

The Application of Artificial Intelligence to Cancer Research: A Comprehensive Guide

Amin Zadeh Shirazi, PhD¹, Morteza Tofiqhi, PhD²,
Alireza Gharavi, BSc³, and Guillermo A. Gomez, PhD¹ 

Technology in Cancer Research & Treatment
Volume 23: 1-20
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: [10.1177/15330338241250324](https://doi.org/10.1177/15330338241250324)
journals.sagepub.com/home/tct



Abstract

Advancements in AI have notably changed cancer research, improving patient care by enhancing detection, survival prediction, and treatment efficacy. This review covers the role of Machine Learning, Soft Computing, and Deep Learning in oncology, explaining key concepts and algorithms (like SVM, Naïve Bayes, and CNN) in a clear, accessible manner. It aims to make AI advancements understandable to a broad audience, focusing on their application in diagnosing, classifying, and predicting various cancer types, thereby underlining AI's potential to better patient outcomes. Moreover, we present a tabular summary of the most significant advances from the literature, offering a time-saving resource for readers to grasp each study's main contributions. The remarkable benefits of AI-powered algorithms in cancer care underscore their potential for advancing cancer research and clinical practice. This review is a valuable resource for researchers and clinicians interested in the transformative implications of AI in cancer care.

Keywords

cancer, artificial intelligence, machine learning, soft computing, machine vision, deep learning

Abbreviations

AI, artificial Intelligence; KNN, K-nearest neighbors; ANN artificial neural network; CNN, convolutional neural network; RNN, recurrent neural network; micro-FTIR, micro-Fourier transform infrared spectroscopy; ABC, artificial bee colony; RC, random committee; MRMR, minimum redundancy maximum relevance; TLBO, teaching learning-based optimization; FCM, fuzzy cognitive maps; WDBC, Wisconsin diagnostic breast cancer; WBC, Wisconsin breast cancer; MRI, magnetic resonance imaging; TCGA, Cancer Genome Atlas; LIDC, Lung Image Database Consortium; ROI, region of interest; mpMRI, multiparametric magnetic resonance imaging; SELDI, surface-enhanced laser desorption and ionization; CEA, carcinoembryonic antigen; GLCM, gray level co-occurrence matrix; WRBC, Wisconsin Recurrence Breast Cancer Database; ABCD, Asymmetry Index (A), Border Irregularity (B), Color score (C) and Diameter (D); MGI, maximum gradient intensity; FNA, fine needle aspiration; DSS, disease-specific survival; NYU, New York University; NN-ET, neural network and extra tree; UCSB, University of California, Santa Barbara, USA; RBCC, red blood cell count; TS, tumor size; HBP, high blood pressure; CRP, C-reactive protein; FSES, fuzzy soft expert system; MIAS, Mammographic Image Analysis Society; T2FIS, type II fuzzy inference system; PCA, principal component analysis; SOM, self-organizing map; BI-RADS, breast imaging reporting and data system; IWOA, improved whale optimization algorithm; PSO, particle swarm optimization; EO, equilibrium optimizer; DNN, deep neural network; WBC, white blood cells; PT, prothrombin time; TT, thrombin time; INR, international normalized ratio; RBC, red blood cells; HMLP, hybrid multi-layered perceptron; TCM, traditional Chinese medicine; BOAALO, butterfly optimization algorithm and the ant lion optimizer; MLP, multilayer perceptron; BiLSTM, bidirectional long short-term memories; CSE, concurrent squeeze and excitation; PSA, prostate-specific antigen; BPH, benign prostatic hyperplasia; FNB, fine needle biopsy; ELM,

¹ Centre for Cancer Biology, SA Pathology and the University of South Australia, Adelaide, SA, Australia

² Department of Electrical Engineering, Faculty of Engineering, Bu-Ali Sina University, Hamedan, Iran

³ Department of Computer Science, Azad University, Mashhad Branch, Mashhad, Iran

Corresponding Author:

Guillermo A. Gomez, Centre for Cancer Biology, SA Pathology and the University of South Australia, Adelaide, SA, 5000, Australia.

Email: Guillermo.Gomez@unisa.edu.au



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

extreme learning machine; TAI-net, triple-attention interaction network; DWI, diffusion-weighted Imaging; XGboost, eXtreme gradient boost; ANFIS, adaptive network-based fuzzy inference system; LSTM, long-term short-term memory networks; ML, machine learning; CS, Cuckoo search; CFS, correlation feature selection; SBO, satin bowerbird optimization; IWO, invasive weed optimization; QDA, quadratic discriminant analysis; TIF, type-I fuzzy logic; WPBC, wisconsin prognostic breast cancer; CT Scan, computed tomography scan; GEO, gene expression omnibus; BreakHis, Breast Cancer Histopathology Database; MCC, Matthew's correlation coefficient; AUC, area under the curve; SEER, surveillance, epidemiology, and end results; CRIncRC, cancer-related long non-coding RNAs; GSI, gemstone spectral imaging; BMI, body mass index; ISIC, international skin imaging collaboration; LBP, local binary patterns; SF, short form; OS, overall survival; CVC, central venous catheter; PLS-ANNDA, partial least squares artificial neural network discriminant analysis; miRNA, micro-ribonucleic acids; FIS, fuzzy inference system; FP, flank pain; VHLG, Von Hippel–Lindau Gene; TCEE, trichloroethylene exposure; AJCC, American Joint Committee on Cancer; DDSM, Digital Database for Screening Mammography; BRATS, multimodal brain tumor image segmentation benchmark; EM, expectation maximization; CART, classification and regression trees; CVNN, complex-valued neural network; ODNN, optimal deep neural network; GONN, genetically optimized neural network; ASO, atom search optimization; GA, genetic algorithm; TNM, Tumor (T), Node (N), and Metastasis (M); FIB, fibrinogen; APTT, activated partial thromboplastin time; ALB, albumin; PLT, platelet count; GLB, globulin; ELM, extreme learning machine; MULTP, modified uniform local ternary patterns; Lung-PET-CT-Dx, large-scale CT and PET/CT dataset for lung cancer diagnosis; SERS, surface-enhanced Raman spectroscopy; PCa, prostate cancer; HUAP, University Hospital Antônio Pedro; TL, transfer learning; EMS-Net, ensemble of multiscale convolutional neural networks; ADC, obvious dispersion coefficient; BEiT, BERT pre-training of image transformers.

Received: October 29, 2023; Revised: March 28, 2024; Accepted: April 8, 2024.

Artificial Intelligence Application to Cancer Research

Cancer continues to be a significant global health challenge, with early diagnosis, accurate prognosis, and personalized treatment being critical for improving patient outcomes. In recent years, the application of artificial intelligence (AI) has emerged as a promising approach to revolutionize cancer care, offering unprecedented opportunities for advancements in cancer research and clinical practice. Figure 1 shows the number of papers published on the application of AI to cancer research and the number of AI-based models applied to different cancers, highlighting the rapidly advancing field of AI in cancer research.

AI has shown promising results in studying different types of human cancers,^{2–5} including but not limited to cervical cancer,^{6,7} pancreatic cancer,^{3,8} breast cancer,^{9,10} colorectal cancer,¹¹ ovarian cancer,¹² laryngeal cancer,¹³ brain cancer,^{14,15} and lung cancer.^{4,16,17} In a scoping review, the extent of the use of AI and ML protocols for cancer diagnosis in prospective settings was explored.² A literature review was also conducted on AI-driven digital cytology-based cervical cancer screening and highlighted the potential of this technology in resource-constrained settings.⁶ Another scoping review was conducted on the use of AI for the prediction and early diagnosis of pancreatic cancer, emphasizing the importance of early detection for improved survival rates.³ Meanwhile, a comprehensive overview of multifunctional magnetic nanostructures integrated with an AI approach for cancer diagnosis and therapy has been conducted.¹⁸ Indeed, a systematic review and meta-analysis have been conducted on the value of AI in lung cancer diagnosis, highlighting the potential of AI-assisted diagnosis in improving accuracy and efficiency.⁴ Another review

was conducted on AI applications to find the most critical features extracted from brain cancer patients' MRI, histopathology, and CT scan images.¹⁹ These studies have shown promising results in improving cancer diagnosis accuracy, sensitivity, and specificity, potentially reducing the burden on healthcare providers and improving patient outcomes. Furthermore, AI has been integrated with other technologies, such as radionics and genomics, to enhance cancer diagnosis and prediction.^{20,21}

This paper reviews AI's role in cancer diagnosis, covering its current use, challenges, and future directions. It focuses on how machine learning, deep learning, and soft computing, illustrated in Figure 2, enhance cancer research by aiding in early detection, accurate diagnosis, treatment prediction, and monitoring tumor recurrence. The aim is to give a thorough overview of AI's application in oncology and demonstrate how these AI subfields can boost diagnostic precision, efficiency, and patient care, with detailed exploration of their applications.

Machine Learning in Cancer Research

Machine learning, a subfield of AI, has been extensively applied in cancer research because of its ability to analyze complex data patterns and make accurate predictions. Several studies have reviewed the application of machine learning in cancer research, providing valuable insights into its potential and limitations.^{22–24}

As a machine learning method, decision trees have demonstrated numerous applications in medicine and public health, particularly in addressing various issues within the field of cancer.^{25–27} Such applications span an array of cancer types, including breast,^{28–32} gastric,^{33,34} thyroid,³⁵ prostate,³⁶ and

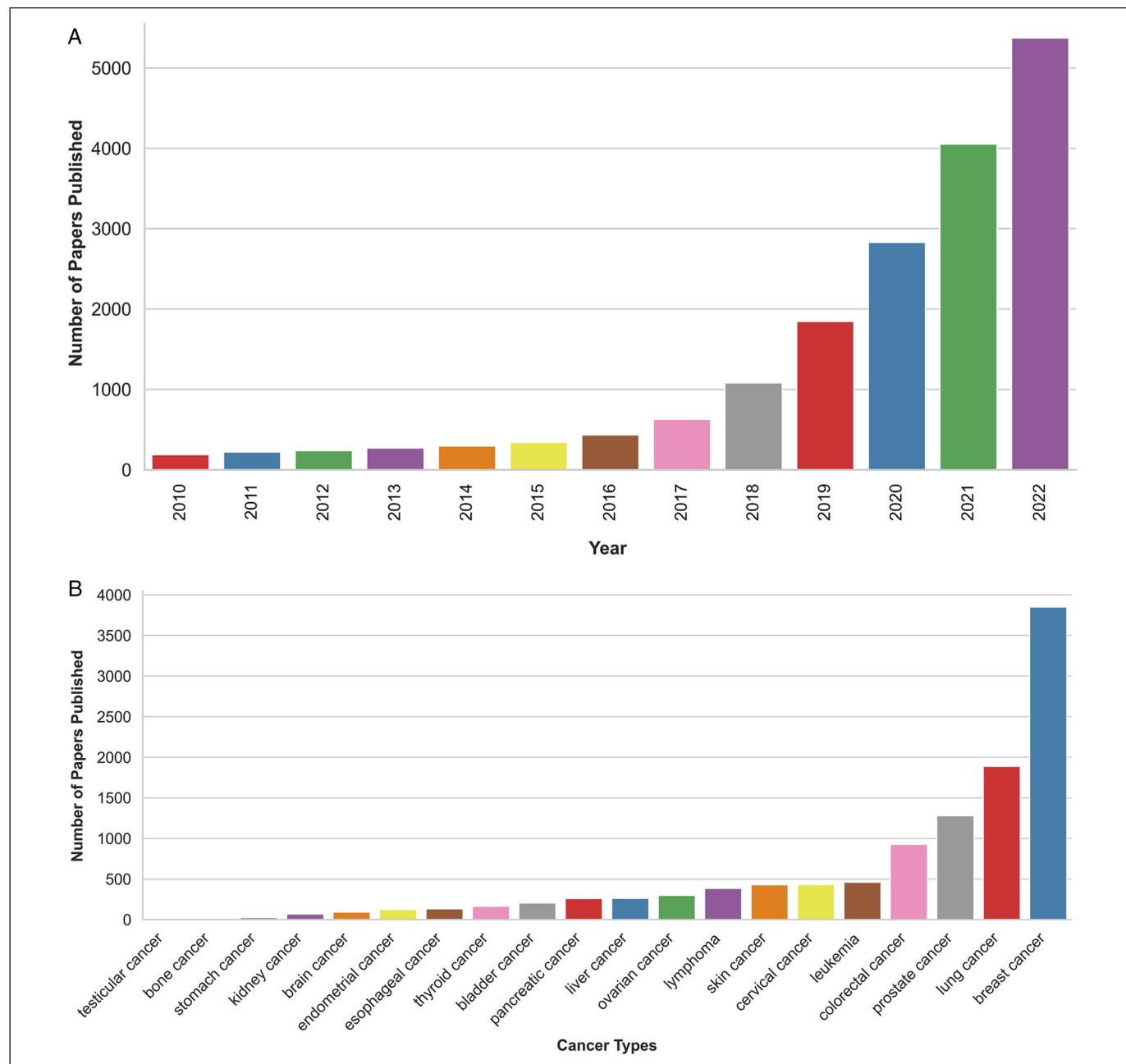


Figure 1. (A) The number of papers published in the field of “*AI applications to cancer research*” between 2010 and 2023 (keywords searched: “AI” and “Cancer”). (B) The number of AI-based models applied to different types of cancers between 2010 and 2022 (keywords searched: “AI” and “Cancer Type”). For A and B, the information was extracted from the PubMed database.¹

colorectal cancer.³⁷ These studies have exhibited significant improvements in the accuracy of cancer diagnoses.^{30,33} Additionally, the k-means machine learning algorithm has been employed in several types of cancer, such as breast^{38,39} and skin cancer.⁴⁰ Another technique, K-nearest neighbors (KNN), has also been used in cancer research,⁴¹ yielding enhanced accuracy in cancer prediction.⁴¹ Furthermore, logistic regression is a widely used technique in cancer research,⁴² aiding in the improvement of various cancer diagnoses, including gastric,⁴³ colon,⁴⁴ and bladder cancer,⁴⁵ among others.

Naïve Bayes, another powerful machine learning technique, has also found applications in the field of cancer.^{46,47} Principal component analysis has also proven valuable for various cancer types, such as lung,⁴⁸ cervical,⁴⁹ and colorectal cancer.⁵⁰ Random Forests, an ensemble machine learning method, has been employed in numerous cancer diagnoses, including breast,^{51,52} ovarian,⁵³ thyroid,⁵⁴ and cervical cancer.⁵⁵ Lastly, eXtreme Gradient Boost, a classification machine learning technique, has been applied to different types of cancer, such as gastric cancer.⁵⁶

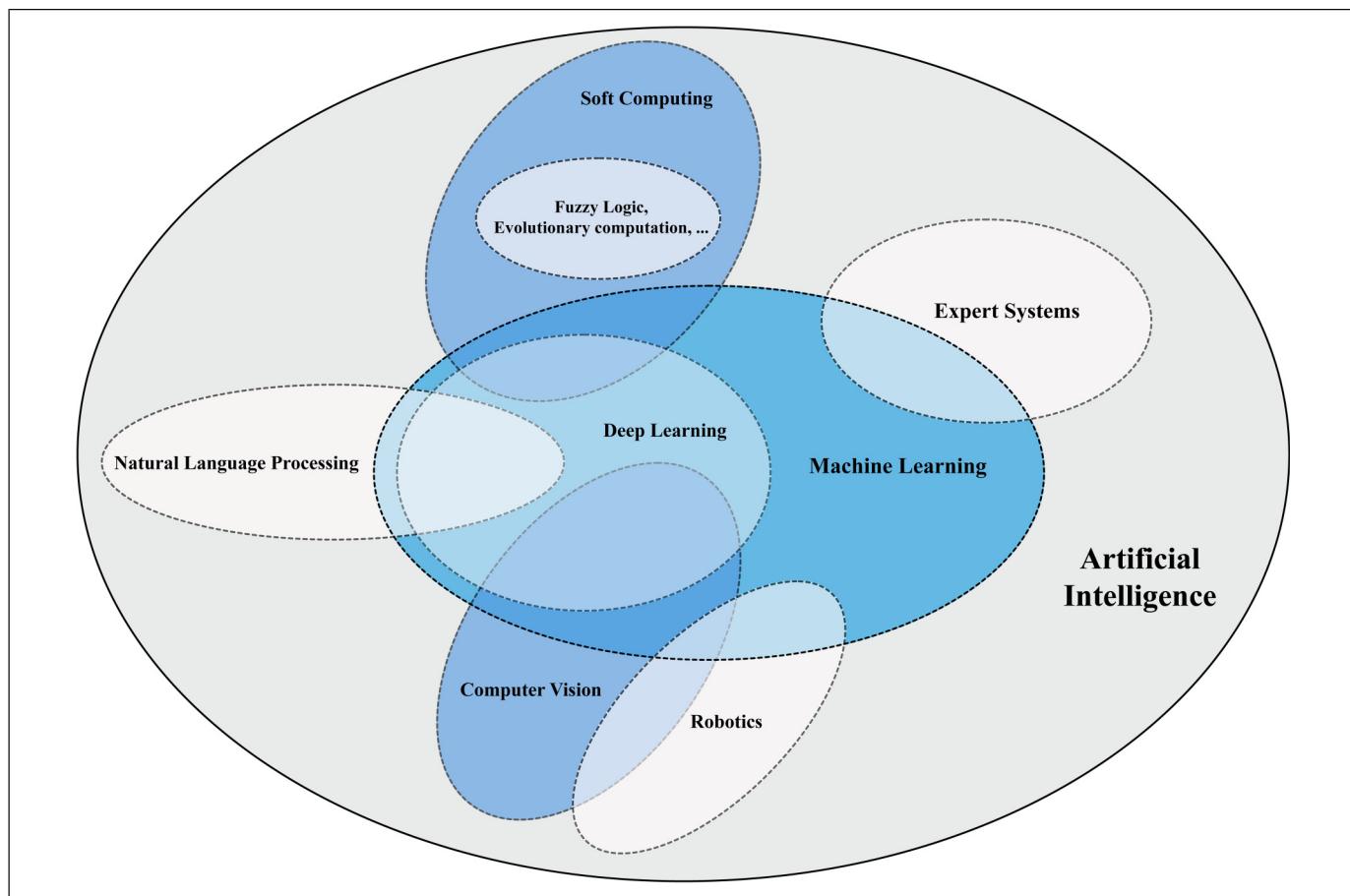


Figure 2. AI subfields and the strategies applied to cancer research in this study. Blue areas emphasize the focus of this review.

Deep Learning in Cancer Research

Deep learning, another subfield of AI, has gained significant attention in cancer research because of its ability to automatically learn complex features from large datasets, making it particularly suitable for tasks such as image analysis and molecular profiling. Several studies have reviewed the application of deep learning in cancer research, highlighting its potential in advancing cancer diagnosis, prognosis, and treatment prediction.^{57,58}

In addition, a large-scale synthetic pathological image dataset called SNOW for breast cancer research, which aids in computational pathology by providing diverse data for deep learning model training, was introduced.⁵⁹ A hybrid technique combining deep learning architectures with machine learning classifiers and fuzzy min–max neural networks for cervical cancer diagnosis, achieving high classification accuracy in Pap smear image classification, was proposed by Kalbhor et al.⁶⁰ Meanwhile, a deep-learning-based method was developed to automatically identify circulating tumor cells and cancer-associated fibroblasts, with significantly better results than conventional computer vision methods.⁶¹ In addition, an enhanced breast histopathology image analysis for cancer detection using variational autoencoders and convolutional neural networks (CNNs) was proposed, resulting in improved prediction accuracy.⁶²

Soft Computing in Cancer Research

Soft computing, a subfield of AI encompassing fuzzy logic and evolutionary algorithms, has also been widely applied in cancer research, offering unique advantages in dealing with uncertainty and imprecision in cancer data. A hybrid approach using the cuckoo search (CS) algorithm and artificial bee colony (ABC) for optimizing feature selection in microarray data for cancer classification was developed.⁶³ The proposed method improved the performance of the Naïve Bayes classifier and provided better results than other feature selection algorithms. In addition, in the work of Gumaiei et al.,⁶⁴ a method for prostate cancer diagnosis using correlation feature selection (CFS) and random committee (RC) ensemble learning was proposed. The technique achieved a 95.098% accuracy rate, outperforming other methods on the same dataset. Indeed, a computer-aided diagnosis system for skin cancer based on soft computing techniques, including CNNs and satin Bowerbird optimization (SBO), was introduced in the work of Xu et al.⁶⁵ The system demonstrated high accuracy, sensitivity, and specificity compared with other methods in the literature. A combined approach to bilevel feature selection techniques using soft computing methods for classifying colon cancer was proposed in the work of Prabhakar et al.⁶⁶ The method combined the multivariate minimum

redundancy–maximum relevance (MRMR) technique with optimization techniques such as invasive weed optimization (IWO) and teaching learning-based optimization (TLBO). The approach achieved a classification accuracy of 99.16% when using quadratic discriminant analysis (QDA). A medical decision support system for assessing gastric cancer risk factors based on fuzzy cognitive maps (FCMs) was presented in the work of Mahmoodi et al.⁶⁷ The system showed higher accuracy than other decision-making algorithms, including decision trees, Naïve Bayes, and artificial neural network (ANN). In addition, a cloud-based breast cancer prediction system that uses soft computing approaches, including Type-1 fuzzy logic (T1F) and SVM, was constructed.⁶⁸ The system achieved a higher precision rate in breast cancer detection than other methods, with BCP-SVM providing 97.06% accuracy. Figure 3 shows the frequency of the use of machine learning, deep learning, and soft computing algorithms in cancer research between 2010 and 2023.

The section details the key algorithms and performance metrics in Machine Learning, Soft Computing, and deep learning, with a systematic overview of studies using these methods in cancer research presented in a table (referenced as Figure 4). It also discusses the challenges and prospects of AI in oncology. The aim was to offer an impartial, informative view on the current AI applications in this field, tailored for a wide audience including researchers and clinicians from various disciplines, without giving personal opinions on individual studies due to space constraints. This approach seeks to provide a clear, objective insight into the extensive application and potential of AI in cancer research.

Performance Metrics

Below, we indicate the formulas for several quantitative measures used to evaluate the performance of the proposed framework in the literature, including accuracy, sensitivity, specificity, precision, recall, F1-score, Matthews's correlation coefficient (MCC), receiver operating characteristic (ROC) curve, and area under the curve (AUC).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{specificity} = \frac{TN}{TN + FP} \quad (3)$$

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

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

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (7)$$

$$AUC = \frac{S_p - n_p(n_n + 1)/2}{n_p \times n_n} \quad (8)$$

where TP, TN, FP, and FN denote true positives, false negatives, and false negatives, respectively. These values are derived from the confusion matrix. For AUC, S_p denotes the sum of the ranks of all positive samples, and n_p and n_n denote the number of positive and negative samples, respectively.

Pre-Processing

The preprocessing step is crucial in preparing the raw data and making it compatible with the machine-learning problem. The main objective of preprocessing is to ensure the model's convergence and improve its training speed. Various techniques and tools are used to preprocess data, including the following steps.

1. **Data cleaning** mainly involves handling missing values, checking for outliers and inconsistent data, eliminating errors, and transforming the data into a uniform format.
2. **Data transformation** is required when attributes are of different types (structured data, medical images, free

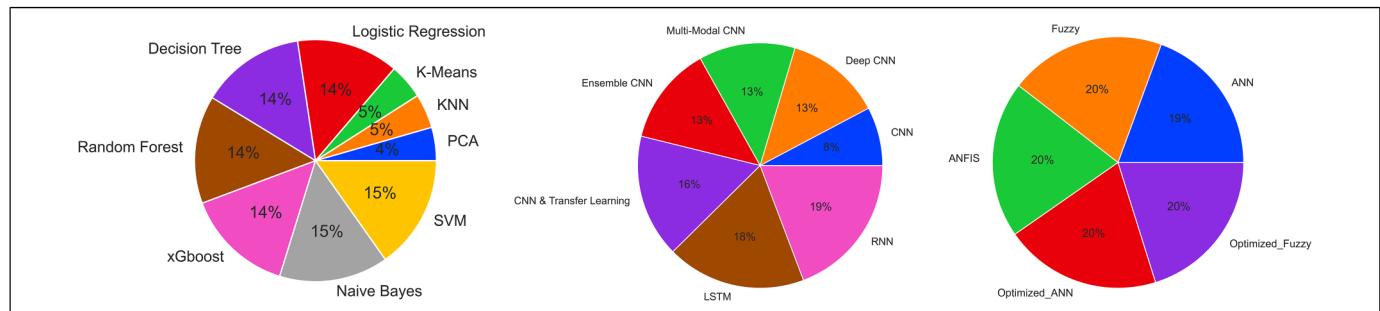


Figure 3. Frequency of algorithms used in different AI subfields in cancer research (obtained from PubMed database¹ between 2010 and 2023).

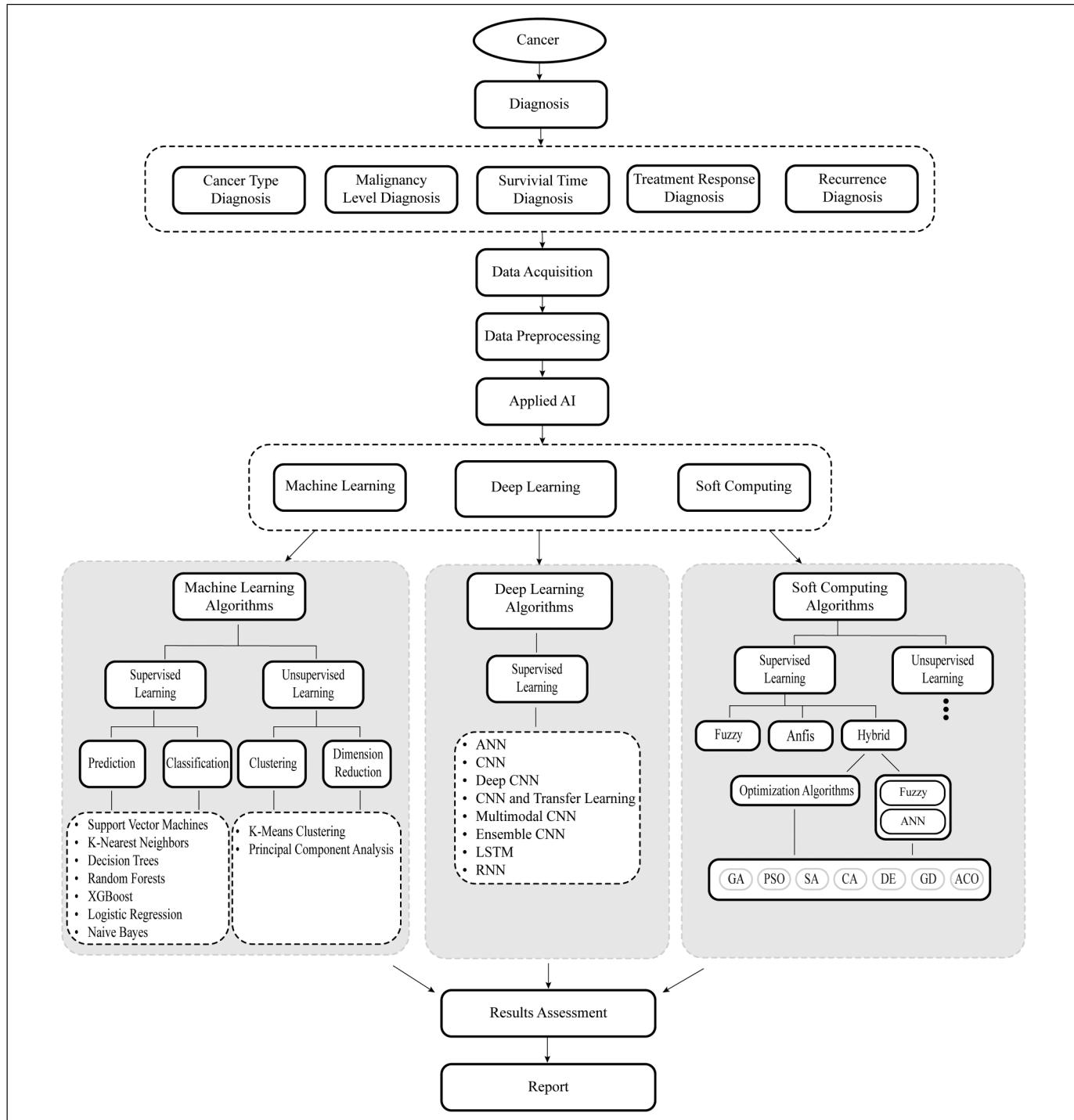


Figure 4. The workflow of each article investigated in this article (AI: artificial intelligence, XGBoost: eXtreme Gradient Boost, ANN: artificial neural network, CNN: convolutional neural network, LSTM: long short-term memory networks, RNN: recurrent neural network, Anfis: adaptive network-based fuzzy inference system).

- text, etc). It converts it to a form that is understandable and acceptable for machine learning models.
3. **Dimensionality reduction** is used to reduce the dimensionality of a dataset having extensive features or variables, leading to a less complex architecture and preventing overfitting issues.
 4. **Data normalization** is used to maintain the general distribution in the dataset, ensure the model's convergence, and improve the model's training speed.
 5. **Feature extraction** produces new low-dimensional features that represent the original high-dimensional

features. This is done by identifying the important features or patterns of the data.

6. **RGB to grayscale conversion** is used to convert input RGB images to grayscale images.
7. **Data augmentation** is a strategy used to extend the data with a set of transformations of the original data, helping overcome overfitting when training on very little data.
8. **Image quality enhancement** is used to improve the visibility of features in an image by changing its attributes, such as contrast and brightness.
9. **Image denoising** is used to restore an image to its original quality by reducing or removing noise to guarantee accurate information analysis while preserving vital details such as edges and texture information.
10. **Image filtering** is a technique to remove unwanted features from images, such as noise, bubbles, and hair.
11. **Image segmentation** is used to divide an image into several segments or objects, which is a crucial stage in computer vision.

Machine Learning Application to Cancer Research

This section explores machine learning algorithms like SVM, Naïve Bayes, Decision Trees, Random Forests, K-means, KNN, Logistic Regression, eXtreme gradient boost (XGBoost), and Hybrid models within cancer research. It examines how these algorithms, with their distinct advantages in accuracy, adaptability, or scalability, analyze vast datasets to uncover new biomarkers, craft personalized treatments, and enhance early detection, ultimately improving patient outcomes. The discussion

will detail each algorithm's features, strengths, and specific uses in oncology.

Support Vector Machines

SVM is a machine learning methodology that essentially functions as a mechanism for sorting data into separate categories by determining the optimal partition or border between them. In the realm of cancer research, SVM serves as a tool for classification, such as differentiating malignant (class 1) from benign tumors (class 2), as shown in Figure 5. Consider a dataset with information on tumor size, shape, and texture collected from many patients. Based on these attributes, the SVM model can use these data to identify the most suitable demarcation between malignant and benign tumors. Based on these features, the SVM algorithm finds a decision boundary (separating hyperplane) that separates malignant and benign tumors in a multi-dimensional space. The support vectors are the tumors closest to this boundary, which "support" the position of the hyperplane. The margin is the distance between the hyperplane and support vectors, which the SVM algorithm seeks to maximize for a more accurate classification. Once the model is trained, it can predict tumor classification for new patients, thus empowering physicians to make more well-informed choices regarding treatment plans.

In the Supplemental materials, Table 1 showcases studies effectively using the SVM algorithm for cancer classification tasks, highlighting its practical impact and value.

Decision Tree and Random Forest

Decision trees and random forests serve as machine-learning algorithms that derive conclusions through a series of inquiries,

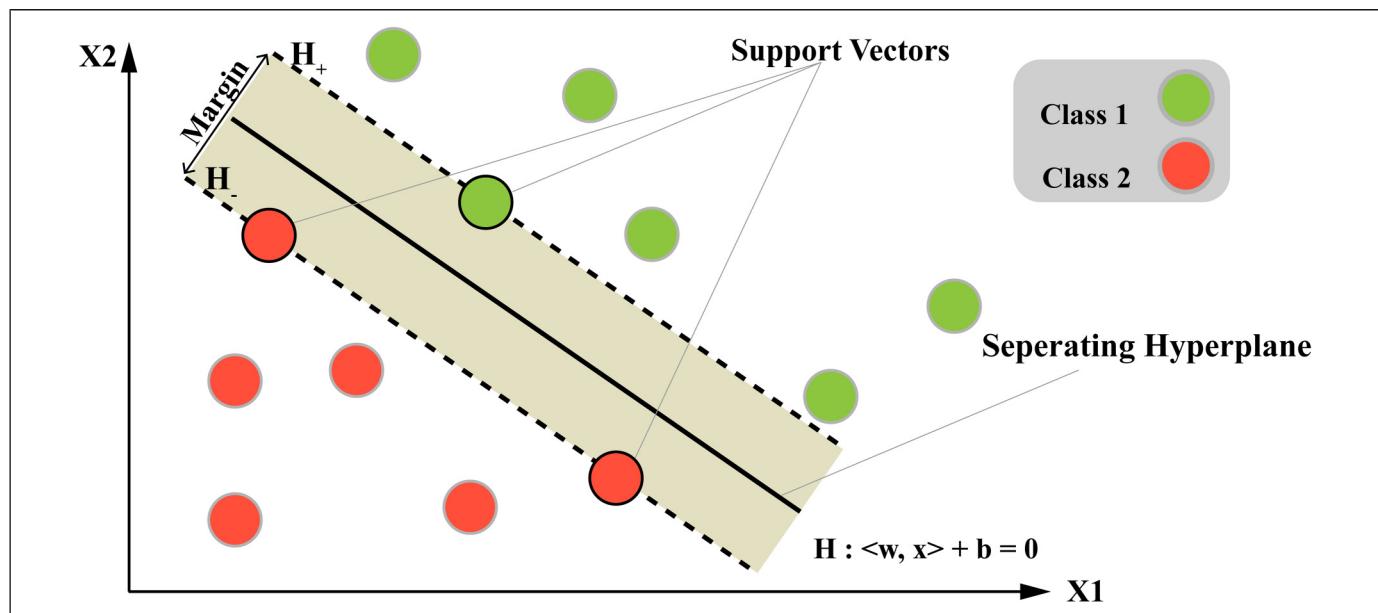


Figure 5. Structure of the support vector machine (SVM) algorithm.

similar to a flowchart. As depicted in Figure 6, decision trees engage a series of binary questions regarding the dataset, branching at each stage (decision nodes) to arrive at a conclusive decision/final result based on majority voting. In contrast, random forests comprise several decision trees, which collectively operate on distinct feature subsets to yield a more precise and consistent prediction. For instance, in cancer research, these models can be employed to categorize tumors as benign or malignant. A decision tree might pose queries such as “Does the tumor exceed 2 cm in size?” or “Do the cells exhibit irregular shapes?” The tree navigates the decision-making pathway based on the responses until a final categorization is determined. In the case of a random forest model, multiple decision trees are utilized, each examining a random data subset, and their outcomes are combined via voting to produce a more reliable classification.

In the Supplemental materials, Table 2 details significant studies that have employed decision trees and random forests algorithms in cancer research, illustrating their practical uses and contributions to classification challenges.

K-Nearest Neighbors

The KNN model is a simple and intuitive machine-learning algorithm that classifies data based on the properties of its closest neighbors. As shown in Figure 7, KNN identifies the K most similar data points to a new, unknown data point and assigns it to the majority class of its neighbors. For instance, in cancer research, KNN can be used for various purposes, such as predicting the response to a specific treatment or determining cancer subtypes. If a new patient’s tumor characteristics are input into the KNN model, it will identify the K most similar cases from the

existing data. Then, it will determine the majority outcome among those neighbors, such as a successful treatment response or a specific cancer subtype, and predict that the new patient will likely experience a similar outcome.

In the Supplemental materials, Table 3 compiles studies that have leveraged the KNN algorithm in cancer research, emphasizing its adaptability and practical applications in classification tasks.

K-Means

The k-means model is a straightforward clustering algorithm that groups data points based on their similarity. This algorithm is illustrated in Figure 8 and works by identifying the “k” number of cluster centroids (centers) and assigning each data point to the nearest centroid, creating distinct clusters. In cancer research, k-means is applied for clustering purposes, such as discovering subgroups of patients with similar gene expression patterns or tumor characteristics. For example, given a gene expression data dataset from different patients, the k-means algorithm would find “k” centroids representing distinct gene expression profiles. Then, each tumor is assigned to the nearest centroid, creating groups of tumors with similar gene expression patterns. This information can help researchers identify new cancer subtypes or understand the underlying biological mechanisms.

In the Supplemental materials, Table 4 gathers key studies employing the k-means algorithm in cancer research, demonstrating its practical uses and contributions to solving clustering issues.

Logistic Regression

Logistic regression is a model that predicts the probability of an event occurring on the basis of one or more input variables. As

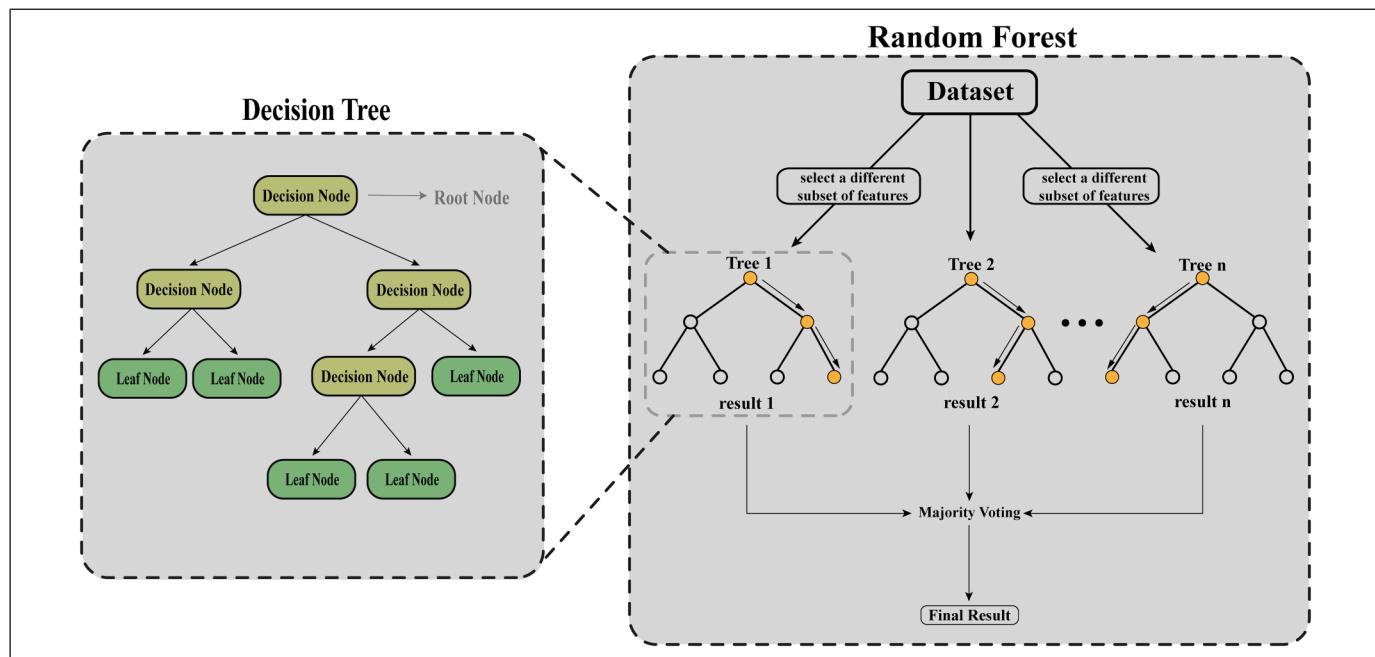


Figure 6. Structure of the decision tree and random forest algorithms.

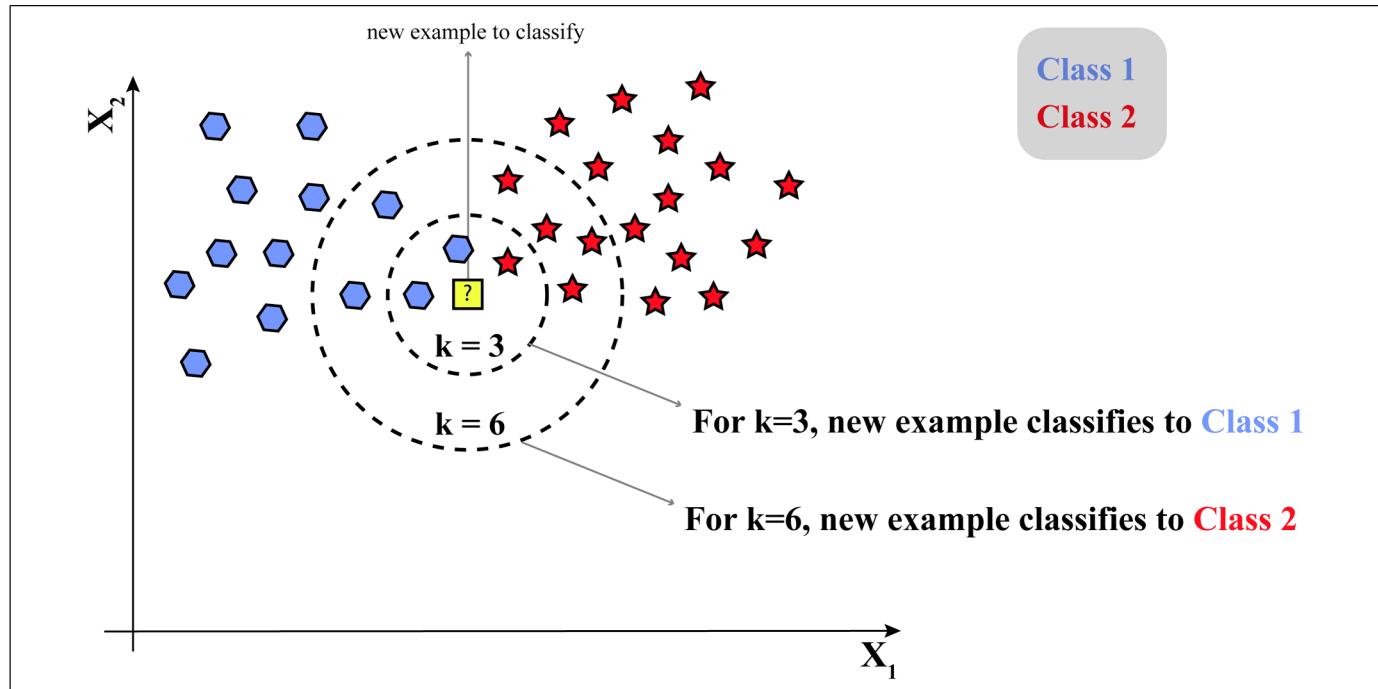


Figure 7. Structure of the K-nearest neighbors (KNN) algorithm.

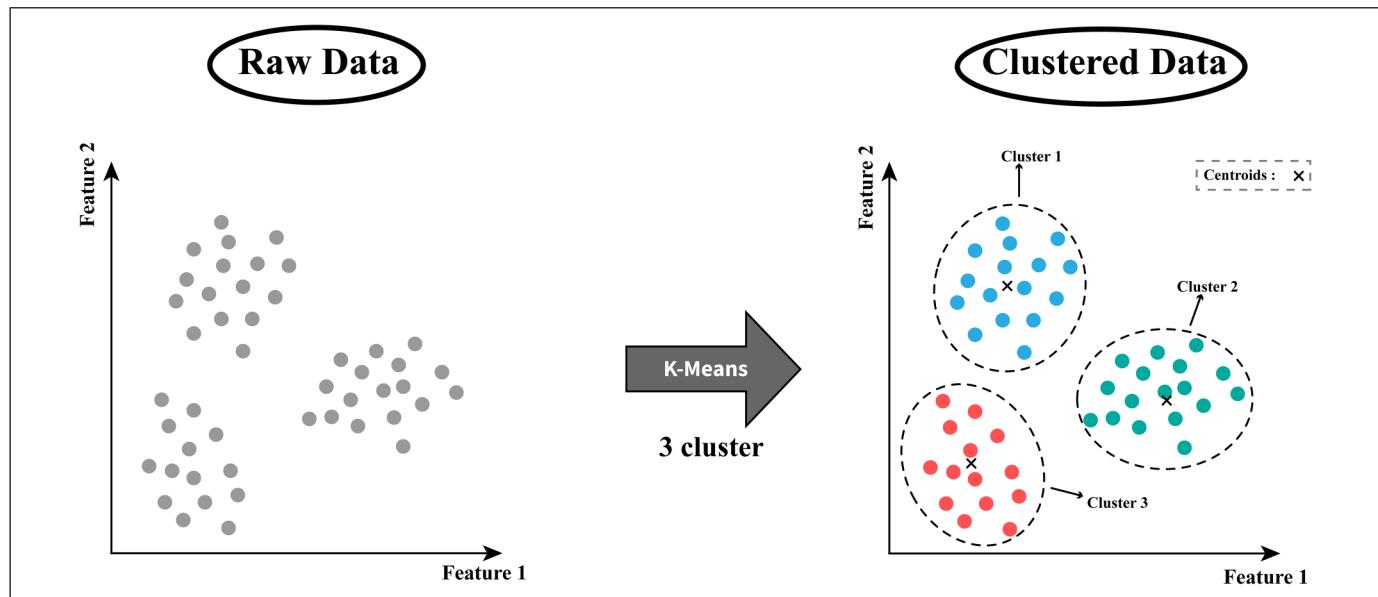


Figure 8. Structure of the K-means algorithm.

shown in Figure 9, logistic regression uses a particular S-shaped curve, called the logistic function, to model the relationship between the variables. This allows it to produce probabilities between 0 and 1, which can be converted into binary outcomes using a threshold value. In cancer research, logistic regression is applied for various purposes, such as predicting the likelihood of a patient's tumor being malignant based on factors such as age, tumor size, and genetic markers. The model generates an

S-shaped curve representing the probability of malignancy as a function of the input variables. Researchers can classify tumors as malignant or benign by setting a threshold on this curve.

In the Supplemental materials, Table 5 outlines important studies that have used the logistic regression model in cancer research, underscoring its practical implementations and importance in addressing classification challenges.

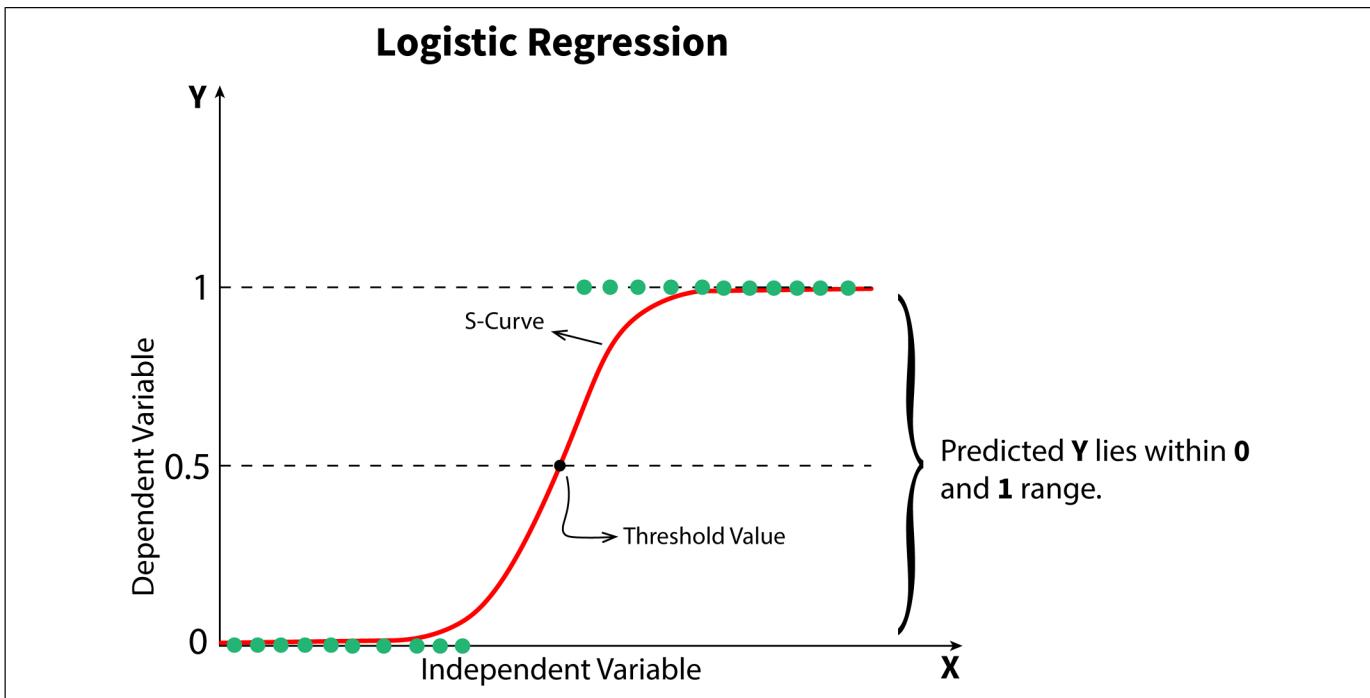


Figure 9. Structure of the logistic regression algorithm.

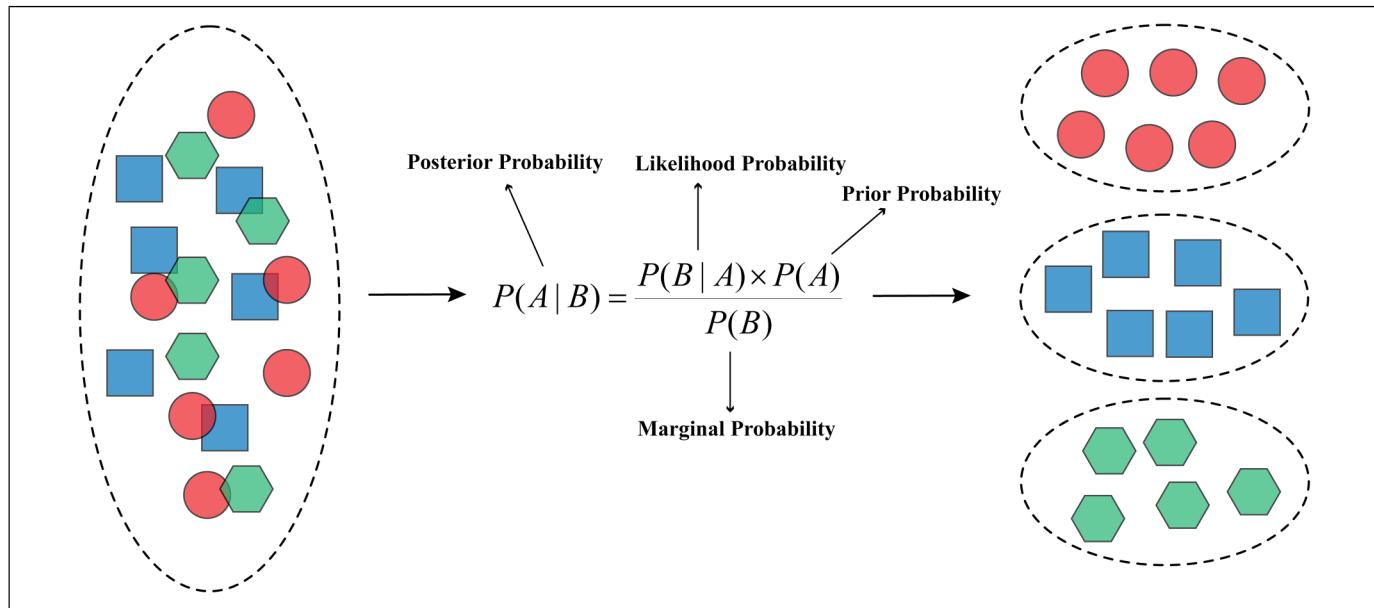


Figure 10. Structure of the Naïve Bayes algorithm.

Naïve Bayes

The Naïve Bayes model is a simple yet powerful probabilistic classifier based on Bayes' theorem. It works by calculating the probability of each class (eg, benign or malignant tumor) given the input features, assuming that the features are conditionally independent. As shown in Figure 10, the model combines prior probabilities (known information about the classes), likelihoods (the probabilities of observing the features

given a class), and marginal probabilities (the probabilities of observing the features) to compute the posterior probabilities, which represent the updated probabilities of each class. In cancer research, the Naïve Bayes model can be used for classification, such as diagnosing cancer based on a set of symptoms or medical test results. Given the input features, the algorithm calculates the posterior probabilities of each class (eg, cancerous or non-cancerous) and assigns the unknown case to the class with the highest probability.

In the Supplemental materials, Table 6 presents studies that have applied the Naive Bayes algorithm in cancer research, showing its practical uses and effects on classification issues.

eXtreme Gradient Boost

The XGBoost model is an advanced machine learning algorithm that builds on the concept of boosting, which combines multiple weak models to create a robust and accurate model. As shown in Figure 11, in XGBoost, a series of decision trees are constructed sequentially, with each tree focusing on correcting the errors (residual values) made by the previous tree in the series. This process continues until a specified stopping criterion is met, resulting in a powerful classifier. In cancer research, the XGBoost model can be employed for classification purposes, such as predicting cancer recurrence based on patient characteristics, genetic data, and treatment history. By iteratively building decision trees that focus on correcting previous errors, XGBoost generates a model that can accurately distinguish between patients with high and low risks of cancer recurrence.

In the Supplemental materials, Table 7 compiles significant studies using the XGBoost algorithm in cancer research, emphasizing its practical applications and efficacy in classification tasks.

Hybrid Models

Hybrid machine learning algorithms refer to the combination of different machine learning methods or models to create intelligent algorithms that leverage the strengths of each individual component. These algorithms improve overall performance,

enhance accuracy, or address specific challenges that cannot be effectively tackled by a single method alone.

For example, a hybrid machine-learning algorithm could combine the power of decision trees, SVM, and ANNs to classify tumor types based on genetic markers, patient characteristics, and medical imaging data. Decision trees can capture simple decision rules, SVMs can handle complex decision boundaries, and ANNs can learn intricate patterns from the data. By integrating these techniques (Figure 12), the hybrid algorithm can leverage the complementary strengths of each method to achieve more accurate and robust tumor classification.

This hybrid approach allows for a more comprehensive analysis of cancer data, taking advantage of multiple machine-learning techniques to tackle the complexity and heterogeneity of cancer-related datasets. Combining different methods intelligently, hybrid algorithms can improve classification and prediction outcomes, leading to better insights and decision-making in cancer research.

In the Supplemental materials, Table 8 features studies that have effectively used hybrid algorithms in cancer research for classification tasks, showcasing their real-world effectiveness and value.

Figure 13 summarizes the steps for applying different types of machine learning algorithms to cancer research, which have already been described one by one.

Soft Computing Application in Cancer Research

Soft computing algorithms, which include fuzzy logic and evolutionary algorithms optimized using evolutionary algorithms,

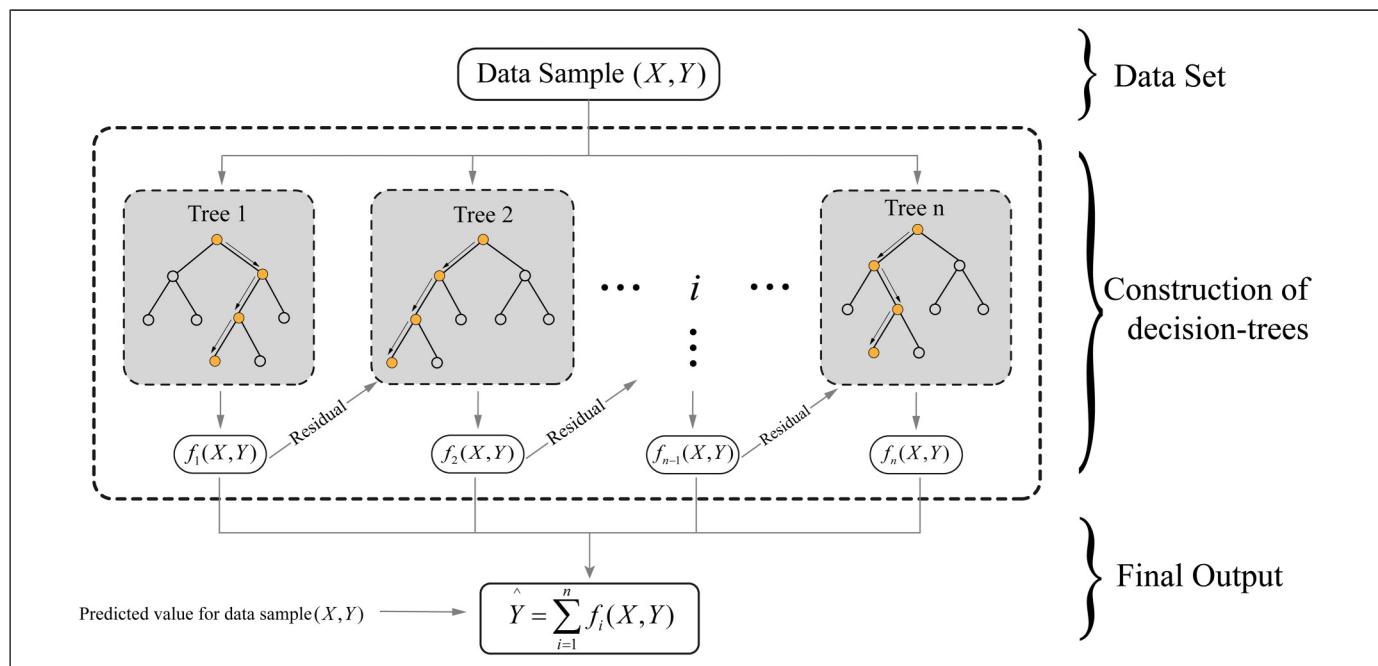


Figure 11. Structure of the eXtreme Gradient Boosting (XGBoost) algorithm.

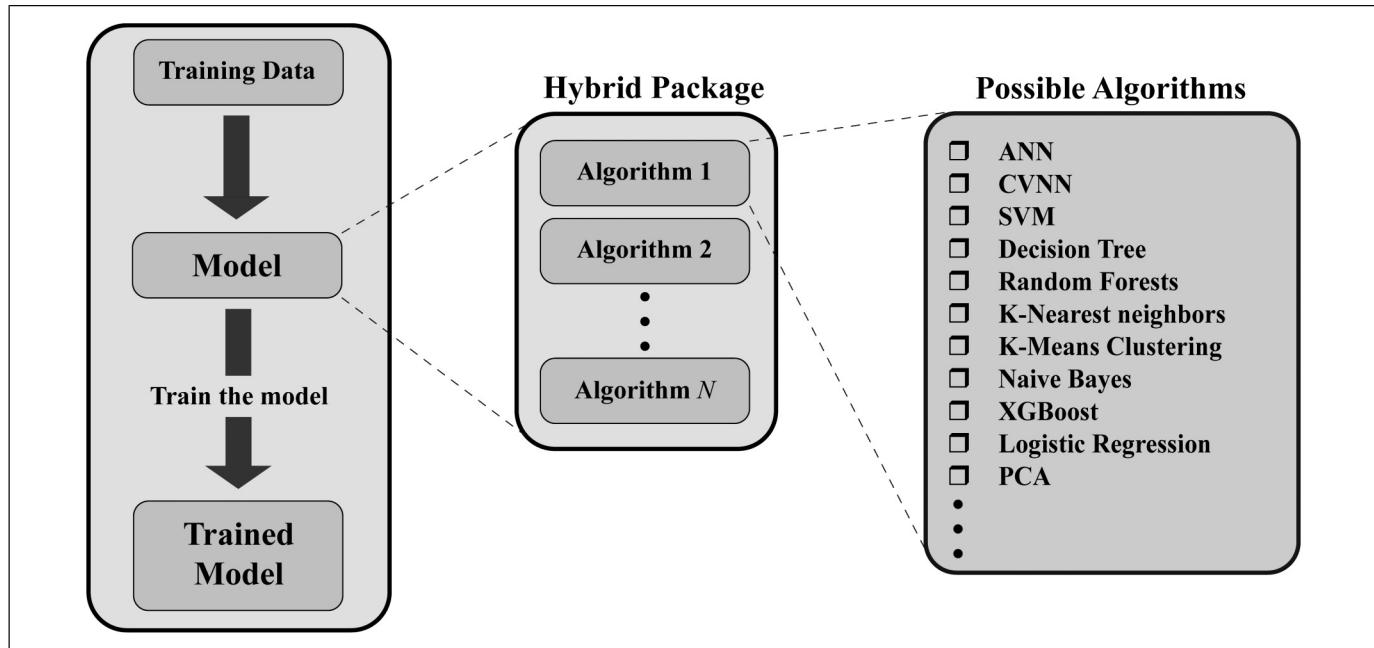


Figure 12. The structure of the hybrid algorithm (ANN: artificial neural networks, CVNN: complex-valued neural network, SVM: support vector machine, XGBoost: eXtreme Gradient Boost, PCA: principal component analysis).

have shown great promise in advancing cancer research. Fuzzy logic provides a means of handling uncertainty and imprecision in medical data, allowing for more accurate diagnoses and personalized treatment plans. Conversely, ANNs can learn complex relationships between input data and output variables, making them helpful in predicting cancer outcomes and identifying novel biomarkers. Evolutionary optimization algorithms, such as genetic algorithms and particle swarm optimization, can efficiently search through large datasets to find optimal solutions, thus improving cancer diagnosis and treatment planning accuracy. By incorporating these soft computing techniques into cancer research, researchers have significantly improved patient outcomes, including earlier detection and more effective treatment options. In the following section, we describe each algorithm and then include a separate table covering the most significant works using each algorithm.

Fuzzy Inference System

The fuzzy inference system (FIS) is a computational framework that deals with uncertainty and imprecision using fuzzy logic, which allows for partial truth values between 0 and 1 rather than binary true or false. As shown in Figure 14, the FIS consists of 4 main steps: “Fuzzification,” “Knowledge Base,” “Inference Engine,” and “Defuzzification.” In cancer research, FIS is used for classification and prediction purposes, such as determining the risk of cancer recurrence or the effectiveness of a specific treatment. For example, input variables such as tumor size, age, and genetic markers may have associated uncertainty. FIS enables the processing of this uncertain information through fuzzification, which converts the crisp input

values into fuzzy sets representing degrees of membership in different categories. The fuzzy inference engine uses a knowledge base to apply rules to the fuzzy input data to generate fuzzy output values. Finally, defuzzification converts the fuzzy output values back into crisp values, providing a precise prediction or classification, such as the low or high risk of cancer recurrence.

In the Supplemental materials, Table 9 outlines prominent studies utilizing the Fuzzy Inference System in cancer research.

Hybrid Soft Computing Algorithms

Hybrid soft computing algorithms that combine optimization methods and ANN or FIS offer a powerful approach to solving complex problems in cancer research. As shown in Figure 15, optimization methods, such as genetic algorithms or particle swarm optimization, are used to fine-tune the ANN or fuzzy systems, thereby enhancing their performance and adaptability. In cancer research, a hybrid algorithm that combines optimization methods and ANN or FIS can be applied for classification and prediction purposes. For example, imagine a scenario in which researchers want to predict the effectiveness of a specific cancer treatment based on a patient’s genetic data and clinical variables. First, the FIS is designed to process and analyze the input data using fuzzy rules that account for the inherent uncertainty and imprecision in the information. The FIS outputs a prediction of treatment effectiveness in the form of a fuzzy membership value. Next, an optimization method, such as a genetic algorithm, is employed to fine-tune the FIS parameters, including the fuzzy rules, membership functions, and

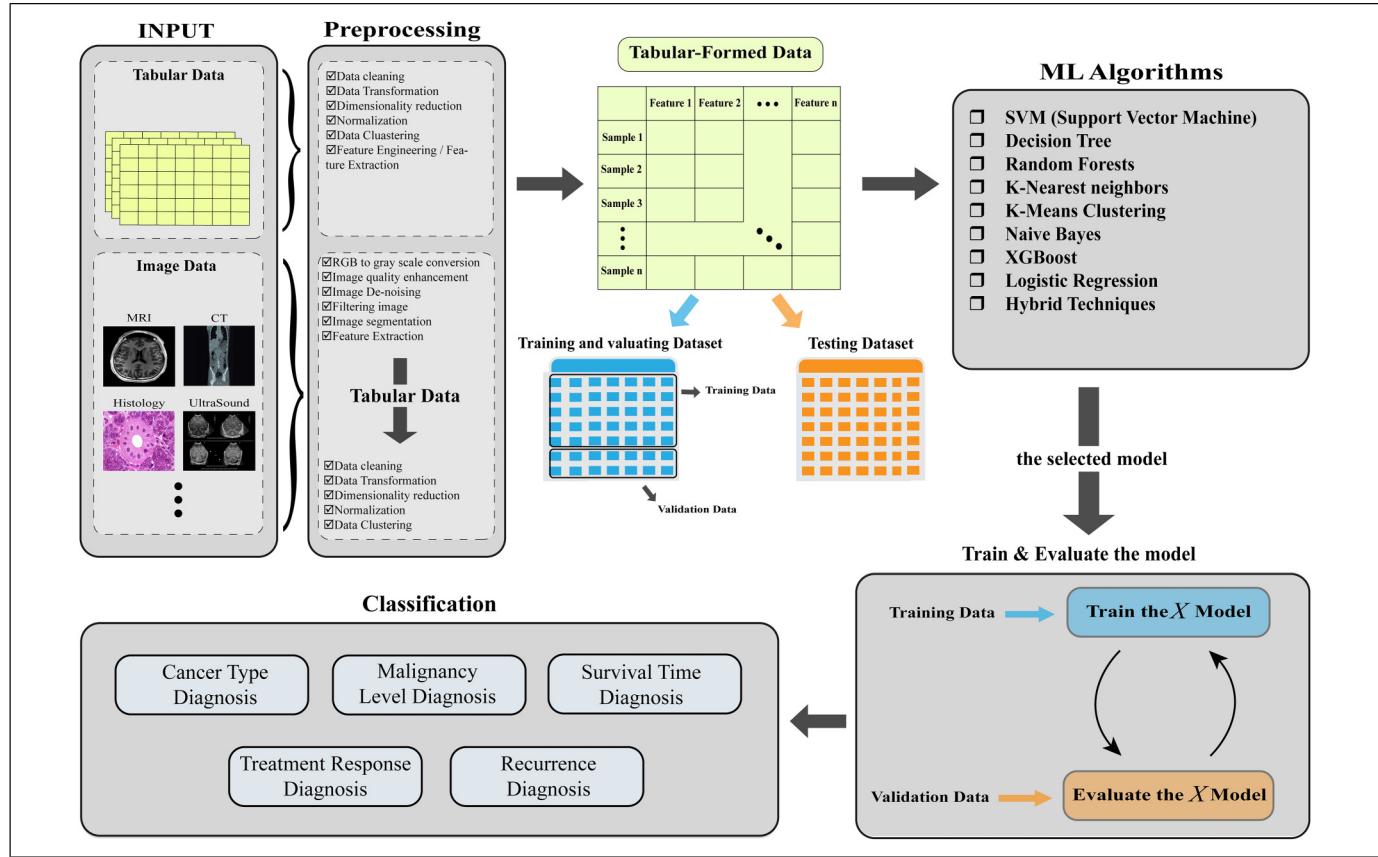


Figure 13. A brief overview of the necessary steps for applying different ML algorithms to cancer research.

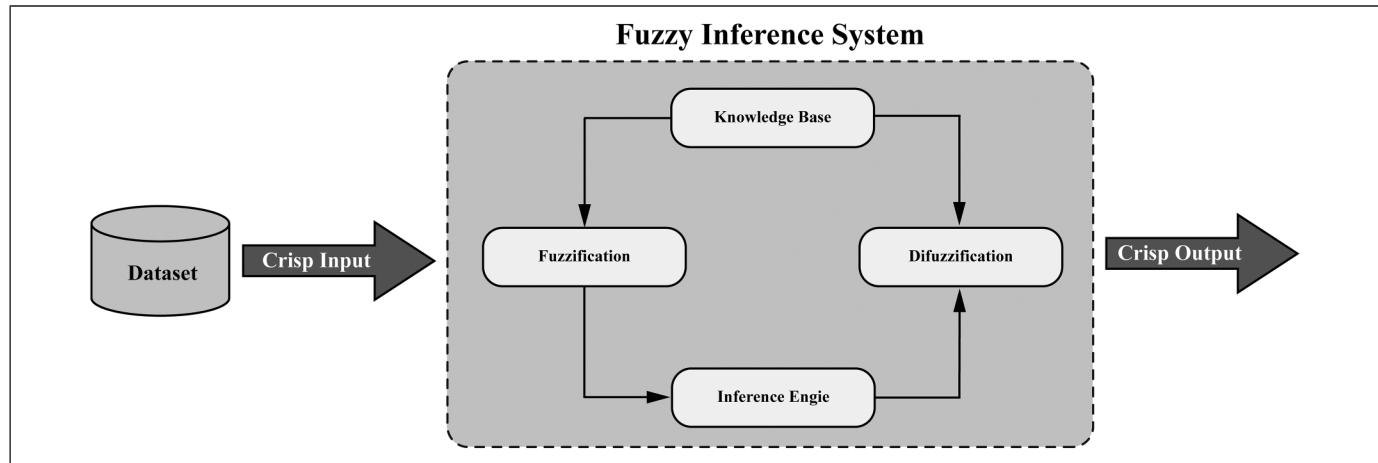


Figure 14. Structure of the fuzzy inference system (FIS) algorithm.

aggregation operators. This optimization process iteratively refines the fuzzy system's performance, resulting in a more accurate and reliable prediction model.

By combining the strengths of FIS and optimization methods, hybrid soft computing algorithms offer a robust and

flexible approach to cancer classification and prediction tasks, capable of handling the complex and uncertain nature of cancer-related data.

In the Supplemental materials, Table 10 compiles important studies that have used optimization algorithms with ANN or

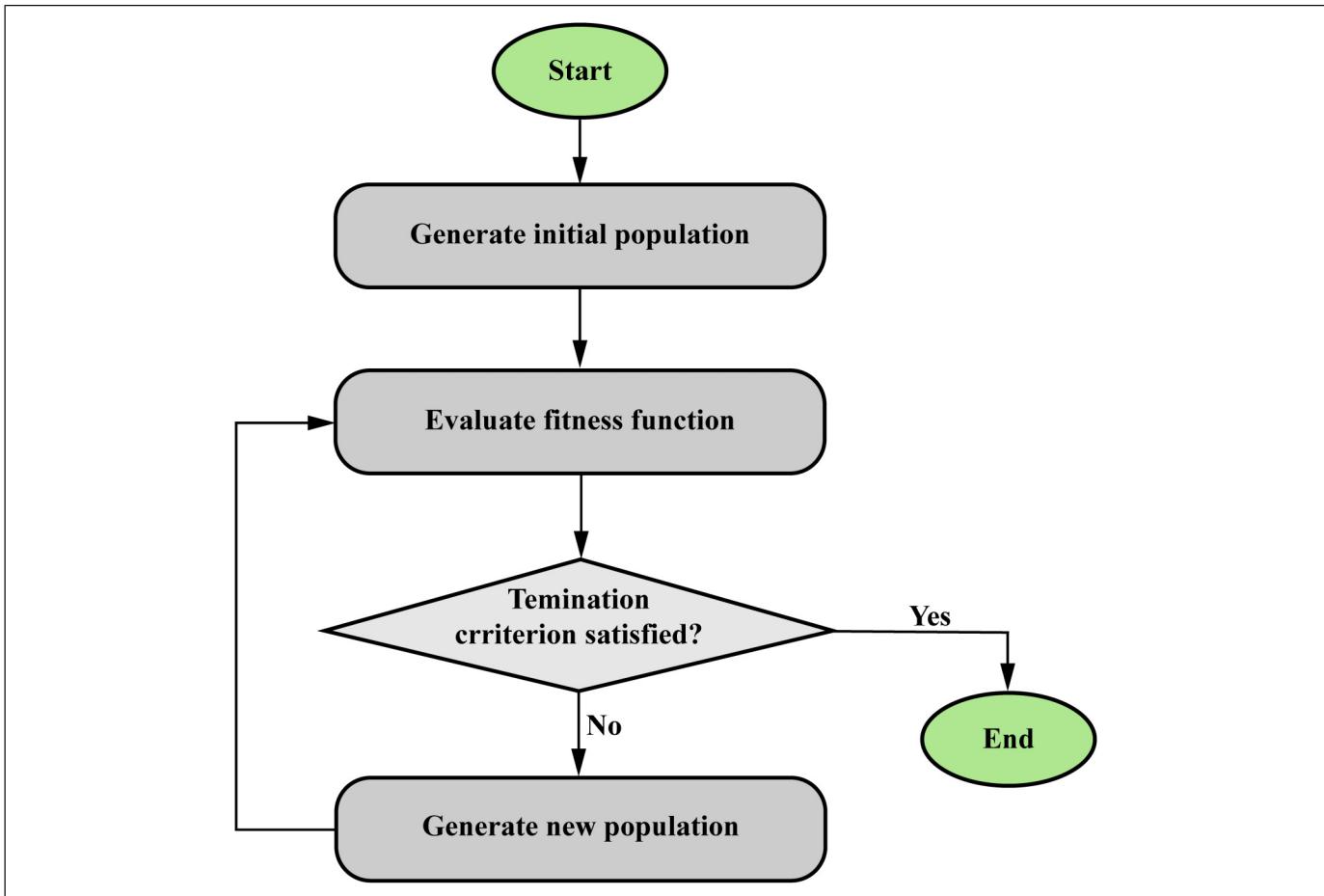


Figure 15. Structure of optimization algorithms (evolutionary algorithms).

FIS in cancer research, emphasizing their practical uses and effectiveness in the field.

Deep Learning

Deep learning algorithms are advanced computational models capable of processing and analyzing complex visual data. These algorithms include ANNs, CNN, multimodal CNN, and recurrent neural networks (RNNs) combined with long short-term memory (LSTM) cells. These models are designed to address specific challenges in processing and learning visual information.

ANNs are computational models inspired by the structure and functioning of the human brain. As shown in Figure 16, in an ANN applied to cancer research, the inputs consist of relevant data features such as patient demographics, tumor characteristics, genetic markers, and medical images. These input features are passed through multiple layers of neurons, where they are processed and transformed. Each layer consists of numerous interconnected neurons that apply weights and biases to the incoming data, enabling the network to learn complex relationships and patterns. After the information passes through the hidden layers, it reaches the output layer, which generates the final predictions or classifications, such

as the likelihood of cancer recurrence or the type of cancer. The output layer is designed to match the desired format, such as a single output for binary classification or multiple outputs for multi-class classification. In summary, ANN processes cancer-related input data through multiple interconnected layers with adjustable weights and biases and generates outputs relevant to cancer classification or prediction tasks.

CNNs (depicted in Figure 17) are specialized deep-learning models designed to process grid-like data, such as images. They use different layers, such as convolutional, pooling, and flattened layers, to scan and identify local patterns or features in an image, such as edges or textures. In cancer research, CNNs can be used for tasks such as classifying histopathological images, identifying cancerous cells, or detecting tumors in medical imaging data, such as MRI or CT scans.

Multimodal CNNs are an extension of traditional CNNs and are designed to process and integrate multiple data types, such as images, text, and audio. As illustrated in Figure 18 by combining information from different sources, these networks can improve classification or prediction performance. In cancer research, a multimodal CNN could be used to simultaneously predict data (eg, patient demographics, and medical history), leading to more accurate and comprehensive predictions.

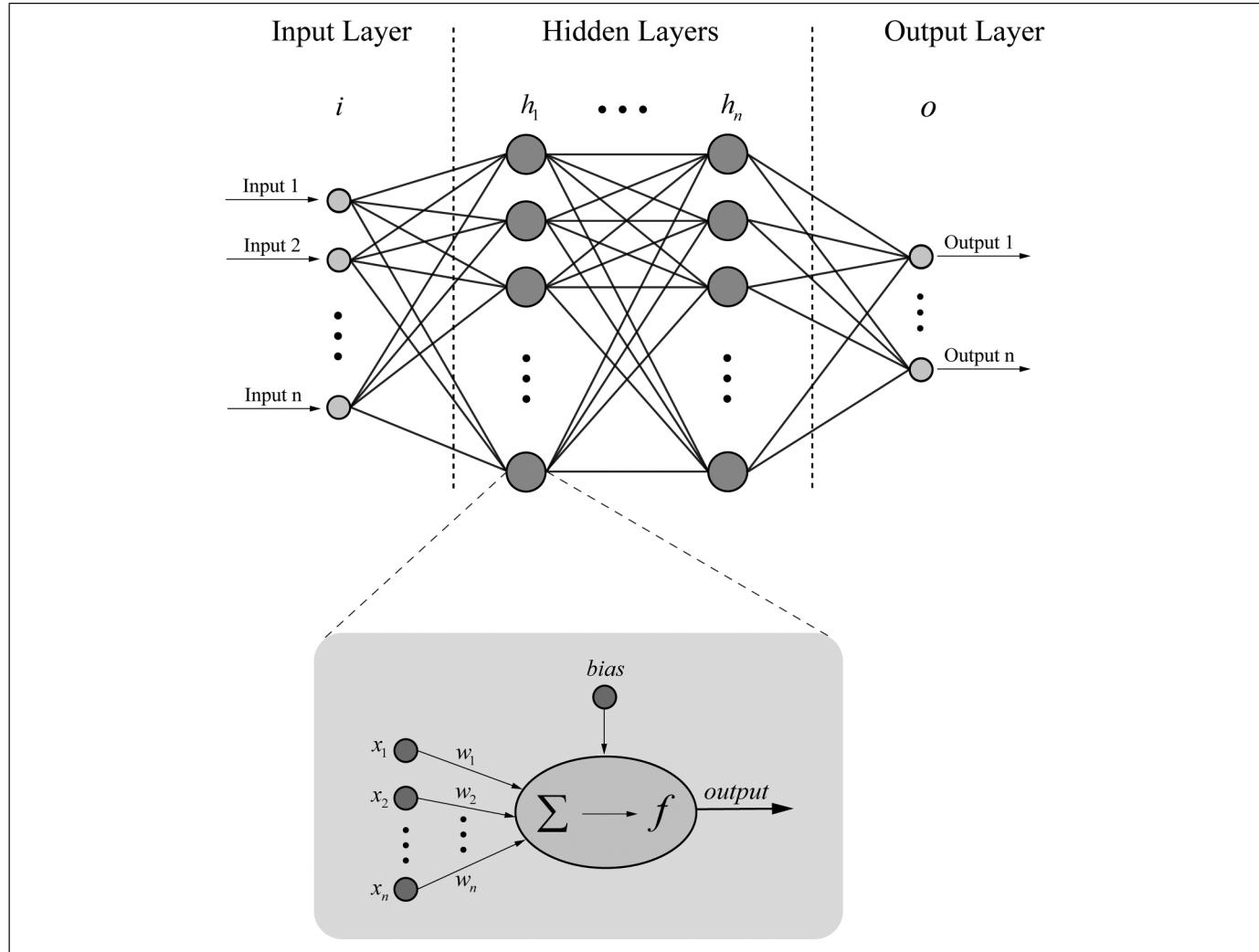


Figure 16. Structure of the artificial neural network (ANN).

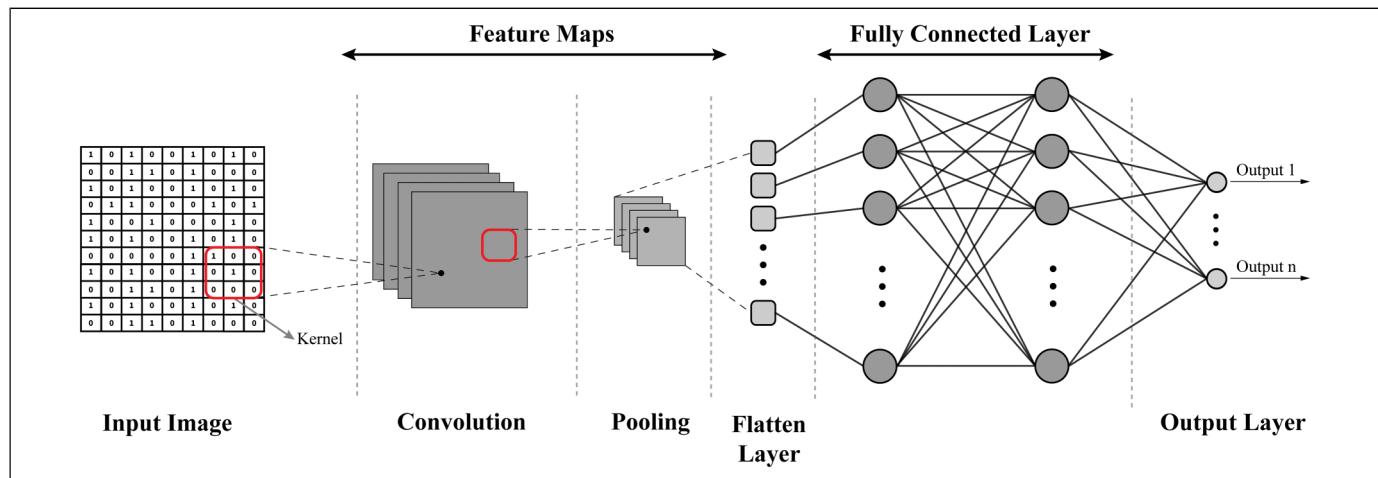


Figure 17. Structure of convolutional neural network (CNN).

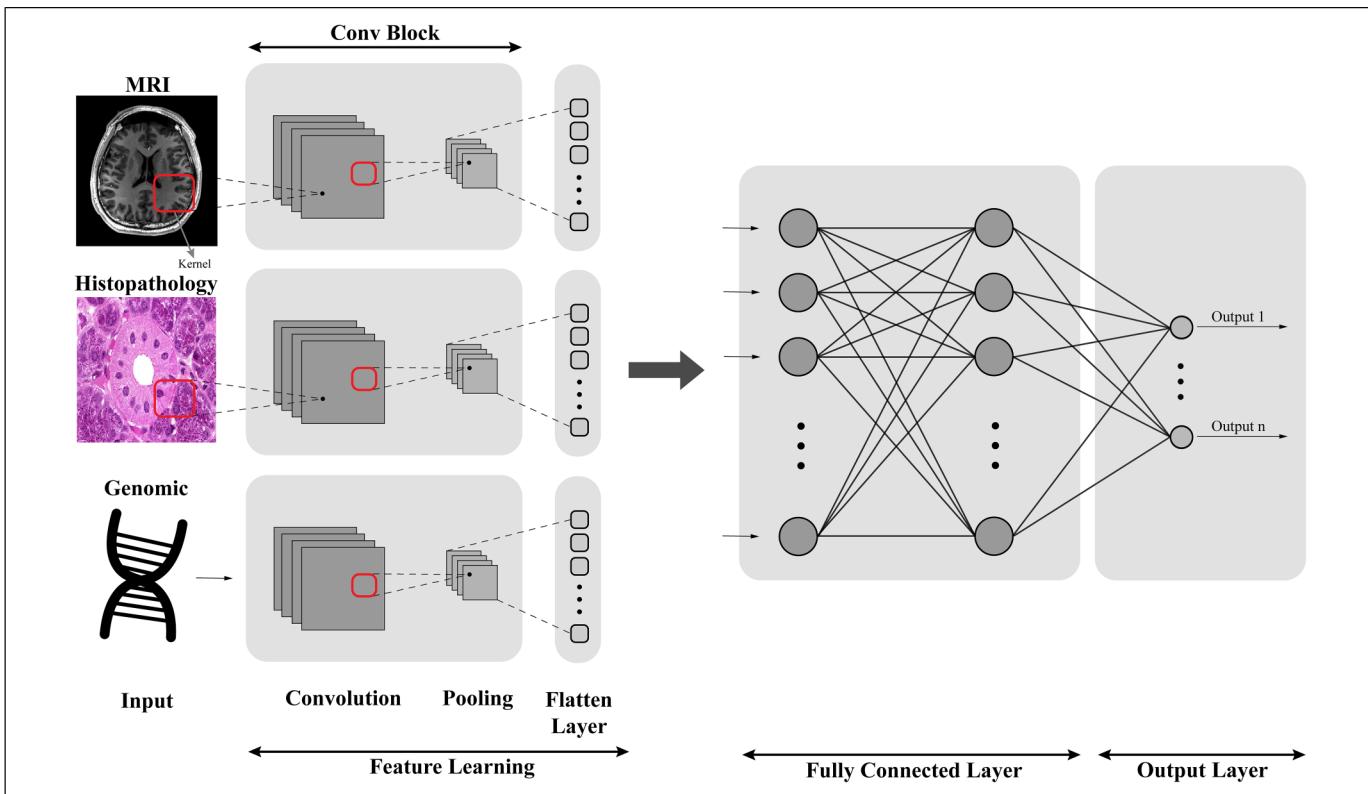


Figure 18. Structure of a multimodal CNN.

Figure 19 depicts the structures of RNNs and LSTM. RNNs are deep learning models designed to process sequential data, making them suitable for tasks involving time series or ordered data. LSTM cells are a type of RNN architecture that addresses the challenge of learning long-term dependencies in the data. In cancer research, RNNs with LSTM cells can be applied to tasks such as predicting cancer progression or recurrence based on time-series data, such as gene expression levels or patient vital signs collected over time.

In summary, deep learning algorithms, including ANNs, CNNs, multimodal CNNs, and RNNs with LSTM cells, offer powerful tools for processing complex visual data in cancer research, enabling accurate classification and prediction tasks that can improve patient outcomes and advance our understanding of cancer biology.

In the Supplemental materials, Table 11 details significant studies using deep learning algorithms in cancer research, demonstrating their practical applications and influence in the field.

Conclusions and Future Work

Rapid progress in AI has profoundly impacted cancer research and treatment, leading to enhanced patient outcomes and healthcare efficiency. Incorporating AI algorithms in cancer diagnosis, prognosis, and treatment response prediction has facilitated early detection, customized intervention approaches, and improved overall patient care. In this comprehensive guide,

we have explored the crucial role of AI in cancer research, with a specific focus on the applications of machine learning, soft computing, and deep learning algorithms. We have offered an in-depth overview of various algorithms' functionality and particular applications, supported by pertinent figures and a tabular summary of key findings from each study with the lowest complexity and high suitability for a better understanding of all readers with different backgrounds.

The impressive advantages of AI-driven algorithms in cancer care emphasize their potential to reshape cancer research and clinical practice. This review serves as an invaluable resource for researchers, clinicians, and healthcare industry stakeholders, offering insights into AI's present state and future potential in cancer care.

Future research in AI for cancer care could explore:

1. Developing advanced AI algorithms to enhance the precision and efficiency of cancer care.
2. Utilizing multi-modal data (eg, genomic, proteomic, imaging, and clinical reports) to gain a comprehensive understanding of cancer, focusing on AI's ability to process and analyze such diverse information.
3. Creating personalized treatment plans using AI to consider individual patient characteristics, aiming for treatments that are both effective and have minimal side effects.
4. Leveraging AI in drug discovery to quicken the identification of drug targets and optimize drug designs,

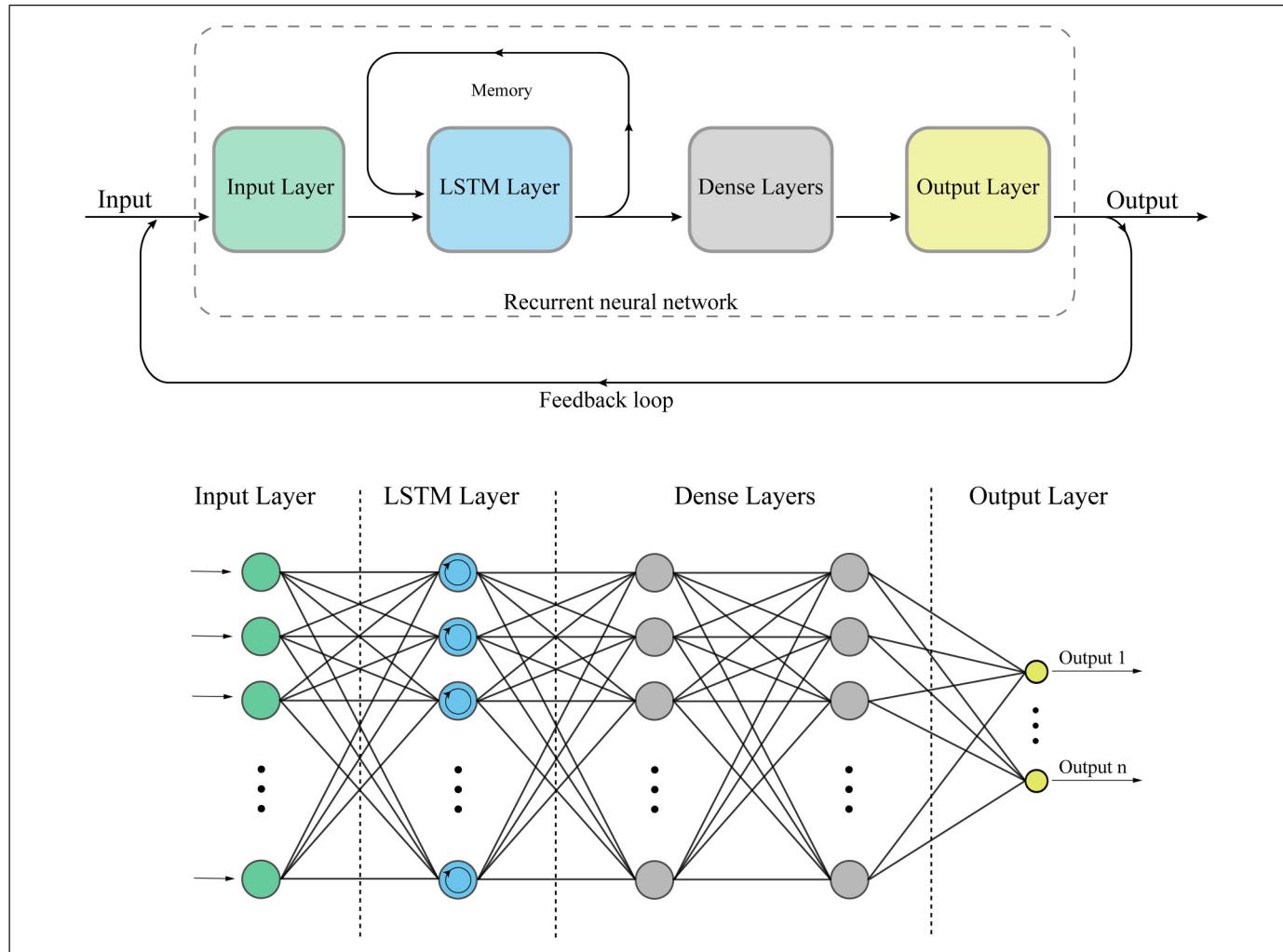


Figure 19. Structure of a recurrent neural network (RNN) and long short-term memory (LSTM) cell.

potentially speeding up the creation of new cancer treatments.

5. Addressing ethical and regulatory challenges associated with AI in cancer care, such as data privacy and algorithmic fairness, to ensure AI's responsible use.

In conclusion, the future of AI in cancer research and treatment holds tremendous promise for enhancing patient outcomes and healthcare efficiency. Ongoing innovation and collaboration among researchers, clinicians, and industry partners will be vital for unlocking the full potential of AI in revolutionizing cancer care.

Author Contributions

AZS, GAG, and MT conceived the idea for the paper manuscript and devised the whole scenario. AZS and AG created the figures. AZS, MT, and AG compiled the tables. AZS and GAG wrote the entire manuscript. AZS, MT, and GAG reviewed and edited the manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a McCleary Murchland Fellowship [to G.A.G.]; grants from the NHMRC [Ideas grant 2021/GNT2013180] to G.A.G.; the Neurosurgical Research Foundation [to G.A.G. and A.Z.S.]; and the Charlie Teo Foundation [Rebel Grant to G.A.G.].

ORCID iD

Guillermo A. Gomez  <https://orcid.org/0000-0002-0494-2404>

Supplemental Material

Supplemental material for this article is available online.

References

1. Available from: <https://pubmed.ncbi.nlm.nih.gov/>.

2. Dang A, Dang D, Vallish BN. Extent of use of artificial intelligence & machine learning protocols in cancer diagnosis: A scoping review. *Indian J Med Res* 2023;157(1):11-21.
3. Jan Z, El Assadi F, Abd-Alrazaq A, Jithesh PV. Artificial intelligence for the prediction and early diagnosis of pancreatic cancer: Scoping review. *J Med Internet Res* 2023;25:e44248.
4. Liu M, Wu J, Wang N, et al. The value of artificial intelligence in the diagnosis of lung cancer: A systematic review and meta-analysis. *PLOS One*. 2023;18(3):e0273445.
5. Chow R, Midroni J, Kaur J, et al. Use of artificial intelligence for cancer clinical trial enrollment: A systematic review and meta-analysis. *J Natl Cancer Inst*. 2023;115(4):365-374.
6. Gupta R, Kumar N, Bansal S, Singh S, Sood N, Gupta S. Artificial intelligence-driven digital cytology-based cervical cancer screening: Is the time ripe to adopt this disruptive technology in resource-constrained settings? A literature review. *J Digit Imaging* 2023;36(4):1-10.
7. Viñals R, Jonnalagedda M, Petignat P, et al. Artificial intelligence-based cervical cancer screening on images taken during visual inspection with acetic acid: A systematic review. *Diagnostics*. 2023;13(5):836.
8. Katta MR, Kalluru PKR, Bavishi DA, Hameed M, Valisseka SS. Artificial intelligence in pancreatic cancer: Diagnosis, limitations, and the future prospects—a narrative review. *J Cancer Res Clin Oncol* 2023;149:6743-6751.
9. Alongi P, Rovera G, Stracuzzi F, et al. Artificial intelligence in breast cancer: A systematic review on PET imaging clinical applications. *Curr Med Imaging*. 2023; 19(8):832-843.
10. Zadeh Shirazi A, Seyyed Mahdavi Chabok SJ, Mohammadi Z. A novel and reliable computational intelligence system for breast cancer detection. *Med Biol Eng Comput* 2018;56(5):721-732.
11. Mehta A, Kumar H, Yazji K, et al. Effectiveness of artificial intelligence-assisted colonoscopy in early diagnosis of colorectal cancer: A systematic review. *Int J Surg* 2023;109(4):946-952.
12. Parpinel G, Laudani ME, Piovano E, et al. The use of artificial intelligence for complete cytoreduction prediction in epithelial ovarian cancer: A narrative review. *Cancer Control*. 2023;30(4):1-8.
13. Bensoussan Y, Vanstrum EB, Johns III MM, Rameau A. Artificial intelligence and laryngeal cancer: From screening to prognosis: A state of the art review. *Otolaryngol–Head Neck Surg*. 2023;168(3):319-329.
14. Zadeh Shirazi A, Fornaciari E, Bagherian NS, et al. Deepsurvnet: Deep survival convolutional network for brain cancer survival rate classification based on histopathological images. *Med Biol Eng Comput* 2020;58(5):1031-1045.
15. Zadeh Shirazi A, McDonnell MD, Fornaciari E, et al. A deep convolutional neural network for segmentation of whole-slide pathology images identifies novel tumour cell-perivascular niche interactions that are associated with poor survival in glioblastoma. *Br J Cancer* 2021;125(3):337-350.
16. Thong LT, Chou HS, Chew HSJ, Laub Y. Diagnostic test accuracy of artificial intelligence-based imaging for lung cancer screening: A systematic review and meta-analysis. *Lung Cancer*. 2022;176:4-13.
17. de Margerie-Mellon C, Chassagnon G. Artificial intelligence: A critical review of applications for lung nodule and lung cancer. *Diagn Interv Imaging* 2022;104(1):11-17.
18. Govindan B, Sabri MA, Hai A, et al. A review of advanced multifunctional magnetic nanostructures for cancer diagnosis and therapy integrated into an artificial intelligence approach. *Pharmaceutics*. 2023;15(3):868.
19. Zadeh Shirazi A, Fornaciari E, McDonnell MD, et al. The application of deep convolutional neural networks to brain cancer images: A survey. *J Pers Med* 2020;10(4):224.
20. Bera K, Braman N, Gupta A, et al. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol* 2022;19(2):132-146.
21. Xu J, Yang P, Xue S, et al. Translating cancer genomics into precision medicine with artificial intelligence: Applications, challenges and future perspectives. *Hum Genet* 2019;138(2):109-124.
22. Alharbi F, Vakanski A. Machine learning methods for cancer classification using gene expression data: A review. *Bioengineering*. 2023;10(2):173.
23. Swanson K, Wu E, Zhang A, et al. From patterns to patients: Advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. *Cell*. 2023;186(8):1772-1791.
24. Bhatt M, Shende P. Advancement in machine learning: A strategic lookout from cancer identification to treatment. *Arch Comput Methods Eng* 2023;30:1-16.
25. Osong B, Bermejo I, Lee KC, et al. Prediction of radiotherapy compliance in elderly cancer patients using an internally validated decision tree. *Cancers* 2022;14(24):6116.
26. Delebarre M, Gonzales F, Behal H, et al. Decision-tree derivation and external validation of a new clinical decision rule (DISCERN-FN) to predict the risk of severe infection during febrile neutropenia in children treated for cancer. *Lancet Child Adolesc Health*. 2022;6(4):260-268.
27. Kurisu K, Inada S, Maeda I, et al. A decision tree prediction model for a short-term outcome of delirium in patients with advanced cancer receiving pharmacological interventions: A secondary analysis of a multicenter and prospective observational study (Phase-R). *Palliat Support Care* 2022;20(2):153-158.
28. Luo Z, Luo B, Wang P, et al. Predictive model of functional exercise compliance of patients with breast cancer based on decision tree. *Int J Womens Health* 2023;21(15):397-410.
29. Im E-O, Yi J-S, Chee W. The characteristics of Asian American breast cancer survivors with low quality of life: A decision tree analysis. *J Cancer Educ* 2023;38(4):1-9.
30. Tian J-X, Zhang J. Breast cancer diagnosis using feature extraction and boosted C5. 0 decision tree algorithm with penalty factor. *Math Biosci Eng*. 2022;19(3):2193-2205.
31. Im E-O, Yi J-S, Chee W. A decision tree analysis on the impact of a technology-based program on symptom distress: Asian American breast cancer survivors. CIN: Computers, informatics, Nursing. 2022;40(7):487-496.
32. Chee W, Yi J-S, Im E-O. Information needs of Asian American breast cancer survivors: A decision tree analysis. *J Cancer Educ* 2022;37(6):1902-1911.

33. Zhao L, Han W, Niu P, et al. Using nomogram, decision tree, and deep learning models to predict lymph node metastasis in patients with early gastric cancer: A multi-cohort study. *Am J Cancer Res* 2023;13(1):204.
34. Chen Q, Zhang J, Bao B, et al. Large-scale gastric cancer susceptibility gene identification based on gradient boosting decision tree. *Front Mol Biosci* 2022;8:1293.
35. Heinzel A, Müller D, van Santen HM, et al. The effect of surveillance for differentiated thyroid carcinoma in childhood cancer survivors on survival rates: a decision-tree-based analysis. *Endocr Connect* 2022;11(12):e220092.
36. Risør BW, Dehbarez NT, Fredsøe J, Sørensen KD, Pedersen BG. Cost-effectiveness analysis of Stockholm 3 testing compared to PSA as the primary blood test in the prostate cancer diagnostic pathway: A decision tree approach. *Appl Health Econ Health Policy* 2022;20(6):867-880.
37. Jin SW, Song CS. Predicting adoption of colorectal cancer screening among Korean Americans using a decision tree model. *Ethn Health* 2022;28(3):358-372.
38. Balaji P, Muniasamy V, Bilfaqih SM, et al. Chimp optimization algorithm influenced type-2 intuitionistic fuzzy C-means clustering-based breast cancer detection system. *Cancers* 2023;15(4):1131.
39. Wisaeng K. Breast cancer detection in mammogram images using K-means++ clustering based on cuckoo search optimization. *Diagnostics*. 2022;12(12):3088.
40. Nawaz M, Mehmood Z, Nazir T, et al. Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. *Microsc Res Tech* 2022;85(1):339-351.
41. Song C, Li X. Cost-sensitive KNN algorithm for cancer prediction based on entropy analysis. *Entropy* 2022;24(2):253.
42. Song X, Liang K, Li J. WGRLR: A weighted group regularized logistic regression for cancer diagnosis and gene selection. *IEEE/ACM Trans Comput Biol Bioinf* 2022;20(2):1563-1573.
43. Lee HD, Nam KH, Shin CM, et al. Development and validation of models to predict lymph node metastasis in early gastric cancer using logistic regression and gradient boosting machine methods. *Cancer Res Treat* 2023;55(4):1240-1249.
44. Ye Z, Zhang H, Liang J, et al. Significance of logistic regression scoring model based on natural killer cell-mediated cytotoxic pathway in the diagnosis of colon cancer. *Front Immunol* 2023;14.
45. Li Q, Luo Y, Liu D, Li B, Liu Y, Wang T Construction and prognostic value of enhanced CT image omics model for noninvasive prediction of HRG in bladder cancer based on logistic regression and support vector machine algorithm. *Front Oncol* 2022;12:7589.
46. Xiong Y, Ye M, Wu C. Cancer classification with a cost-sensitive naive Bayes stacking ensemble. *Comput Math Methods Med* 2021;2021:1-12.
47. Wood A, Shpilrain V, Najarian K, Kahrobaei D. Private naive Bayes classification of personal biomedical data: Application in cancer data analysis. *Comput Biol Med* 2019;105:144-150.
48. Fukai S, Daisaki H, Ishiyama M, et al., Reproducibility of the principal component analysis (PCA)-based data-driven respiratory gating on texture features in non-small cell lung cancer patients with 18F-FDG PET/CT. *J Appl Clin Med Phys*, 2023; 24(5):e13967.
49. Dweekat OY, Lam SS. Cervical cancer diagnosis using an integrated system of principal component analysis, genetic algorithm, and multilayer perceptron. *Healthcare*. 2022;10(10):2002.
50. Park HA, Edelmann D, Canzian F, et al. Predictive polygenic score for outcome after first-line oxaliplatin-based chemotherapy in colorectal cancer patients using supervised principal component analysis. *Cancer Epidemiol Biomarkers Prev* 2022;31(11): 2087-2091.
51. Li J, Shi J, Chen J, Du Z, Huang L Self-attention random forest for breast cancer image classification. *Front Oncol* 2023;13.
52. Gu T, Han Y, Duan R. A transfer learning approach based on random forest with application to breast cancer prediction in underrepresented populations. In PACIFIC SYMPOSIUM ON BIOCOMPUTING 2023, Kohala Coast, Hawaii, USA, 3–7 January 2023. 2022. World Scientific.
53. Zhang H, Chi M, Su D, et al. A random forest-based metabolic risk model to assess the prognosis and metabolism-related drug targets in ovarian cancer. *Comput Biol Med* 2023;153: 106432.
54. Zhou C, Xu C, Yang B, et al. Preoperative US integrated random forest model for predicting Delphian lymph node metastasis in patients with papillary thyroid cancer. *Curr Med Imaging*. 2023;19(9):1031-1040.
55. Zhang J, Yang X, Chen J, et al. Construction of a diagnostic classifier for cervical intraepithelial neoplasia and cervical cancer based on XGBoost feature selection and random forest model. *J Obstet Gynaecol Res.* 2023;49(1):296-303.
56. Li B, Zhang F, Niu Q, et al. A molecular classification of gastric cancer associated with distinct clinical outcomes and validated by an XGBoost-based prediction model. *Mol Ther.-Nucleic Acids*. 2022;31:224-240.
57. Nasser M, Yusof UK. Deep learning based methods for breast cancer diagnosis: A systematic review and future direction. *Diagnostics*. 2023;13(1):161.
58. Maurya S, Tiwari S, Mothukuri MC, et al. A review on recent developments in cancer detection using machine learning and deep learning models. *Biomed Signal Process Control* 2023;80(2).
59. Ding K, Zhou M, Wang H, et al. A large-scale synthetic pathological dataset for deep learning-enabled segmentation of breast cancer. *Sci Data* 2023;10(1):231.
60. Kalbhor M, Shinde S, Popescu DE, Hemanth DJ. Hybridization of deep learning Pre-trained models with machine learning classifiers and fuzzy min–max neural network for cervical cancer diagnosis. *Diagnostics*. 2023;13(7):1363.
61. Shen C, Rawal S, Brown R, et al. Automatic detection of circulating tumor cells and cancer associated fibroblasts using deep learning. *Sci Rep* 2023;13(1):5708.
62. Guleria HV, Luqmani AM, Kothari HD, et al. Enhancing the breast histopathology image analysis for cancer detection using variational autoencoder. *Int J Environ Res Public Health* 2023;20(5):4244.
63. Aziz RM. Cuckoo search-based optimization for cancer classification: A new hybrid approach. *J Comput Biol* 2022;29(6):565-584.
64. Gumaei A, Sammouda R, Al-Rakhami M, et al. Feature selection with ensemble learning for prostate cancer diagnosis from microarray gene expression. *Health Informatics J* 2021;27(1).

65. Xu Z, Sheykhhahmad FR, Ghadimi N, Razmjooy N. Computer-aided diagnosis of skin cancer based on soft computing techniques. *Open Med* 2020;15(1):860-871.
66. Prabhakar SK, Rajaguru H, Kim S-H. An amalgamated approach to bilevel feature selection techniques utilizing soft computing methods for classifying colon cancer. *BioMed Res Int* 2020; 2020:1-13.
67. Mahmoodi SA, Mirzaie K, Mahmoodi MS, Mahmoudi SM. A medical decision support system to assess risk factors for gastric cancer based on fuzzy cognitive map. *Comput Math Methods Med* 2020;2020:1-13.
68. Khan F, Khan MA, Abbas S, et al. Cloud-based breast cancer prediction empowered with soft computing approaches. *J Healthc Eng* 2020;2020:1-16.