REVIEW

from

Prof. Olympia Roeva, PhD

Institute of Biophysics and Biomedical Engineering - BAS

Bioinformatics and Mathematical Modeling Department

for awarding of the educational and scientific degree "Doctor of Philosophy"

Professional field:

4.6 Informatics and computer sciences,

Doctoral programme: 01.01.12 Informatics

with a candidate

Alexander Ognianov Marazov, M.Sc.

PhD thesis title
"Deep Neural Networks for Medical Diagnostics"

1. **Relevance of the problem developed in the PhD thesis in scientific and scientific-applied terms.**

The development and application of mathematical approaches and algorithms for processing medical images for diagnosis is a particularly topical problem. Moreover, the subject of research in the PhD thesis is Alzheimer's disease, the most common form of dementia. Early diagnosis of Alzheimer's disease is an extremely difficult task, on which several scientific teams are actively working. To find effective approaches for diagnosing and determining the stage of Alzheimer's disease, the presented research is focused on the application of deep learning algorithms and fuzzy logic for medical data processing. The PhD thesis is up-to-date both in scientific and scientific-applied terms. The purpose of the PhD

thesis is formulated according to the current state of research and achievements in the field, namely "Improving the diagnostic process in medicine by designing, training, and optimizing software algorithms with convolutional deep neural networks."

To achieve this aim, the following six problems are defined:

1. To implement neural networks for diagnosing the stages of Alzheimer's disease.
2. To determine the specificity and sensitivity of the developed models.
3. To improve the accuracy of the models with a new inference method.
4. To evaluate the accuracy of the inferences in terms of intuitionistic fuzzy sets through InterCriteria analysis.
5. To propose a method for determining threshold values for degrees of membership and uncertainty, thereby enhancing the accuracy of the models.
6. To find an approach to improve the speed of InterCriteria analysis.

2. Degree of knowledge of the state of the problem and creative interpretation of the literature

The PhD student has systematically presented the state and problems in the field under consideration. An overview of neural networks (NNs) and more specifically convolutional neural networks (CNNs) – history and applications in medicine is given. The importance and wide scope of the application of NNs in medical practice and scientific research have been demonstrated.

An overview of intuitionistic fuzzy sets (IFS) is presented – history, extensions, norms and metrics in IFS and an overview of some of their applications for medical diagnostics - basic theoretical statements used in the PhD thesis. Research and development with the application of IFS in NNs to improve the functionality of learning algorithms are also presented.

A brief overview of Intercriteria Analysis (ICrA) provides the foundation that has been built upon to achieve improved analysis speed when dealing with large datasets. There is a lack of analytical review of the applications of ICA, especially in the field of medicine, with which the doctoral student can justify the choice of this approach.

A brief overview of Alzheimer's disease is also presented, with an emphasis on the public health impact of the disease and the main stages of the disease.

3. General analytical characteristics of the PhD thesis

The PhD thesis is well structured and logically consistent according to the tasks to be solved. The thesis is in a volume of 162 pages and contains a list of abbreviations, an introduction, the aim of the PhD thesis and main problems, 5 chapters (one overview chapter and four chapters with original research results), a conclusion – summary of the obtained results, directions for future work, main contributions of the PhD thesis, list of publications, list of citations, declaration of originality of the results, 4 appendices with the developed program codes and bibliography.

Chapter 1 is an overview and gives a picture of part of the state of the considered problems and the used mathematical concepts based on published results. A bibliography of

146 sources was used, including fundamental publications for the field, as well as publications from the last 10 years.

Chapter 2 presents the results of the application of convolutional neural networks for the diagnosis of Alzheimer's disease. The CNMs were trained and tested with about 5000 MRI images. A Python code has been developed that can be adapted for imaging other neurological diseases such as Parkinson's and multiple sclerosis. The results obtained can be used to assist doctors in making a diagnosis, as well as to extract examples from big data.

In Chapter 3, the results of the proposed new ways of inference in machine self-learning, based on the Kemeny-Young method, are presented. An interpretation of the Kemeny-Young method in terms of IFS through the ICrA method is proposed. The quality of the results is assessed by the degrees of agreement and disagreement by ICA. A non-parametric calibration method based on the ICrA method is also presented. The developed approaches are implemented in Python and a practical application of the inference method is presented.

The studies with the application of ICrA to assess the reliability of the proposed in the PhD thesis inference method in classification tasks are presented in Chapter 4. The application of the method to classifiers of Alzheimer's disease is considered. By choosing membership and uncertainty thresholds, models with a desirable balance between coverage and accuracy were found. A coverage of 60.75% was achieved with an accuracy of 90%. The ICrA approach to inference is applied in the One-vs-One procedure. Two theorems are derived and proved. Again, with developed Python code, the application of the approach to Alzheimer's disease staging results is shown.

Chapter 5 presents the results of research aimed at improving the speed of the ICrA algorithm. As a result, a computation of the itercriteria counters with a quasi-linear complexity ($O(n \log n)$) is proposed compared to the existing quadratic complexity of the algorithm – $O(n^2)$. The PhD student remains an open problem – whether intercriteria counters can be computed in $O(n\sqrt{\log n})$.

In conclusion, to achieve the goal of the Thesis, the theoretical framework of two current and well-established mathematical approaches – the IFS and the ICrA – has been expanded.

A good impression is also made by the formulation of several directions for future work – to expand the possible applications of the achieved results and to achieve a further reduction of the complexity of ICrA to work with big data sets.

4. Evaluation of contributions of the PhD thesis and their significance

I accept the contributions formulated in the PhD thesis.

The Scientific contributions are:

1. An inference method based on the Kemeny-Young method is proposed.
2. A method for evaluating the inferences of classification tasks in terms of intuitionistic fuzzy sets, based on InterCriterion analysis, is proposed.
3. A method for applying threshold values to the degrees of agreement and disagreement is proposed, which significantly improves the accuracy of the selected results.
4. An algorithm is proposed to improve the speed of InterCriterion analysis to $O(n \log n)$.

The scientific and applied contributions are:

1. An NN program code was implemented to diagnose the stages of Alzheimer's disease.
2. Model sensitivity of 70% and specificity of 85% were achieved.
3. A Python code was developed for the new inference method based on the Kemeny-Young method. Ortools optimization library was used to achieve practical execution time.
4. The developed inference method gives consistent results that are fault-tolerant by construction.
5. A Python code of the new method for evaluating the inferences of classification tasks in terms of IFS based on ICrA was developed. The code uses a non-trivial equality relation between classifier predictions.
6. A code was developed to apply threshold values to degrees of agreement and disagreement. It was applied to the Alzheimer's disease classification task, achieving a significant increase in the accuracy of the selected results. Model sensitivity of 80% and specificity of 95% were achieved.
7. Python code was developed to improve the speed of ICrA to $O(n \log n)$. Current implementations using an $O(n^2)$ complexity algorithm are compared to the new implementation using the automatic randomized testing package Hypothesis.

5. Assessment of PhD thesis publications

4 publications are indicated for the PhD work. One of them is in a journal with an impact factor (Mathematics, IF = 2.4 (2023), Q1) and one is in a journal with an impact rank (International Journal Bioautomation, SJR = 0.159 (2022), Q3). Two articles were published in the Annual of the Informatics Section, Union of Scientists in Bulgaria. Part of the results of the PhD thesis were presented at two international forums – the 10th European Academy of Neurology Congress, Helsinki, Finland and the 3rd International Symposium on Bioinformatics and Biomedicine, BioInfoMed'2024. The publications reflect the main results obtained in the PhD thesis. They have been published in prestigious journals and show the high scientific level of the research. One citation is presented.

6. Assessment of the compliance of the autoreferate with the requirements for its preparation, as well as the adequacy of reflecting the main points and contributions of the PhD thesis

The autoreferate correctly reflects the content of the PhD thesis and gives an idea of the problems under consideration, the results obtained, as well as the contributions of the thesis.

7. Critical notes on the PhD thesis

Alexander Marazov has taken into account the previously made comments and notes. I believe that all the essential remarks are reflected in the PhD thesis. Some technical errors and inaccuracies are noted, but as omissions of this nature in no way detract from the merits of the thesis, I do not consider it necessary to address them.

I will note that in Chapter 1 the presentation of the applications of IRM in different fields and different tasks is given very mechanically, without analytical discussion. For example: why are IPMs appropriate and how do results obtained using IPMs outperform other results?

I would like to recommend the PhD student to publish in a journal with IF the proposed inference method in machine self-learning and its applications.

8. Conclusion with a clear positive or negative assessment of the PhD thesis

Based on the above, I give a high rating to the scientific work done by Alexander Marazov, M.Sc. and the achieved results.

All the requirements, conditions and criteria of Law on the Development of the Academic Staff in the Republic of Bulgaria, the Internal Regulations for its application, as well as the Regulations for the terms and conditions for acquiring scientific degrees and occupying academic positions in IBPhBME – BAS have been fulfilled. I give a positive assessment of the PhD thesis and recommend to the respected Scientific Jury to award Alexander Ognianov Marazov the educational and scientific degree "Doctor of Philosophy" in the professional field: 4.6 Informatics and computer sciences, doctoral programme: 01.01.11 Informatics.

08.07.2024 г. Scientific Jury member: ... ...

Sofia /Prof. Olympia Roeva/