International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Review of Theoretical Approach: Cancer Cells Identification by Machine Learning

Radhika Vishwakarma[1], Dr. Kailash Patidar[2] and Dr. Damodar Tiwari[3]

[1]*PG Student, Department of CSE, Bansal Institute of Science and Technology, Bhopal, M.P., India*
[2]*Associate Professor, Department of CSE, Bansal Institute of Science and Technology, Bhopal, M.P., India*
[3]*Professor, Department of CSE, Bansal Institute of Science and Technology, Bhopal, M.P., India*

ABSTRACT

The present paper based on Artificial intelligence (AI) and machine learning (ML) which is mostly used for diagnosis for cancer cell identification. Multidisciplinary collaboration is necessary for the creation of an ideal tool in order to guarantee that the right use case is satisfied and to conduct thorough development and testing before the tool is implemented into healthcare systems. The prospects and problems of AI and ML in cancer imaging will be covered in this work, along with suggestions for turning algorithms into broadly applicable tools and building the ecosystem required to support the field's expansion. For the purpose of early cancer detection and staging, the identification of metastatic cancer cells is crucial. Nevertheless, at the outset of the disease, it is very challenging to identify these cells from blood or biopsy samples. On aptamer functionalized substrates, it has been observed that cancer cells, and particularly metastatic cancer cells, exhibit markedly different morphological behavior from their healthy counterparts. Early cancer diagnosis can undoubtedly benefit from the speedy analysis of the data and the capacity to quantify the cell morphology for an instantaneous real-time feedback. In order to quantify and distinguish between the intricate bodily movements of malignant and non-cancerous cells, many feature vectors will be retrieved using machine learning algorithms. The present paper brief discussed about machine learning method and their classification of algorithms and led us to a robust review report the classification problem.

Keywords: *Cancer cell, Cell morphology, Aptamers, Machine learning, Cancer classification*

1. Introduction

One of the main causes of death in the globe is cancer. A better prognosis and successful course of therapy depend on an early and precise diagnosis. Conventional approaches to cancer detection frequently depend on genetic technologies and histological analysis, which can be laborious and prone to human error. Machine learning (ML) advances in recent years have created new opportunities for automated and more accurate differentiation of cancerous cells from normal cells.

Numerous fields of science, including medicine, are fast changing as a result of artificial intelligence (AI) and machine learning (ML). The term artificial intelligence (AI) describes the development of tools or machines that can mimic human thought and behaviour, while machine learning (ML) is a subset of AI in which tools or machines learn from data to produce predictions or classifications with or without human supervision [1]. The advent of high-performance computers in recent times has expedited the progress in these domains.

In the field of medicine, digital areas like imaging are poised to be early adopters of artificial intelligence (AI) and machine learning (ML). The imaging workflow, which includes stages such as image acquisition, reconstruction, interpretation, reporting, and result communication, operates in a digital environment. This facilitates the efficient capture of data for AI and ML applications. Notably, cancer imaging, which constitutes a significant part of many departments' workloads, is a prime candidate for the early implementation of these technologies by radiologists. This is largely due to the repetitive nature of tasks like cancer screening, where large volumes of normal images must be reviewed to identify anomalies, the tedious process of measuring tumors over time, and the laborious job of delineating tumors for disease segmentation. In fact, there are already several commercial products available in the cancer imaging sector aimed at improving work efficiency, minimizing errors, and boosting diagnostic accuracy.

1.1 AI Techniques

Artificial intelligence is the ability of machines, primarily computers, to behave like people. AI allows machines to carry out activities like problem-solving, learning, and face recognition, among others. If a machine is sufficiently knowledgeable about a task, it can function and behave like a human. Thus, knowledge engineering is crucial to artificial intelligence [1], [8], [14]. It is permitted to use the relationship between objects and properties while implementing knowledge engineering. The advantages and potential applications that may be realized with the synergy between AI and clinical decision support system disused by [3], [9]. Below is an explanation of a well-known artificial intelligence technique [20], [21], [22].

Machine Learning

Machine learning is a subset of artificial intelligence (AI) and computer science that concentrates on leveraging data and algorithms to replicate human learning, thereby progressively enhancing accuracy [2]. Machine learning algorithms are capable of self-improvement through training processes. Unlike traditional computing operations that strictly follow predefined steps and produce outputs based on specific inputs without error, machine learning involves scenarios where computers make decisions based on current data samples. In these cases, much like humans, computers can make mistakes during decision-making. Essentially, machine learning is about enabling computers to learn from data and experiences in a manner similar to the human brain. The primary goal of machine learning is to develop models that can self-train, recognize intricate patterns, and solve new problems using past data. Presently, machine learning algorithms are trained using three main approaches: supervised learning, unsupervised learning, and reinforcement learning [4], [7].

Supervised learning

Supervised Learning is a type of machine learning that follows a relatively straightforward approach. This method involves setting specific learning objectives beforehand. During the initial phase of machine training, the machine uses information technology to understand and learn these objectives. To gather fundamental data, learning is conducted in a supervised setting where the required knowledge is gradually acquired. Unlike other methods, supervised learning maximizes the machine's ability to generalize knowledge. Once the system learning is complete, it can assist in solving classification and regression problems in a highly structured manner. Common techniques used in supervised learning include Bayesian Networks (BN), Support Vector Machines (SVM), and k-Nearest Neighbors (KNN). Because the learning process is goal-oriented, it follows a predictable pattern, resulting in more systematic and organized learning outcomes.

Unsupervised learning

Unsupervised learning, in contrast to supervised learning, involves a process where the machine learns without any labeled guidance. In unsupervised learning, the machine independently analyzes data without pre-defined instructions. The approach involves letting the machine understand basic concepts and then giving it the freedom to explore and learn a variety of topics autonomously, including those related to fundamental principles, like tree structures. This ongoing, stage-by-stage learning process broadens the scope of the machine's knowledge. Currently, unsupervised learning encompasses algorithms such as deep belief networks and auto-encoders [19]. These methods are particularly effective for solving clustering problems and have significant applications across various industries.

Reinforcement Learning

Besides supervised and unsupervised learning, reinforcement learning is another significant method in machine learning. Reinforcement learning involves systematic learning of specific content by utilizing data collected over time. In practice, it processes and organizes feedback from previous actions to create a continuous loop of data processing. Overall, reinforcement learning is a method that enhances data collection through statistical analysis and dynamic learning, primarily used to address robot control issues. Notable algorithms in this field include Q-learning and Temporal Difference learning. Unlike supervised learning, reinforcement learning is a behavioral model that does not rely on sample data for training. Instead, it learns through trial and error, where a series of successful outcomes reinforces the development of optimal recommendations or policies for a given problem.

1.2 Machine Learning Techniques

Machine learning (ML) is a scientific discipline that focuses on the use of statistical models and algorithms by computer systems to perform tasks without explicit programming, instead relying on patterns and reasoning. It is a subset of computerized thinking. ML algorithms create a numerical model based on sample data, known as “training data,” to make predictions or decisions without direct programming for the task. This approach is utilized in situations where traditional programming would be impractical or impossible, such as in email filtering and computer vision [12], [13].

Machine learning is closely associated with computational statistics, which is concerned with making predictions using computers. The study of scientific optimization enriches the field with methodologies, theories, and applications related to machine learning. A specific area of machine learning, information mining, focuses on exploratory data analysis through unsupervised learning. In the context of solving business problems, machine learning is often referred to as predictive analytics.

1.3 Classification of Algorithms

There are many types of classification algorithms and machine learning, such as:

1. Decision trees
2. Naive Bayes
3. Logistic regression

4. K-nearest neighbour
5. Support vector machines
6. Random Forest

In machine learning, classification is a supervised learning technique used to categorize unknown items into a predefined set of classes. It involves learning the relationship between a set of feature variables and a target variable of interest. The target variable in classification is a categorical factor with distinct values. Given a set of training data points with known labels, the classification process determines the class label for an unlabeled test case. Various classification algorithms are used in machine learning, including decision trees, naive Bayes, linear discriminant analysis, k-nearest neighbor, logistic regression, neural networks, and support vector machines.

Decision Tree

Recursive partitioning is a technique used in the construction of decision trees for data classification. It involves dividing the training set into discrete nodes, each of which houses the majority of the data in a given category. One can create a decision tree by taking each attribute one at a time.

- First, choose an attribute from our dataset.
- Calculate the significance of the attribute in the splitting of the data.
- Next, split the data based on the value of the best attribute,
- Then go to each branch and repeat it for the rest of the attributes.
- After building this tree, you can use it to predict the class of unknown cases.

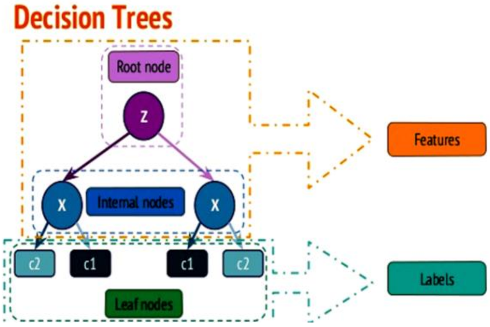


Fig. 1 Decision tree

Decision trees are about testing an attribute and branching the cases based on the result of the test:

1. Each internal node corresponds to a test
2. Each branch corresponds to a result of the test
3. Each leaf node assigns a patient to a class

Naive Bayes

A supervised computation known as the Naïve Bayes classifier uses Bayes hypothesis to classify the dataset. The standard or numerical concept used to calculate likelihood is known as the Bayes hypothesis. The fundamental premise of the Bayes theorem is that autonomous variables and certain free assumptions are needed for the Bayes hypothesis to work.

Bayes theorem on Mathematical Representation:

P (A\B) = Here,

- P (A) = independent probability of A (prior probability)
- P (B) = independent probability of B
- P (B\A) = conditional probability of B given A (likelihood)
- P (A\B) = conditional probability of A given B (posterior probability).

A great and easy computation for predictive modelling is Naive Bayes. This approach is the most effective and practical classification computation that can handle large amounts of messy, non-linear, and inferior data. Specific naïve and Bayes are the two sections that make up naïve. A naïve classifier presumes that the existence of a particular element in a class is independent of the existence of any other element.

Logistic Regression

Logistic regression is a classification algorithm designed for categorical variables. It is similar to linear regression but aims to predict discrete outcomes, such as 0 or 1, yes or no, instead of continuous values. Dependent variables in logistic regression need to be continuous or, if categorical, must be converted into dummy or indicator variables. This transformation involves converting categorical data into numerical values. Logistic regression is applicable for both binary and multi-class classification tasks [6]. A key element in logistic regression is the sigmoid function.

Logistic regression is useful in the following scenarios:

1. When the target variable is categorical or binary: For example, predicting outcomes like zero/one, yes/no, true/false, churn/no churn, positive/negative, etc.
2. When you need probability predictions: Logistic regression provides a probability score between zero and one for a given data sample.
3. When your data is linearly separable: The decision boundary in logistic regression can be a line, plane, or hyperplane. The classifier assigns data points on one side of the boundary to one class, and those on the opposite side to another class.
4. When you want to understand feature impact: Logistic regression allows for the selection of the best features by evaluating the statistical significance of its model coefficients or parameters.

K-Nearest Neighbors Algorithm (KNN)

The K-Nearest Neighbors algorithm is a supervised learning classification algorithm that use a large number of named points to determine how to name new points. Cases are categorized by this computation according to how similar they are to one another. Information points that are close to one another are referred to as neighbors in K-Nearest Neighbors. The foundation of K-Nearest Neighbors is this perspective. Consequently, a measure of how distinct two situations are from one another is their distance [5], [6].

In a classification problem, the K-Nearest Neighbors algorithm works as follows:

Pick a value for K-

- Calculate the distance from the new case hold out from each of the cases in the dataset
- Search for the K-observations in the training data that are nearest to the measurements of the unknown data point
- Predict the response of the unknown data point using the most popular response value from the pick a value for K**.**
- Low estimation of K results in a profoundly complex model and may result in over fitting.
- High estimation of K such as K equals 20, at that point the model becomes excessively summed up.
- Solution is to reserve a piece of your information for testing the exactness of the model. When you've done as such, choose K equals one and afterward use the preparation part for modeling and compute the precision of prediction using all samples in your test set. Rehash this process increasing the K and see which K is best for your model.

Calculate similarities between two data points-

- Use a specific type of Minkowski distance to calculate the distance of these two customers, which is the Euclidean distance.

Euclidean distance:

$$Dis(x_1, x_2) = \sqrt{\sum_{i=0}^{n}(x_{1i} - x_{2i})^2} \qquad (1)$$

Nearest neighbors' analysis can also be used to compute values for a continuous target. In this situation, the average or median target value of the nearest neighbors is used to obtain the predicted value for the new case.

Evaluation metrics in classification

Evaluation metrics explain the performance of a model.

Jaccard Index (or Jaccard similarity coefficient) is defined as size of the intersection divided by the size of the union of two label sets (picture a Venn diagram).

For example, if y is actual labels (10 data points) and $\hat{y}$ is the predicted labels (10 data points), and 8 data points are accurately predicted by the model, than;

$$J(y,\hat{y}) = \frac{|y \cap \hat{y}|}{|y \cup \hat{y}|} = \frac{8}{10+10-8} = 0.66 \qquad (2)$$

Support Vector Machine

Support Vector Machine (SVM) is a supervised algorithm that classifies data by dividing a dataset into two or more classes using a separating boundary. SVM operates by [16]:

- Mapping data into a high-dimensional feature space: This process, known as kernelization, allows data points to be classified even when they are not linearly separable in their original space.
- Finding a separating boundary: A separator is identified between the classes, and the data is transformed so that this separator can be represented as a hyperplane.
- Classifying new data points: The characteristics of new data are used to predict the class to which a new record belongs based on the previously established hyperplane.

The kernel function is a mathematical function used to map data into a higher-dimensional space, making a non-linearly separable dataset linearly separable. Kernel functions come in various forms, such as linear, polynomial, Radial Basis Function (RBF), and sigmoid.

1. Take a 1D linearly inseparable dataset (x)
2. Define a function to map to 2D, $\phi(x) = [x,x^2]$

A sensible choice for the optimal hyperplane is one that maximizes the distance, or margin, between the two classes. The data points closest to the hyperplane are known as support vectors. Only these support vectors are crucial for determining the hyperplane, allowing us to disregard other data points. The goal is to find a hyperplane that has the largest possible distance from the support vectors. This hyperplane is derived from training data using an optimization technique that aims to maximize the margin. Similar to many other optimization problems, this can be solved using gradient descent, although that is beyond the scope of this discussion.

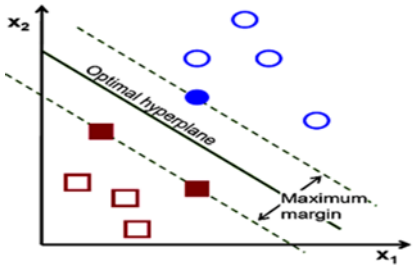


Fig. 2 Support vector machine

Random Forest

Random Forest is a flexible, user-friendly machine learning algorithm that, in most cases, yields exceptional results even in the absence of hyper-parameter tweaking. Due to its ease of use and propensity to be applied to both regression and classification tasks, it is also one of the most popular algorithms. You will now understand the operation of the random forest computation as well as a number of other important aspects of it.

One significant advantage of random forests is their propensity to be applied to both regression and classification problems, which form the foundation of the majority of machine learning systems in use today.

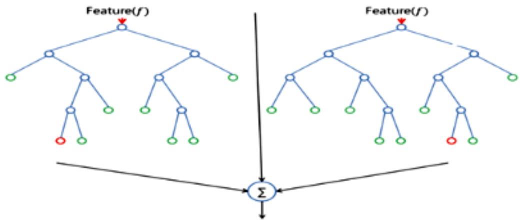


Fig. 3 Random forest

When splitting a hub using Random Forest, only a random subset of the features is taken into account. Additionally, you can use random thresholds for each component rather than looking for the optimal thresholds (like a traditional decision tree does) to further create increasingly unpredictable trees.

1.4 Problem Statement

The differentiation between cancer cells and normal cells is a challenging task that requires expertise and precision. Manual analysis is not only labor-intensive but also susceptible to subjective biases. There is a need for an automated system that can reliably and accurately identify cancerous cells, thereby aiding pathologists and reducing diagnostic errors.

2. Literature Review

Ather et al., 2020, the study focuses on AI in lung nodule detection and classification, highlighting the lesser research in other lung diseases like COPD and pulmonary embolisms. Despite advancements in convolutional neural networks for thoracic imaging, clinical adoption remains limited. The review discusses recent AI developments for pulmonary nodules and the evolving role of radiologists [1].

Bitencourt et al., 2020, this retrospective study examined 311 HER2-positive breast cancer patients undergoing neoadjuvant chemotherapy (NAC), using clinical and MRI radiomic features combined with machine learning to predict pathologic complete response (pCR). The study achieved high sensitivity, specificity, and accuracy in predicting HER2 heterogeneity and pCR, demonstrating the potential of integrating clinical and radiomic data for better outcomes [2].

Bizzo et al., 2019, The paper explores the integration of AI with clinical decision support (CDS) systems in radiology, emphasizing the potential to improve workflows, guide care pathways, and enhance decision-making for both radiologists and referring physicians. AI-enhanced CDS systems could significantly impact patient care by providing more comprehensive and personalized imaging recommendations [3].

Bukowski et al., 2020, the study reviews international efforts towards digital comprehensive diagnostics in healthcare, focusing on integrating various diagnostic data with AI tools to improve accuracy and treatment planning. The paper highlights challenges in data integration and calls for clear governance and strategic investments to advance digital diagnostics [4].

DiCenzo et al., 2020, a predictive model using quantitative ultrasound (QUS) radiomics was developed to assess response to neoadjuvant chemotherapy in breast cancer patients. The study achieved high sensitivity, specificity, and accuracy, showing QUS-based radiomics as an effective tool for predicting treatment response [5].

Dow-Mu Koh et al., 2022, the paper highlights the multidisciplinary efforts needed to develop AI tools for cancer imaging, emphasizing the importance of collaboration in creating and testing these tools to ensure they meet clinical needs [6].

Freeman et al., 2021, a review of AI accuracy in breast cancer screening concluded that while AI shows promise, it is not yet specific enough to replace double reading by radiologists. The review calls for more high-quality prospective studies to evaluate AI's effectiveness in clinical practice [7].

Hagiwara et al., 2020, Similar to Freeman et al., this review evaluates AI's accuracy in breast cancer detection through mammography, finding that AI systems generally lag behind radiologists. The study suggests AI's role in screening remains uncertain and highlights the need for further research [8].

Le Boulc'h et al., 2020, the study assessed the agreement between AI-based breast density assessments and visual assessments by radiologists. The results showed high agreement, suggesting that AI can reliably assist in predicting breast cancer risk [9].

Lo Gullo et al., 2020, this review discusses the potential of radiogenomics, which combines genetic and radiomic data, in cancer diagnosis and treatment. It calls for standardization and larger studies to validate the use of imaging biomarkers in clinical practice [10].

McKinney et al., 2020, An AI system for breast cancer prediction outperformed radiologists in mammogram interpretation, reducing false positives and negatives and showing potential for improving screening accuracy [11].

M. Chandel et al., 2022, the paper provides an overview of machine learning methods and their applications, with a focus on Python programming for AI activities [12].

Prior et al., 2020, The paper emphasizes the importance of open-access data repositories for training machine learning applications in biomedical research, using The Cancer Imaging Archive as an example [13].

Raya-Povedano et al., 2021, The study suggests that AI can significantly reduce workload in breast cancer screening programs while maintaining detection rates, indicating potential efficiency improvements [14].

Rieke et al., 2020, the paper explores federated learning as a solution to data privacy concerns in medical research, advocating for collaborative efforts and standardization in its implementation [15].

Rodriguez-Ruiz et al., 2019, the study found that AI support improves radiologists' cancer detection rates in mammogram readings, suggesting that AI could enhance diagnostic accuracy [16].

Rundo et al., 2020, The study describes a framework for segmenting cancer tissue types based on CT radiomic features, demonstrating high accuracy and potential for clinical research applications [17].

Savenije et al., 2020, The study demonstrates the feasibility of deep learning for automatic organ-at-risk delineation in radiotherapy planning, showing significant time savings compared to traditional methods [18].

Schreuder et al., 2021, the review highlights the potential of AI to improve lung cancer screening efficiency and accuracy, emphasizing the need for further research before widespread implementation [19].

W. T. Tran et al., 2021, the review discusses deep neural networks' promising results in breast cancer screening and diagnosis, particularly with digital mammograms and breast tomosynthesis imaging [20].

Trivizakis et al., 2020, the review examines recent advances in radiogenomics, focusing on deep learning applications in radiology and oncology, and calls for further research to standardize practices [21].

Zwanenburg et al., 2020, the study standardizes radiomic features to ensure reproducibility and calibration across different software, promoting consistency in radiomic analyses [22].

2.1 Problem Identification

There are many of the challenges for android malware detection in this work based on literature review:

- Low accuracy rate of true data prediction from given dataset.
- Traditional System Analysis not sufficient for proper feature extraction.
- More classification error and system analysis does not provide exact results.
- Prone to attacks.

Less effective, Loss of Information and incorrect prediction Results.

3. Methodology

The maximum models show the main steps for preprocessing stage, feature extraction, and classification. The most of system perform with different algorithms have to increase the accuracy of the results by prediction and successful attributes by using machine learning and deep learning algorithms based on AI. It enhances the performance of the overall classification results.

Maximum authors described their work for prediction the malignant cell based on some specified procedure. The common procedure involved in the prediction point of view as following.

- Data selection and loading
- Data Preprocessing
- Splitting Dataset into Train and Test Data
- Classification
- Result Generation
- Compared with similar work

3.1 Data Selection and Loading

- Data selection is the process of determining the appropriate data type and source, as well as suitable instruments to collect data.
- Data selection precedes the actual practice of data collection and it is the process where data relevant to the analysis is decided and retrieved from the data collection.
- Data loading refers to the “load” component.
- After data is retrieved and combined from multiple sources, cleaned and formatted, it is then loaded into a storage system, such as a cloud data warehouse.

Data Preprocessing

The data can have many irrelevant and missing parts. To handle this part, data cleaning is done. It involves handling of missing data, noisy data etc.

- **Missing Data:** This situation arises when some data is missing in the data. It can be handled in various ways.
- **Ignore the tuples:** This approach is suitable only when the dataset we have is quite large and multiple values are missing within a tuple.

- **Fill the Missing values:** There are various ways to do this task. You can choose to fill the missing values manually, by attribute mean or the most probable value.
- **Regression:** Data can be made smooth by fitting it to a regression function. The regression used may be linear or multiple.

Splitting Dataset into Train and Test Data

- Data splitting is the act of partitioning available data into two portions, usually for cross-validator purposes.
- One Portion of the data is used to develop a predictive model and the other to evaluate the model's performance.
- Separating data into training and testing sets is an important part of evaluating data mining models.
- Typically, when you separate a data set into a training set and testing set, most of the data is used for training, and a smaller portion of the data is used for testing. Mostly authors taking the 70-80% data for model training and 30-20% data for testing point of view.
- To train any machine learning model irrespective what type of dataset is being used you have to split the dataset into training data and testing data.

Classification

Classification is the problem of identifying to which of a set of categories, a new observation belongs to, on the basis of a training set of data containing observations and whose categories membership is known.

K-Means: Clustering is an Unsupervised Learning algorithm, which groups the unlabeled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters, and so on**.**

Artificial Neural Network: The term “Artificial neural network” refers to a biologically inspired sub-field of artificial intelligence modeled after the brain. An Artificial neural network is usually a computational network based on biological neural networks that construct the structure of the human brain. Similar to a human brain has neurons interconnected to each other, artificial neural networks also have neurons that are linked to each other in various layers of the networks. These neurons are known as nodes. Artificial neural network tutorial covers all the aspects related to the artificial neural network

3.2 Result Generation

The Final Result have to get generated based on the overall classification and prediction. The performance of basic model approach is evaluated using some measures like [2]-[15],

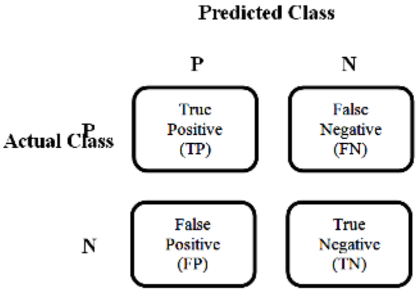
1. Accuracy
2. Precision
3. Recall
4. F1-measure

The final result will get based on the overall classification and prediction and results parameters calculation and generate the confusion matrix. The confusion matrix defined as.

A confusion matrix is a table that is often used to describe the performance of a classification model (or “classifier”) on a set of test data for which the true values are known. The confusion matrix itself is relatively simple to understand, but the related terminology can be confusing.

- **True Positive (TP):** Predicted values correctly predicted as actual positive
- **False Positive (FP):** Predicted values incorrectly predicted an actual positive. i.e., Negative values predicted as positive
- **False Negative (FN):** Positive values predicted as negative
- **True Negative** (**TN**): Predicted values correctly predicted as an actual negative

This framework shows the revised and wrong expectations, in correlation with the real marks. Every disarray network line shows the Real/Genuine marks in the test set, and the segments show the anticipated names by classifier. Something to be thankful for about the disarray grid is that it shows the model's capacity to effectively foresee or isolate the classes.



Predicted Class

		P	N
Actual Class	P	True Positive (TP)	False Negative (FN)
	N	False Positive (FP)	True Negative (TN)

Fig. 4 Prediction class metrics

Accuracy

Accuracy of classifier refers to the ability of classifier. It predicts the class label correctly and the accuracy of the predictor refers to how well a given predictor can guess the value of predicted attribute for a new data.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (3)$$

Precision

Precision is defined as the number of true positives divided by the number of true positives plus the number of false positives.

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (4)$$

Recall

Recall is the number of correct results divided by the number of results that should have been returned. In binary classification, recall is called sensitivity. It can be viewed as the probability that a relevant document is retrieved by the query.

$$\text{Recall}= \frac{TP}{TP+FN} \qquad (5)$$

F-Measure

F measure (F1 score or F score) is a measure of a test's accuracy and is defined as the weighted harmonic mean of the precision and recall of the test.

$$\text{F-measure} = \frac{2TP}{2TP+\text{FP}+FN} \qquad (6)$$

Error Rate

The inaccuracy of predicted output values is termed the error of the method. If target values are categorical, the error is expressed as an error rate. This is the proportion of cases where the prediction is wrong prediction.

Error Rate = 100 - Accuracy

4. Result and Discussion

When creating a machine learning solution that classifies diagnoses as benign or malignant, analytical thinking and strategic thinking were employed. In order to accomplish this, commonly to use a methodical and organized approach. It will begin with exploratory data analysis and work my way up to a deeper grasp of the dataset's features, distributions, and any difficulties. Several methods will be thoroughly evaluated before choosing a model. I want to play around with various machine learning approaches, from more complex ones like ensemble methods and neural networks to more conventional models. This variety in testing enables a comprehensive comprehension of the dataset's compatibility with various algorithms. After a model is chosen, careful hyper parameter tuning will be carried out. The settings of the parameters will be changed to guarantee peak performance and prevent over fitting. The generalization capacity of the model will be rigorously cross-validated, guaranteeing its dependability on unobserved data. The fine needle aspirate (FNA) image of a breast lump that was digitally captured is the source of the features in the dataset. With this technique, a sample of cells from a breast lesion is taken with a fine needle, and the cells are subsequently examined under a microscope. Every characteristic contains important details on the characteristics of the nucleus, aiding in the differentiation of benign and malignant tumors. Metrics such as radius, texture, area, perimeter, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension are examples of these properties. These characteristics reflect the morphological and structural characteristics of the cell nuclei and act as measurable descriptors. Feature extraction in conjunction with digital pictures allows complicated biological features to be translated into numerical attributes. These characteristics serve as the foundation for algorithms used in machine learning and statistical analysis. The dataset enables academics and data scientists to employ computer algorithms for diagnostic classification by converting the intrinsic complexity of biological tissues into quantifiable metrics. Essentially, the features in the Breast Cancer Wisconsin dataset provide a multifaceted view of the properties of cellular nuclei. After digitizing photographs and extracting important characteristics, the images are organized to make analysis and categorization easier. By bridging the gap between quantitative analysis and medical imaging, this method advances the detection and study of breast cancer.

The data cleaning stage is essential to guaranteeing the quality, dependability, and correctness of the information.

In data cleansing, managing missing values is one of the main responsibilities. In order to solve the issue of missing data points distorting analysis, methods such as imputation or removal are used. To avoid having an excessive impact on outcomes, outliers-data points that dramatically depart from the norm-are also recognized and handled. To guarantee uniformity and dependability, discrepancies in values, data formats, or naming conventions are fixed. We find and remove duplicate records, which may result from mistakes made during data gathering. Accurate analysis depends on solving data format problems, such as converting data kinds or standardizing units. In addition, handling typos, special characters, and data input problems that could compromise the integrity of the dataset are all part of the cleaning step.

The data

Taking the data by mostly authors from koggle.com or self-generated based on sample collected by patient in form of .csv or image if data in form of image it has to convert in .csv file based on their characteristics.

Breast Cancer

One kind of cancer that starts in the breast cells is called breast cancer. It happens when genetic abnormalities cause normal cells in the breast tissue to grow and divide uncontrolled, leading to the eventual formation of a lump or tumor. These tumors fall into one of two categories: benign (non-cancerous) or malignant (cancerous). Because they are cancerous, malignant tumors have the ability to infiltrate neighboring tissues and travel via the lymphatic or circulatory systems to other areas of the body. This phenomenon, known as metastasis, might exacerbate the difficulty of treating cancer. On the other hand, benign tumors do not infiltrate neighboring tissues or spread to other bodily parts. They are usually not life-threatening, but they may still need medical attention or treatment, particularly if they cause discomfort.

Method to find Characteristics

The fine needle aspirate (FNA) image of a breast lump that was digitally captured is the source of the features in the dataset. With this technique, a sample of cells from a breast lesion is taken with a fine needle, and the cells are subsequently examined under a microscope. The features of the dataset are the outcome of measuring different aspects of the cell nuclei shown in these pictures.

Every characteristic contains important details on the characteristics of the nucleus, aiding in the differentiation of benign and malignant tumors. Metrics such as radius, texture, area, perimeter, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension are examples of these properties. These characteristics reflect the morphological and structural characteristics of the cell nuclei and act as measurable descriptors.

Essentially, the features in the Breast Cancer Wisconsin dataset provide a multifaceted view of the properties of cellular nuclei. After digitizing photographs and extracting important characteristics, the images are organized to make analysis and categorization easier. By bridging the gap between quantitative analysis and medical imaging, this method advances the detection and study of breast cancer.

Process and cleanse the data

The data cleaning stage is essential to guaranteeing the quality, dependability, and correctness of the information.

In data cleansing, managing missing values is one of the main responsibilities. In order to solve the issue of missing data points distorting analysis, methods such as imputation or removal are used.

To guarantee uniformity and dependability, discrepancies in values, data formats, or naming conventions are fixed. We find and remove duplicate records, which may result from mistakes made during data gathering. Accurate analysis depends on solving data format problems, such as converting data kinds or standardizing units. In addition, handling typos, special characters, and data input problems that could compromise the integrity of the dataset are all part of the cleaning step.

Data information

When attempting to solve a classification problem using just numerical predictors, some difficulties arise because all of the predictors are numerical data. The data's dimensionality raises the possibility of over fitting and complicates computation. It becomes essential to choose pertinent predictors in order to reduce noise caused by unimportant variables.

To guarantee algorithm accuracy, the different scales of numerical predictors necessitate careful scaling or normalization. Complex models, however, may make interpretation more difficult. Innovative feature engineering and non-linear algorithms may be required due to non-linear interactions between predictors and the target.

Given the large number of numeric predictors, over fitting is a potential problem. Regularization, cross-validation, and careful assessment are necessary to strike a balance between model complexity and generalization. Effective feature selection, skillful preprocessing, and appropriate algorithms for high-dimensional numerical data are necessary to overcome these obstacles. Regular testing and validation guarantee that the method works well with fresh data.

Exploratory Data Analysis

A comprehensive investigation of the dataset's complexities is conducted during the foundational phase of data science initiatives, known as exploratory data analysis (EDA). Finding patterns, trends, and anomalies is the goal of this process, which includes a variety of tasks that set the stage for further analysis and modeling. EDA gives a quick overview of the distributions, dispersion, and central patterns of the data using summary statistics. When figuring out how different variables relate to one another, visualizations come in handy since they can show possible dependencies and connections that direct additional research. EDA sheds light on the underlying structure of the data by exploring both numeric and categorical variables, revealing insights into their distributions and relationships. EDA is vital since it deals with data integrity and quality. Determining whether data imputation or removal is required is aided by identifying outliers and evaluating missing values. Additionally, in classification tasks, EDA closely examines the target variable to identify class distributions and any imbalances.

Diagnosis (Target)

Classification of diagnosis (M: Malignant, B: Benign)

From a more general standpoint, there is a significant association between higher feature values and diagnoses of cancer. This finding has important ramifications, especially in light of the possible choice to discretize the data. The discussion here centers on the potential for better results with data discretization.

4.1 Correlated Predictors

Comprehensive examination of every predictor. The trends that the data analysis revealed are really instructive. It is evident that variables with correlations higher than 0.90 have a strong linear relationship. This suggests that when one variable, X, rises, so does another variable, Y. Given that many algorithms specifically take advantage of this kind of strong correlation to increase forecast accuracy, this linear dependency has great potential for machine learning models. It is noteworthy that certain features may not exhibit as strong of a link with the target, yet they might still be useful in the modelling process. They could improve other variables' capacity for prediction or support the general stability of the model. Because these variables might add to a larger prediction framework, it is imperative that they remain in the dataset. As we move forward, we'll continue to analyze these powerful predictors while maintaining a broad perspective on the dataset's potential.

4.2 Modeling the Data

The process of organizing and preparing a dataset for machine learning is known as modeling data. To make the data appropriate for training and assessing machine learning models, it must be cleaned, transformed, and arranged. Prior to training a machine learning model, this phase is essential since the structure and quality of the data directly affect the model's functionality and the accuracy of its predictions. By ensuring that the data is correctly formatted, contains pertinent information, and is error- and inconsistency-free, modeling data helps the model identify patterns and produce precise predictions.

Feature Selection

The process of selecting a subset of the most pertinent and instructive features (or variables) from a broader range of available features is known as feature selection in the context of machine learning and data analysis. By concentrating on the most important predictors, the goal is to increase interpretability, decrease complexity, and improve model performance. Metric-based feature selection is one method that falls under this more general category. It entails calculating each feature's significance or relevance to the target variable using particular metrics or statistical tests. Data scientists can use these metrics to choose or rank features and use them to make well-informed decisions about which characteristics to include in the final model.

Commonly used metrics for feature selection include:

1. **Correlation:** Features are evaluated based on their correlation with the target variable. Higher absolute correlation values indicate greater relevance.
2. **Mutual Information:** This metric measures the mutual dependence between a feature and the target. Higher mutual information suggests greater relevance.
3. **Chi-Squared (χ^2) Test:** Typically used for categorical features, this statistical test assesses the independence between a feature and the target. Smaller p-values indicate higher relevance.
4. **F-Test (ANOVA):** This test is employed for numerical features and evaluates whether there are statistically significant differences in the means of the target variable across different feature categories.
5. **Recursive Feature Elimination (RFE):** RFE is an iterative method that ranks features by recursively training a model and eliminating the least important feature at each step. This continues until a desired number of features is reached.
6. **L1 Regularization (LASSO):** L1 regularization methods, like LASSO (Least Absolute Shrinkage and Selection Operator), encourage sparsity by penalizing the absolute values of feature coefficients. Features with non-zero coefficients are selected.
7. **Information Gain and Entropy:** These metrics, commonly used in decision trees and random forests, assess how well a feature splits or classifies data into different target categories.

8. **Permutation Importance:** This method evaluates the change in model performance (e.g., accuracy or F1 score) when the values of a feature are randomly shuffled. A significant drop in performance indicates the feature's importance.

The type of data and the issue at hand determine which statistic is best. Based on these criteria, feature selection improves interpretability, decreases overfitting, speeds up the modeling process, and frequently results in better model generalization. It's crucial to remember that feature selection should be done cautiously because eliminating crucial features may cause vital data to be lost.

Apart from Metric-Based selection, there are other feature selection methods that can be investigated for diverse datasets and contexts. I invite you to become involved in my project, “Feature Selection for Machine Learning.” I've thoroughly examined a variety of feature selection techniques in this extensive study, painstakingly examining their advantages, disadvantages, and real-world uses.

4.3 Machine Learning Workflow

The main goal in this section is to use several supervised classification algorithms, evaluate their effectiveness, and finally choose which is the best fit for accomplishing our particular objectives. Along with training the models, this phase entails exploring a variety of methodologies to fully assess the models' strengths and weaknesses.

Throughout this process, we will closely examine important performance indicators to determine how well each model matches our particular objectives, including accuracy, precision, recall, and F1-score. We will also look at methods for optimizing model parameters, adjusting hyperparameters, and resolving possible overfitting or underfitting problems.

Model evaluation

Several evaluation metrics are used to evaluate the effectiveness of machine learning models in supervised classification issues, where the objective is to classify data points into predetermined classes or labels. These measures, which might change depending on the kind of problem (binary or multiclass), aid in our understanding of how well a model is performing in terms of producing accurate predictions. These are a few typical measures for evaluating classification issues are Accuracy, Precision, Recall (Sensitivity or True Positive Rate), F1-Score: ROC Curve (Receiver Operating Characteristic Curve).

Baseline Scores

In baseline scores are essential points of reference that shed light on how well our chosen method to the predictive job is working. These scores are essential for differentiating between outcomes that could be the result of chance and significant model performance. Through the computation and assessment of these baseline values, we obtain important insights. If our machine learning model doesn't outperform these arbitrary forecasts, then indicates that there isn't much predictive value to our method. On the other hand, if our model outperforms these baseline scores, it shows that the strategy we have chosen is truly identifying patterns and relationships in the data.

Zero Rate Score

In binary classification tasks, the Zero Rate Score baseline score serves as a straightforward yet crucial reference point. When the model correctly predicts the majority class in every case, essentially making the same prediction regardless of the input data, it represents the accuracy attained. By comparing our model's performance to a method of forecasting the majority class without taking into account any underlying patterns or relationships in the data, this baseline score serves as a fundamental benchmark.

Random Rate Classifier

In binary classification tasks, the Random Rate Classifier-also known as Weighted Guessing-is a standard technique used. Using the class distribution in the training data as a guide, this method randomly assigns class labels to instances, with the likelihood of selecting each class based on its percentage in the dataset. It essentially simulates a situation in which predictions are produced purely at random, with no regard for the characteristics or patterns seen in the data. The Random Rate Classifier is a basic test used to assess how well machine learning models perform; for a model to be deemed effective, it must be able to predict more accurately than this random guessing method.

Model Selection

We are dealing with a classification problem in this case, which is a subset of supervised learning. Our main goal is to correctly classify data items into the appropriate classes or categories by utilizing the prediction potential of different attributes, or predictors. This challenge involves using algorithms to guide us in identifying patterns and relationships in the data so that we may accurately forecast the classes.

1. Logistic Regression
2. Decision Tree
3. Random Forest
4. Support Vector Machine (SVM)
5. K-Nearest Neighbors (KNN)

6. Naive Bayes
7. Gradient Boosting Machines (GBM)
8. XGBoost
9. LightGBM
10. CatBoost
11. AdaBoost
12. LDA (Linear Discriminant Analysis
13. Gaussian Process Classifier
14. MLP (Multi-layer Perceptron)

Each of these models has its strengths and weaknesses, making them suitable for different types of problems and datasets. The choice of the most appropriate model depends on the specific characteristics of the data and the goals of the machine learning task.

Cross-validation

A key method in machine learning for evaluating a model's performance and capacity for generalization is cross-validation. A dataset is divided into several subsets, or “folds,” and the model is iteratively trained and tested on various combinations of these folds. Getting a reliable estimate of a model's performance that is less reliant on the precise random segmentation of the data is the main objective.

A particular kind of cross-validation method called Stratified K Fold makes sure that every fold preserves the same class distribution as the original dataset. This is especially crucial when dealing with imbalanced datasets-that is, datasets in which one class is much more numerous than the others-in classification issues. By preventing situations where one or more folds have too few occurrences from a minority class, Stratified K Fold helps avoid problems that could result in inaccurate or biased performance estimations.

When working with classification tasks, Stratified K Fold is particularly helpful because it makes sure that the model's performance evaluation fairly represents its capacity to manage class imbalances. It is a useful tool in the model evaluation process because it produces more dependable and less biased performance measurements by preserving a constant class distribution across each fold.

Voting Classifier

Often referred to as “base classifiers,” a voting classifier is an ensemble learning technique in machine learning that generates a final prediction by aggregating the predictions of several distinct models. A voting classifier's main concept is to combine the outputs of its basic classifiers to increase the model's overall accuracy and generalizability. There are two main types of voting classifiers: hard voting and soft voting. Depending on the nature of the problem and the properties of the underlying classifiers, one can choose between hard and soft voting. When probability estimates are provided by the basic classifiers, soft voting is favored since it facilitates more complex and probabilistic decision-making. Additionally, it can handle scenarios in which the confidence levels of the basis classifiers' predictions disagree.

After a thorough examination and careful comparison of the chosen approaches, it is clear that Hard Voting does not meet our goals or performance standards. Unfortunately, it is not able to provide the desired level of resilience and prediction accuracy. As a result, I have decided not to consider Hard Voting as a workable ensemble strategy for our particular assignment.

I firmly believe that the Soft Voting approach is a wise selection for our classification issue. It is consistent with our target metrics and gains its strength from the best algorithms that we have carefully examined. This tactical choice demonstrates our dedication to finding the most reliable and efficient solution that takes into account the unique characteristics of our dataset and problem domain.

4.4 Final Solution

Given our selection of the Soft Voting classifier as the linchpin of our machine learning solution for this problem, it's imperative to embark on a thorough exploration of this ensemble model. We're set to delve beneath the surface and scrutinize the inner workings of this algorithm, seeking assurance of its viability and effectiveness. This entails a multifaceted analysis encompassing several aspects, including the distribution of predictions, probabilities, and an in-depth examination of the algorithm's procedural steps.

To begin, we'll closely examine the distribution of predictions generated by the Soft Voting classifier. This serves as an essential step in comprehending how the model distributes its predictions across different classes and the degree of certainty it exhibits in its decisions. Understanding these prediction patterns offers valuable insights into the classifier's decision-making process.

Moreover, we'll employ various techniques to dissect the probabilities associated with the model's predictions. By dissecting these probabilities, we gain a deeper understanding of the confidence levels assigned to each classification outcome. This knowledge is pivotal, as it aids in assessing the model's reliability and its propensity to make accurate predictions.

Creating training and testing data

It operates by dividing the dataset into two distinct parts: the training set and the testing set. The majority of the data, typically around 70-80%, is allocated to the training set. Here, the model learns patterns and relationships within the data, adjusting its parameters to capture these nuances.

The remaining 20-30% constitutes the testing set, which is withheld during training. Once the model is trained, it is used to make predictions on the testing set, and these predictions are matched against the actual values.

Confusion matrix

This matrix is particularly valuable because it helps us quantify the model's performance in terms of accuracy, precision, recall, and F1-score, among other metrics. By examining these metrics, we can gain a comprehensive understanding of how well the model distinguishes between classes. For example, precision measures the model's ability to make accurate positive predictions, while recall assesses its capability to capture all positive instances. Balancing these metrics is often crucial, depending on the specific goals of the classification task. In sum, the confusion matrix serves as a cornerstone for evaluating and fine-tuning classification models, providing critical insights into their strengths and areas of improvement.

It's noteworthy that, in this evaluation, the model's performance leaned more toward making false positive errors rather than false negatives. False negatives would imply failing to identify malignancies that were actually present, which can be considerably more critical in a medical context. However, in this assessment, the model demonstrated an encouraging capability to avoid false negative errors. While there's room for fine-tuning and minimizing false positives, this outcome underscores the algorithm's proficiency in capturing potential cases, even if it occasionally exhibits heightened caution.

Classification report

This report offers a detailed breakdown of how effectively the model has carried out this classification task.

They shed light on its efficacy for individual classes and help identify potential issues like overfitting or underfitting. Such reports are especially instrumental in domains like medical diagnosis, spam detection, sentiment analysis, and various contexts where precise classification is of paramount importance

5. Conclusion

This work presented a theoretical exploration of the Breast Cancer Wisconsin dataset, with the overarching goal of constructing a machine learning solution capable of discriminating between malignant and benign diagnoses. Throughout the work, a consistent emphasis has been placed on achieving a solution characterized by robustness and reproducibility, a paramount requirement given the critical nature of the health dataset under examination.

The exploration and analysis of the dataset have unfolded systematically, unveiling critical insights. Of paramount significance is the discovery that higher feature values consistently correlate with malignant diagnoses. This revelation serves as the foundation for much of the subsequent analysis and underscores the importance of features in characterizing the nature of the disease. Furthermore, the dataset revealed high correlations among specific features, indicating the potential for feature selection, an avenue explored to refine the dataset.

Feature selection and engineering have been pivotal in shaping the analysis. Features were thoughtfully chosen based on their correlations with the target variable, a process of distillation to extract the most salient attributes for modeling. Employing an exhaustive selection approach, the work identified the features most relevant for modeling.

The core of the work lies in the modeling phase, where various classification algorithms were rigorously evaluated. These algorithms encompassed a spectrum of approaches, including Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Naive Bayes, Gradient Boosting, and more. Rigorous assessment occurred through the application of a 10-fold cross-validation approach, ensuring a robust evaluation of model performance. Notably, Linear Discriminant Analysis (LDA) and Multi-layer Perceptron Classifier (MLPC) surfaced as the standout performers among the considered models.

References

[1] Ather, S., Kadir, T. & Gleeson, F. Artificial intelligence and radiomics in pulmonary nodule management: current status and future applications. Clin. Radiol. Vol. 75, pp.13-19, 2020.

[2] Bitencourt, A. G. V. et al. MRI-based machine learning radiomics can predict HER2 expression level and pathologic response after neoadjuvant therapy in HER2 overexpressing breast cancer. EBioMedicine 61, 103042, 2020.

[3] Bizzo, B. C., Almeida, R. R., Michalski, M. H. & Alkasab, T. K. Artificial intelligence and clinical decision support for radiologists and referring providers. J. Am. Coll. Radiol. Vol.16, pp.1351-1356, 2019.

[4] Bukowski, M. et al. Implementation of eHealth and AI integrated diagnostics with multidisciplinary digitized data: are we ready from an international perspective. Eur. Radiol. Vol. 30, pp.5510-5524, 2020.

[5] DiCenzo, D. et al. “Quantitative ultrasound radiomics in predicting response to neoadjuvant chemotherapy in patients with locally advanced breast cancer: Results from multi-institutional study” Cancer Med. 9, pp.5798-5806, 2020.

[6] Dow-Mu Koh et al. “Artificial intelligence and machine learning in cancer imaging,” Communications medicine, 2,133, pp. 1-14, 2022.

[7] Freeman, K. et al. “Use of artificial intelligence for image analysis in breast cancer screening programmes: systematic review of test accuracy,” BMJ 374, n1872, 2021.

[8] Hagiwara, A., Fujita, S., Ohno, Y. & Aoki, S. Variability and standardization of quantitative imaging: monoparametric to multiparametric quantification, radiomics, and artificial intelligence. Invest. Radiol. Vol.55, pp.601-616, 2020.

[9] Le Boulc’h, M. et al. Comparison of breast density assessment between human eye and automated software on digital and synthetic mammography: Impact on breast cancer risk. Diagn. Interv. Imaging Vol.101, pp.811-819, 2020.

[10] Lo Gullo, R., Daimiel, I., Morris, E. A. & Pinker, K. Combining molecular and imaging metrics in cancer: radiogenomics. Insights Imaging 11, 1, 2020.

[11] McKinney, S. M. et al. International evaluation of an AI system for breast cancer screening. Nature 577, pp.89-94, 2020.

[12] Meghna Chandel, Sanjay Silakari, Rajeev Pandey, Smita Sharma, “A Study on Machine Learning and Python’s Framework,” International Journal of Computer Sciences and Engineering, Vol. 10, Issue.5, pp.58-64, May 2022.

[13] Prior, F. et al. Open access image repositories: high-quality data to enable machine learning research. Clin. Radiol. Vol.75, pp.7-12, 2020.

[14] Raya-Povedano, J. L. et al. AI-based strategies to reduce workload in breast cancer screening with mammography and tomosynthesis: a retrospective evaluation. Radiology 300, 57-65, 2021.

[15] Rieke, N. et al. The future of digital health with federated learning. NPJ Digit. Med. 3, 119, 2020.

[16] Rodriguez-Ruiz, A. et al. Detection of breast cancer with mammography: effect of an artificial intelligence support system. Radiology 290, 305-314, 2019.

[17] Rundo, L. et al. Tissue-specific and interpretable sub-segmentation of whole tumour burden on CT images by unsupervised fuzzy clustering. Comput. Biol. Med. 120, 103751, 2020.

[18] Savenije, M. H. F. et al. Clinical implementation of MRI-based organs-at-risk auto-segmentation with convolutional networks for prostate radiotherapy. Radiat. Oncol. 15, 104, 2020.

[19] Schreuder, A., Scholten, E. T., van Ginneken, B. & Jacobs, C. Artificial intelligence for detection and characterization of pulmonary nodules in lung cancer CT screening: ready for practice. Transl. Lung Cancer Res. 10, 2378-2388, 2021.

[20] William T Tran, Ali Sadeghi-Naini, Fang-I Lu, Sonal Gandhi, Nicholas Meti, Muriel Brackstone, Eileen Rakovitch, Belinda Curpen, Computational radiology in breast cancer screening and diagnosis using artificial intelligence. Can. Assoc. Radiol. J. 72, 98-108, 2021.

[21] Eleftherios Trivizakis, Georgios Z Papadakis, Ioannis Souglakos, Nikolaos Papanikolaou, Lefteris Koumakis, Demetrios A Spandidos, Aristidis Tsatsakis, Apostolos H Karantanas, Kostas Marias, Artificial intelligence radiogenomics for advancing precision and effectiveness in oncologic care (Review). Int. J. Oncol. 57, 43-53, 2020.

[22] Zwanenburg, A. et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology Vol.295, pp.328-338, 2020.