ASSESSMENT 1: Rule-based AI scientific Research Paper

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

[*https://github.com/JamesScottBSU/ASSESSMENT-1--Rule-basedAIAlgorithm*](https://github.com/JamesScottBSU/ASSESSMENT-1--Rule-basedAIAlgorithm)

*Abstract*

*The fitness industry's rapid growth has raised demand for personalised gym experiences that meet each person's demands and preferences. This study explores the use of forward chaining and backward chaining, two rule-based AI algorithms, to examine how members of gyms work out and maximise equipment use. The study compares the effectiveness of the two algorithms and looks at fitness and demographic data to find trends in workout preferences using the Gym Members Exercise Dataset from Kaggle.*

Contents

[Introduction 4](#_Toc183191440)

[Literature Review 5](#_Toc183191441)

[Forward Chaining 5](#_Toc183191442)

[Backward Chaining 5](#_Toc183191443)

[Methodology 7](#_Toc183191444)

[Algorithms chosen 7](#_Toc183191445)

[Dataset Selection 7](#_Toc183191446)

[Features: 7](#_Toc183191447)

[Research goals 8](#_Toc183191448)

[Results 9](#_Toc183191449)

[Information gathered 9](#_Toc183191450)

[Workout preferences trends 9](#_Toc183191451)

[Ages of members and the most popular workout type for that age group 9](#_Toc183191452)

[BMI categories of members and that category most popular workout type 9](#_Toc183191453)

[Most popular workouts 10](#_Toc183191454)

[Outcome 10](#_Toc183191455)

[Conclusion 12](#_Toc183191456)

[References 13](#_Toc183191457)

# Introduction

The fitness industry’s rapid growth has led to an increased demand for personalised experiences that meet to the unique needs of gym members. By understanding individual workout preferences, such as session intensity, frequency, and exercise choices, gyms can enhance engagement, optimise resource usage, and improve retention rates. These patterns and trends can be found with the use of rule-based AI approaches like forward and backward chaining, which allow for more personal recommendations for members. These AI algorithms help identify which factors influence workout preferences, allowing gyms to align equipment and resources more effectively with member behaviours and goals.

The research question addressed in this paper is:  
"Applying forward chaining and comparing it with backward chaining to identify trends and patterns, such as gym members' workout preferences, and assist gyms in optimizing equipment usage."

This study compares the efficiency of forward and backward chaining algorithms in spotting important trends while researching how these algorithms might be used to analyse member behaviours and workout preferences. The research aims to find patterns that can enhance gym experiences and raise member happiness by using these techniques to a dataset that contains demographic data, fitness levels, and workout routines. In ultimately, this method will assist fitness centres in making data-driven choices regarding the positioning of equipment, the distribution of resources, and member engagement tactics, resulting in a more personalised and effective workout space.

The study’s objectives are to evaluate each algorithm’s performance in terms of accuracy, clarity, and efficiency, and to determine which approach is more practical for real-world applications in gym management. The project aims to expand knowledge of how AI may be successfully applied in the fitness sector to improve personalisation and encourage member engagement by determining the most effective algorithm for this objective.

# Literature Review

Rule-based systems are a type of AI that uses a series of predefined to make decisions and solve problems. These systems use an “if-then” structure, which mimics human decision making it a reliable and predictable system for various applications. This system stands out because it focuses and relies on the rules created by humans. Because of this it means every operation and decision it makes can is specify for the specific set of guidelines. Because of this it is commonly used in healthcare – for heart disease diagnosis, customer support – chatbots/smart assistants, and banking – fraud detection and risk management. These systems are efficient for decision-making, especially in well-structured domains where rules can be explicitly defined.

## Forward Chaining

Forward chaining is a data-driven process that uses forward deduction or forward reasoning. Forward chaining starts with known facts and applies rules to generate conclusions/endpoint/ goal. For instance, in a gym setting, a Forward Chaining system could start with a member's demographic data and apply rules to suggest suitable equipment based on those attributes.

Advantages includes:

* Suitable to draw multiple conclusions simultaneously
* Higher flexibility than backward chaining
* Reliable for conclusion

Disadvantages include:

* • Time-consuming due to data synchronization
* • A dataset of some sort is needed.

## Backward Chaining

Backward Chaining, on the other hand, is a goal-driven approach. To determine which facts or conditions must be true for a goal or endpoint to be accomplished, it begins with the aim and works backwards. This approach is usually used in reasoning or diagnostic systems, where the system must go back through a series of rules to confirm whether the objective can be achieved. Backward Chaining, for instance, could be used in a gym setting to discover exactly which member attributes (e.g., BMI, age, and preferences) needed to reach a fitness objective, like weight loss.

Advantages includes:

* Swifter than forward chaining
* Efficiently drives correct solutions

Disadvantages includes:

* Provides single answer
* Less flexibility
* Suitable only if the endpoint is known
* Difficult to execute

# Methodology

## Algorithms chosen

This study evaluates two rule-based AI algorithms: Forward Chaining and Backward Chaining. Both algorithms are widely used for decision-making in structured domains, such as fitness equipment recommendations, where decisions are made based on predefined rules derived from specific conditions or goals.

## Dataset Selection

The dataset that I have used to answer this question is **Gym Members Exercise Dataset** from Kaggle. Some uses for this dataset are for analysing fitness patterns and performance across diverse gym experience levels. This dataset includes a detailed overview of gym members routines, fitness metrics, and physical attributes. Contains 973 samples of age, gender, weight, height, max BPM, average BPM, resting BPM, session duration, calories burned, and workout type.  
I chose this dataset due to its quality data, and this dataset included all the basic information for this question. Not only this, but the number of samples will make the answer more accurate. All the features allow me to answer multiple different questions, such as what age group visits the gym the most, or how much time experienced people spend in the gym.

### Features:

* Age: Age of the gym member. (Years)
* Gender: Gender of the gym member (Male or Female).
* Weight (kg): Member’s weight in kilograms.
* Height (m): Member’s height in meters.
* Max\_BPM: Maximum heart rate (beats per minute) during workout sessions.
* Avg\_BPM: Average heart rate during workout sessions.
* Resting\_BPM: Heart rate at rest before workout.
* Session\_Duration (hours): Duration of each workout session in hours.
* Calories\_Burned: Total calories burned during each session.
* Workout\_Type: Type of workout performed (e.g., Cardio, Strength, Yoga, HIIT).
* Fat\_Percentage: Body fat percentage of the member.
* Water\_Intake (liters): Daily water intake during workouts.
* Workout\_Frequency (days/week): Number of workout sessions per week.
* Experience\_Level: Level of experience, from beginner (1) to expert (3).
* BMI: Body Mass Index, calculated from height and weight.

### Research goals

Identify patterns in workout preference – trends in gym members workout behaviours such as preferred exercise and equipment usage, by using forward and backward chaining.

Compare forward chaining and backward chaining techniques: Assess and compare the performance of both algorithms in recognizing trends and generating recommendations for equipment usage.

Enhance gym equipment optimization: Use the findings from the algorithms to identify underused equipment and suggest improvements for gym resource allocation, machine placement, and equipment upgrades.

Provide actionable insights for gym management: Generate clear, practical recommendations for gym managers based on the results of the algorithm analysis, which will help in decision-making related to gym operations.

# Results

## Information gathered

Gender distribution: These statistics will determine the percentages of male and female members, as well as the total number of members. This information can show women prefer different types of exercise to men and if the gym has more male members.

Workout preferences: finds the most common workouts for each gender, BMI category, and age group. This can show the last used workout type and show the most popular type. Therefore, the gym could add more equipment to meet the members’ needs and goals.

Age group preferences: this puts members into age ranges and finds the most common workout type for each age range. This could show what younger members prefer compared to older members. Also, can help advertisements get more engagement from a younger audience.

BMI classification: categorises members as underweight, obese, or healthy based on thresholds from NHS standards. Then it shows what their most common workout type is. This can help create plans for members who are losing weight or building muscle.

## Workout preferences trends

Both algorithms successfully identified distinct patterns related to workout preferences based on member demographics and fitness characteristics. However, the backward chaining did not calculate the percentage.

Ages of members and the most popular workout type for that age group  
Both forward chaining and backward chaining showed that strength was the most popular, with forward chaining showing a percentage of 26% of members. The least popular was HIIT, with 22% of members.   
As for the age range, they both identified calculated that ages 19 to 35 where the most common, and Ages 18 and below were the least common:

(0-18) 27 members or 2.77% of members are within this range. This is the least common age range.  
(19-35) 375 members or 38.54% of members are within this range. This is the most common age range.  
(36-50) 361 members or 37.1 of members are within this range.  
(51+) 210 members or 21.58% of members are within this range.

### BMI categories of members and that category most popular workout type

Both algorithms identified the number of members in each BMI category. They both also got the most common workout done by members in that category. Forward chaining gave percentages of the distribution between members and the category. However, backward chaining did not show percentages:

Normal: 366 members (37.62%) – most common workout was Cardio

Overweight: 238 members (24.46%) – most common workout was Strength

Obese: 201 members (20.66%) – most common workout was Cardio

Underweight: 168 members (17.27%) - most common workout was strength

### Most popular workouts

Both Algorithms found the total amount of members that use that workout type. Again, forward chaining included percentages. But both have the correct data, which as showing that strength is the most popular out of the 4 workout types.

Strength: 258 members

Cardio: 255 members

Yoga: 239 members

HIIT: 221 members

## Outcome

From the information we have gathered we can notice a few things about this gym, and their members. Firstly, what we can say is that there are more members who use strength workout equipment, therefore no equipment needs to be changed for now. However, the gym could add a few more machines if they want.  
Another thing we can notice is that the members are almost equally men and women, which is a positive thing for the gym. The gym can use this for advertising, which can boost engagement.

Another is the age range, which they can improve on. Not many members are 0-18 due to most gyms not allowing under 18s. Most of the members are on the younger side of adulthood, and we can see that the gym should advertise for older adults too.

One of the most important data is the BMI categorisation for members. This puts them into categories and sees what progress they are trying to make. For example, obese members are mostly doing cardio, which tells us they are losing weight. Another example is underweight members doing more strength workouts. This tells us that they are building muscle / gaining weight. This helps the gym notice the numbers and can improve the equipment to tailor for members.

An example of where this can be seen is the US brand planet fitness. Was mostly used by older members, for weight loss. Now most members are interested in strength workouts, and because of this they are adding new machines. They still upgrade cardio as that is the second most popular workout type.

# Conclusion

This study shows how rule-based AI algorithms may be used to monitor gym users behaviour and maximise resource allocation by comparing forward and backward chaining. It was clear that both algorithms, although they had different advantages and disadvantages, were able to recognise patterns in workout preferences based on fitness and demographic data.   
  
Forward chaining proved its ability to obtain multiple findings at once, providing specific trends and percentages that could improve gym management decision-making. Its data-driven approach allows for flexibility in determining comprehensive conclusions, making it suitable for situations where extensive trend analysis is required. On the other hand, backward chaining's goal-oriented design made it effective at finding specific answers, but it lacked the depth and scope that forward chaining offered.

**Key findings include:**

Workout Preferences and Demographics: Across all age groups and BMI categories, strength training was the most popular, followed closely by cardio. Although individuals between the ages of 19 and 35 make up the majority of the gym's membership, older and younger audiences could be drawn in.  
BMI-Based Trends: Members' workout preferences aligned with their fitness goals, such as obese members focusing on cardio for weight loss and underweight individuals favouring strength training to build muscle mass.  
Equipment Optimisation: According to the statistics, the equipment that is now available fits in well with member preferences; however, small changes could improve participation and take into account new trends.

# References

vala khorasani (2024). *Gym Members Exercise Dataset*. [online] Kaggle.com. Available at: https://www.kaggle.com/datasets/valakhorasani/gym-members-exercise-dataset?resource=download.

*Simplilearn.com. (2023). The Ultimate Guide to Forward and Backward Chaining in AI. [online] Available at: https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/forward-and-backward-chaining-in-ai.*

*‌Deepgram. (n.d.). Rule-Based AI. [online] Available at: https://deepgram.com/ai-glossary/rule-based-ai.*

*‌NHS (2023). Obesity. [online] NHS. Available at: https://www.nhs.uk/conditions/obesity/.*

*‌ GeeksforGeeks. (2020). Get first n records of a Pandas DataFrame. [online] Available at: https://www.geeksforgeeks.org/get-first-n-records-of-a-pandas-dataframe/.*

*‌* *Stack Overflow. (n.d.). Backward and forward chaining algorithm for (expert system) in Python. [online] Available at: https://stackoverflow.com/questions/43028336/backward-and-forward-chaining-algorithm-for-expert-system-in-python.*

*‌* *W3 Schools (2019). Python Functions. [online] W3schools.com. Available at: https://www.w3schools.com/python/python\_functions.asp.*

*‌* *GeeksforGeeks. (2024). Forward Chaining and Backward Chaining inference in Rule-Based Systems. [online] Available at: https://www.geeksforgeeks.org/forward-chaining-and-backward-chaining-inference-in-rule-based-systems/.*

*‌* *Radke, P. (2023). Forward vs Backward Chaining in Artificial Intelligence | Built In. [online] builtin.com. Available at: https://builtin.com/artificial-intelligence/forward-chaining-vs-backward-chaining.*

*‌* *www.naukri.com (2024). Code 360 by Coding Ninjas. [online] Naukri.com. Available at: https://www.naukri.com/code360/library/forward-and-backward-chaining-in-ai.*

*Python Software Foundation (2020). csv — CSV File Reading and Writing — Python 3.8.1 documentation. [online] Python.org. Available at: https://docs.python.org/3/library/csv.html.*

*‌* *pandas.pydata.org. (n.d.). pandas.Series.value\_counts — pandas 1.3.4 documentation. [online] Available at: https://pandas.pydata.org/docs/reference/api/pandas.Series.value\_counts.html.*

*‌* *Viacheslav Nefedov (2013). GroupBy pandas DataFrame and select most common value. [online] Stack Overflow. Available at: https://stackoverflow.com/questions/15222754/groupby-pandas-dataframe-and-select-most-common-value.*

*‌* *Bhalla, D. (n.d.). Python : 10 Ways to Filter Pandas DataFrame. [online] ListenData. Available at: https://www.listendata.com/2019/07/how-to-filter-pandas-dataframe.html.*

*‌* *GeeksforGeeks. (2019). Boolean Indexing in Pandas. [online] Available at: https://www.geeksforgeeks.org/boolean-indexing-in-pandas/.*

*GeeksforGeeks. (2018). Python | Output Formatting. [online] Available at: https://www.geeksforgeeks.org/python-output-formatting/.*

*GeeksforGeeks. (2018). Different ways to iterate over rows in Pandas Dataframe. [online] Available at: https://www.geeksforgeeks.org/different-ways-to-iterate-over-rows-in-pandas-dataframe/.*

*ChatGPT (n.d.). ChatGPT. [online] ChatGPT. Available at: https://chatgpt.com/.*

*ChatGPT*

*Question: how would i find the age range of a dataset*

*A screenshot of a computer

Description automatically generated*

*A screenshot of a computer program

Description automatically generated*

*A screenshot of a black screen

Description automatically generated*