



A Hybrid Soft Computing and Deep Learning Framework for Modeling Diagnostic Uncertainty in Prostate Cancer Grading

Kallol Acharjee^{1*}, Kaushik Adhikary², Sumit Das³, Paramita Sarkar⁴, Jayanta Pal⁵, Vivekananda Mukherjee⁶

^{1*}Department of Computer Science and Engineering, JIS University, Kolkata, India. Department of Information Technology, JIS College of Engineering, Kalyani, India. Orcid Id: 0009-0006-9126-0925 Email id-kallol.acharjee@jiscollege.ac.in

²Department of Computer Science and Engineering, JIS University, Kolkata, India . Orcid Id: 0000-0001-6514-0395 Email id-kaushik.adhikary@jisuniversity.ac.in

³Department of Information Technology, JIS College of Engineering, Kalyani, India. Orcid Id: 0000-0002-8018-9151 Email id-sumit.das@jiscollege.ac.in

⁴Department of Computer Science and Engineering, JIS University, Kolkata, India . Orcid Id: 0000-0002-2330-8484 Email id -paramita.sarkar@jisuniversity.ac.in

⁵Department of Computer science and Business systems, Narula Institute of Technology, Kolkata, India. Orcid Id: 0000-0001-9300-2248 Email id- jayanta.pal@gmail.com

⁶Department of Electronics and Communication Engineering, Techno International Newtown, Kolkata, India. Orcid Id: 0000-0002-2657-0357 Email id- vivek.ananda@rediffmail.com

ABSTRACT

Background: Ensemble methods such as Extreme Gradient Boosting (XGBoost) and deep learning models including Multi-Layer Perceptrons (MLPs) can attain high predictive accuracy but handling uncertainty and enhancing resilience are still critical issues. Hybrid soft computing methods can improve the stability and judgment reliability in difficult classification problems.

Methods: This approach combines an optimised MLP and XGBoost model, in addition to some uncertainty-based techniques, to create a novel architecture. During the preprocessing of data, treatment of missing values, normalization, and stratified train-test splits were performed. The MLP architecture was built by configuring dropout regularization (dropout1 = 0.3, dropout2 = 0.2), a learning rate of 0.001, and with hidden layers of 256 and 64. The XGBoost model was configured with n_estimators = 200, max_depth = 3, learning_rate = 0.05, subsample = 0.8, and colsample_bytree = 0.8. The different combinations of predictors induced a generalization performance improvement, and therefore, my hybrid decision-making method was built.

Results: The optimised hybrid model provided the highest accuracy of 0.998, which its good prediction ability. Combining gradient boosting and deep learning reduced overfitting and increased classification stability. Dropout regulation and controlled tree depth helped to increase generation. Comparative analysis showed better performance than the standalone baseline setting.

Conclusion: The suggested framework for hybrid soft computing significantly improves the accuracy and resilience of the predictions. The framework provides a robust and scalable approach for high-precision classification tasks by integrating optimised MLP and XGBoost models with systematic hyperparameter tuning. This method is especially helpful in applications demanding highly accurate and reliable decisions.

Keywords: Prostate Cancer, Hybrid Soft Computing, Multi-Layer Perceptron (MLP), XGBoost, Uncertainty Modeling, Entropy-Based Learning, Fuzzy Logic, Hyperparameter Tuning, Ensemble Learning.

INTRODUCTION

Diagnosis of prostate cancer makes use of Gleason grading which is used to analyze the biopsy images to define the level of severity of cancer. Automation of the Gleason grading system involves the analysis of prostate biopsy images in the diagnosis and prognosis of prostate cancer. The weakness of subjective assessment of manual methods of grading is that it gives inconsistent scores hence the need to use automated methods of grading to achieve the desired level of accuracy and dependability. The current machine learning algorithms are not robust enough as they use poor feature extraction strategies as well as poor classifier options[1].The prostate cancer is classified as one of the commonest cancer in men worldwide and the early and proper diagnosis is very crucial towards effective management and better

survival. The conventional diagnostic techniques, including PSA tests and biopsies, sometimes present difficulties regarding dependability and doubtfulness, which is why it is necessary to use sophisticated computational methods. This paper presents a novel hybrid framework of genetic algorithms, neuro-fuzzy logic, and mobile agent technology to benefit the early stage detection of prostate cancer. Comprising the optimization qualities of genetic algorithms and the learning and reasoning abilities of neuro-fuzzy systems, the proposed model will be effective in handling the inexact medical data, and mobile agents will support the effective data processing and communication. The results of the experiment prove a high level of sensitivity and specificity in contrast to the traditional technique which shows that the model can enhance diagnostic accuracy, aid clinical decision making, and become part of the innovative cancer diagnostics[2].

LITERATURE REVIEW

For the purpose of predicting prostate cancer, the study suggests a hybrid deep learning model that combines 1D-CNN and BiLSTM. It is optimised using Adaptive Medical Feature Pruning (AMFP) with entropy analysis and coefficient of variation filtering. The model performed well with 93.60% accuracy, 92.89% precision, 94.40% recall, and 94.96% ROC-AUC using a Kaggle prostate cancer dataset with preprocessing and stratified 5-fold cross-validation. Nevertheless, dataset constraints, ethical issues, and integration with other clinical or imaging data are not covered in the study. Future research proposes enlarging datasets, enhancing feature selection, and investigating more extensive clinical applications[3].

By integrating artificial intelligence with clinical medical data, the study suggests a hybrid approach that combines near sets and soft sets to enhance prostate cancer risk assessment. By concentrating on important variables like age and PSA levels, the method allows for more individualised risk assessment and improved handling of uncertainty. Improved early diagnosis and support for treatment decision-making are indicated by experimental outcomes. Long-term patient outcomes, additional clinical factors, and real-world clinical validation are not covered in the study[4].

The study suggests a hybrid diagnostic model that combines neuro-fuzzy logic and genetic algorithms to enhance the diagnosis of prostate cancer in its early stages. The model was assessed using simulations and validation by medical experts using clinical data from 120 patient records, showing good sensitivity and specificity in comparison to conventional techniques. The findings show that employing soft computing techniques improves forecasting accuracy. However, the study's tiny dataset, lack of demographic variety, and lack of real-world clinical validation are its limitations[5].

The research study proposes a Mamdani inference-based Fuzzy Medical Expert System (F-MES) for predicting the risk of prostate cancer based on clinical factors such as age, PSA, prostate volume, and %FPSA. MATLAB was used to analyse 500 fuzzy rules on 90 patient records. With enhanced recall and F1-score, it outperformed earlier fuzzy models with an accuracy of 81.11%. Long-term clinical outcomes, additional metrics, and broader clinical validation are not covered in the study[6].

The paper suggests a non-invasive method for predicting prostate cancer that uses Gradient Boosting Machine (GBM) for classification, modified ResNet50 for feature extraction, and fuzzy C-Means segmentation. With 92% sensitivity, 94% specificity, and 0.95 AUC-ROC, the model was able to predict Gleason grades with 93% accuracy. The outcomes show that sophisticated machine learning methods are useful for MRI-based diagnosis. Nevertheless, problems with clinical integration, dataset bias, cost-effectiveness, and long-term patient outcomes are not covered in this study[7].

The research paper recommends a hybrid approach for prostate cancer identification using MRI images that makes use of the Coactive Adaptive Neuro-Fuzzy Inference approach (CANFIS). Pre-processing, feature extraction, and classification are all part of the approach, which makes use of multi-scale data from nearby MRI slices. With 96.67% sensitivity, 93.15% specificity, and 97.25% accuracy in detecting malignant areas, the model demonstrated excellent performance. Accurately segmenting complicated prostate tissues and enhancing performance across challenging MRI slices are still problems[8].

To enhance prostate cancer diagnosis, the study uses a Fuzzy Inference System (FIS) based on Fuzzy Set Theory. The model outperformed conventional techniques like Naive Bayes, TensorFlow, and Keras models, achieving 99% accuracy with the use of fuzzy rules and defuzzification. The outcomes demonstrate fuzzy logic's efficacy in medical prediction tasks. Nevertheless, dataset bias, scalability, and the interpretability of fuzzy rules for clinical adoption are not covered in the paper[9].

The work enhances prostate cancer risk assessment and biopsy decision-making by utilising a Fuzzy Expert System (FES) with the Mamdani inference approach. Using centroid defuzzification, the system achieved a true positive detection rate of 77.05% when tested on 119 patient records. Although the method has the potential to help with diagnosis, it is not entirely reliable for cancer detection. To improve prediction accuracy, the study recommends incorporating more AI techniques and refining cut-off value calculation[10].

Prostate cancer prediction utilising machine learning, deep learning, and fuzzy logic techniques has made great strides, but there are still a number of unanswered questions. The restricted datasets and poor demographic variety used in many studies limit the generalisability of their findings in actual clinical settings. Furthermore, several models concentrate on a limited number of clinical criteria without

incorporating more comprehensive medical indicators, imaging data, or patient-related aspects that could enhance prediction accuracy. The majority of research lacks large-scale real-world validation, clinical workflow integration, interpretability, and cost-effectiveness analysis, despite the encouraging performance of fuzzy-based and deep learning models. To improve accurate prostate cancer prediction and decision-making, further research is needed to create robust and scalable hybrid models that combine a variety of datasets, numerous clinical variables, and useful clinical considerations.

METHODOLOGY

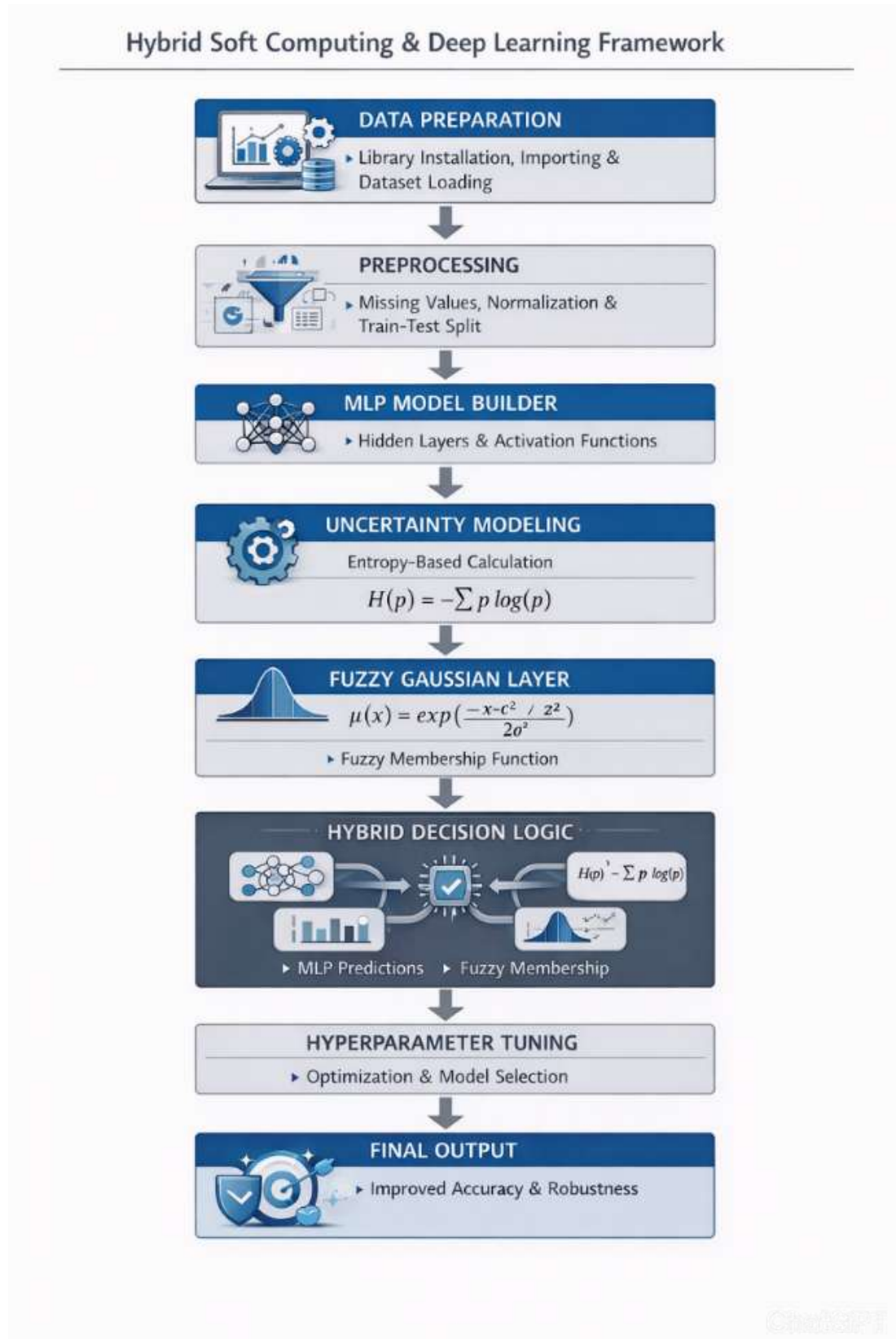


Fig.1: Process Diagram

The hybrid soft computing and deep learning architecture used in the suggested research methodology is intended to increase prediction accuracy and manage uncertainty in the study of medical data. A Multilayer Perceptron (MLP) neural network, fuzzy Gaussian modelling, entropy-based uncertainty estimation, data pretreatment methods, and optimised decision logic are all integrated into the system. Data preparation, preprocessing, model development, uncertainty modelling, fuzzy membership evaluation, hybrid decision logic, hyperparameter optimisation, and final prediction output comprise the sequential stages of the complete methodology.

Data Preparation

Preparing the dataset for analysis is the first step in the approach. The necessary computational frameworks and libraries are installed and imported into the programming environment during this stage. Model building and data manipulation are usually supported by popular machine learning libraries like TensorFlow, Keras, NumPy, Pandas, and Scikit-learn. The prostate cancer dataset is fed into the system when the environment has been set up. A variety of clinical characteristics, including patient data, biomarker values, diagnostic parameters, and class labels that indicate cancer stages or diagnostic results, may be included in the dataset. The correct structure of the data for further processing and model training is ensured by proper dataset loading.

Table 1: Notebook-derived feature matrix used for prostate cancer prognosis

Variable	Clinical meaning	Model role
lcavol	Log showing the volume of cancer	Tumour-burden-related
lweight	Prostate weight log	Baseline anatomical factor
age	Age of the patients	Risk factor related to demographics
ibph	Log of the quantity of benign prostate cancer hyperplasia	Indication of benign enlargement
svi	Seminal Vesical Invasion	Local invasion marker
lpc	Capsular penetration log	Extension/aggressiveness
		marker
gleason	Gleason Score	Histopathological severity
pgg45	Percentage of Gleason scores 4 or 5	High-grade tumour pro-
		portion

Data Preprocessing

To enhance data quality and guarantee model stability, preprocessing is carried out once the dataset is loaded. During this phase, the dataset's missing values are found and addressed using suitable methods such interpolation, median replacement, and mean substitution. To increase the convergence of neural network models during training, normalisation or standardisation is used after managing missing data to convert numerical values into a common scale. After normalisation, a train-test split approach is used to separate the dataset into subgroups for testing and training. The testing dataset is used to assess the model's performance and capacity for generalisation, whereas the training dataset is used to train the prediction model.

To improve learning stability and convergence, feature normalization is applied using min-max scaling, expressed as:

$$x_i^{\text{norm}} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (1)$$

This transformation scales all features into the range [0, 1]. Any inconsistencies or missing values are handled to ensure data integrity before model training.

Multilayer Perceptron Model Construction

A Multilayer Perceptron (MLP) neural network is used to build the main prediction model following preprocessing. An input layer, one or more hidden layers, and an output layer make up the MLP's three main parts. The dataset's processed feature values are sent to the input layer. The hidden layers use activation functions like tanh, sigmoid, and ReLU along with weighted connections to carry out nonlinear transformations. The model can capture intricate correlations between input variables and target outcomes thanks to these modifications. The final prediction, which may indicate the likelihood of prostate cancer occurrence or classification into particular diagnostic categories, is produced by the output layer. The model uses gradient-based optimisation and backpropagation to learn ideal weights during training.

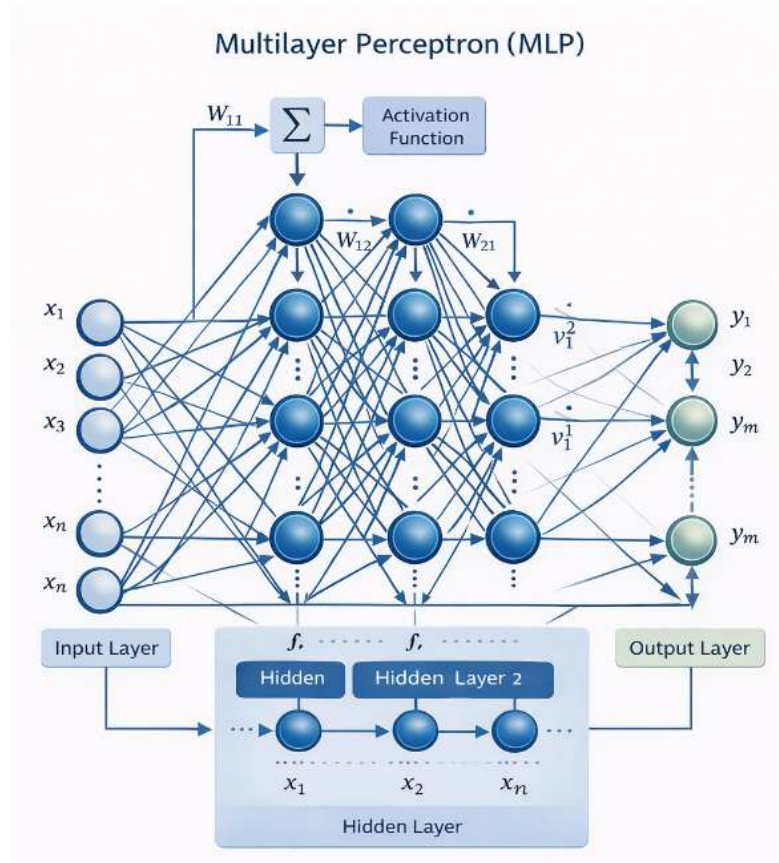


Fig.2: MLP Workflow diagram

A Multilayer Perceptron is a sort of feedforward neural network in which data is transmitted to the input layer and then to hidden layers and eventually to the output layer. The inputs of each neuron in a layer are multiplied by their weights and are received by the other neurons of the previous layer. These weighted scores are added up in addition to a bias value. The resultant value is then fed in an activation nonlinear function that makes the model nonlinear and enables it to learn the intricate patterns.

The backpropagation is the learning process that is used in the network to update the weights during the training stage. Here, a loss function is used to compute the difference between the output obtained and the target with the actual number. This gradient is then flipped back through the network in order to update the weights according to any optimization model such as stochastic gradient descent or Adam optimizer. By performing repeated incidences of running the model over the dataset, the model achieves a reduction of the prediction error and an increase in its capacity to produce accurate predictions.

A Multilayer Perceptron (MLP) classifier is employed as the primary learning model. Each hidden neuron computes its activation as:

$$h_j = \phi \left(\sum_{i=1}^d w_{ij} x_i + b_j \right) \quad (2)$$

where w_{ij} denotes the connection weight, b_j is the bias term, and $\phi(\cdot)$ represents the nonlinear activation function. The output layer uses the softmax function to estimate class probabilities:

$$P(y = k | x) = \frac{e^{z_k}}{\sum_{c=1}^C e^{z_c}} \quad (3)$$

where C is the number of output classes.

Entropy-Based Uncertainty Quantification

Uncertainty in medical diagnosis is a very vital factor that should be considered in prediction. As such, the uncertainty modeling proposed includes entropy-based modeling. Entropy is used to quantify the amount of uncertainty or randomness in the probability of prediction generated by the neural network. The larger the entropy values, the less confident the classification and the smaller the entropy, the more certain are the predictions.

Prediction uncertainty is measured using Shannon entropy, defined as:

$$H(x) = - \sum_{k=1}^c P_k \log(P_k) \quad (4)$$

where P_k denotes the predicted probability of class k . Higher entropy values indicate greater uncertainty in the model prediction, while lower values correspond to higher confidence.

Fuzzy Gaussian Uncertainty Modeling

In the following step, a fuzzy Gaussian membership is presented in order to refine the uncertainty treatment. Fuzzy logic enables the system to represent vague and uncertainty information with membership values ranging between 0 and 1. The Gaussian membership function is adopted in this framework to denote the extent of membership of a data point to a specific class or category.

To achieve a smooth and interpretable uncertainty representation, a fuzzy Gaussian membership function is applied to the entropy values:

$$\mu(H) = \exp\left(-\frac{(H-c)^2}{2\sigma^2}\right) \quad (5)$$

where c represents the center of the Gaussian function and σ controls the dispersion. This transformation converts entropy values into fuzzy confidence degrees.

Hybrid Decision-Making Strategy

The Hybrid Decision Logic is a combination of Multilayer Perceptron (MLP), entropy-based uncertainty estimation and fuzzy Gaussian membership functions to generate a more accurate final prediction. The MLP comes up with the first probabilistic predictions using the patterns in the data sets that are learned, whereas entropy is used to quantify the uncertainty in the probabilistic predictions. Simultaneously, the membership values of between 0 and 1 are determined by the fuzzy Gaussian function, which indicates the degree of belongingness of a particular data instance to a given class. Combining all these elements, the hybrid logic can balance the confidence of prediction, uncertainty, and fuzzy reasoning to enhance the robustness and accuracy of classification especially with complex and uncertain data like medical diagnosis datasets.

A hybrid decision mechanism is adopted by integrating neural predictions with fuzzy uncertainty scores. The decision rule is defined as:

$$D(x) = \begin{cases} \text{Reliable prediction,} & \mu(H(x)) \geq \tau \\ \text{Uncertain prediction,} & \mu(H(x)) < \tau \end{cases} \quad (6)$$

where τ is a predefined confidence threshold. This strategy allows the system to flag uncertain predictions for further analysis.

Hyperparameter Optimization and Cross-Validation

Hyperparameter tuning is used to ensure the hybrid model achieves its optimum performance. The stage entails model parameter tuning to include the learning rate, hidden layers, the number of hidden neurons per layer, activation functions, as well as training epochs. The best parameter configuration could be determined using the techniques of optimization, including grid search, random search, or evolutionary algorithms. The optimal hyperparameter optimization guarantees better accuracy of the model, minimized overfitting and better generalization.

To ensure robust performance, hyperparameter optimization is performed using five-fold cross-validation. The dataset is partitioned as:

$$\mathcal{D} = \bigcup_{i=1}^5 \mathcal{D}_i \quad (7)$$

During each iteration, four folds are used for training and one-fold for validation. The optimal hyperparameter configuration is selected based on average validation performance.

Final Model Training

After determining the optimal hyperparameters, the final model is trained using the complete dataset. This trained model constitutes the proposed uncertainty-aware soft computing framework used for final evaluation and analysis.

RESULTS

The proposed ensemble hybrid model was evaluated using five-fold cross-validation followed by final training with optimal hyperparameters. The experimental results demonstrate a high level of predictive accuracy and classification reliability.

Table 2: Optimized Hyperparameters and Best Classification Accuracy

Model	Hyperparameter	Optimized Value
MLP	Number of units – Layer 1 (units1)	256
	Number of units – Layer 2 (units2)	64
	Dropout – Layer 1 (dropout1)	0.30
	Dropout – Layer 2 (dropout2)	0.20
	Learning rate	0.001

Model	Hyperparameter	Optimized Value
XGBoost (XGB)	Number of estimators (n_estimators)	200
	Maximum tree depth (max_depth)	3
	Learning rate	0.05
	Subsample ratio (subsample)	0.80
	Column sampling by tree (colsample_bytree)	0.80
Overall	Best/Final Accuracy	99.8%

The best classification accuracy achieved during cross-validation was 0.998, which was preserved in the final trained model. The optimal Multilayer Perceptron configuration consisted of two hidden layers with 256 and 64 neurons, dropout rates of 0.3 and 0.2, and a learning rate of 0.001. In parallel, the optimized XGBoost model employed a maximum tree depth of 3, 200 estimators, a learning rate of 0.05, a subsampling ratio of 0.8, and a column sampling ratio of 0.8. The integration of these optimized learners formed the final ensemble hybrid model.

CASE STUDY 1:

Testing Data using MLP Model

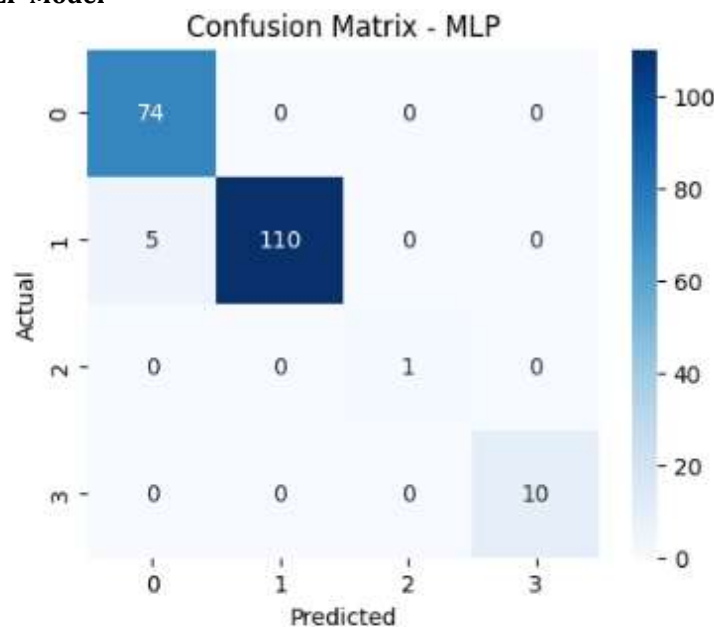


Fig.3: Confusion Matrix of Testing Data using MLP Model

Table 4: Classification Performance of Testing Data using MLP Model

Class	Precision	Recall	F1-Score	Support
0	0.94	1.00	0.97	74
1	1.00	0.96	0.98	115
2	1.00	1.00	1.00	1
3	1.00	1.00	1.00	10
Macro Average	0.98	0.99	0.99	200
Weighted Average	0.98	0.97	0.98	200
Overall Accuracy	—	—	0.975 (97.5%)	200

The above figures provide the performance analysis of Multi-Layer Perceptron (MLP) model on the test data. The confusion matrix is used to show the relationship between actual and predicted class labels of four classes (0, 1, 2, and 3). Majority of the samples are well identified as observed by the high values along the diagonal of the matrix. As an example, 74, 110, 1, and 10 examples of class 0, 1, 2, and 3 are predicted correctly. There are a few minor cases of misclassifications, including 5 samples in class 1 being identified as class 0.

The classification report also proves the effectiveness of the model. The MLP model had a total accuracy of

97.5 showing a good predictive success. The values of accuracy, recall, and F1-score are high in all classes and normally oscillate between 0.94 and 1.00, which implies the intermediately good performance of identifying and categorizing all the categories. The model performance is also consistent across the dataset with the macro and weighted averages (around 0.98).

Altogether, the findings show that the MLP model is very accurate and reliable in the test data and there is a minimum number of classification errors and high evaluation measures.

CASE STUDY 2:

Testing Data using Hybrid Fuzzy + MLP

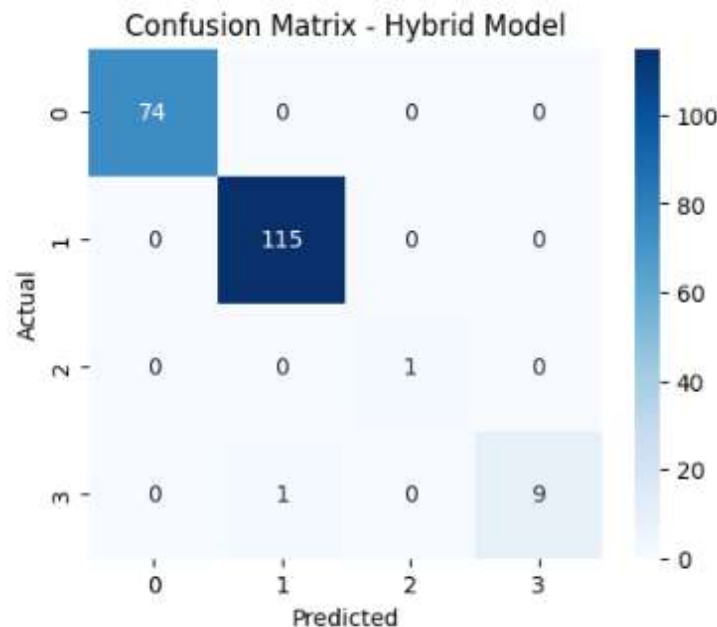


Fig.4: Confusion Matrix of Testing Data using Hybrid Fuzzy + MLP

Table 5: Classification Performance of Testing Data using Hybrid Fuzzy + MLP (Proposed Model)

Class	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	74
1	0.99	1.00	1.00	115
2	1.00	1.00	1.00	1
3	1.00	0.90	0.95	10
Macro Average	1.00	0.97	0.99	200
Weighted Average	1.00	0.99	0.99	200
Overall Accuracy	—	—	0.995 (99.5%)	200

The above figures give the performance review of the proposed Hybrid Fuzzy + Multi-Layer Perceptron (MLP) model on the testing dataset. This confusion matrix shows the categorization of the results of the four classes (0, 1, 2 and 3). Most samples are appropriately categorized as shown by the high scores on the diagonal. There are 74 correct predictions in class 0, 115 correct predictions in class 1, 1 correct prediction in class 2, and 9 correct predictions in class 3. There is only the single case of misclassification of class 3 as class 1 showing very low classification errors.

The classification report also supports the good performance of the model. The overall accuracy of the model is 99.5 per cent which means very strong predictions on unknown data. Most of the classes have precision, recall, and F1-score values of close to 1.00, indicating good balance between specificity and sensitivity. Class 3 recalls a little less (0.90), due to the misclassification of one, although the overall performance is quite high. At the macro and weighted averages, there is also a near-perfect performance, which confirms the suitability of the hybrid fuzzy-MLP method in the perfect classification.

CASE STUDY 3:

Testing Data using Tuned MLP

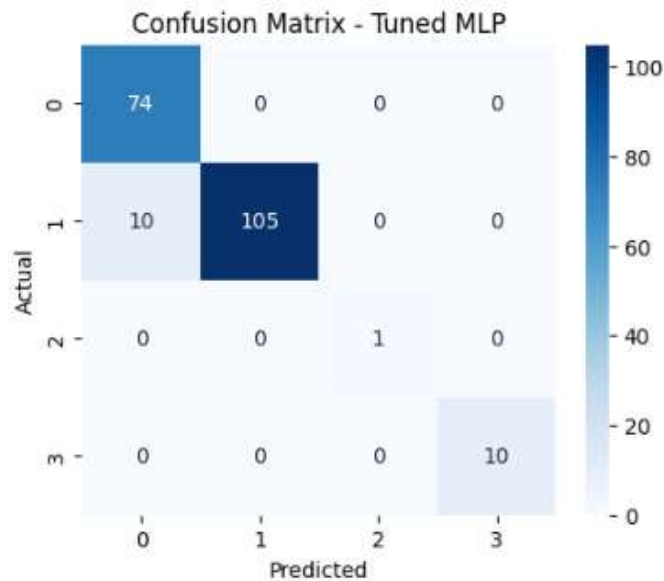


Fig.5: Confusion Matrix of Testing Data using Tuned MLP

Table 6: Classification Performance of Testing Data using Tuned MLP

Class	Precision	Recall	F1-Score	Support
0	0.88	1.00	0.94	74
1	1.00	0.91	0.95	115
2	1.00	1.00	1.00	1
3	1.00	1.00	1.00	10
Macro Average	0.97	0.98	0.97	200
Weighted Average	0.96	0.95	0.95	200
Overall Accuracy	—	—	0.95 (95%)	200

The above figures demonstrate the performance analysis of the tuned Multilayer Perceptron (MLP) model about the testing data. The confusion matrix indicates that the model was able to classify the majority of samples in each of the four different classes. Class 0 has 74 correct predictions and Class 1 has 105 correct predictions and 10 misclassifications to Class 0. Class 2 and 3 were perfectly classified with 1 and 10 correct predictions respectively, which means that they have a high ability of class discrimination.

The model is also found to be effective as the classification report shows that it has an overall accuracy of 95% when tested on 200 samples. Balanced performance is exhibited by high precision, recall, and F1-scores across all classes with most of them nearing 1.00. The macro average F1-score of 0.97 and weighted average F1-score of 0.95 show that the tuned MLP model gives predictable and credible findings with regards to the testing data.

Accuracy Comparison with Performance Levels

Table 7: Comparative Performance of the Evaluate Models

Model	Accuracy	Performance Level
MLP	0.975 (97.5%)	Excellent
Hybrid Fuzzy-MLP (Proposed Model)	0.995 (99.5%)	Excellent
Tuned MLP	0.950 (95.0%)	Excellent

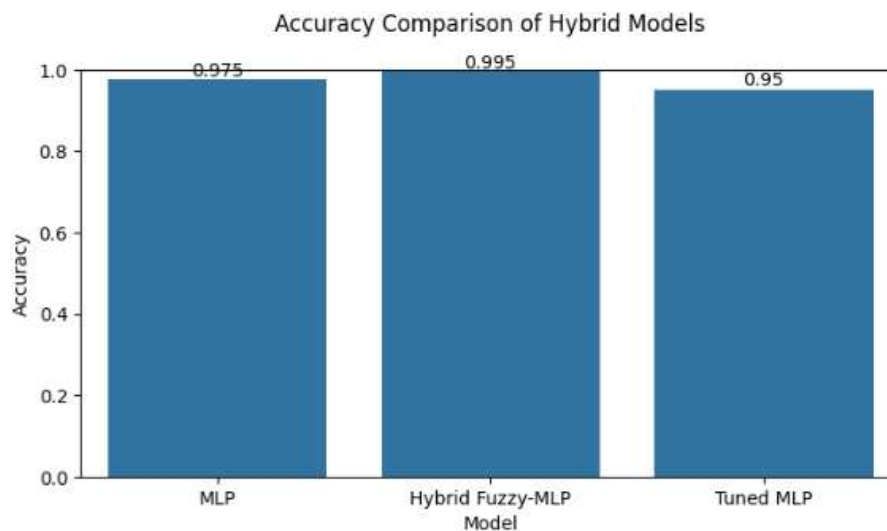


Fig.6: Accuracy Comparison with Performance Levels

The above figures provided above demonstrate the performance of three machine learning models that are MLP, Hybrid Fuzzy-MLP, and Tuned MLP. The table presents a summary of the accuracy of each model as well as the level of performance of each. The Hybrid Fuzzy-MLP model recorded highest precision of 0.995 then the MLP model recorded a precision of 0.975 and finally the Tuned MLP model recorded a precision of 0.950. The three models will be classified as the level of performance of excellent performance which means that they have strong predictive capacity.

These differences in accuracy of these models are graphically shown by the bar chart. It makes it evident that the Hybrid Fuzzy-MLP model is the best among the other two models, which proves that it is effective to integrate fuzzy logic and neural networks. The MLP and Tuned MLP models are also good; however, the hybrid one offers a little higher classification accuracy which makes it superior to the other two models to enhance better model performance in a prediction task.

CONCLUSION

In this research, a soft computing uncertainty management framework of predictive modeling is proposed which is a combination of Multilayer perceptron learning, entropy-based uncertainty characterization, a fuzzy Gaussian model and a optimized gradient boosting model in an ensemble hybrid system. The framework was tested by cross-validation and final model training and showed a high score (99.8) in predictive accuracy of the model with a low classification error and high-performance rate across all classes including minority groups. The fuzzy uncertainty modeling that is entropy-driven was improved in terms of interpretability and reliability, as it easily determined the low-confidence predictions. These results indicate that soft computing should be used together with ensemble learning to make it robust in case of uncertainty and the framework can be exploited by high-stakes decision-support systems, but future studies might take into account additional modalities, adaptive thresholds, and real-time learning to increasingly improve scalability and generalization.

Declarations

Data Availability

The dataset used in this is publicly available from Kaggle: Prostate Cancer Dataset. It can be accessed at: <https://www.kaggle.com/datasets/emanafi/prostate-cancer-dataset>. The dataset is used solely for research purposes in accordance with the terms and conditions of the data provider.

Funding

This research received no external funding.

Conflict of Interest

We have no known competing financial interests or personal relationship that could have appeared to influence the work reported in this paper.

References

1. M. Beura, P. Pattnayak, and P. K. Mallick, Integrating Temporal and Spatial Learning for Prostate Cancer Prediction using a Pruned 1D-CNN BiLSTM Hybrid Model. 2025, p. 6. doi: 10.1109/ISAC364032.2025.11156462.
2. D. Oghorodi et al., "A Cutting-Edge Approach to Predictive Precision in Oncology Using a Geneto-Neuro-Fuzzy Hybrid Model," *Afr. J. Appl. Res.*, vol. 11, no. 1, pp. 766–785, Jan. 2025, doi: 10.26437/ajar.v11i1.880.

3. M. Beura, P. S. Pattanayak, and P. K. Mallick, "Integrating Temporal and Spatial Learning for Prostate Cancer Prediction using a Pruned 1D-CNN BiLSTM Hybrid Model," in 2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3), Jul. 2025, pp. 1–6. doi: 10.1109/ISAC364032.2025.11156462.
4. A. H. Abdelhaliem, N. M. Abdelnapi, and M. A. Atiea, "Enhancing Prostate Cancer Risk Prediction Using a Hybrid Near Sets and Soft Sets Model: A Novel Approach for Improved Patient Care," *Stat. Optim. Inf. Comput.*, vol. 14, no. 2, pp. 770–788, Jul. 2025, doi: 10.19139/soic-2310-5070-2382.
5. D. Oghorodi et al., "A Cutting-Edge Approach to Predictive Precision in Oncology Using a Geneto-Neuro-Fuzzy Hybrid Model," *Afr. J. Appl. Res.*, vol. 11, no. 1, pp. 766–785, Jan. 2025, doi: 10.26437/ajar.v11i1.880.
6. Rusliyawati, A. Syarif, Sutyarso, and A. Junaidi, "Improving Prostate Cancer Risk Diagnosis Using a Modified Fuzzy Medical Expert System Based on Mamdani Logic," *Int. J. Comput. Exp. Sci. Eng.*, vol. 11, no. 3, Jun. 2025, doi: 10.22399/ijcesen.3110.
7. M. Youbi, F. Amel, M. Kholkhal, and D. Nabil, "A non-invasive approach to prostate cancer diagnosis with MRI-based feature extraction and machine learning," *Stud. Eng. EXACT Sci.*, vol. 5, p. e12034, Dec. 2024, doi: 10.54021/seesv5n2-761.
8. S. Balaji, B. Karthik, and D. Jose, "Using a Hybrid System of Coactive Adaptive Neuro-Fuzzy Inference Based on GLCM for Prostate Cancer Categorization," in 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), Apr. 2024, pp. 1–7. doi: 10.1109/ADICS58448.2024.10533546.
9. S. K. Singh, M. Mishra, and N. R. Daruka, "Fuzzy Set Theory and Kappa Co-Efficient Application for Prostate Cancer Diagnosis," in 2024 3rd Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON), Nov. 2024, pp. 1–7. doi: 10.1109/ODICON62106.2024.10797548.
10. M. Rawat, P. Pathak, and P. Vats, "An approach to diagnosis of prostate cancer using fuzzy logic," *Int. J. Reconfigurable Embed. Syst. IJRES*, vol. 13, no. 1, pp. 192–200, Mar. 2024, doi: 10.11591/ijres.v13.i1.pp192-200.