Received: 10 July 2022 | Revised: 24 October 2022 | Accepted: 25 October 2022 | Published online: 14 November 2022

RESEARCH ARTICLE

Artificial Intelligence and Applications
2023, Vol. 1(3) 155–160
DOI: 10.47852/bonviewAIA2202308

BON VIEW PUBLISHING

A Pilot on the Use of Unsupervised Learning and Probabilistic Modelling Towards Cancer Extent Prediction

Ejay Nsugbe[1],*

[1]*Nsugbe Research Labs, UK*

Abstract: Cervical cancer is a global disease that attacks the female reproductive system and remains challenging to diagnose in the early stages. The rise in the application of machine learning models in various aspects of clinical medicine has seen research an increase in research applied toward the diagnosis of cervical cancer using a questionnaire style format. This work explores the application of unsupervised learning for an automated partitioning and learning of class labels, followed by subsequent learning of the designated labels using a supervised learning method. A contrast was conducted between two unsupervised learning methods, namely the agglomerative hierarchical clustering and k-means, where it was seen the k-means provided superior and more accurate clustering capabilities. The labels provided by the k-means were used to train the naïve Bayes (NB) classifier, which is capable of outputting probabilities associated with class predictions, therein allowing for the stage of cancer to be inferred. The results from the NB showed high accuracies in the region of 90+% across the various classifier metrics, therein showing good learning capability for learning from clustered data. The study provides evidence to show that a fully automated platform, which utilizes questionnaire style information, can predict whether a patient has cervical cancer alongside the likely cancer extent. Future work would now involve the use of a rule-based inference system that can automatically provide clinicians with a cancer extent where necessary and is also capable of handling further cancer stage subclasses information.

Keywords: intelligent gynecology, machine learning, unsupervised learning, clinical decision support, cancer, bioinformatics, artificial intelligence

1. Introduction

Cancer manifests itself with an uncontrollable growth spurt of abnormal cells which are capable of destroying healthy tissues and neighboring organs (El-Moselhy et al., 2016; Saha et al., 2010). Cancer of the cervix is a cancer type which targets the cervix within the female reproductive system and is viewed as the fourth major cause of cancer-related deaths worldwide on an annual basis (El-Moselhy et al., 2016; Saha et al., 2010).

This cancer is effectively undiagnosable in its early stage, and studies have shown that the lack of specialist skills and care resources in developing countries have led to the high cancer mortality present in these regions (Wu & Zhou, 2017). Some of the symptoms associated with intermediate to advanced-stage cancer include pelvic pain and considerable vaginal bleeding as the cancer spreads and attacks nearby organs (Wu & Zhou, 2017). In order to improve mortality rate and overall care strategies globally with respect to cervical cancer, an early detection of the cancer is important. Gynecological cancer societies have established guidelines and management strategies in terms of cervical cancer, the majority of which are diagnosed via biopsies and histological methods, although there continues to be a lack of clarity around defining concrete sensitivity and specificity metrics for these methods (American Cancer Society, 2018). An image of a cancer tissue across the surface of the cervix can be seen in Figure 1.

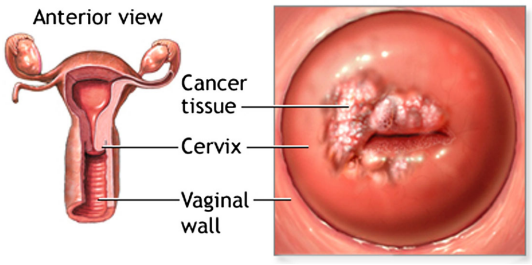
1.1. Background literature and contribution to knowledge

Machine learning has seen broad-scale application in various aspects of medicine due to its broad appeal and potential, likewise in the case of cervical cancer prediction (Nsugbe et al., 2020, 2021). The shortcoming in the literature of the published work in this area mostly present binary machine learning prediction models, which address the problem of cervical cancer to a degree, but give no indication how their proposed system would sit within a clinical framework (Bora et al., 2016; Chandran & Kumari, 2015; Ghoneim et al., 2020; Malli & Nandyal, 2017; Tseng et al., 2014; Zhang et al., 2017).

Strides toward tackling this shortcoming were proposed in a prior work where the author presented a cybernetic system which hosts a cervical cancer prediction machine and allows for a human–machine collaboration between the clinical experts and the said prediction machine within a clinical setting (Nsugbe, 2022). The prediction machine comprised of a combination of an unsupervised learning method to cluster and label the data, followed by supervised learning of the partitioned data (Khanam et al., 2015; Singh et al., 2016). The unsupervised learning method involved the use of probabilistic Gaussian mixture models (GMMs) and

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

Figure 1
Image of cancer tissue across the cervix (MedlinePlus, 2021)



fuzziness (fuzzy c-means)-based learners which not only separate the data into distinct clusters but also produce uncertainty/fuzziness metrics which were translated into stages of cervical cancer to create subclasses for supervised learning and were capable of working in an automated fashion (Nsugbe, 2022; Nsugbe et al., 2020).

As unsupervised learning models represent a relatively stronger artificial intelligence framework, in this work, a different set of unsupervised learning models are investigated in tandem with supervised learning for the solving of the cervical cancer prediction problem (IBM Cloud Education, 2021). By way of investigating different unsupervised learning models with a different learning framework and objective function, it is acknowledged that the ability to predict cancer stage would now be unavailable due to the absence of the GMM and Fuzzy C-Means (FCM) algorithms. To compensate for this, a probabilistic supervised learning method is trained based on the clustered partitioned from the prior process involving unsupervised learning and is capable of providing an associated probabilistic association with its predictions that can be used to infer cancer extent (Ghahramani, 2015). Specifically speaking, the contribution of this paper is as follows:

- A contrastive application of two unsupervised learning algorithms, namely, agglomerative hierarchal clustering and k-means, followed by supervised learning using the naïve Bayes (NB) algorithm, which is capable of providing a probabilistic association with each patient that is predicted to have cervical cancer

2. Cancer Spread Dynamics

As deduced in prior work, the presence of theoretical models allows for a deeper understanding and appreciation of the theoretical underpinning and dynamic interactions between multiple entities during a medical manifestation (Nsugbe, 2022). It is worth mentioning that, although this isn't a theoretical simulation-based work, the subsequent theoretical models have been included to help reinforce the understanding of how the related cancer variables interact amidst treatment, and so on (Araujo & McElwain, 2004; Powathil et al., 2013; Unni & Seshaiyer, 2019; Watanabe et al., 2016). The presence and contributions of these kinds of mathematical models have yielded further medical insights into describing tumor dynamics and immune system response to various therapies, in addition to the tumor therapy responses (Araujo & McElwain, 2004; Powathil et al., 2013; Unni & Seshaiyer, 2019; Watanabe et al., 2016).

The cancer model presented here is one adopted from the work done from Unni and Seshaiyer (2019), which is a model that quantitatively accounts for the interaction between tumors, killer cells, and cytotoxic CD8 + T cells. Unni and Seshaiyer's model comprises four cell candidates, which include the tumor cells $T(t)$, the natural cell killers $N(t)$, the dendritic cells $D(t)$, and the cytotoxic $CD8^+$ T cells $L(t)$, and not only shows the interaction between cells but also their response to cancer treatment (Unni & Seshaiyer, 2019).

An equation describing the interaction for each cell can be seen as follows:

- Tumor cells: using the logistic law to denote proliferation cell rates, we have

$$aT(1 - bT) \tag{1}$$

where a and b are the periodic growth rate and accompanying carrying capacity of the cancer tumor cells, respectively. Accounting for and modeling the interaction between the tumor cells with dendritic cells, naturally occurring killer cells and $CD8^+$ T cells, we have the following (Unni & Seshaiyer, 2019):

$$-(c_1N + jD + kL)T \tag{2}$$

where c_1, j, and k represent the natural killer (NK) cells, dendritic cells, and $CD8^+$ T cells, respectively. Putting equations (1) and (2) together and including variables to account for pharmacokinetic effects and a kill cell parameter for the cell populations, we arrive at the following (Unni & Seshaiyer, 2019):

$$\frac{dT}{dt} = aT(1 - bT) - (c_1N + jD + kL)T - K_T z(M)T \tag{3}$$

where $z(M) = 1 - e^{-M}$ represents a drug-induced cell pharmacokinetic effect, and K_T is the kill parameter for the tumor cell populations.

- NK cells: for a cell recruitment source s_1, alongside the Michaelis–Menten factor, we arrive at the following (Unni & Seshaiyer, 2019):

$$g_1 \cdot \frac{T^2}{h_1 + T^2} \cdot N \tag{4}$$

where g_1 indicates the maximum NK cells recruitment rate by the tumor cells, and h_1 is the steepness of the recruitment curve of the cell. These NK cells also interact with both the tumor and dendritic cells. Accounting for this interaction and introducing parameters for the rate of tumor killings and proliferation of the NK cells c_2 and d_1, we can arrive at the following (Unni & Seshaiyer, 2019):

$$\frac{dN}{dt} = s_1 + \frac{g_1 NT^2}{h_1 + T^2} - (c_2T - d_1D)N - K_N z(M)N - eN \tag{5}$$

where $-eN$ represents the naturally occurring death of the NK cells

- Dendritic cells: for the dendritic cells, we assume a source s_2, d_2, and d_3 to be the rate at which the NK cells kill the dendritic cells and the proliferation rate of the dendritic cells in the presence of tumors, f_1 to be the rate of interaction comprising the dendritic cells and CD8 + T cells, while g denotes an organic death rate of the dendritic cells, all of which superimpose to yield the following (Unni & Seshaiyer, 2019):

$$\frac{dD}{dt} = s_2 - (f_1L + d_2N - d_3T)D - K_D z(M)D - gD \tag{6}$$

- CD8 + T cells: these cells are influential in the removal of tumor cells and typically appear in clusters around tumor cells when they are present, disappearing subsequently afterward. The equation for

the CD8 + T cells and various interactions is as follows (Unni & Seshaiyer, 2019):

$$\frac{dL}{dt} = f_2 DT - hLT - uNL^2 + r_1 NT + \frac{p_1 LI}{g_1 + I} - K_L z(M)L - iL \tag{7}$$

where f_2 symbolizes the interaction between the dendritic cells and the tumor cells, $-hLT$ models the active competition between CD8 + T cells and tumor cells, uNL^2 represents a suppressive regulation of CD8 + T cells which could be unresponsive to cytokine, and $r_1 NT$ is a recruitment term of the CD8 + T cells assembled by the tumor cells which are undergoing lysis by the NK cells. CD8 + T cells, which are potentially activated by cancer treatment medication, can be characterized by the Michaelis–Menten interaction, and is expressed as $\frac{p_1 LI}{g_1+I}$, with I denoting the concentration of the immunotherapy within the bloodstream, and $-iL$ is the death rate of the CD8 + T cells (Unni & Seshaiyer, 2019).

3. Dataset and Methods

3.1. Dataset

The data used for the study emanates from the University of California at Irvine (UCI) database which comprises data collected from the Hospital Universitario de Caracas in Caracas, Venezuela (UCI Machine Learning Repository: Cervical Cancer (Risk Factors) Data Set, 2021). The data comprised a total of 28 features and biomarkers which have been medically proven to contribute toward both human papillomavirus (HPV) and cervical cancer – the dataset consisted of labels which were produced via Hinselmann, Schiller, Cytology, and Biopsy, with the Biopsy chosen as the target label due to being the more accurate method (UCI Machine Learning Repository: Cervical Cancer (Risk Factors) Data Set, 2021). The dataset comprises a questionnaire-style matrix from a total of 858 patients, where the questions spanned areas such as age, number of sexual partners, age of first sexual intercourse, number of prior pregnancies, smoker, use of hormonal contraceptives, presence of intrauterine device (IUD), cervical intraepithelial neoplasia (CIN), and HPV (UCI Machine Learning Repository: Cervical Cancer (Risk Factors) Data Set, 2021). A description of these features can be seen in Nsugbe (2022).

It was noted that some patients did not fill in all aspects of the questionnaire and left a fair number of the features blank. These participants were excluded from the final feature vector. In addition to this, columns 15, 22, 27, and 28 were also excluded from the feature vector due to not containing much information, thus leaving the final feature vector to contain 650+ rows, corresponding to over 650 patients. The SMOTE synthetic sample generator was also used for further sample generation due to class imbalance (Blagus & Lusa, 2013).

3.2. Prediction machines

3.2.1. Unsupervised learning

Agglomerative hierarchical clustering (AHC): this clustering method groups data into cluster dendrograms containing a multilevel hierarchical flow, assembled together to represent clusters (Agglomerative Hierarchical Cluster Tree - MATLAB Linkage - MathWorks United Kingdom, 2021; Davidson & Ravi, 2005; Day & Edelsbrunner, 1984). There are three key steps associated with the AHC, as can be seen as follows:

- Similarity evaluation between objects was done using the default Euclidean distance function, which is expressed as follows, assuming points x_s and x_t:

$$d_{st}^2 = (x_s - x_t)(x_s - x_t) \tag{8}$$

where d denotes the Euclidean distance.

- At this stage, points that are in close proximity are paired together using the linkage function, which hinges upon the information regarding distance obtained from the prior step (Agglomerative Hierarchical Cluster Tree - MATLAB Linkage - MathWorks United Kingdom, 2021; Davidson & Ravi, 2005; Day & Edelsbrunner, 1984). As new clusters are formed, they are subsequently regrouped into bigger clusters until a hierarchy is formed. The Ward's linkage function was used in the implementation of the algorithm and is mathematically expressed as:

$$\sqrt{\frac{2n_r n_s}{(n_r + n_s)}} ||\bar{x}_r - \bar{x}_s||_2 \tag{9}$$

where $||||_2$ denotes a Euclidean distance, $\bar{x}_r$ and $\bar{x}_s$ are centroids for clusters r and s, and n_r and n_s are the number of points in clusters r and s.

- At this stage, the branches of the clustering dendrogram are pruned and assigned to distinct cluster groups, thereby forming a kind of data partitioning. This is primarily done by identifying groupings in the hierarchy of the dendrogram.

K-means: this is an iterative clustering method where data are partitioned into k clusters, based on the Euclidean distance metric (Likas et al., 2003; Nsugbe et al., 2020). The algorithm is centered on the Expectation-Maximization (E-M) where the E step involves the assignment of clusters using the objective function assuming a random initialization $\sum_{i=1}^{m} \sum_{k}^{K} argmin_j ||x^i - \mu_k||^2$, where x^i represents a specific data point and μ_k is a cluster centroid mean; the M-step is an update phase for the cluster centroid represented using $\mu_k = \frac{\sum_{i=1}^{m} w_{ik} x^i}{\sum_{i=1}^{m} w_{ik}}$ w_{ik}, which is a binary variable used to indicate whether a data point belongs in a specific class (Likas et al., 2003; Nsugbe et al., 2020).

3.2.2. Supervised learning

Naïve Bayes (NB) classifier: the NB is a supervised probabilistic learner which is capable of learning data classes from provided labels; it builds on the classical Bayes rule which is mathematically expressed as follows (Nsugbe, 2022; Rish, 2001):

$$P(H_k|E) = \frac{P(E|H_k)P(H_k)}{P(E)} \tag{10}$$

where H_k represents the number of classes, $P(E|H_k)$ is a joint probability with prior H_k, $P(H_k)$ is the initial probability held by the hypothesis, $P(E)$ is a variable for the notion that the training data will potentially be observed from the information provided within the feature vector, and $P(H_k|E)$ is the posterior probability that carries a confidence level of the hypothesis after observing the training data (Nsugbe, 2021; Rish, 2001). The NB makes an assumption of the underlying distribution of the prior data $P(x_i|H_k)$, and a mathematical framework of how data are sorted into the various classes for a sample two class problem, H_a and H_b, assuming class H_a is more likely:

Table 1
Results of the NB classification

Accuracy	Sensitivity	Specificity	Area under the curve
95.80 ± 0.26	97.90 ± 032	93.80 ± 0.42	95.90 ± 0.32

Table 2
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

$$P(H_a)\prod_{i=1}^{N}P(x_i|H_a) > P(H_b)\prod_{i=1}^{N}P(x_i|H_b) \rightarrow P(H_a|x) > P(H_b|x) \tag{11}$$

where, for a specific class, we have

$$\boldsymbol{H} = argmaxP(H_k)\prod_{i=1}^{N}P(x_i|H_b) \tag{12}$$

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

As mentioned, the classifier allows for the expression of an associated probabilistic certainty alongside a class prediction, which in this area of research could prove useful in projecting an idea and inference platform for the potential extent of cancer within a patient.

4. Results

4.1. Unsupervised learning

The credibility and accuracy of the various clustering methods were calculated by comparing the expected number of points in a cluster against the algorithm estimated number and expressed as a percentage. The AHC produced an accuracy of 38%, while the k-means produced an accuracy of 65%. The results suggest that the k-means algorithm is the more suited unsupervised learning approach to be used for this kind of dataset. Due to this, the labels provided by the k-means algorithm were used for the training of the NB algorithm. As the clustering accuracy of the k-means was only as high as 65%, it can be seen that the algorithm produced an unbalanced class labeling; therefore, the SMOTE algorithm was further applied to produce a class balancing effect to the dataset ahead of the supervised learning stage.

4.2. Supervised Learning

As mentioned, the labeled samples from the prior stage were passed on to the next stage involving the NB for supervised learning (with SMOTE applied for class balancing). For the characterization of the performance of the trained NB classifier, four key metrics were employed as used in previous studies, namely the classification accuracy (CA), sensitivity (Sens), specificity (Spec), and area under the curve (AUC) (Nsugbe et al., 2021).

The results of the NB can be seen in Table 1, where it can be noted that the results produced were all above 90%, insinuating good learning capability from the NB model when it comes to learning from samples that have been labeled via an automated process clustering process.

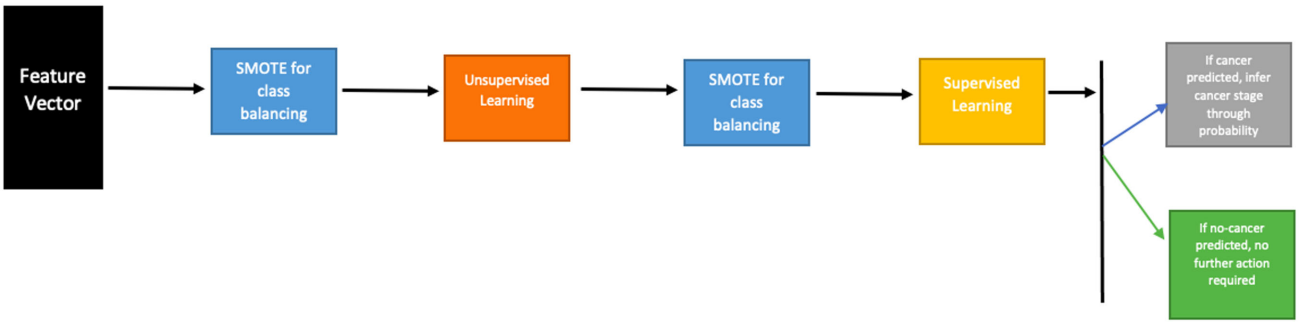
For the application of a probabilistic-based classification system, as the NB associates a probability with each prediction it would be expected that an associated probability score be produced when there is a predicted positive for cervical cancer. Following this, a probabilistic banding system can be put in place to associate the severity of each probabilistic diagnosis with an associated cancer stage. An example of this was proposed in a previous study (Nsugbe, 2022) and can be seen in Table 2.

Note that the probability–cancer stage association could also be automated using fuzzy logic approaches that could even allow for further partitioning of the probability values, while linking them to a potential cancer stage in order to inform care strategies (Mahfouf et al., 2001). A diagrammatic flow diagram of the complete training and learning process adopted as part of this study can be seen in Figure 2.

5. Conclusion

Cervical cancer is a deadly cancer with a high mortality rate worldwide, and with relatively undiagnosable symptoms in its early stages. This study is based around the design of a prediction machine intended to sit within a clinical setting involving human–machine interaction for enhanced cancer care treatments. Using key cervical cancer features, this work has explored the combination of unsupervised learning methods to partition and label data with no external intervention, prior to passing on to a

Figure 2
A diagrammatic flow of the complete training and learning process



probabilistic supervised learning method capable of grading the extent of a cancer, where predicted as a positive.

The AHC and k-means unsupervised learning methods were used for the automated partitioning and labeling of the data samples, where the k-means was seen to outperform the AHC method. The labels provided by the k-means were used to train the NB algorithm which is capable of producing probabilities associated with its predictions, where a high accuracy of 90%+ was produced across all classifier metrics. Where a sample is predicted to be cancerous, a banding for the probabilities was also proposed, as was implemented in a previous study (Nsugbe, 2022), this can serve as a template which can be further expanded and automated using rule-based systems such as fuzzy logic.

Acknowledgments

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

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

References

American Cancer Society. (2018). Cancer facts & figures 2018. Retrieved from https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2018.html

Araujo, R. P., & McElwain, D. L. S. (2004). A history of the study of solid tumour growth: The contribution of mathematical modelling. *Bulletin of Mathematical Biology*, *66*(5), 1039–1091. https://doi.org/10.1016/j.bulm.2003.11.002

Blagus, R., & Lusa, L. (2013). SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*, *14*, 106. https://doi.org/10.1186/1471-2105-14-106

Bora, K., Chowdhury, M., Mahanta, L. B., Kundu, M. K., & Das, A. K. (2016). Pap smear image classification using convolutional neural network. In *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing*, 1–8. https://doi.org/10.1145/3009977.3010068

Chandran, K. P., & Kumari, U. V. R. (2015). Improving cervical cancer classification on MR images using texture analysis and probabilistic neural network. *International Journal of Science, Engineering and Technology*, *4*, 3141–3145.

Davidson, I., & Ravi, S. S. (2005). Agglomerative Hierarchical Clustering with Constraints: Theoretical and Empirical Results. In A. M. Jorge, L. Torgo, P. Brazdil, R. Camacho, & J. Gama (Eds.), *Knowledge discovery in databases: PKDD 2005* 3721, 59–70. https://doi.org/10.1007/11564126_11

Day, W. H. E., & Edelsbrunner, H. (1984). Efficient algorithms for agglomerative hierarchical clustering methods. *Journal of Classification*, *1*(1), 7–24. https://doi.org/10.1007/BF01890115

El-Moselhy, E. A., Borg, H. M., & Atlam, S. A. (2016). Cervical cancer: sociodemographic and clinical risk factors among adult Egyptian females. *Journal of Oncology Research and Treatment*, *1*(1), 106.

Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*, *521*(7553), 452–459. https://doi.org/10.1038/nature14541

Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. *Future Generation Computer Systems*, *102*, 643–649. https://doi.org/10.1016/j.future.2019.09.015

IBM Cloud Education. (2021). What is strong AI?. Retrieved from https://www.ibm.com/cloud/learn/strong-ai

Khanam, M., Mahboob, T., Imtiaz, W., Ghafoor, H., & Sehar, R. (2015). A survey on unsupervised machine learning algorithms for automation, classification and maintenance. *International Journal of Computer Applications*, *119*(13), 34–39. https://doi.org/10.5120/21131-4058

Likas, A., Vlassis, N., & J. Verbeek, J. (2003). The global k-means clustering algorithm. *Pattern Recognition*, *36*(2), 451–461. https://doi.org/10.1016/S0031-3203(02)00060-2

Mahfouf, M., Abbod, M. F., & Linkens, D. A. (2001). A survey of fuzzy logic monitoring and control utilisation in medicine. *Artificial Intelligence in Medicine*, *21*(1-3), 27–42.

Malli, P. K., & Nandyal, D. S. (2017). Machine learning technique for detection of cervical cancer using k-NN and artificial neural network. *International Journal of Emerging Trends & Technology in Computer Science*, *6*(4), 145–149.

MathWork. (2021). Agglomerative hierarchical cluster tree—MATLAB Retrieved from https://uk.mathworks.com/help/stats/linkage.html

MedlinePlus. (2021). Cervical cancer. Retrieved from https://medlineplus.gov/ency/article/000893.htm

Nsugbe, E. (2022). Towards the use of cybernetics for an enhanced cervical cancer care strategy. *Intelligent Medicine*, *2*(3), 117–126. https://doi.org/10.1016/j.imed.2022.02.001

Nsugbe, E., Obajemu, O., Samuel, O. W., & Sanusi, I. (2021). Enhancing care strategies for preterm pregnancies by using a prediction machine to aid clinical care decisions. *Machine Learning with Applications*, *6*, 100110. https://doi.org/10.1016/j.mlwa.2021.100110

Nsugbe, E., Samuel, O. W., Asogbon, M. G., & Li, G. (2020). A self-learning and adaptive control scheme for phantom prosthesis control using combined neuromuscular and brain-wave bio-signals. *Engineering Proceedings*, *2*(1), 59. https://doi.org/10.3390/ecsa-7-08169

Powathil, G. G., Adamson, D. J. A., & Chaplain, M. A. J. (2013). Towards predicting the response of a solid tumour to chemotherapy and radiotherapy treatments: Clinical insights from a computational model. *PLoS Computational Biology*, *9*(7), e1003120. https://doi.org/10.1371/journal.pcbi.1003120

Rish, I. (2001). An empirical study of the Naïve Bayes classifier. In *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, *3*(22), 41–46.

Saha, A., Chaudhury, A. N., Bhowmik, P., & Chatterjee, R. (2010). Awareness of cervical cancer among female students of premier colleges in Kolkata, India. *Asian Pacific Journal of Cancer Prevention*, *11*(4), 1085–1090.

Singh, A., Thakur, N., & Sharma, A. (2016). A review of supervised machine learning algorithms. In *2016 3rd International Conference on Computing for Sustainable Global Development*, 1310–315.

Tseng, C. J., Lu, C. J., Chang, C. C., & Chen, G. D. (2014). Application of machine learning to predict the recurrence-proneness for cervical cancer. *Neural Computing and Applications*, *24*, 1311–1316. https://doi.org/10.1007/s00521-013-1359-1

UCI Machine Learning Repository. (2021). Cervical cancer (risk factors) data set. Retrieved from https://archive.ics.uci.edu/ml/datasets/Cervical+cancer+%28Risk+Factors%29

Unni, P., & Seshaiyer, P. (2019). Mathematical modeling, analysis, and simulation of tumor dynamics with drug interventions. *Computational and Mathematical Methods in Medicine*, *2019*. https://doi.org/10.1155/2019/4079298

Watanabe, Y., Dahlman, E. L., Leder, K. Z., & Hui, S. K. (2016). A mathematical model of tumor growth and its response to single irradiation. *Theoretical Biology & Medical Modelling*, *13*, 6. https://doi.org/10.1186/s12976-016-0032-7

Wu, W., & Zhou, H. (2017). Data-driven diagnosis of cervical cancer with support vector machine-based approaches. *IEEE Access*, *5*, 25189–25195. https://doi.org/10.1109/ACCESS.2017.2763984

Zhang, L., Lu, L., Nogues, I., Summers, R. M., Liu, S., & Yao, J. (2017). DeepPap: Deep convolutional networks for cervical cell classification. *IEEE Journal of Biomedical and Health Informatics*, *21*(6), 1633–1643. https://doi.org/10.1109/JBHI.2017.2705583

How to Cite: Nsugbe, E. (2023). A Pilot on the Use of Unsupervised Learning and Probabilistic Modelling Towards Cancer Extent Prediction. *Artificial Intelligence and Applications* 1(3), 155–160, https://doi.org/10.47852/bonviewAIA2202308