ASSESSMENT 1: Rule-based AI scientific Research Paper

*ARTIFICIAL INTELLIGENCE [CPU5006-20,SEP-BU,SEM1,2024-2025]  
Marker: Andy Gray*

Contents

[Introduction 2](#_Toc182485888)

**Abstract**

This study investigates the effectiveness of two rule-based AI algorithms in classifying gym members by their engagement level, specifically distinguishing between casual users, regular attendees, and high-intensity trainers. With a growing interest in personalized fitness and member retention strategies, categorizing gym members based on attributes such as age, body composition, workout preferences, and session intensity can provide gyms with critical insights for targeted engagement. This research compares a decision tree classifier and a rule-based expert system, evaluating their performance in classifying gym members based on attributes from the Gym Members Exercise Dataset.

The dataset includes detailed information about each member, such as age, gender, BMI (calculated from height and weight), maximum, average, and resting heart rates, and key metrics like workout type, duration, and calories burned per session. Additional attributes such as workout frequency, daily water intake, body fat percentage, and experience level provide further context about each member’s engagement and fitness level. Using these variables, the study assesses the accuracy, interpretability, and computational efficiency of each algorithm in identifying engagement categories.

Initial results indicate that each algorithm shows distinct strengths: the decision tree classifier identifies nuanced, data-driven patterns across multiple attributes, while the expert system offers greater interpretability and transparency through predefined rules. These findings contribute to understanding the practical applications of rule-based AI in fitness management, supporting gyms in optimizing member engagement strategies and improving personalization.

# Introduction

**Introduction**

1. **Problem Background**: Briefly describe rule-based AI and its applications.
2. **Research Objective**: State the purpose, focusing on comparing two rule-based algorithms in terms of performance, accuracy, and computational efficiency.
3. **Research Question**: Clearly state the research question that addresses the effectiveness of the algorithms in a chosen problem domain.

In recent years, the fitness industry has massivly increased in members. Companies are using data to enhance member engagement, include personiased sdervices, and imorove retnetion rates. For fitness centers and gyms, understadning member’s engagment patterns such as workout frequency, session intensity, and exericse preferences this can be critifal for develolping targeted strategies and cater for the client’s needs. Classifying members based on their behavior and fitness level allows companies to gfive more personalised experisnces and allows gyms to make decisions that will impact every client.

This study explores

This study explores the application of two rule-based AI algorithms—a decision tree classifier and a rule-based expert system—to classify gym members by their engagement level. Specifically, the study focuses on identifying categories such as *casual users*, *regular attendees*, and *high-intensity trainers* based on key attributes within the Gym Members Exercise Dataset. This dataset provides a comprehensive view of each gym member, encompassing demographic data (age, gender), physical metrics (BMI, body fat percentage), health indicators (resting and maximum heart rate), and workout details (workout type, duration, calories burned, and frequency).

Rule-based AI algorithms, known for their transparency and interpretability, offer a promising approach for handling structured data like this. Unlike machine learning models that often act as black boxes, rule-based systems use explicit if-then rules, making them suitable for applications where understanding decision criteria is essential. The decision tree classifier used in this study learns rules from data patterns, while the expert system operates on predefined rules based on fitness domain knowledge. Comparing these two approaches provides valuable insights into their strengths and limitations in the context of fitness data analysis.

The research question addressed in this paper is: *How accurately can rule-based AI algorithms classify gym members into engagement levels based on demographic, physical, and workout-related attributes?* The study’s objectives are to evaluate each algorithm’s performance in terms of accuracy, interpretability, and computational efficiency, and to determine which approach is more practical for real-world applications in gym management. By identifying the optimal algorithm for this classification task, the study aims to contribute to a better understanding of how AI can be effectively utilized in the fitness industry to enhance personalization and support member engagement.

# Literature Review

Rule-based systems are a type of artificial intelligence (AI) that utilize a series of predefined rules to make decisions or solve problems. These systems follow an "if-then" structure, where the system evaluates specific conditions and applies corresponding actions. The simplicity and transparency of rule-based systems make them particularly useful in applications requiring human expertise. For example, rule-based systems are commonly used in fields like **medical diagnosis**, **credit scoring**, and **customer support**, where expert knowledge can be codified into logical rules. These systems are efficient for decision-making, especially in well-structured domains where rules can be explicitly defined.

**Decision Tree Classifiers**

Decision tree classifiers are a prominent example of rule-based algorithms. They function by recursively splitting data based on feature values, forming a tree-like structure. Each node in the tree represents a decision rule that partitions the data, and the leaf nodes represent the classification outcome. Decision trees are inherently rule-based because they break down complex decisions into a series of "if-then" conditions. For example:

* **If** income > $50,000, **then** approve loan.
* **If** credit score < 600, **then** deny loan.

Decision trees are widely used in applications such as **loan approvals**, **customer segmentation**, and **predictive maintenance**. However, they can suffer from overfitting, especially when the tree is too deep, which is often mitigated by pruning or using ensemble methods like **Random Forests**.

**Rule-Based Expert Systems**

Rule-based expert systems are designed to simulate the decision-making abilities of human experts in specialized fields. These systems use a knowledge base of rules and an inference engine to draw conclusions. For example, in medical diagnostics, expert systems apply rules like:

* **If** the patient has fever and cough, **then** diagnose flu.
* **If** the patient has chest pain and shortness of breath, **then** diagnose heart attack.

These systems are crucial in domains like **healthcare**, **engineering**, and **agriculture**, where human expertise is difficult to automate through direct calculations. However, rule-based expert systems are heavily dependent on the quality and completeness of the rules and often require ongoing maintenance by domain experts.

*"How accurately can two different rule-based AI algorithms classify gym members by their engagement level, such as casual users, regular attendees, and high-intensity trainers, based on their exercise and demographic data?"*

**Literature Review**

1. **Overview of Rule-Based Systems**: Discuss rule-based AI, its importance, and examples of applications.
2. **Types of Rule-Based Algorithms**: Describe different approaches to rule-based AI, focusing on the types you’ll use. For example:
   * **Decision Tree Classifiers**: These are rule-based because they create decision paths based on logical "if-then" rules.
   * **Rule-Based Expert Systems**: Such systems use predefined rules and are used in domains requiring heuristic knowledge, such as medical diagnosis.
3. **Prior Research**: Summarize relevant prior studies or applications using rule-based algorithms, highlighting their outcomes and limitations.

**Methodology**

1. **Algorithms Chosen**: Provide a rationale for the two specific rule-based algorithms you’re using. For example:
   * **Algorithm 1**: Decision Tree (e.g., ID3 or C4.5), which builds a tree based on information gain or Gini index and generates a set of rules.
   * **Algorithm 2**: Expert System using a rule-based inference engine (e.g., forward or backward chaining) to classify instances.
2. **Dataset Selection and Preprocessing**:
   * Describe the dataset (source, features, target variable).
   * Explain preprocessing steps, such as data cleaning, encoding categorical variables, or feature scaling.
3. **Evaluation Metrics**:
   * Describe the metrics to be used for comparison: accuracy, precision, recall, F1-score, and computational time.
   * Optionally, discuss the use of confusion matrices to gain insights into false positives/negatives.

**Results**

1. **Performance of Algorithm 1 (e.g., Decision Tree)**:
   * Present accuracy, precision, recall, F1-score, and any other chosen metric.
   * Provide insights into the computational efficiency and any unique strengths or weaknesses observed.
2. **Performance of Algorithm 2 (e.g., Expert System)**:
   * Present the same performance metrics.
   * Compare computational costs and describe scenarios where this algorithm performs best.
3. **Comparison and Analysis**:
   * Present a table or graph that compares both algorithms across all metrics.
   * Interpret differences in performance and explain why certain features may influence the outcomes in each model.

**Discussion**

1. **Evaluation of Results**: Discuss why one algorithm may have performed better or worse in specific areas, using evidence from the results.
2. **Strengths and Weaknesses**: Address the strengths and limitations of each algorithm based on the experiment.
3. **Application Potential**: Discuss the potential applications for each rule-based AI system, considering the dataset characteristics and algorithmic trade-offs.

**Conclusion**

1. **Summary of Findings**: Recap the key results and insights from the research question.
2. **Implications**: Address the broader implications, such as when each rule-based system might be preferable based on the characteristics of a given problem.
3. **Future Work**: Suggest further research directions, such as testing additional rule-based methods or hybrid models to improve performance.

**References**