

Fuzzy Expert System to Detect Phishing in Websites

Dániel MÁCSAI
Ismael RUIZ GARCIA
Mauro VÁZQUEZ CHAS

Master in Artificial Intelligence



UNIVERSITAT
ROVIRA i VIRGILI

Planning and Approximate Reasoning
Delivery 3

15th December 2024

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 2 |
| 2 | Definition of the Linguistic Variables | 2 |
| 2.1 | Chosen Features | 2 |
| 2.1.1 | URL-Based Features | 2 |
| 2.1.2 | Content Features | 3 |
| 2.1.3 | External Features | 4 |
| 3 | Output Variable | 5 |
| 4 | Implementation | 5 |
| 5 | Rules | 6 |
| 6 | Implementation | 7 |
| 7 | Testing | 7 |
| 7.1 | Test Cases | 7 |
| 7.1.1 | Case 1: Phishing Website | 7 |
| 7.1.2 | Case 2: Phishing Website | 8 |
| 7.1.3 | Case 3: Legitimate Website | 9 |
| 7.1.4 | Case 4: Legitimate Website | 10 |
| 8 | Complex Fuzzy Expert System | 11 |
| 8.1 | Layer 1: Feature-Based Risk Assessments | 12 |
| 8.2 | Layer 2: Overall Phishing Risk Assessment | 12 |
| 8.3 | Weight Distribution | 12 |

1 Introduction

This report presents a fuzzy expert system designed to detect phishing websites. Phishing attacks are a significant cybersecurity threat, often leading to identity theft and financial losses. Traditional methods for detecting phishing websites have limitations in accurately identifying sophisticated attacks. Fuzzy logic, with its ability to handle uncertainty and imprecision, offers a promising approach to enhance phishing detection.

Our fuzzy expert system leverages a set of carefully selected features to assess the risk of a website being malicious. These features encompass various aspects, including URL characteristics, content analysis, and external factors. By applying fuzzy logic rules, the system can effectively classify websites into different risk categories, ranging from safe to highly suspicious.

This report details the design and implementation of the fuzzy expert system, including the selection of linguistic variables, the definition of membership functions, and the formulation of fuzzy rules. Additionally, we present the results of testing the system on a range of websites, both legitimate and malicious. The evaluation highlights the system's strengths and weaknesses, providing insights for future improvements.

The goal of this work is to contribute to the development of robust and accurate phishing detection techniques, ultimately helping to protect users from online threats.

2 Definition of the Linguistic Variables

To design the fuzzy expert system for detecting phishing websites, we referred to [3]. This paper lists 87 possible features (Boolean, float, and integer types) that could influence the detection of phishing websites. The proposed features are divided into three categories: URL-based features, content features, and external features. From these proposed variables, we selected five features that we consider relevant for detecting phishing websites. For further inspiration and understanding of the problem, we also consulted [6] and [2].

2.1 Chosen Features

2.1.1 URL-Based Features

Phish Hints

- **Description:** Number of words in the URL commonly associated with phishing websites.
- **Type:** Integer
- **Reference:** Feature 51 in the paper

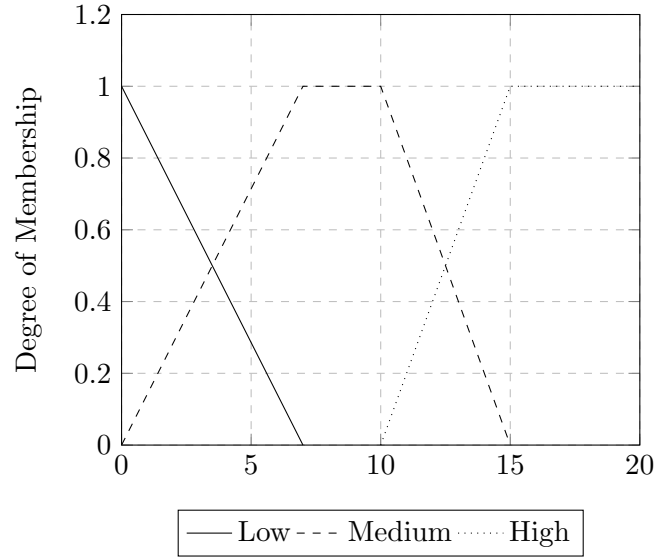


Figure 1: Membership Function for Phish Hints

Domain Age

- **Description:** Age of the website in months.
- **Type:** Integer
- **Reference:** Feature 83 in the paper

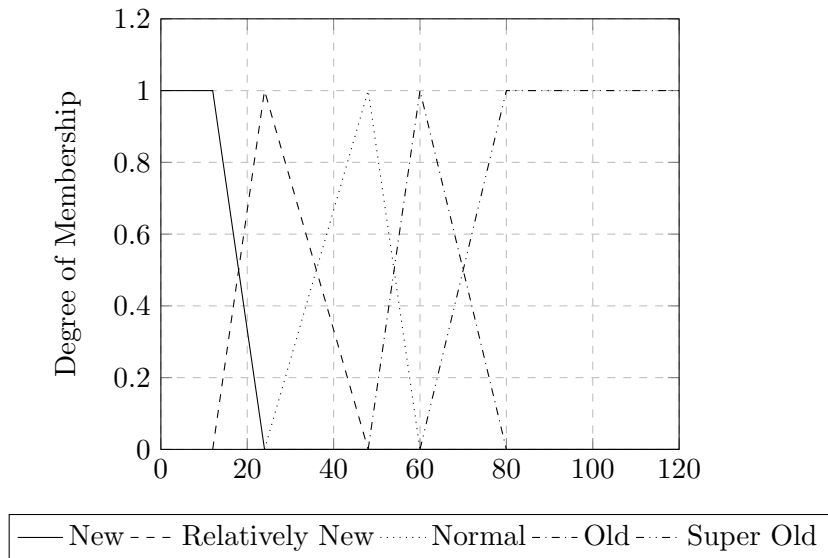


Figure 2: Membership Function for Domain Age (in months)

2.1.2 Content Features

Ratio of External Hyperlinks

- **Description:** The number of external hyperlinks on a webpage divided by the total number of hyperlinks.
- **Type:** Float
- **Reference:** Feature 59 in the paper

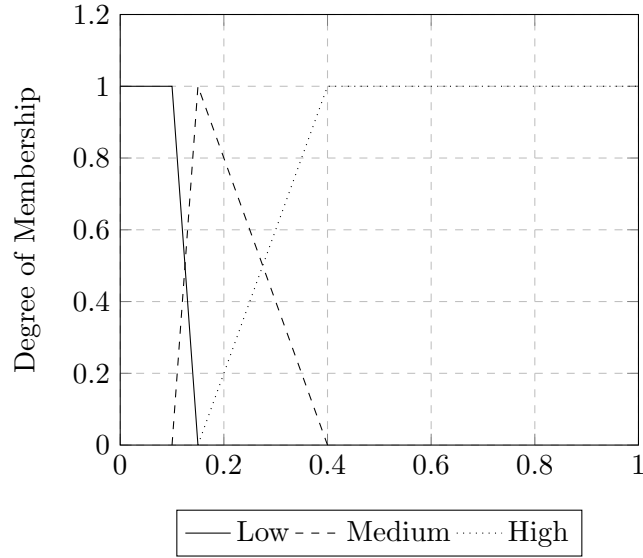


Figure 3: Membership Function for Ratio of External Hyperlinks

2.1.3 External Features

Google Index

- **Description:** Indicates whether a webpage is indexed by Google.
- **Type:** Boolean
- **Reference:** Feature 86 in the paper

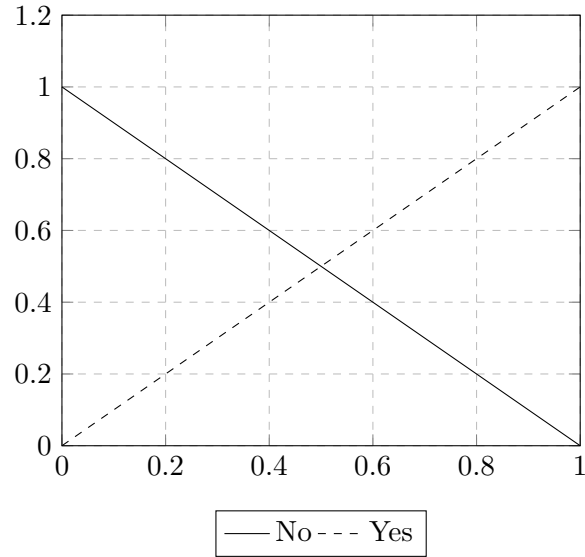


Figure 4: Membership Function for Google Index

GTR

- **Description:** Google Toolbar Rank, a value from 0 to 10 that reflects the importance of the webpage.
- **Type:** Integer
- **Reference:** Similar to feature 87 in the paper, which is PageRank

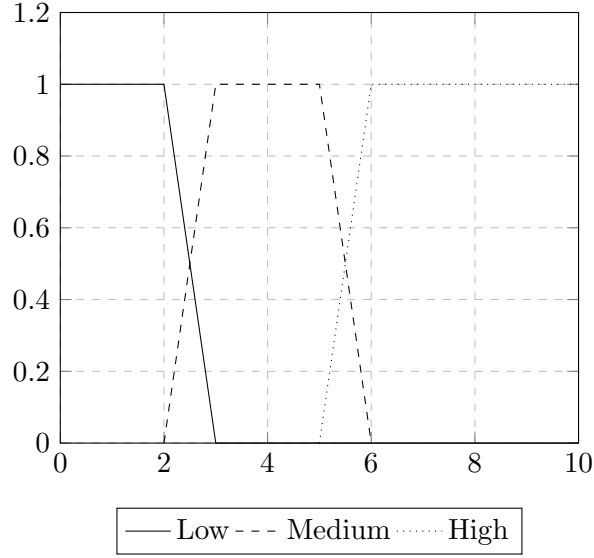


Figure 5: Membership Function for Google Toolbar Rank (GTR)

3 Output Variable

The output variable represents the phishing risk. It is classified into four fuzzy sets, as illustrated in Figure 6.

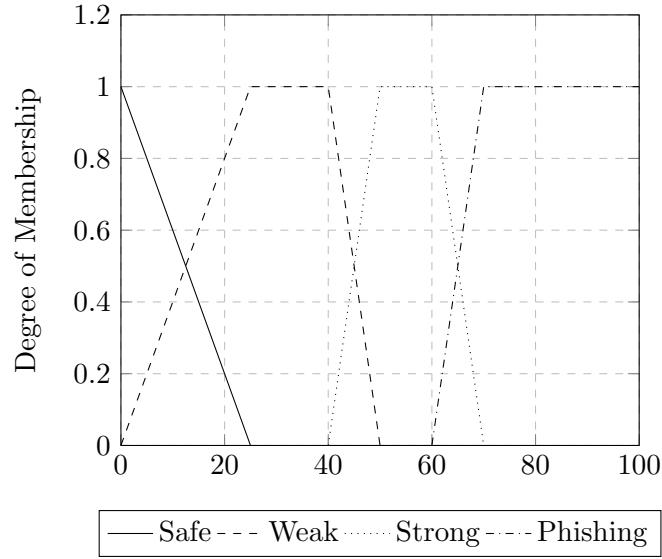


Figure 6: Membership Function for Phishing Risk (Output Variable)

4 Rules

To design the rules involving Google Index, we consulted [3], which states that pages not indexed by Google are more likely to be phishing websites. Consequently, we created the rules listed in Table 1.

| Google Index | Phishing Risk | Weight |
|--------------|---------------|--------|
| NO | PHISHY | 1 |
| YES | WEAK | 1 |

Table 1: Rules for Google Index

For the GTR, we consulted [5], which states:

GTR value is considered as a heuristic because the PageRank value for legitimate sites will be high, whereas for phishing pages, its value will be lower.

Regarding Phish Hints, [3] explains that a higher number of Phish Hints in a URL indicates a greater likelihood of the website being a phishing site. Accordingly, we defined the rules in Table 2.

| GTR | Phish Hints | Operator | Phishing Risk | Weight |
|--------|-------------|----------|---------------|--------|
| HIGH | | | SAFE | 1 |
| MEDIUM | LOW | AND | WEAK | 0.5 |
| MEDIUM | MEDIUM | AND | STRONG | 1 |
| MEDIUM | HIGH | AND | PHISHY | 1 |
| LOW | LOW | AND | STRONG | 0.5 |
| LOW | MEDIUM | AND | STRONG | 1 |
| LOW | HIGH | AND | PHISHY | 1 |

Table 2: Rules for GTR and Phish Hints

To evaluate the effect of the domain age feature, we consulted [4], which states:

The top feature in the list is domain age, which confirms our assumptions that long-running services are statistically more credible.

Thus, we consider older domains less likely to be phishing websites. Similarly, to assess the effect of the ratio of external hyperlinks, we consulted [1], which notes:

Phishing websites often include numerous external hyperlinks pointing to target websites because cybercriminals frequently replicate the HTML code from legitimate websites to construct their phishing sites.

All this information led to the creation of the rules in Table 3.

| Domain Age | Ratio of Hyperlinks | Operator | Phishing Risk | Weight |
|----------------|---------------------|----------|---------------|--------|
| SUPER OLD | | | SAFE | 1 |
| OLD | LOW | OR | SAFE | 0.5 |
| NORMAL | LOW | AND | WEAK | 0.5 |
| NORMAL | MEDIUM | AND | STRONG | 0.5 |
| NORMAL | HIGH | AND | STRONG | 0.5 |
| RELATIVELY NEW | LOW | AND | WEAK | 1 |
| RELATIVELY NEW | MEDIUM | AND | STRONG | 1 |
| RELATIVELY NEW | HIGH | AND | PHISHY | 0.5 |
| NEW | LOW | AND | STRONG | 0.5 |
| NEW | MEDIUM | AND | STRONG | 1 |
| NEW | HIGH | AND | PHISHY | 1 |

Table 3: Rules for Domain Age and Ratio of External Hyperlinks

5 Implementation

In our implementation, we utilized the MATLAB Fuzzy Logic Toolbox to develop the fuzzy expert system. The system was designed using the Mamdani fuzzy inference method, which is well-suited for decision-making processes requiring human-like reasoning. In this system, we employed the minimum (min) operation as the t-norm for the intersection of fuzzy sets, and the maximum (max) operation as the t-conorm for the union of fuzzy sets. For the defuzzification process, we selected the Center of Area (CoA) method, which calculates the centroid of the aggregated fuzzy set to produce a crisp output. This approach ensures that the

output is a balanced representation of the input conditions, providing a reliable decision-making framework for detecting phishing websites.

To validate the fuzzy expert system, we employed MATLAB 3D plots. As these plots only allow the representation of two inputs and the output at a time, we analyzed all ten possible combinations. The results were consistent, with no combinations left without an activated rule.

6 Testing

This section presents the testing of the fuzzy system. Four test cases, representing different scenarios, are evaluated. Some cases are designed to activate multiple rules. The results of each test case are reported with screenshots and explanations to justify the obtained outputs, including the activations of the rules. In the screenshots, we only provide the rules that are activated in each case.

6.1 Test Cases

The dataset used by [3] includes legitimate and phishing websites. Finding phishing websites that are still active was challenging, as they are often unavailable for extended periods. For the domain age, the data provided in the dataset (accurate at the time of creation) was used.

6.1.1 Case 1: Phishing Website

- **URL:** https://dichvuvnpt.com/home/components/com_user/bbtonline.html
- **Personal Opinion:** The site appears suspicious at first glance, as the formatting seems a bit off. The URL is also suspicious, but it is not immediately obvious that it is phishing.
- **Attributes:**
 - Domain Age: 3767 hours = 7.1 months
 - Phish Hints: 0
 - Ratio of External Hyperlinks: 1.0
 - Google Index: 1
 - GTR: 0
- **Rules Activated:**
 - **Rule 5:** If Phish Hints is Low and Google GTR is Low, then Phishing Risk is Strong
 - **Rule 9:** If Google Index is Yes, then Phishing Risk is Safe
 - **Rule 20:** If Domain Age is New and Ratio of External Hyperlinks is High, then Phishing Risk is Phishing
- **Output:** 60.9
- Classification for 60.9: **Strong** (with membership 1)

We see that the Google Index is not a clear indicator of phishing, as the site is indexed but still classified as phishing. The high ratio of external hyperlinks and relatively new domain age are the main factors leading to this classification.

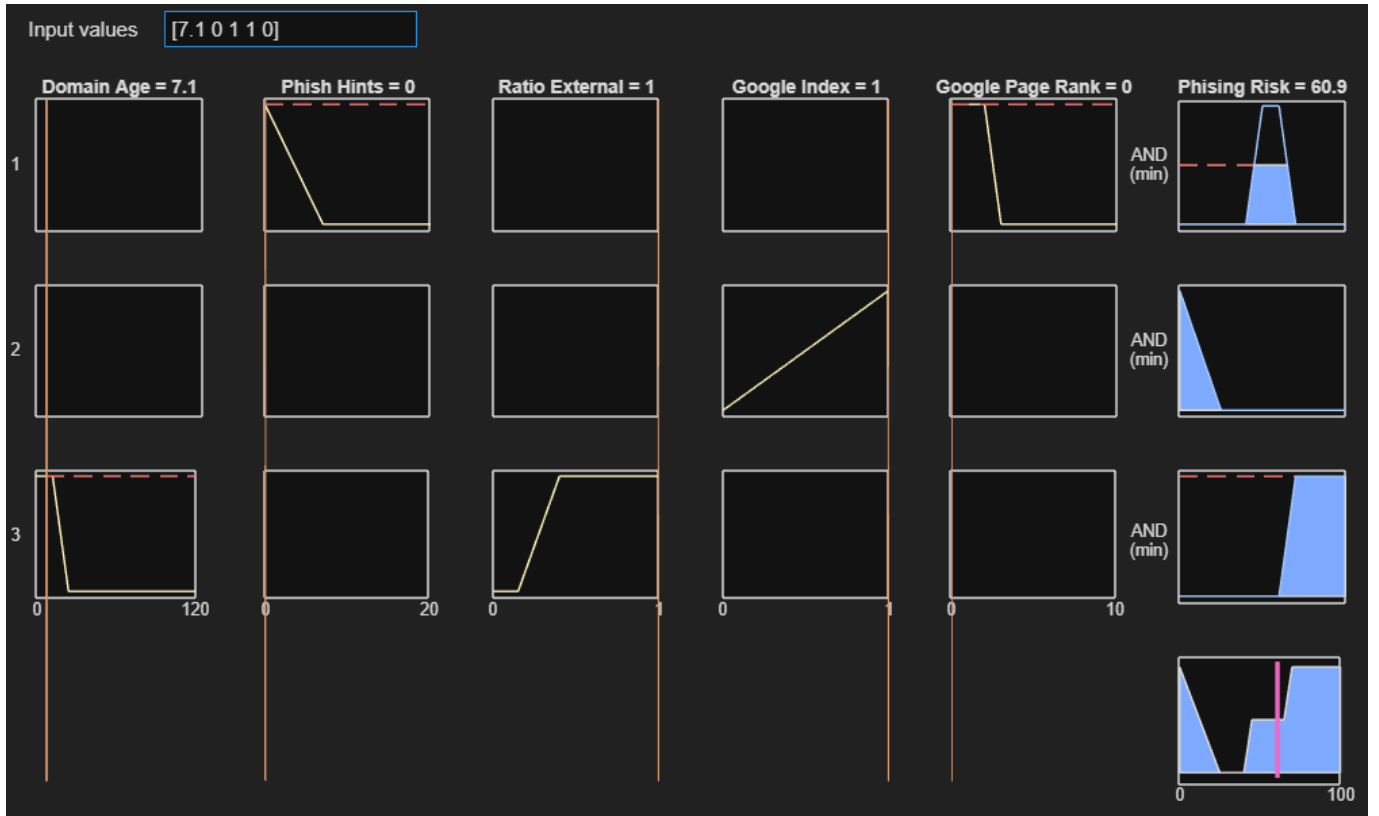


Figure 7: Test Case 1: Phishing Website

6.1.2 Case 2: Phishing Website

- **URL:** <https://www.courageon-immobilier.fr/> (line 6593)
- **Personal Opinion:** The site initially appears professional and trustworthy. However, closer inspection reveals phishing behavior (e.g., fake social media links).
- **Attributes:**
 - Domain Age: 2300 hours = 3.15 months
 - Phish Hints: 2
 - Ratio of External Hyperlinks: 0.4545
 - Google Index: 1
 - GTR: 2
- **Rules Activated:**
 - **Rule 5:** If Phish Hints is Low and Google GTR is Low, then Phishing Risk is Strong
 - **Rule 6:** If Phish Hints is Medium and Google GTR is Low, then Phishing Risk is Strong
 - **Rule 9:** If Google Index is Yes, then Phishing Risk is Safe
 - **Rule 20:** If Domain Age is New and Ratio of External Hyperlinks is High, then Phishing Risk is Phishing
- **Output:** 61.2
- **Classification for 61.2: Phishing** (with membership 1)

The site is classified as phishing due to its relatively new domain age, high ratio of external hyperlinks, and the presence of phishing hints. The Google GTR is low, but the Google Index is active again, which does not seem to be a strong indicator of phishing.

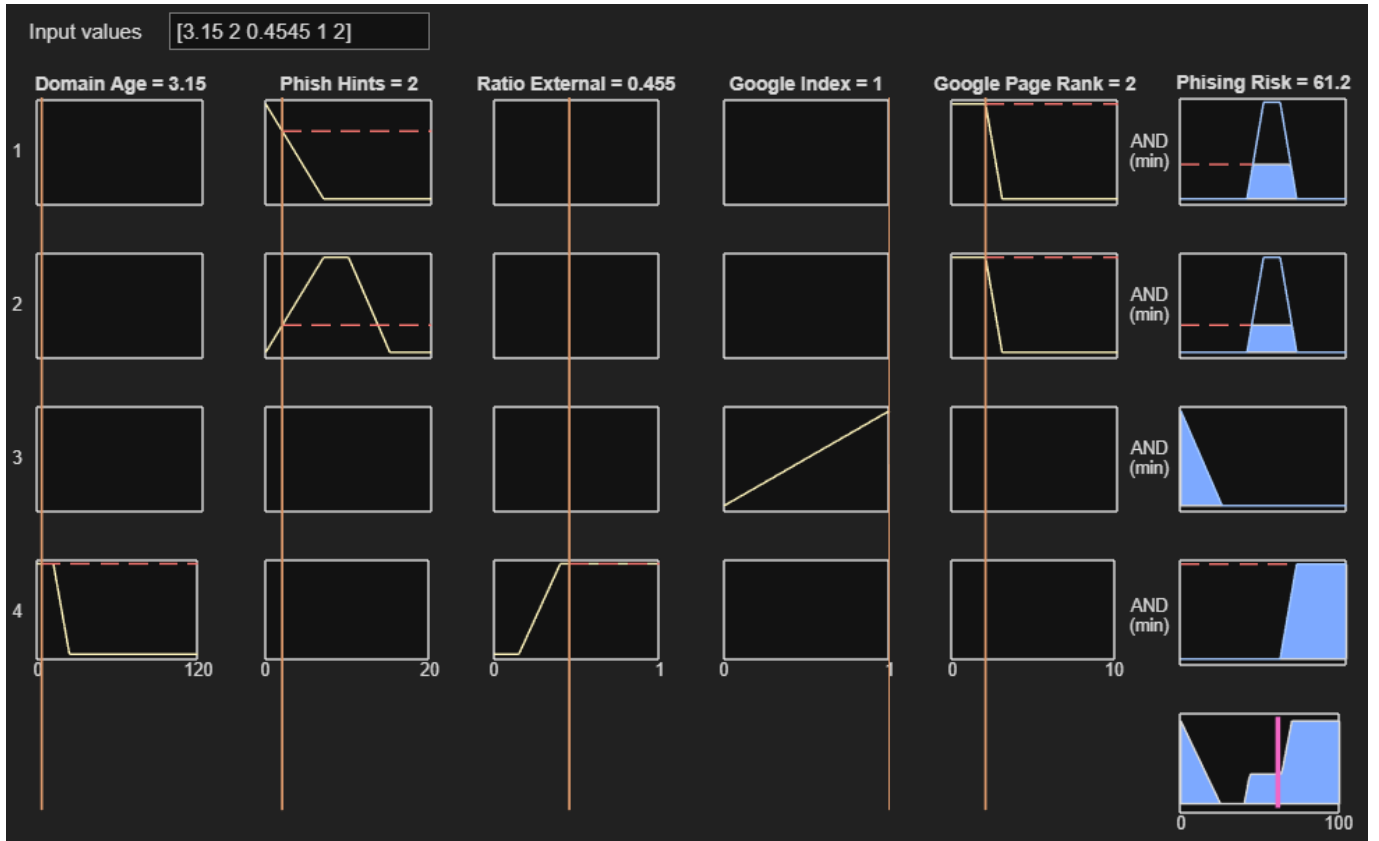


Figure 8: Test Case 2: Phishing Website

6.1.3 Case 3: Legitimate Website

- **URL:** https://www.physiologyweb.com/lecture_notes/membrane_transport/secondary_active_transport.html (line 6598)
- **Personal Opinion:** The site is for educational purposes, and it would be unexpected for it to be phishing.
- **Attributes:**
 - Domain Age: 3596 hours = 4.92 months
 - Phish Hints: 0
 - Ratio of External Hyperlinks: 0.1428
 - Google Index: 0
 - GTR: 4
- **Rules Activated:**
 - **Rule 2:** If Phish Hints is Low and Google GTR is Medium, then Phishing Risk is Weak
 - **Rule 8:** If Google Index is No, then Phishing Risk is Phishing
 - **Rule 19:** If Domain Age is Normal and Ratio of External Hyperlinks is Low, then Phishing Risk is Weak
- **Output:** 60.7
- Classification for 60.7: **Phishing** (with membership 0.9825), but Strong is also active with membership 1

Despite being a legitimate site, the system classifies it as likely phishing due to its relatively recent creation and lack of Google indexing. This is a false positive, as the site is a legitimate educational resource.

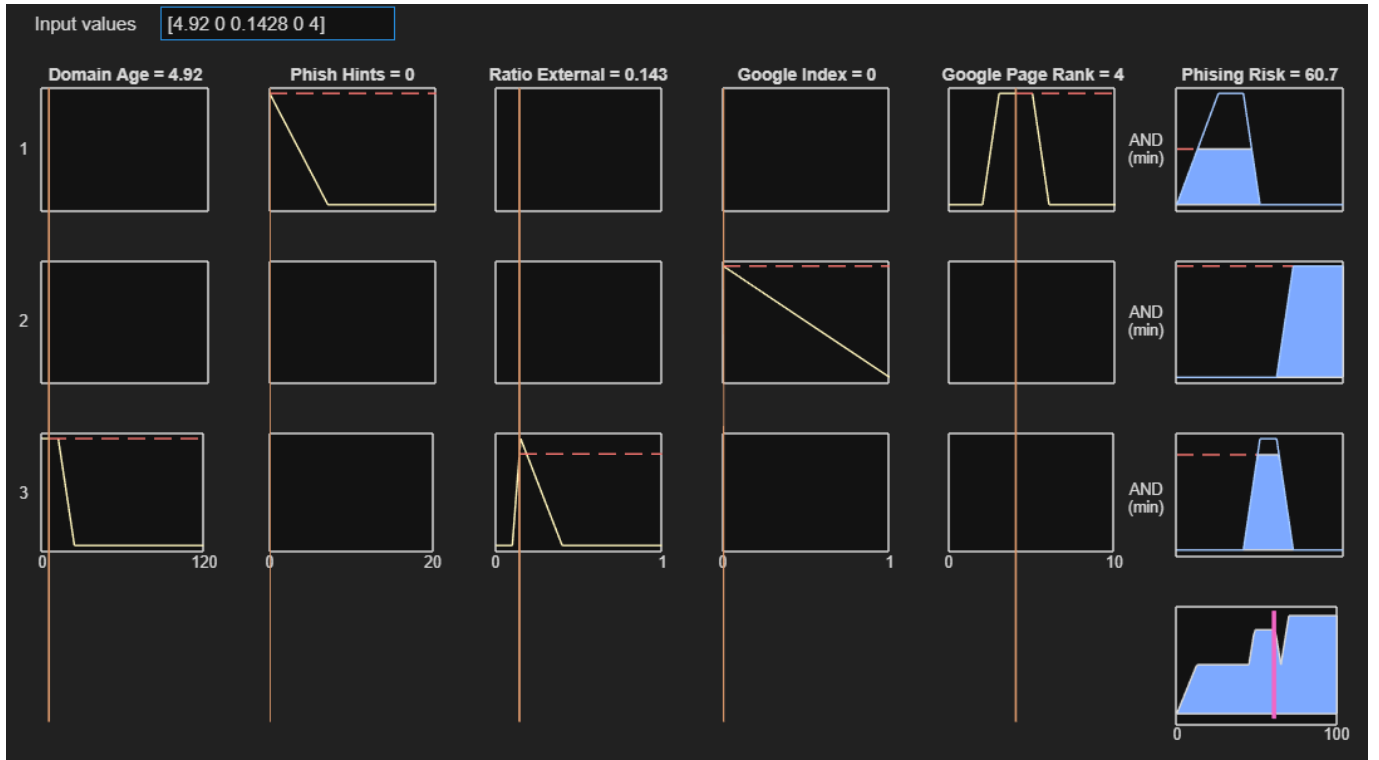


Figure 9: Test Case 3: Legitimate Website (False Positive)

6.1.4 Case 4: Legitimate Website

- **URL:** https://en.wikipedia.org/wiki/Aurelio_Voltaire (line 6638)
- **Personal Opinion:** As Wikipedia is a well-known and trustworthy site, it is expected to be classified as legitimate.
- **Attributes:**
 - Domain Age: 7132 hours = 9.8 months
 - Phish Hints: 0
 - Ratio of External Hyperlinks: 0.1305
 - Google Index: 1
 - GTR: 7
- **Rules Activated:**
 - **Rule 1:** If Google GTR is High, then Phishing Risk is Safe
 - **Rule 9:** If Google Index is Yes, then Phishing Risk is Safe
 - **Rule 11:** If Domain Age is Old or Ratio of External Hyperlinks is Low, then Phishing Risk is Safe
 - **Rule 18:** If Domain Age is New and Ratio of External Hyperlinks is Low, then Phishing Risk is Strong
 - **Rule 19:** If Domain Age is Normal and Ratio of External Hyperlinks is Low, then Phishing Risk is Weak
- **Output:** 32.8
- Classification for 32.8: **Weak** (with membership 0.52)

The high Google GTR is a strong indicator of legitimacy. The site also has a low ratio of external hyperlinks, which is another indicator of trustworthiness. The site is classified as weak, which is expected for a well-known site like Wikipedia.

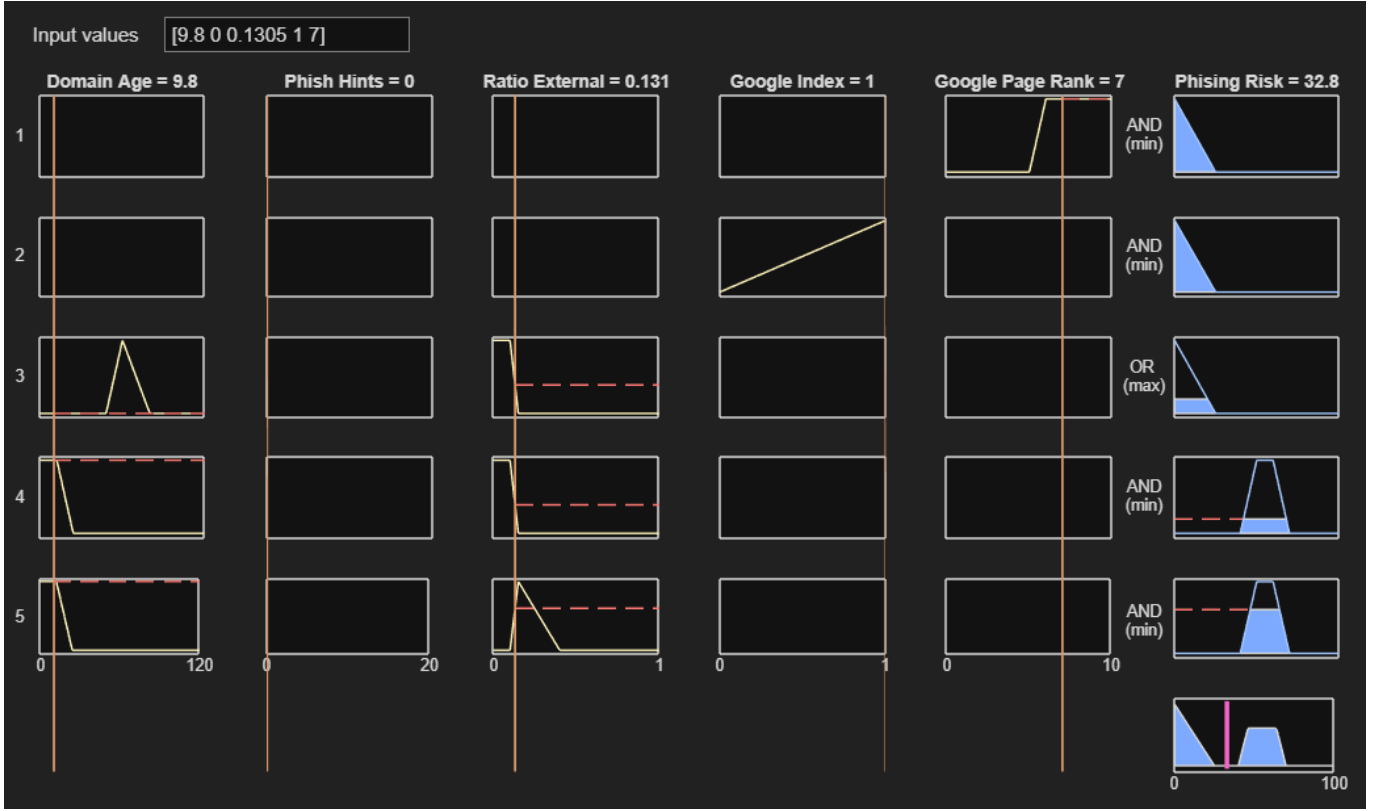


Figure 10: Test Case 4: Legitimate Website

7 Complex Fuzzy Expert System

This fuzzy expert system is structured into two hierarchical layers. The first layer consists of three independent Mamdani fuzzy systems, each focusing on a specific dimension of a website: **URL-based features**, **Content-based features**, and **External-based features**. The outputs from these three systems are then combined in a second-layer Mamdani fuzzy system to compute the overall phishing risk score.

7.1 Layer 1: Feature-Based Risk Assessments

The first layer is composed of three Mamdani fuzzy inference systems, each focusing on a specific feature category. These systems process their respective input features to generate risk levels as outputs:

- URL Risk Mamdani System:

This system evaluates whether characteristics of the URL indicate phishing behavior, such as excessively long URLs or unusual characters.

Examples of outputs include:

- URL Length (f1-f2)
- Special Characters in the URL(f4-f20)

- Content Risk Mamdani System:

This system identifies suspicious behaviors in the webpage content, such as malicious forms or hidden elements designed to deceive users.

Examples of outputs include:

- Number of Hyperlinks (f57)
- Login Form Actions (f66)
- Presence of Invisible Iframes (f73)

- External Risk Mamdani System:

This system assesses external attributes that provide context for the website's legitimacy, such as its age or popularity.

Examples of outputs include:

- Domain Age (f83)
- Google Index Presence (f86)
- Web Traffic Volume (f84)

7.2 Layer 2: Overall Phishing Risk Assessment

The second layer takes the outputs from the first layer (URL Risk, Content Risk, External Risk) and combines them into an overall phishing risk score using another Mamdani fuzzy inference system. Each of these risks carries a specific weight, reflecting its relative importance in detecting phishing. The weights are designed to balance sensitivity (detecting as many phishing websites as possible) and precision (minimizing false positives).

7.3 Weight Distribution

The importance of each risk category (URL, Content, External) varies based on its correlation with phishing behavior. By assigning appropriate weights, the system can prioritize the most reliable and significant indicators:

- URL Risk (40%-50%):

This category carries the highest weight because a suspicious URL alone can often be enough to classify a website as high-risk. However, legitimate websites may also have long or unusual URLs, so this weight must be carefully balanced.

- Content Risk (30%-40%):

This category carries slightly less weight than URL risk because, while content indicators are highly relevant, legitimate websites can occasionally exhibit similar patterns (e.g., a minimalist design with few links). Nonetheless, content risk remains a strong contributing factor.

- External Risk (20%-30%):

External risks carry the lowest weight because they are indirect indicators. While they add valuable context, they are less conclusive compared to URL or content risks.

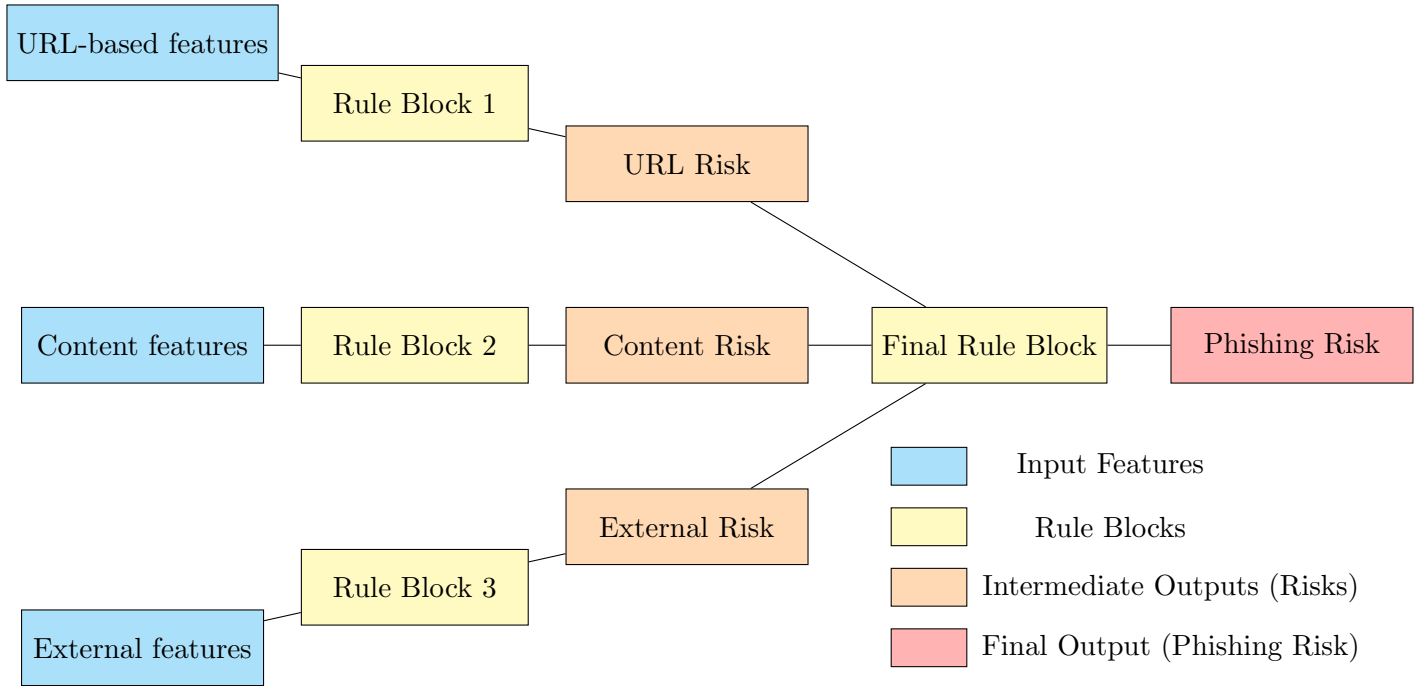


Figure 11: Hierarchical Fuzzy Expert System for Phishing Detection with Rule Blocks

References

- [1] Ali Aljofey et al. “An effective detection approach for phishing websites using URL and HTML features”. In: *Scientific Reports* 12.1 (May 2022), p. 8842.
- [2] Zuochao Dou et al. “Systematization of Knowledge (SoK): A Systematic Review of Software-Based Web Phishing Detection”. In: *IEEE Communications Surveys & Tutorials* 19.4 (2017), pp. 2797–2819. DOI: 10.1109/COMST.2017.2752087.
- [3] Abdelhakim Hannousse and Salima Yahiouche. “Towards benchmark datasets for machine learning based website phishing detection: An experimental study”. In: *Engineering Applications of Artificial Intelligence* 104 (2021), p. 104347. ISSN: 0952-1976. DOI: <https://doi.org/10.1016/j.engappai.2021.104347>. URL: <https://www.sciencedirect.com/science/article/pii/S0952197621001950>.
- [4] Radek Hranický et al. “Unmasking the Phishermen: Phishing Domain Detection with Machine Learning and Multi-Source Intelligence”. In: (2024), pp. 1–5. DOI: 10.1109/NOMS59830.2024.10575573.
- [5] A. Naga Venkata Sunil and Anjali Sardana. “A PageRank based detection technique for phishing web sites”. In: (2012), pp. 58–63. DOI: 10.1109/ISCI.2012.6222667.
- [6] Rasha Zieni, Luisa Massari, and Maria Carla Calzarossa. “Phishing or Not Phishing? A Survey on the Detection of Phishing Websites”. In: *IEEE Access* 11 (2023), pp. 18499–18519. DOI: 10.1109/ACCESS.2023.3247135.