Integrating ML Diagnosis and Multi-Criteria Decision Making for Activity Recommendations in Alzheimer's and Parkinson's Patients

1st Carlos Freitas
Polytechnic School of Porto
School of Engineering
Porto, Portugal
1240521@isep.ipp.pt

2nd Daniel Dias
Polytechnic School of Porto
School of Engineering
Porto, Portugal
1240145@isep.ipp.pt

3rd João Caseiro
Polytechnic School of Porto
School of Engineering
Porto, Portugal
1211334@isep.ipp.pt

4th Pedro Rocha
Polytechnic School of Porto
School of Engineering
Porto, Portugal
1191689@isep.ipp.pt

5th Vitor Castro
Polytechnic School of Porto
School of Engineering
Porto, Portugal
1140547@isep.ipp.pt

***Abstract*—This paper reviews the state of the art for various machine learning techniques applied to Alzheimer's and Parkinson's disease diagnosis. We focus on XGBoost, Random Forests, Logistic Regression, Support Vector Machines (SVMs), and Decision Tree Classifiers, discussing their relevance in healthcare applications and their potential to support personalized activity suggestions for patients. Additionally, while several Multi-Criteria Decision Analysis (MCDA) methods were studied, such as AHP and ELECTRE, only TOPSIS and PROMETHEE II were implemented due to their superior performance for evaluating and recommending activities tailored to the needs of neurodegenerative patients based on multiple criteria, such as cognitive stimulation, safety, and autonomy. By combining ML models with these MADM frameworks, this study proposes a robust approach to improve both diagnosis and personalized care.**

***Index Terms*—Alzheimer, Parkinson, Machine Learning, Multi-Criteria Decision Analysis, Activity Recommendation, Healthcare, XGBoost, TOPSIS, PROMETHEE II.**

I. Introduction

Neurodegenerative diseases such as Alzheimer's and Parkinson's pose significant challenges for healthcare systems worldwide. Both diseases are progressive and often lead to significant cognitive and physical impairments. Traditional diagnostic methods can be invasive, time-consuming, and expensive, necessitating innovative approaches to improve early detection.

Machine Learning (ML) has emerged as a powerful tool for diagnosing complex diseases by analyzing patterns within large datasets [1]. In this paper, we explore the state of the art of ML techniques, including XGBoost [1], Random Forests [2], Logistic Regression, SVMs [3], and Decision Trees [7], highlighting their applications in diagnosing Alzheimer's and Parkinson's. Additionally, we discuss how these predictive models can be extended to suggest personalized activities that support patients' cognitive and physical well-being.

To complement ML predictions, we integrate a Multi-Criteria Decision Analysis (MCDA) framework to evaluate and rank personalized activity recommendations. MADM methods, TOPSIS (Technique for Order Preference by Similarity to Ideal Solution), and PROMETHEE II (Preference ranking organization method for enrichment evaluation), provide structured decision-making tools to consider multiple factors, such as safety, engagement, and required resources.

II. State of the Art

This section provides an overview of the theoretical foundations and existing technologies related to the application of Machine Learning (ML) and Multi-Criteria Decision Analysis (MCDA) in healthcare.

A. *Artificial Intelligence and Machine Learning*

Artificial Intelligence (AI) enables machines to simulate intelligent human behavior, including learning, reasoning, and problem-solving. Machine Learning (ML), a subset of AI, focuses on algorithms capable of learning from data and improving performance over time [1]. These technologies are increasingly used in healthcare to identify patterns in medical data and support decision-making, especially in diagnosing neurodegenerative diseases such as Alzheimer's and Parkinson's.

B. *Learning Paradigms*

ML encompasses various learning paradigms, each suited to different types of problems. This study focuses on supervised learning, a paradigm where models are trained with labeled data, associating inputs with desired outputs [2]. Supervised learning is highly effective for tasks like disease diagnosis, where labeled datasets are available. However, it requires large amounts of high-quality data, which can be costly and time-consuming to obtain.

C. Related Work

Several studies have applied machine learning to the diagnosis of neurodegenerative diseases. Below are representative examples along with a brief explanation of the methods mentioned and the results obtained:

Alzheimer's Diagnosis: Studies such as [4] and [5] explored the use of Random Forests and Support Vector Machines (SVMs) to analyze biomarkers and neuroimaging data. Random Forests, known for their ability to reduce overfitting and handle large datasets [2], achieved accuracies exceeding 90%. Meanwhile, SVMs, effective in high-dimensional datasets due to their regularization capabilities [3], also demonstrated promising results in early detection of this disease.

Parkinson's Prediction: XGBoost, a scalable and efficient implementation of gradient boosting [1], has been successfully applied to clinical and genetic data to identify risk factors associated with Parkinson's disease. This method achieved high predictive accuracy and robustness in real-world datasets [6].

General Health Applications: Logistic Regression and Decision Trees are widely used as baseline models due to their interpretability and simplicity. For instance, Decision Trees, which offer intuitive decision paths [7], proved effective in classifying patient risk levels across various clinical scenarios.

Multimodal Approaches: Recent studies have explored multimodal machine learning techniques to combine different types of data for more robust neurodegenerative disease diagnosis. For example, [15] integrated clinical, genetic, and imaging data using ensemble methods to enhance predictive accuracy in Alzheimer's diagnosis. Similarly, [16] demonstrated that combining vocal metrics with motor assessment data improved the detection of early Parkinson's symptoms. These approaches underscore the importance of leveraging diverse data sources to address the complexity of neurodegenerative diseases.

The integration of these methods and the results obtained from recent studies highlight the potential of machine learning to advance medical diagnostics, particularly in the context of complex diseases such as Alzheimer's and Parkinson's.

D. Multi-Criteria Decision Analysis (MCDA) and MADM

MCDA encompasses a set of methods and processes designed to evaluate and rank alternatives based on multiple criteria, often with conflicting objectives [8]. This approach provides a systematic methodology for decision-making, particularly in scenarios where trade-offs between criteria are required. Within MCDA, the concept of Multi-Attribute Decision Making (MADM) focuses on the evaluation of discrete alternatives based on specific attributes or characteristics, making it a key tool in complex decision-making.

1) Types of MCDA Methods: MCDA methods are broadly classified into three categories:

- **Value-Based Methods:** Methods such as AHP (Analytic Hierarchy Process) and TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) assign scores to alternatives based on weighted criteria, providing a ranking based on the distance to an ideal solution.
- **Outranking Methods:** Techniques such as ELECTRE (Elimination and Choice Expressing Reality) [12]. and PROMETHEE II (Preference Ranking Organization Method for Enrichment Evaluations) compare alternatives pairwise. PROMETHEE II, in particular, uses a preference function to calculate positive and negative flows, generating a complete ranking that allows identifying the best alternative considering the established criteria.
- **Rule-Based Methods:** These methods rely on predefined rules or thresholds to evaluate alternatives. They are particularly useful in decision-making systems in healthcare, where a clear regulatory framework is required.

2) Applications in Healthcare: MCDA methods have been used to optimize healthcare decisions in areas such as:

- **Treatment Planning:** MCDA has been used to prioritize medical interventions based on their effectiveness and cost, facilitating more informed decisions in treatment planning. [17].
- **Resource Allocation:** MCDA tools have been implemented to guide decision-making in the planning, prioritization, and allocation of resources in healthcare, ensuring a more equitable and efficient distribution. [18].
- **Activity Recommendations:** MCDA has also been applied in the evaluation and recommendation of activities tailored to individual patient profiles, improving the personalization of healthcare. [19].

III. Methodology

The methodology of this project integrates Machine Learning (ML) models with Multi-Criteria Decision Analysis (MCDA) techniques to diagnose neurodegenerative diseases and recommend personalized activities for older adults. It is important to note that the data analysis, model training, and results from the ML part are detailed in [20]. and [21]. and the steps to follow are described below:

A. Data Collection

Two datasets were used:

- **Alzheimer's Dataset:** Sourced from Kaggle [4], which contains biomarker and demographic information.
- **Parkinson's Dataset:** Sourced from Kaggle [6], including features such as vocal metrics, demographic details, and clinical data relevant for Parkinson's diagnosis.

The Alzheimer's and Parkinson's datasets were selected for their comprehensive coverage of relevant features and their suitability for binary classification tasks in neurodegenerative disease diagnosis.

B. Data Preprocessing

The datasets were preprocessed to ensure consistency and quality using the following steps:

- **Data Analysis and Cleaning:** Exploratory Data Analysis (EDA) was conducted to visualize feature distributions, investigate patterns, and identify potential outliers. To detect atypical patterns, bar charts and box plots of the data columns or features were generated, and an interquartile range (IQR) analysis was performed. In addition, irrelevant columns, such as DoctorInCharge and PatientID, were removed.

- **Feature Encoding:** Categorical variables, such as demographic details, were transformed into numerical representations using techniques like one-hot encoding or label encoding. However, as the Random Forest model handles categorical data effectively and variables such as ethnicity and gender were not highly influential for classification, encoding had a negligible effect on this specific model.
- **Feature Normalization:** Continuous features were normalized to limit their range between 0 and 1. This step was applied to ensure comparability across variables for models that rely on distance metrics, such as SVM and KNN. However, for tree-based models like Random Forest and Decision Trees, normalization was not necessary, as these models are invariant to feature scaling.
- **Feature Engineering :** Correlation analysis between features was conducted to investigate potential direct relationships among them. However, the results did not reveal significant or strong correlations between features, limiting the opportunities for dimensionality reduction or the creation of composite variables.
- **Train-Test Split:** The datasets were divided into subsets (80% for training and 20% for testing) to effectively evaluate model performance.
- **Dealing with Data Imbalances:** Techniques like SMOTE and Random SMOTE were applied exclusively to the training datasets to address class imbalances. These methods generated synthetic samples for the minority class, ensuring balanced datasets and reducing potential biases during model training.
- **Principal Component Analysis :** PCA techniques were applied to visualize and examine the distributions of the original data, as well as datasets with oversampling and combinations of oversampling and undersampling, providing further insights into the data's structure and its impact on model performance.

It is crucial to remember that although the initial processing pipeline contained normalization and encoding techniques for cross-model compatibility, the Random Forest model did not require these procedures. Because of its natural ability to manage categorical variables and unscaled continuous data efficiently, this algorithm performed well regardless of feature scaling or encoding.

C. Model Selection and Implementation

To determine the best model for the problem, several techniques were implemented using the Scikit-learn library:

- **XGBoost:** Selected for its feature selection capabilities and robustness in handling missing data [1].
- **Random Forests:** An ensemble learning method used to improve prediction accuracy and reduce overfitting [2].
- **Logistic Regression:** A baseline model for binary classification tasks, providing interpretable results [7].
- **Support Vector Machines (SVMs):** Effective in high-dimensional spaces and pattern recognition tasks [3].
- **K-Nearest Neighbors (KNN):** A non-parametric supervised learning classifier that uses proximity to make predictions.
- **Decision Trees:** Intuitive models capable of modeling complex decision boundaries [7].

The models were trained and validated using preprocessed datasets, with hyperparameters tuned through GridSearchCV. After selecting the optimal hyperparameters, the models were trained on the preprocessed datasets, and their performance was compared using techniques such as ROC curves, confusion matrices, and metrics like precision, recall, F1-score, and accuracy to identify the most suitable approach.

D. Performance Evaluation Metrics

To assess model performance, the following metrics were used:

- **Accuracy:** Proportion of correctly predicted samples across all classes [13].
- **Precision, Recall, and F1-Score:** Metrics for evaluating classification performance, particularly for imbalanced datasets [13].
- **ROC-AUC:** Used to measure the ability of the models to distinguish between classes [13].
- **Confusion Matrix :** Used to evaluate model performance by showing correct and incorrect predictions in relation to the actual classes.
- **Learning Curve :** Used to assess how model performance changes as the training dataset size increases, providing insights into model fitting and its generalization capability.

Three datasets (original, oversampled, and undersampled) were evaluated, applying two oversampling techniques, SMOTE and RESAMPLE (similar to Random SMOTE), to determine which would be more effective. After the evaluation, the approach of applying oversampling first, followed by undersampling, proved to be the most effective, consistently outperforming the other datasets and achieving the best values across all metrics.

TABLE I
METRICS FOR THE ORIGINAL DATASET

Algorithm	Accuracy	Precision	Recall	F1-Score	AUC
KNN	0.7197	0.7647	0.8155	0.7893	0.7256
Logistic Regression	0.7767	0.8149	0.8450	0.8297	0.8788
SVM	0.7886	0.8250	0.8524	0.8385	0.8766
Decision Trees	0.8907	0.9518	0.8745	0.9115	0.9172
Random Forest	**0.9026**	0.9389	0.9077	0.9231	**0.9485**

TABLE II
METRICS FOR THE OVERSAMPLED DATASET

Algorithm	Accuracy	Precision	Recall	F1-Score	AUC
KNN	0.7375	0.7890	0.6540	0.7152	0.8309
Logistic Regression	0.7893	0.7909	0.7909	0.7909	0.8722
SVM	0.7969	0.8031	0.7909	0.7969	0.8736
Decision Trees	0.9004	0.8981	0.9049	0.9015	0.9160
Random Forest	**0.9119**	0.9033	0.9240	0.9135	**0.9613**

TABLE III
METRICS FOR THE UNDERSAMPLED DATASET

Algorithm	Accuracy	Precision	Recall	F1-Score	AUC
KNN	0.7229	0.7548	0.6709	0.7104	0.8024
Logistic Regression	0.7857	0.8230	0.7350	0.7765	0.8804
SVM	0.7814	0.8213	0.7265	0.7710	0.8801
Decision Trees	0.8831	0.9286	0.8333	0.8784	0.8896
Random Forest	**0.9286**	0.9427	0.9145	0.9284	**0.9739**

The results presented in the previous tables were generated using the Parkinson dataset as an example to evaluate the performance of the different models. The Random Forest algorithm consistently outperformed other models across all datasets, achieving the highest accuracy, precision, recall, and ROC-AUC scores.

E. Integration with MCDA

Once the diagnoses are obtained from the machine learning models, the MCDA (Multi-Criteria Decision Analysis) approach is used to evaluate and recommend personalized activities. Using multiple criteria, the best activities for the patients are selected, considering their medical conditions and needs. This step includes:

1) Activity Grouping and Criteria Definition: The set of activities, consisting of 43 options, covers areas such as artistic expression, physical exercise, social activities, and recreational activities. These activities are evaluated based on 13 predefined criteria: Cognitive Stimulation, Physical Stimulation, Ease of Execution, Social Interaction, Symptom Control, Activity Duration (hours), Required Resources, Activity Cost (euros), Need for Mobility (0/1), Autonomy Level, Attractiveness and Personal Enjoyment, Safety, and Caregiver Staff (number of caregivers).

2) Creation of the Decision Matrix : The decision matrix was created, with the collaboration of specialists in the field of social education and through detailed research. Some columns of the matrix were initially populated with qualitative values, such as Very High, High, Medium, Low, and Very Low, to evaluate activities based on the predefined criteria. This step was essential, as most Multi-Attribute Decision Making (MADM) techniques begin with the decision matrix, providing the foundation for a structured and objective evaluation of the available alternatives.

TABLE IV
DECISION MATRIX:

Activity	Cog. Stiulat.	Phys. Stimulat.	Ease Exe.	Social Iter.	...
Manual dexterity games	High	Moderate	Low	Very High	...
Spot the differences games	0.7857	High	Moderate	Very High	...
Memory games	Moderate	High	Low	High	...
Bowling	0.8831	Low	High	Very High	...
Numerical comprehension	Low	Very High	Moderate	Moderate	...
...	...	...	...	...	...

Full decision matrix [22]

3) Selection of Weights: The next step was to assign a weight to each criterion, which is essential in multi-criteria problems. These weights reflect the relative importance of each criterion in the decision-making process and influence the calculations of the multi-criteria models to be used. Although this process is typically done by the user, recommended weights were used, determined from prior analysis. To personalize the recommendations based on the patients' health conditions, four sets of weights were defined, based on diagnoses of Parkinson's, Alzheimer's, neither, or both, obtained from the machine learning model. These weights will be used in the initial calculations, adjusting the recommendations to each patient's clinical situation.

TABLE V
WEIGHTS FOR ACTIVITIES BASED ON CRITERIA

Criteria	Alzheimer	Parkinson	Both	None
Cognitive Stimulation	0.22	0.06	0.16	0.10
Physical Stimulation	0.05	0.25	0.16	0.08
Ease of Execution	0.06	0.08	0.08	0.10
Social Interaction	0.15	0.05	0.10	0.10
Symptom Control	0.15	0.15	0.12	0.08
Activity Duration (hr)	0.03	0.03	0.03	0.05
Necessary Resources	0.02	0.02	0.02	0.03
Activity Cost (euros)	0.03	0.04	0.02	0.04
Travel Requirement	0.02	0.02	0.02	0.03
Level of Autonomy	0.02	0.02	0.01	0.06
Attractiveness and Personal Pleasure	0.02	0.02	0.05	0.11
Safety	0.21	0.24	0.22	0.19
Caregiver Personnel (number)	0.02	0.02	0.01	0.03
Total	**1.00**	**1.00**	**1.00**	**1.00**

4) Transformation of Qualitative Data: The qualitative values assigned to activities (e.g., Very High, High, Medium, Low, Very Low) were converted into a 1-5 scale for use in MCDA methods. This numerical transformation allows for objective and quantifiable analysis. This step ensures the data can be processed effectively in decision-making models

5) MCDA Techniques: Two MCDA algorithms were applied to rank activities:

- **TOPSIS (Technique for Order Preference by Similarity to Ideal Solution):** Calculates the relative closeness of each activity to an ideal solution [10]. The steps followed were as follows:
    1) **Step 1: Normalize the decision matrix** Calculate the norm of each column and divide each element of the matrix by the corresponding norm.

2) **Step 2: Construct the weighted matrix** Multiply each element of the normalized matrix by its associated weight.
3) **Step 3: Determine the ideal and anti-ideal solutions**
   - **Positive Ideal Solution (A^+):** The maximum value for criteria to maximize and the minimum value for criteria to minimize.
   - **Negative Ideal Solution (A^-):** The minimum value for criteria to maximize and the maximum value for criteria to minimize.
4) **Step 4: Calculate the distances to the ideal and anti-ideal solutions** Calculate the distance of each alternative to the ideal and anti-ideal solutions.
5) **Step 5: Calculate the relative score of each alternative** Determine the relative score of each alternative based on the calculated distances.
6) **Step 6: Rank the activities** Rank the activities by their scores, from highest to lowest.

- **PROMETHEE II (Preference Ranking Organization Method for Enrichment Evaluations):** It evaluates alternatives based on pairwise comparisons and calculates the net flow to rank the activities [23]. The steps followed were as follows:
  1) **Step 1: Normalize the criteria** The criteria of the alternatives were normalized to make them comparable to each other.
  2) **Step 2: Calculate the difference matrices** For each criterion, a difference matrix was constructed between alternatives, where each element represented the difference between two alternatives.
  3) **Step 3: Calculate the preference matrices** The Type 1 Usual method was used to convert the differences into binary preference values: 1 if there is a preference and 0 if there is no preference.
  4) **Step 4: Calculate the weighted preference index** Each preference matrix was multiplied by its associated weight.
  5) **Step 5: Calculate the outranking flows**
     - **Positive flow:** The sum of the preferences of an alternative over the others, divided by $m-1$ (number of alternatives minus one).
     - **Negative flow:** The sum of the preferences the other alternatives have over the current alternative, divided by $m-1$.
     - **Net flow:** The difference between the positive flow and the negative flow.
  6) **Step 6: Ranking the alternatives** Finally, the alternatives were ranked based on the calculated net flows.

F. Expected Results

The anticipated outcomes of this project were designed around the following key objectives:

- **Model Performance:** It was expected that the Random Forest and XGBoost models would outperform others, given their ensemble learning capabilities. Metrics such as accuracy, precision, recall, F1-Score, and ROC-AUC were used as benchmarks.
- **Feature Importance:** Features such as UPDRS (Unified Parkinson's Disease Rating Scale) and Functional Assessment were expected to have great importance in predicting Parkinson's disease outcomes.
- **Activity Recommendations:** Integration with MCDA techniques aimed to provide actionable and ranked activity recommendations, tailored to individual patient profiles, emphasizing cognitive stimulation, safety, and physical engagement.
- **Impact of Resampling:** Techniques such as oversampling and undersampling were expected to improve the balance of class distributions and enhance model performance on imbalanced datasets.

TABLE VI
ALGORITHMS AND EXPECTED BENEFITS

Algorithm	Expected Benefit
XGBoost	High accuracy, robustness against missing data, and effective feature selection.
Random Forests	Robust ensemble learning to reduce overfitting and enhance prediction accuracy.
Logistic Regression	A simple and interpretable baseline model for binary classification tasks.
SVM	Strong performance in high-dimensional datasets and feature-rich environments.
Decision Trees	Intuitive decision paths and explainability.

G. Obtained Results

1) *Parkinson's Dataset:*

The Random Forest model emerged as the most effective across all datasets—original, oversampled, and undersampled, as we saw before. The performance metrics, importance of features, and graphical results are outlined below.

a) Model Performance: The Random Forest model achieved its highest accuracy and ROC-AUC score using the undersampled dataset with resampling techniques. The model reached an accuracy of **95.67%** and a ROC-AUC score of **97.39%**, confirming its effectiveness in distinguishing between patients with and without Parkinson's disease.

These results also highlight the importance of fine-tuning hyperparameters. The best results for the undersampled dataset were obtained with the following optimized parameters, which were, among other parameters, what obtained the best results using the GridSearchCV technique. *bootstrap: False, class_weight: balanced,max_depth: 20, min_samples_leaf: 2, min_samples_split: 5, n_estimators: 400.*

The feature importance analysis, visualized in Figure 1, underscores the dominance of UPDRS (Unified Parkinson's Disease Rating Scale) as the most critical feature, followed by Functional Assessment, Tremor, Rigidity, and Age. These features collectively captured the key indicators of motor and functional impairments in Parkinson's patients.

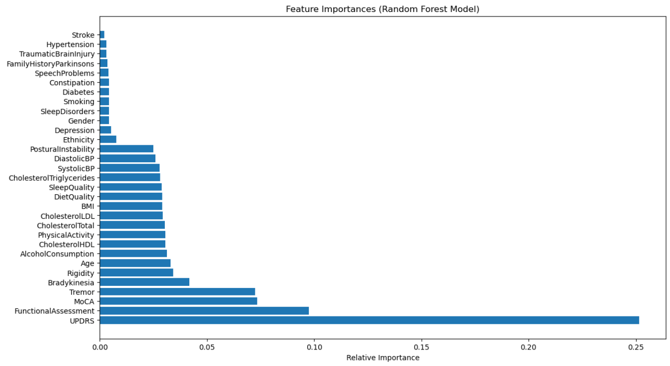

Fig. 1. Feature Importance in the Random Forest Model for Parkinson's Dataset.

b) ROC Curve Analysis: The ROC curves in Figure 2 demonstrate the model's performance across the original, oversampled, and undersampled datasets. The oversampled dataset produced the highest AUC score (**98.5%**), highlighting the efficacy of resampling techniques in improving model robustness.

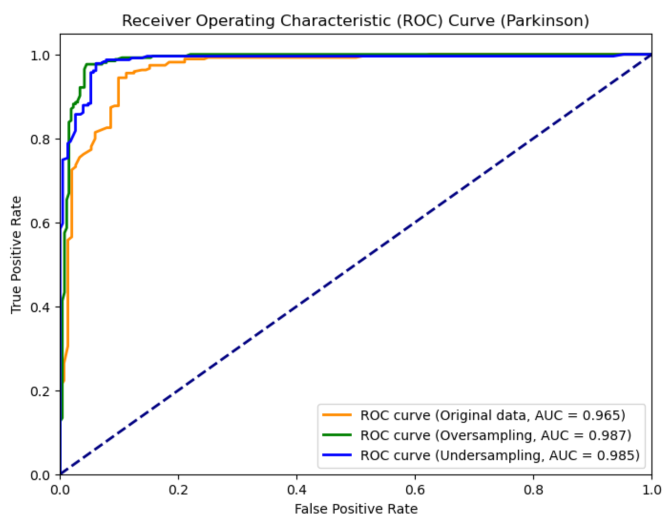

Fig. 2. ROC Curves for the Parkinson's Dataset Using Random Forest.

c) Resampling and Dataset Impact: Comparing the results across datasets reveals the importance of resampling techniques:

- Using the original dataset, Random Forest achieved an accuracy of **92.40%** and ROC-AUC of **96.5%**.
- Undersampling enhanced the model's performance, reaching an accuracy of **95.24%** and ROC-AUC of **98.5%**.
- Oversampling produced the best overall results, with an accuracy of **95.02%** and ROC-AUC of **98.7%**.

d) Impact of Removing the Less Important Features: To evaluate the impact of removing the less important features identified during the feature importance evaluation (*Diagnosis*, *Stroke*, *SpeechProblems*, *TraumaticBrainInjury*, *Hypertension*, *FamilyHistoryParkinsons*, *SleepDisorders*, *Gender*, *Constipation*, *Smoking*, *Diabetes*, *Ethnicity*, *Depression*). These variables were considered less relevant as they demonstrated low contributions to model performance based on their feature importance scores. The results showed that removing these features had minimal impact on the overall accuracy and ROC-AUC scores. Specifically, the accuracy for the original dataset remained almost the same at **93.35%**, for the oversampled dataset at **95.59%**, and for the undersampled dataset at **94.37%**. These findings confirm that removing these variables have a minimal impact in the accuracy across datasets.

e) Concluding Observations: The Random Forest model consistently outperformed other algorithms, demonstrating superior accuracy, interpretability, and robustness. Key observations include:

- **Feature Selection:** UPDRS and Functional Assessment emerged as essential variables, reaffirming the importance of clinical assessments in Parkinson's diagnosis.
- **Resampling Benefits:** The use of undersampling and oversampling improved class balance and performance metrics.
- **Predictive Power:** The integration of optimized hyperparameters maximized the predictive potential of Random Forest.

The integration of these results into the MCDA framework underscores the model's capacity to guide personalized recommendations effectively.

*2) **Alzheimer's Dataset**:*

The Random Forest model was once again the most effective algorithm in all datasets - original, oversampled, and undersampled. The results indicate that this model exceeds in identifying key patterns and features associated with Alzheimer's disease, as evidenced by the high performance metrics and visualized feature importance.

a) Model Performance: The best results were achieved using the undersampled dataset combined with resampling techniques, reaching an accuracy of **95.58%** and a ROC-AUC score of **95.19 %**. These values demonstrate the effectiveness of the Random Forest model in diagnosing Alzheimer's disease.

The optimized hyperparameters for the undersampled dataset were: *bootstrap: True, class_weight: balanced, max_depth: None, min_samples_leaf: 1, min_samples_split: 3, n_estimators: 500*.

The feature importance analysis, illustrated in Figure 3, highlights the dominance of the Functional Assessment as the most influential variable for Alzheimer's diagnosis.

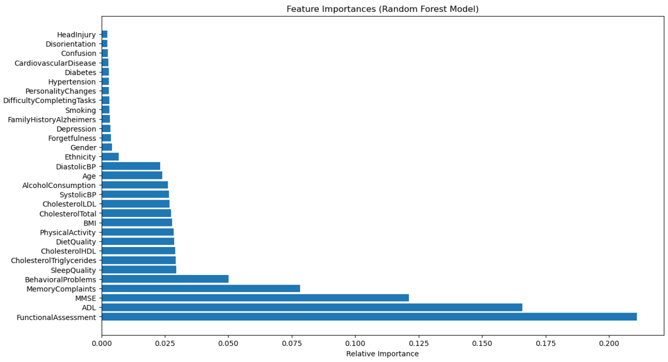

Fig. 3. Feature Importance in the Random Forest Model for Alzheimer's Dataset.

b) ROC Curve Analysis: The ROC curves for the Alzheimer's dataset, shown in Figure 4, demonstrate the model's high sensitivity and specificity across the original, oversampled, and undersampled datasets. The undersampled dataset produced almost the same AUC score, as the original dataset, with a value of (**95.58%**), but the undersampled dataset obtained the best accuracy score emphasizing the importance of resampling techniques to improve model performance.

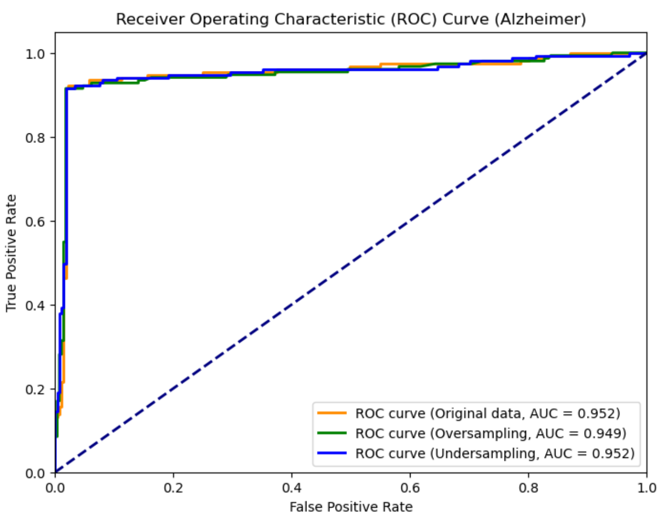

Fig. 4. ROC Curves for the Alzheimer's Dataset Using Random Forest.

c) Resampling and Dataset Impact: Comparative analysis across datasets reveals the following results:

- Using the original dataset, the performance improved slightly compared to the dataset of oversampling, the Random Forest model achieved an accuracy of **95.35%** and ROC-AUC of **95.24%**.
- Oversampling, reaching an accuracy of **94.65%** and ROC-AUC of **94.87%**.
- The undersampling delivered the highest accuracy performance, achieving **95. 58%** and almost the same ROC-AUC as the original dataset **95.19%**.

d) Impact of removing the Less Important Features: To evaluate the impact of removing the less important features identified during the feature importance evaluation (*CardiovascularDisease*, *DifficultyCompletingTasks*, *Hypertension*, *Confusion*, *PersonalityChanges*, *Forgetfulness*, *Depression*, *HeadInjury*, *Gender*, *FamilyHistoryAlzheimers*, *Smoking*, *Diabetes*, *Ethnicity*). These variables were considered less relevant as they demonstrated low contributions to model performance, based on their feature importance scores. The results showed that removing these features had minimal impact on the overall accuracy and ROC-AUC scores. Specifically, the accuracy of the original dataset dropped to **93. 49%**, for the data set sampled at **95. 12%**, and for the data set sampled at **95. 81%**, confirming that if these variables are removed the differences will be few or negative for the learning of the model.

e) Concluding Observations: The results reaffirm the superiority of the Random Forest model for Alzheimer's diagnosis, demonstrating its ability to effectively handle complex feature interactions. Key takeaways include, the most important features, like the Functional Assessment, ADL and MMSE and the predictive accuracy for the implemented model.

IV. Impact Discussion

Combining machine learning algorithms with the multi-criteria decision-making techniques suggested in this work results in a thorough framework for both diagnostic assistance and activity recommendation. Our method goes beyond traditional performance-driven studies by fusing strong classification models, like Random Forest and XGBoost, with decision-making procedures that take into account several aspects of patient care (e.g., safety, cognitive stimulation, required resources). It makes recommendations that take into account the actual context of patient demands by combining clinical, technical, and human variables into a single system.

Improving follow-up and quality of life for individuals with Parkinson's or Alzheimer's disease is one of the main advantages of this integration. The customized character of the suggested activities, which are based on the results of the machine learning models and are improved via the use of analytic hierarchies and outranking techniques, guarantees that every intervention is in line with the patient's cognitive profile, functional evaluation, and general well-being. This preserves motivation and cognitive ability, which improves patient engagement and may slow the progression of the disease. From a wider angle, this kind of human-centered, data-driven approach also makes it easier to optimize healthcare resources. With more information at their disposal, healthcare providers can more effectively assign staff and schedule therapy sessions, which is especially important in environments with a shortage of professionals.

The suggested integration of Random Forest with TOPSIS, and PROMETHEE II has shown great promise throughout this work in assisting decision makers in choosing activities that fit the individual profiles of each patient. Practically speaking, clinical routine decision-making can be greatly streamlined by integrating these algorithms into an intuitive platform or service. Additionally, it has been demonstrated that even class imbalance problems can be reduced by the use of resampling techniques (such as undersampling and oversampling), which enhances the dependability of diagnostic models and the suggestions that follow. With only modest modifications, the resulting framework might be applied to various neurodegenerative illnesses or even chronic ailments.

Recognizing the wider ethical and societal effects of putting such integrated systems into place is similarly crucial. Building confidence between patients and practitioners will depend critically on protecting patient privacy, handling sensitive healthcare data securely, and upholding transparency in model-driven choices. However, by proactively tackling these issues, the field can make significant progress toward a future in which precise and comprehensible machine learning results align with organized criteria for making decisions. In addition to helping physicians find the best therapies, this paradigm opens the door to more fair and inclusive healthcare solutions, which might lessen the strain on overburdened healthcare systems and give patients greater control over their care paths.

V. Future Directions and Conclusion

There is still much room for improvement, even if the suggested combination of machine learning models and multi-criteria decision analysis has demonstrated encouraging outcomes in the diagnosis of neurodegenerative disorders and the recommendation of patient-specific activities. Additional data sources, such sensor-based monitoring and real-time patient feedback, could be incorporated into future research to improve the precision of diagnosis and the applicability of suggested therapies. Furthermore, applying the framework to additional chronic conditions would show how generalizable it is and fortify the basis for all-encompassing patient treatment.

In conclusion, complicated feature interactions were successfully captured and converted into useful activity recommendations by combining Random Forest and other ensemble approaches with decision-making techniques including TOPSIS, and PROMETHEE II. This method provides healthcare providers with useful assistance by striking a balance between predictive insights and a systematic assessment of various criteria, opening the door to more individualized and resource-efficient treatment of patients with Parkinson's and Alzheimer's diseases. Maintaining this model's adaptability, ethics, and usefulness in a variety of clinical settings would need efficient data governance and ongoing methodological upgrades.

References

[1] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[2] L. Breiman, "Random Forests," in *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[3] M. Scholkopf and A. J. Smola, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond," MIT Press, 2002.

[4] R. El Kharoua, "Alzheimer's Dataset," Kaggle, [Online]. Available: https://www.kaggle.com/datasets/rabieelkharoua/alzheimers-disease-dataset. [Accessed: Dec. 2024].

[5] S. Ahmad et al., "Early Detection of Alzheimer's Disease Using Random Forest and SVM Models," *Frontiers in Neurology*, vol. 13, 2023.

[6] R. El Kharoua, "Parkinson's Disease Dataset Analysis," Kaggle, [Online]. Available: https://www.kaggle.com/datasets/rabieelkharoua/parkinsons-disease-dataset-analysis. [Accessed: Dec. 2024].

[7] A. Brown et al., "Decision Tree Analysis in Medical Diagnostics," *Journal of Clinical AI*, vol. 10, no. 2, pp. 123–130, 2021.

[8] R. Belton and T. Stewart, "Multiple Criteria Decision Analysis: An Integrated Approach," Springer, 2002.

[9] W. McKinney, "Data Analysis with Pandas," *O'Reilly Media*, 2017.

[10] C. Hwang and K. Yoon, "Multiple Attribute Decision Making: Methods and Applications," Springer, 1981.

[11] T. Saaty, "The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation," McGraw-Hill, 1980.

[12] J. Roy, "The ELECTRE Method: An Overview," *European Journal of Operational Research*, vol. 10, no. 1, pp. 16–27, 1991.

[13] D. G. Altman et al., "Practical Statistics for Medical Research," Chapman and Hall, 1991.

[14] J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[15] J. Zhang et al., "Multimodal Machine Learning for Alzheimer's Disease Diagnosis Using Clinical, Imaging, and Genetic Data," IEEE Transactions on Biomedical Engineering, vol. 68, no. 2, pp. 527-537, 2021.

[16] S. Arora et al., "Integrating Speech and Motor Data for Early Parkinson's Disease Detection Using Multimodal Learning," Journal of Neural Engineering, vol. 17, no. 3, pp. 1-12, 2020.

[17] Adunlin, G., Diaby, V., & Xiao, H. (2014). Application of multicriteria decision analysis in health care: a systematic review and bibliometric analysis. *Health Expectations*, 18(6), 1894–1905.

[18] N. Zozaya et al., "Multi-Criteria Decision Analysis in Healthcare: Utility and Limitations for Decision-Making," *Journal of Healthcare Policy*, pp. 1–12, 2019.

[19] Kremer, I. E. H., Jongen, P. J., Evers, S. M. A. A., Hoogervorst, E. L. J., Verhagen, W. I. M., Hiligsmann, M. (2021). Patient decision aid based on multi-criteria decision analysis for disease-modifying drugs for multiple sclerosis: prototype development. BMC Medical Informatics and Decision Making.

[20] carlosaft. (2025). Alzheimer Identifier - RandomForestClassifier 95. Kaggle. Recuperado de https://www.kaggle.com/code/carlosaft/alzheimer-identifier-randomforestclassifier-95

[21] carlosaft. (2025). Parkinson Identifier - RandomForestClassifier 93. Kaggle. Recuperado de https://www.kaggle.com/code/carlosaft/parkinson-s-identifier-randomforestclassifier-93

[22] MULTICRITERIA. (s.f.). Hoja de cálculo de análisis de actividades [Hoja de cálculo de Google]. Recuperado de https://docs.google.com/spreadsheets/d/1aeX-XE3EbKk1wrFjbe5km66oAREDI0HeoUZ_LzgAW4E/edit?usp=sharing

[23] Brans, J. P., & Vincke, P. (1985). Note on the PROMETHEE II method. European Journal of Operational Research, 24(2), 228-238. https://doi.org/10.1016/0377-2217(85)90084-8