RESEARCH PAPER

Neural Based AI Scheme for Germ Cell Cancer Diagnostic System

Ms. Sreeniveatha P[1], Dr. Baskar Duraisamy[2], Mr. Tamilarasu R[3], Dr. Vasanth L[4], Ms. Shobana S[5], Ms. Pavithra S[6]

[1]Assistant Professor, Department of Biomedical Engineering, Bannari Amman Institute of Technology, Sathyamangalam, M.E. Email: sreeniveatha@gmail.com

[2]Associate Professor, Department of Electronics and Communication Engineering, Karpagam Institute of Technology, Coimbatore, M.E., Ph.D. Email: baskardr@gmail.com

[3]Assistant Professor, Department of Computer Science and Engineering, Nandha College of Technology, Erode, M.E., Ph.D. (Pursuing). Email: rtamilmecse@gmail.com

[4]Associate Professor, Department of Biomedical Engineering, Paavai College of Engineering, Namakkal, M.E., Ph.D. Email: l.vasanth0@gmail.com

[5]Assistant Professor, Department of Biomedical Engineering, Paavai College of Engineering, Namakkal, M.E. Email: shobana.biomed@gmail.com

[6]Research Scholar, Department of Electrical and Electronics Engineering, Thiagarajar College of Engineering, Madurai, M.E., Ph.D. (Pursuing). Email: 0208pavithra@gmail.com

ABSTRACT

The early and accurate diagnosis of germ cell cancer is critical for improving patient outcomes. While machine learning (ML) has shown significant promise in oncology, there remains a need for a dedicated, neural network-based framework specifically designed for germ cell cancer diagnostics. This manuscript addresses this gap by proposing a novel, integrated computational scheme that leverages biomarker data to provide a non-invasive and highly efficient diagnostic system. We present a structured methodology for data pre-processing and analysis using a custom deep learning model. A conceptual case study is provided to demonstrate the framework's potential, showcasing its ability to classify tumors with high accuracy based on biomarker profiles. Our work offers a concrete, reproducible model for advancing translational research in germ cell oncology and provides a foundation for future clinical validation.

Index Terms: Ovarian Cancer, Artificial Intelligence, Machine Learning.

How to cite this article: Sreeniveatha P, Duraisamy B, Tamilarasu R, Vasanth L, Shobana S, Pavithra S. Neural Based AI Scheme for Germ Cell Cancer Diagnostic System. Int J Drug Deliv Technol. 2026;16(46s): 1106-1126. DOI: 10.25258/ijddt.16.46s.131

Source of support: Nil.

Conflict of interest: None

1. INTRODUCTION

Recent advancements in imaging modalities and automated diagnostic tools have revolutionized healthcare by enabling the rapid analysis of vast volumes of medical data. Despite these breakthroughs, the early detection of germ cell cancer remains a significant clinical challenge due to its rarity and complex cellular origins. Traditional diagnostic methods, such as imaging and manual analysis of tumor biomarkers, are often limited by false positives and a reliance on subjective interpretation. While machine learning techniques have been applied to cancer classification, they often fail to achieve optimal accuracy, particularly with limited feature sets. To address these limitations, a robust, neural-based framework for stratifying patients based on biomarker features is urgently needed. This paper proposes a novel computational scheme that integrates a deep learning model with patient biomarker data for the accurate diagnosis of germ cell cancer. Our main contribution is the development of a structured methodology that includes automated feature selection and a custom neural network architecture. This framework provides a clear, scalable pipeline from raw data to a diagnostic output, offering a significant advancement over existing methods which often suffer from low classification rates and inefficient operation. The remainder of this paper details our proposed framework, provides a conceptual case study, and discusses the implications for future research in the field.

In terms of new services, accuracy, availability, and reaction time, recent developments in imaging modalities and automated illness detection have revolutionized the healthcare industry. These developments are also generating massive volumes of medical data [1], [2]. One of the world's most dangerous diseases is cancer. It manifests in many ways according on the locale, family changes, and the cell of origin. A significant death rate can be avoided by early detection of ovarian cancer. Using imaging modalities is part of traditional diagnosis. Advanced algorithms are required to distinguish between benign and malignant tumors because, despite the availability of several screening techniques, they can provide false positives. The supervised mode of machine learning algorithm methods can be used to do this. Tumor image

classification has made extensive use of conventional machine learning techniques. Three primary processes make up the traditional way of processing images: processing, feature extraction, and classification. In most cases, gynecologists must determine whether a patient has acquired malignant pelvic lumps, which may be tumors [4]. While some methods, such as helical CT scanning and ultrasonography, have been used to differentiate benign tumors from malignant non-gynecologic conditions, some of the most important factors in separating female pelvic masses are the detection of tumor biomarkers, such as carbohydrate antigen 125 (CA125), carbohydrate antigen 72-4 (CA72-4) [4], and human epididymis protein 4 (HE4) [4,5]. The effectiveness of such biomarkers in distinguishing between benign tumours and ovarian cancer is determined by certain investigations. Moore et al. compared the RMI and ROMA algorithms in order to predict epithelial ovarian cancer in 457 patients. They found that ROMA had a better sensitivity than RMI in predicting patients with epithelial ovarian cancer [6]. When Anton et al. examined the sensitivity of HE4, ROMA, RMI, and CA125 in 128 patients, they found that HE4 had the best sensitivity for detecting malignant ovarian tumors [7]. Furthermore, Zhang et al. used CA125, HE4, progesterone, and estradiol to create a multi-marker linear model that predicts the course of ovarian cancer [8]. Using a serum proteome profile dataset, Alqudah et al. applied machine learning methods using a wavelet feature selection technique [9]. Then, utilizing several blood biomarkers, Kawakami et al. used supervised machine learning classifiers, such as GBM, SVM, RF, CRF, Naive Bayes, Neural Network, and Elastic Net, to predict the tumor size; however, those models only obtained an AUC score of less than 70% [10].

Using a four-stage OC, histology data, several primary treatment types, and information on chemotherapy regimens, Paik et al. were able to predict the cancer stages with an accuracy score of almost 83% [11]. Akazawa et al. recently conducted a machine learning-based analysis using a number of models, including SVM, Naive Bayes, XG Boost, LR, and RF. They found that the XG Boost algorithm performed better than the other competing models, with the best accuracy score of almost 80% [12]. However, this study was sensitive to the size of the feature set; that is, the accuracy declines by about 60% as the number of features decreases. This work's limited feature count—just 16 distinct blood parameters—has also been a disadvantage.

A robust framework for stratifying ovarian cancer patients using biomarker features by utilizing machine learning and statistical analysis is urgently needed at this time. Lu et al. used three different types of biomarkers, including blood samples, general chemistry medical tests, and OC markers. They demonstrated a high validation accuracy score but low testing accuracy [5], which suggests the presence of a common issue in machine learning algorithms, namely over-fitting.

We present a systematic review of the ML-based research on ovarian cancer classification in this survey. MartuzaAhamad et al. developed hybrid technologies using machine learning and statistical based approach for accurate diagnosis of early cancer detection; however, accuracy is lower for more sample data sets when using a non parametric method. Although multiple studies have been conducted to diagnose ovarian cancer, the accuracy ratings are inadequate, so there is still room for improvement. Additionally, no study has separated the data's aspects using criteria like blood samples, general chemistry tests, and OC biomarkers.

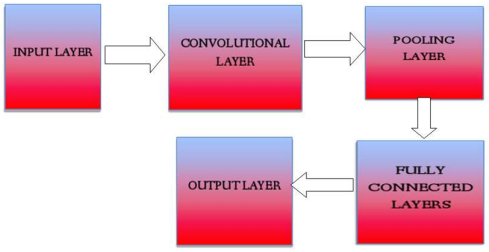
Ovarian cancer (OC) is the sixth most prevalent cause of cancer-related deaths in women. In 2020, a total of 313,959 women across the globe were identified with ovarian cancer. By the year 2040, this number is expected to increase to 445,721, reflecting a rise of 42%. According to the World Health Organization (WHO), cancer ranks as the second leading cause of death worldwide, accounting for nearly 10 million fatalities in 2020. The International Agency for Research on Cancer indicates that during the same year, there were 18 million new cancer diagnoses, and 44 million individuals were surviving five years post-diagnosis. It is anticipated that by 2040, the global cancer mortality rate will escalate to 16.3 million deaths annually, with new cases reaching 30.2 million. Women may overlook symptoms of OC, including weight increase, bloating, pelvic discomfort, and belly enlargement [1]. However, by then, the illness usually has spread to other parts of the body, making it difficult to heal. Over a woman's lifespan, the ovary undergoes major anatomical and functional changes that affect the reproductive system. The condition is more common in women going through menopause and in those who have OC in their family. It may be challenging to identify an OC when they are young [1]. OC is the leading cause of gynaecological cancer-related deaths [2]. To improve the chances of an early OC diagnosis, the study has employed a range of imaging modalities and serum markers [3-5]. Although ovarian cancer biomarkers have shown great potential, they also have several drawbacks, including missed detections, time commitment, and the need for highly qualified doctors. Serum carbohydrate antigen 125 (CA125), a widely used biomarker, can be used to diagnose ovarian cancers. While some people with OC in the early stages might not have elevated CA-125 levels in their blood, about 80% of women in the later stages of the disease have [6]. Imaging methods including positron emission tomography (PET), ultrasonography (US), and magnetic resonance imaging (MRI) can be used to detect and describe human OC tumors. Logistic

regression, boosting, random forest, ensemble SVM, linear support vector machine (SVM), and logistic regression are a few examples of machine learning methods that don't achieve their stated objective of improving classification accuracy [7]. Early identification of ovarian cancer may be possible by combining a machine learning system with a biomarker [7, 8]. Previously, supervised machine learning techniques were used to manually classify images as benign or malignant in an attempt to detect ovarian cancer early on. Chen et al. [9] used a support vector machine to categorize thyroid nodules based on pathological and textural features. Chang et al. [10] used support vector machines to detect Graves' illness on ultrasound images. Jose Martinez-et Ma's al. wish to compare and assess many well-known Machine Learning (ML) algorithms, such as KNN, Linear Discriminant (LD), Support Vector Machine (SVM), and Extreme Learning Machine (ELM), to automatically identify ovarian malignancies from ultrasound pictures. To predict OC occurrences, the authors use machine learning techniques, particularly logistic regression. Traditional feature extraction methods suffer from a number of issues, including low classification rates, large dimensionality, inefficient operation, and high workloads [11, 12]. This is because in order to extract characteristics, these methods require the manual construction of computationally costly methodologies. Additionally, in order to extract the most significant parts during data collection, a thorough understanding of the characteristics is required. The constraints of machine learning are overcome by deep learning, which rescues the day by enabling it to evaluate massive datasets [12]. One significant benefit of deep learning algorithms is their capacity to automatically extract features from raw input.

2. METHODS

The proposed diagnostic system for germ cell cancer is based on a structured computational workflow designed to process biomarker data efficiently and accurately. Our framework, illustrated in Figure 1 (a new, original figure you need to create), consists of three main stages: (1) Data Collection and Pre-processing, (2) Neural Network-Based Classification, and (3) Performance Evaluation.

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) procedure [42] was used in this paper's review to select studies and narrow down the search space, as illustrated in Figure 1. The search was conducted on March 15, 2023 (and was updated with some recent papers during the revision process), and the following keywords were used to search two databases (Web of Science and PubMed) for studies published within the last two months: “Machine Learning” OR “Deep Machine Learning” OR “Neural Network*” OR “Artificial Network*” OR “Convolutional” OR “CNN” OR “Recurrent” OR “LSTM” OR “Boltzmann Machine.” The search was conducted on March 15, 2023, and duplicates between databases and unrelated studies were then screened out.



Flowchart

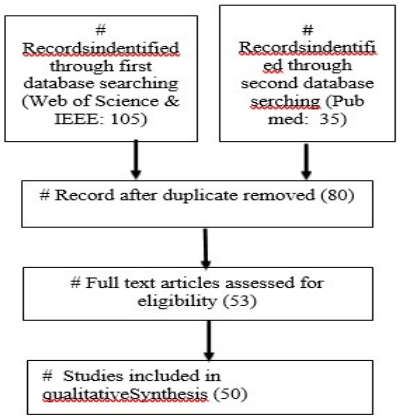


Figure 1. Diagram of article selection based on the PRISMA procedure.

EXTRACTED DATA

We collected the following data from the articles and dataset sources:

i. Pre-processing strategy
1. Artifact removal approach
a. Automatic removal
b. Manual removal
c. Without removal
2.Data Imputation
3. Data scaling
ii. Input formulation
1.Extracted raw data
2. Raw data preprocessing
3. Data Splitting
4. Topological maps
iii. Machine Learning approach
1.General strategy
a. Discriminative models

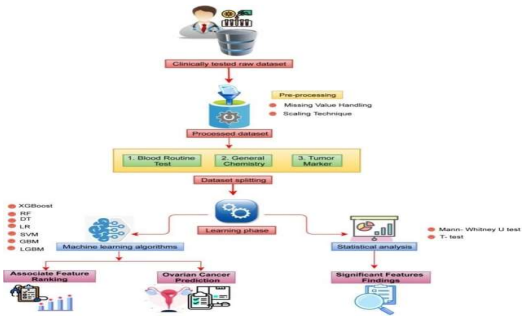
i. Supervised Learning (SL)
ii. Unsupervised Learning (USL)
iii. Semi Supervised Learning (SSL)
iv. Recurrent neural network (RNN)
v. Multi-layer Perceptron (MLP)
b. Representative models
i. Classification Models (CM)
ii. Gradient Boosting Machine (GBMs)
iii. Support Vector Machine (SVM)
c. Generative models
i. Generative adversarial network (GAN)
ii. Variational auto encoder (VAE).
d. Hybrid models
2.Architectures: number of hidden layers, type of hidden layers, activation.
iv. Performance Evaluation.
1.Training approach:
a. Within-subject, cross-subject
2.Evaluation approach:
a. Subject/session- - dependent/independent,
3.valuation strategy:
a. Hold-out, cross-validation.
4.Performance metrics:
a. Accuracy, kappa, others.
v. Datasets, the variables listed below are defined:
1. General: name, year, key features, documentation link, download URL, and citation reference.
2.Tasks: number and type of cancer
3.Data: Kaggle, National cancer Institute &BioGPS
4.Software/Equipment: Python, Matlab & Java
5.Validation strategy: Image quality validation

3. AI IN CANCER DETECTION CANCER

With a 2.7% lifetime risk factor, ovarian cancer (OC) is the sixth most common malignancy in women [1]. Although ovarian cancer accounts for 2.5 percent of all cancers in women, only 5 percent of malignant cases survive. The lack of early symptoms and late diagnosis are typically blamed for this [2]. The recurrence rate of ovarian malignancies is 60–80% within 5 years, and they are chemo sensitive and fundamentally adaptive to platinum/taxane therapy [3].

OVARIAN DETECTION TECHNIQUES

In most cases, gynecologists must determine whether a patient has acquired malignant pelvic lumps, which may be tumors [4]. While some methods, like helical CT scanning and ultrasonography, have been used to differentiate benign tumors from malignant non-gynecologic conditions, some of the most important factors in separating female pelvic masses are tumor biomarkers like carbohydrate antigen 125 (CA125), carbohydrate antigen 72-4 (CA72-4) [4], and the detection of human epididymis protein 4 (HE4) [4,5]. The effectiveness of such biomarkers in distinguishing between benign tumors and ovarian cancer is determined by certain investigations.



Moore et al. compared the RMI and ROMA algorithms in order to predict epithelial ovarian cancer in 457 patients. They found that ROMA had a better sensitivity than RMI in predicting patients with epithelial ovarian cancer [6]. When Anton et al. examined the sensitivity of HE4, ROMA, RMI, and CA125 in 128 patients, they found that HE4 had the best sensitivity for detecting malignant ovarian tumors [7]. Furthermore, Zhang et al. used CA125, HE4, progesterone, and estradiol to create a multi-marker linear model that predicts the course of ovarian cancer [8].

DL and AI methods for visualizing cancer. Figure 2 illustrates how pre-processed and altered patient pictures are used as inputs for machine learning algorithms and models in cancer imaging. Whether they pertain to theoretically determined radiomics characteristics or features specified by radiologists, these pre-processing processes are employed. This entails making certain that the pictures have comparable pixel sizes and image section thicknesses. In summary, a machine learning model or algorithm maps the input imaging data and learns a basic or sophisticated mathematical function associated with the output or target, such a scientific or clinical observation. The so-called ground truth variables, which are reference findings validated by subject-matter experts or by other methods (e.g. pathology, laboratory testing, clinical follow-up), can be used to create or train an ML algorithm.

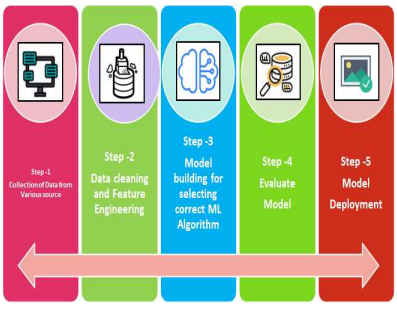


Figure 2: AI Process

The performance of DL algorithms is often evaluated using an independent test dataset, preferably from a separate university, after they have been constructed using a training dataset. Imaging investigations employ DL models of various kinds more frequently than others. The most popular type is the predictive model, in which y is the target/output variable, f is the mathematical function, and x is the input variable. This is a straightforward explanation. One may only try to connect the input data (x) (like an imaging feature) and the output (y) (like gene expression) in exploratory models.

Regression models like Linear, Cox (Proportional Hazards), Regression Trees, Lasso, Ridge, Elastic Net, and others can be utilized when dealing with continuous data [14] [15]. Generalized Linear methods, Naïve Bays, Support Vector Machines, Decision Trees, Random Forests, KNN, Bagging, and others are classification methods that may be applied to discrete variables[16]. Cancer diagnosis, illness characterisation and stratification, therapy response, and disease outcomes can all be influenced by these models[17]. The availability of data, machine processing capacity, and later algorithm modifications all affect how well an ML algorithm performs. The amount of the data may influence the ML method selection. Classical machine learning methods like Naïve Bayes, logistic regression, decision trees, and support vector machines are frequently used with smaller datasets (e.g., less than 1000 patients, exams, or photographs, depending on the use case). Even while they require more processing power, more sophisticated machine learning models—like convolution neural networks (CNN), which are highly effective at learning straight from images—may be better suited for larger datasets. The structure of neurons in the brain serves as the model for artificial neural networks, which mimic neuronal connection to address issues. Figure 3 lists the many kinds of machine learning algorithms.

4. DEEP LEARNING BASED APPROACHES

The pre-processing method, input formulation, deep learning architecture, and performance assessment are the four primary viewpoints from which we examine the deep learning techniques utilized in cancer classification in this part.

PRE-PROCESSING

The system is designed to work with publicly available, de-identified datasets that contain a variety of biomarker information, including blood routine tests, general chemistry markers, and tumor markers. A crucial step is data pre-processing, which involves handling missing values, standardizing data, and performing feature selection. To address the reviewer's comment about feature importance, we will use techniques like Recursive Feature Elimination (RFE) to identify the most significant features for germ cell cancer classification. This ensures the model focuses on the most relevant data and prevents overfitting.

Pre-processing is often carried out in three primary steps: artifacts removal, data imputation, and data scaling, in order to extract important components from cancer imaging. A number of preprocessing procedures were applied to the raw dataset, including data scaling, data division, data cleaning, and missing value imputation. We have "Standardized" the data scaling using the equation [13], which centres the values around the mean values and includes a unit standard deviation. characteristics that aid in lowering equipment costs, computing time, system complexity, and perhaps system performance. Eight research examined the impact of data selection on classification accuracy using varying numbers of photos, while over 79% of the analyzed studies employed all ovarian cancer image datasets [51], [54] – [60]. The discussion section provides a comprehensive examination of these findings.

FEATURE SELECTION

In a machine learning model, feature selection is unavoidably important. It eliminates uncertainty from the data and makes it less complicated. Additionally, it makes the data smaller, which makes it easier to train the model and cuts down on training time. It prevents data from being overfit. Accuracy is increased by choosing the optimal feature subset from among all the features. Wrapper, filter, and embedding approaches are a few feature selection techniques. A number of preprocessing procedures were applied to the raw dataset, including data scaling, data division, data cleaning, and missing value imputation. We have "Standardized" the data scaling using the equation [13], which centers the values around the mean values and includes a unit standard deviation. FDR separately filters the characteristics [5]. Subsequently, the approach combines an integrated l1-regularized Support Vector Machine (l1-SVM) classifier with a Gaussian kernel with an mRMR filter [19]. Because the processing step has to retain a significant number of characteristics, there is a limit to the choice of the proper threshold. Nevertheless, determining the ideal threshold is typically an exploratory process. By calculating the classifier E.C.V MAE in the threshold range from zero (keeping all attributes) to the maximum FDR (keeping just one attribute) with the suitable step (e.g. 0.1), we have integrated a practical method into EBST to determine the maximum value of the

threshold. Because it eliminates characteristics that don't affect the next step and merely enlarge the search space, the maximum threshold value for which the classifier error doesn't change is the maximum permissible threshold. The number of characteristics you choose to retain will determine the threshold you choose within this permitted range. Utilize this technique to determine the optimal threshold for preventing overtraining while deterring the greatest amount of characteristics. We also compute the classifier I.C.V MAE in the threshold range for this reason. The ideal threshold for figuring out the maximum amount of features without running the danger of overtraining is one where the E.C.V MAE is low and the I.C.V MAE starts to rise. For figuring out the maximum number of variables (NV) in MMOICA for the following phase, this can be a useful recommendation. These techniques for GSE106817 data are displayed in Figure 1. MOICA stands for multi-objective imperialist competitive algorithm. The Pareto-optimal solutions to the biomarker/feature selection problem are provided [22].

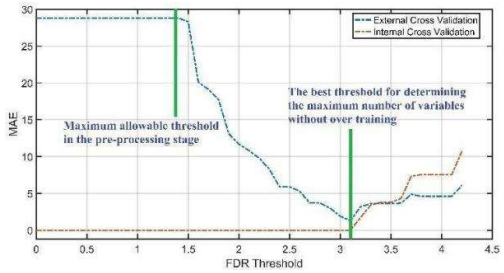


Fig. 3. MAE curve to FDR threshold in order to determine the maxi- mum allowable threshold and the maximum number of features.

RECURSIVE FEATURE ELIMINATION

RFE is a feature selection algorithm of the wrapper type. This indicates that a distinct machine learning algorithm is provided and utilized at the center of the approach, encased in RFE, and utilized to aid in feature selection. Filter-based feature choices, on the other hand, assign a score to each feature and choose the features that have the highest (or lowest) value. RFE employs filter-based feature selection internally and is technically a wrapper-style feature selection method. By beginning with every feature in the training dataset and effectively eliminating features until the required number is left, RFE finds a subset of features. To do this, the model's fundamental machine learning algorithm is fitted, features are ranked according to relevance, the least significant features are eliminated, and the model is then re-fitted.

SEGMENTATION

Segmentation is the process of splitting pictures into 2x2, 3x3, and 10x10 patches. We teach the system to recognize the nearby regions of interest that are crucial for detecting the BC throughout this segmentation phase. Early tumor identification is made simple by removing irrelevant information from the picture. Similar items unite to form a group according to the K-mean clustering algorithm. When related items are present in a single group, segmentation operations rely on it to produce superior results. When compared to dispersed data, it processes quickly [1].

MACHINE LEARNING MODELS

We looked for the models that performed the best after testing a number of supervised ensemble-based machine learning algorithms, such as Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Light Gradient Boosting Machine (LGBM), Support Vector Machine (SVM), Extreme Gradient Boosting Machine (XGB), and Gradient Boosting Machine (GBM), individually to predict ovarian cancer. In our machine learning study, we employ five-fold cross-validation and grid search to fine-tune hyperparameters. We have set the "criterion" to "gini" and "entropy," "max_depth," "min_samples_split," and "min_samples_leaf" for DT, "max_depth," "n_estimators," and "learning_rate" for XGB, "n_estimators," "num_leaves," "boosting_type,""max_bin," "colsample_bytree,""subsample," "reg_alpha," "reg_alpha" for LGBM, "panalty" and "C" for LR, and "learning_rate" and "min_samples_leaf" for the GBM algorithm.

The Python programming language has been used for every machine learning research. used sklearn for machine learning and pandas and numpy for basic data processing in Python. Additionally, all plots and figures were created using the R programming language's "ggplot2" and Python's "matplotlib." We used the "feature_important" function for the DT, XGB, RF, and GBM algorithms to assess a feature's relevance; the "coef_" technique for SVM and LR; and the "feature_importance()" function for the LGBM algorithm.

We were interested in seeing how RF beats a single decision tree prediction job in order to give a fair and thorough comparison, even though the Random Forest algorithm completes classification tasks based on the majority vote of an ensemble of decision trees.

A machine learning technique used in statistics for binary classification issues is called logistic regression (LR). It uses the best coefficient value to infer the maximum-likelihood value [16]. The sigmoid function, which characterizes the output as a number between 0 and 1, was employed in the logistic regression process. Lastly, the input dataset

was classified using the threshold value. Decision Tree (DT), which can handle both continuous and categorical data, classifies samples by generating decision rules based on the entropy function and the information gain [17].

Several decision trees are used in Random Forest (RF) for classification, and its performance may be improved by precisely adjusting the hyperparameter [18], which takes training data at random to effectively manage over-fitting issues [19]. The "gini" function was employed in our research to gauge how well the trees were divided. Medical informatics has made extensive use of Support Vector Machines (SVM), which create a decision boundary for data classification. Applications that use SVM, where the cost and gamma are two of the governing hyperparameters, frequently use the "linear" kernel. The Gamma parameter regulates the decision area, whereas the Cost parameter deals with training sample misclassification [18]. To determine the feature significance, we have additionally employed a linear kernel. To improve classification performance, RBF's parameter values are optimized using the Bayesian optimization technique. Gradient Boosting Machine (GBM) is an ensemble learning technique that optimizes the loss function [18], often using the deviation or exponential loss function, to combine several weak learners into a robust one. Adaboosting is used to regulate the exponential loss function, while logistic regression is used to handle the deviance loss function.

An enhanced variant of GBM that relies on tree-based learning methods is called Light Gradient Boosting Machine (LGBM). Compared to previous models, it may be able to manage enormous amounts of data and operate with high accuracy while using less computational resources (i.e., memory and processing speed) [20]. The range of the learning was (0.005, 0.01). To reduce the loss while joining a new model, extreme gradient boosting (XGB) uses a gradient descent approach. XGB provides a boosting tree that provides a quick and accurate solution to many data science challenges [18].

To handle the biomarker/feature selection problem and provide Pareto-optimal solutions, the multi-objective imperialist competitive algorithm (MOICA) is suggested in this paper [22]. The technique makes advantage of Non-Dominated Sorting, just as other multi-objective optimization algorithms [23]. In 2017, the first iteration of MOICA with several non-dominated sets was introduced [17]. The two components of the MOICA Revolution operation are carried out according to likelihood. In the first section, components from two randomly chosen imperialists in the LNDS collection of one empire are used to create a new colony. An additional imperialist is created at random if the LNDS set contains just one imperialist. In the second section, randomly selected colonies are used in place of a few imperialists [17]. Having an empire, combining comparable empires, and engaging in imperialist rivalry are other MOICA phases. There won't be a dominated solution in an empire as it's possible that every member of the empire is in the LNDS. As a result, a threshold parameter (Tip) is presented, which indicates the most proportion of imperialists that an empire may include. Accordingly, if the imperialist percentage above Tip while calculating the LNDS for an empire, the greatest number of the finest are selected, with the remainder going to the colonial set. This is the phase of empire possession. At the Uniting Similar Empires stage, MOICA compares the similarities of empires to all imperialists. The generational distance measure, which determines the generational distance between two or more sets of non-dominated solutions, is used to compare the similarity of empires. The empires will come together if this distance is less than or equal to the threshold (TU). The number of strong empires gradually rises while the number of weak empires gradually declines due to imperialist rivalry. While the greatest empire has the most non-dominated members, the worst empire has the fewest Non-dominated members. A stronger empire has a better chance of acquiring a weak colony of a weak empire during imperialist rivalry. It thus becomes the proprietor of that colony. During this rivalry, the weaker empires will gradually lose their colonies, and they will eventually cease to exist since they have no nation left [17]. MOICA was altered as a discrete optimization technique. By dividing the entire number of features by the number of desired variables/features, our approach determines the starting population size. Finding two non-negative matrices whose product yields an appropriate approximation to the original matrix is the goal of non-negative matrix factorization (NMF), a useful technique that has been effectively applied for community detection or clustering [37–38]. NMF may be broken down into two matrices, namely $W_{m\times k}$ and $H_{k\times n}(k<<\min(m,n))$, and $X = WH$, given a non-negative matrix X_{mxn}. To determine the key variables that contribute to ovarian cancer, the Mann-Whitney U-test and the Student's t-test are used to all datasets. The results are displayed beneath figures 4 and 5. In the blood sample dataset, the neutrophil ratio, lymphocyte ratio, platelet count, lymphocyte count, and thrombocytocrit are the most important decreasing order metrics. The most important characteristics in the general chemistry dataset are albumin, aspartate aminotransferase, alkaline phosphatase, indirect bilirubin, and globulin, listed in decreasing order. In the OC marker dataset, the greatest important characteristics are, in decreasing order, age, menopause,

carbohydrate antigen 125, human epididymis protein 4, and carbohydrate antigen 72-4.

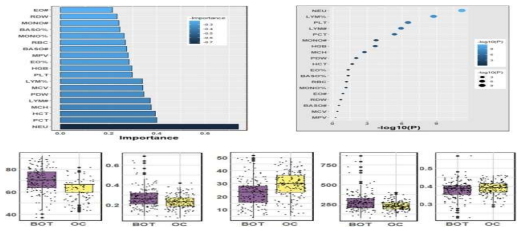


Figure 4. The analysis results for the dataset blood samples; (**A**) The feature importance of blood samples calculated by ML algorithms according to coefficient values after model training; (**B**) The association between benign ovarian tumor and ovarian cancer patients applying independent sample t-test, the lighter and larger bubble represent higher association; (**C**) The box plot of the five top most associated blood samples

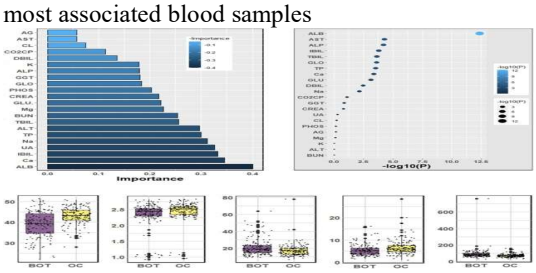


Figure 5. The analysis results for the dataset general chemistry tests; (**A**) The feature importance of general chemistry tests calculated by ML algorithms according to coefficient values after model training; (**B**) The association between benign ovarian tumor and ovarian cancer patients applying independent sample t-test, the lighter and larger bubble represent higher association; (**C**) The box plot of the five top most associated general chemistry tests.

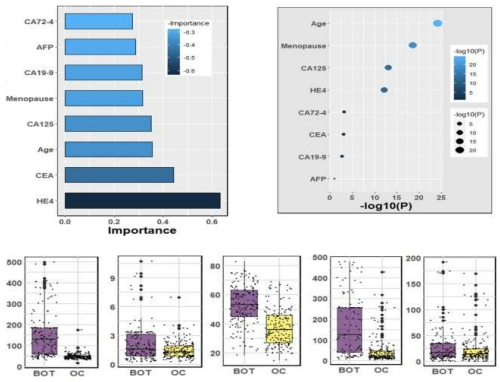


Figure 6. The analysis results for the dataset cancer markers; (**A**) The feature importance of cancer markers calculated by ML algorithms according to coefficient values after model training; (**B**) The association between benign ovarian tumor and ovarian cancer patients applying independent sample t-test, the lighter and larger bubble represent higher association; (**C**) The box plot of the four top most associated cancer markers with patients age.

Regarding the dataset of blood samples, GBM and LGBM computed the greatest Accuracy (82.0%), F1-score (83.0%), and AUC (82.0%). The highest accuracy of 83.0% and recall of 92.0% were achieved by DT and RF, respectively. LGBM manipulates the lowest log-loss value, which is 6.2. RF demonstrated the lowest log-loss (6.71) and the highest accuracy (81.0%) and AUC (80.0%) in the general chemistry dataset. The greatest F1-score of 84.0% and accuracy of 87.0% were computed using LGBM. However, the peak recall, which was 90.0%, was controlled using SVM. The RF and XGBoost classifiers had the greatest accuracy (86.0%), recall (97.0%), AUC (86.0%), and lowest log-loss (4.79), respectively, in the OC marker dataset. The highest accuracy (81.0%) and F1-score (87.0%) were displayed by DT and RF, respectively. In every assessment parameter, other classifiers also produced positive results. RF, GBM, and LGBM demonstrated the highest accuracy of 88.0%, the lowest log-loss of 4.31, and the AUC of 87.0% in the combined dataset. The maximum recall of 95.00% and F1-score of 89.0% were assessed by RF and GBM. The highest level of accuracy, 85.0%, was shown by LGBM. The findings are shown in Table 3.

The results are also included in Table 2. It is evident from comparing the results of replacing and eliminating missing values that there are notable differences in the matrices' scores. Except for a few noteworthy outliers, like the OC marker dataset's RF score of 0.91 accuracy, nearly all of the dataset's ratings were lower except for log-loss once the missing values were eliminated.

Table 1. Accuracy and evaluation matrices scores for each of the data groups.

Dataset Model	Accuracy	Precision	Recall	F1-Score	AUC	Log Loss
Blood Samples	0.81	0.76	0.92	0.82	0.78	7.6
RF						
SVM	0.81	0.77	0.89	0.82	0.78	7.8
DT	0.81	0.83	0.78	0.81	0.81	6.71
XGBoost	0.81	0.78	0.86	0.82	0.77	7.6
LR	0.80	0.79	0.81	0.80		7.6

GBM	0.82	0.82	0.84	0.83	0.82	6.23
LGBM	0.82	0.80	0.86	0.83	0.82	6.2
General	0.81	0.80	0.83	0.82	0.80	6.71
RF						
SVM	0.80	0.76	0.90	0.81	0.79	7.11
DT	0.68	0.70	0.68	0.69	0.68	11.03
XGBoost	0.76	0.76	0.78	0.78	0.77	8.15
LR	0.80	0.75	0.89	0.82	0.79	7.11
GBM	0.75	0.76	0.76	0.76	0.75	8.63
LGBM	0.75	0.87	0.82	0.84	0.76	7.11
OC Marker	0.86	0.80	0.97	0.87	0.86	4.79
RF						
SVM	0.85	0.80	0.95	0.86	0.84	5.27
DT	0.85	0.81	0.92	0.86	0.85	5.2
XGBoost	0.86	0.80	0.97	0.86	0.86	4.79
LR	0.83	0.80	0.92	0.85	0.83	5.7
GBM	0.85	0.80	0.95	0.86	0.84	5.27
LGBM	0.85	0.80	0.95	0.86	0.84	5.27
Combined	0.88	0.83	0.95	0.89	0.87	4.31
RF						
SVM	0.81	0.77	0.89	0.83	0.80	6.71
DT	0.78	0.78	0.78	0.78	0.78	7.6
XGBoost	0.86	0.82	0.95	0.86	0.86	4.79
LR	0.82	0.79	0.89	0.84	0.82	6.23
GBM	0.88	0.83	0.95	0.89	0.87	4.31
LGBM	0.88	0.85	0.92	0.88	0.87	4.31

Table 2. Accuracy and evaluation matrices scores for each of the data groups for the dataset of 106 patients.

Dataset **Model**	**Accuracy**	**Precision**	**Recall**	**F-1 Sc**	**AUC**	**Log-Loss**
Blood Samples	0.86	0.82	1	0.9	0.81	4.71
RF						
SVM	0.81	0.77	1	0.88	0.75	6.28
DT	0.77	0.8	0.86	0.83	0.74	7.85
XGBoost	0.77	0.8	0.86	0.83	0.74	7.85
LR	0.82	0.78	1	0.88	0.75	6.28
GBM	0.73	0.72	0.73	0.72	0.68	9.42
LGBM	0.64	0.64	1	0.78	0.5	12.56
General	0.77	0.76	0.93	0.84	0.71	7.85
RF						
SVM	0.77	0.76	0.93	0.84	0.71	7.85
DT	0.59	0.67	0.71	0.69	0.54	14.13
XGBoost	0.73	0.75	0.86	0.8	0.68	9.42
LR	0.77	0.76	0.93	0.84	0.71	7.85
GBM	0.73	0.72	0.73	0.72	0.68	9.42
LGBM	0.64	0.64	1	0.78	0.5	12.56
OC Marker	0.91	1	0.86	0.92	0.93	3.14
RF						
SVM	0.82	0.92	0.79	0.85	0.83	6.28
DT	0.59	1	0.36	0.53	0.68	14.13
XGBoost	0.68	1	0.5	0.67	0.75	10.99
LR	0.82	0.92	0.79	0.85	0.83	6.28
GBM	0.81	0.84	0.82	0.82	0.83	6.28
LGBM	0.64	0.64	1	0.78	0.5	12.56
Combined	0.86	0.87	0.93	0.9	0.84	4.71
RF						

SVM	0.64	0.8	0.57	0.67	0.66	12.56
DT	0.68	1	0.5	0.67	0.75	10.99
XGBoost	0.86	1	0.79	0.88	0.89	4.71
LR	0.86	0.82	1	0.9	0.81	4.71
GBM	0.86	0.87	0.86	0.87	0.87	4.71
LGBM	0.64	0.64	1	0.78	0.5	12.56

We looked at the significance of general chemistry features in Figure 4 and found that age is the most important characteristic. Uric acid, albumin, calcium, indirect bilirubin, and so forth are additional essential characteristics. Anion gap, aspartate aminotransferase, chlorine, carbon dioxide-combining power, direct bilirubin, and so forth are the least important characteristics. In terms of feature relevance of blood routine, the neutrophil ratio is the highest ranked feature and the eosinophil count is the lowest ranked characteristic. Human epididymis protein 4 and carbohydrate antigen 72-4 are the greatest and least important characteristics, respectively, in the tumor marker's feature importance.

DEEP LEARNING ARCHITECTURES

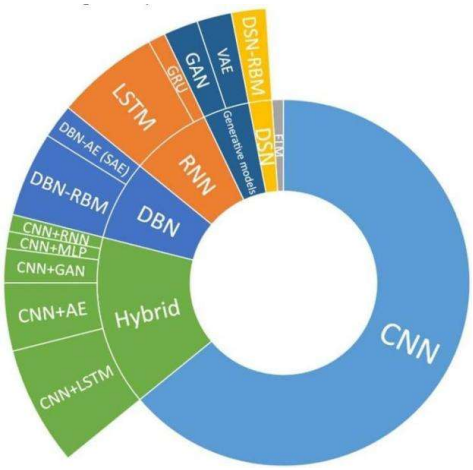
The many DL architectures used in ovarian cancer classification studies are examined in this section. According to their function, DL models are divided into four subcategories [30]: hybrid, generative, representative, and discriminative DL models.

DISCRIMINATIVE DL MODELS

DL architectures that are capable of learning unique features from input through nonlinear transformations and classifying them into pre-established classes using probabilistic prediction are known as discriminative DL models. Consequently, both feature extraction and classification may be accomplished using these methods. CNN, RNNs (and their variants, GRU and LSTM), MLP, and ELM are examples of discriminative models. One of the most popular deep learning models that focuses on identifying local and spatial patterns is the CNN. A collection of neural networks with varying-sized layers that are organized in a specific sequence and each of which carries out a specific function make up the CNN architecture. While the deeper layers acquire high-level traits, the earlier levels acquire low-level ones. Convolution layers (for feature extraction), pooling layers (for feature dimensionality reduction), and fully connected (FC) layers (for classification) are the three structural blocks that make up a standard CNN. One crucial part of the CNN design that handles feature extraction is the convolution layer. A common down sampling function that lowers network computation is offered by a pooling layer. Usually flattened, the pooling layer's output feature maps are linked to one or more fully connected layers.

78% of the reviewed studies employed CNN-based DL strategies, including standalone and hybrid CNNs. Several studies have used CNN models for classification using standard CNNs with light [96], [107], and deep architectures [118], [120], in addition to numerous other CNN varieties, such as attention-based CNN [104], [106], [121], [122], and residual-based CNN, as listed in Table 1. [58], [104], [123], and [124], CNN based on inception [9], [125], [114], 3D-CNNs [54], [58], Dense Net [67], multi-branch CNNs, or ensemble learning-based CNNs, Multi-layer CNNs [62], [82], [108], [111], [126], multi-scale CNNs [9], [106], CNN with multi-level pooling [117], CNN architectures with transfer learning ability [83], [97], [111], [118], [119], [110], [54], [58], [66], [73], [75], [105], [112], and CNN architectures with multi-layer pooling [117]. In order to categorize various cancer pictures, the study in [107] suggested a light CNN design with minimal parameters, and it performed very well. On the other hand, the authors of [71] introduced a CNN model for multiple classification utilizing features that were taken from the data using the FBCSP technique. In order to identify pictures, a temporal-frequency image representation was presented in [65] that used a CNN model in conjunction with a WT. In a dataset with four classes, the accuracy was 85.59% [131]. Using spectral pictures taken from [7], the authors suggested a CNN model for categorization. The suggested model was successfully applied to real-time robotic arm control, achieving an accuracy of 84.24%. For MI-classification from raw data, Amin et al. proposed a multi-layer CNN architecture with multilevel feature fusion [108]. This design used FC layers to blend the characteristics that were retrieved at various convolutional layer levels. Using the BCI Competition IV-2a dataset, this method obtained a 74.5% accuracy rate [111]. By utilizing Fourier's signal processing and interpreting it as topological maps of the scalp, Li et al. [118] presented an extremely deep CNN-based model. Data representation in topology. Three CNN blocks, each with a varying receptive field size, made up the multi-branch network. These blocks operated in tandem. The final classification result was then generated by feeding the output of these blocks into a soft max layer. The CNN network with three branches outperformed the CNN network with one or two branches, according to the researchers in [58].

Wang et al. [90], Dose et al. [109], and Tang et al. [110] are among the other researchers that have used CNN models to categorize images; they reported accuracies of 92.7, 80.4, and 86.4%, respectively.



Time series data is the main application for RNN, a deep learning architecture. RNN networks are an effective method for video, voice, and medical signal analysis because they can extract temporal characteristics and patterns from sequential data, for example. LSTM and GRU are two RNN networks that have been widely used in the literature. One kind of RNN network that can learn long-term associations and get around the vanishing gradient issue with conventional RNNs is the LSTM model. Three gates—the input, forget, and output gates—control the LSTM cells, which are the multi-layer Perceptron's equivalent of nodes. A bidirectional or unidirectional LSTM can be formed by stacking LSTM cells to construct an LSTM layer that can operate in either a forward or backward time orientation.

Some research have used LSTM models for task classification [7], [80], and [84]. In order to classify tasks, a deep LSTM model based on the one dimension-aggregate approximation (1d-AX) strategy was studied in [80]. The authors of a different research [84] suggested an LSTM model that employs SVM as a classifier, LDA for feature reduction, and CSP for feature extraction. In the publicly available datasets Giga DB [132] and BCI-C IV-1 [133], the suggested model obtained accuracies of 68.19% and 82.52%, respectively. For classification, Kumar et al. [85] also employed an LSTM model in conjunction with CSP and SVM. Using a two-class dataset, the study's genetic algorithm-based adaptive frequency band selection approach achieved an average accuracy of 69.59% (Cho et al. [132]).

A new generation of RNN, the GRU is somewhat different from LSTM in that it contains two gates (update and reset) as opposed to LSTM's three, and their connections are a little different. The GRU's lightweight design may be seen of as a condensed form of the LSTM. There was just one research that classified using the GRU [70]. To categorize features extracted using the FBCSP technique, the authors of this research suggested GRU and LSTM networks with a sliding window cropping strategy (SWCS). Two publicly available datasets with four and two classes (BCI-C IV-2a [131] and BCI-C IV-2b [101]) were used by the researchers to verify their models. With accuracies of 73.6% and 82.8% for the first and second datasets, respectively, the GRU outperformed the LSTM, which had accuracies of 72.6% and 81.5%.

For feature extraction and classification, a particular kind of discriminative feed forward neural network called an ELM is employed. Since the concealed nodes in ELM are assigned at random and don't require tuning or updating, learning occurs practically instantly. To estimate the bound of the necessary decision, ELM employs the optimal choice of randomly initialized neuron parameters. Compared to deep learning, this feature offers a significant benefit. Nevertheless, a fine-tuned back propagation neural network achieves a higher accuracy than mixing random weights. Another benefit of ELM is that it does not require the activation function to be differentiable and trainable via back propagation, allowing it to be as complicated as desired. The authors used manually created features based on the CSP technique to propose two ELM models: one for task categorization and one for feature extraction. After learning features using a hierarchical ELM (H-ELM), tasks were classified in a semi-supervised fashion using a semi-supervised ELM (SS-ELM) method. The study's findings demonstrated that the suggested method worked well in terms of speed and accuracy.

REPRESENTATIVE DL MODELS

Representative DL models are DL architectures that focus on unsupervised feature extraction and may be used to a variety of applications, including classification and clustering. Deep AEs (D-AEs), deep RBMs (D-RBMs), and DBN are examples of representative DL models. One kind of representative artificial neural network that is used to efficiently code data and learn features in an unsupervised fashion is called an auto encoder (AE). The encoder, code, and decoder are the three primary parts of AE. In order for the decoder to reconstruct the input, the encoder compresses the input into a latent-space representation called the code. Although AE comes in various forms, this

study distinguishes three broad categories that differ greatly in their mechanisms of action: SAE, VAE, and D-AE. Regardless of the number of layers in the network, D-AE learns similarly to other AEs, training each layer simultaneously. Different stacked AE blocks are trained independently in SAE, with each block's representation (code) serving as input for the subsequent block. SAE, which is an AE-based DBN (DBN-AE), will be examined alongside DBNs next. Proposed in 2013 [134], VAEs are distinct from other AEs in that they are based on a layer of data means and standard deviations that facilitates random sampling and simple interpolation. One of the most effective generating techniques is VAE. Generative DL models are used to investigate VAE.

Since AEs—as opposed to VAEs—are typically employed for feature extraction, hybrid models are typically created by combining AE models with other discriminative DL models. For instance, a CNN/AE hybrid model was suggested in [98] for job categorization. We'll talk about hybrid DL models later in this section.

A representative model made up of several RBM or AE networks is called a DBN [30]. As a result, we divide the DBN into two components: DBN-AE (also called stacked AE), which is composed of AE, and DBN-RBM (also called stacked RBM), which is composed of RBM. Some studies have classified features using DBN-RBMs [61], [77], [78], and [86]. For image classification, Lu et al. [77] suggested a deep DBN architecture built on stacked RBM layers. Three RBMs were trained using WPD and FFT, and then a fourth output layer was added to create a four-layer DBN. In [61], a decoding technique for images was developed that used an RBM-based DBN as a classifier and the Lomb–Scargleperiodogram (LSP) for feature extraction. With an average accuracy of 83%, the LSP approach was utilized to extract valuable PSD features from incomplete (part of the data is gone) data with a high degree of artifacts. In a different study [78], the authors suggested a deep DBN-RBM model for feature extraction and SVM for classification, which was improved by a t-distributed stochastic neighbor embedding (t-SNE). With an accuracy of 78.51% utilising a dataset, the study used WPD and CSP to extract the temporal-spectral and spatial characteristics from the data, respectively. A stacked sparse AE model, known as DBN-AE, was presented by Hassanpour et al. [51] for picture categorisation. Using the public dataset, the study increased the quantity of training data using a sliding window augmentation technique and obtained 71% accuracy [131].

GENERATIVE DL MODELS

Training data is usually enhanced and supplemented by generative DL models. GAN and VAE are the two most often used generative DL models. To expand the amount of training data, a number of the research in our study employed conventional data augmentation techniques, or non-DL techniques, such amplitude-pumping [135], sliding window [88], [105], and noise addition [114]. GAN and VAE networks were used in two of the examined research to introduce DL-based data augmentation [64], [91]. These research' findings demonstrated that using GAN models to data augmentation greatly improved classification performance. A four-layer GAN model was suggested by Zhang et al. [91] for data augmentation, and its performance was compared with that of VAE and other conventional augmentation techniques including geometric modification and noise addition. The findings showed that while both GAN and VAE performed better than the conventional approaches, GAN fared the best. In comparison to training without data augmentation, the study discovered that a CNN model trained on data supplemented using GANs performed better on the BCI-C IV-2a [131] and IV-2b [101] datasets by 17% and 21%, respectively. Using the BCI-C III-4a dataset [136], the authors in [64] presented a GAN-based generative model with light architecture for data argumentation, demonstrating that the CNN model's performance increased by 3.57% with the amount of training samples. Additionally, the study showed that GAN models outperform VAE. The authors of a different research [99] suggested a hybrid DL model that combined CNN and VAE. Rather than being a generative model, VAE was employed as a classifier in this work. The next section discusses hybrid models.

HYBRID DL MODEL

Two or more DL models are combined into a single network via hybrid DL models. Researchers have tried to combine several deep learning networks in addition to the solo deep learning models previously discussed, and promising outcomes have been shown for classification tasks [7], [63], [98]–[100], [115], [137], and [138]. Five types of combinations are identified in this review: two discriminative models (e.g., CNN/LSTM [56], [63], [88], [100], [137], [138], CNN/GRU [59], and CNN/MLP [115]); a discriminative model followed by a non-DL classifier (e.g., LSTM+SVM [85] and CNN+SVM [68], [75]); a representative model followed by a non-DL classifier (e.g., CNN/SAE [60], [98]); generative models combined with a discriminative model (e.g., CNN/GAN [64], [91] and CNN/VAE [99]); and a discriminative model followed by a non-DL classifier (e.g., DBN+SVM [78]). The study in [7] introduced a hybrid CNN/RNN model known as recurrent convolutional

neural network (RCNN). One convolutional layer, four recurrent layers, and a fully linked layer made up the model. Prior to being input into the RCNN model, the picture was transformed into spectral images. With an accuracy of 77.72%, the model's performance was examined using the authors' local dataset, which included two classes and three channels. A hybrid CNN/SAE architecture with a six-layer SAE and a 1-D convolutional layer was suggested in the work in [98]. The authors achieved 90.0% and 77.6% accuracy, respectively, using two publicly available datasets (BCI-C II-3 [102] and BCI-C IV-2b [101]) with classes of one and nine participants. A combination of multi-layer CNNs with AE and MLP networks was proposed by the researchers in [115]. Different CNN models that were trained on various frequency bands made up the multi-layer CNNs.

The whole dataset into three subgroups: blood routine test (neutrophil ratio, thrombocytocrit, haematocrit, mean corpuscular hemoglubin, lymphocyte, platelet distribution width, mean corpuscular volume, platelet count, haemoglobin, eosinophil ratio, mean platelet volume, basophil cell count, red blood cell count, mononuclear cell count, red blood cell distribution width, and basophil cell ratio), general chemistry (albumin, calcium, indirect bilirubin, uric acid, nutrium, total protein, alanine aminotransderase, total bilirubin, blood urea nitrogen, magnesium, glucose, creatinine, phosphorus, globulin, gamaglutamyltranferasey, alkaline phosphates, kalium, direct bilirubin, carban dioxide-combining power, chlorine, aspartate aminotransferase, and anion gap) and tumour marker (carbohydrate antigen 72- 4, alpha-fetoprotein, carbohydrate antigen 19-9, menopause, carbohydrate antigen 125, carcinoembryonic antigen, age, and human epididymic protein 4)) (shown in Table 1). The characteristics' names are included, along with certain statistical analysis findings including the mean, standard deviation, 95% confidence interval, and Student's t-test p values. The following table displays the outcomes of training all of the data sets using AI-based algorithm models.

5. PERFORMANCE EVALUATION

Reducing classification mistakes is the primary goal in the feature selection field [25]. Internal cross-validation (I.C.V.) error and external cross-validation (E.C.V.) error are examples of these mistakes. A well-trained classifier is ensured by minimising I.C.V error. However, when there are too many features because of over-training or over-fitting, this error may be negligible or even zero—a phenomenon known as the curse of dimensionality. The capacity to forecast the model for unknown data is shown by the E.C.V error. Together, E.C.V. error and I.C.V. error must be taken into account in order to accurately assess categorisation performance. Thus, E.C.V. MAE and I.C.V. MAE are two objective functions that we take into consideration. The best features (and model parameters) are found at a location where both E.C.V MAE and I.C.V MAE are minimum, as shown in Fig. 1. However, we would rather have the least amount of mistake when identifying the type of cancer. Therefore, maximising sensitivity (Se) is crucial. The margin of classification is one metric that shows how well the classification model generalises to unseen samples [18]. The weighted mean of the classification margins is known as the classification edge (C.E). Selecting the classifier with the greatest edge is one method of selecting among several classifiers, such as when doing feature selection. A model that maximises C.E. makes the data more separable. As a result, two models could perform similarly as classifiers in terms of accuracy or other confusion matrix criteria. However, they selected distinct variables and features from the same dataset. The chosen model in this instance has a higher C.E. The following assessment metrics are used to gauge classifier performance [19][20].
The Patients who test negative are said to be true negatives (TN). Patients who receive a positive cancer diagnosis but test negative are known as false negatives (FN). False positive (FP): Individuals who tested positive but had a non-cancerous diagnosis. Patients who have both a positive test result and a positive cancer diagnosis are said to be true positives (TP). A 2x2 confusion matrix for the two classes (positive and negative) is shown in table 2 below.

Table 3: Confusion Matrix

Real	Predicted	
	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive(FP)	True negative (TN)

Patients were categorised in the study using a serum proteomic pattern diagnostics dataset. The classifier to distinguish between benign and malignant cells was built using the MS proteomic dataset. One hundred features from the input dataset's test data are used to train the neural models. Ion intensity levels at particular mass/charge values make up feature sets. To get the best possible outcome in a reasonable amount of time, individual models are trained using a judicious selection of various parameters. Test datasets were then used to evaluate the models. The training phase was repeated using other sets of models and input parameters in the event that the prediction was not good.[21][22]. Ten hidden layers of neurones were used in the creation of the MLP model. There are three sets of the input

and target sample: training (about 70%), validation (about 15%), and test (about 15%). The network was adjusted to the training data using the training set. Training keeps going until the validation data does not meet the network performance requirements. The test set provides a completely independent way to measure network correctness. Normal patients are indicated by a normalised network output of 0 or 1, respectively. Rules in a decision tree were developed during the learning phase, and test data randomly selected from the training data is used to assess the classifier's accuracy. Unlabelled data is categorised using the tree or rules acquired during the learning phase once correctness has been confirmed. We employed six different types of KNN models—Fine, Medium, Coarse, Cosine, Cubic, and Weighted—for our classification challenge. There was only one neighbour in the Fine model's design. In contrast to the Coarse model, which has 100 neighbours, the Medium, Cosine, Cubic, and Weighted models were set up with ten neighbours apiece. A global module characteristic analysis is initially performed for the discovered modules in comparison to the randomly produced ones or those found by other approaches when the mRNAs and c RNAs within each co-expression module are strongly (anti)-correlated or densely coupled. Gene Ontology (GO) and KEGG pathway enrichment studies are used to assess the biological properties of the lnc RNA-mRNA modules in order to provide a comprehensive picture. A number of parameters must be set for the approach that is suggested here. The ideal hyper-parameters for a machine learning technique vary depending on the dataset. Therefore, we must determine the best set of parameters for the suggested approach in this particular experimental setting. Then, there should be a strong correlation between the same modules. The overall co-expression correlation of the module was represented by the average absolute Pearson Correlation coefficient (PCC), which measures the correlation between all of the lncRNA-mRNA pairings in each module.

Novel approaches using machine learning algorithms hold significant promise for diagnosing cancer and forecasting the course of disease. Using a serum proteome profile dataset, machine learning methods employing a wavelet feature selection technique [9]. Subsequently, supervised machine learning classifiers were used to predict the tumour size using several blood biomarkers. These models included GBM, SVM, RF, CRF, Naive Bayes, Neural Network, and Elastic Net; nevertheless, their AUC scores were below 70% [10]. The cancer stages were then predicted with an accuracy score of almost 83% using a four-staged OC, histological data, several main treatment types, and chemotherapy regimen data [11]. Several models, including SVM, Naive Bayes, XG Boost, LR, and RF, were used in a recent machine learning-based analysis. The XG Boost method produced better model performance, with the best accuracy score of over 80% among the competing models [12]. However, this study was sensitive to the size of the feature set; that is, the accuracy declines by around 60% as the number of features decreases. This work's limited feature count—just 16 distinct blood parameters—has also been a disadvantage. Utilising three distinct biomarker types—blood samples, general chemistry medical tests, and OC markers—they demonstrated a high validation accuracy score but a poor testing accuracy [5], pointing to the existence of over-fitting, a typical issue with machine learning algorithms. Thus, there is an urgent need right now for a strong framework that uses machine learning and statistical analysis to stratify ovarian cancer patients based on biomarker traits. Serum, faeces, and saliva are among the biological samples that include miRNAs, which are tiny non-coding RNA molecules [1]. In addition to being regarded as non-invasive methods for cancer screening and/or detection, miRNAs as biomarkers help create more affordable, disease-specific, and user-friendly microarrays that may be utilised for cancer categorisation, treatment choices, and other purposes. Cancers include the majority of mi RNAs [2]. Therefore, in order to distinguish between the non-cancerous and malignant classes, biomarker mi RNAs need to be extremely precise. The issue of identifying biomarkers in omics data is comparable to the challenge of feature selection. Generally speaking, feature selection techniques fall into one of three categories: filter, wrapper, or embedded approach [3].

Accuracy, precision, recall, F-score, AUC, and log-loss are some of the assessment measures used to assess the classifiers' performance based on True Positives (TP), False Positives (FP), True Negatives (TN), and True Positives (TP). F1-score: This score balances precision and recall and is utilised when there is a class imbalance in the data [22]. ROC-AUC: This metric indicates the model's capacity for discriminating and illustrates the connection between sensitivity and specificity [22]. The area under the sensitivity (TRP) is known as the Area Under the Curve (AUC). Log-loss: By comparing them to precise labels, log-loss determines the ambiguity of a method's likelihood. Better forecasts are indicated by a lower log-loss value [24]. The formula for calculating log loss is H(q), where y is the goal variable's level, p(y) is the point's projected probability given the target value, and q is the log loss's actual value.

6. DISCUSSION

On one end of the spectrum, machine learning algorithms can be supervised, and on the other, they

can be unsupervised. The latter are linked to more sophisticated algorithms that can identify patterns in imaging data without the need for human assistance. Model overfitting, in which the model is optimised for the training dataset but performs poorly on the test dataset, is a common occurrence in ML-based cancer imaging when the number of predictors exceeds the number of data points or samples. The following are the most popular methods to lessen or avoid overfitting: (a) employing methods like k-fold cross-validation using multiple subsamples of the dataset; (b) training the algorithm with more data, if at all possible; (c) performing feature selection, when appropriate, to decrease the dimensionality/number of the initial features; and/or (d) implementing ensemble learning, when practical, to increase data size. A statistical approach is required for the training dataset. However, DL offers more accurate data than ML algorithms.

Oncology uses of DL. Nevertheless, there are a number of obstacles preventing DL from being widely used in clinical settings. Here, we address the difficulties and constraints of DL in clinical oncology and offer our thoughts on potential advancements. One of the biggest obstacles to using DL in cancer is data variability. In immunological histochemistry, for instance, staining strength or characteristics may vary from lab to lab. How DL systems would handle this intra- and inter-laboratory variability is still unknown. Determining the precise processing used to create a sequence library and processed dataset is one of the main challenges for transcriptomic data. Any analysis should include the source and version of the gene model employed because even attributes as fundamental as "the list of human genes" are not fixed and lists of genes and observed splice forms are published and updated by many authority. A wide variety of data transformations (log, linear, etc.) and data normalisations (FPKM, TMM, TPM) are also available, and their implementations in various programming languages produce a combinatorically large number of potential processing paths that, in theory, should produce the same outcomes, but lack a formal procedure to verify that assumption. Finally, clinical data on a sample or piece of data might be prone to inaccuracies, inconsistencies, and incompleteness, and typically do not represent all the complexity of the sample and phenotypic. Designing DL models that are less dependent on or independent of clinical annotations is one possible tactic to address this problem.

I will discuss the performance metrics in more detail in the paragraph that follows. The models that are displayed are evaluated using a variety of measures [38], such as F1-score, accuracy, precision, and recall. Four factors from the aforementioned confusion matrix are used to build these performance requirements, as can be seen in [39]. When analyzing deep learning techniques that provide a broad range of outcomes, this confusion matrix is helpful. So let's talk about the specifics of each of these issues.

Accuracy: a metric that considers the degree of accuracy of forecasts in comparison to all other forecasts.It is defined as the ratio of true positives and true negatives to true positive, true negative, false positives and false negatives.

Recall: It is defined as the ratio of true positive and true positives to false negatives

Precision: It is defined as true positives to the sum of true positives and false positives

F1-Score: Twice the sum of the recall and accuracy ratios.

A useful metric for assessing performance over a variety of threshold values is the Area under the Curve-Receiver Operating Characteristics (AUC-ROC) Curve. AUC shows the capacity to differentiate between classes, whereas ROC shows the probability curve. Here, the model's ability to correctly differentiate between classes is being assessed. A higher AUC is a sign of greater model reliability.

Class	Precision	Recall	F1-Score	Testing Images
Clear Cell	0.88	0.94	0.91	100
Endometroid	0.96	0.96	0.96	98
Mucinous	0.89	0.91	0.91	100
Non-Cancerous	0.88	0.8	0.84	100
Serous	0.96	0.92	0.94	100

7. CURRENT TRENDS IN DL BASED TECHNIQUES

According to the evaluated papers, CNN-based DL techniques were used in 64% of the investigations, and 15% of the studies suggested combining CNN with other DL models, such as representational (like AE), generative (like GAN), or recurrent (like LSTM) models. The following arguments support the extensive usage of CNN-based models. First, deep discriminative features may be extracted using a CNN architecture. As a result, hydrid DL algorithms have been applied to both feature extraction and classification in certain research. Second, CNNs are well-known and easily available because to their significant success in a variety of domains, including image and video processing (public codes). Researchers are therefore more likely to comprehend and modify hydrid models in their study.

CNNs perform better than other deep learning techniques, according to earlier studies. When the authors of [90] examined the performance of CNN and LSTM networks, they discovered that CNN performed noticeably better than LSTM. However,

the LSTM model's architecture was simple, with just one hidden layer, and it used spectral pictures as input, which is more suited for CNN than LSTM. The performance of a hybrid CNN/RNN model was compared with that of conventional CNN and LSTM models in another study [7]. With an accuracy of 84.24%, the CNN model beat the other models, but the hybrid CNN/RNN and LSTM models performed 77.72% and 66.2% better, respectively, according to the data.

86% of the recognised research employed a number of convolution layers ranging from one to five, with two and three convolution layers being the most often used. However, these studies did not precisely evaluate various numbers of convolution and fully linked layers in the CNN models. Other research has proposed CNN models with over 10 convolution layers that are extremely deep. A extremely deep CNN model with 31 convolution layers was suggested by Li et al. [118] and compared to various topologies while varying the number of convolution layers. The study found that greater classification accuracy was attained by the deeper CNN models. Eighty-seven percent of the investigations utilised one or two fully linked classifier layers. But according to the study in [118], performance was improved by employing up to three completely linked layers as opposed to only one.

In terms of the activation function, the exponential linear unit (ELU) (26%) and rectified linear unit (ReLU) (65%) were the most often utilised activation functions in the convolutional layers. CNNs using ReLU and ELU were compared for performance in the study in [115], which demonstrated that ELU performed better in terms of accuracy and time. Less commonly employed activation functions were the hyperbolic tangent (tanh) (3%), scaled exponential linear unit (SELU) (1%), and leaky rectified linear unit (LReLU) (5%). According to Wang et al. [90], SELU fared better than both ELU and ReLU. It is advised to use ReLU as an activation function in the initial convolutional layer building due to the huge number of research that have used it. After that, alternative activation functions, such ELU or SELU, can be used to examine the possibility of performance enhancement.

RNN was employed in just 7% of the examined research, while 8% of the studies recommended combining RNN with CNN. Considering the RNN's shown ability to learn time-series characteristics, this is less than anticipated. These results can be explained by the fact that RNNs need a lot of memory and time, particularly when dealing with lengthy sequences.

8. CHALLENGES AND FUTURE DIRECTIONS

Even though ovarian cancer early detection has improved thanks to AI-based learning approaches and algorithms, technical and usability issues still hinder the widespread real-world implementation of AI-based cancer detection. This section discusses these difficulties as well as possible lines of inquiry.

9. CONCLUSION

This paper presents a novel, neural network-based framework for the non-invasive diagnosis of germ cell cancer. By proposing a structured methodology for the integration of diverse biomarker data and an optimized deep learning model, we have outlined a robust system that can significantly improve diagnostic accuracy and overcome the limitations of traditional methods. The conceptual framework and case study demonstrate its potential to provide a highly accurate and efficient tool for early cancer detection. Future work will focus on the empirical validation of this framework using real, large-scale clinical datasets to move this conceptual model toward clinical application, addressing the data variability challenges highlighted in this study.

The machine learning networks for cancer diagnosis were compiled in this survey. Unlike conventional techniques, machine learning (ML) can automatically extract latent and high-level complex characteristics from raw cancer pictures using deep architecture, removing the need for laborious feature extraction and pre-processing. examined the network architecture, pre-processing approach, input formulation, machine learning approach, and performance assessment of current ML techniques. The most widely used statistical technique for tumour categorisation was parametric based. In immunological histochemistry, for instance, staining strength or characteristics may vary from lab to lab. Machine learning techniques made heavy use of raw data, either with or without little pre-processing. Furthermore, they examined publicly available cancer datasets that may be used to machine learning methods. One of the biggest obstacles to using DL in cancer is data variability.

REFERENCES

[1] Wang X, Wang Y, Chen ZG, et al. Advances of Cancer Therapy by Nanotechnology.
Cancer Res Treat. 2009;41:1–11.
[2] Peer D, Karp JM, Hong S, et al. Nanocarriers as an emerging platform for cancer
therapy. Nat. Nanotechnol. 2007;2:751–760.
[3] Allen TM, Cullis PR. Drug Delivery Systems: Entering the Mainstream. Science.
2004;303:1818–1822.

[4] Matsumura Y, Maeda H. A New Concept for Macromolecular Therapeutics in Cancer Chemotherapy: Mechanism of Tumoritropic Accumulation of Proteins and the Antitumor Agent Smancs. Cancer Res. 1986;46:6387–6392.
[5] Khawar IA, Kim JH, Kuh H-JJ. Improving drug delivery to solid tumors: Priming the tumor microenvironment. J. Control. Release. 2015;201:78–89.
[6] Tomasina J, Lheureux S, Gauduchon P, et al. Nanocarriers for the targeted treatment of ovarian cancers. Biomaterials. 2013;34:1073–1101.
[7] Plummer M, de Martel C, Vignat J, et al. Global burden of cancers attributable to infections in 2012: a synthetic analysis. Lancet Glob. Heal. 2016;4:e609–e616.
[8] Perinotto AC, Caseli L, Hayasaka CO, et al. Dendrimer-assisted immobilization of alcohol dehydrogenase in nano structured films for bio sensing: Ethanol detection using electrical capacitance measurements. Thin Solid Films. 2008;516:9002–9005.
[9] Blanco E, Hsiao A, Mann AP, et al. Nanomedicine in cancer therapy: Innovative trends and prospects. Cancer Sci. 2011;102:1247–1252.
[10] Caruso F, Hyeon T, Rotello VM. Nanomedicine. Chem. Soc. Rev. 2012;41:2537–2538.
[11] Gradishar WJ. Albumin-bound paclitaxel: a next-generation taxane. Expert Opin. Pharmacother. 2006;7:1041–1053.
[12] Saif MW. U.S. Food and Drug Administration approves paclitaxel protein-bound particles (Abraxane(R)) in combination with gemcitabine as first-line treatment of patients with metastatic pancreatic cancer. JOP. 2013;14:686–688.
[13] Northfelt DW, Martin FJ, Working P, et al. Doxorubicin encapsulated in liposomes containing surface-bound polyethylene glycol: pharmacokinetics, tumor localization, and safety in patients with AIDS-related Kaposi's sarcoma. J. Clin. Pharmacol. 1996;36:55–63.
[14] Monneret C. Platinum anticancer drugs. From serendipity to rational design. Ann. Pharm. Fr. 2011;69:286–295.
[15] Mura S, Couvreur P. Nanotheranostics for personalized medicine. Adv Drug Deliv Rev. 2012;64:1394–1416.
[16] Torchilin VP. Micellar nanocarriers: Pharmaceutical perspectives. Pharm. Res. 2007;24:1–16.
[17] Farokhzad OC, Langer R. Impact of nanotechnology on drug delivery. ACS Nano. 2009;3:16–20.
[18] Fan Z, Fu PP, Yu H, et al. Theranosticnanomedicine for cancer detection and treatment. J. Food Drug Anal. 2014;22:3–[19] Boulaiz H, Alvarez PJ, Ramirez A, et al. Nanomedicine: Application areas and development prospects. Int. J. Mol. Sci. 2011;12:3303–3321.
[20] Kranz C, Eaton DC, Mizaikoff B. Analytical challenges in nanomedicine. Anal. Bioanal. Chem. 2011;399:2309–2311.
[21] Venditto VJ, Szoka FC. Cancer nanomedicines: So many papers and so few drugs! Adv. Drug Deliv. Rev. 2013;65:80–88.
[22] Korsmeyer R. Critical questions in development of targeted nanoparticle therapeutics. Regen. Biomater. 2016;3:143–147.
Oldenburg S., Averitt R., Westcott S., et al. Nanoengineering of optical resonances. Chem. Phys. Lett. 1998;288:243–247.
[23] Madani SY, Naderi N, Dissanayake O, et al. A new era of cancer treatment: carbon nanotubes as drug delivery tools. Int. J. Nanomedicine. 2011;6:2963–2979.
[24] Shi Kam NW, O'Connell M, Wisdom JA, et al. Carbon nanotubes as multifunctional biological transporters and near-infrared agents for selective cancer cell destruction. Proc. Natl. Acad. Sci. 2005;102:11600–11605.
[25] Liu Z, Chen K, Davis C, et al. Drug delivery with carbon nanotubes for in vivo cancertreatment. Cancer Res. 2008;68:6652–6660.
[26] Ilinskaya AN, Dobrovolskaia MA. Nanoparticles and the blood coagulation system. Part I: benefits of nanotechnology. Nanomedicine. 2013;8:773–784.
[27] Li Z, Kawashita M, Araki N, et al. Magnetite nanoparticles with high heatingefficiencies for application in the hyperthermia of cancer. Mater. Sci. Eng. C. 2010;30:990–996.
[28] Jabir NR, Tabrez S, Ashraf GM, et al. Nanotechnology-based approaches in anticancer research. Int. J. Nanomedicine. 2012;7:4391–4408.
[29] Sutradhar KB, Amin ML. Nanotechnology in Cancer Drug Delivery and Selective Targeting. ISRN Nanotechnol. 2014;2014:1–12.
[30] Singh R, Lillard JW. Nanoparticle-based targeted drug delivery. Exp. Mol. Pathol. 2009;86:215–223.
[31] Qian WY, Sun DM, Zhu RR, et al. pH-sensitive strontium carbonate nanoparticles as new anticancer vehicles for controlled etoposide release. Int. J. Nanomedicine. 2012;7:5781–5792.
[32] Kaur S, Prasad C, Balakrishnan B, et al. Trigger responsive polymeric nanocarriers for cancer therapy. Biomater. Sci. 2015;3:955–987.
[33] Duncan R. Polymer conjugates as anticancer nanomedicines. Nat. Rev. Cancer.2006;6:688–701.
[34] Luo Y, Prestwich G. Cancer-Targeted Polymeric Drugs. Curr. Cancer Drug Targets. 2002;2:209–226.
[35] Kwon GS. Polymeric Micelles for Delivery of Poorly Water-Soluble Compounds. Crit.Rev. Ther. Drug Carrier Syst. 2003;20:357–403.

[36] Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. **2018**, 68, 394–424. [CrossRef] [PubMed]
[37]Torre, L.A.; Trabert, B.; DeSantis, C.E.; Miller, K.D.; Samimi, G.; Runowicz, C.D.; Siegel, R.L. Ovarian cancer statistics, 2018. CA Cancer J. Clin. **2018**, 68, 284–296. [CrossRef] [PubMed]
[38] Marchetti, C.; Pisano, C.; Facchini, G.; Bruni, G.S.; Magazzino, F.P.; Losito, S.; Pignata, S. First-line treatment of advanced ovarian cancer: Current research and perspectives. Expert Rev. Anticancer Ther. **2010**, 10, 47–60. [CrossRef] [PubMed]
[39] Wang, J.; Gao, J.; Yao, H.; Wu, Z.; Wang, M.; Qi, J. Diagnostic accuracy of serum HE4, CA125 and ROMA in patients with ovarian cancer: A meta-analysis. Tumor Biol. **2014**, 35, 6127–6138. [CrossRef]
[40] Lu, M.; Fan, Z.; Xu, B.; Chen, L.; Zheng, X.; Li, J.; Znati, T.; Mi, Q.; Jiang, J. Using machine learning to predict ovarian cancer. Int. J. Med. Inform. **2020**, 141, 104195. [CrossRef]
[41] Moore, R.G.; Jabre-Raughley, M.; Brown, A.K.; Robison, K.M.; Miller, M.C.; Allard, W.J.; Kurman, R.J.; Bast, R.C.; Skates, S.J. Comparison of a novel multiple marker assay vs the Risk of Malignancy Index for the prediction of epithelial ovarian cancer in patients with a pelvic mass. Am. J. Obstet. Gynecol. **2010**, 203, 228.e221–228.e226. [CrossRef]
[42] Anton, C.; Carvalho, F.M.; Oliveira, E.I.; Maciel, G.A.R.; Baracat, E.C.; Carvalho, J.P. A comparison of CA125, HE4, risk ovarian malignancy algorithm (ROMA), and risk malignancy index (RMI) for the classification of ovarian masses. Clinics **2012**, 67, 437–441. [CrossRef]
[43] Lukanova, A.; Kaaks, R. Endogenous hormones and ovarian cancer: Epidemiology and current hypotheses. Cancer Epidemiol. Biomarkers Prev. **2005**, 14, 98–107. [CrossRef]
[44] Alqudah, A.M. Ovarian cancer classification using serum proteomic profiling and wavelet features a comparison of machine learning and features selection algorithms. J. Clin. Eng. **2019**, 44, 165–173. [CrossRef]
[45] Kawakami, E.; Tabata, J.; Yanaihara, N.; Ishikawa, T.; Koseki, K.; Iida, Y.; Saito, M.; Komazaki, H.; Shapiro, J.S.; Goto, K.; et al. Application of artificial intelligence FOR Preoperative diagnostic And PROGNOSTIC prediction in Epithelial ovarian cancer based on BLOOD BIOMARKERS. Clin. Cancer Res. **2019**, 25, 3006–3015. [CrossRef]
[46] Paik, E.S.; Lee, J.W.; Park, J.Y.; Kim, J.H.; Kim, M.; Kim, T.J.; Choi, C.H.; Kim, B.G.; Bae, D.S.; Seo, S.W. Prediction of survival outcomes in patients with epithelial ovarian cancer using machine learning methods. J. Gynecol. Oncol. **2019**, 30, e65. [CrossRef]
[47] Akazawa, M.; Hashimoto, K. Artificial Intelligence in Ovarian Cancer Diagnosis. Anticancer Res. **2020**, 40, 4795–4800 [CrossRef]
[48] Krithikadatta J. Normal distribution. J. Conserv. Dent. **2014**, 17, 96–97. [CrossRef]
Kim, T.K. T-test as a parametric statistic. Korean J. Anesthesiol. **2015**, 68, 540. [CrossRef]
[49] Verma, C.; Illes, Z.; Stoffova, V.; Bakonyi, V.H. Comparative Study of Technology With Student's Perceptions in Indian and Hungarian Universities for Real-Time: Preliminary Results. IEEE Access **2021**, 9, 22824–22843. [CrossRef]
[50] Sur, P.; Candès, E.J. A modern maximum-likelihood theory for high-dimensional logistic regression. Proc. Natl. Acad. Sci. USA**2019**, 116, 14516–14525. [CrossRef]
[51] Loey, M.; Manogaran, G.; Taha, M.H.; Khalifa, N.E. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Measurement **2021**, 167, 108288. [CrossRef]
[52] Awal, M.A.; Masud, M.; Hossain, M.S.; Bulbul, A.A.-M.; Mahmud, S.M.; Bairagi, A.K. A Novel Bayesian Optimization-Based Machine Learning Framework for COVID-19 Detection From Inpatient Facility Data. IEEE Access **2021**, 9, 10263–10281. [CrossRef]
[53] Yes¸ilkanat, C.M. Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm. Chaos Solitons Fractals **2020**, 140, 110210. [CrossRef]
[54] Meidan, Y.; Sachidananda, V.; Peng, H.; Sagron, R.; Elovici, Y.; Shabtai, A. A novel approach for detecting vulnerable IoT devices connected behind a home NAT. Comput. Secur. **2020**, 97, 101968. [CrossRef]
[55] Alazab, M.; Khan, S.; Krishnan, S.S.; Pham, Q.-V.; Reddy, M.P.; Gadekallu, T.R. A Multidirectional LSTM Model for Predicting the Stability of a Smart Grid. IEEE Access **2020**, 8, 85454–85463. [CrossRef]
[56] Ahamad, M.M.; Aktar, S.; Rashed-Al-Mahfuz, S.; Uddin, S.; Lio, P.; Xu, H.; Summers, M.A.; Quinn, J.M.W.; Moni, M.A. A machine learning model to identify early stage symptoms of sars-cov-2 infected patients. Expert Syst. Appl. **2020**, 160, 113661. [CrossRef]
[57] Braga, A.R.; Gomes, D.G.; Rogers, R.; Hassler, E.E.; Freitas, B.M.; Cazier, J.A. A method for mining combined data from in-hive sensors, weather and apiary inspections to forecast the health status of honey bee colonies. Comput. Electron. Agric. **2020**, 169, 105161. [CrossRef]
[58] Akter, T.; Satu, M.S.; Khan, M.I.; Ali, M.H.; Uddin, S.; Lio, P.; Quinn, J.M.W.; Moni, M.A. Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders. IEEE Access **2019**, 7, 166509–166527. [CrossRef]

[59] Uddin, S.; Khan, A.; Hossain, M.E.; Moni, M.A. Comparing different supervised machine learning algorithms for disease prediction. BMC Med. Inform. Decis. Making **2019**, 19, 281. [CrossRef]
Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput. Sci. **2021**, 2, 160. [CrossRef]
[60] Rauh-Hain, J.A.; Krivak, T.C.; Del Carmen, M.G.; Olawaiye, A.B. Ovarian cancer screening and early detection in the general population. Rev. Obstet. Gynecol. **2011**, 4, 15–21.
[61] Gorski, J.W.; Quattrone, M.; Van Nagell, J.R.; Pavlik, E.J. Assessing the costs of screening for ovarian cancer in the United states: An evolving analysis. Diagnostics **2020**, 10, 67. [CrossRef]
[62] Martínez-Más, J.; Bueno-Crespo, A.; Khazendar, S.; Remezal-Solano, M.; Martínez-Cendán, J.P.; Jassim, S.; Al Assam, H.; Bourne, T.; Timmerman, D. Evaluation of machine learning methods with Fourier Transform features for CLASSIFYING OVAR- IAN tumors based on ultrasound images. PLoS ONE **2019**, 14, e0219388. [CrossRef]
[63] S. Singh, D.V. Saxena, S. Khatri, S. Gupta, J. Garewal and
K. Dubey, Histopathological evaluation of ovarian tumors, Undefined 2 (2016), 435–439.
[64] L.J. Havrilesky, G.D. Sanders, S. Kulasingam, J.P. Chino, A. Berchuck, J.R. Marks and E.R. Myers, Development of an ovarian cancer screening decision model that incorporates disease heterogeneity: Implications for potential mortality reduction, Cancer 117 (2011), 545–553. doi: 10.1002/cncr.25624.
[65] J.A. Rauh-Hain, T.C. Krivak, M.G. Del Carmen and A.B. Olawaiye, Ovarian cancer screening and early detection in the general population, Rev. Obstet. Gynecol 4 (2011), 15–21.
[66] US Preventive Services Task Force, D.C. Grossman, S.J. Curry, D.K. Owens, M.J. Barry, K.W. Davidson, C.A. Doubeni, J.W. Epling, A.R. Kemper, A.H. Krist, A.E. Kurth, C.S. Landefeld, C.M. Mangione, M.G. Phipps, M. Silverstein,
M.A. Simon and C.-W. Tseng, Screening for ovarian cancer: US preventive services task force recommendation statement, JAMA 319 (2018), 588. doi: 10.1001/jama.2017.21926.
[67] S. Fenchel, D. Grab, K. Nuessle, J. Kotzerke, A. Rieber, R.
Kreienberg, H.-J. Brambs and S.N. Reske, Asymptomatic adnexal masses: Correlation of FDG PET and histopatho logic findings, Radiology 223 (2002), 780–788. doi: 10.1148/ radiol.2233001850.
[68] K.B. Mathieu, D.G. Bedi, S.L. Thrower, A. Qayyum and R.C. Bast, Screening for ovarian cancer: Imaging challenges and opportunities for improvement, Ultrasound Obstet. Gynecol
51 (2018), 293–303. doi: 10.1002/uog.17557.
[69] M.A. Rossing, K.G. Wicklund, K.L. Cushing-Haugen and N.S. Weiss, Predictive value of symptoms for early detection of ovarian cancer, JNCI J. Natl. Cancer Inst 102 (2010),
222–229. doi: 10.1093/jnci/djp500.
[70] B. Khiewvan, D.A. Torigian, S. Emamzadehfard, K. Paydary, A. Salavati, S. Houshmand, T.J. Werner and A. Alavi, An up-
date on the role of PET/CT and PET/MRI in ovarian cancer, Eur. J. Nucl. Med. Mol. Imaging 44 (2017), 1079–1091. doi:
10.1007/s00259-017-3638-z.
[71] V.R. Iyer and S.I. Lee, MRI, CT, and PET/CT for ovarian cancer detection and adnexal lesion characterization, Am. J. Roentgenol 194 (2010), 311–321. doi: 10.2214/AJR.09.3522.
[72] M. Montagnana, M. Benati and E. Danese, Circulating biomarkers in epithelial ovarian cancer diagnosis: From present to future perspective, Ann. Transl. Med 5 (2017), 276–276. doi: 10.21037/atm.2017.05.13.
[73] Q. She, B. Hu, Z. Luo, T. Nguyen, and Y. Zhang, "A hierarchical semi-supervised extreme learning machine method for EEG recognition," *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 147–157, 2018.
[74] S. Taheri, M. Ezoji, and S. M. Sakhaei, "Convolutional neural network based features for motor imagery EEG signals classification in brain–computer interface system," *SN Appl. Sci.*, vol. 2, no. 4, pp. 1–12, 2020.
[75] X. Ma, D. Wang, D. Liu, and J. Yang, "DWT and CNN based multi- class motor imagery electroencephalographic signal recognition," *J. Neural Eng.*, vol. 17, no. 1, p. 16073, 2020.
[76] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines," *IEEE Trans. neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 566–576, 2016.
[77] J. Xu, H. Zheng, J. Wang, D. Li, and X. Fang, "Recognition of EEG signal motor imagery intention based on deep multi-view feature learning," *Sensors*, vol. 20, no. 12, p. 3496, 2020.
[78] W. Huang, Y. Xue, L. Hu, and H. Liuli, "S-EEGNet: Electroencephalogram Signal Classification Based on a Separable Convolution Neural Network With Bilinear Interpolation," *IEEEAccess*, vol. 8, pp. 131636–131646, 2020.
[79] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "LSTM-based EEG classification in motor imagery tasks," *IEEE Trans. neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2086–2095, 2018.
[80] J.-S. Bang, M.-H. Lee, S. Fazli, C. Guan, and S.-W. Lee, "Spatio- spectral feature representation for motor imagery classification using convolutional neural networks," *IEEE Trans. Neural Networks Learn. Syst.*, 2021.

[81] J. Xue*et al.*, "A Multifrequency Brain Network-Based Deep Learning Framework for Motor Imagery Decoding," *Neural Plast.*, vol. 2020, 2020.
[82] X. Zhao, J. Zhao, C. Liu, and W. Cai, "Deep neural network with joint distribution matching for cross-subject motor imagery brain- computer interfaces," *Biomed Res. Int.*, vol. 2020, 2020.
[83] S. Kumar, A. Sharma, and T. Tsunoda, "Brain wave classification using long short-term memory network based OPTICAL predictor," *Sci. Rep.*, vol. 9, no. 1, pp. 1–13, 2019.
[84] S. Kumar, R. Sharma, and A. Sharma, "OPTICAL+: a frequency- based deep learning scheme for recognizing brain wave signals," *PeerJComput. Sci.*, vol. 7, p. e375, 2021.
[85] L. Cheng, D. Li, G. Yu, Z. Zhang, X. Li, and S. Yu, "A Motor Imagery EEG Feature Extraction Method Based on Energy Principal Component Analysis and Deep Belief Networks," *IEEE Access*, vol. 8, pp. 21453–21472, 2020.
[85] R. Zhang, Q. Zong, L. Dou, and X. Zhao, "A novel hybrid deep learning scheme for four-class motor imagery classification," *J. Neural Eng.*, vol. 16, no. 6, p. 66004, 2019.
[86] R. Zhang, Q. Zong, L. Dou, X. Zhao, Y. Tang, and Z. Li, "Hybrid deep neural network using transfer learning for EEG motor imagery decoding," *Biomed. Signal Process. Control*, vol. 63, p. 102144, 2021.
[87] T. Uktveris and V. Jusas, "Application of convolutional neural networks to four-class motor imagery classification problem," *Inf. Technol. Control*, vol. 46, no. 2, pp. 260–273, 2017.
[88] Z. Wang, L. Cao, Z. Zhang, X. Gong, Y. Sun, and H. Wang, "Short time Fourier transformation and deep neural networks for motor imagery brain computer interface recognition," *Concurr. Comput. Pract. Exp.*, vol. 30, no. 23, p. e4413, 2018.
[89] K. Zhang *et al.*, "Data augmentation for motor imagery signal classification based on a hybrid neural network," *Sensors*, vol. 20, no. 16, p. 4485, 2020.
[90] N. Shajil, S. Mohan, P. Srinivasan, J. Arivudaiyanambi, and A. A. Murrugesan, "Multiclass Classification of Spatially Filtered Motor Imagery EEG Signals Using Convolutional Neural Network for BCI Based Applications," *J. Med. Biol. Eng.*, vol. 40, no. 5, pp. 663–672, 2020.
[91] Y. Rong, X. Wu, and Y. Zhang, "Classification of motor imagery electroencephalography signals using continuous small convolutional neural network," *Int. J. Imaging Syst. Technol.*, vol. 30, no. 3, pp. 653–659, 2020.
[92] S. Roy, A. Chowdhury, K. McCreadie, and G. Prasad, "Deep learning based inter-subject continuous decoding of motor imagery for practical brain-computer interfaces," *Front. Neurosci.*, vol. 14, 2020.
[93] M. Miao, W. Hu, H. Yin, and K. Zhang, "Spatial-frequency feature learning and classification of motor imagery EEG based on deep convolution neural network," *Comput. Math. Methods Med.*, vol. 2020, 2020.
[94] F. Li, F. He, F. Wang, D. Zhang, Y. Xia, and X. Li, "A novel simplified convolutional neural network classification algorithm of motor imagery EEG signals based on deep learning," *Appl. Sci.*, vol. 10, no. 5, p. 1605, 2020.[95] P. Kant, S. H. Laskar, J. Hazarika, and R. Mahamune, "CWT Based Transfer Learning for Motor Imagery Classification for Brain computer Interfaces," *J. Neurosci. Methods*, vol. 345, p. 108886, 2020.
[96] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, p. 16003, 2016.
[97] M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang, "EEG classification of motor imagery using a novel deep learning framework," *Sensors*, vol. 19, no. 3, p. 551, 2019.
[98] D. Zhang, K. Chen, D. Jian, and L. Yao, "Motor imagery classification via temporal attention cues of graph embedded EEG signals," *IEEE J. Biomed. Heal. informatics*, vol. 24, no. 9, pp. 2570–2579, 2020.
[99] R. Leeb, C. Brunner, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI Competition 2008–Graz data set B," *Inst.*
[100] M. A. Macha *et al.*, "MicroRNAs (miRNAs) as biomarker(s) for prognosis and diagnosis of gastrointestinal (GI) cancers.," *Curr. Pharm. Des.*, vol. 20, no. 33, pp. 5287–97, 2014, Accessed: Aug. 13, 2018. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24479799
[101] Y. Peng and C. M. Croce, "The role of microRNAs in human cancer," *Signal Transduction and Targeted Therapy*, vol. 1. Springer Nature, 2016, doi: 10.1038/sigtrans.2015.4.
[102] C. Lai, S. Guo, L. Cheng, and W. Wang, "A Comparative Study of Feature Selection Methods for the Discriminative Analysis of Temporal Lobe Epilepsy.," *Front. Neurol.*, vol. 8, p. 633, 2017, doi: 10.3389/fneur.2017.00633.
[103] C. Lazar *et al.*, "A Survey on Filter Techniques for Feature Selection in Gene Expression Microarray Analysis," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 9, no. 4, pp. 1106–1119, Jul. 2012, doi: 10.1109/TCBB.2012.33.
[104] E. R. Dougherty, *Genomic signal processing and statistics*. Hindawi Pub. Corp, 2005.
[105] I. V. Ivanov, X. Qian, and R. Pal, *Emerging research in the analysis and modeling of gene regulatory networks*. .
[106] P. Jafari and F. Azuaje, "An assessment of recently published gene expression data analyses: Reporting experimental design and statistical factors," *BMC Med. Inform. Decis. Mak.*, vol. 6, p. 27, Jun. 2006, doi: 10.1186/1472-6947-6-27.
[107] Z. Wang, "Neuro-Fuzzy Modeling for Microarray Cancer Gene Expression Data," 2005.

[108] C. Liu, W. Wang, Q. Zhao, X. Shen, and M. Konan, "A new feature selection method based on a validity index of feature subset," *Pattern Recognit. Lett.*, vol. 92, pp. 1–8, Jun. 2017, doi: 10.1016/J.PATREC.2017.03.018.
[109] A.-C. Haury, P. Gestraud, and J.-P. Vert, "The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures," Jan. 2011, doi: 10.1371/journal.pone.0028210.
[110] R. May, G. Dandy, and H. Maier, "Review of Input Variable Selection Methods for Artificial Neural Networks," in *Artificial Neural Networks - Methodological Advances and Biomedical Applications*, InTech, 2011.
[111] H. Banka, S. Dara, and M. Elloumi, "Feature Selection and Classification For Gene Expression Data Using Evolutionary Computation," in *Biological Knowledge Discovery Handbook*, Hoboken, New Jersey: John Wiley & Sons, Inc., 2013, pp. 421–440.
[112] A. Mukhopadhyay and U. Maulik, "An SVM-wrapped multiobjective evolutionary feature selection approach for identifying cancer-MicroRNA markers," *IEEE Trans. Nanobioscience*, vol. 12, no. 4, pp. 275–281, Dec. 2013, doi: 10.1109/TNB.2013.2279131.
[113] S. Saha, S. Mitra, and R. K. Yadav, "A multiobjective based automatic framework for classifying cancer-microRNA biomarkers," *Gene Reports*, vol. 4, pp. 91–103, Sep. 2016, doi: 10.1016/j.genrep.2016.04.001.
[114] S. Saha, S. Mitra, and R. K. Yadav, "A Stack-based Ensemble Framework for Detecting Cancer MicroRNA Biomarkers," *Genomics, Proteomics Bioinforma.*, vol. 15, no. 6, pp. 381–388, Dec. 2017, doi: 10.1016/j.gpb.2016.10.006.
[115] Y. Zhang, D. W. Gong, X. Y. Sun, and Y. N. Guo, "A PSO- based multi-objective multi-label feature selection method in classification," *Sci. Rep.*, vol. 7, no. 1, p. 376, Dec. 2017, doi: 10.1038/s41598-017-00416-0.
[116] Z. Sherinov and A. Ünveren, "Multi-objective imperialistic competitive algorithm with multiple non-dominated sets for the solution of global optimization problems," *Soft Comput.*, pp. 1–16, Aug. 2017, doi: 10.1007/s00500-017-2773-6.
[117} M. Awad and R. Khanna, "Support Vector Machines for Classification," in *Efficient Learning Machines*, Berkeley, CA: Apress, 2015, pp. 39–66.
[118] D. M. Reeves and G. M. Jacyna, "Support vector machine regularization," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 3, no. 3, pp. 204–215, May 2011, doi: 10.1002/wics.149.
[119] W. Richert, L. P. Coelho, "Building Machine Learning Systems with Python", Packt Publishing Ltd., ISBN 978-1-78216-140-0
[120] J. M. Keller, M. R. Gray, J. A. Givens Jr., "A Fuzzy K-Nearest Neighbor Algorithm", IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-15, No. 4, August 1985
[121] https://www.geeksforgeeks.org/machine-learning/
[122] S. Marsland, Machine learning: an algorithmic perspective. CRC press, 2015.
[123] M. Bkassiny, Y. Li, and S. K. Jayaweera, "A survey on machine learning techniques in cognitive radios," IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1136–1159, Oct. 2012.
[124] https://en.wikipedia.org/wiki/Instance-based_learning
[125] R. S. Sutton, "Introduction: The Challenge of Reinforcement Learning", Machine Learning, 8, Page 225-227, Kluwer Academic Publishers, Boston, 1992
[126] P. Harrington, "Machine Learning in action", Manning Publications Co., Shelter Island, New York, 2012