# REVIEW

about the PhD thesis for the acquisition of the scientific degree "**doctor**"
**Domain:** 5 Engineering Sciences,
**Professional field 4.6.** „Informatics and computer sciences "
**Scientific specialty**: "Informatics "

**Title of the PhD thesis:** "Optimization Algorithms for Data Management"

**Author of the PhD thesis:** Edjola Naka, Institute of information and communication technologies – Bulgarian Academy of Sciences (IICT-BAS)
**Ph.D. supervisor:** Prof. Ph.D. Vassil Guliashki , IICT-BAS

**Reviewer:** Prof. D.Sc. Ph.D. Eng. Todor Atanasov Stoilov, Institute of Information and Communication Technologies – Bulgarian Academy of Sciences (IICT-BAS), Sofia, Acad. G. Bontchev str., BL.2

## 1. Actuality of the problem developed in the dissertation work in scientific and applied scientific terms. Degree and levels of relevance of the problem and specific tasks created in the dissertation

There is a tendency to complicate the type and content of optimization tasks, which formalize problems from various technical, economic, social, and public domains. The complicated nature of the optimization tasks, the increase in dimensions of source data, and the application of quantitative estimates for non-trivial measurable objects necessitate the development of appropriate algorithms for solving such complex optimization tasks. Such algorithms use terms such as heuristic, metaheuristic, and other formulations that differ from classical terms for optimization in the presence or absence of constraints. The thesis aims to develop and modify a proprietary algorithm called Binary Volleyball Premier League. This algorithm was subsequently applied to data to identify stages of Parkinson's disease.

I find that the usefulness and relevance of dissertation research are easily visible and understandable. I positively assess the thematic focus and current issues of the dissertation research.

## 2. Degree of knowledge of the state of the problem and creative interpretation of the literary material

I positively assess the degree of knowledge of the problem of developing, modifying, and applying heuristic/metaheuristic algorithms for solving optimization tasks of the class of classification and ranking of properties of objects and results. Chapter 1 is a legal analysis and overview of the characteristics and properties of the heuristic/meta-heuristic algorithms that are applied to solve an optimization problem. The contents, operations, and applications of the procedures in Data Management, Machine Learning, and Metaheuristic algorithms were considered and an overview was made.

The thematic orientation of the dissertation research on the development and modification of a metaheuristic method for optimization justifies the presented analysis of the possibilities of individual methods and approaches in data analysis, machine learning, and metaheuristic algorithms. The presented analysis evaluates the capabilities of the various methods applied in these areas for optimization and decision classification. As a result of the overview made, the dissertation work sets its tasks for the development of a specific metaheuristic algorithm, which is based on an existing one titled Volleyball Premier League and to use it for the diagnosis of Parkinson's disease patients.

## 3. Correspondence of the chosen research methodology with the set goal and tasks of the dissertation work

The chosen research methodology is logical and consistently implemented. It contains a modification of the Volleyball Premier League algorithm by the usage of binary optimization arguments. The ways of this binary modification are analyzed since the input data have a general continuous character. This modification and author's development is called the Binary Volleyball Premier League. It has subsequently been applied to various datasets for the diagnosis of Parkinson's disease in patients. The results of the author's calculations have been compared with the application of other

existing classification methods by comparing accuracy (average value, standard deviation), and fitness.

The research methodology is a consequence of the thesis task of developing an algorithm for recognizing features from Parkinson's disease

## 4. Brief analytical description of the nature and assessment of the credibility of the material on which the contributions of the dissertation are based

The Ph.D. thesis is developed logically according to the requirements of an academic research paper. Chapter 1 provides an overview of non-classical optimization methods that are applied to Data Management, Machine Learning, and the development of Metaheuristic algorithms.

Chapter two presents the development of the modified Binary Volleyball Premier League algorithm for feature selection/determination. This algorithm has been applied to diagnose features of Parkinson's disease. In this chapter, the dissertation research also makes an additional version of the developed metaheuristic algorithm for determining properties.

Chapter 3 presents numerical evaluations and comparisons of feature selection results. The proprietary numerical results of the developed method Binary Volleyball Premier League and selected existing classification methods Logistic Regression, Neural Networks, and Naïve Bayes are compared. Comparison evaluations were made according to criteria for Sensitivity, Specificity, Precision, and Accuracy. The developed thesis Binary Volleyball Premier League algorithm is applied to ten Parkinson's disease datasets that relate to different physiological properties: Parkinson, HandPD spiral, HandPD meander, NewHandPD spiral, NewHandPD meander, Early biomarkers of PD based on natural connected speech , Parkinson's Disease Classification speech-based, Replicated acoustic features Parkinson, Parkinson dataset with Multiple Types of Sound Recordings, Gait Data Arm Swing. The reviewer is not familiar with the content of these sets of source data and therefore uses their names declared in the dissertation research. The explicit description does not comment on the features contained in these sets, whose features are identified/selected by solving a defined optimization problem.

The results presented are made in pairs from the input data sets. Applying an initial version of the author's Binary Volleyball Premier League algorithm does not always give superior results in comparison with other methods. Therefore, the Ph.D. research has made further changes and revisions to the currently derived metaheuristic algorithm with additional modifications. This new algorithm is titled hybrid

metaheuristic algorithm BVPL_BALO. The results from its usage to different sets of input data are presented, and assessed according to criteria for duration of calculations, Similarity, Dimensions, and Difference in accuracy.

I assess that the thesis correctly presents its developments for the modifications of the metaheuristic algorithm for feature selection and its application for the diagnosis of Parkinson's disease. In the Ph.D. research, an assessment of the corresponding accuracy of solving the task of feature selection from the application of the developed algorithm was made.

## 5. Scientific and/or scientific and applied contributions of the dissertation work

I find that the Ph.D. work has a scientific contribution. It consists of developing a modified algorithm for solving a feature selection task. I estimate the scientific-applied contribution in the application of the developed metaheuristic algorithm to real data on the features of Parkinson's disease. Additionally, the solutions of the developed algorithm have been compared with other metaheuristic algorithms and it has been evaluated their quantitative performance of computation time and accuracy.

I assess that the results obtained are original and useful. They fulfill the requirements for the development of a Ph.D. work for the acquisition of the educational and scientific degree "Doctor".

## 6. Assessment of the degree of personal involvement of the Ph.D. student in the contributions

The research presented in the Ph.D. thesis and the accompanying scientific publications with the author's participation gives me reason to believe that the dissertation results were made personally by the doctoral student.

## 7. Assessment of the publications on the Ph.D. work: number, nature of the editions in which they were printed. Reflection in science - use and citation by other authors, in other laboratories, countries, etc.

The Ph.D. procedure presents seven publications related to the developed thesis. They were made at scientific conferences in our country, Romania, and Malaysia. All publications are referenced in the electronic library of the IEEE organization. One

paper was also published in the journal Cybernetics and Information Technologies, which has a Q2 quartile, an IF impact factor, and a SCOPUS rank. This is a very good certificate for the Ph.D. research. A list of found citations is also attached, which is proof of the quality of the research and the obtained scientific results.

## 8. Using the results of the Ph.D. work in scientific and social practice

The Ph.D. thesis develops practical feature selection solutions for the diagnosis of Parkinson's disease. This is useful for practice and society. I positively assess the potential of the Ph.D. work for the usage of the obtained results in specific applied tasks.

## 9. Opinions, recommendations, and notes

The reviewer considers that it is appropriate to include sample calculations from certain steps of the feature selection algorithm being developed in the Ph.D. research. This will clarify what is used as input data, and what is calculated and helps to reproduce the results of the dissertation research. Presenting only formal dependencies does not show the internal logic of the quantitative estimates. The input data used for the Ph.D. research are not illustrated in nature. The absence of the type of input data does not allow us to assess the importance, complexity, and need for the development of new algorithms.

These recommendations may be taken into consideration by the Ph.D. student for his future publications and works.

A positive certificate for the Ph.D. work is the large number of references used to present the state of the art of the development and usage of metaheuristic algorithms for solving optimization problems. I also positively assess the publication activity, which has also achieved citations by specialists in this domain from home and abroad.

## 10. Conclusion with a clear positive or negative evaluation of the dissertation work

I give a positive assessment of the scientific-applied and applied results in the PhD thesis of Edjola Naka. I found that the legislative requirements of the Law for academic growth in Bulgaria, and the Regulations for its application and the internal Rules of IICT-BAS are satisfied. This gives me the reason to recommend to the honorable Scientific Jury to award Edjola Naka the educational and scientific degree

"Doctor" in professional field 4.6 „Informatics and computer sciences", scientific specialty "Informatics"..

Date : 22.11.2024 г.
Sofia

REVIEWE
( prof. D.S

*На основание*

*ЗЗЛД*