Turkish Journal of Computer and Mathematics Education Vol.12 No 13 (2021), 5263-5272

Research Article

Expert Advice- Disease prediction using Machine Learning

Kamran Afaq
Dept of Computer Science & Engineering
Galgotias University
Greater Noida, India
kamranakapoet@gmail.com

Jeeshan Uddin
Dept of Computer Science & Engineering
Galgotias University
Greater Noida, India
jeeshu.uddin98@gmail.com
Atul Sharmaf
Dept of Computer Science & Engineering
Galgotias University
Greater Noida, India
atul05992@gmail.com

Article History: Received: 5 April 2021; Accepted: 14 May 2021; Published online: 22 June 2021

Abstract-
Cancer is a type of disease which have now become very frequent and causing very critical damages in recent years and continue to remain its streak as no reliable cure is found as of now and cure that are found are very costly . This leads to an increase in the number of deaths even if a patient is created with a lot of care and given many therapies.The only way to manage it and reduce its damage is to diagnose it at an early stage. CSE(Comp. Science & Engg.) is used in many Bioinformatics and Biomedical fields to carry out many analysis ,prediction and to create automated diagnostic tools to ease the manual work of the front line workers that are our Doctors,nurses and group practices and we think it could also be used to diagnose and prognose disease such as Cancer by training the machine in such a way that it could tell the probability of having it in your system. One of the leading fields which could help us in this process is better known as Machine Learning where various tools and techniques and modules and algorithms are available to forecast the probability of cancer on the basis of collected standard data attributes and using them to predict the probability of extent of diseases such as cancer. The Expert Advice is a research paper which implies Machine learning for self-analytics or (analysis through feedback and reward and award techniques) and prediction of illness like cancer and fortify diseases like cancer. In this paper, we have studied ,researched and surveyed various papers,journals, books and datasets to compare the efficacy of different algos of Machine Learning about cancer prediction based on the given databundel and their attributes like accuracy,probability of prediction comparing different versions of algo and their results..It would be appropriate on our side to state that we can certainly leverage algos like RandomForest (RF), Naïve Bayes (NB), Decision Tree (DT),SVM, FNN.,RBFN,KNN, etc. In order to reach the desired result of forecasting cancer disease based on the given attributes of data, the best and the most accurate practise and paradigm of computer
networks and application in order to fulfill our goals among all here is SVM, which is Support vector machine algorithm.

I. INTRODUCTION

Cancer is a very harmful disease that is prominent by the nature of the behaviour of cells inside the body which has no stable access and has no jurisdiction.. Cancer is not a singular disease but rather more than 100 prominent types of cell division is known as cancer. Cancer dismantles the human body gradually when cells start multiplying uncontrollably to form many clusters of tissue inside the human cell membrane called tumors.Tumors have the potential to lay out to other regions of the body.. Those parts are a very critical part of the human body as the nervous system, digestive glands or circulatory arteries.Infected regions of the human body release harmful hormones, causing behaviour changes in the body's health and alignment.. Mainty There are two types of cancer: Malignant and Benign. Malignant attacks and defects the cell into the surrounding tissues membrane and arteries. Cells multiply thoroughly and destroy the tissue that was into its surroundings that causes other problems by erecting another tumor to develop. So a malignant tumor can be threatening, life changing and destroying and more vicious and dangerous in its true form. Benign tumors are very less harsh then malign as its impact very much less damage but can form into a more dangerous cancer if their growth is at a high rate or they might become malignant after a certain amount of deviation and spreading in the body. Among all the forms of cancer, the very nasty and frequently found and common is Breast cancer. The main risk go-between of breast cancer include sex, overweightness, less physical indulgance in activity, intakes of drugs , hormonal misbalance during menopause, ionizing radio active waves.There could also be different reason for the cause of breast cancer our main motive and our emotional indulgance in the topics tell us to model a machine in sucha way that we could help the needy and by detecting the breast cancer before hand and warning them in advance.There is a need of to detect the dicance forehand and physicians cant tells this just by seeing the patients body so here the ML play a vital important role as we can model and train the data to have us a rough calculation of symptoms and extint of cancer probablity.

II. MACHINE LEARNING TECHNIQUES

Machine Learning (ML) is a submodule of Al. ML is practical in order to deduce the output that comes on the basis of deportment of data sets.The training process is divided into two parts: I On the basis of the dataset provided, the system's unknown dependencies must be determined, and (ii) if estimated dependencies are known, the system's new output must be forecast.In the field of biomedical research, DEEP AND NEURAL NETWORK BASED ML has a wide range of applications. For a given collection of data samples, ML models use a variety of approaches and algorithms to generalise the biological sample of n-dimensional domains.Below are given different areas of ML algorithms

1) *Supervised Learning (Classification Approach)*

It is a way of machine learning in which the the data which is required so that model can learn by itself is available forehand containing different attributes such as desired output for given input.Some of the well known application that is better known as algo for this methods are Decision Tree,ANN,,, SVM, KNN,. Gaussian these are some well known and some others can be made on hands too using mathematics logic.

2) *Unsupervised Learning (Clustering Approach)*

This is a little bit different path then supervised and this technique is used with in sink of supervised machine learning.As we take large data set we need to first identify the different group better known as cluster present in data and it automatically identifies the important attributes using principal component analysis and then we can go the supervised path for specific tasks.The output is not know beforehand some of the popular techniques are K-means,PCA, and lesser knowns are Latent variable model,Hebbian learning.

3) *Semi-Supervised Learning*

This is a techniques between supervised and unsupervised paradigm as its data sets contains the the labelled and the unlabelled attributes in the data.This is best for condition when there is a half known knowledge and is a great discovering techniques for the unknown in the data .This project and process data in a better ways .Few common algos are as follows .Self-training, Generative Models, S3VM, Graph-based method, Multi view Learning, Mixture model

4) *Reinforcement Learning*

Also known as Adaptive Dynamic Programming, an intelligent agent learns by communicating with its surroundings and observing the results of interaction and communications. e.g., Q-Learning, SARSA, Monte Carlo method, Temporal Difference method.

5) *Transduction*

It is a process of the traceback matrix types in which we rely on tracing back out results and deducting intelligence from there.This is a ML paradigm in which we use reverse psychology that is first concluding the result and then tracing it back to find problems and optimization . e.g., Partitioning transduction, Agglomerative transduction, Manifold transduction

6) *Learning to Learn*

Based on training given to the model from input data it tries to identify the new result of input when a new input is setted. Model Regression Network is an example of this type of network. We have agreed that a proper blend of machine learning algorithms for various purposes can be highly advantageous based on the aforementioned decisive terms. In the next paragraphs, the algorithms are described.

A. *Artificial Neural Network (ANN)*

The most intuitive way to learn is to learn the way a human brain thinks and react and how our neurons play a major part in it.So on this assumptions the mathematicians have derived the techniques knows as the Artificial neural network.Neural Nets progress in such a way in a system that are dynamic and allows to reroute the structure of neurons in initial modelling phase too. Relationships between neurons can be easily formed as associations .They can determine the matrix and moderately deploy the corresponding clusters in relation to the data.The ANN learning process makes it easier to separate data and classify patterns. The neurons can be drawn and connected to build network layers, which are layers of neurons within and between layers.The category of ANN are -

.

1.Single layer FF network,
2.Multilayer FF network and lesser and more knows once are
3.Single node with its own feedback
4.Multilayer recurrent network

B. Decision Tree

Decision Tree as its same sounds is a graph based answering algo which contains the weightage of nodes and which ultimately decides which way to proceed.Its job is to find answers to problems based on criteria that aid decision-making, and the solutions are displayed graphically for easier comprehension. Its tree starts at the top with a root node and then shows various branches as alternative solutions.. It gives a systematic implementation of problem solving methods and documentation processes . We can choose a probable answer from a list of options and discover other possibilities.

C. Random Forest

To create a forest, the Random Forest Algorithm generates a large number of decision trees. These trees are reshaped based on a random selection of data and variables. Another classifier called ensemble classifier is utilised to enable this model, which begins by defining a crucial set of qualities to create each decision tree.It calculates the outcome of numerous decision trees. The ultimate selection is based on the highest number of votes from a large number of trees.

D. Support Vector Machine

As the name suggest support vector plays a key role here in order to classify records to the groups they belongs.Support vector parameter is the a parameter by which we can easily classify the clusters of records.Its make use of regressionThere is a dividing plane know as hyperplane to have bundle of input data where SVM divides the data values into two division as the separation is done on the basis of labelled data in the best intuitive way. The distance between the hyperplane and the nearest data point from either set of given datasets is determined.

E. k-Nearest Neighbour (kNN)

As the name suggests it takes the concept of full time neighbours very seriously.

KNN is a data discretion algorithm that filters a new variable to be within the existing case within a defined plane by calculating the nearest values using various mathematical assumptions and formulas with similar features of the kind.. The value of k (where k is a user-specified constant) would detect all similar existing records with comparable attributes with the new case and surround all the cases, allowing a cluster to identify the new attributes for the similar category.

F. Naïve Bayes Classifier

The NB classifier with self-rule assumptions between predictors is described by Bayes. Statistical dividers based on probabilistic computations are known as Bayesian classifiers. This is especially useful for large data sets where segregation can be done based on data points of different types and properties.The Bayesian theorem's goal is to categorise the class indicator for a given tuple.

.

G. K-means Algorithm

K-means is used for cluster components within the data sets.. Here k is a predefined constant that represents the number of clusters or groups we can say made for a looping method.The many K-Means clustering tools crawl through web pages, search them, and then cluster them by studying similarities and determining the advanced search rate of the search results in any search engine such as Google, Yahoo, and others. This is especially useful for huge datasets, where filtration decisions can be made based on data points from several classes and properties, for example.

III. FEASIBILITY ANALYSIS

Financial Feasibility -

As we are developing a project that is all based on web application and web hosting, we are using resources that are freely available which will ultimately make our project cost very low and budget friendly, since most part of the software is made from free resources no extra cost will be imposed on users which will lead to more affordability of our application ultimately leading to mass approach. Bandwidth required for the operation of this application is very low, which signifies that the project is financially feasible

IV. RELATED WORKS/ SURVEY

Sung-Bae Cho and colleagues used a variety of characteristics and classifiers to pick retrieved genes from noisy microarrays. They used three datasets: Leukemia cancer, Colon cancer, and Lymphoma cancer, with 72, 62, and 47 samples, respectively. They used Pearson's and Spearman's correlation coefficients, as well as Euclidean distance, information gain, and other techniques.For feature selection, mutual information and signal to noise ratio are used. They employed MLP, kNN, SVM, and SOM for classification. They tested all of the datasets and found that the greatest accuracy result was 97.1 percent on the Leukemia dataset using all of the classifiers listed above.PSO was utilised by Rui Xu et al. [3]to use gene expression data to predict patient survival To reduce dimensionality, PSO uses Probabilistic NN. PSO/experimental PNN's results on a B-cell Lymphoma dataset of 240 samples were more accurate in predicting survival, with up to 80% accuracy.. In microarray data, Mehdi Pirooznia et al. investigated a variety of classification and feature selection strategies for expressed genes. They were able to determine the efficacy of several classification SVM, Radial Basis Function, Multi-Layer Perceptron, DT, and RF are examples of approaches.. SVMRFE, Chi-Squared, and Correlation-based feature selection were also used to assess the efficiency of the feature selection methods (CFS). Finally, the authors discovered that SVM-RFE To identify the important genes, we used feature selection approaches that were 100 percent accurate. In order to classify high-dimensional cancer microarray data, Barnali Sahu et al. suggested a novel methodology based on feature selection methods. For optimization, this method employs one of the filtering techniques: signal-to-noise ratio (SNR) and PSO.They showed that when PSO is used in conjunction with SVM, k-NN, and PNN, it produces better results. They provided a dataset that included 72 occurrences of leukaemia with 7129 genes, 62 instances of colon cancer with 2000 genes, 77 instances of DLBCL with 6817 genes, and 97 instances of breast cancer with 24481 genes. They were impressed with the accuracy that PSO and other classifiers gave 100% in Breast cancer cases,Gunavathi et al. chose informative genes for categorization from thousands of microarray genes.Swarm intelligence techniques such as PSO, Cuckoo Search, Shuffled Frog Leaping, and Shuffled Frog Leaping with Levy Flight are used to find the informative genes (SFLLF). The samples are classified using the K-Nearest Neighbor (k-NN) classifier. The best result produced using the SFLLF feature of the k-NN classifier feature selection methods.K. Ridges Biomedic Data Repository provided the information. Various datasets, such as CNS, DLBCL, Lung Cancer, Ovarian Cancer, Prostate Outcome, AML/ALL, Colon Tumor, and others, are used to test various swarm intelligence algorithms, such as PSO, CS, SFL, and SFLLF, in order to achieve the highest level of accuracy. The best result is outperformed by SFLLF. Yasodha et al. used a data mining methodology on a huge dataset to discover the valuable knowledgeTo determine the data dependency and decrease the feature set contained in the data set, rough set theory was used. The Hybrid Particle Genetic Swarm Optimization method is used to optimise several aspects of ovarian cancer at various stages. Using an optimised feature set, a multi-class SVM is used to categorise normal or different stages of ovarian cancer.The datasets come from the TCGA ovarian cancer portal (http://tcga-data.Nci.nih.gov/), which contains 12042 genes with 493 occurrences. The classifiers Multiclass SVM, ANN, and Nave Bayes examined their separate trial results of 96 percent, 93 percent, and 90 percent accuracy. The other dataset comes from NUH Singapore (http://www.nuh.com.sg/#) and contains 172 instances of blood tests with 28 features.tool for mining Madeeh Nayer Elgedawy used the Nave Bayes, SVM, and RF machine learning techniques. RF is the most appropriate and useful method, achieving 99.42 percent accuracy, while SVM and NB achieve 98.8 percent and 98.24 percent accuracy, respectively. SVM and SVM ensemble classifier were employed by Min-Wei Huang et al. in breast cancer research.They also used the boosting approach and RBF kernel-based SVM to forecast accuracy, with assessment metrics such as F-measure, ROC curve, and others derived on training data to develop a model. He discovered that an RBF kernel SVM ensemble based on the boosting method outperformed other classifiers in terms of accuracy.Urun Dogan et al. used numerous datasets to create a unified view of multi-class SVM. They mathematically explained multi-class loss function, margin function, aggregator operator, and multi-class loss function fisher consistency. The linear and non-linear SVM accuracy results are also provided. The supervised and unsupervised machine learning classifiers are compared for the breast cancer dataset.They discovered that supervised machine learning requires the use of a training dataset, but unsupervised machine learning does not. Anomaly detection in the dataset is improved using robust one-class SVM and enhanced one-class SVM.

V. PROBLEM FORMULATION

Problem Statement -

To identify cancer symptoms at an early stage to save someone's life by using machine learning techniques and machine learning models on different dataset. Cancer is a disease that can be cured with a series of treatments such as fluid injections, radiation therapies, etc., but has no single medicine or process to cure it. In 2019, about 18 million new cases were reported along with 90.5 million cases reported till 2015. It caused about 8.8 million deaths (15.7% of deaths). Early detection and diagnosis can warn the patients in advance about the harmful development of cell division causing cancer. This paper tries to propose a life saving and life changing method to detect,diagnose and prognose cancer by employing techniques of modern computing power. In the last few decades computational methods have helped dynamically optimize the accuracy of cancer susceptibility, recurrence and survival prediction. A critical survey of the existing literature reveals that a good amount of work has been done in the application of machine learning technique to
develop predictive models for cancer recurrence in patients who had been observed for more than two years. Also, the accuracy of cancer detection probability outcome has boostly improved by 15%–20% the last few years, with the application of modern computing power and techniques. Several research have been mentioned in the journal and are based on different computation and algorithmic methods that could enable the premature cancer diagnosis and prognosis. In this Project we will be exploring various datasets, so that we can create a predictive model that can help diagnose patients and predict the extent of a cancer.

VI. OBSERVATION

Humans are afflicted with various types of cancers, each unique in behavior, symptoms and attributes. Their distinct behaviors create the need for specialized treatment and novel prognoses. The most common types of cancers have been listed below -

A. Breast Cancer

X-Rays can be used to identify malignant cells which manifest as a tumor in the shape of a lump. The cancer spreads continuously by growing into the surrounding tissue.

Table 1: Accuracy measured for breast cancer using ML techniques

Ref	**Types of cancer: Breast Cancer**		
	ML Technique	***Sample***	***Accuracy***
2013	Random Forest	699	99.82%
2015	DT-SVM	699	91%
2017	Random Forest	699	99.24%
	SVM		98.8%
	Naïve Bayes		98.24%
2013	Decision Tree	1189 (Iranian Center)	93.6%
	ANN		94.7%
	SVM		95.7%
2014	RepTree	286	71.32%
	Radial Basis Function Netwrok		73.77%
	Simple Logic		74.47%

Machine learning techniques are employed to prognosticate the diagnosis, using datasets of breast cancer from various resources as shown in the table above.

A few of these techniques are able to yield a 100% accurate prediction based on the sample count and various features available within the datasets.

B. Lung Cancer

Malignant cells of Lung Carcinoma, manifest and spread in the same way as Breast Cancer, they can also be identified by X-Rays. The accuracy of diagnosis has improved significantly in the past few years, with a 15-20% increase.

Table: Accuracy measured for breast cancer using ML techniques

Ref	**Types of cancer: Lung Cancer**		
	ML Technique	***Sample***	***Accuracy***
2003	SAMSON	62	93.6%
2008	Fuzzy Granular Support Vector Machine-Recursive Feather Elimination (FGSVM-RFE)	62	99.71%
2015	KNN using Biogrpahy-Based Optimization	62	80%

C. Colon Cancer

Colon Cancer originates in the end of the human digestive tract by affecting either the colon or the rectum. A primary symptom of this affliction is constant bloody stools. Both colon and rectum exhibit common symptoms.

The table below shows the different machine learning techniques that are used to foretell the possibility of cancer diagnosis along with the accuracy of their prediction.

Table 2: Accuracy measured for colon cancer using ML technique

Ref	**Types of cancer: Colon Cancer**		
	ML Technique	***Sample***	***Accuracy***
2003	SAMSON	62	93.6%
2008	Fuzzy Granular Support Vector Machine-Recursive Feather Elimination (FGSVM-RFE)	62	99.71%
2015	KNN using Biogrpahy-Based Optimization	62	80%

D. Ovarian Cancer

This type of cancer affects the female reproductive system. Ovum produced by the ovaries (reproductive gland) travels to the uterus where it is fertilized into a fetus. Using machine learning techniques on available Ovarian Cancer datasets, we are able to predict its diagnosis with varying degrees of accuracy shown in the table below.

Table: Accuracy measured for ovarian cancer using ML Technique

Ref	**Types of cancer: Ovarian Cancer**		
	ML Technique	***Sample***	***Accuracy***
2014	K-means with harmony search	150	97%
2014	SVM	493	96%
2015	ANN	493	93%
2015	Naïve Bayes	493	90%

E. Liver Cancer

The liver is the largest internal organ of the human body, and it regulates most chemical levels in the blood as well as bile output. After forming in the liver, malignant cells migrate into the bloodstream. Using various machine learning techniques on available Liver Cancer datasets, we are able to predict its diagnosis with varying degrees of accuracy shown in the table below.

Table 3: Accuracy measured for liver cancer using ML technique

Ref	**Types of cancer: Liver Cancer**		
	ML Technique	***Sample***	***Accuracy***
2011	Fuzzy neural network (FNN)	156	95.45%
2013	Particle Swarm Optimization	-	93.33%

Table 4: Various ML Techniques applied on cancer for 100% accuracy

Ref	***ML Technique***	***Types of Cancer***	***Sample***	***Accuracy***
2008	SVM-FREE	Lymphoma	25	100%
		Breast Cancer	84	
		Colon Cancer	45	
		Lung Cancer	72	
		Ovarian Cancer	39	
2011	SVM	Liver Cancer	156	100%
2012	PSO-KNN	Breast Cancer	97	100%
	PSO-SVM			
2014	SFLLF	Colon	62	100%
		Prostate	136	

Various feature selection techniques and their respective classifiers have been listed in the table above, based on previously conducted research in the avenue of cancer prediction. With the help of this table, we can also show that k-fold cross validation and instance of data are key in determining the dataset samples for accurate predictions. This inference is made by assessing the prediction patterns created by using different machine learning algorithms in cases of early onset cancer.

The collection and standardization of datasets is of paramount importance as well. Below is the list of attributes that are used in most of the cases:

Race
Marital Status
Clump Thickness
Uniformity of Cell Size
Uniformity of Cell Shape
Marginal Adhesion
Single Epithelial Cell Size
Histological type
Number of positive nodes
Number of nodes
Number of primaries
Lymph node involvement
Site specific surgery code
Primary site code

Bare Nuclei	Behavior code
Bland Chromatin	Extension of Tumor
Normal Nucleoli	Radiation
Mitoses	Stage of cancer
Class	Grade
Age	Menopause
Deg-malignant	Tumor-size
Breast-quad	Inv-nodes
Irradiation	Node-caps

VII. CONCLUSION

This paper represents an overall scenario of various ML classification algorithms used or can be used for the prediction of different types of cancers,probability of having a cancer of a certain type and standard data models have been used in this study is consider to be cover all the aspects of study in field of cancer detection covering all type and wide variety of cancer mutations such as ovarian cancer, breast cancer. A detailed list of procedures found by many researchers and scholars has been written here to solve the problems by various computational intelligence and computationally optimised techniques, tools and algorithms. The most successful approach identifying and predicting the symptoms and indications of having an cancer is SVM(Support vector machine) and with robust and high level implementation of SVM technique after training the model for prediction ,it gave up to 100% accuracy on a few number of training datasets which is not a perfect prediction in case with large data sets. However, other options are available for the possibilities of optimization of forecasting the cancer at an premature stage. There are many datasets available to explore more for the same. .Our research summary can be seen in the graph given below namely figure1.1 and figure 1.2

VIII. REFERENCES

1) World Health Organization. Available: https://www.who.int/news room/factsheets/detail/cancer

2) Fatima, M. and Pasha, M. (2017) Survey of Machine Learning Algorithms for Disease Diagnostic. Journal of Intelligent Learning Systems and Applications, 9, 1- 16. https://doi.org/10.4236/jilsa.2017.91001

3) Ghumbre S.U., Ghatol A.A. (2012) Heart Disease Diagnosis Using Machine Learning Algorithm. In: Satapathy S.C., Avadhani P.S., Abraham A. (eds) Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (INDIA 2012) held in Visakhapatnam, India, January 2012. Advances in Intelligent and Soft Computing, vol 132. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-27443-5_25

4) Cancer detection techniques using ML by sjjad, available at https://medium.com/datadriveninvestor/breast-cancer-detectionusing-machine-learning 475d3b63e18e#:~:text=Research%20indicates%20that%20most20experienced,achieved%20using%20machine%20learning%20techniques.&text=In%20his%20case%20study%2C%20our,pain%20from%20several%20cell%20images.

5) Nitin Mishra, Saumya Chaturvedi, Garima Pandey. (2020). Diagnosis of Cancer Using Machine Learning. *International Journal of Advanced Science and Technology*, *29*(7s), 5155 -. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/25802

6)T. Saba, A. Sameh, F. Khan, S.A. Shad, M. Sharif **Lung nodule detection based on ensemble of hand crafted and deep features** J Med Syst, 43 (12) (2019), p. 332

7) S. Iqbal, M.U.G. Khan, T. Saba, A. Rehman **Computer assisted brain tumor type discrimination using magnetic resonance imaging features** Biomed Eng Lett, 8 (1) (2017), pp. 5-28, 10.1007/s13534-017-0050-3

9)F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries** CA Cancer J Clin, 68 (2018), pp. 394-424

10)Y. Kurihara, S. Matsuoka, T. Yamashiro, A. Fujikawa, S. Mat ushita, K. Yagihashi, *et al.* **MRI of pulmonary nodules** Am J Roentgenol, 202 (2014), pp. W210-W216

11) C. White **New techniques in thoracic imaging** CRC Press (2001)

12) [Accessed by 2014]. Breast Cancer Surveillance Consortium. https://www.bcscresearch.org/data/mammography_d

13)S.L. MarieSainte, T. Saba, D. Alsaleh, A. Alotaibi, M. Bin **An improved strategy for predicting diagnosis, survivability, and recurrence of breast Cancer** J Comput Theor Nanosci, 16 (9) (2019), pp. 3705-3711

14)K. Ejaz, M.S.M. Rahim, A. Rehman, H. Chaudhry, T. Saba, A. Ejaz, *et al.* **segmentation method for pathological brain tumor and accurate detection using MRI** Int J Adv Comput Sci Appl, 9 (8) (2018), pp. 394-401

15)B. Tahir, S. Iqbal, M.U.G. Khan, T. Saba, Z. Mehmood, A. Anjum, *et al.* **Feature enhancement framework for brain tumor segmentation and classification** Microsc Res Tech (2019), 10.1002/jemt.23224

16)A. Mittal, D. Kumar, M. Mittal, T. Saba, I. Abunadi, A. Rehman, *et al.* **Detecting pneumonia using convolutions and dynamic capsule routing for chest X-ray images** Sensors, 20 (4) (2020), p. 1068

17)L.R. Bareiro Paniagua, D.N. Leguizamón Correa, D.P. Pinto-Roa, J.L. Vázquez Noguera, S. Toledo, A. Lizza **Computerized medical diagnosis of melanocytic lesions based on the ABCD approach** Clei Electron J, 19 (2016) 6-6

18) Z. Lu, Y. Bai, Y. Chen, C. Su, S. Lu, T. Zhan, *et al.* **The classification of gliomas based on a pyramid dilated convolution ResNet model** Pattern Recognit Lett (2020)

19)D.R. Nayak, R. Dash, B. Majhi, R.B. Pachori, Y. Zhang **A deep stacked random vector functional link network autoencoder for diagnosis of brain abnormalities and breast cancer** Biomed Signal Process Control, 58 (2020), p. 101860

20) D.R. Nayak, R. Dash, B. Majhi **Automated diagnosis of multi-class brain abnormalities using MRI images: a deep convolutional neural network-based method** Pattern Recognit Lett (2020)

21)I. Qureshi, M.A. Khan, M. Sharif, T. Saba, J. Ma **Detection of glaucoma based on cup-to-disc ratio using fundus images**
Int J Intell Syst Technol Appl, 19 (1) (2020), pp. 1 16, 10.1504/IJISTA.2020.105172 S. Abbasi, F. Tajeripour **Detection of brain tumor in 3D MRI images using local binary patterns and histogram orientation gradient** Neurocomputing, 219 (2017), pp. 526-535

22)I. Mehmood, M. Sajjad, K. Muhammad, S.I.A. Shah, A.K. Sangaiah, M. Shoaib, *et al.* **An efficient computerized decision support system for the analysis and 3D visualization of brain tumor** Multimed Tools Appl, 78 (2019), pp. 12723-12748

23)D. Das, L.B. Mahanta, S. Ahmed, B.K. Baishya, I. Haque **Automated classification of childhood brain tumours based on texture feature** Songklanakarin Journal of Science & Technology (2019), p. 41

24) M. Arikan, B. Fröhler, T. Möller **Semi-automatic brain tumor segmentation using support vector machines and interactive seed selection** Proceedings MICCAI-BRATS Workshop (2016), pp. 1-3

25)K. Ejaz, M.S.M. Rahim, U.I. Bajwa, N. Rana, A. Rehman **An unsupervised learning with feature approach for brain tumor segmentation using magnetic resonance imaging** Proceedings of the 2019 9th International Conference on Bioscience, Biochemistry and Bioinformatics (2019), pp. 1-7

26)A.A.A. Setio, F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S.J. Vn Riel, *et al.* **Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks**
IEEE Trans Med Imaging, 35 (2016), pp. 1160-1169

27) Q. Dou, H. Chen, L. Yu, J. Qin, P.A. Heng **Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection** IEEE Trans Biomed Eng, 64 (2016), pp. 1558-1567

28)W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, *et al.*

Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification
Pattern Recognit, 61 (2017), pp. 663-673

29)P.G. Tahoces, L. Alvarez, E. González, C. Cuenca, A. Trujillo, D. Santana-Cedrés, *et al.* **Automatic estimation of the aortic lumen geometry by ellipse tracking** Int J Comput Assist Radiol Surg, 14 (2019), pp. 345-355

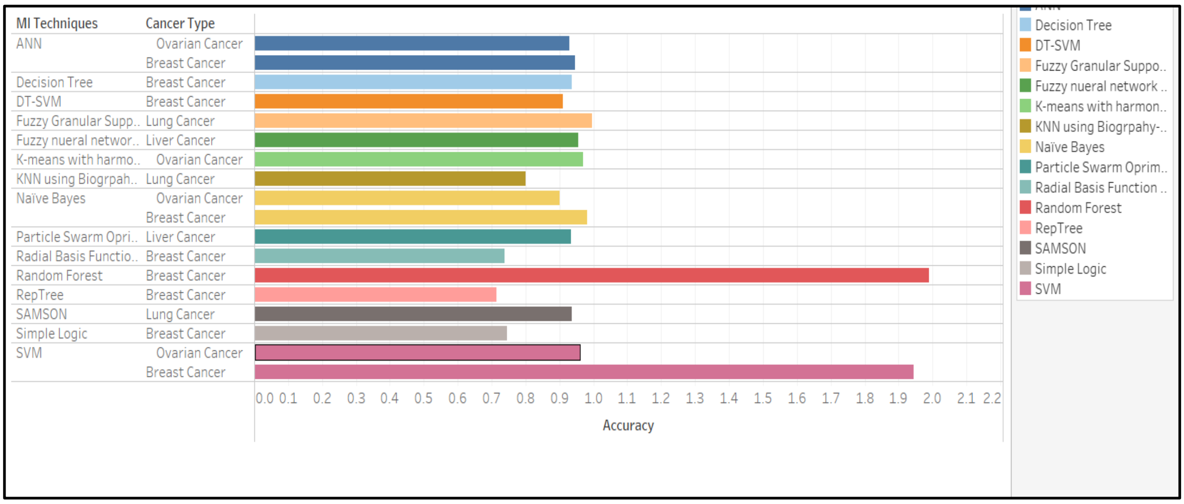

Figure 1: ML techniques vs Accuracy

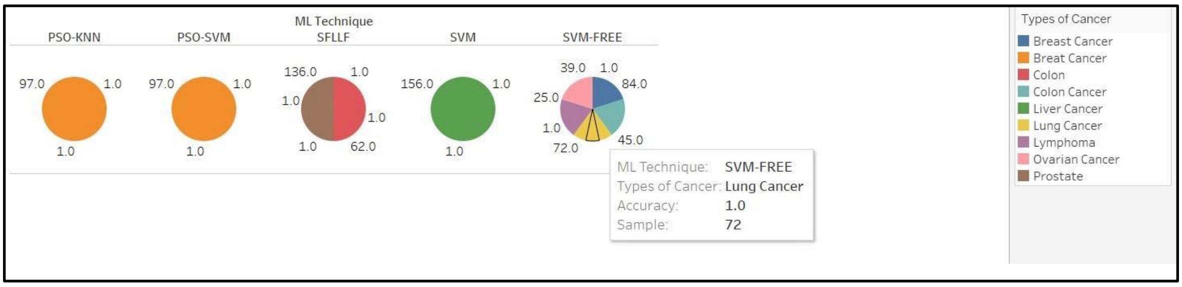

Figure 2: Distribution of sample data