Received: 1 December 2023 | Revised: 7 April 2024 | Accepted: 5 July 2024 | Published online: 24 July 2024

RESEARCH ARTICLE

Artificial Intelligence and Applications
2024, Vol. 2(4) 291–298
DOI: 10.47852/bonviewAIA42022210

BON VIEW PUBLISHING

Towards Unsupervised Learning Driven Intelligence for Prediction of Prostate Cancer

Ejay Nsugbe[1],*

[1]*Nsugbe Research Labs, UK*

Abstract: Prostate cancer is a widespread and global disease which affects adult males – it is said that key causes of the cancer include age, family history, and ethnicity. In this study, the Kaggle prostate cancer dataset, comprising of data from 100 patients with a mixture that both had cancer and did not have cancer, was used alongside machine learning prediction models for the design of unsupervised and automated intelligent systems for the prediction of prostate cancer. Two intelligent systems were designed and underpinned by unsupervised learning algorithms, namely fuzzy c-means and agglomerative hierarchical clustering, where the various intelligent systems were able to make a prostate cancer prediction with accuracies of over 80% for the various classification metrics, alongside being able to predict an associated stage of the prostate cancer. Both designed intelligent systems offer a complimentary alternative to each other, and their relative merits are discussed in the paper.

Keywords: cancer prediction, prostate cancer, unsupervised learning, intelligent system, cybernetics, decision support

1. Introduction

Prostate cancer accounts for a high rate of mortality worldwide; the American Cancer Society, for example, provided a statistic which indicated that over 190,000 men would be diagnosed with prostate cancer in 2020 [1]. The causes of prostate cancer are broad and multifactorial; nonetheless, the consensus reached from dedicated review studies is that there are three key drivers behind the onset of prostate cancer, namely, age, family history, and ethnicity [2–4]. The age of a male has been long viewed as an established driver toward prostate cancer; statistically speaking, males over 70 years old have a 1 in 8 chance of being diagnosed with prostate cancer, while under 40-year-old males are statistically unlikely to be diagnosed with prostate cancer. In terms of family history, for prostate cancer to be fully deemed as a hereditary case, it is said that it needs to have affected either three generations, three first-degree relatives, or two relatives to have been diagnosed prior to the age of 55.[1] It is also said that men with a first-degree relative who has been diagnosed with prostate cancer have up to 3 times likelihood of being diagnosed with prostate cancer when compared with an individual with no family history [5].

Importantly, prostate cancer diagnoses and incidence have been seen to be related to ethnicity, where males of African ethnicity carry the highest number of prostate cancer diagnoses worldwide. In the United States of America – despite a drop in the overall mortality rate since the early '90s – prostate cancer-related deaths among African Americans appear to be over twice the number of Caucasian males [6, 7]. With further investigation into this discrepancy, it has been noted on a genetic scale that different gene expression exists between a case study involving African Americans and Caucasian patients, therein implying racial subtleties in the synthesis and metabolism of androgens [8].

The human prostate is the size of a walnut, and it forms a functional part of both the endocrine and reproductive systems, while one of its functions is the secretion of seminal fluids which are used to nurture and facilitate the transportation of sperm [1, 9]. The atrophy and proliferation of prostate cells are dependent upon androgens (i.e., testosterone) which are effectively steroid hormones that develop the male characteristics via binding to androgen receptors [1, 9, 10]. Androgens are generally produced via negative feedback involving the testicular-hypothalamic-pituitary, which ensures the required levels of androgen are produced in a balanced fashion [1]. The rapid increase in the number of prostate cells gives rise to a prostate-specific antigen (PSA), which is normally contained inside the prostate wall, except during the creation of prostate cancer cells, which ultimately leads to the disruption of cell walls and therein a leakage of PSA [10]. This serves as a biomarker for prostate cancer and serves as a surrogate for the estimation of tumor volumes computationally [10]. An image showing the various stages of development of prostate cancer can be seen in Figure 1.[2]

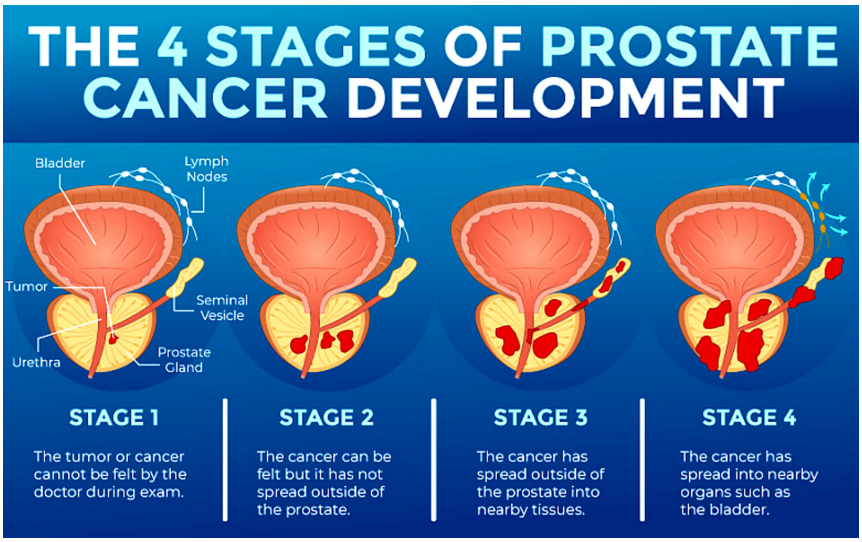
Due to the phenomenon of PSA leakage, the onset of prostate cancer diagnosis usually commences with a blood test to measure the PSA levels and is closely followed up by digital rectal examinations such as prostate biopsy and transrectal examinations [1]. Pending the positive diagnosis of a cancer, a Gleason index score is calculated which informs both the risk group of the patient and the potential treatment sources. Urodynamic-based methods have also been proposed by various authors as a means of inferring growth of the prostate via the closure of the urinary tract and therein the flow rate of urine; but despite the success of this method, it continues

[1]https://www.urotoday.com/library-resources/advanced-prostate-cancer/111566-epidemiology-and-etiology-of-prostate-cancer.html.

[2]https://drjockers.com/prostate-cancer/.

*Corresponding author: Ejay Nsugbe, Nsugbe Research Labs, UK. Email: ennsugbe@yahoo.com

Figure 1
An image of the various stage of prostate cancer



to remain mostly an auxiliary source in the diagnosis and monitoring of prostate cancer [11, 12].

Machine learning has been extensively applied in various aspects of clinical medicine over the years and has been viewed as a useful tool which can contribute toward bridging the gap between the limits of medical knowledge and patient needs and contributing toward enhancing clinical care strategies via the use of varied information sources toward predicting risks for various outcomes [13, 14]. From the literature, various machine learning methods have been applied toward the prediction of prostate cancer. These methods primarily hinge on a supervised learning architecture which is deemed a form of "weak AI" and relies upon a form of external expert-based intervention to label the training sample set before the designated machine learning algorithm learns for the various data classes [15–22].

Furthermore, these applied methods in the literature show a continued trend of utilizing a binary-based prediction framework (i.e., cancer/no-cancer) for the training of the machine learning classifier, which limits the insight provided by the classifier to the clinician, especially for a disease like cancer where the treatment delivered is dependent on the stage of the cancer. As a means toward providing a solution to this shortcoming, this study uses the prostate cancer dataset from the Kaggle website [23], containing geometric features from the digital examinations performed across 100 patients, to design two contrastive predictive systems. These systems are underpinned with unsupervised learning algorithms and provide a predictive insight into the stage of the cancer, which would carry high appeal in a clinical setting [14].

Specifically, the contributions made in this manuscript are as follows:

1) The application of fuzzy c-means unsupervised learning to create class labels for the data, followed by supervised learning to identify whether prostate cancer is present, and a subsequent prediction of the stage of cancer for cases where the patient does indeed have cancer
2) The application of the agglomerative hierarchical unsupervised learning method, alongside probabilistic supervised learning and fuzzy logic, to create an intelligent system capable of inferring a cancer stage given an estimated probability for a patient who is predicted to have prostate cancer

The intention is that either of these intelligent prediction models can play the role of a prediction machine in a previously proposed clinical cybernetic system involving human-machine collaboration that is poised toward enhancing care for cancer patients via the augmentation of both human and Artificial Intelligence within a clinical setting.

2. Mathematical Prostate Cancer and Dataset

2.1. Mathematical prostate cancer

Mathematical models have a high appeal in the study of cancer due to the complexity of the disease and help with the typical "what" and "how" questions, which include "the effectiveness of a care therapy for a specific patient", "optimal dosage", and "least vs most effective treatments" to name a few [1]. Although this paper is not a theoretical study, the intention is to present a mathematical model around prostate cancer and immunotherapy to show a mathematical relationship between the different variables and immunotherapy action as part of the prostate cancer disease. The models presented here are adopted from the work done by Kronik et al. [24] and Badziul et al. [25], as follows:

For the cellular vaccine V:

$$\frac{\mathrm{dV}}{\mathrm{dt}} = -\mathrm{K_i n_v V} \tag{1}$$

where is the rate of cell maturation after a vaccine, and is the number of vaccine cells necessary to take up maturation of a single dendritic cell.

For the antigen-presenting dermal dendritic cells D_m:

$$\frac{\mathrm{d}D_m}{\mathrm{dt}} = k_i\left(V + V_p\right) - k_m D_m \tag{2}$$

where V_p is the natural influx of mature dendritic cells, and k_m is the migration rate of dendritic cell migration from the skin to lymph nodes.

For the mature dendritic cells D_c:

$$\frac{\mathrm{d}D_c}{\mathrm{dt}} = \alpha_1 k_m D_m - k_{CR} D_C \tag{3}$$

where α_l is the fraction of antigen-presenting dendritic cells entering the lymph nodes.

For the "exhausted" dendritic cells D_R:

$$\frac{dD_R}{dt} = k_{CR}D_C - \mu_D D_R \tag{4}$$

where k_{CR} is the rate of exhaustion of mature dendritic cells, and μ_D is the death rate of the "exhausted" mature dendritic cells.

For the antigen-specific effector cell C:

$$\frac{dC}{dt} = a_c D_C - \mu_C C - k_R CR \tag{5}$$

where a_c is the effector cell recruitment by mature dendritic cells, μ_C is the effector cell death rate, and k_R is the rate of effector cell inactivation by a set of inhibitory cells.

For the regulatory/inhibitory cells R:

$$\frac{dR}{dt} = a_R D_R - \mu_R R \tag{6}$$

where a_R is the rate of inhibitory cell recruitment by exhausted dendritic cells, and μ_R is the death rate of the inhibitory cells.

For the prostate cancer cells P:

$$\frac{dP}{dt} = rP - a_p CP\frac{h_p}{h_p + P} \tag{7}$$

where r is the tumor growth rate, a_p represents the maximal prostate cancer killing efficacy, and h_p is the effector cell efficacy damping coefficient.

2.2. Dataset

The dataset used as part of this study is the open-source data available on the Kaggle website containing data collected from 100 patients involving digital examination of their prostates and contains post-processed values from the digitized images of the prostate, namely, Radius, Perimeter, Texture, Smoothness, Area, Compactness, Symmetry, and Fractal Dimension [23]. The dataset contained 68 cases where the patient was without prostate cancer and 32 cases where there was cancer present (malignant); the various classes were balanced using the SMOTE algorithm, as applied in various prior studies [14].[3] Due to the various scales, ranges, and values of the features within the dataset, the matrix containing all the features was standardized prior to the classifier training. It is worth noting that data integration and heterogeneity are not an immediate concern for this kind of dataset, due to the source of the data emanating from a clinical imaging modality.

3. Methods

3.1. Intelligent system 1

1) Fuzzy C-Means and supervised learning
2) Fuzzy C-Means (FCM)

Given a sample group of elements defined as $X = \{x_1, x_2, \ldots x_n\}$, the FCM is an unsupervised learning method that can partition the data into various groups based on its objective function.[4] In the FCM, membership functions are used to express the extent to which a sample data point belongs to a certain cluster, with a value of 1 implying a maximal value and full compatibility between a sample and a cluster.

The objective function used for the partitioning of data points into various clusters can be seen as follows:

$$J_m = \sum_{i=1}^{D}\sum_{j=1}^{N} \mu_{ij}^m ||x_i - c_j||^2 \tag{8}$$

where D represents the number of data points, N is the total number of clusters, m is the fuzzy partition matrix exponent, x_i is an ith data point, c_j is a cluster centroid, and μ_{ij} is the degree of membership of a candidate point x_i in a jth cluster.

For the work done in this paper, the fuzzy partition matrix exponent was chosen to be 2, while the soft clustering option was utilized with a set membership function value of 0.6 in order to help identify candidate patients who have eligibility in both clusters (cancer and no-cancer) and thus could potentially be an early-stage cancer.

3) Supervised learning algorithms

The labeled data from the FCM were passed into a set of supervised learning algorithms, where in this stage the following sets of supervised learning algorithms were applied and contrasted.

Decision Tree (DT): DTs are a class of non-parametric learners which apply a Boolean-like logic toward sorting data into various classes, and are typically regarded as a white-box-based modeling approach due to the transparency associated with their classification process [14]. The mathematical underpinning of decision trees can be seen in Nsugbe et al. [14].

Quadratic Support Vector Machine (QSVM): this classifier is of an iterative form and works toward finding an optimal separation boundary between data classes in a high dimensional subspace, while utilizing a portion of the data known as support vectors for this [13, 14]. The classifier utilizes a "kernel" trick which is a computationally effective means of projecting the data into a high dimensional subspace where the class boundaries are applied and the overall structure is preserved [13, 14]. The quadratic kernel was used and therein a non-linear separation boundary between the respective data classes; the mathematical formulation and prior applications of the classifier in related studies can be seen in Nsugbe et al. [13, 14].
The classification exercises were conducted using the MATLAB Classification Learner Application which, for a selected model and options, iteratively tunes for the best parameters that fulfill the selected architecture; the classifiers were validated using a k-fold cross-validation framework with k chosen to be 10, while the data were split into 80% for training and 20% for validation purposes.

3.2. Intelligent system 2

1) Agglomerative hierarchical clustering, probabilistic learning, and fuzzy logic
2) Agglomerative hierarchical clustering (AHC)

This is an unsupervised method that utilizes a hierarchical tree-based format to produce distinct classes for unlabeled data.[5] The three key steps associated with the computation of the AHC include:

A similarity calculation using the default Euclidean distance, assuming points x_s and x_t, expressed as follows:

[3] https://www.analyticsvidhya.com/blog/2020/10/overcoming-class-imbalance-using-smote-techniques/

[4] https://uk.mathworks.com/help/fuzzy/fcm.html-Supervised%20learning

[5] https://uk.mathworks.com/help/stats/linkage.html

$$d_{st}^2 = (x_s - x_t)(x_s - x_t) \tag{9}$$

where d denotes the Euclidean distance.

After the first step, data points within proximity of each other are associated together using a linkage function in an iterative fashion. The Ward's linkage function was applied in the implementation of the algorithm in this study, and is mathematically expressed as follows:

$$\sqrt{\frac{2n_r n_s}{(n_r + n_s)}}||\bar{x}_r - \bar{x}_s||_2 \tag{10}$$

where $||\square||_2$ is the Euclidean distance metric, $\bar{x}_r$ and $\bar{x}_s$ are centroids for candidate clusters r and s, and n_r and n_s are the number of points assigned to those two clusters.

The final stage involves the pruning of the dendrogram clusters into distinct groups as part of the final partitioning process and involves the identification of groupings within the hierarchy of the dendrogram.

3) Supervised learning algorithms

In this scenario, the labels produced by the AHC were passed toward a set of probabilistic supervised learning algorithms, as follows:

Probabilistic Support Vector Machines (PSVM): this algorithm works in the same way as the standard support vector machine but utilizes the Platt scaling mechanism to convert the classification scores from the model, which yield binary class labels into probability distributions across various classes [26].[6] This Platt scaling allows for the obtaining of a probabilistic score associated with a class prediction which, unlike the classical classification, provides an associated degree of certainty about the prediction [26]. The Platt scaling works with a fitting process where the classification scores from a model are converted to a probabilistic reflection via the fitting of a logistic regression model, which is expressed mathematically as follows:

$$\frac{1}{1 + \exp(Af(x) + B)} \tag{11}$$

where $f(x)$ represents the classifier scores, and parameters A and B are learned as part of the algorithm's fitting process.

The medium Gaussian kernel was used for the support vector machine's computations prior to the probabilistic conversion.

Naïve Bayes (NB): underpinned by the Bayes probabilistic rule, and expressed in Equation (12), the NB can provide associated probabilities for label class predictions [14, 27].

$$P(H_k|E) = \frac{P(E|H_k)P(H_k)}{P(E)} \tag{12}$$

where H_k represents the number of classes, $P(E|H_k)$ is the joint probability with a prior H_k, $P(H_k)$ is an initial probability held by the hypothesis, $P(E)$ is a variable for the assumption that the training data can potentially be observed with the supplied information within the feature vector, while $P(H_k|E)$ is the estimated posterior probability which conveys a certain confidence level of the hypothesis after the training data has been observed. The NB algorithm assumes that the data are normally distributed, $P(x_i|H_k)$, a mathematical framework of how the data are sorted into their various classes for a sample 2-class problem, H_a and H_b, assuming H_a carries a higher likelihood, and can be seen as follows:

$$P(H_a)\prod_{i=1}^{N} P(x_i|H_a) > P(H_b)\prod_{i=1}^{N} P(x_i|H_b) \rightarrow P(H_a|x) > P(H_b|x) \tag{13}$$

where, for a specific class:

$$\mathrm{H} = \mathrm{argmaxP}(H_k)\prod_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{P}(x_i|H_b) \tag{14}$$

where k∈{1,2, ...k}, and H is the most likely class given a feature vector x.

4) Fuzzy logic (FL)

FL contributes toward the use of linguistic variables to model expert knowledge on a certain process, and in a sense extends classical logic toward a form of truth modeling [28, 29]. Looking at this from a set theory perspective, where classical logic typically requires an element to be either in a set or not, fuzzy logic allows for partial membership within a set. The fuzzy sets (FS) are a component of FL that carries the extent to which an element belongs to a particular set.

The type 1 FL is adopted for use in this work and utilizes linguistic variables such as small, medium, and large with respect to the size of a building, for example, based on expert knowledge. The term "membership function" (MF) is currently used in FL systems and refers to the extent of the similarity of an element within a set to the FS. For the deployment of an FL system into an engineering or clinical setting, a complete system comprising the following constituent parts is required, a diagram of which can be seen in Figure 2.

Fuzzification: the fuzzification step takes in information about the process at hand, which contains a degree of vagueness and forms an internal representation of this information [28, 29].
This process makes the crisp input quantities fuzzy. The premise behind fuzzification is that many quantities which are presumed crisp, carry a significant amount of uncertainty [28, 29]. For example, the quantity "about 6 feet tall" carries uncertainty in terms of vagueness of information. The ammeter usually outputs crisp current values, but this could be subject to several errors which would result in imprecision. Fuzzification provides a mechanism to represent, and consequently handle, these types of uncertainties.

Rule base: the rule base aspect of the FL system is responsible for the mapping of linguistic input variables to a respective output action, which forms a key component of what makes FL systems interpretable and thus of high appeal to clinical applications [28, 29]. The rule base generically works with an IF-THEN architecture.

Inference engine: the inference engine is responsible for the conversion of the input data into a respective action based on a set of defined rules.

Defuzzification: the defuzzification process can be viewed as an interpolation process that produces a final output – methods for defuzzification include maxima, centroid, and center of sums methods [28, 29].

For the implemented FL system in this work, the inputs were modeled using Gaussians while the outputs were modeled using triangular functions, the Mamdani type 1 fuzzy system was implemented, and the defuzzification was achieved using a centroid approach.

[6] https://uk.mathworks.com/help/stats/classificationsvm.fitposterior.html

Figure 2
FL system diagram showing the various system components

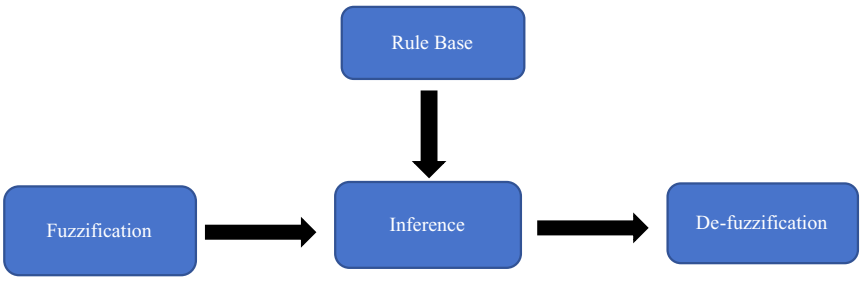


Table 1
Results from the implementation of intelligent system 1

	ACC	SEN	SPEC	AUC
FCM-DT	88.3 ± 0.95	93.7 ± 0.95	86.0 ± 0.0	90.0 ± 0.0
FCM-QSVM	88.7 ± 0.48	87.0 ± 0.0	80.4 ± 0.97	83.7 ± 0.48

4. Results

The following classifier metrics were used to assess the performance of the designed intelligent systems, as adopted from previous studies, namely, accuracy (ACC), sensitivity (SEN), specificity (SPEC), and area under the curve (AUC) [14].

4.1. Intelligent system 1

The clustering accuracy was calculated as the number of correctly classified samples using the accompanying labels of the dataset as ground truth, and it was seen that the FCM produced a clustering accuracy in the region of 70%. The results for the FCM labeled data for a 3-class cancer classification problem comprising Cancer, Early-Stage Cancer, and No-Cancer can be seen in Table 1. From this, it can be noted that both classifiers produced accuracies greater than 80% across the four chosen classifier metrics, with the DT seen to produce the best set of results, thereby showcasing the notion that a tree-based splitting and classification method is superior to the kernel method for this kind of dataset.

A diagrammatic flow of the proposed Intelligent System 1 can be seen in Figure 3.

4.2. Intelligent system 2

The AHC clustering exercise produced a clustering accuracy of 69%. The result of the 2-class exercise to probabilistically predict cancer vs no-cancer can be seen in Table 2. Both candidate classifiers produced accuracies around 90%, with the PSVM performing slightly better than the NB, which can be attributed to the superiority of the complexity of kernel-based classifiers when compared to NB, which has a simpler classification architecture.

As mentioned, for cases where cancer has been predicted by a classifier, the associated probability estimate is to be passed through to an FL system primed toward associating a cancer stage based on a probabilistic rule set, as seen in Table 3. A screenshot of the rules in MATLAB's FL toolbox in symbolic format can be seen in Figure 4.

A diagrammatic representation of the Intelligent System 2 framework can be seen in Figure 5.

Figure 3
Diagrammatic representation of intelligent system

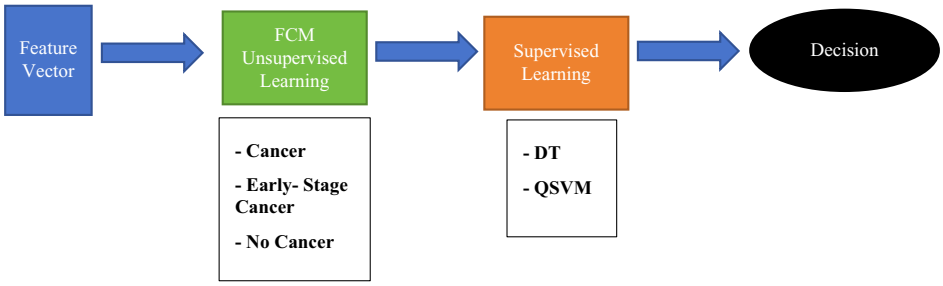


Table 2
Results from the implementation of intelligent system 2

	ACC	SEN	SPEC	AUC
AHC-PSVM	94.2 ± 0.79	93.4 ± 1.10	96.0 ± 0.0	94.0 ± 0.0
AHC-NB	90.9 ± 0.74	89.4 ± 0.70	92.5 ± 0.85	90.8 ± 0.63

Table 3
Probability values and associated potential cancer stage

Probability value	Potential cancer stage
0.4–0.5	Early stage
0.5–0.7	Medium stage
0.7+	Advanced stage

Contrasting both sets of intelligent system architectures, it can be seen that Intelligent System 1 produces slightly lower classifier performance metrics, but provides more insight on various stages of the prostate cancer, while posing a simpler computational architecture, which also implies a simpler hardware demand and maintenance requirement from an implementation perspective.

While Intelligent System 2 provides a higher set of classification metrics, there are more models involved in the overall architecture; hence, it is envisaged to be relatively more expensive to implement and maintain from a practicality perspective. The mechanism through which a cancer stage is estimated in Intelligent System 2 is through an indirect means and therein an inference approach via probability estimates; thus, it may be possible that this aspect of the system may be a cause for concern from a clinical application perspective.

It can also be noted from the architecture of both intelligent systems that neither one contains a black box prediction architecture; thus, both systems have model interpretability incorporated within them as part of their respective designs.

4.3. Feature ranking exercise

In order to identify the key drivers within the dataset driving the differentiation between the prediction of the prostate cancer vs no-cancer, a feature ranking exercise was conducted using the Relief algorithm, which is a filter-based feature selection algorithm [30]. The following were the five top-ranked features in order of importance from the algorithm:

1) Compactness
2) Fractal dimension
3) Smoothness
4) Radius
5) Perimeter

Given the scope of this paper, it is not immediately apparent how these can be linked directly to the structure and physiology of the male prostate in order to translate this information into tangible biomarkers. Future work will involve a further study around this basis.

Figure 4
FL representation of the rule set in symbolic format from the MATLAB FL toolbox

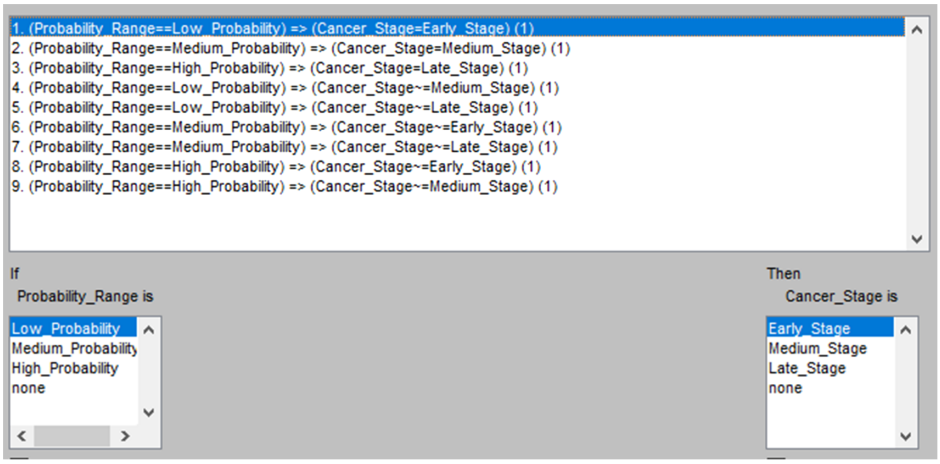
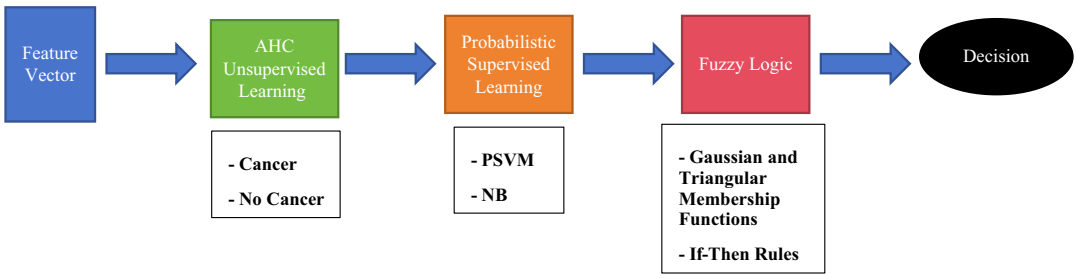


Figure 5
Diagrammatic representation of intelligent system 2



5. Conclusion

Prostate cancer is a widespread cancer which accounts for a high mortality in adult males. Medically it has been seen that factors such as age, family history, and ethnicity are some of the main drivers behind the cancer, alongside the notion that Black and African ethnicity have a greater predisposition toward this variant of cancer.

Machine learning models have been investigated in this study toward the prediction of prostate cancer. The source of the data is the Kaggle dataset, which contains features extracted from digital examination of 100 patients. This study showcases the application of a relatively "Strong AI" framework underpinned by unsupervised learning for the design of a contrastive set of Intelligent Systems to self-sort the prostate cancer dataset prior to the application of a supervised learning model. Two sets of Intelligent Systems were designed, where the first candidate system works with the FCM followed by supervised learning for the prediction of whether a patient has cancer, alongside the detection of a potential early-stage cancer which is medically difficult to diagnose with current methods. The second Intelligent System is underpinned by AHC alongside probabilistic learning, which allows for a cancer stage inference via the use of probabilistic estimates and is further automated with the application of a type 1 Mamdani FL system.

The results from both sets of Intelligent Systems showcased various aspects toward predicting prostate cancer alongside various associated cancer stages using different prediction frameworks. Both of the systems can be seen to carry interpretability due to the refrainment from the use of black box prediction architectures and are therein likely to carry a high degree of clinical appeal for a candidate decision support system. From a practical perspective, it is envisaged that Intelligent System 2 is likely to carry a greater resource demand for the implementation of the system due to its model architecture, while Intelligent System 1 offers a lower hardware implementation requirement.

Future work in this area is set to involve the investigation of the use of regression techniques that will allow for a continuous cancer stage estimation, as has been shown in separate studies [31, 32]. The complexity of the FL system can also be expanded by including further rule sets and clinical expert intuition in order to further robustify the system and perhaps contribute toward making it more informative.

It should be noted that the proposed prediction models are intended toward forming a key component in a cybernetic architecture which facilitates human-machine collaboration between clinical experts and a prediction machine as part of strides toward an enhanced platform for improving care strategies within clinical medicine.

Furthermore, other aspects of extending the proposed work could involve the development of a cloud-based "human health monitoring solution", where further considerations such as data storage and networking need to be considered.

Acknowledgement

In loving memory of Sodi's dad who was taken from us far too early due to prostate cancer. We were too young to understand what this was when it happened, but you left behind a strong lineage and legacy through your work and the strong children which you raised.

The author would like to thank Brian Kerr from Kerr Editing for proofreading this manuscript.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by the author.

Conflicts of Interest

The author declares that he has no conflicts of interest to this work.

Data Availability Statement

Data is available on request from the corresponding author upon reasonable request.

Author Contribution Statement

Ejay Nsugbe: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

References

[1] Phan, T., Crook, S. M., Bryce, A. H., Maley, C. C., Kostelich, E. J., & Kuang, Y. (2020). Mathematical modeling of prostate cancer and clinical application. *Applied Sciences*, *10*(8), 2721. https://doi.org/10.3390/app10082721

[2] Etzioni, R., Gulati, R., Tsodikov, A., Wever, E. M., Penson, D. F., Heijnsdijk, E. A., . . . , & Mariotto, A. B. (2012). The prostate cancer conundrum revisited: Treatment changes and prostate cancer mortality declines. *Cancer*, *118*(23), 5955–5963. https://doi.org/10.1002/cncr.27594

[3] Potosky, A. L., Miller, B. A., Albertsen, P. C., & Kramer, B. S. (1995). The role of increasing detection in the rising incidence of prostate cancer. *JAMA*, *273*(7), 548–552. http://doi.org/10.1001/jama.1995.03520310046028

[4] Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians*, *68*(1), 7–30. https://doi.org/10.3322/caac.21442

[5] Gandaglia, G., Karakiewicz, P. I., Abdollah, F., Becker, A., Roghmann, F., Sammon, J. D., . . . , & Sun, M. (2014). The effect of age at diagnosis on prostate cancer mortality: A grade-for-grade and stage-for-stage analysis. *European Journal of Surgical Oncology*, *40*(12), 1706–1715. https://doi.org/10.1016/j.ejso.2014.05.001

[6] Klaassen, Z., Howard, L., Vidal, A., Moreira, D., Andriole, G., Castro-Santamaria, R., & Freedland, S. (2018). *AUA 2018: Black race predicts poor compliance but higher prostate cancer risk: Results from the REDUCE Trial*. Retrieved from: https://www.urotoday.com/conference-highlights/aua-2018/aua-2018-prostate-cancer/104427-aua-2018-black-race-predicts-poor-compliance-but-higher-prostate-cancer-risk-in-a-multinational-repeat-biopsy-cohort-results-from-reduce.html

[7] Wang, B. D., Yang, Q., Ceniccola, K., Bianco, F., Andrawis, R., Jarrett, T., . . . , & Lee, N. H. (2013). Androgen receptor-target genes in African American prostate cancer disparities. *Prostate Cancer*, *2013*(1), 763569. https://doi.org/10.1155/2013/763569

[8] Mitra, A. V., Bancroft, E. K., Barbachano, Y., Page, E. C., Foster, C. S., Jameson, C., . . . , & Eeles, R. A. (2011). Targeted prostate cancer screening in men with mutations in BRCA1 and BRCA2 detects aggressive prostate cancer: Preliminary analysis of the results of the IMPACT study. *BJU International*, *107*(1), 28–39. https://doi.org/10.1111/j.1464-410X.2010.09648.x

[9] Ekman, P. (2000). The prostate as an endocrine organ: Androgens and estrogens. *The Prostate*, *45*, 14–18. https://doi.org/10.1002/1097-0045(2000)45:10+%3C14::AID-PROS4%3E3.0.CO;2-7

[10] Kuang, Y., Nagy, J. D., & Eikenberry, S. E. (2018). *Introduction to mathematical oncology*. USA: CRC Press.

[11] Litwin, M. S., & Tan, H. J. (2017). The diagnosis and treatment of prostate cancer: A review. *JAMA*, *317*(24), 2532–2542. https://doi.org/10.1001/jama.2017.7248

[12] Provost, M. J. (2014). Everything works wonderfully: An overview of servitization and physical asset management. *Asset Management & Maintenance Journal*, *27*(5), 43–45. https://search.informit.org/doi/10.3316/informit.581600748488062

[13] Nsugbe, E., Phillips, C., Fraser, M., & McIntosh, J. (2020). Gesture recognition for transhumeral prosthesis control using EMG and NIR. *IET Cyber-Systems and Robotics*, *2*(3), 122–131. https://doi.org/10.1049/iet-csr.2020.0008

[14] Nsugbe, E., Samuel, O. W., Asogbon, M. G., & Li, G. (2021). Phantom motion intent decoding for transhumeral prosthesis control with fused neuromuscular and brain wave signals. *IET Cyber-Systems and Robotics*, *3*(1), 77–88.https://doi.org/10.1049/csy2.12009

[15] Abraham, B., & Nair, M. S. (2019). Automated grading of prostate cancer using convolutional neural network and ordinal class classifier. *Informatics in Medicine Unlocked*, *17*, 100256. https://doi.org/10.1016/j.imu.2019.100256

[16] Chiu, P. K. F., Shen, X., Wang, G., Ho, C. L., Leung, C. H., Ng, C. F., . . . , & Teoh, J. Y. C. (2022). Enhancement of prostate cancer diagnosis by machine learning techniques: An algorithm development and validation study. *Prostate Cancer and Prostatic Diseases*, *25*(4), 672–676. https://doi.org/10.1038/s41391-021-00429-x

[17] Erdem, E., & Bozkurt, F. (2021). A comparison of various supervised machine learning techniques for prostate cancer prediction. *European Journal of Science and Technology*, (21), 610–620. https://doi.org/10.31590/ejosat.802810

[18] Janssen, F. M., Aben, K. K. H., Heesterman, B. L., Voorham, Q. J. M., Seegers, P. A., & Moncada-Torres, A. (2022). Using explainable machine learning to explore the impact of synoptic reporting on prostate cancer. *Algorithms*, *15*(2), 49. https://doi.org/10.3390/a15020049

[19] Li, H., Lee, C. H., Chia, D., Lin, Z., Huang, W., & Tan, C. H. (2022). Machine learning in prostate MRI for prostate cancer: Current status and future opportunities. *Diagnostics*, *12*(2), 289. https://doi.org/10.3390/diagnostics12020289

[20] Mokoatle, M., Mapiye, D., Marivate, V., Hayes, V. M., & Bornman, R. (2022). Discriminatory Gleason grade group signatures of prostate cancer: An application of machine learning methods. *PLOS ONE*, *17*(6), e0267714. https://doi.org/10.1371/journal.pone.0267714

[21] Nematollahi, H., Moslehi, M., Aminolroayaei, F., Maleki, M., & Shahbazi-Gahrouei, D. (2023). Diagnostic performance evaluation of multiparametric magnetic resonance imaging in the detection of prostate cancer with supervised machine learning methods. *Diagnostics*, *13*(4), 806. https://doi.org/10.3390/diagnostics13040806

[22] Srivenkatesh, M. (2020). Prediction of prostate cancer using machine learning algorithms. *International Journal of Recent Technology and Engineering*, *8*(5), 5353–5362. https://doi.org/10.35940/ijrte.E6754.018520

[23] Saifi, S., & Mahmood, S. (2018). *Prostate cancer* [Data set]. Kaggle. https://kaggle.com/sajidsaifi/prostate-cancer

[24] Kronik, N., Kogan, Y., Elishmereni, M., Halevi-Tobias, K., Vuk-Pavlović, S., & Agur, Z. (2010). Predicting outcomes of prostate cancer immunotherapy by personalized mathematical models. *PLOS ONE*, *5*(12), e15482. https://doi.org/10.1371/journal.pone.0015482

[25] Bądziul, D., Jakubczyk, P., Chotorlishvili, L., Toklikishvilie, Z., Traciak, J., Jakubowicz-Gil, J., & Chmiel-Szajner, S. (2020). Mathematical prostate cancer evolution: Effect of immunotherapy based on controlled vaccination strategy. *Computational and Mathematical Methods in Medicine*, *2020*(1), 7970265. https://doi.org/10.1155/2020/7970265

[26] Platt, J. C. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In A. J. Smola, P. Bartlett, B. Schölkopf, & D. Schuurmans (Eds.), *Advances in large margin classifiers* (pp. 61–74). MIT Press.

[27] Sjöqvist, H., Längkvist, M., & Javed, F. (2020). An analysis of fast learning methods for classifying forest cover types. *Applied Artificial Intelligence*, *34*(10), 691–709. https://doi.org/10.1080/08839514.2020.1771523

[28] Mahfouf, M., Abbod, M., & Linkens, D. (2001). A survey of fuzzy logic monitoring and control utilisation in medicine. *Artificial Intelligence in Medicine*, *21*(1–3), 27–42. https://doi.org/10.1016/S0933-3657(00)00072-5

[29] Obajemu, O. (2016). *Predictive dynamic risk mapping and modelling of patients diagnosed with bladder cancer.* PhD Thesis, University of Sheffiled.

[30] Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based feature selection: Introduction and review. *Journal of Biomedical Informatics*, *85*, 189–203. https://doi.org/10.1016/j.jbi.2018.07.014

[31] Nsugbe, E., Starr, A., Jennions, I., & Carcel, C. R. (2017). Particle size distribution estimation of a mixture of regular and irregular sized particles using acoustic emissions. *Procedia Manufacturing*, *11*, 2252–2259. https://doi.org/10.1016/j.promfg.2017.07.373

[32] Nsugbe, E., Ruiz-Carcel, C., Starr, A., & Jennions, I. (2018). Estimation of fine and oversize particle ratio in a heterogeneous compound with acoustic emissions. *Sensors*, *18*(3), 851. https://doi.org/10.3390/s18030851

How to Cite: Nsugbe, E. (2024). Towards Unsupervised Learning Driven Intelligence for Prediction of Prostate Cancer. *Artificial Intelligence and Applications*, *2*(4), 291–298. https://doi.org/10.47852/bonviewAIA42022210