

# A Process for Human Resource Performance Evaluation Using Computational Intelligence: An Approach Using a Combination of Rule-Based Classifiers and Supervised Learning Algorithms

ANDERSON SILVA DE OLIVEIRA GÓES<sup>1</sup> AND ROBERTO CÉLIO LIMÃO DE OLIVEIRA<sup>1</sup>

Technology Institute, Federal University of Para, Belém 66075110, Brazil

Corresponding author: Anderson Silva de Oliveira Góes (agoes@ufpa.br)

**ABSTRACT** This paper proposes a process for human resource performance evaluation using computational intelligence techniques. The human resource (or employee's) performance evaluation is essentially a regular assessment and review of an employee's performance on the job. This evaluation can be performed in different ways, depending on the kind of job of the employee and on the company's politics or business area. The process proposed on this research combines Fuzzy logic, text sentiment analysis and supervised learning classification techniques, such as a multi layer perceptron artificial neural network, decision tree algorithms and naïve bayes into ensemble classifiers, in an attempt to provide a fair evaluation process, minimizing or even eliminating common problems caused by simple objective or subjective approaches. The data provided for this research was originated from several evaluations applied in two Brazilians institutions. Simulation results shows consistence on the data generated by this proposed process, indicating a good perspective for applications on companies of most business areas.

**INDEX TERMS** Ensemble classifiers, fuzzy logic, human resource management, human resource performance evaluation, decision trees, supervised learning.

## I. INTRODUCTION

Human resource management (HRM) is defined as the process that organize, manage and leads a team [1]. Always an important process, the HRM has gone through a major transformation in form and function in the past 3 decades [2]. In nowadays, such a competence is considered a source of sustained competitive advantage for organizations operating in a global economy [3]. The importance of this process is felt in both public and private companies and, considering companies that work with projects as an example, the importance of this area becomes extremely critical. Not only the HRM is considered key to the success or not of a given project [4], but such kind of company also has to have a special care about the evaluation of senior project managers, a position which often lacks of a proper tool for evaluation [5].

Given this context, it becomes clear that HRM implies directly in competitive advantage to a particular company [6],

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Imran Tariq<sup>1</sup>.

[7], [8] indicating if the business will have a good chance of prosperity or if the current performance is just good enough to survive, or, in more critical cases, not even that [9].

The HRM is composed by several competences. An evaluation of these competences seeks to identify if an employee has all the required knowledge for a determined job position [10].

This paper presents an approach related to the way how each employee performs in these competences, which is a process known as "Human resource performance evaluation".

In such a process, it is vital for success, to establish the right elements to be evaluated, in order to ensure a model that is reliable enough to control this process. It is also important to understand that, most likely, each kind of company may have different needs to be satisfied by a model of this nature.

This paper presents an approach using computational intelligence techniques as an alternative to deal with the common variations on this kind of process, combining ruled based classifications, such as regular crisp functions and fuzzy logic scripts, as well as supervised learning algorithms based

classifiers, such as decision trees, naïve bayes and artificial neural networks. The implemented process also uses text sentiment analysis and classification committees (ensemble classifiers) combining all the approaches.

The proposed process seeks to achieve an original contribution using an ensemble classifiers strategy. Using this approach, the process attempts to combining the strengths of each used methodology and, at the same time, minimize their weaknesses, making the employee's final evaluation result as fair as possible.

The sub- topic A of this introduction presents the bibliographic review and related works considered on the areas presented on this research. Following, the paper is composed by another 5 sections. Section 2 presents an overview of the human resource performance evaluation process, applied to a basic business context. Section 3 explains the origin of the data used on this research. Section 4 presents the process applied on the research. Section 5 presents a look over the results compiled so far, after the conducted experiments. Finally, section 6 presents a conclusion of the study.

#### A. BIBLIOGRAPHIC REVIEW AND RELATED WORKS

With the understanding that this research consists in a multidisciplinary work, the search for related works needs to consider a reasonable number of subjects. In order to cover most of the knowledge used in this paper, the bibliographic review considered the following areas:

- Human resource management.
- Employee's performance evaluation.
- Fuzzy logic applied to employee's performance evaluation.
- Other computational intelligence techniques applied to employee's performance evaluation.
- Sentiment Analysis.

Each of these areas consists in a field worth of a complex study only by itself. With that in mind, a good number of researches can be found in each subject. Some of them, however, present some concepts that can be used to support some of the ideas presented in the process to be proposed on this paper and, therefore, should be mentioned in this section.

The work of [11] presents a study about a capability evaluation system for a company, that is based on a staff competency model. Such a concept is interesting, because it shows how a capability of a company is directly tied to the competency of their workers. That idea is one important aspect for considering the relevance of studies in the employee's performance evaluation field.

The fuzzy logic study needs to consider a few different aspects. Being that such a technique works depending of rules and that these rules require at least reasonable knowledge of the problem domain, the search for related work in this area needed to be specifically focused in the employee's performance evaluation domain or human resource management.

The available literature regarding the use of fuzzy logic in a such a context is particularly rich. There are papers, such as

[12] and [2], that propose the use of a fuzzy system for the selection of applicants to a specific position in an institution. On a similar field, the work of [13] presents a strategy for recommending human resources for project leaders using adaptive fuzzy logic. The work of [14], by other hand, proposes a strategy for HRM classification combining Fuzzy sets, naïve bayes, decision tree and data mining.

The paper [9], presented by the authors at an earlier time, shows the preparation of two fuzzy scripts, used to apply an employee's performance evaluation simple process and, later on, compare the results to a simple crisp evaluation process. The fuzzy logic scripts have an important role in this process, dealing in a special way with the self-evaluation factors and considering different rules for each job position. This paper expands the work presented on [9], creating more fuzzy scripts, with a more complex set of rules, integrating these scripts to an information system and later combining their exit to other classification strategies.

The fuzzy approach, due to its own nature, requires a reasonable knowledge of the problem being treated. There are other approaches, however, that can be applied to such a context basically by extracting standards from a coherent database.

The work of [15], for instance, presents an approach using genetic algorithms to dynamically distribute human resources according to their age. The work of [16] proposes an approach using decision tree algorithms such as random forest, combining then with Support Vector Machine (SVM) and Naive Bayes as an approach to predict employees churn with data mining.

The sentiment analysis bibliographic review requires a look into researches involving data mining and text classification, in addition to works related to the use of these techniques on this specific context.

The work of [17] presents a look into employee's performance review using sentiment analysis. The research studies a big amount of text originated from peer analysis and proposes a way to identify certain aspects that would be difficult to appear on a superficial analysis. The work of [18] and [19] also present aspects regarding the use of text analysis in a human resource management context. [18] presents an approach for team member selections based on contextual sentiment closeness. The work of [19] presents an approach to detect subjectivity on teacher's performance through text analysis. The work of [20], by other hand, presents a strategy for text classification that adopts a Bagging ensemble classifiers strategy based on a genetic algorithm.

The related works present interesting aspects to be considered on this research proving aspects such as the viability of the use of Fuzzy logic on human resource performance evaluation and the relevance in the use of sentiment analysis as a way of identify aspects that would be very difficultly observed considering only a simple objective evaluation. In addition, the use of other computational intelligence strategies such as decision trees and other data mining techniques can also be found in the researched bibliography.

Even so the individual use of these strategies can be quite frequently found in literature, the research for a process that combines these on a same evaluation shows a considerable gap. It is also important to point that, most of the approaches found in literature tend to deal with the strengths and weakness of each individually chosen technique.

This paper presents an original human performance evaluation process that aims to combine the most positive aspects of each used strategy, such as:

- The flexibility of the fuzzy logic scripts.
- The ability of knowledge extraction of a decision tree algorithm.
- The complexity of a text sentiment analysis, providing bigger safety to the process.
- The use of ensemble classifiers as a strategy for not trusting on a single classify.

The next section will present a basic overview of a human resource performance evaluation process, in other for a better understanding of the used methodology.

## II. HUMAN RESOURCE PERFORMANCE EVALUATION

The human resource performance evaluation is a complex process, that has a considerable number of challenges, difficulties, methodologies and, if well executed, rewards. This research focus on the main aspect of such a process, which seeks to compare the original expectations regarding an employee's performance and the results that were actually produced.

There are quite a few possible variations in a process of such a kind. A few aspects, however, can be highlighted as being of extreme importance to any kind of company.

First, the necessity to understand what are the expectations of both company and employees regarding the applying of such a process. The company needs to know how to process the information collected on the evaluation and the employees need to be aware of the possible benefits and consequences of such a process.

It is also necessary to understand that in a process so much sensitive to the company management and to their employees, a reasonable number of difficulties are expected to be found. With that in mind, it is also important for all parties involved in the process, to know exactly what kind of difficulties are to be expected.

The expectations that a company may have about this kind of evaluation should suffer a direct influence from the business main characteristics, such as size, business area, sector (private or public) and other factors. The work of [11] points out to the fact that an employee's performance evaluation should not consider only qualitative and capacity aspects. In addition of those, there must also be evaluated the employee's closeness to the mission, view and main values defended by the company. In addition, [11] says that a basic evaluation model should consider a combination of factors, taking in consideration the kind of position occupied by the employee in the organization.

The idea of customized evaluations for different job positions is particularly important for this research, in order to better establish a comparison between the original expectation of an employee's performance and the work that was actually delivered.

An analysis on the company's expectations shows that, most of the time, there is a consolidated idea of the necessity of having capable human resources, in other to achieve the level of competitiveness that the market requires [5].

The look on the employee's expectations, however, is a bit more complex. In an analysis about the private sector, for instance, such a motivation evolves, obviously, the own maintenance in the company. A competitive environment, after all, usually doesn't allow the continuity of employees that are not well evaluated. Good evaluations, by other hand, are usually linked to better opportunities such as job promotions and salary increase.

The public sector has a few similar aspects and other very different, in a comparison with the private one. Employees in high job positions, usually linked to management functions, most of the times need to show results to their superiors or, in some cases, directly to the population. In these cases, even so the public sector does not usually seek for financial profit, these employees still need to constantly prove themselves as being worth of the position currently occupied and, as such, have similar expectations regarding to this process as their colleagues of the private sector.

The reality for those occupying more technical positions in the public sector, however, is different. The operational positions in this sector are usually more stable to external interferences. With the risk of resignation being considerably lower in comparison to the private sector, the employees in this kind of position tend to look to a performance evaluation process as a way of seeking for improvement in their conditions (financial bonus for instance).

On the next sub-sections, this paper will present methodologies frequently used on human resource evaluation process's, as well as the most common difficulties faced by those who participate in then.

### A. HUMAN RESOURCE EVALUATION METHODOLOGIES

A good performance evaluation process usually works with factors and aspects that are of big relevance for the core of the company's business, in an attempt to make the process as fair (and correct) as possible. Let us then see a brief explanation and comments about some of the most common techniques.

A very common alternative used in a process of this kind is to link the employee's productivity to performance goal indicators. In that case, if the worker fails to reach the defined goal, actions may be taken by the company. Actions that, in this case, may involve a position change, considering that one of the goals of the workers performance evaluation is to provide grounds for the decision makers to put the right people in the correct positions [21].

Another possible alternative for such a process is the so called "pair evaluation" or "peer review". In this scenario,

the employee's performance is evaluated by the colleagues that work close to him or her [17] (possibly on the same hierarchical level).

Another very common resource used in this kind of process is the self- evaluation (or self- assessment) strategy. In that case, as the own name states, the employees have the opportunity to appraise their own performance [22].

The probably most frequent and simple form of evaluation consists on a process where the employee is evaluated by his or hers hierarchical superior. In that case, the superior (usually some kind of manager), evaluates the employees there are under his responsibility. It is important to remember that, in that case, the manager is also evaluated by a superior [23].

In many cases, despite the strategy chosen for this evaluation process, the employees may understand that they are not being appraised in a way that is the fairest. This kind of complain is usually linked to arguments such as the difficulty for managers to evaluate a big team of workers or even a possible difficulty in understanding technical aspects of the employee's position.

Another essential aspect of the employee's performance evaluation process is the definition of the evaluation factors. Desirably, the evaluation criteria should be decided considering the particularities of each job, and even period in which the evaluation takes place, being that an intrinsic part of the evaluation [21].

The model applied in this paper presents aspects that are important to basically any kind of business, such as punctuality, professional posture and commitment. That characteristic, contributes for the possibility of application of such a system on a larger number of business area.

Once the evaluation factors have being decided, it is necessary to analyze the best way to rate them. A simple way to do that is to grade then between a scale (0 to 10, for instance). After all the factors are rated for a specific employee, a system may simply calculate the average between then and that will be the employee's final performance result. Such an approach is adequate for a lot of cases. After all, it is easy to be implemented and comprehended by everyone.

In some cases, however, the application of a simple crisp logic such as that may cause problems on the evaluation process. Consider aspects in which the employee obtains extreme evaluations (positive or negative) in one factor. A 10 or a 0 grade in one factor applied to a simple mathematical formula may unbalance the worker's evaluation, causing a result that, most likely, will not be as fair as intended.

The scale method with pre – defined concepts instead of numerical grades is also an interesting alternative [24]. Such a methodology is very simple to understand and apply, with that probably being the main positive characteristic of such alternative. This strategy, however, also has problems that most of time have negative influences on the evaluation's final result.

One very common issue in such a methodology is the very own definition of the evaluation concepts. That being said,

**TABLE 1. Common rules applied for a scale evaluation method.**

Grade	Description
Insufficient	The employee performed underneath the original expectations.
Regular	The employee performed his basically expected obligations.
Good	The employee went beyond his obligations and exceeded the original expectations.
Excellent	The employee went far beyond his obligations. It is a grade applied to the workers, who are, usually an example to the rest of the company.

if a company chooses to work with this scale, it is essential to clarify to those who are evaluating and being evaluated what exactly is an insufficient, regular, good or excellent performance. These definitions may vary from one company to another, but, generally, the table 1 rules are most commonly applied.

Considering the risks and strengths of the methodologies presented in this paper, this research proposes on the next sections a process that uses computational intelligence techniques as a way of preserve the best aspects of each of them, and, in the same time, minimize their weakness.

### III. DATABASE

The database applied on this research uses data originated from two Brazilian institutions from different business areas (tourism and information technology).

The original collected data consists in close to 2000 records. These records used data provided by different kind of processes, which are going to be referenced here as methodologies 1, 2, 3 and 4.

The methodology 1 had 3 kind of appraisals. These were the self, superior and pair evaluation. Methodology 2 applied only the self and superior evaluation. The methodology 3 also considered the self and superior evaluation results using, using, however, the self-evaluation factors only as a reference. The methodology 4 also uses only the self and superior evaluation factors. The difference in this methodology is that the hierarchical superior only becomes aware of the employee's self- evaluation on a later moment (after evaluating the employee).

The analysis performed on the data seeked out to identify a set with enough consistence and coherence to enable the Knowledge Discovery in Database process (KDD) [25]. Such a process was a crucial step in order to apply the computational intelligence ruled based strategies and for a reasonable performance of the induction algorithms.

**TABLE 2. Frequent identified problems in the appraisal process.**

Problem	Description
1	An appraisal focused only in the hierarchical superior evaluation pleases the evaluators and the company. The evaluated, however, believe that this process is not fair.
2	An appraisal that gives big strength to the self-evaluation pleases the evaluated. Such a process, however, often brings to unrealistic evaluations, that displeases the company and the evaluators.
3	An evaluation focused in results is generally a good approach, but can only be applied to a reduced number of companies.
4	The company often finds difficult to interpret the evaluation results and actually translate then into useful actions.

The KDD and Data Mining have a big correlation in their tasks. For that reason, a lot of authors consider both as part of a same process, which is basically formed by the following steps:

1. Problem identification.
2. Pre-Processing.
3. Patter Extraction.
4. Post-Processing.
5. Knowledge application.

**A. PROBLEM IDENTIFICATION**

The KDD first step took some work. Even so the problem domain was well established, as well as the relevance of the human resource appraisal process. Generally speaking, the guarantee of fairness in the process proved to be an object of main concern for the main actors involved on it.

To understand what exactly this fairness would be, however, the very own identification of these group of actors needed to be clarified. In order to do that, after research and conversations with area specialists, the parties involved in the process were divided in 3 categories, being that the **Evaluators**, the **Evaluated** and the **Company** itself.

The Evaluators group is formed by managers, supervisors, coordinators, directors and other management positions. The Evaluated group is composed by most of the company’s employees. It represents those that have their performance appraised. The Company group is formed mostly by human resource managers, company directors, CEOs and other members of the company’s board. Table 2 presents the most critical problems identified after the conclusion of this task.

**B. PRE-PROCESSING**

The original collected data came from companies that used similar, yet not completely equal processes. The first step was to unify the evaluation factors into a coherent and reusable set. Such a task was concluded building a database composed

**TABLE 3. Job positions descriptions.**

E.F	Position Name	Description
1	<b>Junior Technician</b>	Position that considers those who are in lower positions in the company’s hierarchy. Depending on the company, may be considered from this category trainees or workers in the early stages of a career plan.
	<b>Full Technician</b>	Position that involves employees with a larger experience and usually with a specific formation for that position. In an eventual career plan, these workers would occupy a considerable position.
	<b>Senior Technician</b>	Position that involves the company’s top technicians. These workers usually have big experience on their field and occupy the top positions in an eventual career plan.
	<b>Manager</b>	Position that involves the managers, coordinators, supervisors and other management positions of the company.

of 21 attributes, which are going to be briefly explained on the next paragraphs.

The first attribute of this database is called “jobPosition”. As the name states, such an attribute indicates what kind of job the employee performs in the company. Obviously, the companies have different jobs and the idea of consider then all during the build of this database would not only be exhausting, but would also contribute for a process with few generalizations and, therefore, harder to be applied in a bigger number of companies.

That way, inspired on a hierarchy commonly used on project oriented companies, the employee’s positions were divided into four groups, which are explained on table 3.

The attributes 2 to 10 present the superior hierarchical evaluation factors. These factors, described on tables 4 and 5, were chosen considering a criteria that took in consideration the original collected data, literature review and experience from the author, in a way that would result on a dataset possible to be used on a human resource evaluation process that could be applied on a varied types of companies and institutions.

The factors 11 to 19 represent attributes similar to the factors 2 to 10, considering, however, the self-evaluation attributes. The attribute number 20 represents the text

TABLE 4. Evaluation factors 2 to 6 descriptions.

E.F	Name	Description
2	<b>Commitment</b>	Evaluates whether the employee has committed to the values, actions and defined objectives of the institution, including booth individual and group activities.
3	<b>Problem Identification</b>	Evaluates the employee's ability to identify non-conformities in technical or strategic levels.
4	<b>Task Performance</b>	Evaluates the speed and quality in which the employee performs the tasks assigned to he or she.
5	<b>Interpersonal Relationship</b>	Evaluates the manner in which the employee relates to the co-workers and superiors.
6	<b>Professional Posture</b>	Evaluates the employee's ability to comply with the rules established by the company, maintaining the proper etiquette required in a professional environment.

regarding the justification of the hierarchical superior to the provided evaluation. Such a text will inform why the manager is evaluating the employee under his responsibility like that. The attribute number 21 is similar to the 20<sup>th</sup>, this time, however, representing the self-evaluation justification. Once the pre-processing stage was done, the resulting database was ready for analysis regarding to the kdd's pattern extraction stage.

C. PATTERN EXTRACTION

The pattern extraction played a very important role on this research, especially regarding the build of the fuzzy scripts, considering that these require the knowledge of the area of expertise.

The first analysis, which was made using the machine learning workbench tool Weka [27], considered the appraisals final results, as seen on figure 1.

The figure shows a clear concentration of results around good performances. Such a concentration is inconvenient from the data mining point of view and presents a few aspects that can be expected on a real work environment. Such an analysis can be better understood after looking back into the scale definitions.

The number of insufficient evaluations is relatively low, as expected. Considering that an "insufficient" worker is someone under the expectations, it stands to reason that not many performance evaluations should present this final

TABLE 5. Evaluation factors 7 to 10 descriptions.

E.F	Name	Description
7	<b>Initiative</b>	Evaluates the employee's ability to put itself available to solve tasks without advance request.
8	<b>System Overview</b>	Evaluates the employee's ability to make itself aware of the events and processes that affect the company as a whole, not limited to his or hers specific job function.
9	<b>Punctuality</b>	Evaluates the employee's ability of respect the institution's work hours.
10	<b>Attendance</b>	Evaluates the employee's ability to attend on work days.

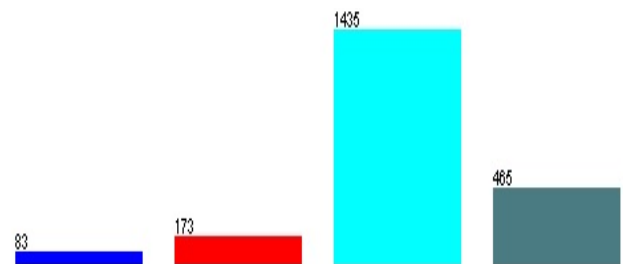


FIGURE 1. Original performance appraisal results (dark blue = insufficient, red = regular, light blue = good, gray = excellent).

grade. The number of excellent results, however, presents a considerably high concentration, relatively speaking. Opposing to the insufficient grade, an excellent evaluation presents the other extreme of the scale, representing a work that serves as example to the colleagues as someone very above the performance average. The regular results, by other hand, present a considerably lower concentration then expected considering the scale definition.

One very reasonable explanation for such a concentration around good performances results is the very own performance evaluation methodology. In a simple crisp evaluation, considering an average between the evaluation factors individual grades, especially if these do not have different weights, higher and lower individual performances in each factor tend to cancel each other, converging to a final good evaluation.

The analysis of a few individual performance evaluation factors, as seen in figures 2, 3 seems to endorse such a theory. These pictures show the results, respectively in the "commitment" and "professional posture" evaluations.

The commitment evaluation, for example, demonstrates a concentration of most records with 100% achievement, meaning an excellent concept in this factor. The professional



**TABLE 8.** Proposed indicators for a fair human resource evaluation process.

Indicator	Description
1	The general final results concentration should be between Good and Regular results.
2	The appraisals should have variation approaches depending on the employee's job position.
3	The self-evaluation factors should be considered on the evaluation formula, but in a way that the final result does not get contaminated by then.
4	The evaluation performance needs to be confirmed and validated, to avoid discrepancies.

#### IV. APPLIED METHODOLOGY

The development of the proposed methodology needed to consider objective indicators in order to point, by looking at data, what would in fact be a fair and reliable human performance evaluation result. After a careful look into the original data and the first classification's performance, the 4 indicators described on table 8 were proposed.

The fact that the supervised learning classifiers could not achieve a perfect result and the fuzzy scripts did not present an adequate performance for all the records shows the importance of the fourth indicator (described on table 8), which presents the need of some kind of validation for the employee's evaluation.

The attempt to make such a validation was first made applying sentiment analysis on the performance evaluation text justifications and, later, comparing the result with the evaluation factors performances. Even so this strategy showed promising results, however, it has two main problems. First, language and context variations make it for a very hard implementation of this resource. Second, a result discrepancy with the evaluation performance factors doesn't exactly shows where the problem is (in the factors performance or in the justification).

These attempts took the research for the development of such a performance evaluation process on the direction to use a committee (or ensemble) of classifiers strategy [28]. This strategy is encouraged in problems that do not want to rely on one single classifier and, for that matter, perform a strategy that seeks for a second and third or more opinions, in order to combine them into a final most appropriated opinion [29].

As shown in figure 5, the process uses variations to perform committee classifications with simple crisp and fuzzy logic, supervised learning classifiers (Multi layer perceptron, Decision Tree C4.5, Cart, Random Forest and Naïve Bayes) and text sentiment analysis. The next sub-sections will provide more details about the implementation of each part of the process.

#### A. SENTIMENT ANALYSIS

The first lane presented on figure 5 shows the processes necessary to assess the sentiment analysis regarding the text justification provided for an employee's performance evaluation. The sentiment analysis strategy applied on this research utilizes a traditional methodology, which judges the sentiment polarity building an emotional dictionary [30]. The process of building such a dictionary was based on a standard lexicon sentiment analysis database, providing a positive and negative score to each word presented on the performance evaluation justifications. The words with higher positive analysis would be most likely linked to good and excellent performances and negative ones to insufficient and regular grades.

The process called "calculate justification polarity", is responsible for identify the sentiment regarding the text presented and then link this result to the most appropriated evaluation grade. This process involves a technique called "Bag of Words" and the "Boosting" algorithm as a form of apply an ensemble of classifiers, as presented on figure 6.

The sub-process (1a) basically receives a string formed by the evaluation's justification and proceeds to next step (1b), when the string is split into a vector data structure, according to the bag of words [31] technique.

The step (1c) updates the general words polarity table. This task considers the entire evaluation database and calculates the frequency in which each word appears on a positive (excellent or good) and negative (regular or insufficient) evaluation, therefore, assigning a positive and negative probability score for each word on the database.

The steps (1d) and (1e) comprehend the text classification utilizing the committee of classifiers algorithm know as Boosting [32].

On this research, the individual classifiers forming this ensemble are the random forest algorithm, multi-layer perceptron artificial neural network and naïve bayes. Table 9 presents the basic steps for the implement of this algorithm.

Once the step (1d) completes the sentiment analysis classification ensemble, step (1e) calculates the positive and negative score of the presented performance evaluation text, formed by the word vector already created on step (1b). The polarity of the vector is then classified by the classification ensemble and registered as the sentiment analysis employee's performance evaluation result, on step (1f), and registered on a variable for further use on the continuity of the process, on step 2.

#### B. RULED CLASSIFICATIONS

The next step (step 3, in the Ruled Classifications lane), into this human resource evaluation process, as shown in figure 5, is the application of the committee of classifiers 1 and 2, responsible for define a final performance evaluation grade for the employee, after analyzing the hierarchical superior and self-evaluation results.



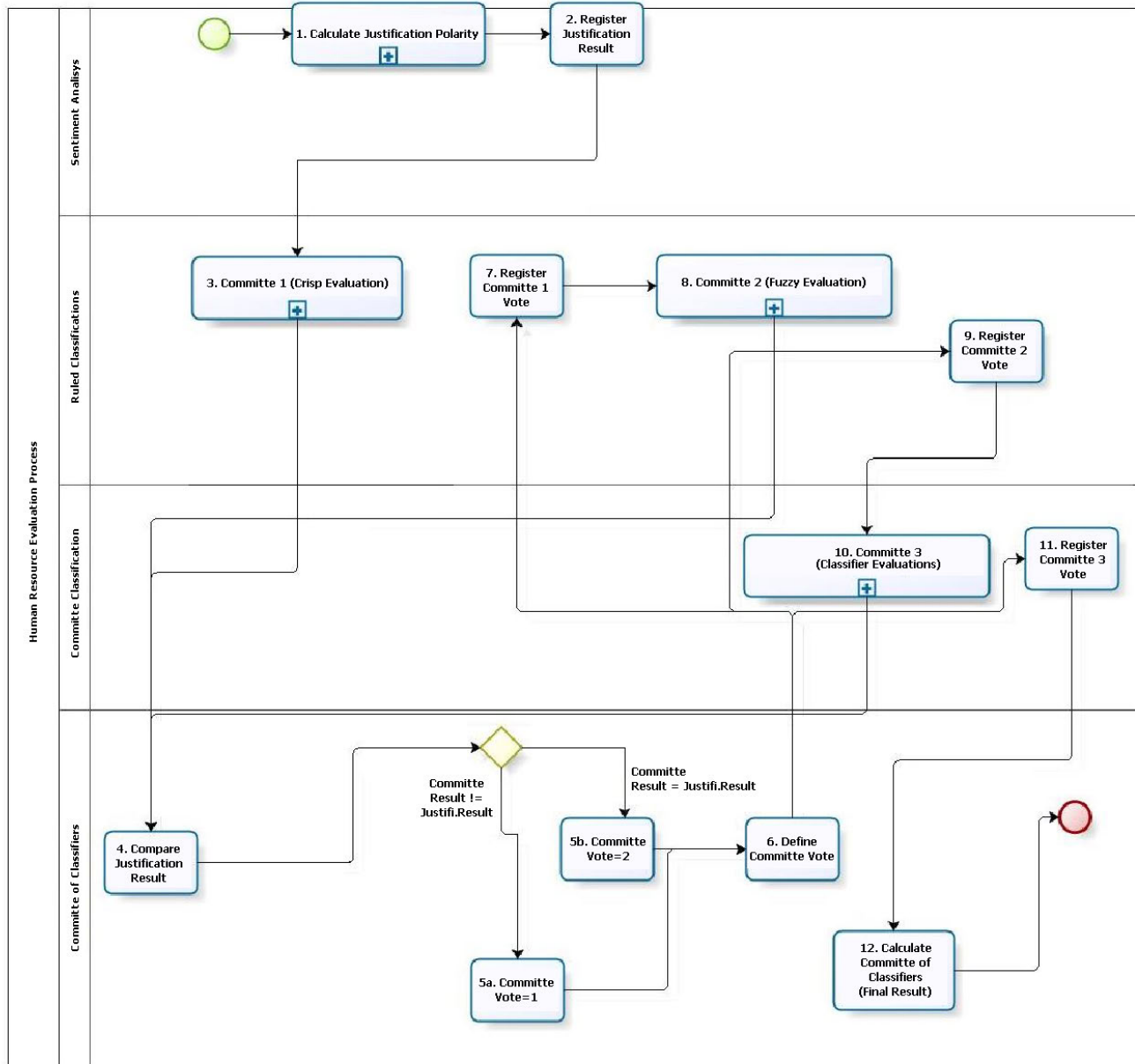


FIGURE 5. Human performance evaluation process. Source: Author. provided by Bizagi [33].

The committee 1, represented by process 3 is composed by the basic mathematical simple crisp evaluations. These evaluations apply a weighted averaged function to the sum of superior and self-evaluation inputs, composed by 4 different averages, as described on table 10.

The committee 1 also applies a majority vote strategy, with each of these functions voting on a resultant class. This committee, as the table 10 states very clearly, is the most impacted by the defined weights of the self-evaluations.

The step 4, in the Committee of classifiers lane, applies a result comparison between this committee and the justification sentiment analysis result. If they have different results, the exit class provided by the classifier receives a vote of

value 1 for being the employees final evaluation. If the results are equal, however, the class receives a vote of value 2. This voting process is represented by the steps 5a and 5b of the process, with the committee voting being registered on steps 6 and 7.

The step 8 represents the Committee 2, which is formed by the Fuzzy logic scripts. These scripts are based on a set of rules in an If- Then shape, which forms a knowledge base and have a few words with values represented from membership functions or fuzzy sets [34].

This characteristic makes this committee the part of the process that most drastically differs the performance evaluation logic according to the job position occupied by the employee.

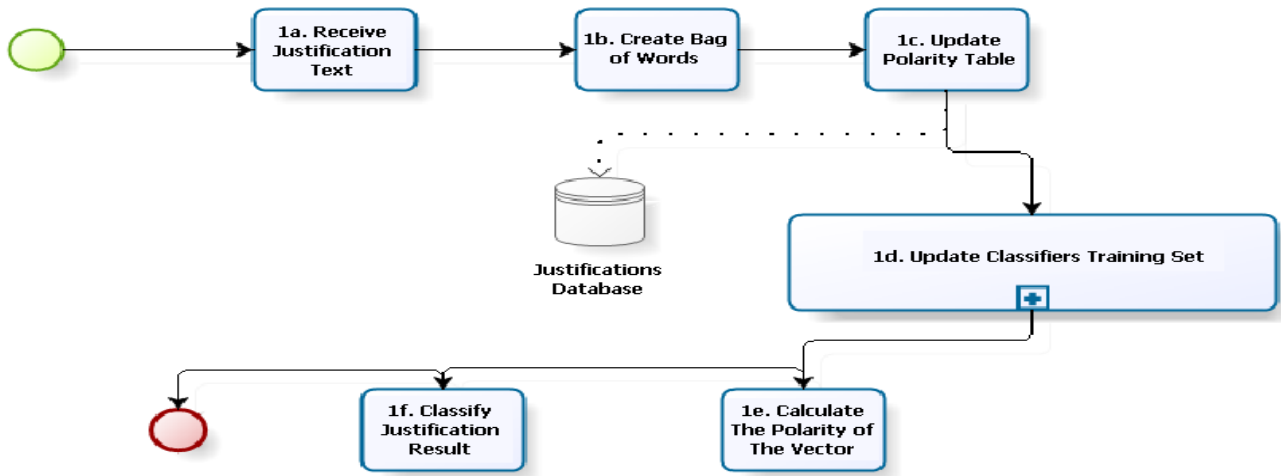


FIGURE 6. Sentiment analysis sub-processes. Source: Author. provided by Bizagi [33].

TABLE 9. Boosting algorithm implementation.

```
// Algorithm a1 – Random Forest
// Algorithm a2 – Multi Layer Perceptron
// Algorithm a3 – Naïve Bayes
1. Receive original training set T.
2. Specify the fraction F of each new training subset to be created from T.
3. Create data subset t1 from a fraction F of T without replacement.
4. Call the algorithm a1 and train with t1 to create the classifier c1.
5. Create dataset t2 with half of the instances being correctly classified by c1 and the other half misclassified.
    5.1 Flip a coin.
    5.2 If head, select samples from T and present them to c1. Add to t2 the first to be incorrectly classified.
    5.3 If tail, select samples from T and present them to c1. Add to t2 the first to be correctly classified.
    5.4 Continue flipping coins until t2 reaches F size.
6. Call a2 and train the classifier c2 with t2.
7. Create t3 by selecting those records for which c1 and c2 disagree.
8. Call a3 and train the classifier c3 with t3.
9. Receive a Test instance X.
10. Classify X with c1 and c2. If they agree, considers this it is final classification.
11. If they disagree, classify X with c3 and considers this it is final classification.
```

The Fuzzy evaluation performance applied on this particular research expands the work presented on [9], utilizing of 4 different scripts, built with the JFuzzyLogic framework [35].

TABLE 10. Crisp evaluation weights.

Function	Description
User Defined	The user defines the corresponding weights for the hierarchical superior and self-evaluation results.
1x0	The superior hierarchical results possess weight 1 and the self-evaluation results possess weight 0.
1x1	Both evaluation results possess weight 1.
2x1	The superior hierarchical results possess weight 2 and the self-evaluation results possess weight 1.
3x1	The superior hierarchical results possess weight 3 and the self-evaluation results possess weight 1.

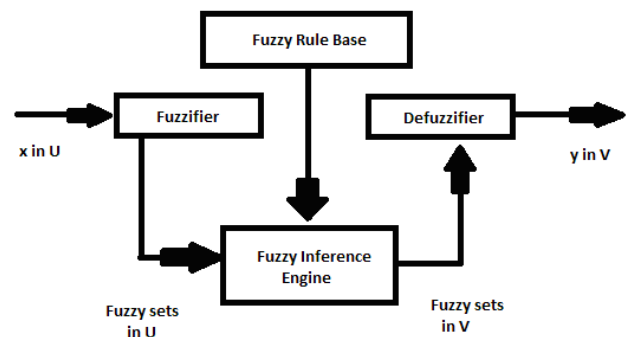


FIGURE 7. Basic fuzzy system of Mandani type.

The scripts implement a Fuzzy system based on a Mamdani type [36], and have their basic structure presented on figure 7 and table 11.

The input and output variables were the same in all the four scripts. The importance that each of them carries on the set of rules, however, had differences on each of them.

TABLE 11. Fuzzy scripts structure.

Field	Description
Input variables	The 18 numerical evaluation factors.
Output variables	The employee’s final evaluation.
Membership functions	Defines how to Fuzzify the input variables and Defuzzify the output.
Rule blocks	Blocks of If-Then rules that define the behavior of the script, and ultimately, the employee’s final grade.

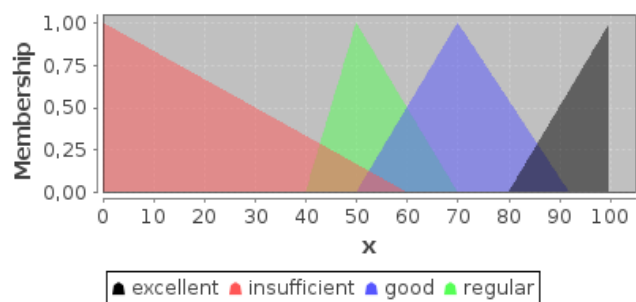


FIGURE 8. Membership function for an input value.

The “Fuzzy script 1” introduced on this process the idea of considering different weights for the evaluation factors, something that was not applied to the crisp evaluation. That way, the input variables were divided into 2 groups, with the group 1 variables having a higher importance on the set of rules than the ones on the group 2. In this script, the superior hierarchical evaluation of the factors “commitment” and “task performance” were part of the group 1 and the other hierarchical superior evaluated factors were part of the group 2. The self-evaluation inputs were not considered on this script.

The variables membership functions have the same standard values for all the inputs in scripts 1 and 2, mapping the numerical value, into a fuzzy value corresponding to an insufficient, regular, good or excellent grade, such as presented on figure 8. A similar membership function was also adopted for defuzzify the output value.

The script 1 contains the smallest number of rules (28 total) between the 4 scripts produced. These rules present the behavior expected from the fuzzy system in case the majority of groups 1 and 2 inputs have different grades.

The “Fuzzy script 2” was the first evaluation strategy applied on this process to explicit consider the difference regarding the employee’s job position, as well as consider part of the self-evaluation inputs. The result was a considerably larger number of If-Else rules, being a total of 144.

The differential treatment for the employees depending on the job position also caused a change related to the

TABLE 12. Scripts membership function comparison (commitment, punctuality and professional posture factors).

Term	Scripts 1 and 2	Script 3 and 4
Insufficient	(0, 1) (60, 0)	(0, 1) (60, 0)
Regular	(40, 0) (50,1) (70,0)	(40, 0) (65,1) (100,0)
Good	(50, 0) (70,1) (92,0)	(80, 0) (85,1) (90,0)
Excellent	(80, 0) (100, 1)	(88, 0) (100, 1)

variable groups that were used on Fuzzy Script 1. That way, the employees occupying Senior technician and manager positions had their group 1 factors changed to the “problem identification” and “systemic view” evaluation factors. The junior and full technician job positions did not had alterations in these groups. The self-evaluation factors related to the group 1 inputs formed the new evaluation factors group, called group 3.

The Fuzzy Script 2 introduced the self-evaluation group 3 attributes to the rules as a way of analyze how the employee perceives his own performance on the evaluation factors that have the highest weight on his evaluation.

The “Fuzzy script 3” was created utilizing a different approach, modifying the fuzzify blocks of the entries related to “commitment”, “punctuality” and “professional posture”, as shown on table 12, with 0 representing no pertinence and 1 total pertinence to the class.

The modify in the membership functions of these classes was experimented after research with specialists and the author’s own experience pointed these factors to often (implicitly) being considered in a more extreme way during the evaluations.

The inclusion of the third Group made the second and third scripts lightly sensitive to the self-evaluation factors, considering the ones with the highest weight on the rule set. The doubt remained at this point, however, of how the Fuzzy logic would behave considering all the self-evaluation factors and what kind of impact this would have on an employee’s final grade generated by this logic. To answer this question, the “Fuzzy script 4” was created.

The fourth fuzzy script considered as the group 4 of attributes, the self-evaluation factors related to the group 2 superior evaluation factors. The result was a script with a total of 170 rules. Figures 9 and 10 present the surface of the fuzzy scripts 2 and 4, with the Job Position and Commitment input values and the final evaluation output.

After the employee’s performance is calculated by the Fuzzy scripts, the most voted class between their results is registered as the fuzzy committee vote. The figure 5 diagram’s 4,5 and 6 steps are once again applied to compare the committee’s class exit to the sentiment analysis and, after, register the committee 2 final vote on step 9 (similar to the process explained for committee 1). Next, the process proceeds to the next step (10), comprehending the third ensemble classify, which will be explained on the next section.

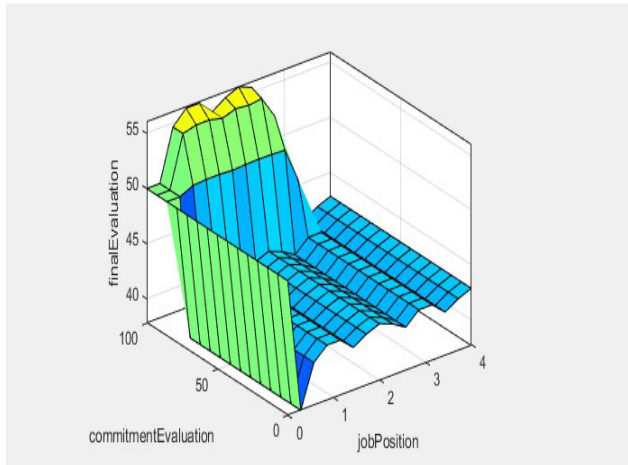


FIGURE 9. Fuzzy script 2 surface.

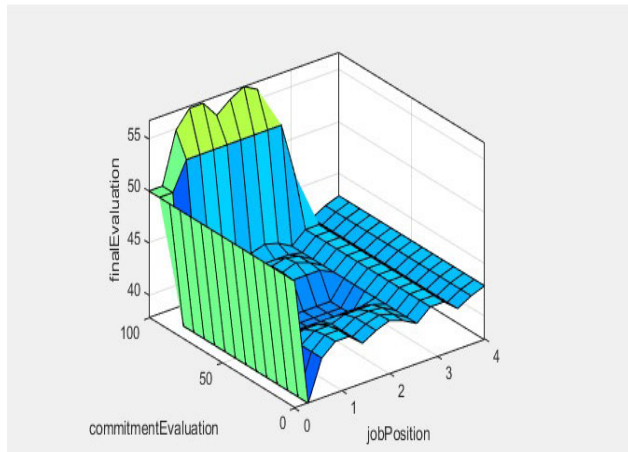


FIGURE 10. Fuzzy script 4 surface.

**C. COMMITTEE CLASSIFICATIONS**

The Committee classification’s lane comprehend the execution of an ensemble classifier formed by the supervised learning algorithms C4.5, Cart and Random Forest Decision Trees, Multi layer perceptron and Naïve bayes.

This committee is built using the ensemble algorithm know as Bagging [37], as presented on table 13. After the algorithm is performed, each classifier votes on the corresponding exit class, with the highest one voted being chosen in the end of the process. Once again, like with the other two committees, the committee 3 final evaluation result is compared to the justification’s sentiment analysis and the committee’s vote is registered.

After the third committee vote is registered, the process moves on to step number 12. This final last step of the process is better described on the next sub-section.

**D. THE COMMITTEE OF CLASSIFIERS AND AN EVALUATION’S FINAL RESULT EXAMPLE**

The final step on this performance evaluation process basically comprehends the election of the highest voted exit class as the employee’s final evaluation, as exemplified on

TABLE 13. Bagging algorithm implementation.

```
// Algorithm a1 – C4.5
// Algorithm a2 – Random Forest
// Algorithm a3 – Multi Layer Perceptron
// Algorithm a4 – Naïve Bayes
// Algorithm a5 – Cart
1. Receive original training set T.
2. Specify the number I of interactions.
3. Specify the fraction F of each new training set to be created from T.
4. Do it from j=1 to I
    4.1 Create a t1 training set from T, randomizing F percent from T.
    4.2 Call algorithm aj with tj and receive the classifier cj.
    4.3 Add cj to ensemble E.
5. Foreach classifier from E do
    5.1 Classify unhandled instance X.
    5.2 Compute l vote for the classified class.
6. Obtain total of votes received by each class.
7. Choose the class with the highest vote as the final classification
```

TABLE 14. Employee’s evaluation sample (hierarchical superior’s results).

Input	Value
Job Position	Senior Technician
Commitment	85
Problem identification	80
Task performance	88
Interpersonal relationship	70
Professional posture	90
Initiative	90
Systemic overview	60
Punctuality	80
Attendance	100
Superior evaluation Weight (on Crisp evaluation)	2

tables 14, 15, 16 and 17, which presents an evaluation sample of a random employee.

As table 14 indicates, this evaluation sample belongs to a senior technician employee. He or she was mostly well evaluated by the hierarchical superior, with the exception of the systemic overview factor. Another important observation is the fact that only one factor achieved a 100% performance. The self-evaluation results on table 15, as expected, shows a better performance. With no performance bellow 80% and 4 below 100%, that actually represents a somewhat rigorous self-analysis.

The results presented on tables 16 and 17 show the differential analysis performed by each technique. The well balanced (and probably coherent), self-evaluation, led to an unanimous result in the crisp committee, which voted for a

**TABLE 15. Employee’s evaluation sample (self-evaluation results).**

Input	Value
Commitment	100
Problem identification	100
Task performance	100
Interpersonal relationship	90
Professional posture	100
Initiative	88
Systemic overview	80
Punctuality	90
Attendance	100
Self-evaluation Weight (on Crisp evaluation)	1

**TABLE 16. Employee’s evaluation sample (individual technique’s results).**

Input	Results
User Defined Weights	Good
Crisp 1x0	Good
Crisp 1x1	Good
Crisp 2x1	Good
Crisp 3x1	Good
Fuzzy Script 1	Good
Fuzzy Script 2	Regular
Fuzzy Script 3	Regular
Fuzzy Script 4	Regular
C4.5	Excellent
Random Forest	Good
MLP	Excellent
Naïve Bayes	Good
Cart	Excellent
Justification text Sentiment Analysis	Good

final good evaluation. The Fuzzy committee was more rigorous, with most of the scripts resulting on a regular evaluation. Such a difference of results happened due to the employee’s position. As a senior technician, the “problem identification” and “systemic overview” results were decisive for the performance result. The exception was the first script, the only one that resulted in a good evaluation, due to the fact that it considers no difference between the job positions on the evaluation.

The third committee was the one with the biggest number of different votes. Classifying the employee’s performance as excellent, the committee presented 3 votes for this result, against 2 for a good performance.

With each committee choosing one different class, the sentiment analysis ended up being decisive for this employee’s final evaluation. With the justification text sentiment analysis classified as good, the first committee was able to issue 2 votes for good performance, against 1 vote for a regular performance, issued by the committee 2 and 1 vote for an

**TABLE 17. Employee’s evaluation sample (committees results).**

Input	Results
Committee 1 (Crisp)	Good (2 votes)
Committee 2 (Fuzzy)	Regular (1 vote)
Committee 3 (Classifiers)	Excellent (1 vote)
Committee of classifiers (final result)	Good

excellent performance, issued by committee 3, therefore classifying the employee’s performance final evaluation as good.

The next section will discuss in a more detailed manner, the general results presented after the application of this human resource evaluation process on this entire database.

## V. RESULTS

The conducted research generated a large range of results which indicate interesting aspects about each applied technique, and their combined use into this proposed human resource performance evaluation process.

The first difficulty in this aspect was the very own parameter definition for this analysis. In many similar researches, the computational intelligence techniques often try to simulate the human behavior. That is not the case here. This process proposes a methodology to mitigate mistakes caused by human subjective appraisal. This premise indicates that the new results are expected to be different from the original ones and, therefore, an accuracy measure, for instance, is not indicated for this analysis.

The processing speed of the algorithm, an extremely important indicator for tasks that deal with large amount of data, at least at this point, is not being considered either. After all, unless applied to a huge company, with thousands of employees being simultaneously analyzed, this shouldn’t be a problem.

There are, however, other indicatives that point for the success or not, of this general set of evaluations. On the results collected so far, there were considered aspects such as the general distribution of performances (quantitative of excellent, good, regular and insufficient evaluations), the correct class classification (in order to make sure that the algorithms are not failing and presenting an exit without any chosen class), how rigorous or light each technique is being and how sensitive that is for the employee’s final performance evaluation result.

The following figures will present a summary of the results achieved individually, by each technique, and the committees results. Figure 11 presents a comparison between the ruled classify strategies, while figure 12 presents a summary of the learning induction algorithms results.

These results present interesting data to be discussed. Regarding the crisp evaluations, as it was to be expected, the rigor of the appraisal appears directly linked to the self-evaluation weight. As a result, the evaluation with the highest number of Excellent and Good evaluations was

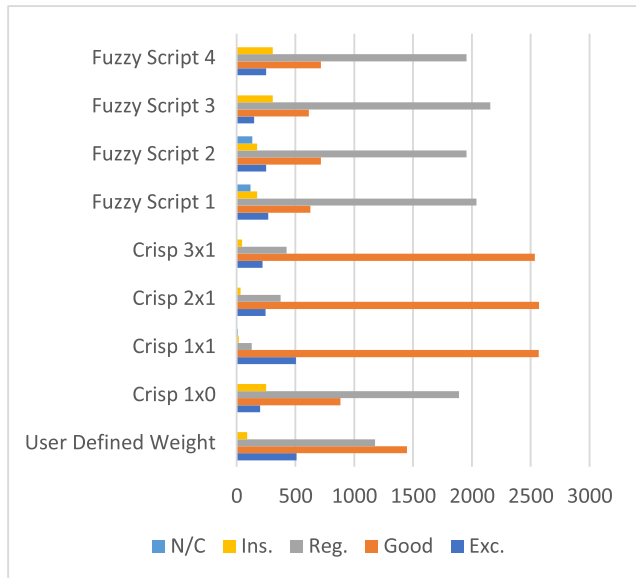


FIGURE 11. Ruled strategies result comparison.

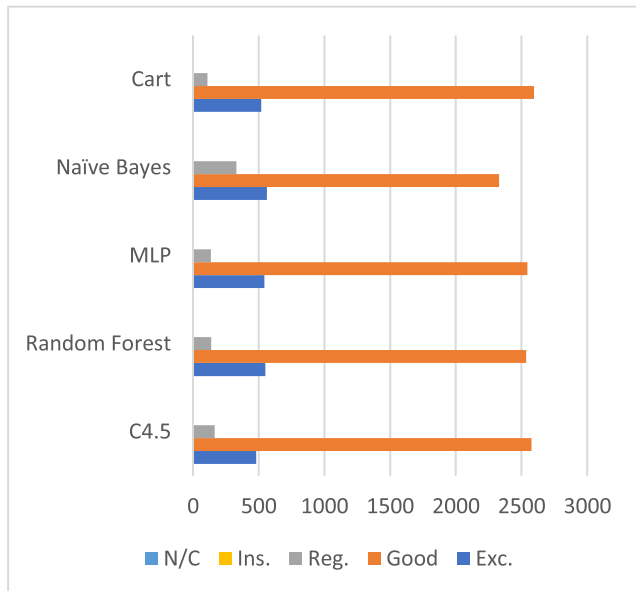


FIGURE 12. Learning induction algorithms result comparison.

“Crisp 1 × 1”, where both the hierarchical superior and self-evaluation present the same weight. The analysis on the classifications performed by weights defined by the user are not being analyzed due to the big number of possible weight combinations.

Another aspect that calls the attention is the number of insufficient evaluations, considerably higher in the evaluation that did not consider the self-evaluations on the average calculation (“Crisp 1 × 0”). Such an observation contributes to the hypothesis that, on this model, once the self-evaluation has a considerable weight it becomes a lot difficult to find an insufficient evaluation for an employee.

As previously discussed, usually, the general expected balance from a set of human resource performance evaluation

TABLE 18. Employee’s evaluation summary (Classification’s difference).

Comparison	Equally classified	Different classified	Difference Perc.
Crisp X Classifiers	2526	701	21.7
Crisp x Fuzzy	1009	2218	68,73
Fuzzy x Classifiers	594	2633	81.59
Crisp X Fuzzy X Classifiers	547	2680	83.04

results with this type of scale should present a larger result concentration around good and regular results, with less employees achieving excellent or insufficient evaluations. In that case, “Crisp 1 × 0” was the closest of the crisp evaluations to achieve such a goal.

The Fuzzy results, presented on figure 12, show a considerably different scenario. Unlike the crisp results, this time, the result concentration occurred around the regular performance, with the excellent and good results presenting small variations.

The Fuzzy scripts 1 and 2, less complex and with fewer rules, presented respectively 117 and 132 non classified instances. On that aspect, the raise of complexity on the third and fourth script clearly improved the classifications, resulting on just a single non-classified instance.

The Fuzzy script 3 appears to be the most rigorous of the 4 scripts, presenting the smallest number of excellent and good performances and the highest of regular and insufficient. Once this script was built with the goal of applying an extremer evaluation, as explained on the last section, the results appear to indicate that this goal was achieved.

The only script to consider all the available performance inputs, Fuzzy script 4 presented similar results from the third script, with a slightly closer balance between good and regular performances.

Much like the crisp evaluations, the learning algorithms results also concentrated most of the evaluation around a good performance (as seen in figure 12). Such a behavior can be explained considering that these classifiers were built using training sets originated from the original crisp analysis.

Another interesting data is the difficulty presented by these classifiers in qualify an insufficient performance. Such a behavior can also be described by the lack of original insufficient results, as presented on figure 1.

As figures 11 and 12 present, there is reasonable difference on the appraisals depending on which technique is applied on such an evaluation. The committee of classifiers strategy attempts to analyze such a difference, and choses the one that best fits the employee’s performance, in a fair way. Table 18 presents an analysis of how different these evaluations were for each instance presented on the database.

As table 18 indicates, the vast majority (83.04%) of the appraisals suffered variations from committee to committee.

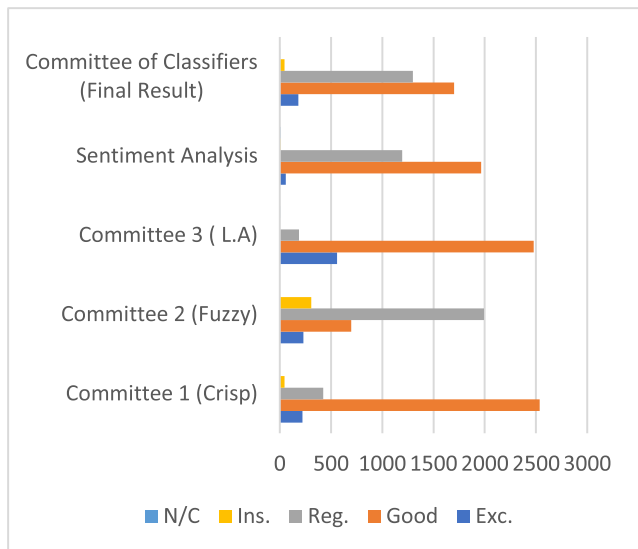


FIGURE 13. Employee's evaluation summary (Final result).

The smallest variation took place in the analysis between the Crisp and Classifiers committees. That makes sense, considering that the classifiers were built considering the original database results, which were basically crisp evaluations. The variation of both crisp and classifiers to the fuzzy committee is considerably larger, which also makes sense, considering the larger rigor and personalization of the fuzzy scripts.

Such variations maximize the importance of the sentiment analysis on the justification text, which will often have a key role on the decision of the employee's final performance evaluation result, as the data on figure 13 indicates.

The Crisp Committee (or Committee 1) results present a clear concentration of results around the good performance result, which can be explained by the presence of 3 functions that take in consideration the self-evaluations, against one that does not. The number of non-classified instances is, as expected, almost zero. Considering that this technique works with simple mathematical functions, such cases probably represent errors during the information process.

As a result of the individual scripts, the Fuzzy Committee (or Committee 2) results, presents a closer balance from good and regular performances. In a comparison with the first committee, however, presents a majority of regular performances.

The comparison between the committees 1 and 2 results shows just how sensible an evaluation can be on a flotation between good and regular results, depending on a few adjustments made on the appraisal formula.

The classification of a text into a performance evaluation result is not an easy job. A lot of context and continual needs of optimization needs to be considered.

These aspects were decisive for the application of the sentiment analysis as a consideration of weight for the votes of each individual technique and not having a single vote of its own. That strategy combines both an objective and subjective evaluation, providing one single final classification.

After the consolidation of the final results, once this process was applied to the entire database, it became possible to relate this data with the success indicators described on the begging of this sub-section.

The distribution of performances, for instance, shows a clear difference from the original data. With the performance classifications divided between good and regular performances. That is, without doubt, a first indication of success on this process. The number of excellent results also appears promise for the process validation. Considering that such a performance should be reserved for an employee's elite, it stands to reason that such a classification should not have a big percentual rate.

Being that this performance corresponds to 5.60% of the database, such a goal appears to have being achieved. The insufficient performance results are harder to be interpreted. With only 1.45% of classifications, however, that also appears to be coherent with the statistical difficulties for most of the employees to achieve this negative mark.

The correct classification is probably the easiest of the indicators to be validated. On a final account of 3227 records, only 1 failed to have received a final classification, giving the process an almost 100% success rate in that indicator.

The rigorousness and lightness of each technique and the sensibility of each of then to the process final result is an important indicator to prove the presence of actual diversity on the techniques and provide confidence that the employee will not have to trust on only one (quite possibly failed), methodology. The results on table 18 validate such a diversity, showing that only 16.96% of the records had a 100% of coincidence in all of the committee's classifications.

After the analysis of the indicators, it is possible to state that this proposed human resource evaluation process shows promise results and proves to be a worthy contribution both to the human resource field and to the application of computer intelligence techniques.

## VI. CONCLUSION

This paper presented a proposal for a human resource performance evaluation process using both ruled based and supervised learning algorithms. Giving that such a task is of the highest importance for the human resource management area, such a process has, as a main goal, to achieve a fairer final performance result for each employee, in a way that meets the expectations of those who evaluate and of the evaluated. In other to reach such goal, this process applied computational intelligence techniques know as text sentiment analysis, fuzzy logic, artificial neural network, naïve bayes, decision tree algorithms and the ensemble of committees using algorithms known as bagging and boosting.

The applied methodology seeks a way of combining the strategies of each used technique, in order to create a process that does not rely on just one classifier, but, instead, collects the opinion of each one of them, applying to then the best possible judgment to define the employee's final performance evaluation.

Working initially with an original database composed of about 2000 records, the proposed process was applied to both old and new records, comprehending a total of 3228 performance evaluations. The result analysis showed, in a comparison between each of the used techniques, how sensitive a performance evaluation of an employee can be, depending of a number of choices that are made during the course of a subjective process and how that may lead to a not so fair final result.

The final result compilation shows a well succeeded process, being that virtually all the database records were correctly classified in a distribution that differs from the original presented evaluations, balancing the performances around good and regular grades, with a few records qualified on the extreme classifications (excellent and insufficient). Giving the successful application of the proposed methodology, as well as the coherent results presented on section V, it is believed by the authors that such a paper presents a worthy contribution both to computational intelligence and human resource management fields.

## REFERENCES

- [1] *A Guide to the Project Management Body of Knowledge (PMBOK Guide)*, 4th ed, Project Management Institute (PMI), Newtown Square, PA, USA, 2008.
- [2] P. Luukka and M. Collan, "Fuzzy scorecards, FHOWA, and a new fuzzy similarity based ranking method for selection of human resources," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Manchester, U.K., Oct. 2013, pp. 601–606.
- [3] G. R. Ferris, W. A. Hochwarter, M. R. Buckley, G. Harrell-Cook, and D. D. Frink, "Human resources management: Some new directions," *J. Manage.*, vol. 25, no. 3, pp. 385–415, 1999.
- [4] M. Huemann, A. Keegan, and J. R. Turner, "Human resource management in the project-oriented company: A review," *Int. J. Project Manage.*, vol. 25, no. 3, pp. 315–323, 2007.
- [5] X. Gui, J. Xiao, J. Zhang, and Y. Bao, "Designing the expert systems for senior Managers' performance evaluation," in *Proc. 7th Int. Joint Conf. Comput. Sci. Optim.*, Beijing, China, Jul. 2014, pp. 607–610, doi: 10.1109/CSO.2014.117.
- [6] M.A. Huselid, "The impact of human resource management practices on turnover, productivity, and corporate financial performance," *Acad. Manage. J.*, 38, 3, pp. 635–672, 1995.
- [7] R. Amit, and M. Belcourt, "Human resources management processes: A value-creating source of competitive advantage," *Eur. Manage. J.*, vol. 17, no. 2, pp. 174–181, 1999.
- [8] V. Shahhosseini and M. H. Sebt, "Competency-based selection and assignment of human resources to construction projects," *Scientia Iranica*, 18, no. 2, pp. 163–180, 2011.
- [9] A. S. de Oliveira Góes and R. C. Limão, "Fuzzy System for Human Resource Performance Evaluation," in *Proc. Int. Conf. Comput. Sci. Comput. Intell.*, Las Vegas, NV, USA, Dec. 2017, pp. 315–321.
- [10] O. Pektor, B. Walek, and I. Martinik, "Proposal of a system for evaluating competencies," in *Proc. 19th Int. Carpathian Control Conf. (ICCC)*, Szilvasvarad, Hungary, May 2018, pp. 502–507, doi: 10.1109/CarpathianCC.2018.8399682.
- [11] Y. Chang, "Study on capability evaluation system of a company based on staff competency model—Take Company A as an Example," in *Proc. 6th Int. Conf. Inf. Manage., Innov. Manage. Ind. Eng.*, Xi'an, China, Nov. 2013, Art. no. 109114, doi: 10.1109/ICIII.2013.6703095.
- [12] A. Fahmi, A. Derakhshan, and C. Kahraman, "Human resources management using interval valued intuitionistic fuzzy analytic hierarchy process," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Istanbul, Turkey, Aug. 2015, pp. 1–5, doi: 10.1109/FUZZ-IEEE.2015.7338094.
- [13] S. Ajoudanian and M. N. Abadeh, "Recommending human resources to project leaders using a collaborative filtering-based recommender system: Case study of GitHub," *IET Softw.*, vol. 13, no. 5, pp. 379–385, Oct. 2019, doi: 10.1049/iet-sen.2018.5261.
- [14] A. L. Jabar and T. A. Rashid, "Combining fuzzy rough set with salient features for HRM classification," in *Proc. IEEE Int. Conf. Comput. Inf. Technol., Ubiquitous Comput. Commun., Dependable, Autonomic Secure Comput., Pervas. Intell. Comput.*, Liverpool, U.K., Oct. 2015, pp. 244–251, doi: 10.1109/CIT/IUCC/DASC/PICOM.2015.35.
- [15] N. Hampornchai, N. Chakpitak, T. Chandarasupsang, T.-A. Chaikijkos, and K. Dahal, "Dynamic adjustment of age distribution in human resource management by genetic algorithms," in *Proc. IEEE Congr. Evol. Comput.*, Singapore, Sep. 2007, pp. 1234–1239, doi: 10.1109/CEC.2007.4424611.
- [16] I. O. Yigit and H. Shourabizadeh, "An approach for predicting employee churn by using data mining," in *Proc. Int. Artif. Intell. Data Process. Symp. (IDAP)*, Malatya, Turkey, Sep. 2017, pp. 1–4, doi: 10.1109/IDAP.2017.8090324.
- [17] A. Maurya, L. Akoglu, R. Krishnan, and D. Bay, "A lens into employee peer reviews via sentiment-aspect modeling," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Barcelona, Spain, Aug. 2018, Art. no. 670677, doi: 10.1109/ASONAM.2018.8508512.
- [18] H. Kulkarni, "Thought process based team member selection using contextual sentiment closeness," in *Proc. 3rd Int. Conf. Conver. Technol. (ICT)*, Pune, India, Apr. 2018, pp. 1–5, doi: 10.1109/I2CT.2018.8529734.
- [19] A. A. Abed and A. M. El-Halees, "Detecting subjectivity in staff performance appraisals by using text mining: Teachers appraisals of palestinian government case study," in *Proc. Palestinian Int. Conf. Inf. Commun. Technol. (PICICT)*, Gaza City, Palestinian, May 2017, pp. 120–125, doi: 10.1109/PICICT.2017.25.
- [20] X. Li, T. Shi, P. Li, and W. Zhou, "Application of bagging ensemble classifier based on genetic algorithm in the text classification of railway fault hazards," in *Proc. 2nd Int. Conf. Artif. Intell. Big Data (ICAIBD)*, Chengdu, China, May 2019, pp. 286–290, doi: 10.1109/ICAIBD.2019.8836988.
- [21] J. Yuan, "Worker evaluation using FCE and AHP," in *Proc. 5th Int. Conf. Intell. Hum.-Mach. Syst. Cybern.*, Hangzhou, China, Aug. 2013, pp. 568–571, doi: 10.1109/IHMSC.2013.141.
- [22] M. Bohlouli, N. Mittas, G. Kakarontzas, T. Theodosiou, L. Angelis, and M. Fathi, "Competence assessment as an expert system for human resource management: A mathematical approach," *Expert Syst. Appl.*, vol. 70, pp. 83–102, Mar. 2017.
- [23] Y. Yamazaki and J. Yoon, "A cross-national study of fairness in Asia: How perceptions of a lack-of-group bias and transparency in the performance evaluation system relate to job satisfaction," *Hum. Resource Manage.*, vol. 55, no. 6, pp. 1059–1077, 2016.
- [24] I. Chiavenato, *Human Resources Administration*, 2nd ed. São Paulo: Editora Atlas, 1981.
- [25] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery in databases," *AI Mag.*, vol. 17, no. 3, pp. 37–54, 1996.
- [26] F. Lobato, "Evolutionary strategies to optimize the treatment of missing data by multiple data imputation," Ph.D. dissertation, Inst. Technol., Graduate Program Elect. Eng., Federal Univ. Pará, Belém, Brazil, 2016.
- [27] G. Holmes, A. Donkin, and I. H. Witten, "Weka: A machine learning workbench," in *Proc. Austral. New Zealand Intell. Inf. Syst. Conf.*, Brisbane, QLD, Australia, Nov./Dec. 1994.
- [28] R. E. Schapire, "The strength of weak learnability," in *Machine Learning*, vol. 5, no. 2. Springer, Jun. 1990, pp. 197–227. [Online]. Available: <https://link.springer.com/content/pdf/10.1007/BF00116037.pdf>, doi: 10.1007/BF00116037.
- [29] R. Polikar, "Ensemble based systems in decision making," *IEEE Circuits Syst. Mag.*, vol. 6, no. 3, pp. 21–45, 2006.
- [30] R. Liu, Y. Shi, C. Ji, and M. Jia, "A survey of sentiment analysis based on transfer learning," *IEEE Access*, vol. 7, pp. 85401–85412, 2019, doi: 10.1109/ACCESS.2019.2925059.
- [31] S. Zirpe and B. Joglekar, "Polarity shift detection approaches in sentiment analysis: A survey," in *Proc. Int. Conf. Inventive Syst. Control (ICISC)*, Coimbatore, India, Jan. 2017, pp. 1–5, doi: 10.1109/ICISC.2017.8068737.
- [32] R. E. Schapire, "A brief introduction to boosting," in *Proc. 16th Int. Joint Conf. Artif. Intell.*, 1999, pp. 1–6.
- [33] *Bizagi Modeler*. Accessed: Nov. 4, 2019. [Online]. Available: <https://www.bizagi.com/pt/produtos/bpm-suite/modeler>
- [34] L.-X. Wang, *A Course in Fuzzy and Systems Control*. Upper Saddle River, NJ, USA: Prentice-Hall, 1997.
- [35] P. Cingolani and J. Alcalá-Fdez, "JFuzzyLogic: A robust and flexible fuzzy-logic inference system language implementation," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Jun. 2012, pp. 1–8.



- [36] V. C. Gerogiannis, E. Rapti, A. Karageorgos, and P. Fitsilis, "A fuzzy linguistic approach for human resource evaluation and selection in software projects," in *Proc. Int. Conf. Ind. Eng. Oper. Manage. (IEOM)*, Dubai, United Arab Emirates, Mar. 2015, p. 19, doi: [10.1109/IEOM.2015.7093709](https://doi.org/10.1109/IEOM.2015.7093709).
- [37] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, p. 123, 1996. [Online]. Available: <https://link.springer.com/content/pdf/10.1023/A:1018054314350.pdf>, doi: [10.1023/A:1018054314350](https://doi.org/10.1023/A:1018054314350).



**ANDERSON SILVA DE OLIVEIRA GÓES** received the B.S. and master's degrees in computer science from the Federal University of Para, Brazil, in 2008 and 2014, respectively, where he is currently pursuing the Ph.D. degree. His current research interests include computational intelligence, data mining, and fuzzy logic.



**ROBERTO CÉLIO LIMÃO DE OLIVEIRA** was born in Belem, Para, Brazil, in 1965. He received the Ph.D. degree in electrical engineering from the Federal University of Santa Catarina, Florianópolis, Brazil, in 1999. He is currently a Full Professor at the Federal University of Para, Brazil. He published more than 200 refereed articles in journals and conferences, and published a book. His research interests include natural computing, computational intelligence, artificial neural networks, and evolutionary computing.

• • •