https://github.com/Willthetitan/ARTIFICIAL-INTELLIGENCE--CPU5006-20-SEP-BU-SEM1-2024-2025--Willthetitan  
  
ARTIFICAL INTELLIGENCE

# Introduction

## AI and its influences

Artificial Intelligence is seeming to be increasingly involved in our everyday lives and rightly so. Throughout humanity we have made tools to optimize our lives, and this Tool is no different. However, AI in of itself is not just a tool but a collection of tools with different purposes. Within this essay we will be focusing on Rule Based AI. What is Rule Based AI? “A system designed to achieve artificial intelligence (AI) via a model solely based on predetermined rules is known as a rule-based AI system.”(Mario Grunitz 2021 – Ref 1.1) . This method of Operation is very successful and hence been used in many formats but its predominant utility is data manipulation Due to its simplicity of implementation and Consistency. This is due to its “if-then” coding statements allowing the user to have full control in the models pre-defined outcomes insofar requiring no form of training.

## AI and a Goal in Mind

With These Utilizations in mind and its predominant strength of data manipulation, I believe there is potential for this AI model to have success within the Medical Field more specifically in the use of determining potential patients with heart disease via Common attributions like Age , Chest Pain , Blood Pressure , Cholesterol , ECG and more.

This is due to these attributes having identifiers that can assist with diagnosis via diagnostic data. And as discussed before Rule Based Algorithms Preferred use is Data Manipulation. This Scientific paper will explore how this approach, rooted in clear “if-then” logic can be of enhancement to modern medical diagnostics, potentially offering a practical solution in an already increasingly data driven healthcare landscape.

# Literature Review

Although Rule Based AI is as simple as “if-then” we need to better explore its inner workings as to how it handles these “if then” Operations. This section explores the Essential Components that underpin Rule Based systems.

**1. Rules: The Core Mechanism**

At the core of any rule-based system are the rules, which define the system’s outcome. Each rule follows the general format of:

* **IF** [condition(s)] **THEN** [action].

For example:

* **IF** an individual is tired **AND** it is nighttime, **THEN** sleep.

These rules incorporate logical relationships between conditions and corresponding actions, forming the basis for the systems operation.

**2. Knowledge Base: Repository of Rules and Facts**

To evaluate conditions and execute actions, the system utilizes a **Knowledge Base**, which serves as a repository for two critical components:

* **Rules**
* **Facts**

In rule-based artificial intelligence, Data is denoted as FACTS. For instance, to conclude that an individual should sleep, the system must first verify that the FACTS "tired" and "nighttime" are present in the Knowledge Base and are True.

**3. Inference Engine: Applying Rules to Derive Conclusions**

The process of applying rules to available facts is needed bringing us to the **Inference Engine**, a key component responsible for Conclusions. The Inference Engine functions by iteratively evaluating the conditions specified via the RULES against the FACTS stored in the Knowledge Base.

The Inference Engine can utilize different reasoning strategies to reach a conclusion, such as forward chaining (starting from known facts to derive conclusions) or backward chaining (working backward from a desired conclusion to find supporting facts).

**4. Working Memory: Temporary Storage for Dynamic Processing**

The use of Iteration requires the Inference Engine to use an extra component known as Working Memory to temporarily store current facts.

**5. User Interface (Optional)**

Many rule-based systems incorporate a User Interface (UI) to enable user interaction. This interface allows users to:

* Input new facts or modify existing ones within the Knowledge Base.
* Receive conclusions generated by the system’s reasoning process.

**Benefits of Rule-Based Systems in Healthcare**

Now that we have a firm understanding of rule-based systems, it is essential to explore their application within the healthcare industry. Currently, Rule Based systems are recognized as effective tools specifically in their integration into electronic health records and clinical decision support systems. These systems streamline clinical workflows. However, significant gaps remain in understanding how AI, could be incorporated within the healthcare industry As highlighted by ([Jiang et al., 2017](https://www.sciencedirect.com/science/article/pii/S0952197623000787#b65), [Tekkeşin et al., 2019](https://www.sciencedirect.com/science/article/pii/S0952197623000787" \l "b148)) “ It is all too common for humans to be overlooked when discussing AI’s role in real-world applications”

Within the same scientific paper Kumar emphasizes the potential of AI-driven systems to analyze vast amounts of medical data, transforming them into diagnostic information. This Reflects the diagnostic reasoning by human clinicians, suggesting that rule-based models could effectively mimic human decision making. This resemblance indicates the possibility of constructing rule sets that try to emulate human logic in clinical diagnosis, which is perfectly suited towards Rule Based AI systems.

However before we can begin we must determine what Format of Rule Based AI would be most suitable towards our goal of determining potential patients that could be susceptible to heart Disease

# Chaining Methods in Rule-Based AI

Forward Chaining  
Forward chaining is a data-driven inference method that begins with known facts and applies rules iteratively until a specific goal is achieved. For example:

* + Fact: A patient has a fever.
  + Goal: A fever indicates the possibility of the flu.

Backward Chaining  
backward chaining is the opposite being a goal-driven approach that starts with a goal and works backward to gather supporting data. For instance:

* + Goal: A fever may indicate the presence of a cold.
  + Fact: The patient presents with a fever.

# Viability of Forward and Backward Chaining

Both forward and backward chaining can be viable in clinical diagnosis, depending on the data available and the diagnostic use. Forward chaining is well-suited for scenarios where observable symptoms need to be linked to potential diagnoses, while backward chaining is useful when a hypothesis-driven approach is needed to confirm or refute specific conditions.

# Hybrid Chaining

A third approach hybrid chaining is another chaining method combining the strengths of forward and backward chaining. Hybrid systems are particularly valuable in complex diagnostic scenarios where both data-driven reasoning and goal-oriented reasoning are required. These systems can dynamically shift between reasoning strategies.

# Machine Learning

Although machine learning is not rule-based, it warrants consideration in the medical domain due to its powerful pattern recognition capabilities. Machine learning models can analyze large datasets to identify patterns and make predictions about future data in this case future patient problems, such as predicting the likelihood of heart disease in patients who have not yet presented with symptoms.

# Methodology

Within this Section we will be implementing and applying our current Understanding in context to Rule Based AI systems and our goal. With these in mind we must first gather a dataset that is best suited to our needs. With potential Heart Disease patients the main Factors to evaluate a potential risk our Age , Blood Pressure , Cholesterol , chest pain type , blood sugar and Resting ECG. A dataset I found on the web via Kaggle has these Factors, and patient data consisting of over

1026 patients and whether these patients got heart Disease or not, We can set this outcome as our target so that we can assess the accuracy of our Rule set This allows us to implement a Confusion Matrix

## **Confusion Matrix**

## A confusion matrix summarizes model performance on a set of test data by comparing predictions to actual outcomes:

|  |  |  |
| --- | --- | --- |
|  | Prediction True | Prediction False |
| Actual True |  |  |
| Actual False |  |  |

## This matrix enables us to evaluate the effectiveness of our rule set by comparing predicted outcomes with actual diagnoses.

## **Rule Set Approach**

## To determine the most effective Chaining method, both forward chaining and backward chaining. In medical diagnostics symptoms lead to conclusions meaning forward chaining. Is suitable for Diagnostics However, in our case, since patients do not display symptoms and require goal driven evaluation, backward chaining is more suitable. This approach involves refuting wether or not a patient could be at risk of heart disease.

## **Rule Set Design and Implementation**

## We can approach this form of chaining in two different ways, one that identifies factors associated with increased heart disease risk (positive rule set) and another that identifies factors associated with a lower likelihood of heart disease (negative rule set).

## **Positive Return Rule Set**

## The positive rule set returns a result of **1** when conditions indicative of increased heart disease risk are met:

## **If** chest pain == Angine , colored blood vessels == Two or Three, oldpeak >= 2, and slope == downsloping: **Then** return 1

## **If** chest pain == Asymptomatic, blood sugar > 120, cholesterol > 240, and resting ECG == Abnormal: **Then** return 1

## **If** chest pain == Non anginal, colored blood vessels == One, slope == flat, and cholesterol >240: **Then** return 1

* If Age < 45, thalassemia == Reverse able, and rest ECG == Wave abnormality

## **Anything else** **Then** return 0

## This rule set aims to identify anomalies in key attributes that are positively linked to heart disease.

## **Negative Return Rule Set**

## The negative rule set returns a result of **0** when conditions Lead to a healthy patient:

## **If** chest pain, age == asymptomatic , colored blood vessels == 0 , and cholesterol <200: **Then** return 0

## **If** age > 45, cholesterol < 200, and chest pain == asymptomatic: **Then** return 0

## **If** chest pain == Non anginal , oldpeak < 0.2, and resting ECG == Normal: **Then** return 0

## **Anything else** **Then** return 1

Unlike the positive rule set, this approach focuses on identifying conditions that are less likely to be associated with heart disease, therefore ruling out potential heart disease.

# RESULTS

## Now that we have implemented our rule sets, we can now Critically Evaluate the results we received via the Rule sets we created, and compare the contrast.

# Accuracy

The accuracy of our models were based via the total ratio of correct instances to the total instances.

For our inclusive(Positive Return) Rule set we received a 39.32% accuracy rating whereas our Exclusive(Negative Return) Rule set gave us a 46.24% accuracy rating this discrepancy may stem from the broader scope in the inclusive rule set, which considers attributes that, while indicative of potential heart disease, could also be associated with other unrelated conditions. In contrast, the exclusive rule set avoids this pitfall by narrowly focusing on ruling out factors unrelated to heart disease, therefore enhancing its accuracy.

# F1-Score

Another Great Mathematical Model to assess our Rule sets is by evaluating the overall performance of a model. For our Inclusive Rule set we received a 8.25% Performance Rating Whereas for our Exclusive Rule set we received a 63.24% Performance Rating. This allows us to Further recognize the Exclusive Rule set, this time in its Precision and Recall Techniques to provide a better-balanced performing model further reinforcing its status as a more reliable and well rounded potential Diagnostic Model

# DISCCUSION

## The results highlight a large disparity between the two models although they are both backward chaining methods, this shows the significance of a effective approach towards the rule set

# Strengths of Positive Rule set

# Although the positive rule set proved less effective for diagnosing heart disease, it could excel when applied to datasets involving a broader range of potential conditions. For example, instead of focusing exclusively on heart disease, this rule set might perform better in identifying general symptoms connected to heart-related issues, making it valuable for more broad diagnostic tasks.

# Strengths of Negative Rule Set

# For the negative ruleset the results demonstrate that an exclusionary approach offers superior accuracy when targeting a specific diagnosis. The negative rule set excels because it focuses on eliminating conditions unrelated to the target, minimizing the risk of miss diagnosing similar symptoms as heart disease. This precision makes it particularly effective for narrowly defined diagnostic objectives.

# Limitations of backward Chaining

While backward chaining was appropriate for our goal of ruling out heart disease, it has limitations In a broader diagnostic goal , forward chaining is often more suitable due to its hypothesis-driven approach, which allows multiple potential diagnoses to be explored. For example, if a patient presents with a fever, forward chaining could generate multiple possible diagnoses, such as:

* + Viral infection
  + Bacterial infection
  + Covid 19
  + Sinusitis
  + Heat exhaustion
  + Etc…

### whereas backward chaining focuses on a single hypothesis, returning only one possible diagnosis. However, for the goal of ruling out heart disease backward chaining was the optimal choice.

# Implications of future work

The findings of this study suggest that the selection of a chaining method should be carefully tailored to the clinical context, depending on whether the goal is future diagnostics or immediate diagnosis. Future research should further investigate the potential of integrating forward and backward chaining into a hybrid system, inheriting the strengths of both approaches. Additionally, exploring the application of machine learning algorithms could offer new insights and enhance diagnostic accuracy, providing a more robust and adaptable framework for medical decision-making.

# Conclusion

## This study compared the effectiveness of backward chaining rule sets in diagnosing potential heart disease patients. By implementing and evaluating both inclusive and exclusive approaches, the results revealed significant disparities in accuracy and performance. The inclusive rule set achieved an accuracy of 39.32% and a performance rating of 8.25%, whereas the exclusive rule set demonstrated superior results with an accuracy of 46.24% and a performance rating of 63.24%.

## These findings strongly suggest that, in clinical diagnoses focused on solely potential future symptoms, an exclusionary approach eliminating other possible conditions is more effective than attempting to diagnose based solely on current symptoms. Rule-based AI systems are particularly well-suited for such tasks due to their clarity and well-defined outcomes.

## The results underscore the importance of aligning rule sets with the specific requirements of the diagnostic context. However, future research should explore the potential benefits of hybrid systems that integrate both forward and backward chaining, as well as the application of machine learning algorithms, to enhance diagnostic accuracy and flexibility.

## References

### 1.1

Gruntiz, M. (2021) ‘Rule-based AI vs machine learning: what’s the difference?’, https://wearebrain.com/blog/rule-based-ai-vs-machine-learning-whats-the-difference/, 13 September. Available at: https://wearebrain.com/blog/rule-based-ai-vs-machine-learning-whats-the-difference/ (Accessed: 27 November 2024).

1.2

GeeksforGeeks, P. (2024) Rule-based system in Ai, GeeksforGeeks. Available at: https://www.geeksforgeeks.org/rule-based-system-in-ai/ (Accessed: 27 November 2024).

1.3

Kumar, P. (2023) Artificial Intelligence in Healthcare: Review, Ethics, Trust Challenges & Future Research Directions, https://www.sciencedirect.com/science/article/pii/S0952197623000787. Available at: https://www.sciencedirect.com/science/article/pii/S0952197623000787 (Accessed: 27 November 2024).

1.4

GeeksforGeeks                                                                                                                                                                                                                                                  and GeeksforGeeks (2024) *Introduction to machine learning: What is and its applications*, *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/introduction-machine-learning/ (Accessed: 28 November 2024).

1.5

GeeksforGeeks (2024) *Rule-based system in Ai*, *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/rule-based-system-in-ai/ (Accessed: 28 November 2024).

1.6

Alowais, S.A. *et al.* (2023) *Revolutionizing Healthcare: The role of Artificial Intelligence in Clinical Practice*, *BMC medical education*. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC10517477/ (Accessed: 28 November 2024).