Appl Intell (2006) 25:305–319
DOI 10.1007/s10489-006-0109-9

Data-driven fuzzy rule generation and its application for student academic performance evaluation

Khairul A. Rasmani · Qiang Shen

Abstract Several approaches using fuzzy techniques have been proposed to provide a practical method for evaluating student academic performance. However, these approaches are largely based on expert opinions and are difficult to explore and utilize valuable information embedded in collected data. This paper proposes a new method for evaluating student academic performance based on data-driven fuzzy rule induction. A suitable fuzzy inference mechanism and associated Rule Induction Algorithm is given. The new method has been applied to perform *Criterion-Referenced Evaluation* (*CRE*) and comparisons are made with typical existing methods, revealing significant advantages of the present work. The new method has also been applied to perform *Norm-Referenced Evaluation* (*NRE*), demonstrating its potential as an extended method of evaluation that can produce new and informative scores based on information gathered from data.

Keywords Data-driven learning · Fuzzy rule induction · Student performance evaluation

1 Introduction

Evaluation of student performance can be made based on Criterion-Referenced Evaluation (CRE) and Norm-Referenced Evaluation (NRE). In CRE, evaluation is carried out with respect to established criteria of performance [6, 7], i.e. student scores are implicitly referred to a set of specific criteria of achievement. Although existing methods have been used as a tool to double check student performance in CRE, evaluators often use *ad hoc* inference methods, which lack a formal mechanism, to support their derivation of a final mark or grade. It is therefore desirable to have an alternative, and systematic, method to help the user (students, parents, decision-makers, etc.) to confirm or refute the final result.

One of the drawbacks of CRE is the lack of its ability in reflecting the knowledge that has been used to support the evaluation, unable to show what criteria the 'final result' or 'score' refers to. Instead of using CRE, evaluation may also be made on the basis of NRE, a method of assessment based on comparison and utilizing information gathered from previous student performance data [6, 7]. Examples below show cases where NRE is necessary:

- *Case 1:* A student was awarded a 70% score on a subject. According to CRE, this belongs to the grade 'Excellent'. However, if compared to other student performance 70% is among the lowest marks and most other students get more than this, whilst CRE still considers this as 'Excellent' which is clearly not the case.
- *Case 2:* Based on CRE, a lecturer conducting a course has given all of his students good grades. Data from previous years however suggest that such achievement is rare. Thus, results obtained from CRE may not reflect the true performance of the students but NRE may provide some additional information on the student performance of the class when compared to other students' performance in previous years.

It is therefore helpful to present results obtained by NRE alongside those obtained by CRE to provide additional information about a student's achievement. Currently, statistical methods have been used to make comparisons of individual achievements with achievements in the *norm group* (e.g. larger student population). Such approaches

K. A. Rasmani · Q. Shen (✉)
Department of Computer Science, University of Wales,
Aberystwyth SY23 3DB, United Kingdom
e-mail: qqs@aber.ac.uk

however, have not been widely adopted, possibly because they produce numerical values that are less meaningful to the user.

The use of fuzzy approaches for educational evaluation is in general fairly new. However, it has reached a wide range of application areas in educational systems in addition to evaluation of student academic performance, including the evaluation of curriculum and that of the educators (e.g. lecturers and tutors). In student performance evaluation in particular, fuzzy techniques have been adapted for evaluation based on numerical scores obtained in an assessment [1, 3, 12] and for assessing prior educational achievement based on evidence such as academic certificates [2, 5].

Much attention has also been given to adopting fuzzy approaches for the evaluation of teaching using a computer, in particular in Intelligent Tutoring Systems (ITS) and Computer Assisted Instruction (CAI). For instance, in [8, 19, 25] fuzzy approaches were proposed for determining the level of a student's understanding of a certain subject matter in the context of ITS; and in [27] a fuzzy approach was proposed to assess student performance based on several criteria with a strong suggestion that the method be applied to CAI. Interesting work has been reported along this line of research. This includes evaluation of journal grades [29]; evaluation of vocational education performance [13]; collaborative assessment [11]; and performance appraisal systems of academics in higher education [23].

The focus of attention of this work is an evaluation of student academic performance. It proposes the use of a data-driven fuzzy rule induction approach to obtain user-comprehensible knowledge from historical data to justify any evaluation. This paper shows the advantages of the approach in student performance evaluation as it can be built not only based on information in a given dataset but also allowing expert knowledge to be added if such knowledge is available. Information induced from the dataset, especially that not formerly known by experts in the domain, can be very useful in developing fuzzy models for practical applications.

The rest of this paper is structured as follows: Section 2 reviews typical existing fuzzy techniques that have been used for aggregation of student scores to produce an evaluation of student performance. Section 3 presents a proposed technique based on data-driven fuzzy rule induction to perform evaluation of student performance. Section 4 gives experimental results, contrasting to the present work as outlined in Section 2 and, finally, conclusions are summarized in Section 5 with further work pointed out.

2 Background

The main characteristic of evaluation related to student performance is that theevaluation tasks require consideration of evidence collected via several modes of assessment such as practicals, examinations and observations. Such evaluation usually involves awarding scores as numerical values and grades that may often be expressed in linguistic terms such as good, bad, satisfactory, excellent, etc. These linguistic terms carry imprecision that may arise from different human interpretations and from different means of implementing the evaluation.

The use of linguistic terms in assessing performance has been the main reason for researchers applying fuzzy techniques to student performance evaluation. It has been argued that one of most appropriate ways of handling multiple variables that contain imprecise data is to use fuzzy logic reasoning which reflects the way of human-thinking. For example, in [1], Biswas (1995) states that the reasons behind the use of a fuzzy approach in their work are that an educational grading system usually involves a substantial amounts of vagueness and fuzzy theory can provide a possible model of subjective judgements. Also, in [5], Fourali (1994) states that the reason for adopting a fuzzy approach is that academic competence is a fuzzy concept and a decision on evidence is fuzzy because different assessors may have different standards. In [12], Law (1996) reinforces such views in supporting the use of fuzzy techniques for student performance evaluation by giving a list of reasons: (a) Scores/marks given by teachers for student performance are not very precise, (b) examinations consist of vague data, and (c) a common method of grading students is the use of linguistic variables.

The development of fuzzy approaches for evaluation of student performance involves three important tasks: fuzzification, inference and defuzzification. In general, student scores or marks (crisp values) have to be transformed into fuzzy input values before aggregation can be done using a fuzzy inference mechanism. Fuzzy values can also be obtained directly from domain experts, avoiding the need for fuzzification in this case. The outputs of fuzzy inference are typically also in terms of fuzzy values, representing a student's performance. These fuzzy values need to be transformed again into crisp values in order to produce an output, often a percentage mark, that can be easily understood by the user. The returned fuzzy values may also be used directly to describe the levels of performance [29].

To put present development in context and facilitate comparisons to be presented later, the remaining of this section reviews three existing approaches which have been used for evaluation of student performance, namely Biswas' Approach, Law's Approach, and Chen and Lee's Approach. For the purpose of simplicity, without losing generality, all the three methods will be explained with their application to perform evaluation of different scores obtained from several questions, Q_i.

Table 1 Standard Fuzzy Sets (SFS) to represent student performance

Linguistic terms	Fuzzy sets
Excellent	{0/0, 0/20, 0.8/40, 0.9/60,1/80, 1/100}
Very good	{0/0, 0/20, 0.8/40, 0.9/60, 0.9/80, 0.8/100}
Good	{0/0, 0/20, 0.8/40, 0.9/60, 0.9/80, 0.8/100}
Satisfactory	{0.4/0, 0.4/20, 0.9/40, 0.6/60, 0.2/80, 0/100}
Unsatisfactory	{1/0, 1/20, 0.4/40, 0.2/60, 0/80, 0/100}

Table 2 Grade and their corresponding mid-grade points

Linguistic terms	Grade/score	Mid-grade points
Excellent	$(90 \leq \mathrm{A} < 100)$	95
Very good	$(80 \leq \mathrm{B} < 90)$	85
Good	$(50 \leq \mathrm{C} < 70)$	60
Satisfactory	$(30 \leq \mathrm{D} < 50)$	40
Unsatisfactory	$(0 \leq \mathrm{F} < 30)$	15

2.1 Biswas' approach

Biswas (1995) proposed a fuzzy technique to perform evaluation based on student's answerscripts. It employs the idea of fuzzy similarity which is specifically defined as follows:

For two discrete fuzzy sets Q and M their similarity is:

$$S(Q, M) = \frac{\sum_i \mu_Q(x_i)\mu_M(x_i)}{\max\left(\sum_i \mu_Q(x_i), \sum \mu_M(x_i)\right)}, \tag{1}$$

where $i = 1, 2, \ldots,$ are the domain elements. Obviously $S(Q, M) \in [0, 1]$. Also, the larger the value of $S(Q, M)$, the greater the similarity between fuzzy sets Q and M.

In this work, the above measure is used to compare the similarity of a student's performance, expressed in fuzzy values, with *Standard Fuzzy Sets* (SFS), which are predefined with membership values corresponding to different levels of student performance. The SFS are devised by experts according to the standard fixed by educational authority, for example a department in a university. Table 1 shows an example of SFS used in [1] which refer to the following levels of student performance: Excellent (A), Very Good (B), Good (C), Satisfactory (D) and Unsatisfactory (F).

At the initial stage of evaluation, the evaluator needs to award fuzzy marks for each question (Q_i) into a *fuzzy grade sheet*, a table containing rows for question numbers and columns for awarding marks in term of fuzzy values. A matching operation is then performed according to Definition (1) for each question (Q_i), to each level of performance A, B, C, D and F, to obtain similarity values $S(Q_i, A)$, $S(Q_i, B)$, $S(Q_i, C)$, $S(Q_i, D)$ and $S(Q_i, F)$. The grade for each question is determined based on the maximum similarity value among the level of performance.

The calculation of the total score involves the use of marks allocated for each question and the so-called *mid-grade points* according to each grade awarded. An example of *mid-grade points* used in [1] is shown in Table 2. Different grades obtained from each question are used to calculate the total score based on the definition:

$$TS = \frac{1}{100}\left[\sum T(Q_i) \times P(g_i)\right] \tag{2}$$

where $T(Q_i)$ are marks allocated for each question and $P(g_i)$ are the *mid-grade points*. The total score (TS) will be in the form of crisp values $\in [0, 100]$ and the new final grade will be determined based on crisp interval values referring to the level of performance.

Although this technique shows the usefulness of using fuzzy membership values for aggregating marks from different questions, it has several disadvantages. In particular, the use of a *fuzzy grade sheet* to obtain fuzzy marks is very confusing because the fuzzy marks are not referred to each level of performance. In addition, this method may take a large amount of time to compute the matching operations between the fuzzy marks and each of the *SFS* [3]. This method also suffers from the use of *mid-grade points* in the calculation of the total score. These values may greatly influence the total score and thus can create unexpected results.

2.2 Chen and Lee's approach

Chen and Lee (1999) proposed a technique for evaluation of student answerscripts with an intention to resolve drawbacks of the method outlined above. In this approach, the *degrees of satisfaction* is defined in advance by experts with respect to levels of performance, from which the *maximum degree of satisfaction* per level is obtained. Examples of *degrees of satisfaction* and the *maximum degree of satisfaction* given in [6] is summarised in Table 3, which also shows the eleven levels of student performance that have been proposed and used.

Similar to the method of [1], the evaluator has to award fuzzy marks into the *fuzzy grade sheet* which is an extended version of that given in [1]. The fuzzy marks for each question

Table 3 Degrees of satisfaction according to performance level

Satisfaction levels	Degrees of satisfaction	Maximum degree of satisfaction
Extremely good (EG)	100%	1.00
Very very good (VVG)	91–99%	0.99
Very good (VG)	81–90%	0.90
Good (G)	71–80%	0.80
More or less good (MG)	61–70%	0.70
Fair (F)	51–60%	0.60
More or less bad (MB)	41–50%	0.50
Bad (B)	25–40%	0.40
Very bad (VB)	10–24%	0.24
Very very bad (VVB)	1–9%	0.09
Extremely bad (EB)	0%	0.00

(Q_i) are awarded according to each level of performance. From this, the degree of satisfaction for each individual is calculated such that

$$D(Q_i) = \frac{\sum \mu_{Q_i}(x_i) F(x_i)}{\sum \mu_{Q_i}(x_i)} \tag{3}$$

where $\mu_{Q_i}(x_i)$ are membership values awarded to each level of performance and $F(x_i)$ is the respective *maximum degree of satisfaction*.

The final step of the method is to calculate the total score *TS* based on several questions as follows

$$TS = \left[\sum T(Q_i) D(Q_i) \right] \tag{4}$$

where $T(Q_i)$ are marks allocated for each question by the evaluator and $D(Q_i)$ is the computed degrees of satisfaction for $Q_{i.}$. From *TS* a grade is awarded based on the satisfaction level that has been predefined.

As pointed out in [3], this technique is less complex compared to the approach presented in [1], whilst still able to produce useful estimation of student performance. Although the proposed method seems simple, the usage of the *maximum degree of satisfaction* is very confusing and the results of the aggregation are biased towards the number of satisfaction levels created. Fewer satisfaction levels means that the difference between the original score and the new score is bigger. The use of an *extended fuzzy grade sheet* to award fuzzy marks may not be practical when the problem scales up, as it involves awarding too many fuzzy values to evaluate each of the questions. This can become worse in cases where the number of questions or modes of assessment increases.

2.3 Law's approach

Law (1996) proposed an alternative approach to student performance evaluation, based on the notion of *fuzzy expected values.* The *fuzzy expected value* of a fuzzy set A is defined as:

$$E(A) = \frac{\int_{R^n} x \mu_A(x) f(x)\, dx}{\int_{R^n} \mu_A(x) f(x)\, dx} \tag{5}$$

with $\mu_A(x)$ being the membership function of x in A and $f(x)$ being the distribution function of x in A.

Contrary to the methods proposed in [1] and [3], in Law's approach, the original student scores are represented in crisp values. Fuzzification is used to transform such scores into fuzzy values. The fuzzy partitions underlying the fuzzification are defined in advance by experts based on an expectation of the percentage of students who will receive a certain level of performance (being one of the following five grades: *A, B, C, D* and F). A *fuzzy assessment matrix*, M, is created using the fuzzified values, in the form of:

$$M = \begin{bmatrix} \mu_A(Q_1) & \mu_B(Q_1) & \mu_C(Q_1) & \mu_D(Q_1) & \mu_F(Q_1) \\ \mu_A(Q_2) & \mu_B(Q_2) & \mu_C(Q_2) & \mu_D(Q_2) & \mu_F(Q_2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mu_A(Q_n) & \mu_B(Q_n) & \mu_C(Q_n) & \mu_D(Q_n) & \mu_F(Q_n) \end{bmatrix}$$

The matrix is employed, in conjunction with the *fuzzy expected values* for each level of performance to compute an intermediate new score vector (one new score per question):

$$NS = M \times [E(A), E(B), E(C), E(D), E(F)]^t \tag{6}$$

where the expected values for each level of performance *E*(*A*), *E*(*B*), *E*(*C*), *E*(*D*), and $E(F)$ are calculated using definition (5) and the same fuzzy partitions mentioned above. This new vector is then used to calculate the core of the total score (*CTS*),

$$CTS = \sum_{j=1}^{n} D(Q_j) NS_j \tag{7}$$

where $D(Q_j)$ are the full percentage marks allocated for each question. Since $CTS \in [0, 1]$, the final total score, *TS* is set to $CTS \times 100$ to obtained a readily understandable mark on student performance.

These approaches demonstrate the advantage of using *fuzzy expected value* in student performance evaluation. However, although it may be useful to obtain evaluation results according to expert expectation, the resulting new total score and grade may not reflect actual performance of the student on the subject matter. This is because the initial fuzzy partitions may not be specified with regard to students performance but to the expectation of, say how many students out to pass certain examination. Furthermore, as pointed out in [27], this method works with respect to single evaluation criterion; it cannot assess a student's performance based on multiple criteria. In addition, the method involves extensive computation which may limit the take-up of the approach in practice.

In summary, methods presented in [1, 3, 12] show that fuzzy approaches are potentially useful for student performance evaluation. However, apart from the previously discussed individual disadvantages, it can be observed that these methods also have several common shortcomings. Firstly, these methods produce a new total score in terms of crisp values before a new grade can be awarded. This can be a substantial setback as the difference of the new total score with the original score may be very large and thus create confusion for the user, especially the students. Secondly, all the methods are wholely based on expert opinions without offering the possibility of making direct use of information gathered from data. Newly developed fuzzy approaches should look into ways of avoiding, or at least reducing such disadvantages. The following section proposes such an approach.

3 Data-driven fuzzy rule based approach

Reasoning based on fuzzy approaches has been successfully applied for the inference of multiple attributes containing imprecise data; in particular, fuzzy rule-based systems (FRBS) which provide intuitive methods of reasoning have enjoyed much success in solving real-world problems. Recent developments in this area also show the availability of FRBS which allow interpretation of the inference in the form of linguistic statements whilst having high accuracy rates. The use of linguistic rule models such as "If assignment is very poor and exam is average then the final result is poor" helps capturing the natural way in which humans make judgements and decisions. Furthermore, historical data that is readily available in certain application domains [14] can be used to build fuzzy models which integrate information from data with expert opinions. It is also important that the designed fuzzy models are interpretable by, and explainable to, the user [24]. This section describes a newly proposed data-driven fuzzy rule induction method that achieves such objectives, and shows how the method can be applied to the classification of student performance. Description of Neuro-Fuzzy Classification (NEFCLASS) algorithm, which will be used later for comparison, is also given briefly in this section.

3.1 Weighted subsethood-based algorithm (WSBA)

Simplicity in generating fuzzy rules and the ability to produce high classification accuracy are the main objectives in the development of WSBA. To achieve these objectives, fuzzy subsethood measures and weighted linguistic fuzzy modelling are employed.

3.1.1 Fuzzy subsethood values

Fuzzy subsethood values represent the degree to which a fuzzy set is a subset of another fuzzy set. For example, for two fuzzy sets A and E, fuzzy subsethood values [4, 28] of fuzzy set A to fuzzy set E, denoted $S(E, A)$ can be defined as follows:

$$S(E, A) = \frac{M(E \wedge A)}{M(E)} = \frac{\sum_{x \in U} \nabla(\mu_E(x), \mu_A(x))}{\sum_{x \in U} \mu_E(x)} \quad (8)$$

where $S(E, A) \in [0, 1]$ and ∇ denotes a *t-norm* operator.

Fuzzy subsethood values have been used to address different problems, including to measure the *degree of truth* of learned fuzzy rules [28], and to promote certain linguistic terms as part of the antecedent of an emerging fuzzy rule [4].

3.1.2 Weight calculation

As with many existing techniques for representing weights, in this work, measures of weighting are limited to the range of 0 to 1, with 0 representing the lowest weight (or of least importance) and 1 the highest (or of most importance). Such weights can be calculated from fuzzy subsethood values as follows. Note that the meaning of *subsethood* is herein extended to allow fuzzy sets associated with different linguistic variables to be related.

Suppose that the subsethood value for a certain linguistic term A_i of linguistic variable A with regard to classification E is $S(E, A_i)$, and that the linguistic variable A has the following possible linguistic terms: $A_1, A_2, \ldots, A_l$. Then, the relative weight for linguistic term A_i, with regard to classification E is:

$$w(E, A_i) = \frac{S(E, A_i)}{\max_{j=1..l} S(E, A_j)} \quad (9)$$

Clearly, $w(E, A_i) \in [0,1]$ and $i = 1, 2, \ldots, l$. This allows the creation of a weight for each linguistic term per condition attribute. Intuitively, the linguistic term with the highest subsethood value will be the most important and that with the lowest will be the least important.

The resulting weights are attached to the linguistic terms associated with conditional attributes. Therefore, for each conditional attribute A, the compound weight $T(A)$ of the weighted conjunction of linguistic terms associated with it can be calculated such that

$$T(A) = \left(\frac{w_1}{w}(A_1)\nabla \cdots \nabla \frac{w_m}{w}(A_m)\right) \quad (10)$$

where ∇ is the *t-norm*, $A_i, i = 1, 2, \ldots, m$ are the linguistic terms of variable A, which are conjuctively combined, and w is the largest amongst the m associated weights: $w(E, A_i), i = 1, 2, \ldots, m$.

Similarly, the compound weight $T(B)$ of the weighted disjunction of linguistic terms associated with variable B is

$$T(B) = \left(\frac{w_1}{w}(B_1)\Delta \cdots \Delta \frac{w_n}{w}(B_n)\right) \quad (11)$$

where Δ is the *t-conorm*, and $A_i, i = 1, 2, \ldots, n$ are the linguistic terms of variable B, which are disjunctively combined.

3.1.3 Rule generation

Without losing generality, consider fuzzy rules with multiple conditional attributes and a single conclusion attribute. These rules could be written in the default form of fuzzy general rule, with each corresponding to one possible class:

Rule 1 **IF** A is (A_1 *OR* A_2 *OR* … *OR* A_i) **AND** B is (B_1 *OR* B_2 *OR* … *OR* B_j) **AND**… **AND** H is (H_1 *OR* H_2 *OR* … *OR* H_k) **THEN** the class is E_1
Rule 2 **IF** A is (A_1 *OR* A_2 *OR* … *OR* A_i) **AND** B is (B_1 *OR* B_2 *OR*… *OR* B_j) **AND** … **AND** H is (H_1 *OR* H_2 *OR* … *OR* H_k) **THEN** the class is E_2
.
.
.
Rule n **IF** A is (A_1 *OR* A_2 *OR* … *OR* A_i) **AND** B is (B_1 *OR* B_2 *OR*… *OR* B_j) **AND** … **AND** H is (H_1 *OR* H_2 *OR* … *OR* H_k) **THEN** the class is E_n (12)

In the above definition, 'OR' and 'AND' are fuzzy logical operators and are interpreted by minimum and maximum operator respectively. All linguistic terms of each attribute are used to describe the antecedent of each rule initially. This may look tedious, but the reason for keeping this complete form is that every linguistic term may contain important information that should be taken into account. Otherwise, there is no need for adopting the given fuzzy partitions of the underlying domains in the first place. Of course, during training, some of such terms may be omitted due to no evaluated contribution (or with a relative weight of 0) with regard to the training data (see later).

However, the above default rules do not tell any differences between the relative contributions made by the individual linguistic terms of each variable towards the eventual conclusion drawn. It is here that relative weights computed via subsethood values can help. Following this idea, by multiplying each linguistic term by its respective weight, the fuzzy rules to be generated will be of the form:

Rule 1 **IF** A is w(E_1, A_1)A_1 *OR* w(E_1, A_2)A_2 *OR* … *OR* w(E_1, A_i)A_i **AND** B is w(E_1, B_1)B_1 *OR* w(E_1, B_2)B_2 *OR*… *OR* w(E_1, B_j)B_j **AND** … **AND** H is w(E_1, H_1)H_1 *OR* w(E_1, H_2)H_2 *OR* … *OR* w(E_1, H_k)H_k **THEN** the class is E_1
Rule 2 **IF** A is w(E_2, A_1)A_1 *OR* w(E_2, A_2)A_2 *OR* … *OR* w(E_2, A_i)A_i **AND** B is w(E_2, B_1)B_1 *OR* w(E_2, B_2)B_2 *OR*… *OR* w(E_2, B_j)B_j **AND** … **AND** H is w(E_2, H_1)H_1 *OR* w(E_2, H_2)H_2 *OR* … *OR* w(E_2, H_k)H_k **THEN** the class is E_2
.
.
.
Rule n **IF** A is w(E_n, A_1)A_1 *OR* w(E_n, A_2)A_2 *OR* … *OR* w(E_n, A_i)A_i **AND** B is w(E_n, B_1)B_1 *OR* w(E_n, B_2)B_2 *OR*… *OR* w(E_n, B_j)B_j **AND** … **AND** H is w(E_n, H_1)H_1 *OR* w(E_n, H_2)H_2 *OR* … *OR* w(E_n, H_k)H_k **THEN** the class is E_n (13)

Computationally, the ruleset can be simply represented by

$$Y_k = \underset{i=1..m}{\Delta}\left(\underset{j=1..n}{\nabla}(w_{A_{ij},E_k} \times \mu_{A_{ij}}(x)\right), \quad k = 1, 2, \ldots, n. \tag{14}$$

where w_{A_{ij},E_k} denote the weights of atomic linguistic propositions and $\mu_{A_{ij}}(x)$ represent the membership function of the linguistic terms modified by the weights, with Δ and ∇ denoting the interpretation of logical disjunction and conjunction operators respectively.

This method does not require any threshold value and generates a fixed number of rules according to the number of classes of interest (i.e. one rule will be created for each class). In the process of generating fuzzy rules, linguistic terms that have a weight greater than zero will automatically be promoted to become part of the antecedents of the resulting fuzzy rules. Any linguistic term that has a weight equal to 0 will of course be removed from the fuzzy rule. This will make the rules simpler than the original default rules (13). In running WSBA for classification tasks, the concluding classification will be that of the rule whose overall weight is the highest amongst all. The structure of WSBA approach is shown in Fig. 1. Example applications of WSBA can be found in [20, 21].

3.2 Neuro-Fuzzy Classification (NEFCLASS)

Neuro-Fuzzy Classification (NEFCLASS) is an FRBS which combines a neural network learning approach with a fuzzy rule-based inference method [16–18]. NEFCLASS can be encoded as a three-layer *feedforward* neural network. The first layer represents the fuzzy input variables, the second layer represents the fuzzy rulesets and the third layer represents the output variables. The functional units in this network implement *t-norms* and *t-conorms*, replacing the activation functions that are commonly used in conventional neural networks. NEFCLASS is a data-driven FRBS that has the ability to create fuzzy membership functions and fuzzy rules automatically from training instances. Prior knowledge in the form of fuzzy rules can also be added to the rule base and used alongside new rules created using the training dataset.

Fuzzy rules are generated based on overlapping rectangular clusters that are created by the grid representing fuzzy sets for the conditional attributes. Clusters that cover areas where training data is located are added to the emerging rule-base. The system allows the user to choose the maximum number of rules, otherwise the number of rules are restricted to that of just the best performing ones. The firing strength of each

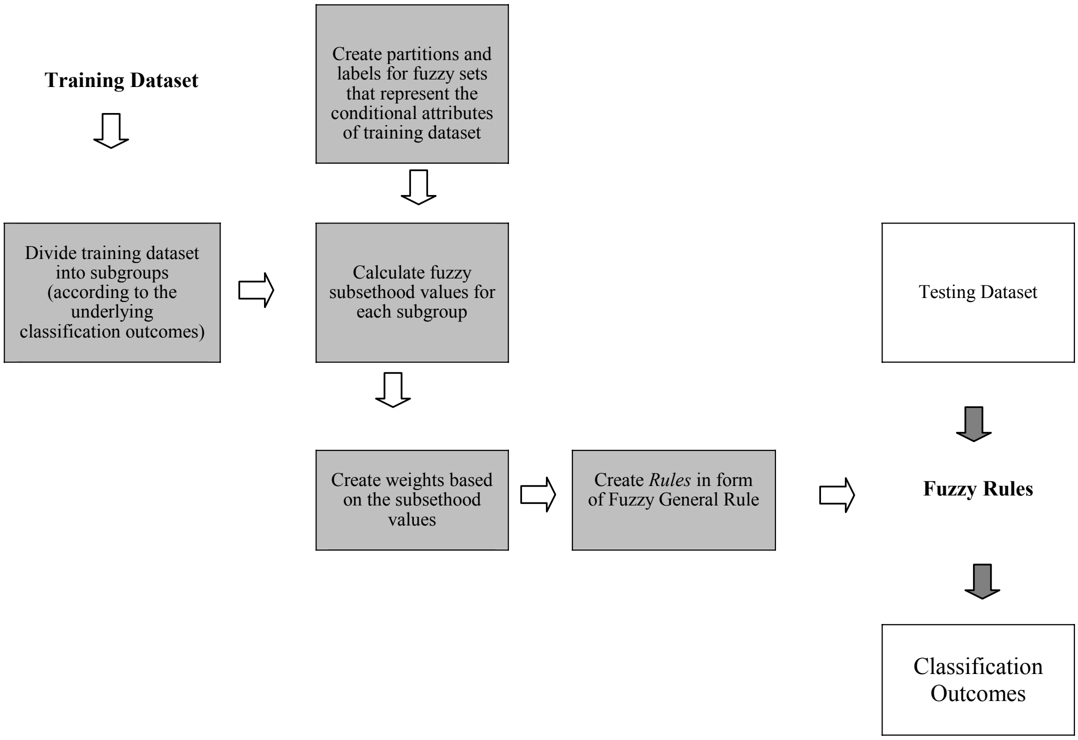


Fig. 1 Structure of WSBA approach

rule is used to reach the conclusion on the decision class of new observations.

The number of partitions and the shape of membership functions of the conditional attributes are user-defined. The rule learning process can be started, for example, using a fixed number of equally distributed triangular membership functions. A simple heuristic method is used for the optimization of membership functions. The optimization process results in changes to the membership function's shape by making the supports of the fuzzy set larger or smaller. Constraints can be employed in the optimization process to make sure that the fuzzy sets overlap each other.

NEFCLASS has undergone through several refinements over the years. For example, to enhance the interpretability of the induced fuzzy rules, NEFCLASS offers additional features such as rule pruning and variable pruning. The system has also been tested not only for classification of benchmark datasets but also for real world problems such as presented in [16].

4 Experimental results

The experiments presented in this section served as examples to illustrate the potential of WSBA for the evaluation of student performance. Note that a wide range of assessment methods are available and have been used (see for example [9]), depending on the purpose to conduct the assessment. In this paper, only CRE and NRE are considered for the implementation. The objective of the experiment involving CRE is to provide evidence that the proposed algorithm will produce results similar to the original grades obtained using statistical methods, if an ideal and representative training data is available. The objective of the experiment involving NRE is to show that WSBA is able to produce grades that can be used to provide additional information on the achievement of the students. In conducting these experiments, the following aspects have been taken into account:

- In data-driven rule based systems, decision classes of the training instances are typically those given by experts. In students' performance evaluation, such decisions are normally given by experts based on an aggregation of numerical crisp scores. This method is used to obtain the decision class for the training data (SAP50A and SAP50B, as presented in Appendices A and B respectively).
- The small training data (SAP50A and SAP50B) is used as an example and in the form of numerical crisp scores, which is the most popular way to measure student performance. Note that the fuzzy approach allows the possibility of utilizing data in the form of fuzzy values such as those proposed in [1, 3] or in terms of linguistic labels that represent the fuzzy sets such as those shown in Table 1. In such cases, the decision class for the training data is determined by fuzzy values (see for example [4]).

To avoid confusion, 'original score/grade' in this section will refer to the score and grade obtained from the use of the standard statistical mean and 'new score/grade' will refer to the score or grade obtained from existing fuzzy approaches, including WSBA and NEFCLASS. Note that both datasets used include only numerical scores, to facilitate comparison with other approaches. This need not be the case in general, the scores of individual assessment components may be given in fuzzy terms (as often the case for coursework grading for instances).

4.1 Criterion referenced evaluation (CRE)

Three existing methods are selected to support the comparative studies, as outlined in Section 2. Additionally, NEFCLASS is used for further comparison, employing a fuzzy rule-based approach. The dataset used for the purpose of training WSBA and NEFCLASS models is a set of student performance records (labeled SAP50A). It consists of 50 instances, involving three conditional attributes: assignment, test and final exam, and five possible classification outcomes: Unsatisfactory (E), Satisfactory (D), Average (C), Good (B) and Excellent (A). Note that the term 'Average' describing students' performance used in this paper is not referring to the statistical average. For the sake of simplicity, only five linguistic labels similar to the classification outcomes are used to represent student achievements. The fuzzy partitions and labels (shown in Fig. 2) are based on expert opinions representing the students' performance. The primary assumption is that the partitions chosen by experts are those best possible to represent the training data (SAP50A). Clearly, better fuzzification, if available will help improve the experimental results reported below. Note that the given definition of the fuzzy sets is obtained solely on the basis of the normal distribution of the crisp marks given. This ensures their comparison with other approaches.

The classification of the grades in this experiment is based on an interval that refers to the level of performance given by experts as shown in Table 4. To facilitate a fair comparison, the same dataset consisting of 15 instances and having the same features as the training dataset is used for all of the methods. The details of the testing dataset are shown in Table 5.

Table 4 Marks and their associated original grade and level of achievement

Marks	Grade	Level of achievement
0–25	E	Unsatisfactory
26–45	D	Satisfactory
46–55	C	Average
56–75	B	Good
76–100	A	Excellent

Table 5 Testing dataset

Case	Assignment	Test	Final exam
1	10.00	23.33	20.00
2	5.00	16.67	12.00
3	15.00	13.33	18.00
4	45.00	26.67	40.00
5	35.00	33.33	30.00
6	35.00	50.00	38.00
7	45.00	43.33	54.00
8	50.00	40.00	50.00
9	45.00	50.00	58.00
10	50.00	70.00	62.00
11	65.00	70.00	74.00
12	85.00	60.00	76.00
13	95.00	76.67	86.00
14	85.00	83.33	96.00
15	90.00	90.00	98.00

The experimental results obtained from the three existing fuzzy approaches are shown in Table 6 and the results for WSBA and NEFCLASS are listed in Table 7. It can be seen that the conventional fuzzy approaches produce different scores from the original (that is obtained by statistical mean). Thus, it is expected that when the new scores are translated into new grades, some of them may be different from the original grades. In particular, the results returned by the method of Biswas (1995), give rise to unexpected

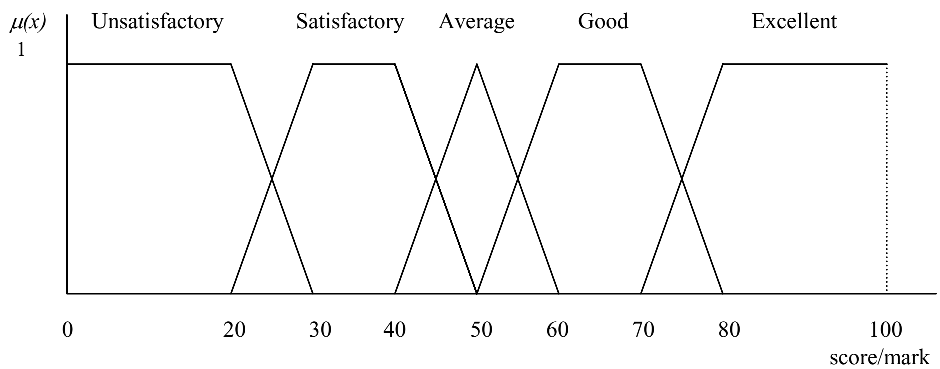


Fig. 2 Fuzzy partition for five levels of student performance

Table 6 Comparison of scores and grades obtained by Biswas', Chen and Lee's and Law's approaches based on CRE

	Statistical mean		Biswas' approach		Chen and Lee's approach		Law's approach	
Case	Final marks	Grade	New score	New grade	New score	New grade	New score	New grade
1	17.78	E	12.5	E	27.22	D*	15.29	E
2	11.22	E	12.5	E	25	E	13.17	E
3	15.44	E	12.5	E	25	E	13.17	E
4	37.22	D	37.5	D	44.44	D	33.82	D
5	32.78	D	50	C*	45	D	34.38	D
6	41.00	D	45	D	48.33	C*	39.07	D
7	47.44	C	45	D*	53.78	C	48.44	C
8	46.67	C	40	D*	51.67	C	45.31	D*
9	51.00	C	40	D*	58.67	B*	54.69	C
10	60.67	B	35	D*	68.33	B	62.50	B
11	69.67	B	45	D*	78.33	A*	69.87	B
12	73.67	B	70	B	88.33	A*	76.23	A*
13	85.89	A	87.5	A	97.22	A	84.71	A
14	88.11	A	87.5	A	100	A	86.83	A
15	92.67	A	87.5	A	100	A	86.83	A

*Indicates that the new grade is different from the original.

Table 7 Comparison of results obtained using WSBA and NEFCLASS based on CRE

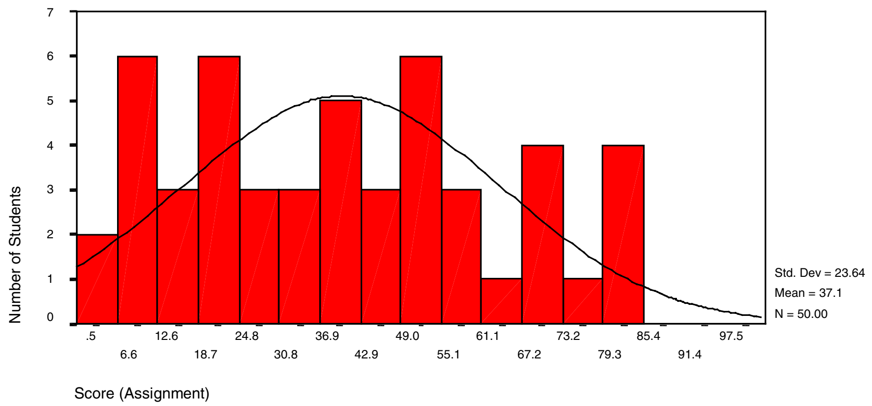
	Original score		WSBA						NEFCLASS					
			Membership value degree						Membership value degree					
Case	Final marks	Grade	E	D	C	B	A	New grade	E	D	C	B	A	New grade
1	17.78	E	0.667	0	0	0	0	E	0.983	0	0	0	0	E
2	11.22	E	1	0	0	0	0	E	1	0	0	0	0	E
3	15.44	E	1	0	0	0	0	E	1	0	0	0	0	E
4	37.22	D	0.125	0.5	0.172	0	0	D	0	0.53	0	0	0	D
5	32.78	D	0.190	0.75	0.172	0	0	D	0.105	0.404	0	0	0	D
6	41.00	D	0	0.136	0.172	0	0	C*	0	0	0.166	0	0	C*
7	47.44	C	0	0	0.227	0.066	0	C	0	0	0.315	0	0	C
8	46.67	C	0	0	0.125	0.045	0	C	0	0.067	0.087	0	0	C
9	51.00	C	0	0	0.5	0.197	0	C	0	0	0.477	0	0	C
10	60.67	B	0	0	0.125	0.833	0	B	0	0	0.167	0.151	0	C*
11	69.67	B	0	0	0.063	0.6	0.111	B	0	0	0	0.151	0	B
12	73.67	B	0	0	0	0.444	0.111	B	0	0	0	0.409	0.458	A*
13	85.89	A	0	0	0	0.333	0.667	A	0	0	0	0	0.609	A
14	88.11	A	0	0	0	0.273	1	A	0	0	0	0	1	A
15	92.67	A	0	0	0	0.273	1	A	0	0	0	0	1	A

*Indicates that the new grade is different from the original.

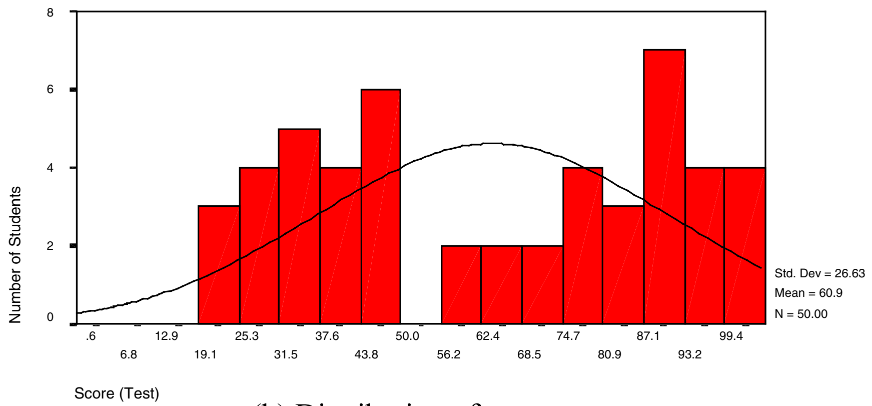
new scores such as case 10 where the original score of 61.67 (grade B) was downgraded to 35 (grade D). This is due to the approximation that is used in creating *mid-grade points*, and partly due to the use of fuzzy input values. Note that the use of *mid-grade points* has also resulted in a minimum score of 12.5 and a maximum score of 87.5, narrower than the original range.

Using Chen and Lee's method, all of the new scores are higher than the original. This is due to the use of maximum values of the *degree of satisfaction* created for each level of achievement. As for the results produced by Law's method, it is expected that the new scores will be different because the expected value for each grade has been predefined in advance according to the percentage of students who will receive a certain grade. Thus, results produced by this method may not reflect the students' true performance and they will be different if the expert evaluator changes the setting for the percentage.

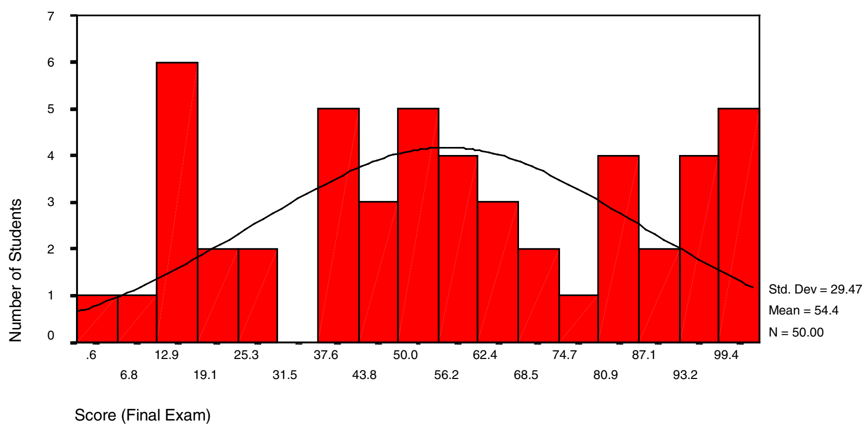
Fig. 3 (a) Distribution of assignment score (b) Distribution of test score (c) Distribution of final exam score



(a) Distribution of assignment score



(b) Distribution of test score



(c) Distribution of final exam score

In Table 7, it can be seen that by using the data-driven fuzzy rule-based approaches, fuzzy membership values obtained from fuzzy rules can be used to determine the new grade. Thus, it can be observed that the use of membership values in describing a student result has several advantages. First, these membership values can be interpreted as how strong the student's performance belongs to a specific grade. This can be very useful in differentiating smoothly student performances over boundary cases, giving a second opinion in deciding on borderline performances. An example of such a case is shown in case number 6 of Table 7. Second, with the use of fuzzy values, further analysis of estimated performance can be carried out directly, without the need for fuzzification. Third, the success of those methods in performing

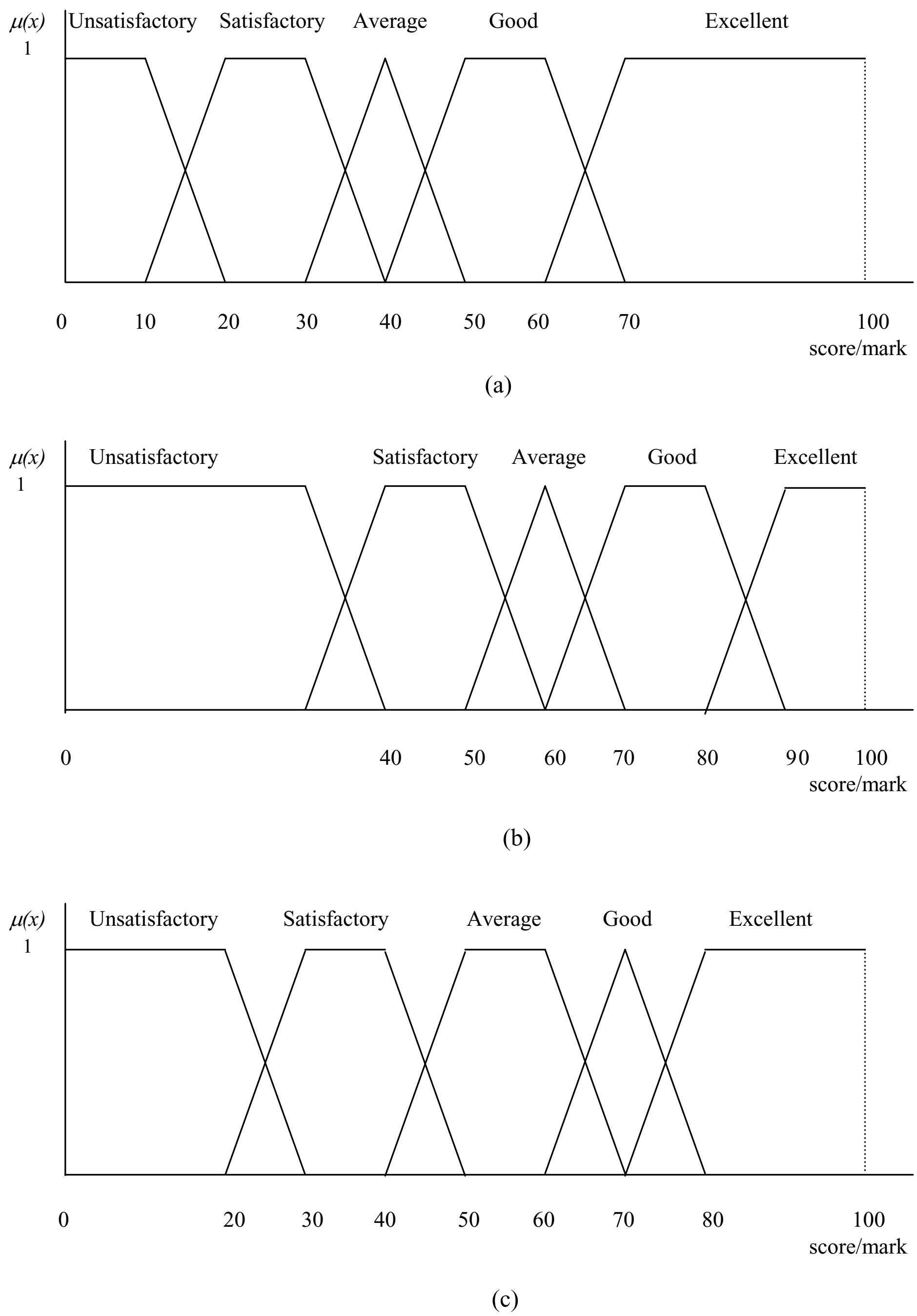


Fig. 4 (a) Partition for assignment score (b) Partition for test score (c) Partition for final exam score

CRE will allow them to be used for NRE. This also provides the possibility that student performance evaluation can be carried out properly using fuzzy values and linguistic terms (Good, Excellent, etc.) rather than the traditional numerical crisp values.

On comparison of the results produced by WSBA with NEFCLASS, from Table 7 it can be seen that WSBA has the ability to produce better classification (in terms of less grades changed). This is an advantage in addition to its computational simplicity.

4.2 Norm-Referenced Evaluation (NRE)

The purpose of this experiment is to demonstrate how the NRE can be conducted using WSBA. The training dataset for this experiment is a student performance dataset, labeled SAP50B (refer to Appendix B). This dataset is different from the dataset used for CRE in terms of distribution of data in each of the assessment components. This is to reflect the fact that students' performance in an assessment component does not always distribute normally. The distribution of the scores

Table 8 Comparison between original grades, standardized scores and the grades obtained by WSBA

Case	Original score		Statistical standardized score	WSBA					
				Membership value degree					
	Final marks	Grade	z-score values	E	D	C	B	A	New grade
1	17.78	E	−1.432	0.844	0.444	0	0	0	E
2	11.22	E	−1.717	0.844	0.444	0	0	0	E
3	15.44	E	−1.534	0.5	0.444	0	0	0	E
4	37.22	D	−0.589	0.023	0.409	0.5	0	0	C*
5	32.78	D	−0.782	0.023	0.667	0.5	0	0	D
6	41.00	D	−0.425	0.023	0.469	0.5	0	0	C*
7	47.44	C	−0.146	0	0.409	0.467	0.129	0	C
8	46.67	C	−0.179	0	0.182	0.7	0	0	C
9	51.00	C	0.009	0	0.409	0.5	0.386	0	C
10	60.67	B	0.429	0	0.182	0.765	0.8	0.006	B
11	69.67	B	0.819	0	0.091	0.191	0.3	0.351	A*
12	73.67	B	0.993	0	0	0	0.135	0	B
13	85.89	A	1.523	0	0	0	0.135	0.351	A
14	88.11	A	1.619	0	0	0	0.135	0.333	A
15	92.67	A	1.817	0	0	0	0.135	1	A

*Indicates that the new grade is different from the original.

for assignment, test and final exam are shown in Figs. 3(a)–(c). The fuzzy partitions to represent each of the assessment components (Figs. 4(a)–(c)) are created mainly on the basis of statistical distribution of the data and also partly on expert opinions (for the sake of simplicity). Any available methods to construct the fuzzy membership functions from training data can be utilized. These partitions are used to transform crisp values of conditional attributes into fuzzy values for both training and testing. The method to identify the decision class for the training data is similar to the one used for CRE.

The same testing dataset used in the experiment on CRE is employed here. The results obtained using WSBA are compared with the result produced by the popular statistical z-score method [26]. In calculating the statistical standardized score, the mean and standard deviation of the training dataset (SAP50B) are used.

Note that as with any data-driven learning problems, the issue of choosing which 'norm' group should be used as the basis for comparison is very important. Learned rules can only be as good as the data provided for learning. Thus, this approach has an inherent limitation, regardless of the method employed (statistical or fuzzy). Nevertheless, for real applications of the work presented herein, it is reasonable to assume that there is considerable amount of historical data which is representative to use (as is the case for any established educational organization), even though for the matter of illustrative convenience a relatively small dataset is adopted for training.

The results of this experiment is presented in Table 8, comparing the original scores and grades, the values calculated using a statistical standardized score, and the membership values and grades obtained by WSBA. Note that NRE is an assessment method that refers to other student's performance. Thus, the results obtained by this approach are not necessarily similar to those obtained by CRE, depending on which data the assessment refers to.

It can be seen from Table 8 that the results of the statistical standardized-score approach show how much the original score is diverted from the mean and standard deviation of the training dataset. However, it is often the case that these z-score values do not help the student or user to understand the evaluation that has been made (even though they may make good sense for expert evaluators).

The experimental results show that there are three cases where the new grades produced by WSBA are different from the original. Suppose that the training data represents students' results from the previous year. The new grade created by the system induced by WSBA can then be interpreted as the "grade obtained when the original results were compared to last year's performance". Thus in cases 4 and 6, for example, the students were awarded grade D using the CRE but this result is equivalent to grade C when the score is compared to the previous year's result. This kind of information is very helpful when the user wants to compare one student's achievement with the achievements of a (possibly different) group of students or with those of a larger population of students. Importantly, as suggested earlier, these new grades can be presented alongside the grades obtained via CRE. Thus, it will provide additional information that can be very useful for decision-making on degree classification or for better interpretation of classified degrees, for instances.

5 Conclusion

This paper has presented examples of how a fuzzy rule-based approach can be used for aggregation of student academic performance. It has been shown that the proposed approach has several advantages compared to existing fuzzy techniques for the evaluation of student academic performance. In CRE, the use of fuzzy membership values to determine the decision is very helpful for the user to understand why the new grade was awarded. In CRE, the proposed method has the potential to be developed further for use as an extended method of evaluation by providing new grades that refer to achievements of other groups. The membership values produced by this method are also more meaningful compared to the values produced by statistical standardized-score. However, it is worth noting that the newly proposed fuzzy approach is not to replace the traditional method of evaluation; instead it is meant to help strengthen the system that is commonly in use, by providing additional information for decision making by the user.

In this paper, WSBA is proposed to be employed for this purpose because of the simplicity of the method. It has been shown that although WSBA employs a simple approach, the proposed method is able to provide classification similar to that produced by more sophisticated algorithm such as NEFCLASS. Of course, more complex fuzzy rule-based methods such as those based on Evolutionary Computation, Fuzzy Clustering and Neural Networks may also be used [10, 15, 22]. However, the simpler approach has an advantage in terms of transparency and understandability of the methods and its results.

The proposed method also provides room for other improvements. In particular, interpretability of learned fuzzy rules has always been regarded as a very important factor in FRBS but has not been sufficiently addressed in this paper. Thus, further research should include this very important issue. As an approximate modellling approach, WSBA has the advantage in producing fuzzy systems of high classification accuracy, but the use of crisp weights to modify fuzzy terms is rather unnatural and may lead to confusion regarding the semantics of the resulting systems. However, the structure of WSBA rulesets enables the system model to be adapted with fuzzy quantifiers [20], making the model more interpretable whilst maintaining its accuracy. Also, the creation of fuzzy partitions to be used for WSBA are currently based on expert opinion and partly from statistical information on the training data. The fuzzification is not in any way optimized. Further research should include the use of methods that generate better fuzzy partition automatically from data.

Appendix 1 Student performance dataset (SAP50A)

Case	Assignment	Test	Final exam	Final marks	Grade
1	5	37	18	20.00	E
2	10	23	16	16.33	E
3	15	13	6	11.33	E
4	40	13	20	24.33	E
5	25	31	14	23.33	E
6	15	10	26	17.00	E
7	10	13	30	17.67	E
8	10	17	8	11.67	E
9	25	23	4	17.33	E
10	5	17	12	11.33	E
11	12	32	34	26.00	D
12	25	33	30	29.33	D
13	30	30	34	31.33	D
14	40	20	38	32.67	D
15	50	40	30	40.00	D
16	65	17	38	40.00	D
17	50	26	38	38.00	D
18	55	35	38	42.67	D
19	50	40	40	43.33	D
20	45	51	36	44.00	D
21	40	60	44	48.00	C
22	35	60	48	47.67	C
23	32	50	65	49.00	C
24	55	60	48	54.33	C
25	30	70	54	51.33	C
26	45	47	60	50.67	C
27	40	40	64	48.00	C
28	35	50	58	47.67	C
29	35	63	58	52.00	C
30	25	47	72	48.00	C
31	40	67	64	57.00	B
32	35	61	76	57.33	B
33	60	70	54	61.33	B
34	50	60	66	58.67	B
35	80	73	62	71.67	B
36	55	75	76	68.67	B
37	75	57	84	72.00	B
38	50	87	72	69.67	B
39	70	47	86	67.67	B
40	85	57	76	72.67	B
41	70	82	76	76.00	A
42	80	87	74	80.33	A
43	85	90	80	85.00	A
44	75	83	84	80.67	A
45	85	87	88	86.67	A
46	90	67	96	84.33	A
47	95	87	90	90.67	A
48	95	97	98	96.67	A
49	90	93	94	92.33	A
50	100	83	98	93.67	A

Appendix 2 Student performance dataset (SAP50B)

Case	Assignment	Test	Final exam	Final marks	Grade
1	5	34	16	18.33	E
2	2	45	46	31.00	D
3	23	45	19	29.00	D
4	34	43	46	41.00	D
5	5	23	11	13.00	E
6	17	96	48	53.67	C
7	61	98	94	84.33	A
8	29	97	57	61.00	B
9	74	90	93	85.67	A
10	52	34	69	51.67	C
11	33	39	37	36.33	D
12	6	21	22	16.33	E
13	15	74	35	41.33	D
14	48	76	50	58.00	B
15	81	89	97	89.00	A
16	79	92	98	89.67	A
17	28	66	87	60.33	B
18	23	84	23	43.33	D
19	8	39	14	20.33	E
20	19	33	64	38.67	D
21	58	64	98	73.33	B
22	39	25	65	43.00	D
23	43	39	65	49.00	C
24	52	94	66	70.67	B
25	68	79	94	80.33	A
26	48	77	51	58.67	B
27	1	43	13	19.00	E
28	21	31	81	44.33	D
29	45	75	53	57.67	B
30	65	97	79	80.33	A
31	34	71	49	51.33	C
32	13	25	7	15.00	E
33	16	23	78	39.00	D
34	27	59	35	40.33	D
35	51	31	58	46.67	C
36	48	89	73	70.00	B
37	67	63	92	74.00	B
38	57	88	85	76.67	A
39	66	96	99	87.00	A
40	43	79	41	54.33	C
41	78	87	78	81.00	A
42	55	21	56	44.00	D
43	7	38	36	27.00	D
44	21	87	23	43.67	D
45	78	78	97	84.33	A
46	16	98	36	50.00	C
47	15	45	12	24.00	E
48	39	21	12	24.00	E
49	37	59	57	51.00	C
50	6	45	3	18.00	E
Minimum	1.00	21.00	3.00	13.00	–
Maximum	81.00	98.00	99.00	89.67	–
Mean	37.12	60.90	54.36	50.79	–

Acknowledgments This research is partially funded by the Universiti Teknologi MARA, Malaysia and The Public Service Department, Government of Malaysia. The authors are grateful to the anonymous referees, and to Knox Haggie and Richard Jensen for their comments and suggestions in producing and revising this article.

References

1. Biswas R (1995) An application of fuzzy sets in students' evaluation. Fuzzy Sets Syst 74:187–194
2. Carlsson C, Fuller R, Fuller S (1997) OWA operators for doctoral student selection problem. In: Yager RR, Kacprzyk J (eds) The ordered weighted averaging operator: Theory, methodology and applications. Kluwer Academic Publisher, London.
3. Chen SM, Lee CH (1999) New methods for students' evaluation using fuzzy sets. Fuzzy Sets Syst 104:209–218
4. Chen SM, Lee SH, Lee CH (2001) A new method for generating fuzzy rules from numerical data for handling classification problems. Appl Artif Intell 15:645–664
5. Fourali C (1994) Fuzzy logic and the quality of asses sment of portfolios. Fuzzy Sets Syst 68:123–139
6. Gipps C, Stobart G (1993) Assessment: A teachers' guide to the issue. Hodder and Stoughton, London.
7. Goldstein H (1996) Statistical and psychometric models for assessment. In: Goldstein H, Lewis T (eds) Assessment: Problems, developments and statistical issues, John Wiley & Sons, Chichester. pp 41–55
8. Graesser AC, Person N, Harter D (2001) Teaching tactics and dialog in autotutor. Int J Artif Intell Educ 12:257–279
9. Hounsell D, McCulloch M, Scott M (1996) The ASSHE inventory: changing assessment practices in scottish higher education. Edinburgh: Centre for Teaching and Learning Assessment, The University of Edinburgh
10. Jang J-SR (1993) ANFIS: Adaptive-network-based fuzzy inference systems. IEEE Trans Syst, Man and Cybern 23:665–685
11. Kwok CW, Ma J, Vogel D, Zhou D (2001) Collaborative assessment in education: An application of a fuzzy GSS. Inform Manag 39:243–253
12. Law CK (1996) Using fuzzy numbers in educational grading system. Fuzzy Sets Syst 83:311–323
13. Lee C, Liu L, Tzeng GH (2001) Hierarchical fuzzy integral evaluation approach for vocational education performance: case of Junior Colleges in Taiwan. Int J Fuzzy Syst 3:476–485
14. Marin-Blazquez JG, Shen Q (2002) From approximate to descriptive fuzzy classifiers. IEEE Trans Fuzzy Syst 10:484–497
15. Mitra S, Hayashi Y (2000) Neuro-fuzzy rule generation: Survey in soft computing framework. IEEE Trans Neur Netw 11:748–768
16. Nauck D (2003) Measuring interpretability in rule-based classification systems. Proceedings of the International Conference on Fuzzy Systems, St. Louis, Missouri
17. Nauck D (1997) Neuro-fuzzy systems: review and prospects. In: Proceedings of the Fifth European Congress on Intelligent Techniques and Soft Computing, Aachen
18. Nauck D, Kruse R (1997) What are neuro-fuzzy classifiers? Proceedings of the Seventh International Fuzzy Systems Association World Congress IFSA'97, Academia, Prague
19. Papanikolaou KA, Magoulas GD, Grigoriadou M (2000) Computational intelligence in adaptive hypermedia. In: Proceedings of the INNS-IEEE International Conference on Neural Networks, Como, Italy
20. Rasmani KA, Shen Q (2004) Modifying fuzzy subsethood-based rule models with fuzzy quantifiers. In: Proceedings of the IEEE International Conference on Fuzzy Systems, Budapest, Hungary

21. Rasmani KA, Shen Q (2003) Weighted linguistic modelling based on fuzzy subsethood values. In: Proceedings of the International Conference on Fuzzy Systems, St. Louis, Missouri
22. Ross TJ (1995) Fuzzy logic with engineering applications, 2nd edn., John Wiley & Sons
23. Shaout A, Al-Shammari M (1998) Fuzzy logic modeling for performance appraisal systems: A framework for empirical evaluation. Expert Syst with Appl 14:323–328
24. Shen Q, Chouchoulas A (2002) A rough-fuzzy approach for generating classification rules. Patt Recogn 35:2425–2438
25. Stathacopoulou R, Magoulas GD, Grigoriadou M (1999) Neural network-based fuzzy modeling of the student in intelligent tutoring systems. In: Proceedings of the International Joint Conference on Neural Networks, Washington
26. Vogt WP (1999) Dictionary of statistics and methodology: A non-technical guide for the social sciences, 2nd edn. Sage Publication, London
27. Weon S, Kim J (2001) Learning achievement evaluation strategy using fuzzy membership function. In Proceedings of the 31st ASEE/IEEE Frontiers in Education Conference, October 10–13, Reno, NV
28. Yuan Y, Shaw MJ (1995) Induction of fuzzy decision trees. Fuzzy Sets Syst 69:125–139
29. Zhou D, Ma J, Turban E, Bolloju N (2002) A fuzzy set approach to the evaluation of journal grades. Fuzzy Sets Syst 131:63–74

Khairul Rasmani is a lecturer at the Faculty of Information Technology and Quantitative Sciences, Universiti Teknologi MARA, Malaysia. He received his Masters Degree in Mathematical Education from University of Leeds, UK in 1997 and his Ph.D. degree from University of Wales, Aberystwyth, UK in December 2005. His research interests include fuzzy approximate reasoning, fuzzy rule-based systems and fuzzy classification systems.

Qiang Shen is a Professor and the Director of Research with the Department of Computer Science at the University of Wales, Aberystwyth, UK. He is also an Honorary Fellow at the University of Edinburgh, UK. His research interests include fuzzy systems, knowledge modelling, qualitative reasoning, and pattern recognition. Prof. Shen serves as an associate editor or editorial board member of a number of world leading journals, including the IEEE Transactions on Systems, Man, and Cybernetics (Part B), the IEEE Transactions on Fuzzy Systems, and Fuzzy Sets and Systems. He has acted as a Chair or Co-chair at a good number of major conferences in the field of Computational Intelligence. He has published a book and over 170 peer-refereed articles in international journals and conferences in Artificial Intelligence and related areas.