Comments for computers-2922494

# REVIEWER 1

## 1. Comment

Page 3 , In Table 1, in the 8th step, the output says:

"in this step must be: String: (x2+cos()); Chromosome: 23,8,14; Operation: 23 mod 2 = 1",

must say:

"in this step must be: String: (x2+cos()); Chromosome: 23,8,14; Operation: 23 mod 3 = 2",

## Response

Corrected.

## 2. Comment

Explain that the given example helps to understand how instead of using just the raw data (features) (x1, x2, ..., xN), with GE they are transformed into new data (new features) input to be applied in a given NN.

## Response

The following text has been added at the end of subsection 2.2:

*“This transformation of the initial 3 features of a hypothetical problem into just 1 artificial feature also shows the dynamics of the feature construction method, since with its use there can be a dramatic reduction in the dimension of the objective problem while preserving all the necessary information. ”*

## 3. Comment

Page 4,  "an Radial Basis Function" "**a** Radial Basis Function"

## Response

Corrected.

## 4. Comment

Explain the selected fitness function and its importance related to the objective problem.

## Response

We have added the following text in subsection 2.2:

*“**This fitness function can be thought of as a performance measure of using the artificial features produced by Grammatical Evolution. The most promising features will have a lower price than the worst features. ”*

## 5. Comment

Please, in line 110, explain how C being a NN can be subtracted from t\_j and then squared.

## Response

We have re – write this paragraph to the following to clarify the output value of model C(x).

## 6. Comment

Page 5, the mutation procedure explains how an element from the chromosome is selected to apply the mutation operation, but it is not explained how in fact the mutation happens in this element selected.

## Response

The following text has been added in Subsection 2.3 to provide an example of the mutation procedure using the example chromosome of 2.2:

*“**As an example of the mutation procedure consider the chromosome x=[ 9,8,6,4,16,10,17,23,8,14 ] used in the previous subsection that constructs the artificial feature f( x ) = x 2 + cos ( x 3 ) . If the mutation procedure alters the value 17 to 16 the new chromosome will be the following: x'=[ 9,8,6,4,16,10,16,23,8,14 ] and subsequently the new feature will be f( x ) = x 2 + sin ( x 3 ) ”*

## 7. Comment

It is recommended that the example given for the chromosome be used in the following examples, for instance to explain the crossover operation.

## Response

**(Paradeigma one – point crossover me xrisi tou paradeigmatos kai akoma enos xromosomatos. Ayto na dosei ena egkyro xromosoma)**

## 8. Comment

Page 6,

In the incise (a) "Set **as** IC as the total number ...", (a), correct the phrase "Set IC as the total number ..."

"as IG the number of generations", correct to: "IG **as** the number of generations", please review similar cases.

## Response

Corrected.

## 9. Comment

Please, introduce the concept of "margin"  (initial margin, set of margins, etc.) is used in the algorithm proposed and why it is needed.

## Response

## 10. Comment

Please, review the phrase "Every chromosome contains a set for intervals randomly initialized in N\_M", if the intention is to explain that each element of the chromosome is brought alive from a pool of possible elements contained in a certain interval, then the phrase could be rewritten in order to bring the right understanding to the reader.

## Response

The following text has been added at this point to clarify the procedure.

*“**This interval will be used as the initial value interval for the chromosome elements. Progressively, the above interval will be adjusted and its range reduced in order to improve the performance of the process controlled by Grammatical Evolution. ”*

## 11. Comment

Please, review the concept of chromosome , for instance, when speaking about the "elements for the chromosomes" it sounds that those elements are the genes, then it could be related to the genotype. Now, it could be said that each gene lives in a defined region which its limits depends on the nature of the given features.

## Response

## 12. Comment

In Algorithm 1, in the evaluation step: "iv. Denote by Emin(g\_i) the maximum train error obtained by the Feature Construction procedure", it seems that the right function name must be "Emax(g\_i)."

## Response

Corrected.

## 13. Comment

Usually a fitness score is a single value, please explain why an interval is used instead: fi = [Emin(gi), Emax(gi)], particularly when "the chromosomes are sorted with respect to their fitness values."

## Response

**(margin of promising values)**

## 14. Comment

Please, state the tournament size used in the tournament selection.

## Response

The tournament size is denoted as T\_s in the revised text and it is defined in Table 2.

## 15. Comment

In Equation (3), N is not defined.

## Response

Done.

## 16. Comment

In Algorithm 2, the explanation about how to get an offspring is not clear, it seems that the selection is considering just only the indexes and not the values themselves, then the index selected acts as a point where to apply the scissors to cut the chromosome.

## Response

## 17. Comment

Please, explain if the library OPTIMUS contains an implementation of GE or it is an original implementation.

## Response

The following statement has been added in Subsection 3.3: *“This library contains optimization methods, optimization examples as well as a series of examples utilizing the Grammatical Evolution technique.”*

## 18. Comment

Results does not show a comparison between standard methods against the proposed one.

## Response

We have performed additional experiments with two techniques

1) The ADAM optimizer, used to train a neural network with h=10 hidden nodes

2) The NEAT method (NeuroEvolution of Augmenting Topologies )

## 19. Comment

The conclusions are not related to the abstract and not properly related in the introduction: "In the current work, an efficient technique was proposed to identify the bounding box 266 of the values of chromosomes in Grammatical Evolution."

## Response

**(BETTER CONCLUSIONS. SAME AS REVIEWER3)**

# REVIEWER 2

## 1. Comment

The abbreviation BNF must be spelled out when first mentioned.

## Response

Done.

## 2. Comment

The explanation of artificial features should be expanded. What types of features are used.

## Response

The following text has been added in the Introduction section:

*“These artificial features can be considered as functions of the original features and in most cases these features are fewer in number than the original.”*

## 3. Comment

The FC method in experiments requires NG generations at the genetic stage, IFC requires IC generations at the evaluation stage, but step 2.ii in IFC also requires NG generations. It's unclear how this saves time? It should be clarified.

## Response

We state in the Conclusion section: “*However, a major drawback of the present technique is that it requires a significant amount of computing resources for its execution, but this can be alleviated by using parallel computing techniques, such as the MPI programming library [*[*MPI*](#LyXCite-MPI)*] or the OpenMP library [*[*OpenMP*](#LyXCite-OpenMP)*].”*

Hence, the proposed modification is expected to be slower than the original method, since it requires and additional step that involves another Genetic Algorithm.

## 4. Comment

If the authors claim time efficiency, comparisons should also be made in terms of execution time or the total number of iterations.

## Response

Yes you are correct. We have modified the text in the Conclusion sections to the following:

*“The proposed method precedes Grammatical Evolution and aims to discover a promising value interval for the chromosome values and, in this way, will significantly improve the efficiency of the Grammatical evolution procedure by limiting the optimization to more narrow intervals for the values of the chromosomes. ”*

## 5. Comment

In section 3.1, when describing datasets, it would be better to specify what type of data they contain. What are the inputs and what are the outputs for the algorithm in these tasks.

## Response

A separate table indicating the number of classes for each classification dataset has been added in Subsection 3.1 of the revised manuscript.

## 6. Comment

Figure 5 all dots are flattened because of outline in MLP column. Is there a way to improve it? Could it be an error in experiment or bad hyperparameters that it just didn't converged? Maybe rerun the experiment or use log scale for y axis to improve it?

## Response

**(ALEXANDROS)**

## 7. Comment

In Section 3.3, the authors compare their method with MLP and RBF networks. It would be appropriate to compare the proposed methods with algorithms of the same class, other genetic algorithms.

## Response

a) The MLP column stands for the results obtained by the application of a genetic algorithm with the same parameters as in the case of Feature Construction (Number of chromosomes and Number of generations). We have updated the revised manuscript to clarify this.

b) We have performed additional experiments with two techniques

1) The ADAM optimizer, used to train a neural network with h=10 hidden nodes

2) The NEAT method (NeuroEvolution of Augmenting Topologies )

## 8. Comment

Typos:

a) Line 216 - name of the column is skipped.

b) Lines 97 and 104, i believe N\_G and N\_g means the same.

c) Algorithm 1. Check the usage of "as"

## Response

a) The word MLP has been added in the revised text

b) Corrected to N\_G

c) Corrected.

# REVIEWER 3

## 1. Comment

The introduction outlines the proposed technique and its application without explicitly stating the research objectives. Clarifying the objectives would provide a more precise roadmap for readers.

## Response

We have added the following paragraph in the Introduction section to clarify this:

*“The purpose of this technique is to identify an interval of values for the gene values within which the search for the global minimum of the objective problem will be most efficient. This technique starts from a wide range of values, which is gradually narrowed until a promising configuration is identified. ”*

## 2. Comment

In the sentence "Any BNF grammar can be described as a set) G=(N, T, S, P)", the variable G is defined as a set, but it should be defined as a tuple.

## Response

Corrected.

## 3. Comment

In the sentence "The original BNF grammar is extended by enumerating the production rules", it would be more precise to mention how the production rules are enumerated or extended. In the sentence "The constant N denotes the dimension of the input data", it's unclear what N exactly represents. It should be explained more explicitly.

## Response

a) The following text has been added in subsection 2.1: *“For each non-terminal symbol the production rules are given a sequence number starting from 0, which will help the Grammar Evolution process in choosing the next production rule.”*

b) Also, the following text has been added in 2.1: “The constant N denotes the number of original features in the used dataset.”

## 4. Comment

In Table 1, the operation for the string "<expr>" states 9 mod 3=0, but it should be 9 mod 3=0, followed by the selection of the production rule, which is not explicitly shown.

## Response

The production rule: <expr> ::=  (<expr> <op> <expr>) is selected in the second row and it replaces the non – terminal symbol <expr> of the first row.

## 5. Comment

In the last row of Table 1, is there a missing step that shows the final production of the expression x2 +cos (x3) from the chromosome?

## Response

In the final row, there are no more non-terminal symbols in the expression and hence the production terminates with a valid expression.

## 6. Comment

Algorithm 1 does not incorporate elitism, where the best-performing individuals from one generation are preserved in the next. Without elitism, promising solutions can be lost throughout the optimization process.

## Response

The algorithm 1 uses the Algorithm 2 to perform crossover. In the second step of the algorithm 2 we state:

*“**The first ( 1- p s ) × I c chromosomes are copied intact to the next generation and the rest of the population are replaced by offsprings created. ”*

## 7. Comment

The range of the random number ai is specified as [−0.5, 1.5]. This range seems unusual for a probability or crossover parameter. Typically, probabilities are in the range [0, 1]. If this range is intentional, it should be justified or explained.

## Response

This range was specified in the relevant paper of Kaelo and Ali (reference 46) and the following text has been added:

*“**This process and the range of a i random values are suggested in [*[*doublecrossover*](#LyXCite-doublecrossover)*]. ”*

## 8. Comment

The algorithm lacks details on how the tournament selection process is conducted and how the random number ai is generated. These details are crucial for understanding and implementing the algorithm correctly.

## Response

We have added the following text in subsection 2.2 regarding the tournament procedure:

*“**The procedure of tournament selection has the following steps: A set of T\_s >1 randomly selected chromosomes is created and the chromosome with the lowest fitness value in this set is selected as parent. ”*

## 9. Comment

Using a single-point crossover method may restrict offspring diversity, potentially leading to premature convergence in complex optimization problems.

## Response

This crossover method was used also in the original paper of the Grammatical evolution technique, as we state in the revised manuscript.

## 10. Comment

In step 1.ii, "the the" appears to be a duplication of the word "the". It should likely be corrected to "the upper value".

## Response

Corrected.

## 11. Comment

The algorithm mentions the mutation rate (pm) without explaining its significance or how it is determined. It should clarify how pm is chosen and its effect on the mutation operation.

## Response

The following text has been added in subsection 2.2 to clarify this:

*“**In most cases of Genetic Algorithms, the mutation rate takes small values, e.g. 1-5%, in order for random changes to occur in the chromosomes, on the one hand, leading the optimization to search for the total minimum, but on the other hand, not to lose good values that have been discovered in some generation of the algorithm. ”*

## 12. Comment

The experimental setup involves several parameters (e.g., number of chromosomes, number of generations, mutation rate) that are manually set. The robustness of the reported results to variations in these parameters is not explored, and the findings may be sensitive to parameter choices.

## Response

We have performed additional experiments in the revised text (TABLE 6), where the number of chromosomes changes from 100 to 200 and the added text reads:

*“An additional experiment was conducted by altering the number of chromosomes in the proposed method from 100 to 500 and the results are outlined in Table* [*tab:experiments\_nc*](#tab_experiments_nc) *for the classification datasets. In this experiment, the stability of the proposed technique is revealed, since any changes in the population size do not significantly affect its experimental performance, as measured by the classification error in the test set. ”*

## 13. Comment

While the 10-fold cross-validation method was used to validate the experimental results, other validation techniques (e.g., leave-one-out cross-validation, bootstrapping) were not explored. The choice of validation method may affect the reliability and generalizability of the reported performance metrics.

## Response

The following statement has been added in subsection 3.3

*“This method has been utilized in many research papers and it is considered as a standard validation method. In the event that small datasets are provided, other validation techniques may be used, such as the leave - one out method. ”*

## 14. Comment

The methodology used for statistical testing and the significance thresholds are not explicitly stated, making it difficult to assess the robustness of the reported findings.

## Response

**(ALEXANDROS)**

## 15. Comment

The conclusions lack discussion on specific challenges in generalization. A brief discussion on transferability and necessary adaptations would enhance completeness.

## Response

**(BETTER CONCLUSIONS. SAME AS REVIEWER1)**