INSTITUTE OF BIOPHYSICS AND BIOMEDICAL ENGINEERING
BULGARIAN ACADEMY OF SCIENCES

Deep Neural Networks for Medical Diagnostics

Alexander Marazov

ABSTRACT OF

DISSERTATION THESIS

for acquiring the educational and scientific degree of Doctor
Professional field: 4.6 “Informatics and Computer Science”
Doctoral program: “Informatics”

Scientific supervisor:
Assoc. Prof. Lyudmila Pavlova Todorova, Ph.D.

Sofia, 2024

The dissertation has been discussed and directed for defense by an extended session of the "Bioinformatics and Mathematical Modeling" section at the Institute of Biophysics and Biomedical Engineering – BAS on 25.04.2024.

The present dissertation is structured into an introduction, five chapters, a conclusion – a summary of the obtained results, a declaration of originality, bibliography, a list of the author's publications, a list of noticed citations, four appendices, and comprises 162 pages. 146 sources are cited.

The defense of the dissertation will take place on ....2024 at .... in the conference hall of the Institute of Biophysics and Biomedical Engineering – Bulgarian Academy of Sciences (Acad. Georgi Bonchev St., Bl. 21, Sofia) at an open meeting of the scientific jury composed of:

Regular members:

- Prof. Maria Nisheva (Sofia University "St. Kl. Ohridski")
- Prof. Petya Koprinkova (Institute of Information and Communication Technologies – BAS)
- Prof. Sotir Sotirov (University "Prof. Asen Zlatarov", Burgas)
- Prof. Olympia Roeva (Institute of Biophysics and Biomedical Engineering – BAS)
- Assoc. Prof. Petar Vasilev (Institute of Biophysics and Biomedical Engineering – BAS)

Alternate members:

- Prof. Evdokia Sotirova (University "Prof. Asen Zlatarov", Burgas)
- Assoc. Prof. Simeon Ribagin (Institute of Biophysics and Biomedical Engineering – BAS)

Author: Alexander Marazov

Title: Deep Neural Networks for Medical Diagnostics

List of Abbreviations

- **AD** – Alzheimer's disease
- **IFP** – intuitionistic fuzzy pair
- **IFS** – intuitionistic fuzzy sets
- **CNN** – Convolutional Neural Networks
- **FS** – fuzzy set
- **CT** – computed tomography
- **MRI** – Magnetic Resonance Imaging
- **ResNet** – residual neural network
- **ILSVRC** – ImageNet Large Scale Visual Recognition Challenge
- **IFNN** – intuitionistic fuzzy neural network
- **IFDNN** – intuitionistic fuzzy deep neural network
- **LSTM** – long short-term memory
- **RBFNN** – radial basis function neural network
- **OvO** – One-vs-One
- **OvR** – One-vs-Rest
- **SVM** – Support Vector Machines
- **TP** – True Positives
- **TN** – True Negatives
- **FP** – False Positives
- **FN** – False Negatives

Introduction

Approaches based on fuzzy and intuitionistic fuzzy logic are increasingly being applied for effective segmentation of brain structures and for extracting suitable features from brain images [1] to achieve correct classification. The classification task in machine learning is the process of assigning a category or class to new objects based on their features or attributes.

Dementia, especially Alzheimer's disease (AD), is the most critical neurodegenerative disease affecting memory and cognitive functions. Over 55 million people worldwide have dementia, with AD accounting for 60-70% of cases. This fatal progressive disease requires early identification and intervention to improve future health outcomes [2, 3].

Alzheimer's disease is a global problem requiring effective diagnostic approaches, and deep learning is promising for accurate and early detection of the disease. The aim of this thesis is to design, train, and optimize algorithms with deep neural networks to improve the diagnosis and staging of Alzheimer's disease, using intuitionistic fuzzy sets and intercriteria analysis.

The dissertation is structured as follows.

In chapter 1 of the presented work is divided into four sections. In section 1.1, a review of neural networks is provided, with a focus on convolutional neural networks. Section 1.2 briefly presents the theoretical foundation of intuitionistic fuzzy sets (IFS) and their application for medical diagnostics. section 1.3 examines the method of intercriteria analysis, based on the apparatus of indexed matrices and intuitionistic fuzzy logic, from a theoretical perspective.

In chapter 2 presents the application of convolutional neural networks for the diagnosis of Alzheimer's disease. The program code, model selection and steps, model training, and analysis of the obtained results are described. The conclusion of the chapter outlines the advantages of the developed code.

In chapter 3 of the dissertation presents new inference methods developed based on intercriteria analysis and the Kemeny-Young method. A new machine learning inference method is proposed that minimizes the difference between the rankings of an ensemble of classifiers. A method for determining the degrees of membership, non-membership, and uncertainty of preference in the sense of intuitionistic fuzzy logic is presented, using intercriteria analysis.

In chapter 5 is dedicated to the computational complexity of calculating intercriteria counters. A proposition for a one-to-one correspondence between the number of discrepancies in the intercriteria counters and the number of inversions in a vector is defined and proven. An approach for reducing computations, which has been implemented programmatically, is developed.

Part of the results obtained in the dissertation have been published in four works. One publication is in the international journal "Mathematics" (IF = 2.4, Q1). One publication is in the international journal "International Journal Bioautomation" with SJR (SJR = 0.159). Two publications are in the Yearbook of the "Informatics" section (one of which is a studio), a publication of the Union of Scientists in Bulgaria. Results obtained in the dissertation will be presented at the Congress of the European Academy of Neurology in Helsinki, Finland, and at the international conference BioInfoMed'2024 in Burgas, Bulgaria.

Aim of the Dissertation

Improving the diagnostic process in medicine by designing, training, and optimizing software algorithms with convolutional deep neural networks.

To achieve this aim, the following tasks are set:

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.

Chapter 1

Literature Review

1.1 Neural Networks

Neural networks are a powerful tool for analyzing, modeling, and solving various problems in the field of artificial intelligence. From the first models inspired by the human brain to modern implementations, they have undergone significant development and have established themselves as one of the main technologies for machine learning.

1.1.1 History of Neural Networks

In 1957, Rosenblatt introduced a neural network model known as the perceptron (see Figure 1.1).

Despite the initial enthusiasm, the limitations of single-layer perceptrons were revealed in [4]. They demonstrated that single-layer perceptrons cannot solve problems such as the XOR function. To overcome this limitation, the necessity for multi-layer architectures was highlighted.

1.1.2 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a type of artificial neural network designed for processing data with a spatial structure, such as images. Their architectures are inspired by the biological organization of the visual cortex of the human brain and are optimized for extracting features from visual data, such as edges, textures, shapes, and more.

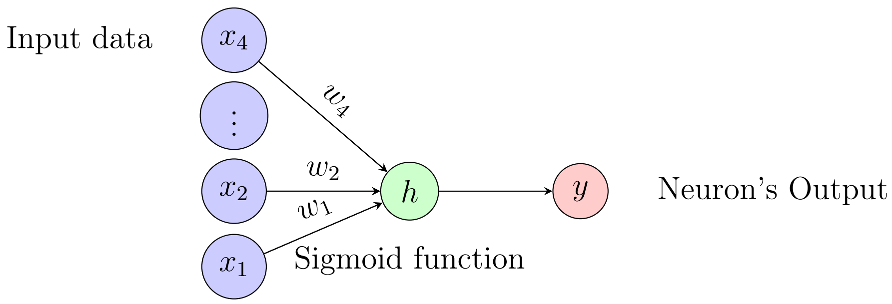


Figure 1.1: Representation of a single-layer perceptron with a sigmoid activation function

With the development of CNNs, the emergence of deep architectures has become increasingly significant. ResNet (Residual Network), introduced by He et al. in 2015 [5], features residual blocks of the form $B(x) = F(x) + x$. This allows the network to learn more effectively and avoid the vanishing gradient problem.

1.2 Intuitionistic Fuzzy Sets

1.2.1 Theoretical Basis of Intuitionistic Fuzzy Sets

Let E be a fixed set. The set A is called an intuitionistic fuzzy set (IFS) if it has the form [6]:

$$A = \{\langle x, \mu_A(x), \nu_A(x)\rangle \mid x \in E\}$$

where the functions $\mu_A : E \to [0, 1]$ and $\nu_A : E \to [0, 1]$ respectively denote the degree of membership and non-membership of the element $x \in E$ to the intuitionistic fuzzy set A, which is a subset of E and for every $x \in E$:

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1$$

The function π_A, defined by the formula:

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$$

represents the degree of uncertainty of the membership of the element $x \in E$ to the set A. In cases where E is a fuzzy set, the degree of uncertainty $\pi_A(x) = 0$ for every x. The degree of membership of an element in IFS is determined by the degree to which the element belongs to a given set, while the degree of non-membership represents the degree to which the element does not belong to that set. The third parameter introduced in IFS is the degree of uncertainty, which quantitatively determines the degree of uncertainty associated with the degrees of membership and non-membership. Introducing this parameter makes them more suitable for representing imprecise and unclear information [7, 8, 9, 10, 11].

1.2.2 Norms and Metrics in Intuitionistic Fuzzy Sets

All IFS over a fixed universe E form a metric space with a metric that applies not to the elements of E, but to the values of the functions μ and ν, as the "norm" of a given element x of the IFS associates a number with that element, related to the values of the functions μ and ν calculated for that element x. The presence of the second functional component ν leads to the definition [12] of different norms for each x with respect to a fixed $A \subset E$.

1.2.3 IFS for Medical Diagnostic Purposes Based on MRI Images

Medical imaging is an active area of research where anomalies are detected non-invasively. Due to the complexity of the images, many structures are difficult to visualize, and various computer techniques are applied by researchers to process medical images.

1.2.4 IFS and Neural Networks

In 2023, Atanassov et al. [13] developed the concept of Intuitionistic Fuzzy Deep Neural Network (IFDNN), which combines artificial neural networks and intuitionistic fuzzy sets. The goal is to combine the advantages of both methods.

1.3 Method of Intercriteria Analysis

The method of intercriteria analysis is a method for studying the dependencies between criteria that evaluate a set of objects. It is based on intuitionistic fuzzy sets as a mathematical tool for treating uncertainty, as well as on indexed matrices of generalized networks.

Let a matrix $M = (a_{C_p,O_q})$ with dimensions m by n be given. For any two integers p and q, satisfying the inequalities $1 \leq p \leq m$ and $1 \leq q \leq n$:

- C_p is a criterion;
- O_q is an object;
- a_{C_p,O_q} is the assessment of the q-th object with respect to the p-th criterion.

Each such assessment is determined as a real number or as another object that is comparable to the other assessments with respect to a relation R. Formally, we can define for each i, k, l the relation $R(a_{C_i,O_k}, a_{C_i,O_l})$.

Let $\bar{R}$ be the dual relation of R, so that $\bar{R}$ is satisfied if and only if R is not satisfied. For example, if R is the relation "$<$", then $\bar{R}$ is the relation "$\geq$".

1.3.1 Intercriteria Counters

Based on the previous relations, we can construct counters to obtain assessments between the criteria.

Let S^{μ}_{ij} be the number of cases where $R(a_{C_i,O_k}, a_{C_i,O_l})$ and $R(a_{C_j,O_k}, a_{C_j,O_l})$ are satisfied simultaneously. Let S^{ν}_{ij} be the number of cases where $R(a_{C_i,O_k}, a_{C_i,O_l})$ and $\bar{R}(a_{C_j,O_k}, a_{C_j,O_l})$ are satisfied simultaneously.

Since the total number of pairwise comparisons between the objects is $(n^2 - n)/2$, it can be seen that:

$$S^{\mu}_{ij} + S^{\nu}_{ij} \leq \frac{(n^2 - n)}{2}$$

For each i, j, such that $1 \leq i < j \leq m$, and for $n \geq 2$, we define:

$$\begin{aligned} \mu_{ij} &= \frac{2}{(n^2 - n)} S^{\mu}_{ij}, \\ \nu_{ij} &= \frac{2}{(n^2 - n)} S^{\nu}_{ij}. \end{aligned} \tag{1.1}$$

We construct an intuitionistic fuzzy pair $\langle \mu_{ij}, \nu_{ij} \rangle$, through these two values, for every pair of criteria C_i and C_j. This pair represents an intuitionistic fuzzy assessment of the similarity between the two criteria.

Chapter 2

Alzheimer's Diagnosis with Convolutional Neural Networks

This chapter explores the application of convolutional neural networks for diagnosing Alzheimer's disease. The approach follows a study by the author published in [14].

2.1 Data

For the purpose of this study, we use freely available data from the *kaggle* platform [15]. Around 5000 MRI (magnetic resonance imaging) images are divided into a training set and a test set randomly.

Each image is annotated with one of the following four classes:

1. NonDemented - without dementia
2. VeryMildDemented - very mild dementia
3. MildDemented - mild dementia
4. ModerateDemented - moderate dementia

All images and annotations are collectively referred to as a dataset.

2.2 Imaging Diagnosis Using Convolutional Neural Networks

A solution with minimal code (*low code*) is proposed, utilizing the *fastai* library. The approach is universal and with minimal changes, it can be applied to other diseases that have imaging diagnosis.

2.3 Program Code

The architecture of a convolutional neural network *resnet34* is selected. *ResNet* [5] [5] has variants with fewer parameters (e.g., *ResNet18*) as well as with more (e.g., *ResNet101*).

The model loaded is trained on images from [16], which have nothing to do with MRI, but still improve the results and make training much faster possible.

2.3.1 Choosing a Step

The step in optimization tasks for artificial intelligence is called the learning rate. It is the most important parameter when training neural networks.

Fastai provides visual assistance in choosing the learning rate through the method.

Traditionally, determining a suitable learning rate has required trial and error. In [17], Smith proposes an alternative approach.

Figure Figure 2.2 shows a graph of the losses for the *resnet34* model trained on MRI images of Alzheimer's disease. In this case, the point of steepest descent of the loss (valley) is shown. For the learning rate, a value close to this point, or the smallest achieved value, is chosen. In this case, we choose 0.00144, which is not the smallest value, but is close both to the local minimum around 10^{-3} and to the valley point.

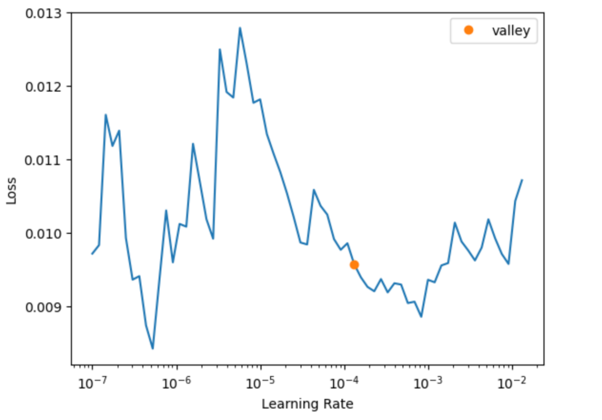


Figure 2.2: Loss graph for the *resnet34* model trained on MRI images of Alzheimer's disease. The valley is the point of steepest descent of the loss. It is recommended to choose a learning rate around it.

2.4 Model Training

We train the model on an NVIDIA Tesla P100 GPU for about 30 minutes. Training on a GPU accelerates the training time several times compared to training on a regular CPU [18].

For the actual training, we use the fine-tuning technique. This allows us to start training with pre-trained parameters that we fine-tune.

2.5 Results

The confusion matrix, depicting the actual and predicted classes, is presented in Table 2.1. Values on the main diagonal represent correct predictions. It is noticeable that errors are concentrated in close categories:

- NonDemented and VeryMildDemented are confused in 218 examples.
- VeryMildDemented and MildDemented are confused in 374 examples.

This can be interpreted in two ways. Firstly, there is no consensus among annotators on where the boundaries of individual classes lie. Or secondly, the important features distinguishing the two states were not learned by the model. Upon reviewing the data by an Alzheimer's expert, it was found that it is difficult to give a definitive answer about the progression of the disease solely based on the images. This could cause both of the aforementioned problems.

	Prediction			
Annotation	NonDemented	VeryMildDemented	MildDemented	ModerateDemented
NonDemented	543	94	3	0
VeryMildDemented	124	309	15	0
MildDemented	65	64	50	0
ModerateDemented	2	5	1	4

Table 2.1: Confusion matrix on the test set.

The model's sensitivity, defined as TP/(TP+FN), is 70%, while the specificity, defined as TN/(TN+FP), is 85%.

The results from the *resnet34* model on nine randomly selected images are shown in Figure 2.3.

In Figure 2.4, the results from the *resnet34* model for 9 misclassifications with the highest confidence are presented. The model's output is close to 1, indicating high confidence in the prediction. On the other hand, the actual classes are adjacent classes: NonDemented and VeryMildDemented; VeryMildDemented and MildDemented; MildDemented and ModerateDemented.

2.6 Conclusion

In section 2.5, we demonstrate that with minimal code and short training time, we can obtain a model with very good results. It should also be noted that the code can be easily adapted for image diagnostics of other diseases.

Similar automated systems cannot replace a doctor's assessment. Their main applications are (1) assisting doctors in diagnosis or (2) automated extraction of examples from large volumes of data.

In chapter 3, we will explore an approach that can significantly improve the results of the classification task by aggregating predictions from multiple classifiers.

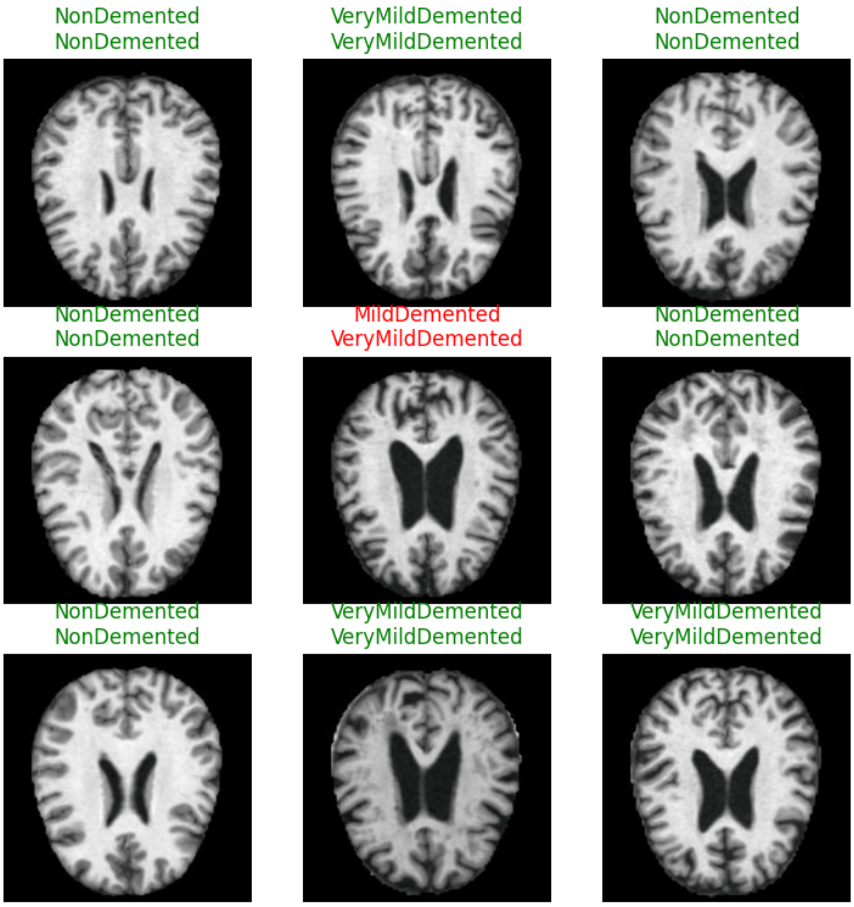


Figure 2.3: A sample of nine random images. Above each image, the predicted class/annotated class is indicated. Most predictions are correct.

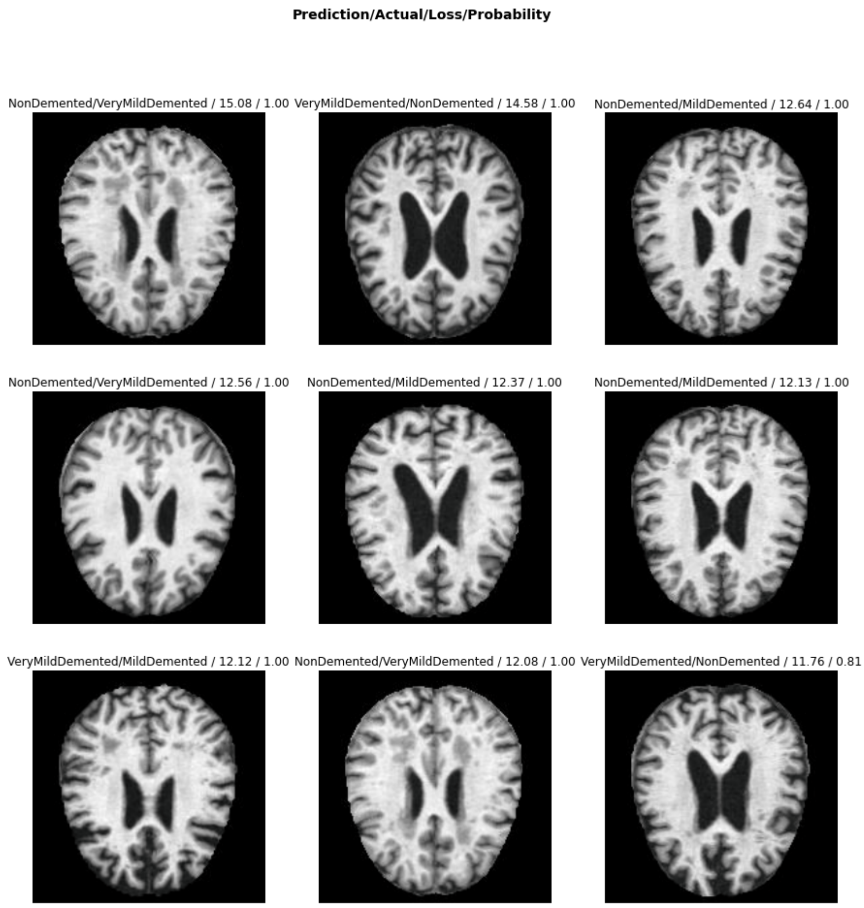


Figure 2.4: A selection of nine misclassified images with the highest confidence in predictions. Above each image, the prediction class/annotated class/loss for the example/model confidence are indicated. The loss for the example is defined as the value contributed by the particular example to the target function.

Chapter 3

An Inference Method based on Kemeny-Young Method

In Chapter 3, new inference methods based on the Kemeny-Young method in machine learning are proposed. The approach extends the results published in [19].

3.1 The Kemeny-Young Method

This method for aggregating election results was proposed by John Kemeny in 1959 [20]. The idea behind the method is to determine a rank order of alternatives that minimizes the difference between this ranking and the individual preferences of each participant. For this purpose, a preference matrix is used, reflecting the degree of preference for each pair of alternatives by each participant. This matrix is used to construct alternative rankings, which are then compared to determine the overall ranking.

Let's assume we have a set of n alternatives $X = \{x_1, x_2, ..., x_n\}$ and a preference matrix P, where P_{ij} denotes the preference of alternative x_i over alternative x_j. These preferences can be measured, for example, by the number of votes for a given pair of alternatives.

The goal is to find a rank order σ of the alternatives that minimizes the sum of the differences between the participants' preferences and the rank order:

$$\min \sum_{i<j} P_{\sigma(i)\sigma(j)}$$

subject to the condition that σ is a permutation of the indices $\{1, 2, ..., n\}$.

This can be formulated as a problem of finding a minimum Hamiltonian cycle, which is an NP-hard problem [21]. Let each edge have a weight associated with the preferences between two alternatives. For practical solutions to such problems, integer programming optimizers (MIP solvers) are used [21].

3.2 Intuitionistic Interpretations of the Kemeny-Young Method

Intercriterion analysis provides an opportunity to give an interpretation in terms of intuitionistic fuzzy sets of the Kemeny-Young method. The degrees of uncertainty and non-membership in intuitionistic fuzzy sets are a convenient tool for assessing the quality of the predicted results.

The Kemeny-Young method uses ballots on which voters rank the alternatives according to their preference order. A voter may rank more than one alternative at the same level of preference. Unranked alternatives are usually interpreted as the least preferred. To establish correspondence with the terms of intercriterion analysis (see Section 1.3.1), we will use the following rules:

- Equally ranked alternatives will increase the uncertainty counter.
- Preferences will increase the membership counter.

The first step is to create a matrix that counts the bilateral preferences of the voters. The second step is to test all possible rankings, compute a score for each ranking, and compare the scores. Each score for a ranking is equal to the sum of the bilateral preferences pertaining to that ranking. The ranking with the highest score is chosen according to the Kemeny-Young method.

3.2.1 Counting and Ranking via the Kemeny-Young Method

A formal definition for ranking is introduced.

Definition 3.1. We call a ranking an ordered n-tuple $(X_1, X_2, \ldots, X_n)$, arranged from the lowest ranked element X_1 to the highest ranked element X_n.

In general, there does not exist a ranking $(X_1, X_2, \ldots, X_n)$ such that for every $i < j$, every voter has $X_i \preceq X_j$. This motivates the need for a procedure to provide a ranking that is optimal in some sense.

3.2.2 Computing the Consensus Ranking

After traversing all possible permutations of the candidates' rankings and computing a score for each, the ranking with the highest score can be identified, and this becomes the consensus ranking according to the Kemeny-Young method.

3.2.3 Consensus Matrix

Once the consensus ranking is computed, the number of pairwise comparisons can be organized into a consensus matrix.

3.3 Inference in Machine Learning

Inference in machine learning is the process of obtaining results or predictions from a model that has been trained using data. After the model has been trained using an appropriate algorithm on the training dataset, inference is used to obtain predictions for new or unseen data that was not part of the training.

It is important to note that inference does not involve updating the model's parameters. These are fixed after training, and inference uses these parameters to generate predictions. This makes inference faster than training, as it does not require any changes to the model's parameters.

3.4 Challenges in Multi-Class Classification

Let's consider the generalization of the binary classification task to K classes. Challenges in multi-class classification can present specific difficulties that differ from those in binary classification tasks. Classifiers may produce a K-dimensional vector with values corresponding to the confidence for each class. Additionally, there are two other approaches to handling multi-class classification: "One-vs-One" (OvO) and "One-vs-Rest" (OvR), each with its own advantages and limitations.

3.4.1 One-vs-One (OvO):

In OvO (One vs One), a binary classifier is trained for each pair of classes. For N classes, this results in K(K-1)/2 classifiers. Although it requires more classifiers, OvO can be computationally efficient for algorithms that scale well with the number of samples. An example of such an algorithm is the support vector machines (SVM) method. However, OvO may suffer from class distribution imbalance in the resulting binary data.

3.4.2 One-vs-Rest (OvR):

In OvR (One vs Rest), a binary classifier is trained for each class, considering instances of that class as positive examples, and instances of all other classes as negative examples. OvR usually results in N classifiers for N classes. While OvR can more effectively handle class imbalances than OvO, it may lead to less accurate results if the classifiers are biased towards the dominant class.

3.4.3 Normalized Probabilities

Neural networks and other classifiers may produce multi-dimensional output vectors. A common practice in classification is to normalize the output vectors. In this approach, for each example, we generate a vector with a length equal to the number of classes, whose elements sum to 1. Each element of the vector is interpreted as the "probability" of the class at the respective position being the true class. Probability is in quotes because without calibration, these numbers are not close to true probabilities.

3.5 Classifier Calibration

Calibration in machine learning is the process of adjusting the predictions or probabilities provided by models so that they are as close as possible to the actual values or probabilities of the events we are trying to predict [22]. This process is carried out with the aim of maximizing confidence in the model and improving its accuracy.

Calibration methods are applied after the model is trained. They use validation data to compose a calibration function for a pre-trained model, which transforms the model's predictions to be better calibrated.

Generally, methods can be divided into parametric and non-parametric [23]. Parametric methods find the parameters of a function that best approximates the model's predictions to the realized probabilities. An example of a parametric method is the Brier method [24]. Non-parametric methods only consider the ordering of the predicted classes. Isotonic calibration is an example of a non-parametric method [25].

It has been shown how the likelihood estimation of the result can be obtained in terms of intuitionistic fuzzy sets. This represents a non-parametric calibration method, based on the method of intercriteria analysis.

3.6 Aggregating Results

The new approach for aggregation is obtained by taking the rankings from different trained classifiers and selecting a winner using the Kemeny-Young method. To use the new approach for outputting results quickly enough, a standard procedure for transforming the Kemeny-Young method into an optimization problem is applied. The input to the procedure consists of the different rankings from the classifiers that are subject to aggregation.

3.6.1 Application

An ensemble of classifiers has been used, on which the method described in section 3.6 has been applied. Nine classifiers have been trained on the MRI image dataset for diagnosing Alzheimer's disease. The type of classifiers is not essential, as the aggregation method is universal.

The confusion matrices for the aggregated model using the new method and for a single classifier are shown in Figure 3.2. Aggregation significantly improves the number of correct predictions (the main diagonal of the confusion matrix) compared to individual classifiers shown below. We observe improvement ranging from 27 to 103 more correct predictions compared to the original nine models.

3.7 Practical implementation of the inference method

For ranks of small length, one way to compute the optimal aggregation is by comparing the results of all possible ranks – an approach known as brute force method [26]. This approach has algorithmic complexity $o(n!)$ due to the need to traverse all permutations of the categories.

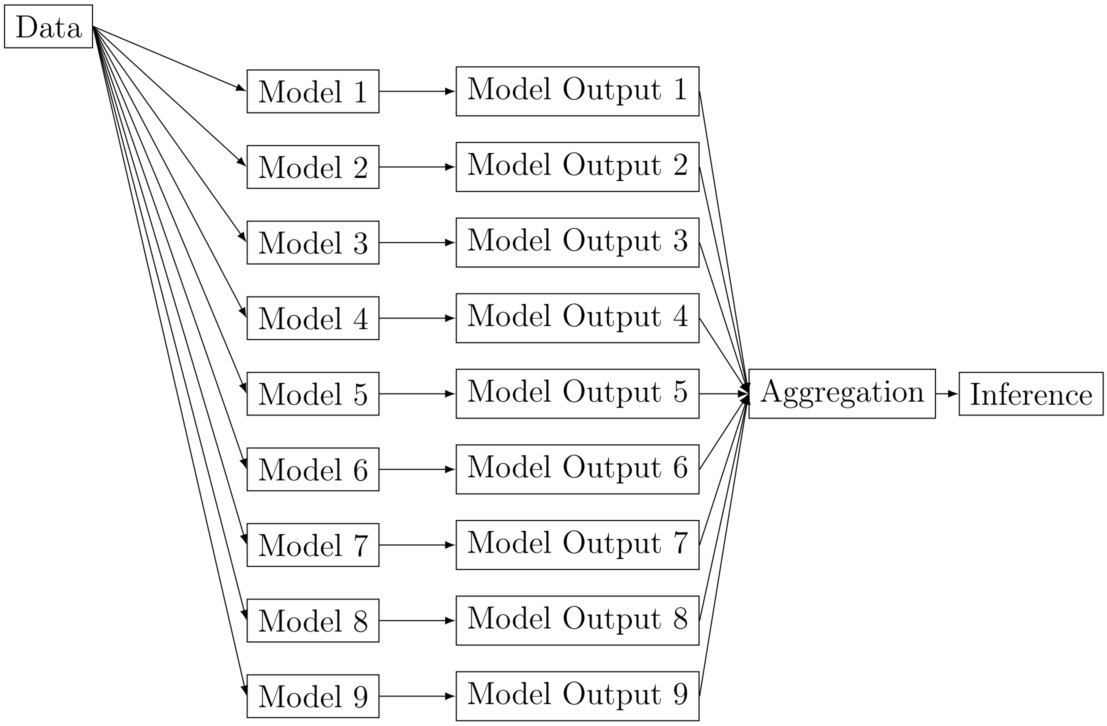


Figure 3.1: Illustration of the inference approach. Three models are trained using resnet18, resnet34, and resnet50, but this is not essential for the aggregation procedure before inference.

3.7.1 Formulating an Integer Optimization Problem

The Kemeny-Young method is NP-hard even for only four voters [27]. For practical application, a directed graph with weights $G = (V, E)$ is constructed, with the categories as vertices. The edges are defined as follows: for each pair of candidates i, j, let #>{i>j} denote the number of voters who rank i higher than j. We place an edge between each pair i,j with weight $w_e = |\#\{i > j\} - \#\{j > i\}|$ (if non-zero). The orientation of the edge is from the less preferred to the more preferred vertex.

The formulation is based on an alternative interpretation of the optimal aggregation via Kemeny-Young, which minimizes the weights of the edges that disagree:

$$\min \sum_{e \in E} w_e x_e$$

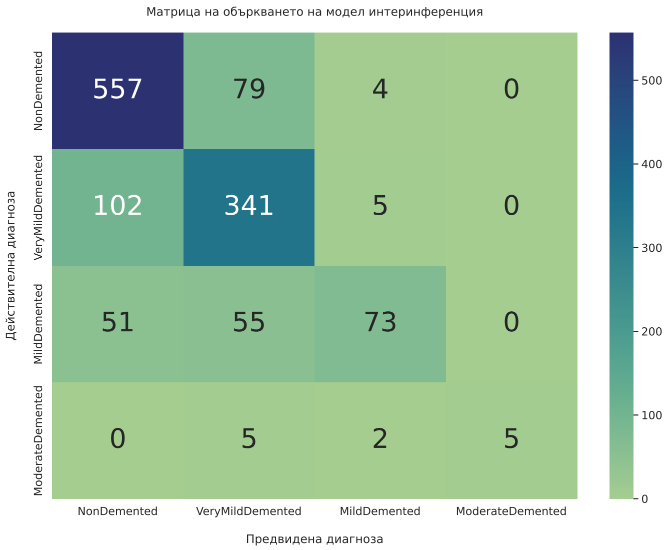
subject to:

$$\forall i \neq j \in V, x_{ij} + x_{ji} = 1 \forall i \neq j \neq k \neq i \in V, x_{ij} + x_{jk} + x_{ki} \geq 1$$

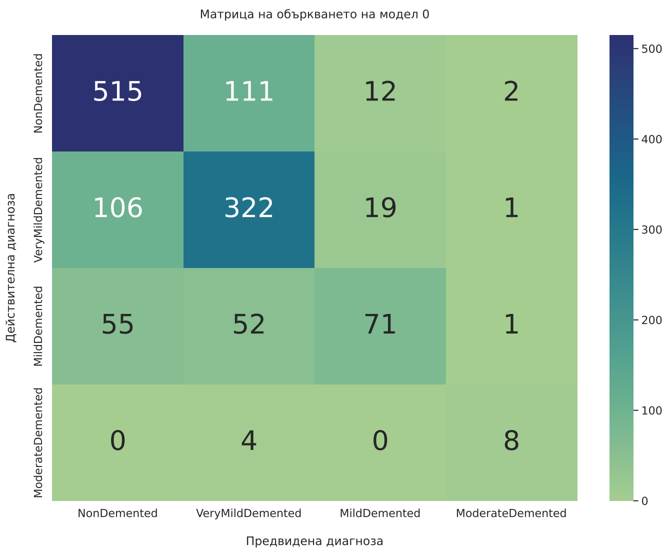
In the above problem, all variables are integer binary numbers. They have the following interpretation:

$x_{ij} = 1$ if in the aggregated rank i is ranked lower than j.

The constraints essentially enforce that the variables define a total order. The first set of constraints impose antisymmetry and completeness: either i is ranked lower than j, or vice versa. The second set of constraints impose transitivity [26].



(a) Results after aggregation using the method described in section 3.6. Aggregation significantly improves the number of correct predictions (the main diagonal of the confusion matrix) compared to individual classifiers shown below.



(b) The *ResNet18* model has more errors compared to Figure 3.2a.

Figure 3.2: Results of the different inference methods.

Chapter 4

Assessment with Inter-Criteria Analysis of Machine Learning Outputs

In this chapter, a procedure based on inter-criteria analysis is applied to the method introduced in section 3.6. The specific method of outputting is not essential for the presented procedure. It is sufficient for pairs of preferences to exist for each example.

Let us have an aggregated ranking produced by models c_1, c_2, ..., c_m. These models will be the objects in the inter-criteria analysis (see section 1.3).

With i, j, we denote classes from the classification task. Let the total number of classes be K. We will associate different classes with the criteria in the inter-criteria analysis (see section 1.3).

Figure 4.1 presents classifiers c_k and c_l. Also shown are the predicted values from classifier c_k – $a_{i,k}$, $a_{j,k}$ for classes i and j, respectively.

Let the aggregated inference have chosen the winning class j. We compute $\mu_{i,j}$, $\nu_{i,j}$ as defined in Formula 1.1, for every class i, different from j. That is, we count in what fraction of cases two randomly chosen classifiers k and l will agree or disagree when ranking classes i and j, respectively for $\mu_{i,j}$, $\nu_{i,j}$. For the relation in the inter-criteria analysis, we can use a standard comparison of the output values, but a more complex relation can be used that better reflects the degree of uncertainty, as done in Section 4.1.

We average $\mu_{i,j}$ and $\nu_{i,j}$ over i to obtain an intuitionistic fuzzy estimate of the predicted class after aggregation j, (μ_j, ν_j).

4.1 Application on Alzheimer's Disease Classifiers

An application of the method proposed in Section 4 is considered on the aggregation of the nine Alzheimer's disease classifiers used in Section 3.6.1. To obtain more reliable results for uncertainty, we define predictions for two classes as equal if their difference is less than 0.25. The Python implementation can be found in the IntervalNumber class.

The results from the applied method are presented using intuitionistic fuzzy interpretation triangles. The examples from each cell of the confusion matrix in Figure 3.2a are represented by separate intuitionistic fuzzy interpretation triangles. On Figure 4.2, the lower triangle corresponds to the cell with row MildDemented and column NonDemented in Figure 3.2a with a total of 51 examples.

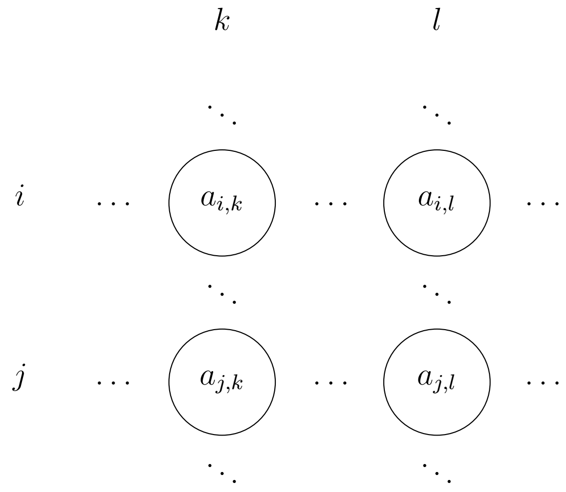


Figure 4.1: Normalized probabilities for classes i, j, and for models k, l. The element of the matrix $a_{i,k}$ is the prediction of model k for class i. We apply inter-criteria analysis by counting the number of agreements and disagreements on the relations $R_l(a_{i,k}, a_{j,k})$ and $R_k(a_{i,l}, a_{j,l})$. The relations R_l and R_k, in general, may be different.

It is interesting to note that examples with incorrect predictions have higher uncertainty. This shows that assessments with inter-criteria analysis can be used to detect model errors. Section 4.2 discusses how we can use this fact more specifically.

4.2 Results on Threshold Values

A method is presented to increase the accuracy of predictions by selecting only those predictions that are deemed reliable, using threshold values similar to the classification introduced in [28]. Predictions that are not considered reliable can be further examined by an expert. The proportion of examples that do not pass through the expert is called coverage.

For this purpose:

1. We set a preliminary accuracy for the derived results.

2. We find a degree of membership and uncertainty above which the predictions meet the specified accuracy, with the greatest coverage.

For the Alzheimer's predicting model, a precision of 90% is fixed. It is found that a membership threshold to the class above 0.9 and an uncertainty threshold below 0.3 fulfill this accuracy with the greatest coverage.

In Figure 4.8, the confusion matrix obtained when applying the specified threshold values is presented. Comparative analysis shows that the results are significantly more accurate compared to the way they were processed before the introduction of these thresholds.

With these thresholds, predictions were obtained for 777 out of a total of 1279 examples in the test dataset. This results in coverage of 60.75% with an accuracy of 90%.

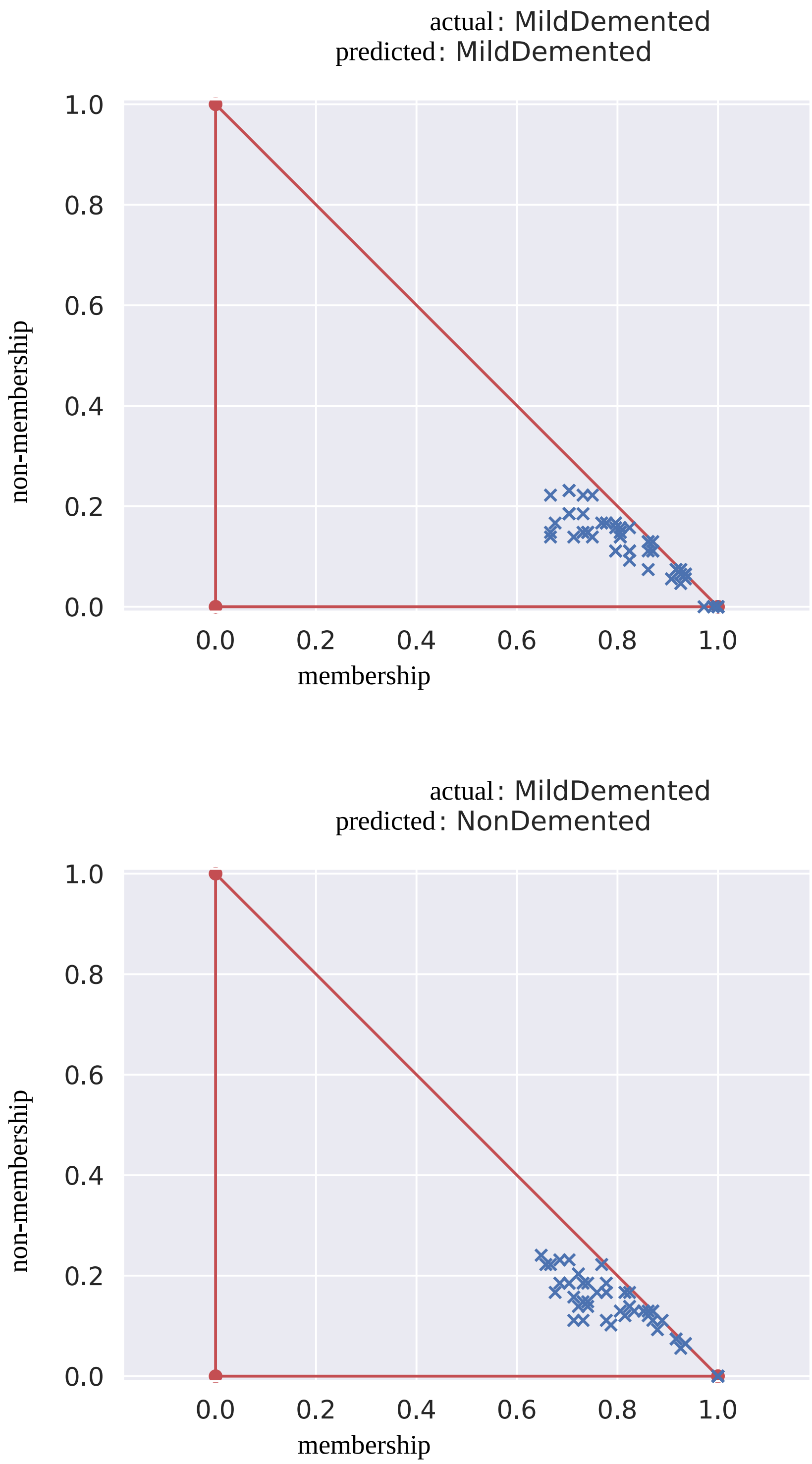


Figure 4.2: Intuitionistic fuzzy interpretation triangles of the examples from the confusion matrix in Figure 3.2a. Above each graph, it is indicated which row and column of the matrix the examples are from.

Figure 4.8: Confusion matrix after applying a membership above 0.9 and uncertainty below 0.3.

4.3 Aggregation of Results and Evaluation with Inter-criterion Analysis in OvO

The inter-criterion approach for inference can be applied in the OvO procedure. In the standard inference procedure in the OvO method, aggregation is performed by averaging the predicted scores of all $K(K-1)/2$ classifiers. This is the procedure in *scikit-learn* [29]. This approach does not require much computation compared to the proposed approach in section 3.6. In most cases, this is not a problem, because the final model is an ensemble of small constituent models. To apply section 3.6, we find an order that minimizes the disagreements in the $K(K-1)/2$ classifiers using the Kemeny-Young method.

4.4 Achieving Order

In [30], a result concerning the ordering on intuitionistic fuzzy pairs, generated by the arithmetic mean power (M_p) for $p > 0$, is proven. A family of orderings on intuitionistic fuzzy pairs, generated by the weighted arithmetic mean power (M_p^α), is introduced, and it is proven that a similar result holds for them. The considered orderings naturally extend the classical partial order and allow comparison of previously incomparable alternatives. It is shown that there exists $p > 0$ for which under the new ordering all elements become comparable.

Definition 4.1 ([31])**.** We denote intuitionistic fuzzy pairs (IFP) as $\langle a, b\rangle$.

$$a + b \leq 1 \tag{4.1}$$

In [32], the following definition is given for comparing intuitionistic pairs.

Definition 4.2. For two intuitionistic fuzzy pairs: $u = \langle u_1, u_2\rangle$ and $v = \langle v_1, v_2\rangle$, we say that u is less than or equal to v, i.e.:

$$u \leq v \Leftrightarrow \begin{cases} u_1 \leq v_1 \\ u_2 \geq v_2 \end{cases} .$$

Definition 4.3 ([33], p. 175)**.** The power mean of two non-negative numbers x and y is defined as follows:

$$M_p(x, y) = \left(\frac{x^p + y^p}{2}\right)^{\frac{1}{p}}. \tag{4.2}$$

The power mean has the following interesting property [33] (p. 175):

$$M_p(x, y) \leq M_q(x, y) \text{ за } p \leq q.$$

Definition 4.4 (see [33], p. 175)**.** The weighted power mean with weight α of two non-negative numbers x and y is defined as follows:

$$M_p^\alpha(x, y) = (\alpha x^p + (1 - \alpha) y^p)^{\frac{1}{p}}, \tag{4.3}$$

където $\alpha \in (\frac{1}{2}, 1)$.

Definition 4.5 ([31])**.** Given two IFPs $u = \langle u_1, u_2\rangle$ and $v = \langle v_1, v_2\rangle$, we say that u is preferred to v by the geometric mean power p with the first component preferred, and we write $u \preceq_{\mu;M_p} v$, if

$$\begin{cases} 1 - u_1 \geq 1 - v_1 \\ M_p(1 - u_1, u_2) \geq M_p(1 - v_1, v_2). \end{cases} \tag{4.4}$$

Definition 4.6. Given two IFPs $u = \langle u_1, u_2\rangle$ and $v = \langle v_1, v_2\rangle$, we say that u is preferred to v by the geometric mean power p with weight α, and we write $u \preceq_{\mu;M_p^\alpha} v$, if

$$\begin{cases} 1 - u_1 \geq 1 - v_1 \\ M_p^\alpha(1 - u_1, u_2) \geq M_p^\alpha(1 - v_1, v_2). \end{cases} \tag{4.5}$$

The following lemma has been proven and used:

Lemma 4.1. For every constant $c \in (\frac{1}{2}, 1)$ and for all $t \in (0, 1-c)$, the following inequality holds:

$$(2c - 1)\ln(2c - 1) > (c - t)\ln(c - t) + (c + t)\ln(c + t). \tag{4.6}$$

The following lemma has been proven:

Lemma 4.2. Let $a \in (\frac{1}{2}, 1)$. Then, for all $k \in \left(1, \frac{a}{1-a}\right)$ and $t \in (0, a - k(1 - a)]$, it holds that

$$ka\ln(a) + (k(1 - a) + t)\ln(k(1 - a) + t) < t\ln(t). \tag{4.7}$$

4.4.1 Main Results

The main result is formulated and proven in the following theorem.

Theorem 4.1. Let $u = \langle u_1, u_2\rangle$ and $v = \langle v_1, v_2\rangle$. If

$$u \preceq_{\mu;M_p} v \tag{4.8}$$

for some $p > 0$, then

$$u \preceq_{\mu;M_q} v \tag{4.9}$$

for all $q > p$.

This result allows the generation of all orderings with $p > 0$ as a transition from partial to total ordering. This is achieved at the boundary transition $p = \infty$. From (4.9), we know that this ordering preserves the order, i.e., we increase the number of comparable elements consistently. Introducing the weight α allows fine-tuning of the ordering, moving closer and closer to the linear ordering generated with a larger weight on the first component. The following theorem has been proven.

Theorem 4.2. Let $u = \langle u_1, u_2\rangle$ and $v = \langle v_1, v_2\rangle$. Let $\alpha \in (\frac{1}{2}, 1)$. Then, if

$$u \preceq_{\mu;M_p^\alpha} v \tag{4.10}$$

for some $p > 0$, we have

$$u \preceq_{\mu;M_q^\alpha} v \tag{4.11}$$

for all $q > p$.

Let $u = \langle u_1, u_2\rangle$ and $v = \langle v_1, v_2\rangle$. Let $\alpha_1 \in (\frac{1}{2}, 1), \alpha_2 \in (\alpha_1, 1)$. Then, if for some $p > 0$,

$$u \preceq_{\mu;M_p^{\alpha_1}} v \tag{4.12}$$

it follows that

$$u \preceq_{\mu;M_p^{\alpha_2}} v. \tag{4.13}$$

4.4.2 Application on the Results for Staging Alzheimer's Disease

Python code is presented for finding the value of p that converts intuitionistic fuzzy pairs into a total ordering. Thus, a value of $p = 12.81$ is found achieving a total ordering for the results from section 4.1.

Chapter 5

Improving the Efficiency of the Inter-criterion Analysis Algorithm

This chapter demonstrates that computing the inter-criterion counters can be done in $O(n \log n)$ (quasilinear complexity). Up to this point, all implementations have used $O(n^2)$ computations [34, 35, 36], which does not allow processing data over hundreds of thousands. The results are presented for the first time in [37].

5.1 Notation

All vectors are n-dimensional, whose elements can be ordered using the relation "$\leq$". The elements of different vectors do not necessarily belong to the same set, but nevertheless we will write "$\leq$" for each of them.

Definition 5.1. Let k and l be two n-dimensional vectors. We say that indices i, j, $i < j$ are in **disagreement** [38], if and only if

$$k_i \leq k_j \wedge l_i > l_j \vee k_i > k_j \wedge l_i \leq l_j$$

Definition 5.2. We denote by count_disagreements(k, l) the number of disagreements between k and l.

There are $n(n-1)/2$ such combinations of indices i, j, $i < j$, which can be trivially traversed in $O(n^2)$.

It has been shown that the number of disagreements between k and l, count_disagreements(k, l), can be computed in $O(n \log n)$. Previous implementations run in $O(n^2)$, as many as the combinations of indices.

Definition 5.3. We say that indices i, j, $i < j$ are in **equality (indefiniteness)** [38], if and only if

$$k_i = k_j \wedge l_i = l_j$$

Equality between elements can be defined naturally in the presence of an ordering.

Definition 5.4. We say that indices i, j, $i < j$ are in **agreement** [38], if neither the definition for disagreement nor for indefiniteness is satisfied.

Definition 5.5. An **inversion** in a vector v refers to combinations of indices i, j, $i < j$, for which $v_i > v_j$. [39]

The inversions of a vector can be computed in $O(n \log n)$, as shown in [39]. The algorithm is based on a modification of Merge Sort, which has a complexity of $O(n \log n)$.

For vectors that allow equalities between their elements, we will introduce the following convenient notation.

Definition 5.6. We call $\hat{v}$ the **enumerated vector** of vector v[1], and we define its elements with ordered pairs:

$$\hat{v}_i = (v_i, i)$$

We extend the ordering of the elements of v to a lexicographic ordering for $\hat{v}$ in a natural way. Since the indices are unique, equalities among the elements of enumerated vectors are not possible, even if there are equalities in the original vector.

Definition 5.7. The number of inversions in a vector v will be denoted by $\text{count\_inversions}(v)$.

Definition 5.8. Let k and l be two vectors. Let us introduce the notation $\text{sort}_k(l)$, which sorts the elements of l according to the ordering of k.

Let us express the above definition using a sorting permutation. Consider the permutation σ that sorts k. For indices i, j, $i < j \Leftrightarrow k_{\sigma(i)} \leq k_{\sigma(j)}$. Then for the i-th element we have

$$\text{sort}_k(l)_i = l_{\sigma(i)}$$

Sorting is an operation that can be performed in $O(n \log n)$ operations [39].

The following proposition has been proven:

Proposition 5.1.

$$\text{count\_disagreements}(k, l) = \text{count\_inversions}(\text{sort}_{\hat{k}}(\hat{l}))$$

5.2 Computing the Inter-criterion Counters

The definitions of the inter-criterion counters from [38] for two criteria k and l, which for brevity we identify with the vectors of values over the n objects. S^{μ}_{kl} is the count of agreements, S^{ν}_{kl} is the count of disagreements, S^{π}_{kl} is the count of equalities.

It is shown how to compute S^{π}_{kl} in $O(n \log n)$.

5.3 Program Implementation

The algorithm has been implemented and tested in Python. Automatic testing has been conducted with the Hypothesis library for randomized testing, covering a wide range of cases, establishing that the two implementations in the code with $O(n \log n)$ and $O(n^2)$ give identical results.

With this approach, the time for applying the inter-criterion analysis in [40] is reduced.

[1]Similar to the enumerate function in Python, which gives (i, v_i)

5.4 Previous Results

The degree of disagreement in inter-criterion analysis [38] differs from the Kendall's τ metric [41] in that it allows for equality between elements. It is known that Kendall's τ metric can be reduced to counting inversions, for example [42]. The equalities in the definition of inter-criterion analysis do not hinder reducing count_disagreements to counting inversions. In [41], Chan and Patrascu's result [43] is also mentioned, which allows even faster counting of inversions with $O(n\sqrt{\log n})$.

Since inversions are not used in the computation of the equality counter S^{π}_{kl}, the question of whether inter-criterion counters can be computed in $O(n\sqrt{\log n})$ remains open for now.

Conclusion – Summary of Obtained Results

In this dissertation work, a software method has been implemented, which provides good results in the staging of Alzheimer's disease. The code has the following significant advantages:

- easy to adapt for imaging diagnostics of other diseases
- short training time
- small amount of code

Automatic diagnostics allows for preventive monitoring of a large number of patients.

A new inference method has been developed, based on the Kemeny-Young method, which provides stable aggregation of an ensemble of classifiers. The new method has been implemented with software code that solves the NP-hardness problem in multi-class classification tasks.

When using classification models, it is important not only to obtain the most accurate result possible, but also to have an assessment of its accuracy. For this purpose, a new method for evaluating the accuracy of classification models has been developed, based on inter-criterion analysis. The new method provides an interpretation of the obtained results in terms of intuitionistic fuzzy logic.

Applying inter-criterion analysis to data with multiple dimensions becomes impractical due to current implementations with quadratic complexity. In response to this problem, an algorithm with quasi-linear accuracy has been proposed.

Directions for Future Work

- Transfer the developed diagnosis using deep convolutional neural networks for staging Alzheimer's disease to other diseases that require imaging diagnostics. Suitable candidates include Parkinson's disease, multiple sclerosis, and other neurological disorders.
- Implement the new inference method based on the Kemeny-Young method, not only for ensembles but also for OvO classifiers. This will enable its application to a wide class of classification models.
- Confirm or reject the possibility of further acceleration of inter-criterion analysis to $O(n\sqrt{\log n})$. Finding a faster algorithm will allow inter-criterion analysis to be computed faster and for larger volumes of data.

Main Contributions of the Dissertation

Scientific Contributions

1. Proposed a inference method based on the Kemeny-Young method.
2. Proposed a method for evaluating the inferences of classification tasks in terms of intuitionistic fuzzy sets, based on inter-criterion analysis.
3. Proposed a method for applying threshold values to the degrees of membership and uncertainty, which significantly improves the accuracy of the selected results.
4. Proposed an algorithm to improve the speed of inter-criterion analysis to $O(n \log(n))$.

Scientific and Practical Contributions

1. Implemented neural network code for diagnosing stages of Alzheimer's disease.
2. Achieved model sensitivity of 70%, and specificity of 85%.
3. Developed Python code for the new inference method based on the Kemeny-Young method. Used the optimization library *ortools* to achieve practical execution times.
4. The developed inference method provides consistent results, which are construction-wise error-tolerant.
5. Developed Python code for the new method for evaluating the inferences of classification tasks in terms of intuitionistic fuzzy sets, based on inter-criterion analysis. The code defines a non-trivial equality relation between classifier predictions.
6. Developed Python code for applying threshold values to the degrees of membership and uncertainty. Applied to the Alzheimer's disease classification task, achieving significant improvement in the accuracy of the selected results. Achieved model sensitivity of 80%, and specificity of 95%.
7. Developed Python code to improve the speed of inter-criterion analysis to $O(n \log(n))$. Current implementations using $O(n^2)$ complexity algorithms have been compared with the new implementation using the package for automated randomized testing *Hypothesis*.

List of Publications Related to the Dissertation Work

Scientific Articles in Journals with Impact Factor

1. Vassilev P., T. Stoyanov, L. Todorova, A. Marazov, V. Andonov, N. Ikonomov. Orderings over Intuitionistic Fuzzy Pairs Generated by the Power Mean and the Weighted Power Mean. Mathematics 2023, 11(13), 2893; https://doi.org/10.3390/math11132893. IF = 2.4 (2022). Category Q1

Scientific Articles in Journals with SJR Rank in SCOPUS

2. Marazov A., Shanon, A. Improved Speed of InterCriteria Analysis. International Journal Bioautomation, (inpress); SJR (SCImago Journal Rankings) (2022): 0.159

Scientific Articles in Journals with SJR Rank in SCOPUS

3. Marazov A. Diagnostic Imaging of Alzheimer's Disease with Convolutional Neural Networks: Implementation with fastai. Yearbook of the "Informatics" Section, Union of Scientists in Bulgaria; Volume 11, 2021, 15–23; ISSN 1313–6852.

4. Marazov A. "Inference in Machine Learning based on the Kemeny-Young Method". In: Annual of the "Informatics" Section, Union of Scientists in Bulgaria 12 (2022/2023), pp. 72–94. ISSN: 1313-6852. (in Bulgarian)

Participation in Scientific Forums

5. Ignatova V., Marazov. A. Spectrum of vascular cognitive impairment according to the type of brain vascular damage. 10th EAN Congress in Helsinki, Finland on 29 June-02 July, 2024. (in press).

6. Marazov A., Vassilev P. Assessment with InterCriteria Analysis of Outputs in Machine Learning. 3rd International Symposium on Bioinformatics and Biomedicine, BioInfoMed'2024, 4–6 July 2024 (in press).

List of Citations Related to the Dissertation Work

One citation has been observed for the scientific publications related to the dissertation work:

1. Vassilev, P, T. Stoyanov, L Todorova, A Marazov, V Andonov, N Ikonomov. Orderings over Intuitionistic Fuzzy Pairs Generated by the Power Mean and the Weighted Power Mean. https://www.mdpi.com/2227-7390/11/13/2893 Mathematics 2023, 11(13), 2893; https://doi.org/10.3390/math11132893.

Cited in:

Todorova, S. Overview of Publications on Indexed Matrices. Annual of the "Informatics" Section, Union of Scientists in Bulgaria, Volume XII, 2022-2023, pp. 32-62.

Bibliography

[1] K. S. Neetha and D. L. Narayan. "Segmentation and Classification of Brain Tumour Using LRIFCM and LSTM". In: *Multimedia Tools and Applications* (2024), pp. 1–26.

[2] R. K. Singh and D. K. Choubey. "Alzheimer's Dementia and its Various Biomarker Details with Several Detection Methodologies-A Comprehensive Review". In: *2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. IEEE. Jan. 2024, pp. 657–662.

[3] S. Suganyadevi et al. "Deep Learning Based Alzheimer Disease Diagnosis: A Comprehensive Review". In: *SN Computer Science* 5.4 (2024), p. 391.

[4] Marvin Minsky and Seymour Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, 1969.

[5] Kaiming He et al. "Deep Residual Learning for Image Recognition". In: *CoRR* abs/1512.03385 (2015). arXiv: `1512.03385`. URL: `http://arxiv.org/abs/1512.03385`.

[6] K Atanassov. "Intuitionistic fuzzy sets". In: *VII научна сесия на ИТКР, София, юни* (1983).

[7] K.T. Atanassov. "Research on intuitionistic fuzzy sets in Bulgaria". In: *Fuzzy Sets Syst.* 22 (1987), p. 93.

[8] K T Atanassov and G Gargov. "Interval-valued intuitionistic fuzzy sets". In: *Fuzzy Sets Syst* 31 (1989), pp. 343–349.

[9] K.T. Atanassov. "Remarks on the intuitionistic fuzzy sets". In: *Fuzzy Sets Syst.* 51 (1992), pp. 117–118.

[10] K.T. Atanassov. "Research on intuitionistic fuzzy sets, 1990–1992". In: *Fuzzy Sets Syst.* 54 (1993), pp. 363–364.

[11] K.T. Atanassov. "New operations defined over the intuitionistic fuzzy sets". In: *Fuzzy Sets Syst.* 61 (1994), pp. 137–142.

[12] К Атанасов. *Актуални проблеми на науката*. БАН, Център за научна информация, 1989.

[13] Krassimir Atanassov, Sotir Sotirov, and Tania Pencheva. "Intuitionistic Fuzzy Deep Neural Network". In: *Mathematics* 11.3 (2023), p. 716. DOI: `10.3390/math11030716`.

[14] Alexander Marazov. "ОБРАЗНА ДИАГНОСТИКА НА БОЛЕСТТА НА АЛЦХАЙМЕР С КОНВОЛЮЦИОННИ НЕВРОННИ МРЕЖИ: ИМПЛЕМЕНТАЦИЯ С *FASTAI*". In: *Annual of "Informatics" Section Union of Scientists in Bulgaria* 11 (2021/2022), pp. 1–13. ISSN: 1313-6852.

[15] S Dubey. *Alzheimer's Dataset (4 class of Images)*. 2020. URL: `https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images`.

[16] J Deng et al. "Imagenet: A large-scale hierarchical image database". In: *2009 IEEE conference on computer vision and pattern recognition*. 2009, pp. 248–255.

[17] Leslie N. Smith. "Cyclical Learning Rates for Training Neural Networks". In: *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2017, pp. 464–472. DOI: `10.1109/WACV.2017.58`.

[18] Anusha Jayasimhan and P Pabitha. "A comparison between CPU and GPU for image classification using Convolutional Neural Networks". In: *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*. 2022, pp. 1–4. DOI: `10.1109/IC3IOT53935.2022.9767990`.

[19] А. Маразов. "Инференция с интеркритериален анализ и метода на Kemeny-Young". In: *Annual of "Informatics" Section Union of Scientists in Bulgaria* 12 (2022/2023), pp. 72–94. ISSN: 1313-6852.

[20] John Kemeny. *Mathematics without Numbers*. 1959.

[21] Tim Roughgarden. *Algorithms Illuminated (Part 4): Algorithms for NP-Hard Problems*. English. Soundlikeyourself Publishing, LLC, July 2020, p. 273. ISBN: 0999282964.

[22] Bianca Zadrozny and Charles Elkan. "Transforming classifier scores into accurate multiclass probability estimates". In: *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (2002), pp. 694–699.

[23] Telmo Silva Filho et al. "Classifier Calibration: A Survey on How to Assess and Improve Predicted Class Probabilities". In: *Machine Learning* 112.9 (Sept. 2023), pp. 3211–3260. ISSN: 1573-0565. DOI: `10.1007/s10994-023-06336-7`. URL: `https://doi.org/10.1007/s10994-023-06336-7`.

[24] G. W. Brier. "Verification of Forecasts Expressed in Terms of Probability". In: *Monthly Weather Review* 78.1 (1950), pp. 1–3.

[25] R. E. Barlow and H. D. Brunk. "The Isotonic Regression Problem and Its Dual". In: *Journal of the American Statistical Association* 67.337 (1972), pp. 140–147.

[26] V Conitzer, A Davenport, and J Kalagnanam. "Improved Bounds for Computing Kemeny Rankings". In: 2006.

[27] JJ Bartholdi III, CA Tovey, and MA Trick. "The Computational Difficulty of Manipulating an Election". In: (1989).

[28] Krassimir Atanassov, Vassia Atanassova, and Georgi Gluhchev. "Notes on Intuitionistic Fuzzy Sets". In: *InterCriteria Analysis: Ideas and Problems* 21.1 (2015), pp. 81–88. ISSN: 1310–4926.

[29] Fabian Pedregosa et al. "Scikit-learn: Machine learning in Python". In: *Journal of machine learning research* 12.Oct (2011), pp. 2825–2830.

[30] P. Vassilev et al. "Orderings over Intuitionistic Fuzzy Pairs Generated by the Power Mean and the Weighted Power Mean". In: *Mathematics* 11.13 (2023), p. 2893. DOI: 10.3390/math11132893.

[31] P. Vassilev and T. Stoyanov. "On Power Mean Generated Orderings between Intuitionistic Fuzzy Pairs". In: *Advances in Fuzzy Logic and Technology 2017. EUSFLAT IWIFSGN 2017.* Ed. by J. Kacprzyk et al. Vol. 643. Advances in Intelligent Systems and Computing. Cham, Switzerland: Springer, 2018, pp. 476–481. DOI: 10.1007/978-3-319-66827-7_44.

[32] Krassimir T. Atanassov. *On Intuitionistic Fuzzy Sets Theory*. Berlin: Springer, 2012.

[33] P.S. Bullen. *Handbook of Means and Their Inequalities*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2003.

[34] D. Mavrov et al. "InterCriteria Software Design: Graphic Interpretation within the Intuitionistic Fuzzy Triangle". In: *Proceedings of the Fifth International Symposium on Business Modeling and Software Design*. 2015, pp. 279–283.

[35] D. Mavrov. "Software for intercriteria analysis: Working with the results". In: *Annual of "Informatics" Section, Union of Scientists in Bulgaria* 8 (2015-2016), pp. 37–44. ISSN: 1313-6852.

[36] N. Ikonomov, P. Vassilev, and O. Roeva. "ICrAData – Software for InterCriteria Analysis". In: *Int J Bioautomation* 22.1 (2018), pp. 1–10.

[37] A. Marazov A. Shanon. "Improved Speed of InterCriteria Analysis". In: *International Journal Bioautomation* (2024).

[38] Krassimir Atanassov, D. Mavrov, and Vassia Atanassova. "Intercriteria Decision Making: A New Approach for Multicriteria Decision Making, Based on Index Matrices and Intuitionistic Fuzzy Sets". In: *Issues in Intuitionistic Fuzzy Sets and Generalized Nets* 11 (2014), pp. 1–8.

[39] Tim Roughgarden. *Algorithms Illuminated (Part 1): The Basics*. Soundlikeyourself Publishing, 2017. Chap. 3.2, p. 62. ISBN: 9780999282908.

[40] L. Todorova, J. Surchev, and A. Marazov. "Assessment of the Factors Influencing the Survival of the Ventricular Shunt in Infantile Hydrocephalus". In: *International Journal of Bioautomation* 27.2 (2023), pp. 107–116. DOI: 10.7546/ijba.2023.27.2.000926.

[41] Wikipedia. *Kendall tau distance*. URL: https://en.wikipedia.org/wiki/Kendall_tau_distance.

[42] W. Knight. "A Computer Method for Calculating Kendall's Tau with Ungrouped Data". In: *Journal of the American Statistical Association* 61.314 (1966), pp. 436–439. DOI: 10.2307/2282833. JSTOR: 2282833.

[43] Timothy M. Chan and Mihai Pătraşcu. "Counting Inversions, Offline Orthogonal Range Counting, and Related Problems". In: *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*. 2010, p. 161. ISBN: 978-0-89871-701-3. DOI: 10.1137/1.9781611973075.15.