier in Bayesian).

III. ARTIFICIAL IMMUNE SYSTEMS FOR WEIGHTED NAIVE BAYES CLASSIFICATION

A. Problem Definition

Given a training set $D = \{x_1, \cdots, x_N\}$ with $N$ instances, each of which contains $n$ attribute values and a class label. We use $x_i = \{x_{i,1}, \cdots, x_{i,j}, \cdots, x_{i,n}, y_i\}$ to denote the $i$th instance in the data set $D$, with $x_{i,j}$ denoting the $j$th attribute value and $y_i$ denoting the class label of the instance. The class space $Y = \{c_1, \cdots, c_k, \cdots, c_L\}$ denotes the set of labels that each instance belongs to and $c_k$ denotes the $k$th label of the class space. For ease of understanding, we also
use \((x_i, y_i)\) as a shorthand to represent an instance and its class label, and use \(x_i\) as a shorthand of \(X_i\). We also use \(a_j\) as a shorthand to represent attribute \(j\). For an instance \((x_i, y_i)\) in the training set \(D\), its class label satisfies \(y_i \in \mathcal{Y}\), whereas a test instance \(x_t\) only contains attribute values and its class label \(y_t\) needs to be predicted by the WNB model, which is formally defined as

\[
c(x_i) = \arg \max_{c_k \in \mathcal{Y}} P(c_k) \prod_{j=1}^{n} P(x_{i,j} | c_k \wedge w_j)
\]

(1)

In Eq. (1), \(P(c_k)\) denotes the probability of class \(c_k\) in the whole training set. \(P(x_{i,j} | c_k)\) denotes the joint distribution of \(x_{i,j}\) conditioned by the the given class \(c_k\). \(w_j\) denotes the weight of \(j\)th attribute.

In this paper, we focus on the calculation of the conditional probability \(P(x_{i,j} | c_k \wedge w_j)\) by using optimal attribute weight value \(w_j\). While all existing attribute weighting approaches define the weight without considering the uniqueness of the underlying training data, we intend to solve the optimal \(w\) value selection problem as an optimization process. Assume that the calculation of each conditional probability value \(P(x_{i,j} | c_k \wedge w_j)\) has an optimal \(w_j\) value, there are \(n\) \(w\) vectors needed to finish the classification process. As a result, the WNB classification can be translated to an optimization problem as follows.

\[
c(x_i) = \arg \max_{c_k \in \mathcal{Y}} P(c_k) \prod_{j=1}^{n} P(x_{i,j} | c_k \wedge w_j) \quad \text{s.t.} \quad 0 < w_j \leq 1
\]

(2)

B. AISWNB

In this paper, we propose to use AIS to learn attribute weight values for NB classification. In our solution, antigens in AISWNB are simulated as feature attribute vectors which are presented to the system during the training and the testing process. Antibodies as candidate, presented by attribute weight vector \(w\) which has good affinity, will experience a form of clonal expansion after being presented with input data sets (analogous to antigens). When antibodies are cloned they will undergo a mutation process, in which a test instance \(x_t\) only contains attribute values and its class label \(y_t\) needs to be predicted by the WNB model, which is formally defined as

\[
c(x_t) = \arg \max_{c_k \in \mathcal{Y}} P(c_k) \prod_{j=1}^{n} P(x_{t,j} | c_k \wedge w_j)
\]

(1)

In Eq. (1), \(P(c_k)\) denotes the probability of class \(c_k\) in the whole training set. \(P(x_{t,j} | c_k)\) denotes the joint distribution of \(x_{t,j}\) conditioned by the the given class \(c_k\). \(w_j\) denotes the weight of \(j\)th attribute.

In this paper, we focus on the calculation of the conditional probability \(P(x_{t,j} | c_k \wedge w_j)\) by using optimal attribute weight value \(w_j\). While all existing attribute weighting approaches define the weight without considering the uniqueness of the underlying training data, we intend to solve the optimal \(w\) value selection problem as an optimization process. Assume that the calculation of each conditional probability value \(P(x_{t,j} | c_k \wedge w_j)\) has an optimal \(w_j\) value, there are \(n\) \(w\) vectors needed to finish the classification process. As a result, the WNB classification can be translated to an optimization problem as follows.

\[
c(x_t) = \arg \max_{c_k \in \mathcal{Y}} P(c_k) \prod_{j=1}^{n} P(x_{t,j} | c_k \wedge w_j) \quad \text{s.t.} \quad 0 < w_j \leq 1
\]

(2)

Algorithm 1 AISWNB (Weighted NB by AIS)

Input:
- Clone Factor \(c\); Threshold \(T\);
- Maximum Evolution Generation \(MaxGen\);
- Antibody Population \(W\);
- Antigen Population \(D^a\); Test affinity set \(D^b\);

Output:
The target class label \(c(x_t)\) of test instance \(x_t\):
1: \(W\) ← \(w_{i,j}\) value of \(w_i\) for each individual is set a random number distributed between \([0, 1]\).
2: \(t \leq MaxGen\) and \(f[\overline{w^{t+1}}] - f[\overline{w^t}] \leq T\) do
3: \(f[\overline{w^t}]) ← \) Apply antigen population \(D^a\), test affinity set \(D^b\) to antibody \((\overline{w}^t)\), and calculate the affinity of \((\overline{w})^t\).
4: \(W^t \leftarrow \) Apply the sequence of each \(f[\overline{w}^t]\) to the whole antibody population \(W^a\) and find the \(\overline{w}^t\) with the best affinity.
5: \((V^t) \leftarrow \) Select the temporary antibodies set with the lowest affinity with clone factor \(c\).
6: \((V^t) \leftarrow \) Clone \(\overline{w}^t\) with clone factor \(c\) and obtain clone antibody set.
7: \(W^a \leftarrow [W^a - (V^t)] \cup (V^t);\)
8: for all each \(w^t_i\) in \(W^a\) do
9: \(v^t_i+1 \leftarrow \) Apply \(w^t_i\) and a normally distributed random variable \(N(0,1)\) to \(w^t_i\) and obtain the mutation individual.
10: \(w^t_i+1 \leftarrow \) Apply \(v^t_i\) to \(w^t_i\) and obtain the new individual in \(t+1\)th generation.
11: end for
12: end while
13: \(c(x_t) \leftarrow \) Apply \(w_i\) to instance \(x_t\) to predict the underlying class label .

AISWNB is achieved through the following two major steps: (1) Use AIS algorithm to train models from selected instances, with the purpose of achieving the adaptive weight combination with global optimum; and (2) The test instances are classified by the AISWNB classifiers by using the learned attribute weight values. The detailed process is described as follows:

1) Initialization: For individuals in \(W = \{w_1, \cdots, w_L\}\) with its population size \(L\), we should make sure that every individual \(w_i = \{w_{i,1}, \cdots, w_{i,j}, \cdots, w_{i,n}\}\) in antibody population is generated through certain random mechanisms. So we set \(w_{i,j}\) value of \(w_i\) for each individual as a uniformly distributed random number within range \([0, 1]\). Besides, in the experiment 80% of train instances \(D\) are used as the antigens set \(D^a\) to learn the \(w_i\) and the remaining instances are used as the test set \(D^b\), with the population size \(L\) being set to 50.

2) Evaluation of AISWNB:
- Calculation of affinity function: The affinity of the \(i\)th individual of the \(t\)th generation \(\overline{w}^t_i\) is the classification accuracy that is obtained by AISWNB using the \(\overline{w}^t_i\) to carry out the probability estimation. Calculation of affinity function can be described as

\[
f[\overline{w}^t_i] = \frac{1}{N_a} \sum_{j=1}^{N_a} [c(x^a_j), y^a_j]
\]

(3)

800
AISWNB algorithm adopts a greedy search strategy. Only if the affinity of $w_{t+1}^c$ is better than that of the target individual $w_t$, it is chosen as the offspring. Otherwise, the individual $w_t$ is maintained in the $t+1$th generation. The system chooses the individual $w_{t+1}$ with the best affinity performance in $t+1$th generation as the new memory antibody.

An unabridged evolutionary process for the population includes Evaluation and Update, which continuously repeats until (1) the algorithm surpasses the pre-set maximum number MaxGen, or (2) the same result is obtained for a number (e.g. T) of consecutive iterations. After obtaining the best individual $w_c$, corresponding to the obtained attribute weight values, we use the weight values to build a WNB classifier to classify test data.

IV. Experiments

A. Experimental Conditions

We implement the proposed method using WEKA [21] data mining tool and validate its performance on 36 benchmark data sets from UCI data repository. The data characteristics are described in Table I. Because Naive Bayes classifiers are designed for categorical attributes, in our experiments, we first replace all missing attribute values using the unsupervised attribute filter ReplaceMissingValues in WEKA. Then, we apply unsupervised filter Discretize in WEKA to discretize numeric attributes into nominal attributes.

![AISWNB classification system](image)

**Fig. 1.** AISWNB classification system.

where, $c(x_i^t)$ is the classification result of the $i$th instance in test data set $D^t$ with $N_t$ instances, using the AISWNB classifier based on individual $w_t^i$, $y^t_i$ is the actual class value of the $i$th instance. $\delta(c(x_i^t), y^t_i)$ is one if $c(x_i^t) = y^t_i$ and zero otherwise.

- **Antibody Selection**: We sort the individuals in the initial antibody population according to the affinity of each individual, and choose the individual $w^c_t$ with the best affinity performance in $t$th generation as the memory antibody.

- **Antibody Clone**: To ensure that the population size of every generation is fixed, the best individual $w^c_t$ will be cloned under the clone factor $c$. After that, we use the clone set to replace the individuals with low affinity according to the same rate $c$.

- **Antibody Mutation**: Using the mutate operation to treat the individuals in $t$th generation $\forall w^t_i$. It means that we get the middle generation composed with the new variation individuals from the parent generation. For any individual $w^t_i$ from the $t$th generation, the new variation individual $w^{t+1}_i$ can be generated as follows:

$$w^{t+1}_i = w^t_i + F \ast N(0,1) \ast (w^c_t - w^t_i) \quad (4)$$

Among them, N(0,1) is a normally distributed random variable within the range [0,1]. $F$, as the variation factor during the process of evolution, can be adaptively obtained according to the different clones [20].

$$F = 1 - f[w^t_i] \quad (5)$$

where $f[w^t_i]$ denotes the affinity of the $i$th individual of the $t$th generation.

3) **Update of AISWNB**: To determine whether the variation individual $v^{t+1}_i$ can replace the target individual vector $w^t_i$ to be the new individual $w^{t+1}_i$ in the $t+1$th generation,
B. Baseline Methods

We continue by introducing baseline algorithms and their abbreviations in our experiments.

1. **NB** : The standard Naive Bayes classifier with attribute independence assumption [2].
2. **CFSWNB** : Attribute weighted Naive Bayes based on correlation-based feature selection [7].
3. **GRWNB** : Attribute weighted Naive Bayes based on gain ratio [6].
4. **MIWNB** : Attribute weighted Naive Bayes using mutual information weighted method [8].
5. **ReFWNB** : Attribute weighted Naive Bayes using a feature selection method based on attribute estimation [9].
6. **TreeWNB** : Attribute weighted Naive Bayes with the weighting method according to the degree to which they depend on the values of other attributes [5].
7. **SBC** : A bagged decision-tree based attribute selection filter for naive Bayes [22].
8. **RMWNB** : Attribute weighted Naive Bayes with the attribute weights selected randomly from \([0, 1]\).
9. **AISWNB** : Attribute weighted Naive Bayes based on artificial immune system in which we use the AIS to calculate the weight for each attribute adaptively.

The initial important task is to analyze the performance in terms of classification accuracy between the generic NB and WNB with the related attribute weighted methods in literature. We compare the effect of NB with CFSWNB, GRWNB, MIWNB, ReFWNB and TreeWNB. The purpose of the second experiment is to compare the proposed attribute weighted Naive Bayes, namely AISWNB, with each single weighted approach in literature. The classification accuracy of each algorithm on each data set is obtained via 10 runs of 10-fold cross validation, and our algorithm is carried out on the same training data sets and evaluated on the same test data. Finally, we compare related algorithms via two-tailed \(t\)-test with a 95% confidence level [23]. We use the classification accuracy as the criterion to assess the performance of algorithms. The higher the classification accuracy is, the better the corresponding classification effect [24].

C. Weighted NB vs. Standard NB

Table II reports the detailed results (the classification accuracy and standard deviation) of AISWNB and other baseline algorithms. Besides, Table III illustrates the compared results of two-tailed \(t\)-test, in which each entry \(w/t/l\) means that the algorithm in the corresponding row wins in \(w\) data sets, ties in \(t\) data sets and loses in \(l\) data sets on the 36 benchmark data sets, compared to the algorithm in the corresponding column. Overall, the results can be summarized as follows:

1. TreeWNB shows the best performance. Compared with NB, TreeWNB wins 7 data sets, ties 26 data sets, loses 3 data sets. The average classification accuracy on 36 standard data sets for TreeWNB \((82.45\pm4.78)\) is higher than NB \((82.16\pm4.88)\).
2. CFSWNB outperforms simple NB. Compared with NB, CFSWNB wins 8 data sets, ties 25 data sets and loses 3 data sets. But the average classification accuracy on 36 standard data sets for CFSWNB \((82.15\pm4.80)\) is slightly lower than NB.
3. SBC’s performance is comparable with NB (6 wins and 5 losses), though the average classification accuracy on 36 standard data sets for SBC \((81.90\pm4.99)\) is lower than NB.
4. MIWNB and GRWNB are all inferior to NB. Compared with NB, MIWNB (5 wins and 9 losses), GRWNB (5 wins and 6 losses). And the average classification accuracy on 36 standard data sets for MIWNB \((80.85\pm5.12)\) and GRWNB \((81.96\pm4.78)\) are both lower than NB.
5. ReFWNB is an inappropriate attribute method for improving accuracy performance. It is inferior to NB (5 wins and 12 losses), SBC (3 wins and 10 losses) and all other attribute weighted method: CFWSNB (4 wins and 13 losses), GRWNB (2 wins and 10 losses), MIWNB (5 wins and 8 losses) and TreeWNB (2 wins and 11 losses). The average classification accuracy on 36 standard data sets for ReFWNB (79.69 ± 3.95) is lower than all the compared methods.

D. Accuracy Comparisons Between AISWNB and Baselines

The above results demonstrate that WNB (CFSWNB and TreeWNB) can indeed outperform NB for classification. In Fig. 2, we report the performance of the proposed AISWNB, which uses artificial immune system principles to calculate attribute weight values for NB classification. In both figures, data points below the x = y diagonal line are data sets with average classification accuracy, which is significantly higher than the different algorithms. The detailed results are further presented in Table II. Our experimental results indicate that AISWNB has very significant accuracy gain compared to other weighted methods. In summary, our experimental results show:

1. AISWNB greatly outperforms NB (11 wins and 0 loss). The average classification accuracy on 36 data sets for AISWNB (83.43 ± 4.92) is higher than NB (82.16 ± 4.88).

2. AISWNB significantly outperforms SBC (10 wins and 0 loss). And the average classification accuracy on 36 data sets for AISWNB (83.43 ± 4.92) is higher than SBC (81.90 ± 4.99).

3. AISWNB outperforms CFSWNB (8 wins and 0 losses), GRWNB (9 wins and 0 losses), MIWNB (11 wins and 0 losses), ReFWNB (14 wins and 0 losses) and TreeWNB (8 wins and 0 losses), with average classification accuracy in 36 standard data sets for AISWNB (83.43 ± 4.92) is higher than CFSWNB (81.90 ± 4.99), GRWNB (81.96 ± 4.78), MIWNB (80.85 ± 5.12), ReFWNB (79.69 ± 3.95) and TreeWNB (82.45 ± 4.79).

E. Convergence and Learning Curves

In order to investigate the convergence of the AISWNB algorithm, we report the relationship between the number of iterations and classification accuracy on the 12 data sets (half for data sets with a large number of instances, the remaining half with a large number of attributes), and the results are shown in Figs. 3 and 4. Each point in the curves corresponds to the mean accuracy from 10-fold cross validation under the underlying iteration with the current optimal attribute weight values. Figs. 3 and 4 show that AISWNB converges quickly, with a higher classification accuracy than other algorithms. For additional insight into our experiment, we observe the “kr-vs-kp” for example. It is a high-dimensional data set (37 attributes) with 3196 instances. In addition, strong attribute dependencies have been found in this data set by Kohavi [25]. Our results show that AISWNB achieves 96.9% classification accuracy, which is significantly higher than NB’s accuracy on the same data set (which is 88.1%). The accuracy of the final convergence is better than CFSWNB (0.913), GRWNB (0.9), MIWNB (0.908), ReFWNB (0.904), TreeWNB (0.944), RMWNB (0.809) and NB (0.881).
levels of improvement can also be observed from other data sets.

In some situations, the proposed AISWNB can achieve a better accuracy than a number of attribute weighted baseline methods in only one iteration. This demonstrates the convergence performance of AISWNB. Meanwhile, in order to determine whether the improvement of AISWNB is attributed to random attribute weights or simple random weight values can have high accuracy performance, we use random weight values and denote the results by RMWNB in Tables II and III. The results clearly show that random attribute weight selection is non-effective to improve the accuracy of WNB. Detailed explanations can be discussed as follows according to the t-test results in Table III.
1. RMWNB is significantly inferior to AISWNB (0 wins and 15 losses). The average classification accuracy on 36 benchmark data sets for RMWNB (76.59±5.80) is much lower than AISWNB (83.43±4.92).

2. RMWNB is greatly inferior to CFSSWNB (0 wins and 14 losses), GRWNB (0 wins and 8 losses), MIWNB (3 wins and 8 losses), TreeWNB (1 wins and 15 losses) and SBC (1 wins and 12 losses). Average classification accuracy on the benchmark data sets for RMWNB is lower than all algorithms listed above. Compared with ReFWNB, RMWNB (6 wins and 6 losses) seems competitive with ReFWNB. But the average accuracy of RMWNB is lower than ReFWNB (79.69±3.95). It is worth noting that in Section IV-C we have investigated accuracy performance for a variety of attribute weighting methods for WNB in the literature. ReFWNB has shown to be an inappropriate method to improve the classification accuracy of WNB.

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed to use artificial immune system (AIS) for weighted Naive Bayes classification. Our method, namely AISWNB, uses immunity theory in artificial immune systems to search optimal weight values for NB classification. The experiments and comparisons on 36 benchmark data sets show that AISWNB outperforms state-of-the-art attribute weighted approaches. For high-dimensional data sets, AISWNB also demonstrates very good performance. Currently, AISWNB mainly focuses on improving the classification accuracy. Future work may include designing new evolution models to improve the performance of WNB in terms of AUC and classification accuracy. In addition, some other evolutionary algorithms and weighting methods can also be used for weighted naive Bayes classification.

REFERENCES


