IST-Africa 2025 Conference Proceedings
Miriam Cunningham and Paul Cunningham (Eds)
IST-Africa Institute and IIMC, 2025
ISBN: 978-1-905824-74-8

Fuzzy Logic Predictive Model for Early Diagnosis and Prognosis of Breast and Prostate Cancer

Aggrey SHITSUKANE[1], Michael ODHIAMBO[1], Walter KIHUYA[2], Samuel KARIUKI[1]
Mathew EGESSA[1]
[1]*Technical University of Mombasa Box 90420-80100, Kenya*
Tel: +254724963044, Email: shitsukane@tum.ac.ke; md.odhiambo@tum.ac.ke; samuelkariuki@tum.ac.ke; megessa@tum.ac.ke
[2] *The Management University of Africa, P.O Box 29677-00100, Nairobi, Kenya*
Tel: +254720590532, Email: wkihuya@mua.ac.ke

Abstract: Breast and prostate cancers are major global health concerns, where early diagnosis and accurate prognosis are critical for improving survival rates. Traditional diagnostic methods, like mammography for breast cancer and prostate-specific antigen (PSA) testing for prostate cancer, have limitations in specificity and sensitivity. Predicting cancer prognosis involves complex clinical and pathological factors, which traditional models struggle to integrate effectively. Fuzzy logic offers a promising approach to enhance diagnostic and prognostic accuracy by handling imprecise information. Operating on degrees of truth, fuzzy logic is well-suited for complex medical systems. It has been applied in cardiovascular risk assessment and diabetes management. In breast cancer, factors including tumor size, lymph node involvement, and genetic markers are critical, while in prostate cancer, PSA levels, Gleason score, and tumor stage are key. Integrating fuzzy logic models into clinical decision support systems (CDSS) can enhance early diagnosis and prognosis prediction by providing interpretable, rule-based recommendations that align with clinical reasoning.

Keywords: Artificial intelligence (AI), Cancer prediction, Fuzzy logic, Clinical decision support system (CDSS), Medical data uncertainty.

1. Introduction

Breast and prostate cancers are among the most rampant and impactful menaces worldwide. Early diagnosis and accurate prognosis are crucial for effective treatment and improving survival rates [1]. Traditional diagnostic methods and statistical models have limitations in handling complexity and uncertainty inherent in medical data [2]. Fuzzy logic, which allows for modelling of imprecise and ambiguous information, offers a promising approach to enhance diagnostic and prognostic accuracy.

In the context of this paper, prediction is the application of fuzzy logic models to estimate the likelihood of developing breast or prostate cancer, using clinical variables like tumour size, lymph node involvement, hormone receptor status, PSA levels, and Gleason scores. Diagnosis is evaluating whether a patient has breast or prostate cancer using clinical and pathological evidence. The research focuses on utilizing fuzzy logic to increase cancer detection accuracy by accounting for ambiguity in medical data. The research examines how fuzzy logic models assess disease severity and anticipate cancer's possible path, including recurrence risk. Cancer risk is the likelihood of a patient getting or having cancer based on specific diagnostic signs. The research uses fuzzy logic to define risk categories (low, moderate, high, and extremely high) and provides a more nuanced and interpretable assessment.

Breast and prostate cancers are common cancers worldwide, with significant impacts on public health. Breast cancer is a frequently diagnosed cancer among women, with over 2.3 million new cases and 685,000 deaths reported globally in 2020 [3]. Prostate cancer is the most common cancer in men, with an estimated 1.4 million new cases and 375,000 deaths in the year 2020 [4]. Early diagnosis and accurate prognosis are critical for improving survival rates and reducing the burden of these cancers.

Traditional methods for diagnosing breast and prostate cancer include imaging techniques, laboratory tests, and biopsies. Mammography is the primary screening tool for breast cancer, with high sensitivity but also a notable rate of false positives and negatives [5]. Prostate-specific antigen (PSA) testing and digital rectal examinations (DRE) are standard for prostate cancer screening, but they too have limitations in specificity and sensitivity [6]. Magnetic resonance imaging (MRI) and ultrasound provide additional diagnostic information but not universally available.

Accurately predicting prognosis of breast and prostate cancer patients involves evaluating various clinical and pathological factors. For breast cancer, factors such as tumor size, lymph node involvement, hormone receptor status, and genetic markers are critical [7]. Prostate cancer prognosis depends on PSA levels, Gleason score, tumor stage, and patient age [8]. Despite the availability of these factors, existing statistical models often struggle with the inherent complexity and variability of cancer progression.

Artificial intelligence (AI) techniques have shown promise in enhancing cancer diagnostics. Methods such as logistic regression, decision trees, and neural networks have been applied to predict cancer outcomes with varying degrees of success [9] [10]. However, these methods are limited by their "black-box" nature, which reduces interpretability and clinical acceptance.

Fuzzy logic, introduced by Zadeh (1965), is a mathematical framework for dealing with uncertainty and imprecision. Unlike classical logic that operates on binary true/false values, fuzzy logic allows for degrees of truth represented by values between 0 and 1. This flexibility makes fuzzy logic particularly suitable for modelling complex systems where data may be incomplete or ambiguous [11]. Advantages of Fuzzy Logic in Oncology includes handling Uncertainty and variability inherent in medical data. Fuzzy logic models can incorporate clinical expertise through the creation of intuitive if-then rules, enhancing their interpretability and acceptance among clinicians. Fuzzy systems can be easily updated as new medical knowledge and data become available, ensuring that they remain relevant and accurate over time [12].

Fuzzy logic has been applied successfully in various medical domains, including diagnostics, treatment planning, and decision support systems. For instance, fuzzy logic models have been used to assess the risk of cardiovascular diseases [13], manage diabetes [14], and predict patient outcomes in intensive care units [15]. These applications demonstrate the potential of fuzzy logic to improve clinical decision-making by incorporating expert knowledge and handling uncertainty effectively.

Fuzzy logic models have shown potential in cancer diagnostics and prognostics. For breast cancer, fuzzy logic has been used to interpret mammographic findings to improve diagnostic accuracy [16]. In prostate cancer, fuzzy models have been applied to PSA levels, biopsy results, and MRI findings to enhance the prediction of cancer presence and progression [8]. These studies highlight the ability of fuzzy logic to synthesize clinical inputs into a coherent diagnostic assessment.

Integrating fuzzy logic models into clinical decision support systems (CDSS) can provide clinicians with valuable tools for early diagnosis and prognosis prediction. CDSS equipped with fuzzy logic can offer interpretable, rule-based recommendations that align with clinical reasoning, facilitating better patient management [17]. By continuously updating the rule base with new clinical data and expert input, fuzzy logic models can remain adaptive and relevant in dynamic healthcare environments.

2. Objectives

1. To design fuzzy logic predictive models for diagnosis of breast and prostate cancer.
2. To develop fuzzy logic predictive model for diagnosis of breast and prostate cancer.
3. To validate fuzzy logic predictive model for diagnosis of breast and prostate cancer

3. Methodology

The methodology outlines the steps for developing and validating fuzzy logic predictive models for early diagnosis and prognosis prediction of breast and prostate cancer. The general structure of the fuzzy rule-based system is illustrated in Figure 1. The fuzzification block takes specific, real-world inputs and transforms them into fuzzy sets. In this process, a single precise value is converted into a spectrum of possible values, each with varying degrees of membership. The fuzzy inference engine is the heart of the system, applying logical rules to the fuzzified inputs to make decisions. Using IF-THEN rules, it processes these fuzzy inputs to generate corresponding fuzzy outputs. It contains expert knowledge that drives the system’s behaviour. The database stores membership functions and other key parameters necessary for the fuzzy logic system to function. These membership functions are applied during both the fuzzification and defuzzification stages. Once the fuzzy inference engine generates fuzzy outputs, defuzzification block converts them back into precise values. This step is essential since real-world applications require exact outputs. The input refers to the specific, real-world data that the fuzzy system processes, while the output is the final, crisp result of the process.

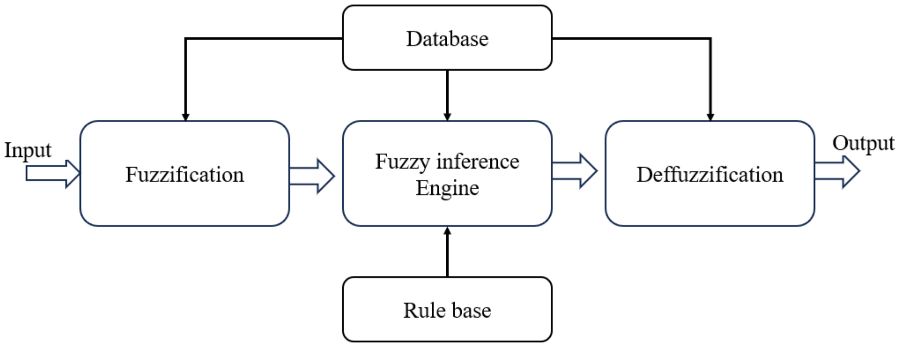


Figure 1: General structure of fuzzy rule-based system

Defuzzification step was done using the Centroid of Area (COA) that computed the crisp output from the aggregated fuzzy set.

$$y_{crisp} = \frac{\int y.\mu B_{agg}\,(y)dy}{\int \mu B_{agg}\,(y)dy}$$

Where

y_{crisp} = Variable representing the output

Numerator = $\int y.\mu B_{agg}\,(y)dy$ represents the weighted sum of the output values.

Denominator = $\int \mu B_{agg}\,(y)dy$ represents the total area under the aggregated membership function.

The formula calculates area under the aggregated fuzzy set, providing a crisp, single-valued output that reflects the overall inference of the fuzzy system. Clinical experts were used to developed the set of if-then 15 rules for prostate and 36 rules for breast cancer.

We implemented a fuzzy inference system (FIS) to process the input data and apply the rule base, Combined the results of multiple rules to produce a fuzzy output. Combined rules allow for a detailed, layered approach to assessing cancer risk.

Detailed explanation of the factors is explained next: Prostate-Specific Antigen (PSA) is a protein produced by cells in the prostate gland, a small gland in men that produces seminal fluid. PSA is present in small amounts in the blood of healthy men. However, its levels can increase due to various conditions, including prostate cancer. The PSA levels are between 0 and 2.5 ng/mL, prostate cancer is low chance that you would require further testing, A PSA level between 2.5 and 4.0 ng/mL is normal and PSA levels between 4.0 and 10.0 is higher risks.

The Gleason Score is a grading system used to evaluate the aggressiveness of prostate cancer based on the microscopic appearance of cancer cells in a biopsy sample. It helps determine how likely a prostate cancer is to grow and spread, guiding treatment decisions Gleason score, Grade 1: The cancerous cells look a lot like normal cells. Grades 2-4: Cancerous cells in the tissue look less like normal cells. Grade 5: Cancerous cells look very abnormal.

Tumor Staging is a system used to describe the size and extent of cancer within the body. It helps determine the prognosis, guide treatment decisions, and assess the likelihood of cancer spreading Tumor stage, stage 0 carcinoma in situ, a precancerous change, stage 1 the tumour is usually small and has not grown outside of the organ where it started, in stages 2 and 3 the tumour is larger or has grown outside of the organ it started in to nearby tissue and stage 4 the cancer has spread through the blood or lymphatic system to a distant site in the body (metastatic spread)

Patient Age is a significant factor in the diagnosis, treatment, and prognosis of cancer, including prostate cancer. It influences decisions regarding the aggressiveness of treatment, potential side effects, and overall management strategies patient age: older age more than 70 years, Middle aged 40-69 years, and Young aged less than 40 years.

Breast cancer the rule base combines the tumor size, lymph node involvement, and hormone receptor status to provide a comprehensive prediction of cancer risk. Each rule captures the interaction of these factors, allowing for an assessment of the overall cancer risk. Tumor size refers to the physical dimensions of the breast tumor, usually measured in centimetres. It is a crucial prognostic factor in breast cancer as it often correlates with the stage of cancer and potential spread. tumor size is categorized into linguistic variables such as "small," "medium," and "large." The boundaries between these categories are not strict, allowing for smooth transitions. A tumor less than 2 cm is considered stage T1, between 2 and 5 cm to be stage T2, greater than 5 cm, it's typically stage T3 and at stage T4 has spread into the chest wall or skin.

Lymph node involvement refers to whether cancer cells have spread to the lymph nodes near the breast, particularly the axillary lymph nodes. The presence of cancer cells in lymph nodes is a strong indicator of cancer spread. Lymph node status can be expressed in fuzzy terms like "negative," "low involvement," and "high involvement" based on the number of affected nodes and their spread. Level one lymph nodes are lateral to the pectoralis major, level two lymph nodes are posterior to the pectoralis major and level three lymph nodes are medial to the pectoralis major. Hormone receptor status refers to the presence of estrogen receptors (ER) and progesterone receptors (PR) on the surface of cancer cells. These receptors indicate how the cancer might respond to hormone therapy. This factor is divided into categories such as "positive," "negative," or "mixed." Each category affects treatment decisions and prognosis. Fuzzy logic models used these factors as inputs to create rules to predict the likelihood of cancer recurrence, response to treatment, or overall prognosis. These rules resemble clinical decision-making patterns, accounting for the imprecision and variability inherent in medical data.

In prediction of prostate cancer as in table 1 and breast cancer shown in table 2, rule base combines variables to provide output cancer risk using MATLAB fuzzy logic toolbox

Table 1. Membership functions and linguistic terms for prostate cancer

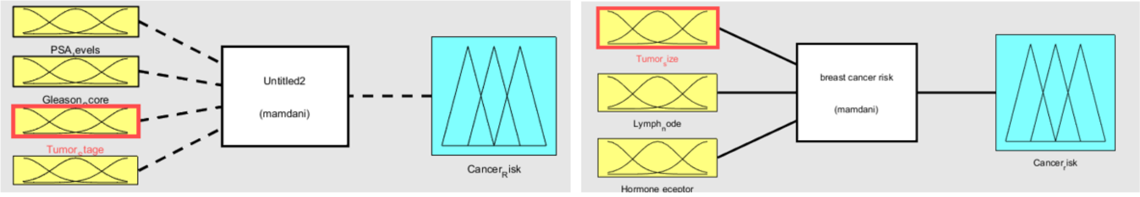
Variables	Categories	Parameters
PSA levels (ng/mL)	Low :0-2.5	[-5.4 -0.6 1.896 2.5]
	Moderate 2.5-4.0	[1.279 2.409 4.242 4.0]
	High 4.0-10.0	[2.66 4.92 8.054 10]
	Very high >10	[6.962 9.89 12.3 16]
Gleason Score	Grade 1: cells look similar to normal cells	[-2 0 2]
	Grade 2-4: cells look less like normal cells	[0.7 1.579 3.37 4.54]
	Grade 5: cells look very abnormal.	[3.019 5.019 7.019]
Tumor Stage	Stage 0: precancerous change	[-1.992 0.01303 1.208]
	Stage 1: small tumor hasn’t spread	[0.118 1.007 2.33]
	Stage 2-3: tumor grown outside the original organ	[1.425 1.92 2.85 3.29]
	Stage 4: metastatic	[2.761 3.952 4.972 7.632]
Age	Young: <40 years	[-36 -4 28.6 40.0]
	Middle aged: 40-69 years	[14 39.4 68.8 91.44]
	Older: > 70 years	[64.2 80.66 104 136]
Cancer Risk	Low risk	[-0.348 0.00347 0.0783 0.27]
	Moderate risk	[0.1 0.321 0.496]
	High risk	[0.371 0.537 0.7783]
	Very high risk	[0.68 0.769 1.01 1.19]

Table 2. Membership functions and linguistic terms for breast cancer

Variables	Categories	Parameters
Tumor size	T1-less than 2cm	[-5.4 -0.6 1.896 2.5]
	T2-2-4 cm	[1.279 2.409 4.242 4.0]
	T3- greater than 5cm	[2.66 4.92 8.054 10]
	T4-spread into the chest wall or skin	[6.962 9.89 12.3 16]
Lymph node involvement	Level 1 lateral to the pectoralis major	[-2 0 2]
	Level 2 – Posterior to the pectoralis major.	[0.7 1.579 3.37 4.54]
	Level 3- Medial to the pectoralis	[3.019 5.019 7.019]
Hormone receptor status	ER positive: Estrogen receptor positive	[-1.992 0.01303 1.208]
	PR positive: progesterone receptor positive	[0.118 1.007 2.33]
	HER2 positive: Human epidermal growth factor receptor 2-positive	[0.6 1 1.4]
Cancer Risk	Low risk	[-0.348 0.00347 0.0783 0.27]
	Moderate risk	[0.1 0.321 0.496]
	High risk	[0.371 0.537 0.7783]
	Very high risk	[0.68 0.769 1.01 1.19]

Sample rules:
"IF Tumor Size is T2 AND Lymph Node is Level 2 AND Hormone Receptor is PR Positive THEN Cancer Risk is **High**".
“IF PSA is High AND Gleason is Grade 4-5 AND Tumor Stage is Stage 3 AND Age is Middle-Aged THEN Cancer Risk is **High”.**

Using Kaggle datasets, specifically the Breast Cancer Wisconsin Dataset [19] that includes features such as tumor size, shape, and other clinical measurements and Prostate Cancer Data [20] with patient-specific details like PSA levels, Gleason scores, and other relevant features for predicting prostate cancer. Quantitative analysis was performed to evaluate the model's accuracy and determine the impact of individual features on cancer risk using a sample dataset of 1000 patients. Figure 2 shows the fuzzy logic models used to predict cancer risk. Mamdani fuzzy inference engine was used because of its ability to handle complex, multi-dimensional problems with interpretability and transparency essential in the medical field. Its capacity to provide fuzzy outputs also better aligns with the uncertainty inherent in predicting diseases like cancer.



(a) (b)

Figure 2: Mamdani FIS (a) prostate cancer (b) breast cancer.

4. Technology Description

Fuzzy logic is a powerful tool in enhancing prediction and prognosis of breast and prostate cancers by effectively managing complexities and uncertainties inherent in medical data. Unlike traditional diagnostic methods, such as mammography for breast cancer and PSA testing for prostate cancer, which often have limitations in specificity and sensitivity, fuzzy logic allows for modelling of imprecise information through a system of "if-then" rules derived from clinical expertise. This approach considers critical factors like tumor size, lymph node involvement, and hormone receptor status in breast cancer, as well as PSA levels, Gleason score, and tumor stage in prostate cancer. The fuzzy logic system uses fuzzification to convert these clinical inputs into fuzzy sets, processes them through a rule base using the Mamdani inference method, and then defuzzifies the results into crisp values that predict cancer risk. This flexible, rule-based system mirrors clinical reasoning, making it more interpretable and adaptable to new data, thus enhancing early diagnosis.

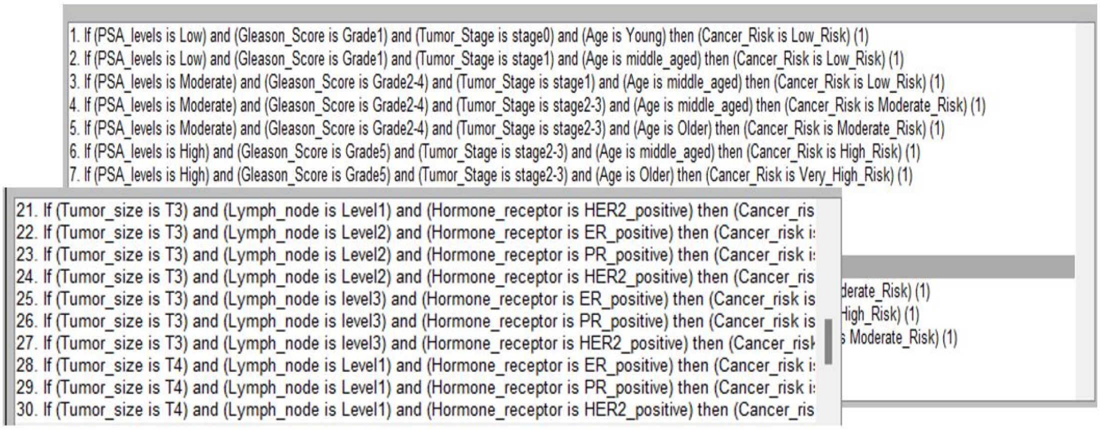
Integrating fuzzy logic models into Clinical Decision Support Systems (CDSS) can offer a significant advantage in oncology by providing clinicians with interpretable, adaptable, and accurate recommendations for patient management. The fuzzy logic approach allows for the continuous updating of rules as new clinical insights emerge, ensuring the models remain relevant in a rapidly evolving medical landscape. By handling the inherent variability in patient data, fuzzy logic facilitates risk assessments that traditional statistical models struggle with, supporting more informed decision-making. This technology has potential in improving diagnostic accuracy and prognosis prediction, ultimately contributing to better patient outcomes by guiding tailored treatment plans and early interventions in breast and prostate cancers.

5. Developments

The project effectively integrated critical clinical factors for both breast cancer (tumor size, lymph node involvement, hormone receptor status) and prostate cancer (PSA levels, Gleason score, tumor stage) into a fuzzy logic framework to enhance accuracy of cancer diagnosis and prognosis. Using Mamdani inference engine, the study processed these inputs through a rule-based system, providing exact analysis of cancer risks that aligns with clinical decision-making processes. By embedding fuzzy logic models into clinical decision support systems (CDSS), the system can facilitate personalized cancer risk predictions and treatment recommendations tailored to individual patient data, advancing precision medicine. The rule-based structure of fuzzy logic ensured that the model’s outputs were both interpretable and clinically relevant, maintaining accuracy as medical knowledge evolves. The model’s ability to manage diverse datasets makes it suitable for deployment across various healthcare environments, from large hospitals to smaller clinics. Furthermore, the fuzzy logic model demonstrated robustness in comparison with other machine learning models, confirming its effectiveness in cancer risk prediction.

6. Results

In this section, we present the outcomes of applying the proposed fuzzy logic model for cancer prediction. The model was developed to handle the inherent uncertainty and imprecision in medical data, providing a robust framework for early cancer diagnosis. Through the integration of clinical parameters and fuzzy inference systems, the model demonstrated its ability to predict cancer risk. The results were assessed in terms of key performance metrics, including sensitivity, specificity, and overall prediction accuracy. Additionally, we compared the model’s predictions with actual clinical diagnoses to evaluate its effectiveness in real-world scenarios.



```
1. If (PSA_levels is Low) and (Gleason_Score is Grade1) and (Tumor_Stage is stage0) and (Age is Young) then (Cancer_Risk is Low_Risk) (1)
2. If (PSA_levels is Low) and (Gleason_Score is Grade1) and (Tumor_Stage is stage1) and (Age is middle_aged) then (Cancer_Risk is Low_Risk) (1)
3. If (PSA_levels is Moderate) and (Gleason_Score is Grade2-4) and (Tumor_Stage is stage1) and (Age is middle_aged) then (Cancer_Risk is Low_Risk) (1)
4. If (PSA_levels is Moderate) and (Gleason_Score is Grade2-4) and (Tumor_Stage is stage2-3) and (Age is middle_aged) then (Cancer_Risk is Moderate_Risk) (1)
5. If (PSA_levels is Moderate) and (Gleason_Score is Grade2-4) and (Tumor_Stage is stage2-3) and (Age is Older) then (Cancer_Risk is Moderate_Risk) (1)
6. If (PSA_levels is High) and (Gleason_Score is Grade5) and (Tumor_Stage is stage2-3) and (Age is middle_aged) then (Cancer_Risk is High_Risk) (1)
7. If (PSA_levels is High) and (Gleason_Score is Grade5) and (Tumor_Stage is stage2-3) and (Age is Older) then (Cancer_Risk is Very_High_Risk) (1)
```

```
21. If (Tumor_size is T3) and (Lymph_node is Level1) and (Hormone_receptor is HER2_positive) then (Cancer_ris
22. If (Tumor_size is T3) and (Lymph_node is Level2) and (Hormone_receptor is ER_positive) then (Cancer_risk i
23. If (Tumor_size is T3) and (Lymph_node is Level2) and (Hormone_receptor is PR_positive) then (Cancer_risk i
24. If (Tumor_size is T3) and (Lymph_node is Level2) and (Hormone_receptor is HER2_positive) then (Cancer_ris
25. If (Tumor_size is T3) and (Lymph_node is level3) and (Hormone_receptor is ER_positive) then (Cancer_risk is
26. If (Tumor_size is T3) and (Lymph_node is level3) and (Hormone_receptor is PR_positive) then (Cancer_risk is
27. If (Tumor_size is T3) and (Lymph_node is level3) and (Hormone_receptor is HER2_positive) then (Cancer_risk
28. If (Tumor_size is T4) and (Lymph_node is Level1) and (Hormone_receptor is ER_positive) then (Cancer_risk i
29. If (Tumor_size is T4) and (Lymph_node is Level1) and (Hormone_receptor is PR_positive) then (Cancer_risk i
30. If (Tumor_size is T4) and (Lymph_node is Level1) and (Hormone_receptor is HER2_positive) then (Cancer_ris
```

```
derate_Risk) (1)
High_Risk) (1)
s Moderate_Risk) (1)
```

Figure 3: Expert rules in the database.

Figure 3 shows the rule base snapshot for both prostate and breast cancer defining how fuzzy inputs are mapped to fuzzy outputs. This shows how the systems is making decisions by providing more human-like reasoning.

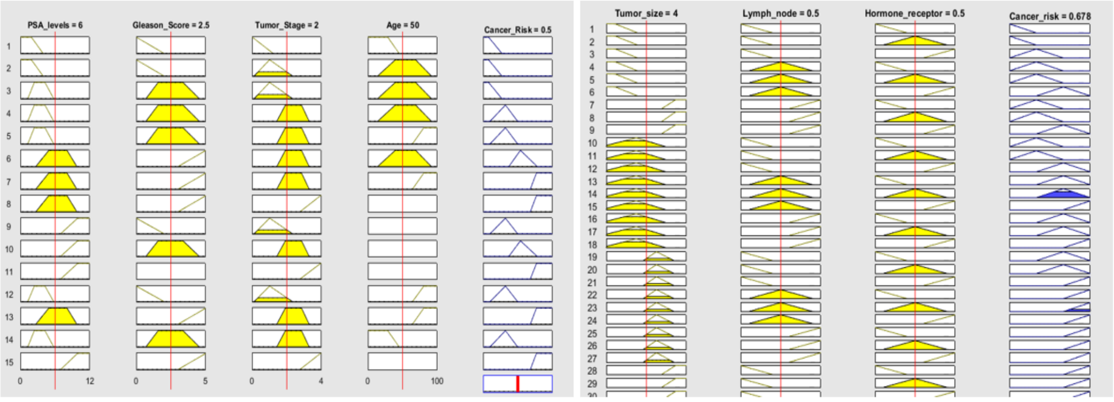


Figure 4. Rule viewer

The Rule Viewer shown in Figure 4 was used in MATLAB for cancer prediction to interpret, visualize, and tune fuzzy logic rules that handle the inherent uncertainty in medical data. It aided in making the predictive model more transparent which is crucial for clinical applications and decision-making.

Confusion matrix for breast cancer in Table 3 was used to compute Accuracy, precision, sensitivity, specificity, and the F1 score using a prostate cancer dataset of 1,000 patients.

Table 3. *Confusion matrix breast cancer.*

	Predicted positive	**Predicted negative**
Actual Positive	True positive (TP)- **450**	False Negative (FN)-**80**
Actual negative	False positive (FP)-**70**	True Negative (TN)-**400**

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP + TN}}{\boldsymbol{TP + TN + FP + FN}} = \frac{\mathbf{450 + 400}}{\mathbf{450 + 400 + 70 + 80}} = \mathbf{0.85}.$$

$$Precision = \frac{TP}{TP + FP} = \frac{450}{450 + 70} = \mathbf{0.865}$$

$$Sensitivity = \frac{TP}{TP + FN} = \frac{450}{450 + 80} = \mathbf{0.849}$$

$$Specificity = \frac{TN}{TN + FP} = \frac{400}{400 + 70} = \mathbf{0.851}$$

$$F1\ Score = 2\ x\ \frac{Precision\ x\ Sensitivity}{Precision + sensitivity} = 2\ x\ \frac{0.865\ x\ 0.849}{0.865 + 0.849} = \mathbf{0.856}$$

Table 4. *Confusion matrix prostate cancer.*

	Predicted positive	**Predicted negative**
Actual Positive	True positive (TP)- 300	False Negative (FN)-5**0**
Actual negative	False positive (FP)-100	True Negative (TN)-550

The matrix for prostate cancer gave an Accuracy of 85%, Precision of 75%, Sensitivity/Recall of 85.7%, Specificity of 84.6% and F1 Score of 79.9% The F1 Score balances precision and recall, indicating the model’s ability to correctly predict cancer while minimizing false positives and false negatives.

The 3D surface plots provide insights into how different clinical variables interacted to impact cancer risk. Figure 5 shows a sample analysis based on interactions between key variables.

Patients with a Gleason score greater than 4 and PSA levels above 5 ng/mL had a cancer risk, significantly higher than those with lower Gleason scores and PSA levels. Patients aged above 65, a PSA level above 6 ng/mL result in a higher cancer risk, compared to patients under 50, where the cancer risk was approximately less for the same PSA level. The model indicated a sharp increase in cancer risk after age 60, with a rise in risk for patients with elevated PSA.

Tumor larger than 5 cm in size combined with lymph node involvement are associated with a cancer risk. For tumor smaller than 2 cm, the risk was significantly lower, regardless of lymph node involvement. Hormone receptor-positive tumor larger than 4 cm had a high cancer risk probability. In hormone receptor-negative cases, the risk was lower for the same tumor size.

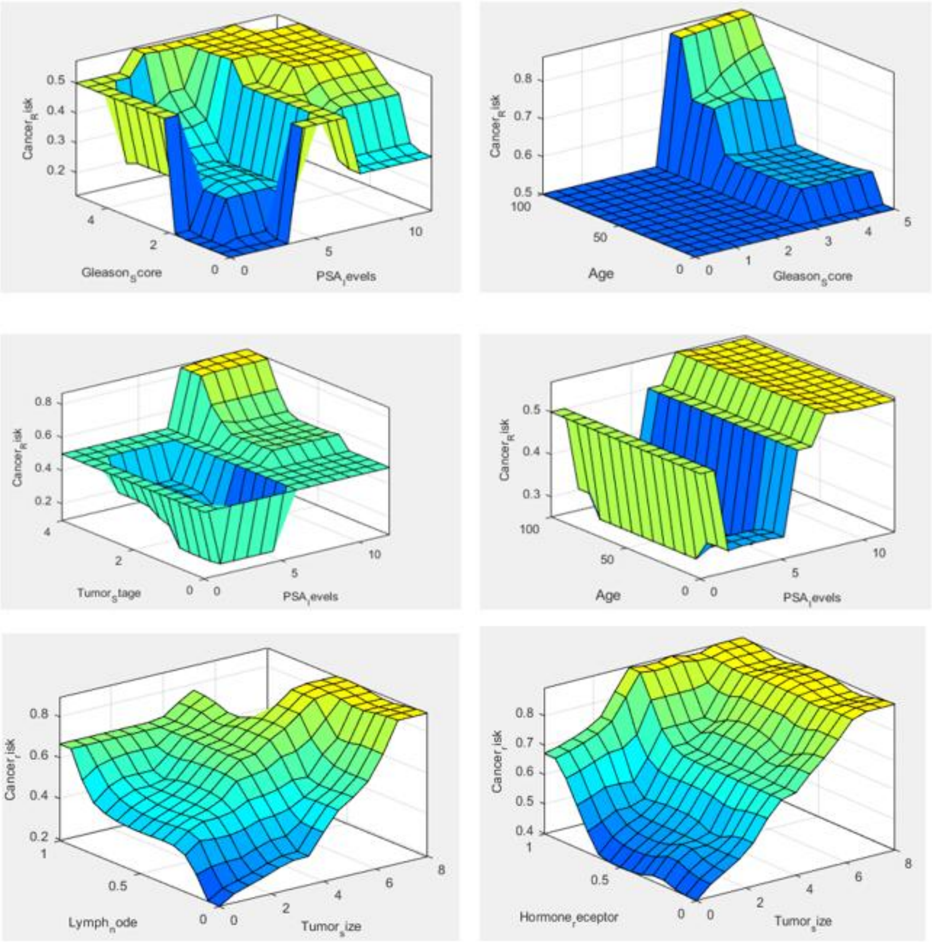


Figure 5:Cancer prediction surface plots.

The Model was compared with Baseline Algorithms to validate the robustness of the developed breast cancer prediction model using the same dataset.

Table 5: Breast cancer baseline algorithms comparison.

Algorithm	**Accuracy**	**F1-Score**
Fuzzy Logic	85%	85.6%
Logistic Regression	81.3%	83.5%
Support Vector Machine (SVM)	82.7%	84.1%
K-Nearest Neighbors (KNN)	80.5%	81.2%

The analysis of Table 5 indicates that Fuzzy Logic model outperforms the other algorithms in both accuracy and F1-Score, making it the best-suited model for cancer prediction in this scenario. Its ability to handle uncertainty and imprecision in the data makes it particularly valuable in medical applications. While SVM and Logistic Regression also offer competitive performance, KNN appears to be the least effective for this task, likely due to its sensitivity to data variability. To ensure model's generalizability, 10-fold cross-validation was performed and the model maintained consistent performance across all folds: Average Accuracy: 84.5%, Standard Deviation: 1.2%, Cross-Validation F1-Score: 85.1%. This shows that the model is stable and less likely to overfit to the training data.

Table 6. Prostate cancer baseline algorithms comparison.

Algorithm	Accuracy	F1-Score
Fuzzy Logic	85%	80%
Logistic Regression	84%	80%
Support Vector Machine (SVM)	82.7%	84.1%
K-Nearest Neighbors (KNN)	83%	81%

Fuzzy Logic Model performs better comparably to traditional machine learning models, as shown in Table 6 showing robustness in its predictions. The model has competitive performance with good accuracy of 85% and an F1 Score of 80%. To ensure model's generalizability, 10-fold cross-validation was performed and the model maintained consistent performance across all folds with standard deviation 0.03% indicates a slightly higher variance in performance compared to Logistic Regression and Random Forest.

7. Business Benefits

The fuzzy logic-based cancer prediction project offers several significant business benefits, particularly for healthcare providers, technology companies, and stakeholders in the medical AI industry. Integrating fuzzy logic into clinical decision support systems (CDSS), healthcare providers can significantly improve diagnostic accuracy and efficiency. This reduces the time and costs associated with misdiagnosis and unnecessary tests, leading to streamlined patient care pathways and better resource management in hospitals and clinics. Accurate early diagnosis and personalized prognosis prediction enable tailored treatment plans, which can improve patient outcomes and reduce recurrence rates. The project’s models will help identify cancer at earlier, more treatable stages translating to significant cost savings for healthcare systems and insurance companies, while also lowering out-of-pocket expenses for patients. The integration of fuzzy logic in cancer prediction will enhance clinical decision-making process by offering transparent and interpretable outcomes. The model can generate valuable data insights that can inform strategic decisions, such as identifying trends in cancer diagnosis and treatment efficacy. These insights can be used by healthcare providers, pharmaceutical companies, and insurers to optimize their services, improve drug development, and develop targeted healthcare policies.

8. Conclusions

The fuzzy logic-based cancer prediction model demonstrates an advancement in enhancing diagnostic and prognostic capabilities for breast and prostate cancers. By effectively handling complexities and uncertainties inherent in medical data, the project’s models offer an accurate, interpretable, and adaptable approach to cancer risk assessment. Integrating these fuzzy logic systems into clinical decision support tools can allows for real-time, rule-based recommendations that align with clinical reasoning, improving early diagnosis, patient outcomes, and healthcare efficiency. The project not only addresses critical gaps in traditional cancer prediction methods but also provides scalable, data-driven solutions that benefit both clinicians and patients. Its successful implementation holds great potential to transform oncology care, setting a new standard in the integration of fuzzy logic in medical practice. However, the model’s deployment in clinical settings must be guided by rigorous validation, continuous refinement, and adherence to ethical standards to ensure reliability, and safety of patients. Future work should focus on expanding datasets, hybridization, incorporating diverse populations, and integrating the model into clinical workflows to maximize its impact on patient care and outcomes.

Acknowledgment

The authors acknowledge Technical University of Mombasa for the invaluable support and contributions to this project.

Declaration of use of content generated by Artificial Intelligence (AI) (including but not limited to Generative-AI) in the paper

The authors confirm that there has been no use of content generated by Artificial Intelligence (AI) (including but not limited to text, figures, images, and code) in the paper entitled "Fuzzy Logic Predictive Model for Early Diagnosis and Prognosis of Breast and Prostate Cancer".

References

[1] T. Khan *et al.*, “Understanding of PSA biology, factors affecting PSA detection, challenges, various biomarkers, methods, and future perspective of prostate cancer detection and diagnosis,” *Adv. Cancer Biol. - Metastasis*, vol. 5, no. June, p. 100059, 2022, doi: 10.1016/j.adcanc.2022.100059.

[2] M. Soleimani, F. Campean, and D. Neagu, “Diagnostics and prognostics for complex systems: A review of methods and challenges,” *Qual. Reliab. Eng. Int.*, vol. 37, no. 8, pp. 3746–3778, 2021, doi: 10.1002/qre.2947.

[3] C. Shang and D. Xu, “Epidemiology of Breast Cancer,” *Oncologie*, vol. 24, no. 4, pp. 649–663, 2022, doi: 10.32604/oncologie.2022.027640.

[4] D. Withrow *et al.*, “Current and projected number of years of life lost due to prostate cancer: A global study,” *Prostate*, vol. 82, no. 11, pp. 1088–1097, 2022, doi: 10.1002/pros.24360.

[5] J. M. Seely and T. Alhassan, “Screening for breast cancer in 2018—what should we be doing today?,” *Curr. Oncol.*, vol. 25, no. June, pp. S115–S124, 2018, doi: 10.3747/co.25.3770.

[6] A. B. Galosi *et al.*, “Detection limits of significant prostate cancer using multiparametric MR and digital rectal examination in men with low serum PSA: Up-date of the Italian Society of Integrated Diagnostic in Urology,” *Arch. Ital. di Urol. e Androl.*, vol. 93, no. 1, pp. 92–100, 2021, doi: 10.4081/AIUA.2021.1.92.

[7] S. Łukasiewicz, M. Czeczelewski, A. Forma, J. Baj, R. Sitarz, and A. Stanislawek, “Łukasiewicz.Pdf,” *Cancers (Basel).*, vol. 13, p. 4287, 2021.

[8] R. Boadh *et al.*, “Study and prediction of prostate cancer using fuzzy inference system,” *Mater. Today Proc.*, vol. 56, no. January, pp. 157–164, 2022, doi: 10.1016/j.matpr.2022.01.040.

[9] V. R. Mudunuru and L. A. Skrzypek, “A comparison of artificial neural network and decision trees with logistic regression as classification models for breast cancer survival,” *Int. J. Math. Eng. Manag. Sci.*, vol. 5, no. 6, pp. 1170–1190, 2020, doi: 10.33889/IJMEMS.2020.5.6.089.

[10] M. Johnson, A. Albizri, and S. Simsek, “Artificial intelligence in healthcare operations to enhance treatment outcomes: a framework to predict lung cancer prognosis,” *Ann. Oper. Res.*, vol. 308, no. 1–2, pp. 275–305, 2022, doi: 10.1007/s10479-020-03872-6.

[11] O. Castillo and P. Melin, “Type-2 fuzzy logic systems,” *SpringerBriefs Appl. Sci. Technol.*, vol. 1, no. 6, pp. 7–12, 2012, doi: 10.1007/978-3-642-28956-9_2.

[12] X. Zhang, M. Wang, B. Bedregal, M. Li, and R. Liang, “Semi-overlap functions and novel fuzzy reasoning algorithms with applications,” *Inf. Sci. (Ny).*, vol. 614, pp. 104–122, 2022, doi: 10.1016/j.ins.2022.10.017.

[13] G. Casalino, R. Grassi, M. Iannotta, V. Pasquadibisceglie, and G. Zaza, “A Hierarchical Fuzzy System for Risk Assessment of Cardiovascular Disease,” *IEEE Conf. Evol. Adapt. Intell. Syst.*, vol. 2020-May, no. June, 2020, doi: 10.1109/EAIS48028.2020.9122750.

[14] K. Saxena and U. Banodha, “A Fuzzy Logic Based Cardiovascular Disease Risk Level Prediction System in Correlation to Diabetes and Smoking,” *Adv. Intell. Syst. Comput.*, vol. 1042, no. Cvd, pp. 29–40, 2020, doi: 10.1007/978-981-32-9949-8_3.

[15] A. A. Sadat Asl, M. M. Ershadi, S. Sotudian, X. Li, and S. Dick, “Fuzzy expert systems for prediction of ICU admission in patients with COVID-19,” *Intell. Decis. Technol.*, vol. 16, no. 1, pp. 159–168, 2022, doi: 10.3233/IDT-200220.

[16] F. Li, C. Shang, Y. Li, and Q. Shen, “Interpretable mammographic mass classification with fuzzy interpolative reasoning,” *Knowledge-Based Syst.*, vol. 191, 2020, doi: 10.1016/j.knosys.2019.105279.

[17] L. Souza-Pereira, N. Pombo, S. Ouhbi, V. Felizardo, and N. Garcia, “Clinical decision support systems for chronic diseases: A Systematic literature review,” *Comput. Methods Programs Biomed.*, vol. 195, 2020, doi: 10.1016/j.cmpb.2020.105565.

[18] L. A. Zadeh, “Fuzzy Sets,” *Inf. Control*, no. 8, pp. 338–353, 1965, doi:

10.1017/CBO9781107415324.004.
[19] https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic
[20] https://www.kaggle.com/datasets/sajidsaifi/prostate-cancer