Machine Learning in Medical Applications: A review of state-of-the-art methods

Mohammad Shehab[a], Laith Abualigah[b,c,*], Qusai Shambour[d], Muhannad A. Abu-Hashem[e], Mohd Khaled Yousef Shambour[f], Ahmed Izzat Alsalibi[g], Amir H. Gandomi[h]

[a]*Information Technology, The World Islamic Sciences and Education University. Amman, Jordan*
[b]*Faculty of Computer Sciences and Informatics, Amman Arab University, Amman, Jordan*
[c]*School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang, 11800, Malaysia.*
[d]*Department of Software Engineering, Al-Ahliyya Amman University, Amman, Jordan*
[e]*Department of Geomatics, Faculty of Architecture and Planning, King Abdulaziz University, Jeddah, Saudi Arabia*
[f]*Department of Scientific Information and Services, Umm Al-Qura University, Mecca, Saudi Arabia*
[g]*Information System College, Israa University, Gaza-Palestine*
[h]*Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo, NSW 2007, Australia.*

Abstract

Applications of machine learning (ML) methods have been used extensively to solve various complex challenges in recent years in various application areas, such as medical, financial, environmental, marketing, security, and industrial applications. ML methods are characterized by their ability to examine many data and discover exciting relationships, provide interpretation, and identify patterns. ML can help enhance the reliability, performance, predictability, and accuracy of diagnostic systems for many diseases. This survey provides a comprehensive review of the use of ML in the medical field highlighting standard technologies and how they affect medical diagnosis. Five major medical applications are deeply discussed, focusing on adapting the ML models to solve the problems in cancer, medical chemistry, brain, medical imaging, and wearable sensors. Finally, this survey provides valuable references and guidance for researchers, practitioners, and decision-makers framing future research and development directions.

Keywords: Machine learning, Medical field, Diagnosis, Healthcare, Medical applications

1. Introduction

Machine learning (ML) refers to the tools, methods, and techniques employed in various fields (e.g., medical). It can help solve diagnostic problems in different medical domains [1], such as medical imaging, cancer diagnosis, and wearable sensor. ML is used to examine significant clinical parameters, such as extracting medical information and predicting diseases and development stages. Therefore, it helps in planning and supporting the patient's status. Furthermore, it ensures efficient healthcare monitoring by helping in data analysis and sending intelligent warnings if necessary [2].

In specialized hospitals, the data of diagnoses for patients are considered medical records [3]. Executing a learning algorithm requires accurate encoding of patients' information. Encoding is a simple step, but ML must start adequately; it should analyze information automatically and match it with similar problems solved previously. Thus, it helps the physician achieve an accurate, easy, and quick diagnosis of new cases. Nonspecialists and students can also use it to diagnose patients.

---

*Corresponding author
Email addresses: `moh.shehab12@gmail.com` (Mohammad Shehab), `Aligah.2020@gmail.com` (Laith Abualigah), `q.shambour@ammanu.edu.jo` (Qusai Shambour), `mabohasm@kau.edu.sa` (Muhannad A. Abu-Hashem), `myshambour@uqu.edu.sa` (Mohd Khaled Yousef Shambour), `asalibi@israa.edu.ps` (Ahmed Izzat Alsalibi), `Gandomi@uts.edu.au` (Amir H. Gandomi)

This review presents an overview of ML methods, techniques, and tools. Articles on approaches and theories supporting ML's medical field application have been collected. The following sections discuss medical diagnosis via ML techniques, ML in medical applications, and utilizing ML methods in a medical context. The paper is organized as shown in Figure 1.
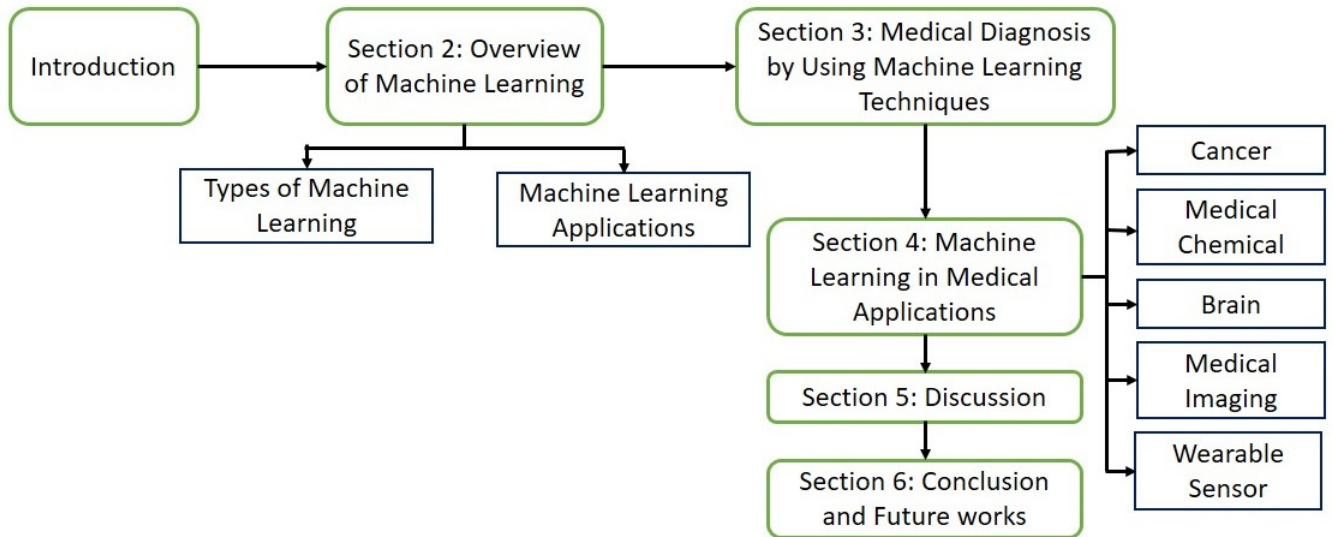

Figure 1: The structure of the review paper

2. Overview of Machine Learning

Here we introduce a background for ML types, methods, and techniques, followed by a discussion of ML's application in the medical area.

2.1. Types of Machine Learning

Some variations exist as to how the types of ML methods are defined (i.e., based on the problem and content of the results required). Table 1 illustrates and briefly describes the most common methods and techniques. Figure 2 shows the detailed relationships among ML types, methods, and techniques.

Table 1: Machine Learning

| Type	Method	Technique	Description	Publication
Supervised	Classification	Support Vector Machine (SVM)	Find a hyperplane in N-dimensional space that distinctly classifies the data points	[4][5][6][7]
		Naive Bayes (NB)	Based on Bayes' theorem with the independence assumptions between predictors	[8][9][10][11]
		Artificial Neural Network (ANN)	Group of linked input/output units where each link is weighted by psychologists and neurobiologists to improve and examine computational analogs of neurons	[12][13][14][15]
		k-Nearest Neighbor (k-NN)	Find the data points that are spread into a number of classes to recalculate the classification of a next sample point	[16][17][18][19]
		Decision Tree (DT)	Divides the dataset into smaller units while a related decision tree is accumulatively built in parallel	[20] [21] [22] [23]
	Regression	Back-Propagation Neural Network (BPNN)	Finding the gradient of the cost function in a neural network with short for the backward propagation of errors	[24] [25] [26] [27]
		Support Vector Regression(SVR)	Same principles as the SVM for classification, with only a few minor differences	[28] [29] [30] [31]
		Multiple Linear Regression(MLR)	Utilizing several explanatory variables to predict the outcome of a response variable for modeling the linear relationship between the response (dependent) variable and explanatory (independent) variables	[32] [33] [34] [35]
		Partial Least Squares (PLS)	Bears some relation to principal components regression; instead of finding hyperplanes of maximum variance between the response and independent variables	[36] [37] [38] [39]
Unsupervised	Clustering	k-Means	Spots the centroids for k data points in order to allocate them to the nearest cluster, while minimizing the number of centroids	[40] [41] [42] [43]
		Hierarchical Algorithm (HA)	A method of cluster analysis which seeks to build a hierarchy of clusters. Where it includes of two types; agglomerative and divisive	[44] [45] [46] [47]
		Mean-Shift	Nonparametric clustering technique which does not require prior knowledge of the number of clusters, and does not constrain the shape of the clusters	[48] [49] [50] [51]
		Density-Based Spatial Clustering of Application with Noise (DBSCAN)	groups points together which are close to each other based on a distance measurement and a minimum number of points	[52] [53] [54] [55]
	Dimensionality Reduction	Feature Selection	Find a subset of the original variables (features). There are three strategies: the filter strategy, the wrapper strategy, and the embedded strategy (i.e., features are selected to add or be removed while building the model based on the prediction errors)	[56] [57] [58] [59]
		Feature Extraction	Transforms the data in the high-dimensional space to space of fewer dimensions, which may be linear, but many nonlinear dimensionality reduction techniques also exist	[60] [61] [62] [63]
Reinforcement	Reinforcement Learning	Q-learning	Provides agents with the capability of learning to act optimally in Markovian domains by experiencing the consequences of actions, without requiring them to build maps of the domains	[64] [65] [66] [67]
		Temporal Difference	An unsupervised technique in which the learning agent learns to predict the expected value of a variable occurring at the end of a sequence of states	[68] [69] [70] [71]
		Value Iteration	Progressively enhancing the value function every iteration until it converges	[72] [73] [74] [75]
		Markov Decision	Dynamic programming, with two variants seen in RL: Policy Iteration and Value Iteration	[76] [77] [78] [79]
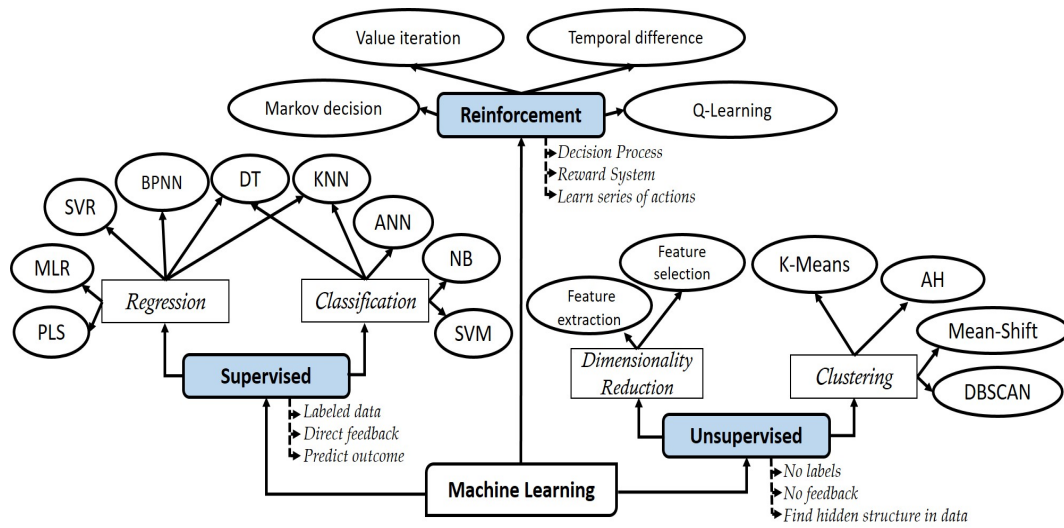
Figure 2: Types of Machine Learning

2.2. Machine Learning Applications

As mentioned above, ML is considered a popular AI application in which devices, software, and computers perform via cognition (i.e., very close to the principal of the human brain). Recently, almost all fields contain at least one of the ML methods, such as email spam and malware filtering, that we deal with every day without realizing it. Additional examples of ML applications are shown in Figure 3. This review discusses ML's employment in the medical field, including healthcare, image processing, and diagnosis.
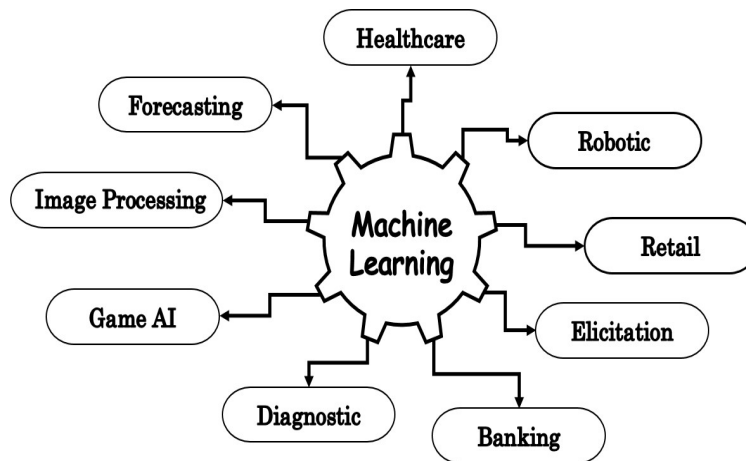

Figure 3: Machine Learning Applications

Several papers on ML's application in the medical field have been widely published. Figure 4 demonstrates published articles amount in the period between 2000 and December 2021. The materials are gathered depending on the keyword "machine learning in the medical field" First, published articles were collected from well-known publishers, including Springer, Elsevier, IEEE, and some other journals via Google Scholar search. Second, the search outcomes were classified per publishing date to show the growth of ML usage in the medical field.
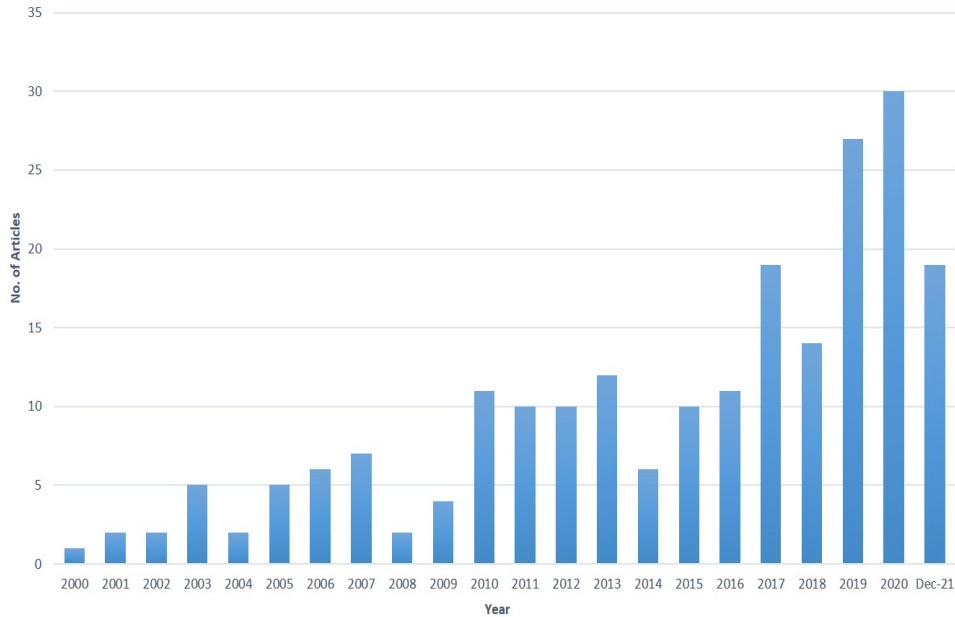
Figure 4: Number of publications for the ML in medical field from 2000 till Dec/2021

3. Medical Diagnosis by Using Machine Learning Techniques

ML algorithms were utilized in computer-aided diagnosis (CAD) applications. Such algorithms learn from many diagnosed samples collected from medical test reports in company with the experts' diagnoses to support medical experts in predicting and diagnosing diseases in the future. The use of ML can assist in enhancing the reliability, performance, and accuracy of diagnosing systems for specific diseases [80, 81, 82].

Wei et al. investigated ML models' applicability for automatically categorizing clustered microcalcifications in digital mammograms. The classifiers were used in computer-aided diagnosis to support radiologists in making precise breast cancer diagnoses on mammograms. They were trained and examined depending on a database of 697 clinical mammograms to determine whether a cluster of microcalcifications is malignant or benign based on quantitative image features extracted from the MCs [83].

Shoeb and Guttag proposed a support vector machine method as a patient-specific classifier to precisely recognize the onset of an epileptic seizure by studying scalp electroencephalogram (EEG) signals [84]. However, such detection is challenging due to the brain's electrical activity. Thus, the critical resolution involves identifying the key features that distinguish a seizure from other types of brain activity.

Li and Zhou proposed a co-training style method called Co-Forest, which can be utilized in computer-aided diagnosis systems. As an ensemble algorithm, random forest is used to deal with the problem of learning a hypothesis from a few samples diagnosed by medical experts. Experiments on three medical datasets confirm the usefulness of the Co-Forest algorithm in building computer-aided diagnosis systems [80].

Ye et al. proposed using a supervised ML method named compound covariate predictor to classify metastatic Hepatocellular carcinoma (HCC) patients, and they recognized genes related to metastasis and patient survival. The algorithm was applied to mixed pairs of 50 HCC tests gathered from 30 patients and proved to produce accurate classification results [85].

Guvenir et al. developed a supervised feature projection-based algorithm for diagnosing cardiac arrhythmia [86]. The training dataset comprises 452 patient records with 279 defined features. Each record contains the decision of an expert cardiologist and clinical measurements, including electrocardiogram (ECG) signals, gender, weight, and age. The authors used pre-classified examples in the learning stage of the algorithm.

A cross-validation model was used to assess the proposed algorithm, and the results were compared with those of other standard classification algorithms.

A predicted model of heart failure using ML techniques was presented by Wu et al. [87]. The authors aimed to provide the right diagnosis of heart failure at least six months before it occurs. 536 data records were collected from 41 outpatient clinics in Pennsylvania from 2001 to 2006. The authors used three ML methods for prediction, logistic regression, support vector machine (SVM), and boosting. Comparative analysis was performed to investigate the performance of each method using 10-fold cross-validation.

4. Machine Learning in Medical Applications

ML is broadly applied to medical and healthcare problems [88]. The next sections will present some areas in which CSA was applied, including cancer (i.e., types and diagnosis), brain problems, medical imaging, and wearable sensors. The classifications in this section are adopted from the contributions of previous works (i.e., showing data based on the problems and, in other cases, based on the ML techniques) [89, 90].

4.1. Cancer

4.1.1. Prediction of cancer

Cancer research is a significant area with considerable societal impact. The use of ML in cancer studies shows high potentials in various aspects, including cancer-related problems' benchmark, such as the classification and prediction of cancer types, drug response, and treatment strategies [91, 92].

Shi et al. presented a hybrid ML method that incorporates the k top-scoring pair (TSP) feature ranking algorithm with SVM and the k-nearest neighbors (KNN) algorithms [93]. The proposed method was motivated by the capabilities of feature selection approaches in improving the performance of the categorization and prediction of complex diseases throughout gene analysis. The hybrid approach is validated using the simulated and the four real cancer prognostic datasets, and it has shown good accuracy performance.

ML models may also be used for predicting the clinical effectiveness of specific medicine and appropriate treatment techniques for particular patients. Accordingly, Menden et al. developed ML models, particularly ANNs, to calculate the reaction of cancer cell lines to medical treatment, which measured throughout IC50 values [94]. The prediction process is based on the chemical properties and the genomic properties of the cell lines of the drugs under consideration. The genomic background of each cell was considered in predicting IC50 profile. Based on their study, the potential efficacy of thousands of drugs can be tested through silico, as anti-tumor agents depending on their formation. This finding can lead to the discovery of innovative drug repositioning opportunities, and it will eventually be valuable for tailored medicine by connecting the sensitivity of the drug to the patients' genomic traits. Likewise, Borisov et al. presented a new methodology for drug scoring and/or tailored medication in which three ML algorithms, namely, SVM, binary tree (BT), and random forest (RF), are used to predict the clinical effectiveness of cancer drugs by transferring attributes attained using the expression-based data from cell lines [95]. The algorithms were tested on different datasets of cancer-like diseases, including lung adenocarcinoma, renal carcinoma, and chronic myeloid leukemia.

By contrast, Huang et al. introduced an open-source software platform that can anticipate the responses of a customized drug from gene expression profiles by utilizing SVM algorithm in conjunction with recursive feature elimination (RFE) model [96]. Specific models were built based on drug response and gene expression data from the panel of cell lines of NCI-60. The models accurately predict the drug receptiveness of a range of cancer cell lines. The prediction results of applying the models were consistent with previously observed responses in the literature. The proposed open-source software platform is expected to undergo tests in various cancer types.

A critical challenge in employing ML in cancer studies is the selection of appropriate software packages to be applied to mitigate cancer-related problems. Wozniak et al. proposed a framework called CANDLE/Supervisor, which deals with hyper-parameter exploration in deep NNs. The proposed framework provides several features to maintain ML in cancer research. First, it allows users to substitute the optimizer or ML problem. Second, it allows the use of diverse-scale datasets. Finally, it allows users to benefit

from numerous computational concurrencies on several leadership-class systems. The software evaluation on clusters and individual workstations revealed good results in terms of scaling and multiplatform execution [91].

Way et al. presented the employment of an ML method to detect faulty genes and pathway activity in cancer tumors, which can help in the identification of hidden responders as their transcriptome may expose responsive molecular states [97]. The algorithm combines copy number, RNA-seq, and mutations from 33 cancer to recognize abnormal molecular states in tumors. The proposed algorithm detects Ras activation throughout cancer types and recognizes phenocopying variations by applying it to the tRas pathway. The algorithm has an unsteady performance across different types of cancer, but it can be generally applied to other pathways and genes. Du et al. proposed a two-phase approach to classify patients' level of fatigue [98]. A limited subset of genes reckoned most predictive by a standardized linear regression approach known as elastic net is selected in the first phase. The second phase distinguishes patients with high from low fatigue using the RF, a broadly used classifier. The proposed approach was trained with cancer patients receiving external beam radiotherapy (T1) and one-month post external beam radiotherapy (T2). Numerous genes were selected, including GPX4, FHL2, and PRDX5, which provided clues concerning the reason of radiotherapy-related fatigue.

Deep learning (DL) considers using conventional ML methods to achieve the goal of learning from data representations. With CT scan pictures and histology, DL has been routinely utilized to detect and classify a malignancy. Thus, Gupta and Malhi used multilayer feedforward NN as a DL technique included in the H2O framework, which is a fast, scalable open-source DL framework, using CT scan pictures to identify the neck and head cance areasr. About 26,000 CT scan images were gathered from the National Cancer Institute's Cancer Imaging Archive (TCIA). The images were preprocessed using the Weiner filter technique and then segmented using the fuzzy c means algorithm; and finally, the gray level co-occurrence matrix technique was used to extract relevant characteristics to train and test the DL classifier [99]. Additionally, to cope with the challenges of feature dimensionality and the limited size of available datasets, Fakoor et al. introduced an approach to improve the diagnosis and classification of cancer types using unsupervised and deep learning methods on gene expression data [92]. The ability to use data from multiple types of cancer to automatically delineate features is a real benefit of the suggested strategy. It helps enhance the diagnosis and categorization of a particular one. The proposed strategy can enhance the classification accuracy in cancer problems, but it can also propose a scalable approach to deal with gene expression data obtained from diverse cancer types.

The development of a computer-aided detection system aimed at diagnosing breast cancer has been receiving increasing interest in recent years. Shaikh and Ali used four classifiers, k-NN, SVM, NB, and J48, on two different cancer datasets: the Wisconsin breast cancer diagnosis and the Breast Cancer Digital Repository. A dimensionality reduction method called WrapperSubsetEval was applied on both datasets before the classification process to improve the classification performance. The experiments were carried out using the well-known Weka tool [100].

Pati employed ML techniques, including multilayer feedforward ANN, random subspace, and minimal sequential optimization, to analyze the gene expressions of cancerous samples and discover the set of genes considered the main possible cause of cancer. The techniques were applied to gene expression data of lung cancer. The author argued that adding ecological context to the advanced ML techniques on the enormous gene expression data available on the Internet may help improve our understanding of the causes of cancer [101]. Table 2 summarizes the ML techniques for prediction the cancer.

Table 2: Summary of cancer's Prediction using ML techniques

Description	Technique	Author(Ref)
Improving the performance of the classification and prediction of intricate diseases throughout gene expression analysis	SVM and KNN	Shi et al. ([93])
determine the reaction of cancer cell lines to medicine treatment	ANNs	Menden et al. ([94])
determine the clinical effectiveness of anti-cancer drugs	SVM, BT, and RF	Borisov et al. ([95])
determine personalized drug restrain from gene declaration description	SVM and RFE	Huang et al. ([96])
Proposed framework provides several features to maintain machine learning in cancer research	ANNs	Wozniak et al. ([91])
recognize the genetic trouble using pattern the disease diagnosis and DNA copy number variation	BHM and GM	Manogaran et al. ([102])
Detect faulty genes and pathways activity in cancer tumors	Gene expression	Way et al. ([97])
Classify patients having high fatigue from low fatigue	regularized random forest	Du et al. ([98])
Identify the locations of the head and neck cancer from the CT scan images	ANN	Gupta and Malhi ([99])
Enhancing the diagnosis and classification of cancer types	ANN	Fakoor et al. ([92])
Developing computer-aided detection/diagnosis system for the diagnosis of breast cancer	k-NN, SVM, and NB	Shaikh and Ali ([100])
Analyze the gene expression of cancerous samples and predict the subset of genes which are the most possible of causing cancer	ANN	Pati ([101])

4.1.2. The Cancer's types
  1. *Lung cancer:*

Lung cancer remains one of the most well-known cancers in the world in regards of new cases (about 0.13 of a year) and mortality (nearly 0.2 cancer death). Errors in lung cancer sample or deadly growth judgment lead to the ineffective operations because the anticancer approach depends on tumor morphology. An attempt to assess the performance of ML methods in the duty of lung cancer analysis and classification based on gene expression stages was performed in [103]. In that work, four publicly accessible datasets were processed (i.e., datasets from the Dana-Farber Cancer Institute, the University of Michigan, the University of Toronto, and Brigham and Women's Hospital, which included 203, 96, 39, and 181 units, respectively). The k-nearest neighbor technique, naive Bayes with the assumption of a couple of normal attribute distribution, and distribution through histograms, SVM, and decision tree were used. The performance of ML techniques was assessed, and SVM showed the best results among all datasets. All algorithms except the decision tree showed the highest possible performance. However, the decision tree showed the best outcomes for the datasets of the University of Toronto. In sum, ML techniques can be employed for lung cancer morphology analysis and related jobs based on gene expression level validation. The problem of distinguishing between 24 diseased and 15 healthy individuals based on protein mass colors was addressed in [104]. To preprocess the data, the volume of charge ratio (m/z) normalization, baseline removal, and conversion of certain peak height ratios to top

ratios were performed. After preprocessing the data, the major challenge faced was a large number of variables (1676 m/z values) opposite the number of cases (forty-one). Dimension reduction technique was employed as an essential part of the analysis. Decision selection was linked with model production in one 10-fold cross-validation circle. Different experimental setups were explored, including two peak top representations, couple decision selection methods, and six selection methods-all on the original 1676-mass and the 124-mass dataset. Important prediction correctness was obtained by a multilayer perceptron and NB, with the latter representing harmonious effectiveness across different experimental options. They attempted to distinguish the most distinct proteins based on rates indicated by the pair variable selection methods and by NN based sensibility judgment. These three plans listed the following four peaks as the more important discriminators: 11683, 1403, 17350, and 66107.

A big data tool for clinical decisions was proposed in [105] to diagnose pneumonitis after stereotactic frame emission treatment (SBRT). A total of 61 characteristics were reported for 201 following cases with SBRT, in which 8 cases (4.0%) were broadcast pneumonitis (RP). Utilizing decision projections, pneumonitis ratios were determined independently for every feature. The effectiveness of three various algorithms was evaluated. On multivariate investigation, the 3 various essential features chosen were the dose 15 of the heart, dose 4 of the bronchus, and culture. Heightened precision was obtained if the proposed algorithm was employed with standardization. To determine the distribution pneumonitis within an error of less than 10%, they predicted that an individual size of 800 cases was needed. The correctness of the classification was restricted by the number of cases in the investigation and not by the opinions collected.

2. *Breast cancer:*
Breast cancer is the most common cancer in the world [106] and the leading cause of cancer mortality in women [107]. It develops in the cells of the breasts. Early recognition and precise diagnosis, on the other hand, can aid to improve the chance of survival [108].
Breast disease cancer (BDC) is the second biggest problem of cancer deaths among women. Simultaneously, it is also one of the most temporary cancer types if diagnosed early [109]. The authors attempted to have reported, with growing confirmation, that SVM has numerous accurate diagnosis experiences. It attempts to confirm that SVM has accurate analytical ability [110]. This paper proposed breast cancer analysis based on the SVM combined with feature selection technique. Experiments were carried out on different common datasets, including the Wisconsin breast cancer dataset (WBCD). These kinds of datasets are often used in ML techniques for breast cancer analysis. . Specificity, sensitivity, classification performance, positive, negative predictive values, and receiver running characteristic confusion matrix and curves are used to evaluate the efficacy of the proposed algorithm. The results revealed that the greatest classification accuracy is 99.5%, which is obtained from the SVM model containing five selected features. This algorithm is more encouraging than the previously published one.
[111] proposed that investigating breast cancer information order assistant can improve processing and save more patients by knowing the genes biomarker to promote the proper treatment model. Thus, producing additional breast cancer studies is now a big problem for researchers, particularly with the great improvement in the fields of bio-informatics, data mining, medical image, and ML techniques, which are a modern reconstruction in cancer therapy.
A growing volume of data demonstrate the importance of patient-oriented communication in the control of healthcare. The decrease in the incidence of denying emotions improves patient well-being and increases primary results. The Verona coding donations of emotional sequences (VR-CoDES) is a strong scheme for classifying and coding patient effects and the corresponding HCP answers. Coded transcripts can be utilized to teach and guide health concern experts (HCPs) and trainees in an aimful and evidence-driven way. Notwithstanding this highest potential, the use of VR-CoDES is still limited for working reasons. The coding system itself is confused, and training is needed to guarantee consistent transcript explanation. Furthermore, the method of explanation is time consuming and difficult even for a demoted expert. Recent progress in ML, especially its use in text examination, could be an essential benefit [112]. This paper explained the literature that examined this instance, and

used 200 discussions between radiotherapists and adult female breast cancer cases. It evaluated various state-of-the-art classifiers in terms of their capacity to predict the answers of HCPs to sensitive cues and attention. The results confirmed a highly encouraging performance, with the greatest classifiers performing forecast efficiencies of over 82%.

The ML method was adopted in [113] for classification. Using feature selection technique, the algorithms supplied data with normal dimensionality and provided precise results. In this paper, experiments were conducted using four distinct feature selection techniques and four classifiers on four datasets. Artificial NNs increase the classification efficiency of breast cancer when utilizing feature selection. The highest improvement in efficiency after using feature selection Entropy was 51%. The precision obtained by artificial NNs did not present any relationship with a particular feature selection approach. With NB, SVM, and decision trees, no development in efficiency using feature selection was statistically determined to respect all datasets. However, in several regards, these classifiers showed heightened classification precision with feature selection as opposed to utilizing all features of the dataset.

[114] proposed 13 cytokine predictors in breast cancer survivors (BCS). This study was concerned with chemotherapy-related cognitive impairment (CRCI), chemotherapy conclusion via multivariate after 6 months to 10 years, and nonparametric procedure. A cross-sectional dataset involved the achievement of a review, cognitive examination, and nonfasting plasma from 66 cases.Data were investigated by utilizing irregular cover modeling to recognize the far more relevant variables for every one of the cognitive analysis rates. A distinct cytokine form divined each cognitive examination. For each reduction, the adjusted R^2 ranged from 0.71-0.77 (p's < 9.50-10). All of the cytokine predictors and cognitive analysis scores had nonlinear similarities. The results showed a novel CRCI field and recommended nonlinear cytokine specificity to NNs cognitive functions evaluated in this investigation. The performance of adjuvant chemotherapy tools for BDC has been compared to variations in the genomic characterization of tumors in [115]. That work investigated equivalence between, increased inhibitory compounds of paclitaxel and gemcitabine (GI50), and gene model number, variety, and expression first in BDC cell lines and then in the specific cases. Genes encoding the immediate aims of these medications, metabolizing proteins, transporters, and those beforehand correlated with chemo resistance to paclitaxel ($n = 31$ genes) or gemcitabine ($n = 18$) were also examined. Multifactorial, principal component investigation showed character was the powerful pointer of consciousness for paclitaxel, and copy amount and expression were useful for gemcitabine. The agents were mixed using SVM. A composition of 15 genes divined cell line irritability to paclitaxel with 82% precision. Copied whole profiles of three genes concurrently with an expression of 7 genes divined gemcitabine acknowledgment with 85% precision. The diagnosis and copy number investigations of two autonomous collections of cases with known acknowledgments were then investigated with these patterns. These injected tumor segments from 21 cases used paclitaxel and gemcitabine and 319 cases used paclitaxel and anthracycline treatment. However, a unique paclitaxel SVM obtained from an 11-gene subset for data for 4 of the primary genes was unavailable. The precision of this SVM was comparable in cell lines and tumor blocks (70%-71%). The gemcitabine SVM showed 62% forecast precision for the tumor blocks due to the occupation of individuals with weak nucleic acid honor. Nevertheless, the paclitaxel SVM divined sensation in 84% of cases with no or smallest remaining infection.

An interactive machine system assesses and makes diagnoses based on cytologic characteristics obtained straight from a digital scan of fine-needle aspirates (FNA) slides in [116]. A continuous chain of 569 cases contributed data to improve the system, and the additional 54 cases were connected. These new cases presented examples to examine the system. The proposed planned precision of the system determined by 10-fold cross-validation was 97%. The real precision of 54 unique samples was 100%. A digital image report linked with ML procedures can help develop the diagnostic precision of breast fine director aspirates.

Missing data accusation is an essential task in situations, where all accessible data and not divorce records with avoiding values must be utilized. This work assessed the effectiveness of various statistical and ML techniques that were employed in [117] to prophesy recurrence in cases in huge physical breast cancer dataset. The correctness of prophecies on early cancer recurrence was estimated using ANNs,

in which several ANNs were evaluated by using datasets with assigned absent values. The allegation systems based on ML algorithms improved accusation statistical techniques in the prophecy of reliable result. Friedman's test exhibited a good variety (p = 0.0091) in the respected area, and the pairwise correlation examination revealed that the AUCs for MLP, KNN, and SOM were exceeding the AUC from the LD-based prediction paradigm. The techniques derived from ML systems were commonly employed to fix the issue of avoiding values and achieve the significant development of prognosis precision related to allegation systems based on statistical systems.

In [118], the suggested paradigm is the compound of controls and various ML procedures. ML paradigms can assist physicians in decreasing the number of wrong judgments. They utilized models and associations among numerous problems and predicted the result of a disease by utilizing historical problems collected in datasets. This study aimed to introduce a rule-based analysis system with ML procedures for the prophecy of various types of BCD. A dataset was studied with eight cases that involve the reports of 900 cases, in which 876 (97.3%) and 24 (2.7%) cases were women and men, respectively. ML procedures include 10 cross-fold systems were performed together with the introduced paradigm for predicting BCD. The effectiveness of ML procedures was assessed with efficiency, correctness, sensibility, specificity, and area under the ROC circuit. This investigation showed that TRF, a rule depends on analysis paradigm, was the greatest model with the largest stage of efficiency, among others. Therefore, this paradigm is suggested as a helpful agent for BCD prophecy as well as medical judgment construction.

A mechanical Allred Scoring methodology is introduced in [119] for estrogen receptors (ER). Clear equation was utilized to stamp the color photo receiving into evidence color modification through smearing in various laboratory. The Markov stochastic domain paradigm with expectancy-most optimization was applied to the division of the ER cells. The suggested classification approach was determined to produce F-measure 0.95. ANN was consequently utilized to achieve an intensity-based rate for ER cells of pixel color depth characteristics. Concurrently, the relationship rate (i.e., the rate of ER-positive cells) was calculated by numbering of cells. The last ER rate was calculated using the power and dimension scores followed by pathologists. The analysis precision for the analysis of cells through the categorizing in regards of F-measure was 0.9626. The difficulty of individual inter-observer knowledge was discussed by quantifying ER number from a couple of specialist pathologists and suggested method. The interclass relationship performed was higher than 0.90. The investigation has the possible benefit of supporting the pathologist in determination making over standard method. It could be developed as a portion of an electronic judgment backing method and another sense achieving/summary scheme.

The use of ML mechanisms in medical analysis is growing continuously. This growth is largely due to the fact that the performance of analysis and recognition methods has improved in terms of helping therapeutic specialists in diagnosing infections [120]. In [121], the BCD study was carried out by applying the Least Park Support Vector Machine (LS-SVM) classifier method. The robustness of the LS-SVM was examined utilizing analysis precision, analysis of sensibility and specificity, k-fold cross-validation approach, and confusion model. The achieved analysis precision of 98.53% is very encouraging, especially when compared with the earlier detailed analysis methods. Consequently, by LS-SVM, the received outcomes revealed that the applied system can create an active discussion and showed the head of the purpose of a distinct creative support diagnosis method.

[122] used gene modification profiles to analyze, identify, and predict the subgroups of BCD. The full exome sequencing data were examined using 358 ethnically comparable BCD in the TCGA design. The physical and nonsynonymous particular nucleotide modifications distinguished through all cases were used to determine a quantitative result. Applying these rates with positive template factorization approach, they grouped the cases into three subsets. This investigation showed that gene modification profiles can be efficiently applied with unsupervised ML techniques to recognize clinically distinct BC subgroups. The analysis model improved in this design could supply a fair prediction of the degree of cancer based solely on the mutation characterizations. This research represented the primary employment of only somatic modification form data to recognize and predict the BCD subgroups. This generic approach can also be used to other cancer populations.

11

Microarray breast cancer data were utilized in [123] to classify the cases that applied ML systems. First, eight various ML methods were employed to the data, without using any feature selection technique. A couple of various feature selection techniques were then employed. The results of the analysis were matched with one another and with the completion of the first state. After implementing the feature selection techniques using the best 50 features, SVM yielded the best outcomes. MLP used numerous layers and neurons to investigate the impact of the number of layers and neurons on the analysis precision. However, the improvement in the number of layers was only slight and precision was not enhanced in some cases.

3. *prostate cancer:*

Diagnosing cancer using computer capabilities has grown rapidly in recent years. ML is one of the techniques broadly used to diagnose and predict prostate cancer. In this section, we discuss the methods that applied ML to improve diagnosing prostate cancer.

First, Nguyen et al. proposed an automated Gleason scoring system for diagnosing prostate cancer. The proposed method combined Quantitative Phase Imaging (QPI) techniques to report the unlabeled samples and ML methods to classify the tissues and diagnose tissue biopsies [124]. Gu et al. aimed to utilize the capabilities of the ML method to achieve the timely diagnosis of prostate cancer right after prostatectomy [125].

Zhu et al. proposed an online adaptive radiotherapy tool for evaluating adaptive intensity-modulated radiotherapy (IMRT) using the capabilities of ML methods. Their proposed method aimed to identify quality standard and requirements of prostate adaptive IMRT [126].

Another prostate cancer diagnosing method using ML techniques was proposed by Hussain et al. [127]. Multi-ML techniques, such as SVM and Bayesian, were used to efficiently diagnose prostate cancer. Some feature extraction methods were also used for further efficiency enhancement.

An analysis for magnetic resonance radiomics to enhance the performance of PI-RADS v2 was proposed by Wang et al. [128]. The major purpose of the work was to test the capabilities of ML methods in enhancing the performance of detecting prostate cancer.

4. *Pancreatic cancer:*

Circulating exosomes include a wealth of proteomic and transmitted information, which highly benefit cancer diagnostics. While microfluidic methods are employed to strongly divide cells from multiple samples, estimating these procedures for exosome separation has been defined by the mean throughput and susceptibility to clogging of nanofluidics. Furthermore, the study of exosomal biomarkers features large heterogeneity between patients and within a growth itself. To tackle these challenges and in order to investigate crude clinical individuals, a multichannel nanofluidic method was developed in [129]. Using these principles, exosomes from healthy, infected murine, clinical groups, and profiled the RNA cargo inside of these exosomes were isolated. ML algorithm was employed to generate predictive boards that could recognize samples obtained from various cancer-bearing people. This method was applied on cancer and precancer mice from healthful handles as well as pancreatic cancer cases from healthy handles in blinded reports.

Pancreatic cancer is one of the primary terminal cancers, ranking 4th among all cases of cancer-related mortalities. Pancreatic cancer cases suffer from bad prognosis with a 5-year remnant rate of just 6%. Predicting remaining pancreatic cancer is tested due to various tumor characteristics, therapies, and patient groups. Certain predictions helped in improving personalized attention and control. The effectiveness of ML was tested in [130] to foretell pancreatic cancer durability. Pancreatic cancer cases were classified through the Monitoring, Epidemiology and End Outcomes database (2010-2014). Clinical data, such as age, sex, and culture, were selected for the cases. Patients' studies were incidentally broken into parts for training (80%) and testing sets (20%) to foretell durability at 6, 12, and 24 months. Several supervised ML paradigms were examined to classify the patterns with best forecasts. A total of 14,631 cases had a median survival of 13 months. Forest algorithm obtained better

outcomes in comparison with other examined models. The trained paradigm admitted AUCs of 85.3%, 84.6%, and 83.2% at six, 12, and 24 months, respectively. The most critical features that inspired prediction were age at diagnosis (19.9%), surgery (14.6%), tumor size (18.5%), and tumor extension (8.4%). Thus, ML was better able to predict the survival of cases with pancreatic cancer compared with other methods. Improved survival prediction helped improve decision handling and social and care demand organization. Table 3 shows the different types of cancer with the ML techniques which have been used for each type.

Table 3: Summary of Cancer's types using ML techniques

Types of Cancer	Techniques	Authors(Ref)
Breast Cancer	SVM	Akay ([110])
	Feature Selection	Abou Tabl et al. ([111])
	VR-CoDES	Barracliffe et al. ([112])
	ANN	Reinbolt et al. ([131])
	Feature Selection	Lindqvist and Price ([113])
	RF	Kesler et al. ([132])
	RF and ANN	Henneghan et al. ([114])
	SVM	Dorman et al. ([115])
	Feature Selection	Wolberg et al. ([116])
	ANN	Jerez et al. ([117])
	TRF	Montazeri et al. ([118])
	ANN	Mungle et al. ([119])
	LS-SVM	Polat and Günes ([121])
	RF, SVM, NB, and k-NN	Vural et al. ([122])
	SVM	Turgut et al. ([123])
Lung Cancer	Gene expression	Podolsky et al. ([103])
	SVM and DT	Hilario et al. ([104])
	DT and RF	Valdes et al. ([105])
Prostate Cancer	RF	Nguyen et al. ([124])
	SVR	Zhu et al. ([126])
	SVM and NB	Hussain et al. ([127])
	SVM	Wang et al.([128])
Urinary Bladder Cancer	SVM	Garapati et al. ([133])
	BPNN, NB, KNN, SVM, RBFN, ELM, and RELM	Wang et al. ([134])
Pancreatic Cancer	LDA	Ko et al. ([129])
	RF	Osman ([130])

4.1.3. Machine learning techniques used in cancer

1. *Classification*

As the applicability of ML methods in medical research field is growing rapidly, they are used as an alternative to traditional methods for diagnosing a wide range of diseases. For example, Chan et al. proposed a method to diagnose oral cancer using ML. The proposed method presents a combination between feature selection and ML methods. To test the proposed feature selection methods, they were applied on four proposed classifiers. Experiments showed the superiority of the traditional methods if clinicopathologic and genomic markers were taken in consideration [135].

Furthermore, Liu et al. proposed gene classification method to eradicate gene redundancy and increase classification quality [136]. The proposed method is divided into two main stages: gene selection and finding optimal gene group. To manage the gene selection stage, they proposed a method called recursive feature addition, where statistical similarity measures and ML were merged to improve the selection process. Another algorithm called Lagging Prediction Peephole Optimization was proposed to identify the best gene group for disease prediction and categorization.

Another classification method using ML techniques was proposed by Fuery et al. [137]. Their proposed method aimed to analyze gene expressions to find and classify harmful tissues using SVM. The analysis considered mislabeled and abnormal tissue outcomes and classified experimented tissues.

Parmar et al. evaluated feature selection ML methods to find the most accurate method for anticipating head and neck cancer. The selected methods were tested in terms of performance efficiency and robustness [138]. In addition, Cho and Won examined studies that aimed to assess cancer classification and feature selection methods. The assessment was conducted using three dataset benchmarks for evaluating efficiency. In addition, a hybrid between several classification methods was proposed to enhance classification performance [139].

To efficiently classify epithelial ovarian cancer (EOC), Klein et al. proposed a preliminary study that aimed to utilize the capabilities of ML algorithms and MALDI-Imaging mass spectrometry. The main purpose of the proposed hybridization was to detect EOC histological subtypes in the hope that such a step would lead to the discovery of new factors to predict and diagnose EOC [140]. Furthermore, Way et al. presented a method to anticipate and classify cancer with low neurofibromin 1 (NF1) activity depending on RNAseq data acquired from TCGA (The Cancer Genome Atlas) [141].

2. *Gene expression*

Gene expression data are vast due to the emergence of recent powerful methods in gene level data extraction. Accordingly, gene expression data classification for precisely detecting correspondent patterns, such as cancer species, has become a research hotspot. Recently, several ML algorithms have been employed to build precise classification models. In particular, the cancer classification based on gene expression data using ML techniques has attracted much attention recently as a promising research field [142, 143, 144, 145].

Guyon et al. proposed a gene selection method by using SVM based on recursive feature elimination in order to address the problem of selecting a small subset of genes, recorded on DNA micro-arrays, from broad patterns of gene expression data [146]. The authors built a classifier that was appropriate for genetic diagnosis and drug discovery. The experimental results proved the effectiveness of the proposed selection method in gene selection because it yielded improved classification performance.

Tan and Gilbert proposed the use of supervised ML techniques in correctly classifying cancerous and normal tissues from the gene expression profiles [147]. The authors performed classification tasks using the C4.5 decision tree, after which they bagged and boosted decision trees on seven publicly available cancerous microarray. They confirmed that ensemble ML often performed better than single ML techniques in classifying gene expression data on cancer classification problems.

To reduce the dimensionality and extract valuable gene information from cancer microarray data, Wang et al. proposed the use of a set of feature selection algorithms for the extraction of significant information in microarray data analysis [148]. The use of feature selection algorithms, namely, wrappers, correlation-based feature selection, and filters together with ML algorithms, such as naive Bayes, decision trees, and SVM, has led to their promising applications in choosing relevant genes that are highly correlated to cancers with high confidence.

Jin et al. focused on classifying binary and multicategory cancer types. The authors first dealt with the high-dimensional problem using the Chi-square method for tag selection of the serial analysis of gene expression before building the classification model. They then applied five different ML algorithms (C4.5, SVM, nearest neighbor, naive Bayes, and RIPPER) for classifying cancer types. The experimental results confirmed that the gene selection can enhance the performance of all classifiers [149].

Chen et al. demonstrated a novel supervised ML model based on Monte Carlo methods, local field, and

SVM theory. The proposed model was applied to accurately find patterns in high-dimensional gene datasets of colon cancer [142]. Meanwhile, Ghanat Bari et al. proposed a hybrid ML-assisted network inference that exploited the capability of ML and network biology to improve the understanding of the existence of Class II cancer genes by uncovering it in cancer networks [143]. The authors constructed more than 2×108 SVM models for reforming a cancer network using 8807 expression arrays, corresponding to 9 cancer types.

Ayyad et al. proposed two efficient classification techniques for gene expression data, namely, smallest modified k-nearest neighbors and largest modified k-nearest neighbors [144]. Both techniques were developed to enhance the performance of nearest neighbors by employing robust neighbors via new weighting strategy.

Xu et al. proposed a deep flexible neural forest model to classify cancer subtypes for the diagnosis and treatment of cancer. Their model aimed to transform a multiclassification problem into several binary classification problems. In order to remove redundancy and noise from the gene expression dataset, the authors proposed a dimensionality reduction method that used the fisher ratio and the neighborhood rough set to select the most informative genes among a given gene expression, thus attaining good classification performance [145].

4.2. Medicinal Chemistry

There are many studies have been introduced about utilizing organic chemistry to discover the drug. Some of the efficient studies used machine learning models which aid to introduce predictions. However, the results may not accurate and fuzzy (i.e., outcomes of black-box). Thus, Kovacs et al. [150] introduced new mechanism to retrieve the expected results and send them to specific locations in reactants. Also, the authors determined Clever Hans predictions which based on dataset bias (i.e., use the wrong reason to find the correct prediction).

In [151], Patsilinakos et al. utilized machine learning on the chemical compositions of essential oils (EOs) to explain the achieved experimental results. Moreover, the authors developed accurate ML models which contributed to determining the Staphylococcus aureus and Staphylococcus epidermidis which are responsible to produce, stimulate, or inhibit biofilm.

Examining the reproductive toxicity on the human directly to assess the health risks is very important. However, it face many challenges such as time consuming, expensive, and need a large number of animals. Therefore, Jiang et al. [152] used six machine learning approaches and 1823 chemicals for reproductive toxicity characterized. The experimental results illustrated that SVM model achieved best performance. As well as, rational boundary has been determined which classified the accurate predictions and inaccurate predictions.

Drevinskas et al. [153] used machine learning methods (classification and regression tree and deep neural network classification) to find antiviral properties of medicinal plant extracts. The authors introduced three chemical analysis on 16 medicinal plants have been chosen and investigated their antiviral properties and chemometric characteristics. The experimental results showed that the possibility of ensure chemometric attributes that affect antiviral activity in medicinal plant extracts.

4.3. Brain

The brain is the most complex organ in the human body because it consists of about 100 billion neurons and one million billion (1015) interconnections [154]. It controls sensorimotor functions, such as walking and breathing; cognitive functions, such as talking, reasoning, and memorizing; and other complex functions, such as emotions and feelings. The brain is also subjected to many diseases requiring surgery, which could result in either the deterioration of the cited functions or even in permanent disability. Therefore, functional magnetic resonance imaging (fMRI) has emerged as a powerful new instrument for collecting vast quantities of data about activity in the human brain.

Davatzikos et al. applied high-dimensional nonlinear pattern classification methods to fMRI images in order to distinguish between the spatial patterns of brain activity associated with lie and truth [155]. The authors performed a forced-choice deception task on 22 participants. The predictive accuracy was 88% from

99% of the true and false responses were discriminated correctly. Nonlinear ML techniques were used in lie detection and other possible clinical applications of fMRI in individual subjects. The results indicated that accurate clinical tests could be based on measurements of brain function with fMRI.

The paper of Zacharaki et al. [156] focused on MRI texture and shape by using the classification of brain tumor type and grade. Given that the automated computer analysis tool proved to be better than human readers, the reliability and reproducibility of brain tumor diagnostic procedures were improved. Thus, a computer-assisted classification method combining conventional and perfusion MRI was used and developed for differential diagnoses. The authors applied the feature extraction of the tumor shape and intensity characteristics and then utilized feature selection through SVM with recursive feature elimination.

EEG is a noninvasive tool that deals with ML approach. It is an electrophysiological monitoring method for recording the electrical activity of the brain. Consequently, [157] focused on biometrics to understand and distinguish humans based on their behavioral characteristics [158]. A framework was established for the Visual Evoked Potential-based biometrics, and all the results passed through classification and feature extraction stages.

In [84], Shoeb and Guttag introduced and evaluated the ML approach (i.e., SVM) by analyzing the scalp EEG to detect the onset of an epileptic seizure by constructing patient-specific classifiers. The proposed approach included shaping the problem into an appropriate ML framework, then separating seizure from other types of brain activity by identifying the critical features. The proposed approach was applied on two or more seizures per patient and tested on 916 hours of continuous EEG from 24 patients. The results showed that the proposed approach detected 96% of 173 test seizures with a median detection delay of 3 seconds and a median false detection rate of two false detections per each 24-hour period.

The efficient EEG-based brain-computer interfacing (BCI) and mental state monitoring applications have been briefly reviewed by preprocessing and classification techniques [159]. BCI allows for communication that is solely based on brain signals and is independent from muscles or peripheral nerves [160]. Furthermore, the authors introduced an outline of the Berlin brain-computer interface (BBCI), widely considered to be a noninvasive tool. The results were presented with a real-time arousal monitoring experiment.

Müller et al. presented linear and nonlinear ML techniques to BCI [161]. These techniques are a salient ingredient of the BBCI online feedback system. Their work focused on selecting the feature selection and highlighting the common weaknesses when validating ML methods in the context of BCI. It provided a brief overview of BBCI, which can be used for helping individuals not trained heavily to enhance and broaden the spectrum of practical applications in BCI.

Three EEG datasets were employed in [162] to evaluate ML algorithms for use in BCI. Each of these datasets was collected for a particular task. The first dataset is called EEG self-paced key typing for detecting explicit left/right (L/R) button press. The second dataset is called EEG synchronized imagined movement for predicting imagined L/R button press. The third dataset is called closed-loop cursor control for vertical cursor control. The evaluation was performed by submitting 10 entries to the competition, where the winning results reported for two of the three datasets (first and second datasets).

In an interesting research, Gravesteijn et al. [163] investigated the added value of various machine learning algorithms, such as SVM, NN, and random forests (RF), in predicting the outcome of moderate and severe traumatic brain injury (TBI). In a low-dimensional setting, the results showed that the performance of ML algorithms is no better than that of traditional regression-based models, and that there was bias when comparing ML algorithms to other traditional regression approaches in earlier studies that supported ML performance.

Lv et al. [164] used improved CSP and transfer learning algorithms to enhance the accuracy of classifying EEG signals and speed up training time. Experiments were performed on real and imaginary movements, performed by 7 participants, in which five different subjects were applied with the right and left hand. The results reveal that movements of left and right hand are more precise at varying speeds than those with the same speed. Furthermore, the proposed algorithm outperforms standard classification algorithms in terms of data accuracy, as evidenced by the findings.

Khan et al. [165] introduced a multimodal classification approach using deep learning mechanism to classify the tumors of brain. The proposed approach begins with utilizing the linear contrast stretching. Following that, conducting feature extraction using deep learning. Thereafter, applying a correntropy-based

16

joint learning approach and the extreme learning machine to choose best features. The component features were then merged into a single matrix. Finally, the matrix is fed into an extreme learning machine to be classified. The results show high performance in terms of classification accuracy and computational time.

Flanders et al. [166] present a very large and heterogeneous dataset of brain computed tomography studies to help diagnose intracranial hemorrhages using machine learning applications. There are 874,035 photos in the dataset, which were gathered from a variety of sources, including scanner manufacturers, organizations, and countries. The dataset includes the classification of intracranial hemorrhages by a broad number of neuroradiologists volunteers. Finally, this dataset was entered into the 2019 Machine Learning Challenge organized by the RSNA Radiology Informatics Committee.

4.4. Medical Imaging

The contents of this section are close to those of the previous section (i.e., Brain) because both sections are based on the same concepts, such as fMRI, MRI, and EEG [167, 168, 169]. The patterns of the medical images can be recognized patterns through ML technique,allowing radiologists to make informed decisions based on radiological information, like basic radiography [170], computed tomography (CT) [171], MRI [172], positron emission tomography (PET) images [173], and radiology reports [174].

Zhang et al. presented a sequential enhancement learning technique for increased performance when using SVM to detect microcalcification (MC) clusters in digital mammograms.The authors proposed a novel tumor segmentation method by exploring a one-class SVM, which was capable of learning the nonlinear distribution of the tumor data without employing prior knowledge [175]. By contrast, El-Naqa et al. employed SVM to reduce a structural risk [176]. They also turned MC detection into a supervised-learning problem and utilized SVM to see if the MC was present at each location in the image.

ML and pattern recognition algorithms have a significant impact on brain imaging as they help gain novel biological insights. However, ML techniques are not usable for nonexperts. This is due to the presence of pitfalls that can lead to false and illogical interpretations. Meanwhile, Lemm et al. introduced a review article to provide an easy and straightforward introduction to algorithmic principles by presenting the most frequent algorithmic topics from both a theoretical and a practical neuroscience data analyst's perspective [177].

[178] highlighted six types of applications in ML and radiology. All applications were distributed accurately based on their contents. The authors attempted to provide all possible categories that could be compared to a well-trained and professional radiologist, such as, diagnosis, computer-aided detection, segmentation of the medical image, and registration. The development of technology in ML areas and radiology can mutually benefit in the long run. This work benefits patients by improving accuracy, lowering expenses, or expanding experiences.

A set of interested algorithms have appeared recently while training the deep artificial neural network (ANN). However, these algorithms faced some limitations in fixing the actual problems because of overfitting and vanishing gradient, as well as shortage of computing power (i.e., the lack of enough data to train the computer system) [179].

DL is a branch of ML that deals with algorithms that are inspired by the biological and function of the brain (i.e., ANN) [180].DL has quickly become the method of choice for assessing medical images in the field of medical imaging, which has led to the increasing the number of related studies. Litjens et al. collected around 300 contributions-most of them published in 2016 and the first month of 2017 [181]. The aim of this study was to show how DL approaches have penetrated the entire field of medical image processing, to identify barriers in applying deep learning to medical imaging tasks, and to highlight unique contributions that address or overcome these challenges. Among the applications covered in this study are neuropathology, abdominal, pulmonary, digital, cardiac, retina, musculoskeletal applications, and breast.

[182] presented the basis of DL methods, such as (i) photo recording successes to find the correspondence of the model point;; (ii) reveal anatomical/cellular structures by applying the assembly process after each layer so that features of larger input regions are progressively compressed; (iii) tissue segmentation for quantitative brain assessment at all ages; and (iv) computer-aided detection to observe abnormal or suspicious areas and thus alert clinicians to draw attention.

A dimensional reduction strategy for medical imaging classification was introduced by [183]. The authors formulated the task as a constrained optimization problem with generative and discriminative objectives, show how to extend it to semi-supervised learning (SSL) setting. They also introduced accuracy rates that overcome other well-known algorithms on different datasets by showing a representation of the set distinction, which is compatible with the previous clinical studies.

The classification is considered as the most common use of ML. For instance, nonlesions or lesions and normal or abnormal classifications depend on the features of input (i.e., the features of the segmented object). Thus, Suzuki collected and compared the pixel/voxel-based ML (PML), which directly utilized the values of pixel/voxel in the medical images [184]. The contributions of the work are as follows: observing the PMLs to clarify their classes, defining the differences and similarities between PMLs and the ML features, determining the merits and weaknesses of PMLs, and illustrating the PML applications in medical imaging.

Wang et al. used medical images and ML techniques to introduce a new method for high-dimensional pattern regression [185]. The uniqueness of this work is that it estimates continuous (i.e., for pattern regression) instead of categorical variables, which are considered important in clinical diagnosis. In other words, the proposed method used the images to predict the disease in the earlier stages. The authors employed the basic feature extraction with the common feature extraction techniques to determine the performance of optimal regression. The Relevance Vector Machine was then used to generate regression models of the selected characteristics in order to simplify model parameter selection and reduce the effects of outliers.

ML methods for medical imaging frequently perform poorly on important subgroups of the population that are not discovered during training or testing. As a response, Luke Oakden-Rayner et al. [186] evaluate the effectiveness of different potential strategies for assessing hidden stratification impacts and define these impacts using synthetic trials on the CIFAR-100 testing set as well as Adelaide Hip Fracture and MURA datasets. They explore the findings' clinical significance and propose that assessing hidden stratification should be a crucial element of every ML implementation in medical imaging.

Medical imaging data preparation is a time-consuming and expensive procedure that yields algorithms with low value and poor generalization. As a response and to reap the benefits of ML in medical imaging, Martin J. Willemink et al. [187] present an overview and clarify the essential methods for arranging and applying medical imaging data in ML algorithms. They discuss the existing constraints of data management and propose novel solutions to the challenge of data availability.

In addition, a number of review articles have recently been published to look into various areas of machine learning in medical imaging [188]. The recent advancement of ML in COVID19 identification and diagnosis is reviewed by Rokaya Rehouma et al. [189]. In the review, they concentrate on ML methods based on CT and Xray pictures published in reputable journals, as well as a discussion of the most common aspects of medical imaging in COVID19 patients. Furthermore, a review article of the employment of AI and ML in medical imaging is presented by Ana Barragán-Montero et al. [190] to introduce the fundamental technological foundations of AI, as well as cutting-edge ML algorithms and their implementation to medical imaging. They explore state of the art trends and upcoming directions to aid the reader in comprehending how AI algorithms are turning into a commonplace component in any medical image analysis process, paving the road for AI-based solutions to be implemented clinically. In addition, a review article on ML security, privacy, and unity issues in medical imaging is introduced by Georgios A. Kaissis et al. [191]. The review concentrates on medical imaging implementations as well as possible attack vectors and upcoming opportunities in medical imaging.

4.5. Wearable Sensor

Recently, there are many necessary wearable devices in our daily lives, for instance, wearable sensors to measure physiological and mobile phones. This section illustrates the related works of ML types and techniques used in the wearable sensors. At the end of the section, Table 4 summarizes and highlights each type and technique of ML, a brief description of the problem, and the mechanism of collecting the datasets.

4.5.1. Supervised
A. *Deep learning and Artificial Neural Networks*

The spread of IoT healthcare, sport, and other mobile devices greatly contributes in collecting massive amount of data. Many methods have been proposed and investigated to help cope with the rapid growth of the collected data for their efficient utilization and management.

Ravi et al. proposed a method to classify massive amounts of information collected through IoT wearable tools using DL. The proposed method merged properties obtained from data of inertial sensor with information acquired by a collection of shallow properties in order to improve the classification process especially when on-node calculation was involved [192].

In addition, the DL technique was used by Hammerla et al. to track IoT wearable devices in order to recognize and capture human daily activities [193]. Deep, convolutional, and recurrent approaches were investigated and applied on movement datasets that were acquired using wearable devices to capture human movement. A description of recurrent approach training was provided as well.

Eskofier et al. [194] also used the DL technique to utilize the data collected through wearable sensors. The proposed method used the DL technique to observe the movement of patients suffering from Parkinson's disease (PD). The research mainly aimed to study and reveal the capabilities of DL in managing and analyzing such data. Furthermore, Kwon et al. proposed an unsupervised learning approach to detect user movements using data gathered from smart phones. The proposed method came as a response to solve the weaknesses of former methods, such as creating a training dataset and increasing the number of recognized movements [195].

Recently, ANN techniques have witnessed rapid development in pattern recognition as a result of their learning ability to separate nonlinearly separable classes. ANNs can autonomously extract a nonlinear combination of features by learning the complex mappings [196]. Therefore, ANNs have been widely explored as an ideal solution to recognize human activity using wearable sensors.

Mantyjarvi et al. proposed a multilayer perceptron classifier to recognize different human motions [197]. The authors first used the principal component analysis (PCA) and independent component analysis with a wavelet transform for ANN. They then used a multilayer perceptron classifier to recognize the different activities of wearable device user. The experimental results using acceleration sensors demonstrated the proper classification accuracy of the proposed method.

Given that multilayer ANNs can produce complex discriminating surfaces for recognition problems, Yang et al. proposed a systematic framework using ANNs and acceleration data to recognize human activities [196]. The proposed framework first used the common PCA to minimize the feature sets dimension for static and dynamic activities. A neural classifier using a constant threshold criterion was then adopted to differentiate static and dynamic activities. The experimental results proved the strength of the proposed recognition framework in creating competent classifiers with satisfactory accuracy.

Meanwhile, Zhu and Sheng proposed a human daily activity recognition method that combined hidden Markov and ANNs models [198]. First, the data of human daily activities from multiple wearable sensors were fused. Second, the proposed recognition method was applied for the fine-grained classification of the activity type: strong displacement activity, transitional activity, and zero displacement activity. Experimental results proved the validation and accuracy of the proposed approach.

In another work, the same authors proposed a resource-aware and lightweight ANN technique for hand gesture recognition [199]. The proposed method can take into account contextual information as well as sequential limitations between different gestures. It can also designate the start and end locations of a gesture automatically.

Motivated by the potential high performance of the deep convolutional ANNs (CNNs) due to its ability in feature extraction from raw sensor inputs, Jiang and Yin proposed a deep convolutional ANN archi-tecture that can learn the optimum discriminative features for activity recognition [200]. Low-level to high-level features must be learned to achieve efficient activity recognition, the proposed architecture considered a 2D activity image, which is an ensemble of time-series sensor signals of accelerometers and gyroscopes. The proposed architecture achieved great performance in both computation cost and recogni-tion accuracy. By contrast, Ordóñez and Roggen proposed a DL framework that combined convolutional

and long-short-term memory recurrent ANNs [201]. The proposed framework can automatically uncover features attached to the dynamics of human motion production. The experimental results, which relied on realistic wearable sensor data collected from periodic and sporadic activities, showed good results in terms of recognition accuracy and computation time.

Stuntebeck et al. proposed a framework for the automated tagging of health-related events using NNs. The proposed framework collected patient's data and sent them to a server for storage and analysis. Thereafter, the system required users to provide positive classification by asking them about their conditions being examined. The accuracy of the classifier's decisions was improved by using user feedback. Apache Derby and Weka data mining applications were supported on the remote server to archive collected data, feature extraction, and perform statistical analysis [202].

B. *Support Vector Machine*

This section presents studies that use a supervised ML method, called (SVM classifier), which can be used for classification and regression issues.

In [203], Authors provided a biomedical processor that incorporates a CPU with configurable settings for ML accelerators for real-time and low-energy recognition algorithms. The proposed accelerators of ML contained a SVM accelerator (SVMA) and an active-learning data selection accelerator (ALDSA). SVMA can be adapted to several classification algorithms through several energy-scalable kernels. ALDSA was used for online active learning of signal models through improving connection with clinical professionals who are located in different parts of the country. The proposed system minimized the energy consumption of two applications:ECG-based arrhythmia detection and EEG-based seizure detection.

Mannini et al. presented a probabilistic modeling method to classify different pathological gaits using SVM classifier along with the derivative information of HMM. The experimental data were acquired utilizing measuring portions positioned at the shank and waist from two pathological populations: post-stroke participants and Huntington's disease, as well as the elderly group.Huntington's disease and post-stroke subjects as well as the elderly group using measurement parts located at the shank and waist. The authors used HMM likelihoods and features of time-frequency domain with post-processing classification to improve the accuracy of the classification process [204].

Lara and Labrador presented a survey in human activity recognition based on wearable sensors [205]. They evaluated 28 systems in terms of energy consumption, obtrusiveness,recognition accuracy, and flexibility. Moreover, they categorized learning scheme and response time into two levels of taxonomy, namely, supervised and semi-supervised learning methods, and offline and online response times.

A fall detector prototype using feature selection and SVM was proposed by Shan and Yuan [206]. The authors aimed to provide a system that can detect impending falls before they occur. Human motion data were gathered using a wearable microelectromechanical accelerometer. Thereafter, the data were sent to a PC via wireless transceivers for analysis. Feature selection was used to observe the majority discrimination power features and SVM was used for training, resulting in the creation of the classifier, which can further enable the system to detect falls before it occur.

A detection scheme for mental stress using physiological sensors was proposed by Sun et al. [207]. A total of 20 participants were asked to perform three human activities, namely, sitting, standing, and walking. The physiological measurements of every participant using ECG, galvanic skin response (GSR), and accelerometer were collected before they were subjected to mental stressors. The authors applied three classifier algorithms, namely, the decision tree, naive Bayes, and SVM, to define the psychological condition under stressed or relaxed states. To evaluate the influence of activity information on stress inference, the authors then divided the training dataset to two sets. The first set contained features of ECG and GSR, and the second set contained the accelerometer information. The analysis results demonstrated that enhancing mental stress detection in a mobile environment required the use of accelerometer data.

An activity recognition using a wearable acoustic sensor was presented by Yatani and Truong [208]. The authors developed a wearable acoustic sensor to recognize sound activities coming from the user's throat. A total of 12 sound activities were observed, including sitting and deep breath, eating, drinking,

speaking, whistling, and laughing. A total of 10 participants were asked to wear device sensors around their necks to record the throat sounds, and every participant performed 10 samples for each activity. After collecting the sound data, the proposed system analyzed the features for each sound sample and normalized the features across values for all gathered samples. SVM, naive Bayes, and 5-NN were applied for classification. The authors focused on the results produced by the SVM classifier, which were better than the results of other classifiers.

Meanwhile, Suutala et al. proposed a learning technique using SVM classifier to recognize different individual activities. The data set included signals from 13 subjects engaged in 17 different daily activities. The authors collected the features of time domain for every sensor device. The proposed algorithm was based on SVM and combined generative learning with discriminative learning of individual to smooth sequences of time-dependent activity. After recognizing the human activities using VSM, a forward-backward algorithm was applied to train the probability matrix of a global transition. In the end, the final classification was achieved via a Viterbi algorithm that used the probability matrix of transition and the confidence values [209].

Zhang et al.[210] suggested a one-class SVM-based fall detection system for older individuals. The authors used a wearable tri-axial accelerometer to record daily human activities. The SVM model was learned by performing positive falling samples of 12 participants from the younger and older populations. The observed results show the ability of SVM to classify the fall data and daily activities from the testing dataset.

Varkey et al. proposed an algorithm to recognize different human activities and movements in a particular activity using wearable sensors. Three participants were asked to perform six activities (at least eight times) in their preferred ways. The authors trained SVM by various activities so that they can be used for recognizing activities and their movements. Once SVM was trained, the proposed algorithm began recognizing the types of activities or the movement. The proposed algorithm was evaluated across different cases based on the training way of SVM [211].

Patel et al. employed SVM in their proposed method to evaluate the severity of symptoms and motor difficulties from accelerometer data collected from PD patients. Twelve participants with a PD diagnosis, with age range between 46 and 75 years, were asked to complete a series of motor tasks. Five basic feature types were extracted from the accelerometer data for each motor activity task. SVM with three different kernels was implemented, and the results showed that, in terms of misclassification cost value, the third-order polynomial kernel provided satisfactory results in comparison with other kernels [212].

Wearable monitoring system was proposed in [213] to provide early warning of dangerous physiological conditions. The researchers aimed to integrate observations from healthcare staff with data collected via wearable sensors. A one-class SVM formulation was applied to detect the free parameters of the proposed system. SVM was trained with the "normal" dataset in the training stage. It was then used to define the "abnormal" and the "normal" dataset in the validation stage.

An investigation of the current predictive algorithms for activity recognition was offered by Huỹnh et al. [214]. The researchers studied the performance of four methods (K-means, nearest neighbor, SVM, and HMM) for the recognition of high-level activity containing a small-scale set activities. A participant was asked to wear three sensors on different body parts, including wrist, thigh, and hip. Data were collected from the accelerometer while performing various high-level activities. When compared to other methods, the SVM method produced better results in terms of classification accuracy.

A classification study based on SVM for the daily living activities in a smart home was presented by Fleury et al. [215]. Multiple sensor types, including microphones, accelerometers and magnetometers, door contact on appliances, temperature and hygrometry, and location sensors, were placed in a flat. The features for each sensor were extracted and integrated to create a feature vector. A total of 13 participants were asked to perform seven different things that they do on a regular basis. The acquired data were utilized to create the training dataset using the SVM method, and the testing step was completed using the "leave-one-out" cross-validation model.The preliminary results showed good classification rate for polynomial and Gaussian kernels with an adapted parameter.

Huỹnh and Schiele proposed an integrated method that combined the discriminative classifiergenerative model with a discriminative classifier. The generative part used multiple eigenspace approached to find

a low-dimensional representation of sensor data without supervision, whereas the discriminant part used the produced eigenspaces to build features for training the SVM classifier. High recognition rate scales were registered by the offered algorithm although a subset of training dataset was applied for training [216].

Stikic et al. reported a weakly supervised recognition of human actions using wearable sensors devices. [217]. The authors proposed two recognition approaches based on graph structure using SVM classifier and multi-instance learning. The first approach grouped the sensor data to bags-of-activities and labels were provided at the bag level. Thereafter, numerous multi-instance learning extensions were proposed to label activity data. The second approach used graph structures to represent activity data and transfer the provided labels to the unlabeled ones. The presented approaches were applied for training the SVM classifier using the initial and new obtained labels. The researchers employed the TU Darmstadt and PLCouple1 datasets, both of which contained long-term records of daily activities for a person in non-laboratory settings. The results of their proposed approach were close to those of fully supervised techniques.

Tucker et al. proposed a monitoring system to detect and diagnose patient movements to determine compliance with their medication procedures. The proposed method used data mining technique with SVM to classify and analyze the data acquired through nonwearable devices. The main core of the proposed method was to observe patients while practicing daily walking then compare the observation with their routine walking, which helped in inferring useful information regarding the patients' health condition [218].

Sano et al. studied the behavioral and physiological patterns of 66 students in 30 days. The Mental Health Composite Score, the Perceived Stress Scale, and the Pittsburg Sleep Quality Index score are four essential performance variables that affect academic achievement, according to the authors. Wearable sensors and mobile phone data were used to identify personality traits, sleep, mood, and stress, among other things. The best pairings of characteristics were originally seen using sequential forward feature selection. The results of classification techniques illustrated the relationships between the key performances that impact the behavioral choices on GPA [219].

Sano and Picard used ML techniques to recognize the stress markers. The data was gathered via mobile phone usage, a wrist sensor and 18 participant surveys. SVM, SVM with Radial basis function, k-nearest neighbors, PCA and SVM with RBF kernel, PCA and k-nearest neighbors, and PCA and SVM with linear kernel were among the six classifier types employed. Correlation analysis was carried out to observe the features that were statistically significantly associated with stress. The ML methods were then applied to classify the participants into two groups according to their stress postures [220].

C. *Decision Tree*

The elderly are vulnerable to instances of falling, and this has inspired researchers to utilize massive amounts of information acquired by IoT wearable devices to observe the elderly's activities. jetola et al. proposed a human collapse detector depending on the information acquired by healthcare IoT wearable devices. The proposed method applied and utilized ML methods' capabilities along with DT to sense human collapses in four directions [221].

An efficient algorithm was introduced in [222] for the mechanical identification (automated recognition) of natural activities utilizing a wireless heart scale director and five triaxial wireless accelerometers. The proposed method was assessed on several datasets containing 30 natural gymnasium exercises obtained from 21 people at two various labs. The results showed that the proposal obtained an identification accuracy rate of 94.6% and 56.3% utilizing dependent and independent training, consequently. The increasing heart scale data increased the accuracy of dependent and independent identifications by 1.2% and 2.1% respectively. When agreeing to the type of activity without discrimination of intensity levels, an independent performance of 80.6% was obtained.

D. *Gaussian Process*

As IoT has led to the emergence of devices like sensors and E-health devices with the power to communicate through networks, several research studies have been conducted to find ways to manage and utilize the massive information transformation acquired by this technology. A large number of studies have been carried out in order to extract useful information from the massive amounts of data produced by E-health devices and sensors. It is a feasible process when working with clear or clean data, but it is a difficult process when dealing with data uncertainty caused by missing data or device problems. To fulfill this purpose, Gaussian process (GP) [223] framework has been employed by Clifton et al. [224]. They presented a novel patient-personalized method for physiological data analysis and deduction in the case of data uncertainty. Furthermore, 200 patients were observed to demonstrate the practicality of the suggested technique in a real-world clinical setting. For the first time, a systematic, probabilistic approach to patient vital sign monitoring is used in this study, which is established within a GP framework. The GP paradigm provides a logical approach to inferring from noisy, potentially artifactual data collected by sensor devices used in e-health monitoring. The concept of the GP as a distribution over functions, which is well suited to the analysis of time series of patient physiological data, in which the authors execute inference over functions, is of essential relevance. Later in 2018, Leo L. Duan et al.[225] proposed a dual level GP joint hierarchical model in order to enhance personalized medical monitoring data prediction. The model employs three independent GPs: two at the population level and one at the individual level. The two GPs are employed at the population level to obtain nonlinear variations in a continuous biological marker and a binary result. The proposed model is used to analyze compelling medical data on cystic fibrosis progression of the disease, in which lung function measures and the beginning of acute respiratory episodes are both observed at the same time across every patient's clinical course.

Furthermore, Li-Fang Cheng et al. [226] presented MedGP, a statistical framework that combines 24 clinical parameters and provides a large reference data set from which correlations between monitored parameters may be inferred and exploited for high-quality patient state inference across time. To accomplish so, they create a highly structured sparse GP kernel that allows for tractable computation over vast numbers of time points while predicting clinical parameters, patients, and patient observation periodicity. In addition, a GP-based model is proposed by Jean-Rassaire Fouefack et al. [227]. The proposed method, which named Dynamic multi feature-class Gaussian process models (DMFCGPM), is a statistical modeling method for machine learning of several characteristics in medical images.

4.5.2. Unsupervised
A. *Feature selection*

A variety of feature selection algorithms have been used to detect human movements using the information acquired via wearable devices [228, 229]. Zhang and Sawchuk, for example, proposed a method for detecting human movement based on the analysis of data obtained from wearable devices. The proposed approach used feature selection to detect the features indicating types of human motion. Physical features of human movements were set at early stage of the proposed method, and a framework of feature selection was implemented to evaluate the effect of physical feature set [230]. Another method that used feature selection methods to detect human movements was proposed by Lester et al. [231]. The proposed method combined hidden Markov models (HMMs) and boosting. Boosting was used to filter the features and detect human movements, whereas HMM was used to recognize the regularities and harmony of the movements.

Furthermore, Pirttikangas et al. proposed another human movement detector based on Wearable devices placed in different parts of the human body. The proposed method minimizes the amount of features required to detect and recognize movement [232]. In addition, a multiuser movement detection platform was proposed by Wang et al. [233]. The proposed platform used Coupled HMM and Factorial Conditional Random Field to model sensors' data transformation and communication.

Ugulino et al. investigated the literature of human activity detection methods and proposed a classifier for human movements [234]. The proposed classifier used Correlation-based Feature Selection algorithm

as a feature selection method, with the goal of avoiding redundancy and finding the most valuable features.

B. *Hidden Markove Model*

A combination of sensing techniques and computational modeling for behavior recognition was proposed by Brugarolas et al. [235]. The authors offered a canine body-area-network to give trainers improved interpretation for dog training. A wireless sensing system was employed on a canine vest to collect data. The proposed system used decision tree classifiers to detect the static posture of canine, such as standing, and lying down. By contrast, HMMs were used for dynamic activities, like stairs climbing and walking. Various dynamical ML approaches were produced and assessed in [236] to trace the appearance and hardness of shaking and dyskinesia with the 1-s presentation by investigating flags obtained from PD cases affecting small numbers of heterogeneous sensors with 3D accelerometric and cover electromyographic modalities. The proposed approaches were examined on a 44-h signal database constructed from heterogeneous sensors worn by 8 PD cases and 4 healthy controls who conducted daily actions of being in an apartment-like condition. A comparison of the effectiveness of the 4 ML approaches with independent annotations of disease behavior and severity showed that, notwithstanding their differing procedures to effective pattern classification, effective NNs, effective SVM, and HMMs were fairly useful in preventing error rates of the effective tracking well under 10%. A standard set of experimentally determined signal characteristics was utilized to externally reduce the algorithm's demand for particular learning. They also found that error rates under 10% were obtainable when the proposed algorithms were examined on data from a sensor area distinct from those applied in algorithm training.

On-body wearable sensors have been extensively used in various educational and mechanical areas. Of high importance are their treatments in mobile monitoring and pervasive computing methods. In relation to this, the unusual quantitative investigation of human movement and its electronic classification are the foremost computational optimization assignments to have proceeded. A discussion of how human bodily movement can be analyzed was presented in [237]. On-body accelerometers with significant importance applied to the computational optimization algorithms were used for this scheme. In terms of the mentioned problem, the foremost steps regarding sensor choice, data recovery, feature selection, classification, and extraction were analyzed by tracking the diagram of a generic classification scheme with supervised training. As for the ML systems required for analysis, the special importance of MM was highlighted in this work. In particular, the proposed method was triggered in classifiers based on HMMs. An instance was also explained and examined by investigating a dataset of accelerometer time range.

C. *Linear Discriminant and Extreme Learning Machine*

Wearable sensors that typically contain accelerometers and/or gyroscopes describe an encouraging technology for anticipating and decreasing the impacts of falls. At present [238] aimed to address the aforementioned goal of recognizing the appearance of a fall and alert application providers to this performance. Upcoming systems also give information on the reasons and events of falls in order to assist clinical investigation and the targeting of interruptions.The goal of the present research was to improve and test the efficiency of a wearable sensor system for determining the reasons for falls. A total of 16 young people engaged in test trials, including falls due to drops, drugs, and "other" irregular causes. Three-dimensional dispatch data collected during the falling actions were used in the linear discriminant investigation procedure. This system produced 96% sensibility and 98% specificity in identifying the problems of falls by applying acceleration data from three brands. By contrast, a unique marker presented 54% sensitivity and two markers produced 89% sensitivity. These results showed the efficiency of a three-node accelerometer pattern for identifying fall-related problems.

One of the most common causes of injury among the elderly is falling. Utilizing wearable devices for fall exposure has a large cost and may induce difficulty in the day-to-day lives of the elderly. To help in resolving this issue, a mechanical fall detection method was proposed in [239], which only required a low-

cost camera. To recognize falls originating from various daily activities, the proposed method combines two computer vision algorithms based on form fall identification and a learning model classifier. They chose the curvature scale space (CSS) characteristics of human silhouettes at each shape from a fall video clip and recreated the development using a case of CSS words (BoCSS). The classifier was then used to distinguish the BoCSS illustration of a fall from other activities, depending on ELM. A variable-length based on PSO was proposed to identify the appropriate amount of shielded neurons, matching input impacts, and ELM biases in order to reduce the feeling of ELM to its hyperparameters. An activity dataset was created with 6 types of behaviors (bending, walking, sitting, falling, squatting, and lying) from Ten cases using a low-cost Kinect depth camera. The results showed that proposed approach obtained up to 77.14% specificity, 91.15% sensitivity, and 86.83% accuracy. On a general dataset, the proposed approach showed comparable results to other well-known approaches of fall detection requiring multiple cameras.

D. *Clustering Applications*

In recent years, ML algorithms have played an important role in DNA sequencing to resolve various biomedical and healthcare applications issues [240, 241, 242]. Such algorithms mainly aim to distinguish the genetic disorder. Accordingly, Manogaran et al. proposed a hybrid approach that includes the Bayesian Hidden Markov (BHM) model with Gaussian Mixture (GM) clustering algorithms to diagnose cancer diseases by modeling the DNA copy number change throughout the genome. The proposed approach was compared with various existing approaches to demonstrate its effectiveness [102].

[132] aimed to conclude whether holding element functional magnetic resonance imaging (fMRI) taken at the original pretreatment could perfectly determine character cancer. A total of 31 cases were tested with breast cancer (aged 34-65 years) earlier to any therapy, post-chemotherapy, and one-year succeeding. Cognitive examination rates were normalized based on data obtained from 43 normal woman-like instruments and then worked to classify cases as injured or not injured based on longitudinal variations. They included the clustering technique, a measure of insufficient connectivity, by utilizing design theory to original resting state fMRI and recorded these measures along with significant patient-related and medical variables into irregular forest analysis. Using these methods, the authors proposed that resting-state fMRI is an encouraging tool for divining expected cognitive impairment linked with breast cancer. This knowledge could identify treatment decision-making by recognizing cases at the highest risk for cognitive impairment.

Rahman et al. utilized ML and statistical similarity matching techniques for introducing a medical image retrieval framework [243]. The proposed framework included the following: (i) ML, supervised and unsupervised learning techniques, such as fuzzy c-mean clustering and SVM for image pre-filtering and classification, and (ii) statistical distance measures and a relevance feedback (RF) for similarity matching, Where statistical distance measures are used to integrate better perceptive subjectivity and RF to dynamically update query parameters. The simulation results revealed that the proposed framework has improved in terms of efficiency and effectiveness.

Wang et al. argued that the majority of the existing classification techniques lacked incremental learning ability because the monitoring of human activities, in many real-life scenarios, may last for a long time [244]. In such a case, the reliability and accuracy of the classifiers are far from convincing as they are only trained once from sensor data gathered over several days or weeks. In order to address the practical needs of sensor-based human activity recognition, the authors proposed an incremental learning technique. The proposed technique was designed based on an adjustable fuzzy clustering algorithm and probabilistic ANNs. it can simply learn more information from new training data to improve accuracy of recognition process. It also has the ability to freely remove or add new activities to be detected. In addition, the process of updating the presented technique does not need previous training data. The proposed method was based on realistic wearable sensor data been collected from various daily life activities, and the experimental results indicated a reasonable ability to incrementally learn, in addition to the high precision of recognition.

Table 4: Summary of using ML techniques in wearable sensors

Type	Technique	Description	Data set	Author(Ref)
Supervised	DL and ANN	Classify the massive amount of information collected through IoT wearable tools	Raw signal First derivative	Ravi et al. ([192])
		Track IoT wearable devices to recognize human daily activities	Opp, PAMAP2 and DG	Hammerla et al. ([193])
		Observe the movement of patients who suffer from Parkinson's syndrome	IMU	Eskofier et al. ([194])
		Solve the weaknesses of former methods for increasing the number of recognized user movements	smartphone*	Kwon et al. ([195])
		Extracting a non-linear combination of features by learning the complex mappings	MMA7260Q*	Yang et al. ([196])
		Classification the recognition of different human motion	ADXL202*	Mantyjarvi et al. ([197])
		Classification the activity type: strong displacement, transitional, and zero displacement activities	IMU	Zhu and Sheng ([198])
		Focusing on the hand gesture recognition and daily activity recognition	IMU	Zhu and Sheng ([199])
		Meet the practical requirements of sensor-based human activity recognition and improve its accuracy	S-HAR	Wang et al. ([244])
		Automatically identify the most effective discriminative characteristics for activity recognition, automatically	UC1, USC and SHO	Jiang and Yin ([200])
		Uncovering the features that are attached to the dynamics of human motion production, automatically	OPPORTUNITY	Ordóñez and Roggen ([201])
	SVM	Used for online active learning of signal models through enhancing the communications with remote clinical experts	CHB-MIT and MIT-BIH	Lee and Verma ([203])
		Classify different pathological gaits using SVM classifier along with the derivative information of HMM	IMU	Mannini et al. ([204])
		Evaluated twenty-eight systems in terms of obtrusiveness, recognition accuracy, flexibility, and energy consumption	Accelerometer*	Lara and Labrador ([205])
		Provide a system that is able to detect the impending falls before it occurs	NEC 78K0547*	Shan and Yuan ([206])
		Define the psychological condition to stressed or relaxed states and evaluate the influence of activity information on stress inference	Accelerometer*	Sun et al. ([207])
		Analyzed and classified the features for each sound sample, and normalize the features across values for all gathered samples	Wearable Acoustic*	Yatani and Truong ([208])
		Recognize different individual activities and to smooth temporal time-dependent activity sequences	Wearable acceleration*	Suutala et al. ([209])
		Performing positive falling samples of twelve participants of younger and elder	MMA7260Q*	Zhang et al. ([210])
		Recognize different human activities and movements in a particular activity using wearable sensors	Gyroscopes*	Varkey et al.([211])
		Evaluate the severity of symptoms and motor difficulties from accelerometer data for Parkinson disease patients	wearable sensor*	Patel et al. ([212])
		Determine the free parameters of the early warning of serious physiological determination	wearable sensor*	Clifton et al. ([213])
		Recognition of high-level activity that contained a set of small scale activities	wearable sensor*	Huỹnh et al. ([214])
		Identify daily live activities in a smart home and used the features for each sensor were extracted and put together to create a feature vector	Smart Home's TU sensors*	Fleury et al. ([215])
		Construct a low-dimensional model of sensors data without supervision, whereas the discriminant part uses the produced eigenspaces to build features	IPAQ*	Huỹnh and Schiele ([216])
		Apply two recognition approaches based on multi-instance learning and graph structure	PLCouple1 and TU Darmstadt	Stikic et al. ([217])
		Recognize several factors including personality traits, sleep, mood, and stress and compare how these features performed	flight sensors*	Tucker et al. ([218])
		Detect and diagnose patient movements in order to check whether they are following their medication procedure or not	smartphone*	Sano et al. ([219])
		Observe significant features associated with stress and classify the participants into two groups according to their stress posture	smartphone*	Sano and Picard ([220])
	DT	Using the massive amount of information to observe old people activities	SHIMMER	Ojetola et al. ([221])
		Recognition the natural activities in some instances, their strengths, using wireless heart scale director	wearable accelerometers*	Tapia et al. ([222])
	GP	Managing and using the massive information transformation for observing the patients' health condition	wearable sensor*	Clifton et al. ([224])
Unsupervised	Feature selection	Detecting the features which indicate to a human motion	IMU	Zhang and Sawchuk ([230])
		Filtering the features, detecting human movements, and recognizing its regularities	wearable sensor*	Lester et al. ([231])
		Reduce the number of features that required to recognize and detect the movement	wearable accelerometers*	Pirttikangas et al. ([232])
		Using a new platform to model sensors' data transformation and communication	ITS400*	Wang et al. ([233])
		Minimize features redundancy and determining the most useful features	wearable accelerometers*	Ugulino et al. ([234])
	HMM	Detect the static posture of canine such as standing and lying down, and dynamic activities, like stairs climbing and walking	inertial sensors*	Brugarolas et al. ([235])
		Trace the appearance and hardness of shaking and dyskinesia and investigating flags obtained from PD cases	EMG*	Cole et al. ([236])
		Analyze feature selection, classification, and extraction by tracking the diagram of a generic classification scheme with supervised training	MEMS*	Mannini and Sabatini ([237])
	ELM	Classify and recognize falls from different daily activities	SDUFall	Ma et al. ([239])

*The authors didn't use the traditional data set, they collected the data using different devices (As mentioned above)

4.5.3. Other ML Techniques

This section presents different studies that analyze the performance of classification methods for detecting human daily living activities.

Akhavian and Behzadan proposed a system to explore the prospect of using mobile phone sensors and transferring nodes to identify detailed activities of construction equipment [245]. The authors employed data logger apps for Android and iOS devices to give time-stamped data, making it easier to record data and video in synchronization. The fast Fourier transform was used in feature extraction to convert time domain signals to frequency domain signals, whereas ReliefF and Correlation-based Feature Selection were used in the feature selection process. Moreover, the authors applied five classification methods, namely, logistic regression, k-nearest neighbor, SVM, NN, and decision tree, to recognize the action of equipment.

A detection fall system using ML methods was proposed by Özdemir and Barshan. The proposed system required fitting motion sensors (accelerometer, gyroscope, and compass) on six positions on the body. The raw data used in this research contained 2,520 activities of daily living trials. The authors applied six classifier methods to recognize falls from daily living activities, including the least-squares method, k-nearest neighbor, Bayesian decision-making, SVM, ANNs, and dynamic time warping [246].

Zhang and Sawchuk proposed a framework to recognize daily human activity via wearable sensors [247]. Six equally divided axes of accelerometer and gyroscope were used for sensing nine types of human activities, such as forward walking, going upstairs, jumping up, and sitting. The study contained 110 input feature spaces extracted from the accelerometer and gyroscope, such as average velocity, skewness, and mean crossing rate, to name a few. The proposed sparse representation-based approach achieved better performance than other methods, including NN, SVM, and naive Bayesian classifier, when the feature dimensions were more significant than or equal to 40.

Attal et al. studied classification techniques used to identify human activities in daily life using three wearable sensors placed at three points on the body: the right thigh, left ankle, and chest. The process of identifying human activity is comprised of three steps, including sensor placement, data pre-processing, and classification. The study compared the performance of four supervised classification methods and three unsupervised classification methods to recognize 12 activities, including transition, dynamic, and static activities. The comparative experiments showed that the supervised methods provided more accurate results than the unsupervised methods when applying extracted traits or source data. However, the unsupervised methods were more computationally efficient than the supervised methods and did not need any label. The experimental results demonstrate that the k-nearest neighbor and HMM classifiers achieved the best performance among other supervised and unsupervised classification methods, respectively [248].

A comprehensive study on the classification methods for human activities was performed by Barshan and Yüksek [249]. Three orientation tracker units were used. Each has a triaxial accelerometer, gyroscope, and magnetometer. The data sensors were positioned at five places on the body: both knees' sides, chest, and wrists. The authors classified 19 body activities, including lying, standing, sitting, ascending, and descending stairs. Eight participants were asked to perform the classified activity in their way. The PCA classification mechanism was applied to minimize the feature set extracted from the sensors from 1,170 to 30 features. Comparisons among eight classifier methods were performed in terms of rates of correct differentiation, computational cost, and confusion matrices. Other beneficial swarm intelligence techniques can be used in this domain refer to [250, 251, 252, 253, 254].

5. Conclusion and Future works

This paper focuses in depth on the application of machine learning in medical diagnosis and covers recent developments. Therefore, more than 200 research publications were gathered, studied, and analyzed, from January 2000 to December 2022, to highlight the strengths, limitations, benefits, and drawbacks of employing machine learning applications for various medical challenges. According to the raised challenges studied in this review, we provide potential research possibilities and future initiatives in the field as follows:

- The majority of studies used supervised ML for cancer detection because it is sufficient for classifying future inputs [91, 92, 100]. However, there are some limitations, such as the possibility of receiving

incorrect class labels after classification [255], and the challenge of big data classification [111, 113]. Thus, feature selection should be included to aid in the improvement of these challenge issues [256, 113].

- ML algorithms are employed to detect the onset of some brain issues, such as trauma and epileptic seizures. To improve ML performance, AI techniques can be employed in conjunction with ML techniques to aid in adequately tracking regions of interest [257].

- To improve the diversity of component classifiers, we recommend using fuzzy classifiers in medical diagnosis, which have not yet been utilized before [258].

- Due to a large number of medical images, such as in [259], we strongly recommend adopting unsupervised learning approaches due to their ability to acquire and retain massive amounts of data [260, 261].

- Finally, the massive amount of data acquired by wearable sensors necessitates an effective method for extracting the most relevant data and dealing with massive online data processing requirements quickly and efficiently.

Overall, this survey is intended to serve as a resource for researchers and industry professionals, and decision-makers in their potential research and application projects in the coming years. This survey also supports future works by presenting the most comment researches in the medical field using the learning methods. Finally, creating a website that includes up-to-date publications is the future work to avoid outdated quickly and keep abreast of scientific developments.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] I. Kononenko, Machine learning for medical diagnosis: history, state of the art and perspective, Artificial Intelligence in medicine 23 (1) (2001) 89–109.

[2] M. Gharaibeh, D. Alzu'bi, M. Abdullah, I. Hmeidi, M. R. Al Nasar, L. Abualigah, A. H. Gandomi, Radiology imaging scans for early diagnosis of kidney tumors: A review of data analytics-based machine learning and deep learning approaches, Big Data and Cognitive Computing 6 (1) (2022) 29.

[3] K. Y. Bilimoria, A. K. Stewart, D. P. Winchester, C. Y. Ko, The national cancer data base: a powerful initiative to improve cancer care in the united states, Annals of surgical oncology 15 (3) (2008) 683–690.

[4] F. Ahmad, et al., An fmri study of neural decision making using support vector machine, in: Proceedings of Business and Management Conferences, no. 6809558, International Institute of Social and Economic Sciences, 2018.

[5] B. Manavalan, T. H. Shin, G. Lee, Pvp-svm: sequence-based prediction of phage virion proteins using a support vector machine, Frontiers in microbiology 9 (2018) 476.

[6] A. Jaafari, H. R. Pourghasemi, Factors influencing regional-scale wildfire probability in iran: An application of random forest and support vector machine, in: Spatial Modeling in GIS and R for Earth and Environmental Sciences, Elsevier, 2019, pp. 607–619.

[7] I. Aljarah, A. Al-Zoubi, H. Faris, M. A. Hassonah, S. Mirjalili, H. Saadeh, Simultaneous feature selection and support vector machine optimization using the grasshopper optimization algorithm, Cognitive Computation 10 (3) (2018) 478–495.

[8] C.-z. Gao, Q. Cheng, P. He, W. Susilo, J. Li, Privacy-preserving naive bayes classifiers secure against the substitution-then-comparison attack, Information Sciences 444 (2018) 72–88.

[9] A. Wood, V. Shpilrain, K. Najarian, D. Kahrobaei, Private naive bayes classification of personal biomedical data: Application in cancer data analysis, Computers in biology and medicine 105 (2019) 144–150.

[10] W. Jiang, Y. Shen, Y. Ding, C. Ye, Y. Zheng, P. Zhao, L. Liu, Z. Tong, L. Zhou, S. Sun, et al., A naive bayes algorithm for tissue origin diagnosis (tod-bayes) of synchronous multifocal tumors in the hepatobiliary and pancreatic system, International journal of cancer 142 (2) (2018) 357–368.

[11] S. Xu, Bayesian naïve bayes classifiers to text classification, Journal of Information Science 44 (1) (2018) 48–59.

[12] G. Carleo, M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science 355 (6325) (2017) 602–606.

[13] M. Van Gerven, S. Bohte, Editorial: Artificial neural networks as models of neural information processing, Artificial Neural Networks as Models of Neural Information Processing 5.

[14] G. Villarrubia, J. F. De Paz, P. Chamoso, F. De la Prieta, Artificial neural networks used in optimization problems, Neurocomputing 272 (2018) 10–16.

[15] Z. Zhang, Artificial neural network, in: Multivariate Time Series Analysis in Climate and Environmental Research, Springer, 2018, pp. 1–35.

[16] P. Thanh Noi, M. Kappas, Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery, Sensors 18 (1) (2018) 18.

[17] A.-J. Gallego, J. Calvo-Zaragoza, J. J. Valero-Mas, J. R. Rico-Juan, Clustering-based k-nearest neighbor classification for large-scale data with neural codes representation, Pattern Recognition 74 (2018) 531–543.

[18] D. Adeniyi, Z. Wei, Y. Yongquan, Automated web usage data mining and recommendation system using k-nearest neighbor (knn) classification method, Applied Computing and Informatics 12 (1) (2016) 90–108.

[19] N. García-Pedrajas, J. A. R. del Castillo, G. Cerruela-García, A proposal for local k values for k-nearest neighbor rule, IEEE transactions on neural networks and learning systems 28 (2) (2017) 470–475.

[20] H. Parvin, M. MirnabiBaboli, H. Alinejad-Rokny, Proposing a classifier ensemble framework based on classifier selection and decision tree, Engineering Applications of Artificial Intelligence 37 (2015) 34–42.

[21] J. Tanha, M. van Someren, H. Afsarmanesh, Semi-supervised self-training for decision tree classifiers, International Journal of Machine Learning and Cybernetics 8 (1) (2017) 355–370.

[22] Q. He, Z. Xu, S. Li, R. Li, S. Zhang, N. Wang, B. T. Pham, W. Chen, Novel entropy and rotation forest-based credal decision tree classifier for landslide susceptibility modeling, Entropy 21 (2) (2019) 106.

[23] S. A. Parikh, R. Gomez, M. Thirugnanasambandam, S. S. Chauhan, V. De Oliveira, S. C. Muluk, M. K. Eskandari, E. A. Finol, Decision tree based classification of abdominal aortic aneurysms using geometry quantification measures, Annals of biomedical engineering 46 (12) (2018) 2135–2147.

[24] M. Sornam, M. Prabhakaran, Logit-based artificial bee colony optimization (lb-abc) approach for dental caries classification using a back propagation neural network, in: Integrated Intelligent Computing, Communication and Security, Springer, 2019, pp. 79–91.

[25] A. Gautam, V. Bhateja, A. Tiwari, S. C. Satapathy, An improved mammogram classification approach using back propagation neural network, in: Data Engineering and Intelligent Computing, Springer, 2018, pp. 369–376.

[26] M. Monisha, A. Suresh, B. T. Bapu, M. Rashmi, Classification of malignant melanoma and benign skin lesion by using back propagation neural network and abcd rule, Cluster Computing (2018) 1–11.

[27] S. Samanta, S. S. Ahmed, M. A.-M. M. Salem, S. S. Nath, N. Dey, S. S. Chowdhury, Haralick features based automated glaucoma classification using back propagation neural network, in: Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014, Springer, 2015, pp. 351–358.

[28] J.-Y. Hsia, C.-J. Lin, Parameter selection for linear support vector regression.

[29] M. Zahid, F. Ahmed, N. Javaid, R. A. Abbasi, Z. Kazmi, H. Syeda, A. Javaid, M. Bilal, M. Akbar, M. Ilahi, Electricity price and load forecasting using enhanced convolutional neural network and enhanced support vector regression in smart grids, Electronics 8 (2) (2019) 122.

[30] T. Benkedjouh, K. Medjaher, N. Zerhouni, S. Rechak, Health assessment and life prediction of cutting tools based on support vector regression, Journal of Intelligent Manufacturing 26 (2) (2015) 213–223.

[31] A. Soualhi, K. Medjaher, N. Zerhouni, Bearing health monitoring based on hilbert–huang transform, support vector machine, and regression, IEEE Transactions on Instrumentation and Measurement 64 (1) (2015) 52–62.

[32] A. K. Yadav, S. Chandel, Identification of relevant input variables for prediction of 1-minute time-step photovoltaic module power using artificial neural network and multiple linear regression models, Renewable and Sustainable Energy Reviews 77 (2017) 955–969.

[33] F. Khademi, S. M. Jamal, N. Deshpande, S. Londhe, Predicting strength of recycled aggregate concrete using artificial neural network, adaptive neuro-fuzzy inference system and multiple linear regression, International Journal of Sustainable Built Environment 5 (2) (2016) 355–369.

[34] S. Johansen, B. Nielsen, Tightness of m-estimators for multiple linear regression in time for multiple linear regression in time series, Univ. of Copenhagen Dept. of Economics Discussion Paper (16-05).

[35] L. Melie-Garcia, B. Draganski, J. Ashburner, F. Kherif, Multiple linear regression: Bayesian inference for distributed and big data in the medical informatics platform of the human brain project, BioRxiv (2018) 242883.

[36] H. Liu, Z. Ma, J. Han, Z. Chen, Z. Zheng, Regularized partial least squares for multi-label learning, International Journal of Machine Learning and Cybernetics 9 (2) (2018) 335–346.

[37] W. Song, H. Wang, P. Maguire, O. Nibouche, Nearest clusters based partial least squares discriminant analysis for the classification of spectral data, Analytica chimica acta 1009 (2018) 27–38.

[38] W. Wang, D. C. Ackland, J. A. McClelland, K. E. Webster, S. Halgamuge, Assessment of gait characteristics in total knee arthroplasty patients using a hierarchical partial least squares method, IEEE journal of biomedical and health informatics 22 (1) (2018) 205–214.

[39] S. Ge, R. Wang, Y. Leng, H. Wang, P. Lin, K. Iramina, A double-partial least-squares model for the detection of

steady-state visual evoked potentials, IEEE Journal of Biomedical and Health informatics 21 (4) (2017) 897–903.

[40] Y. Marom, D. Feldman, k-means clustering of lines for big data, arXiv preprint arXiv:1903.06904.

[41] M. A. Z. Alvarez, K. Agbossou, A. Cardenas, S. Kelouwani, L. Boulon, Demand response strategy applied to residential electric water heaters using dynamic programming and k-means clustering, IEEE Transactions on Sustainable Energy.

[42] E. Dickinson, M. J. Rusilowicz, M. Dickinson, A. J. Charlton, U. Bechtold, P. M. Mullineaux, J. Wilson, Integrating transcriptomic techniques and k-means clustering in metabolomics to identify markers of abiotic and biotic stress in medicago truncatula, Metabolomics 14 (10) (2018) 126.

[43] L. Zeinalkhani, A. A. Jamaat, K. Rostami, Diagnosis of brain tumor using combination of k-means clustering and genetic algorithm, Iranian Journal of Medical Informatics 7 (2018) 6.

[44] X. Zhang, Pythagorean fuzzy clustering analysis: a hierarchical clustering algorithm with the ratio index-based ranking methods, International Journal of Intelligent Systems 33 (9) (2018) 1798–1822.

[45] V. Cohen-Addad, V. Kanade, F. Mallmann-Trenn, C. Mathieu, Hierarchical clustering: Objective functions and algorithms, in: Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2018, pp. 378–397.

[46] X. Zhang, Z. Xu, Hesitant fuzzy agglomerative hierarchical clustering algorithms, International journal of systems Science 46 (3) (2015) 562–576.

[47] A. Bouguettaya, Q. Yu, X. Liu, X. Zhou, A. Song, Efficient agglomerative hierarchical clustering, Expert Systems with Applications 42 (5) (2015) 2785–2797.

[48] K. Matsuo, M. Yamada, K. Bylykbashi, M. Cuka, Y. Liu, L. Barolli, Implementation of an iot-based e-learning testbed: Performance evaluation using mean-shift clustering approach considering four types of brainwaves, in: 2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA), IEEE, 2018, pp. 203–209.

[49] N. Kim, S. Park, J. Lee, J. Choi, Load profile extraction by mean-shift clustering with sample pearson correlation coefficient distance, Energies 11 (9) (2018) 2397.

[50] W. Chen, X. Hu, W. Chen, Y. Hong, M. Yang, Airborne lidar remote sensing for individual tree forest inventory using trunk detection-aided mean shift clustering techniques, Remote Sensing 10 (7) (2018) 1078.

[51] Y. Guo, A. Şengür, Y. Akbulut, A. Shipley, An effective color image segmentation approach using neutrosophic adaptive mean shift clustering, Measurement 119 (2018) 28–40.

[52] M. M. R. Khan, M. A. B. Siddique, R. B. Arif, M. R. Oishe, Adbscan: Adaptive density-based spatial clustering of applications with noise for identifying clusters with varying densities, in: 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), IEEE, 2018, pp. 107–111.

[53] C. Peng, S. Guoyou, L. Shuang, Y. Jian, An incremental density based spatial clustering of application with noise algorithm based on partition index, Journal of Computational and Theoretical Nanoscience 13 (12) (2016) 10273–10280.

[54] W. Zhang, X. Zhang, J. Zhao, Y. Qiang, Q. Tian, X. Tang, A segmentation method for lung nodule image sequences based on superpixels and density-based spatial clustering of applications with noise, PloS one 12 (9) (2017) e0184290.

[55] M. Wen, S. Cho, J. Chae, Y. Sung, K. Cho, Range image-based density-based spatial clustering of application with noise clustering method of three-dimensional point clouds, International Journal of Advanced Robotic Systems 15 (2) (2018) 1729881418762302.

[56] D. Mladenić, Feature selection for dimensionality reduction, in: International Statistical and Optimization Perspectives Workshop" Subspace, Latent Structure and Feature Selection", Springer, 2005, pp. 84–102.

[57] M. Masaeli, J. G. Dy, G. M. Fung, From transformation-based dimensionality reduction to feature selection, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 751–758.

[58] H.-L. Wei, S. A. Billings, Feature subset selection and ranking for data dimensionality reduction, IEEE transactions on pattern analysis and machine intelligence 29 (1) (2007) 162–166.

[59] P. E. Rauber, R. R. O. d. Silva, S. Feringa, M. E. Celebi, A. X. Falcão, A. C. Telea, Interactive image feature selection aided by dimensionality reduction, Proc. EuroVA (2015) 54–61.

[60] B. Mojaradi, H. Abrishami-Moghaddam, M. J. V. Zoej, R. P. Duin, Dimensionality reduction of hyperspectral data via spectral feature extraction, IEEE Transactions on Geoscience and Remote Sensing 47 (7) (2009) 2091–2105.

[61] X. Wang, K. K. Paliwal, Feature extraction and dimensionality reduction algorithms and their applications in vowel recognition, Pattern recognition 36 (10) (2003) 2429–2439.

[62] S. De Backer, A. Naud, P. Scheunders, Non-linear dimensionality reduction techniques for unsupervised feature extraction, Pattern Recognition Letters 19 (8) (1998) 711–720.

[63] L. M. Bruce, C. H. Koger, J. Li, Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction, IEEE Transactions on geoscience and remote sensing 40 (10) (2002) 2331–2338.

[64] C. J. C. H. Watkins, P. Dayan, Q-learning, Machine Learning 8 (3) (1992) 279–292.

[65] C. J. Watkins, P. Dayan, Q-learning, Machine learning 8 (3-4) (1992) 279–292.

[66] V. Mnih, A. Puigdomènech Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, K. Kavukcuoglu, Asynchronous Methods for Deep Reinforcement Learning, arXiv e-prints.

[67] M. L. Littman, Markov games as a framework for multi-agent reinforcement learning, Morgan Kaufmann, 1994, pp. 157 – 163.

[68] X. Gu, S. Ghiassian, R. S. Sutton, Should all temporal difference learning use emphasis, arXiv preprint arXiv:1903.00194.

[69] H. Van Seijen, A. R. Mahmood, P. M. Pilarski, M. C. Machado, R. S. Sutton, True online temporal-difference learning, The Journal of Machine Learning Research 17 (1) (2016) 5057–5096.

[70] C. Keerthisinghe, G. Verbič, A. C. Chapman, A fast technique for smart home management: Adp with temporal difference learning, IEEE Transactions on smart grid 9 (4) (2018) 3291–3303.

[71] V. Pong, S. Gu, M. Dalal, S. Levine, Temporal difference models: Model-free deep rl for model-based control, arXiv preprint arXiv:1802.09081.

[72] P. Poupart, N. Vlassis, J. Hoey, K. Regan, An analytic solution to discrete bayesian reinforcement learning, in: Proceedings of the 23rd international conference on Machine learning, ACM, 2006, pp. 697–704.

[73] F. L. Lewis, D. Vrabie, Reinforcement learning and adaptive dynamic programming for feedback control, IEEE circuits and systems magazine 9 (3) (2009) 32–50.

[74] M. L. Littman, Markov games as a framework for multi-agent reinforcement learning, in: Machine learning proceedings 1994, Elsevier, 1994, pp. 157–163.

[75] L. P. Kaelbling, M. L. Littman, A. W. Moore, Reinforcement learning: A survey, Journal of artificial intelligence research 4 (1996) 237–285.

[76] M. van Otterlo, M. Wiering, Reinforcement learning and markov decision processes, in: Reinforcement Learning, Springer, 2012, pp. 3–42.

[77] T. Degris, O. Sigaud, P.-H. Wuillemin, Learning the structure of factored markov decision processes in reinforcement learning problems, in: Proceedings of the 23rd international conference on Machine learning, ACM, 2006, pp. 257–264.

[78] T. Jaakkola, S. P. Singh, M. I. Jordan, Reinforcement learning algorithm for partially observable markov decision problems, in: Advances in neural information processing systems, 1995, pp. 345–352.

[79] G. Baird, (not) throwing the game-an application of markov decision processes and reinforcement learning to optimising darts strategy.

[80] M. Li, Z.-H. Zhou, Improve computer-aided diagnosis with machine learning techniques using undiagnosed samples, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 37 (6) (2007) 1088–1098.

[81] A. Choudhury, D. Gupta, A survey on medical diagnosis of diabetes using machine learning techniques, in: Recent Developments in Machine Learning and Data Analytics, Springer, 2019, pp. 67–78.

[82] Z. Obermeyer, E. J. Emanuel, Predicting the future big data, machine learning, and clinical medicine, The New England journal of medicine 375 (13) (2016) 1216.

[83] L. Wei, Y. Yang, R. M. Nishikawa, Y. Jiang, A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications, IEEE transactions on medical imaging 24 (3) (2005) 371–380.

[84] A. H. Shoeb, J. V. Guttag, Application of machine learning to epileptic seizure detection, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 975–982.

[85] Q.-H. Ye, L.-X. Qin, M. Forgues, P. He, J. W. Kim, A. C. Peng, R. Simon, Y. Li, A. I. Robles, Y. Chen, et al., Predicting hepatitis b virus–positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning, Nature medicine 9 (4) (2003) 416.

[86] H. A. Guvenir, B. Acar, G. Demiroz, A. Cekin, A supervised machine learning algorithm for arrhythmia analysis, in: Computers in Cardiology 1997, IEEE, 1997, pp. 433–436.

[87] J. Wu, J. Roy, W. F. Stewart, Prediction modeling using ehr data: challenges, strategies, and a comparison of machine learning approaches, Medical care (2010) S106–S113.

[88] B. Pogorelc, Z. Bosnić, M. Gams, Automatic recognition of gait-related health problems in the elderly using machine learning, Multimedia Tools and Applications 58 (2) (2012) 333–354.

[89] M. Gharaibeh, M. Almahmoud, M. Ali, A. Al-Badarneh, M. El-Heis, L. Abualigah, M. Altalhi, A. Alaiad, A. H. Gandomi, Early diagnosis of alzheimer's disease using cerebral catheter angiogram neuroimaging: A novel model based on deep learning approaches, Big Data and Cognitive Computing 6 (1) (2022) 2.

[90] A. H. Gandomi, F. Chen, L. Abualigah, Machine learning technologies for big data analytics (2022).

[91] J. M. Wozniak, R. Jain, P. Balaprakash, J. Ozik, N. T. Collier, J. Bauer, F. Xia, T. Brettin, R. Stevens, J. Mohd-Yusof, C. G. Cardona, B. V. Essen, M. Baughman, Candle/supervisor: a workflow framework for machine learning applied to cancer research, BMC Bioinformatics 19 (18) (2018) 491.

[92] R. Fakoor, F. Ladhak, A. Nazi, M. Huber, Using deep learning to enhance cancer diagnosis and classification, in: Proceedings of the International Conference on Machine Learning, Vol. 28, ACM New York, USA, 2013.

[93] P. Shi, S. Ray, Q. Zhu, M. A. Kon, Top scoring pairs for feature selection in machine learning and applications to cancer outcome prediction, BMC Bioinformatics 12 (1) (2011) 375.

[94] M. P. Menden, F. Iorio, M. Garnett, U. McDermott, C. H. Benes, P. J. Ballester, J. Saez-Rodriguez, Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties, PLoS one 8 (4) (2013) 1–7.

[95] N. Borisov, V. Tkachev, M. Suntsova, O. Kovalchuk, A. Zhavoronkov, I. Muchnik, A. Buzdin, A method of gene expression data transfer from cell lines to cancer patients for machine-learning prediction of drug efficiency, Cell Cycle 17 (4) (2018) 486–491.

[96] C. Huang, R. Mezencev, J. F. McDonald, F. Vannberg, Open source machine-learning algorithms for the prediction of optimal cancer drug therapies, PLoS one 12 (10) (2017) 1–14.

[97] G. P. Way, F. Sanchez-Vega, K. La, J. Armenia, W. K. Chatila, A. Luna, C. Sander, A. D. Cherniack, M. Mina, G. Ciriello, Machine learning detects pan-cancer ras pathway activation in the cancer genome atlas, Cell reports 23 (1) (2018) 172–180.

[98] W. Du, K. Dickinson, C. A. Johnson, L. N. Saligan, Identifying genes to predict cancer radiotherapy-related fatigue with machine-learning methods, in: Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, ACM, 2018, pp. 527–527.

[99] P. Gupta, D. A. K. Malhi, Using deep learning to enhance head and neck cancer diagnosis and classification, in: 2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA), IEEE, pp. 1–6.

[100] T. A. Shaikh, R. Ali, Applying Machine Learning Algorithms for Early Diagnosis and Prediction of Breast Cancer Risk, Vol. 46, Springer Singapore, 2019, pp. 589–598.

31

[101] J. Pati, Gene expression analysis for early lung cancer prediction using machine learning techniques: An eco-genomics approach, IEEE Access 7 (2019) 4232–4238.

[102] G. Manogaran, V. Vijayakumar, R. Varatharajan, P. Malarvizhi Kumar, R. Sundarasekar, C.-H. Hsu, Machine learning based big data processing framework for cancer diagnosis using hidden markov model and gm clustering, Wireless Personal Communications 102 (3) (2018) 2099–2116.

[103] M. D. Podolsky, A. A. Barchuk, V. I. Kuznetcov, N. F. Gusarova, V. S. Gaidukov, S. A. Tarakanov, Evaluation of machine learning algorithm utilization for lung cancer classification based on gene expression levels, Asian Pacific Journal of Cancer Prevention 17 (2) (2016) 835–838.

[104] M. Hilario, A. Kalousis, M. Müller, C. Pellegrini, Machine learning approaches to lung cancer prediction from mass spectra, Proteomics 3 (9) (2003) 1716–1719.

[105] G. Valdes, T. D. Solberg, M. Heskel, L. Ungar, C. B. Simone II, Using machine learning to predict radiation pneumonitis in patients with stage i non-small cell lung cancer treated with stereotactic body radiation therapy, Physics in Medicine & Biology 61 (16) (2016) 6105.

[106] M. Ghoncheh, Z. Pournamdar, H. Salehiniya, Incidence and mortality and epidemiology of breast cancer in the world, Asian Pac J Cancer Prev 17 (S3) (2016) 43–6.

[107] R. L. Siegel, K. D. Miller, A. Jemal, Cancer statistics, 2017, CA: a cancer journal for clinicians 67 (1) (2017) 7–30.

[108] A. Gonzales, M. Alzaatreh, M. Mari, A. A. Saleh, A. Alloubani, Beliefs and behavior of saudi women in the university of tabuk toward breast self examination practice, Asian Pacific journal of cancer prevention: APJCP 19 (1) (2018) 121.

[109] A. D. Trister, D. S. Buist, C. I. Lee, Will machine learning tip the balance in breast cancer screening, JAMA oncology 3 (11) (2017) 1463–1464.

[110] M. F. Akay, Support vector machines combined with feature selection for breast cancer diagnosis, Expert systems with applications 36 (2) (2009) 3240–3247.

[111] A. Abou Tabl, A. Alkhateeb, W. ElMaraghy, A. Ngom, Machine learning model for identifying gene biomarkers for breast cancer treatment survival, in: Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, ACM, 2017, pp. 607–607.

[112] T. Haamann, D. Basten, The role of information technology in bridging the knowing-doing gap: an exploratory case study on knowledge application, Journal of Knowledge Management.

[113] N. Lindqvist, T. Price, Evaluation of feature selection methods for machine learning classification of breast cancer (2018).

[114] A. M. Henneghan, O. Palesh, M. Harrison, S. R. Kesler, Identifying cytokine predictors of cognitive functioning in breast cancer survivors up to 10 years post chemotherapy using machine learning, Journal of neuroimmunology 320 (2018) 38–47.

[115] S. N. Dorman, K. Baranova, J. H. Knoll, B. L. Urquhart, G. Mariani, M. L. Carcangiu, P. K. Rogan, Genomic signatures for paclitaxel and gemcitabine resistance in breast cancer derived by machine learning, Molecular oncology 10 (1) (2016) 85–100.

[116] W. H. Wolberg, W. N. Street, O. Mangasarian, Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates, Cancer letters 77 (2-3) (1994) 163–171.

[117] J. M. Jerez, I. Molina, P. J. García-Laencina, E. Alba, N. Ribelles, M. Martín, L. Franco, Missing data imputation using statistical and machine learning methods in a real breast cancer problem, Artificial intelligence in medicine 50 (2) (2010) 105–115.

[118] M. Montazeri, M. Montazeri, M. Montazeri, A. Beigzadeh, Machine learning models in breast cancer survival prediction, Technology and Health Care 24 (1) (2016) 31–42.

[119] T. Mungle, S. Tewary, D. Das, I. Arun, B. Basak, S. Agarwal, R. Ahmed, S. Chatterjee, C. Chakraborty, Mrf-ann: a machine learning approach for automated er scoring of breast cancer immunohistochemical images, Journal of microscopy 267 (2) (2017) 117–129.

[120] J. Wang, X. Yang, H. Cai, W. Tan, C. Jin, L. Li, Discrimination of breast cancer with microcalcifications on mammography by deep learning, Scientific reports 6 (2016) 27327.

[121] K. Polat, S. Güneş, Breast cancer diagnosis using least square support vector machine, Digital signal processing 17 (4) (2007) 694–701.

[122] S. Vural, X. Wang, C. Guda, Classification of breast cancer patients using somatic mutation profiles and machine learning approaches, BMC systems biology 10 (3) (2016) 62.

[123] S. Turgut, M. Dağtekin, T. Ensari, Microarray breast cancer data classification using machine learning methods, in: 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), IEEE, 2018, pp. 1–3.

[124] T. H. Nguyen, S. Sridharan, V. Macias, A. Kajdacsy-Balla, J. Melamed, M. N. Do, G. Popescu, Automatic gleason grading of prostate cancer using quantitative phase imaging and machine learning, Journal of biomedical optics 22 (3) (2017) 036015.

[125] Y. Gu, X. Lin, A. Kapoor, W. Mei, D. Tang, Attempt to predict early recurrence of prostate cancer following prostatectomy through machine learning, AME Medical Journal 3.

[126] X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.-F. Yin, Q. J. Wu, A planning quality evaluation tool for prostate adaptive imrt based on machine learning, Medical physics 38 (2) (2011) 719–726.

[127] L. Hussain, A. Ahmed, S. Saeed, S. Rathore, I. A. Awan, S. A. Shah, A. Majid, A. Idris, A. A. Awan, Prostate cancer detection using machine learning techniques by employing combination of features extracting strategies, Cancer Biomarkers (Preprint) (2018) 1–21.

[128] J. Wang, C.-J. Wu, M.-L. Bao, J. Zhang, X.-N. Wang, Y.-D. Zhang, Machine learning-based analysis of mr radiomics can help to improve the diagnostic performance of pi-rads v2 in clinically relevant prostate cancer, European radiology 27 (10) (2017) 4082–4090.

[129] J. Ko, et al., Combining machine learning and nanofluidic technology to diagnose pancreatic cancer using exosomes, Acs Nano 11 (11) (2017) 11182–11193.

[130] M. Osman, 748p predicting survival of pancreatic cancer using supervised machine learning, Annals of Oncology 29 (suppl_8) (2018) mdy282–131.

[131] R. E. Reinbolt, S. Sonis, C. D. Timmers, J. L. Fernández-Martínez, A. Cernea, E. J. de Andrés-Galiana, S. Hashemi, K. Miller, R. Pilarski, M. B. Lustberg, Genomic risk prediction of aromatase inhibitor-related arthralgia in patients with breast cancer using a novel machine-learning algorithm, Cancer medicine 7 (1) (2018) 240–253.

[132] S. R. Kesler, A. Rao, D. W. Blayney, I. A. Oakley-Girvan, M. Karuturi, O. Palesh, Predicting long-term cognitive outcome following breast cancer with pre-treatment resting state fmri and random forest machine learning, Frontiers in human neuroscience 11 (2017) 555.

[133] S. S. Garapati, L. Hadjiiski, K. H. Cha, H.-P. Chan, E. M. Caoili, R. H. Cohan, A. Weizer, A. Alva, C. Paramagul, J. Wei, et al., Urinary bladder cancer staging in ct urography using machine learning, Medical physics 44 (11) (2017) 5814–5823.

[134] G. Wang, K.-M. Lam, Z. Deng, K.-S. Choi, Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques, Computers in biology and medicine 63 (2015) 124–132.

[135] S.-W. Chang, S. Abdul-Kareem, A. F. Merican, R. B. Zain, Oral cancer prognosis based on clinicopathologic and genomic markers using a hybrid of feature selection and machine learning methods, BMC Bioinformatics 14 (1) 170.

[136] Q. Liu, A. H. Sung, Z. Chen, J. Liu, L. Chen, M. Qiao, Z. Wang, X. Huang, Y. Deng, Gene selection and classification for cancer microarray data based on machine learning and similarity measures, BMC Genomics 12 (5) S1.

[137] D. Haussler, D. W. Bednarski, M. Schummer, N. Cristianini, N. Duffy, T. S. Furey, Support vector machine classification and validation of cancer tissue samples using microarray expression data , Bioinformatics 16 (10) (2000) 906–914.

[138] C. Parmar, P. Grossmann, D. Rietveld, M. M. Rietbergen, P. Lambin, H. J. W. L. Aerts, Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer, Frontiers in Oncology 5 (2015) 272.

[139] S.-B. Cho, H.-H. Won, Machine learning in dna microarray analysis for cancer classification, in: Proceedings of the First Asia-Pacific Bioinformatics Conference on Bioinformatics 2003 - Volume 19, Australian Computer Society, Inc., 2003, pp. 189–198.

[140] O. Klein, F. Kanter, H. Kulbe, P. Jank, C. Denkert, G. Nebrich, W. D. Schmitt, Z. Wu, C. A. Kunze, J. Sehouli, S. Darb-Esfahani, I. Braicu, J. Lellmann, H. Thiele, E. T. Taube, Maldi-imaging for classification of epithelial ovarian cancer histotypes from a tissue microarray using machine learning methods, PROTEOMICS Clinical Applications 13 (1) (2019) 1700181.

[141] G. P. Way, R. J. Allaway, S. J. Bouley, C. E. Fadul, Y. Sanchez, C. S. Greene, A machine learning classifier trained on cancer transcriptomes detects nf1 inactivation signal in glioblastoma, BMC Genomics 18 (1) (2017) 127.

[142] H. Chen, H. Zhao, J. Shen, R. Zhou, Q. Zhou, Supervised machine learning model for high dimensional gene data in colon cancer detection, in: 2015 IEEE International Congress on Big Data, IEEE, 2015, pp. 134–141.

[143] M. G. Bari, C. Y. Ung, C. Zhang, S. Zhu, H. Li, Machine learning-assisted network inference approach to identify a new class of genes that coordinate the functionality of cancer networks, Scientific reports 7 (1) (2017) 6993.

[144] S. M. Ayyad, A. I. Saleh, L. M. Labib, Gene expression cancer classification using modified k-nearest neighbors technique, Biosystems 176 (2019) 41–51.

[145] J. Xu, P. Wu, Y. Chen, Q. Meng, H. Dawood, M. M. Khan, A novel deep flexible neural forest model for classification of cancer subtypes based on gene expression data, IEEE Access 7 (2019) 22086–22095.

[146] I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene selection for cancer classification using support vector machines, Machine learning 46 (1-3) (2002) 389–422.

[147] A. C. Tan, D. Gilbert, Ensemble machine learning on gene expression data for cancer classification.

[148] Y. Wang, I. V. Tetko, M. A. Hall, E. Frank, A. Facius, K. F. Mayer, H. W. Mewes, Gene selection from microarray data for cancer classificationï¿½a machine learning approach, Computational biology and chemistry 29 (1) (2005) 37–46.

[149] X. Jin, A. Xu, R. Bie, P. Guo, Machine learning techniques and chi-square feature selection for cancer classification using sage gene expression profiles, in: International Workshop on Data Mining for Biomedical Applications, Springer Berlin Heidelberg, 2006, pp. 106–115.

[150] D. P. Kovács, W. McCorkindale, A. A. Lee, Quantitative interpretation explains machine learning models for chemical reaction prediction and uncovers bias, Nature communications 12 (1) (2021) 1–9.

[151] A. Patsilinakos, M. Artini, R. Papa, M. Sabatino, M. Božović, S. Garzoli, G. Vrenna, R. Buzzi, S. Manfredini, L. Selan, et al., Machine learning analyses on data including essential oil chemical composition and in vitro experimental antibiofilm activities against staphylococcus species, Molecules 24 (5) (2019) 890.

[152] C. Jiang, H. Yang, P. Di, W. Li, Y. Tang, G. Liu, In silico prediction of chemical reproductive toxicity using machine learning, Journal of Applied Toxicology 39 (6) (2019) 844–854.

[153] T. Drevinskas, R. Mickienė, A. Maruška, M. Stankevičius, N. Tiso, A. Šalomskas, R. Lelešius, A. Karpovaitė, O. Ragažinskienė, Confirmation of the antiviral properties of medicinal plants via chemical analysis, machine learning methods and antiviral tests: a methodological approach, Analytical Methods 10 (16) (2018) 1875–1885.

[154] F. A. Azevedo, L. R. Carvalho, L. T. Grinberg, J. M. Farfel, R. E. Ferretti, R. E. Leite, W. J. Filho, R. Lent, S. Herculano-Houzel, Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain, Journal of Comparative Neurology 513 (5) (2009) 532–541.

[155] C. Davatzikos, K. Ruparel, Y. Fan, D. Shen, M. Acharyya, J. Loughead, R. Gur, D. D. Langleben, Classifying spatial patterns of brain activity with machine learning methods: application to lie detection, Neuroimage 28 (3) (2005) 663–668.

[156] E. I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E. R. Melhem, C. Davatzikos, Classification of brain tumor type and grade using mri texture and shape in a machine learning scheme, Magnetic Resonance in Medicine: An Official

Journal of the International Society for Magnetic Resonance in Medicine 62 (6) (2009) 1609–1618.

[157] R. Basnet, S. Mukkamala, A. H. Sung, Detection of phishing attacks: A machine learning approach, in: Soft Computing Applications in Industry, Springer, 2008, pp. 373–383.

[158] A. Jain, D. Maltoni, D. Maio, J. Wayman, Biometric systems technology, design and performance evaluation, Springer-Verlag London limited, 2005.

[159] K.-R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, B. Blankertz, Machine learning for real-time single-trial eeg-analysis: from brain–computer interfacing to mental state monitoring, Journal of neuroscience methods 167 (1) (2008) 82–90.

[160] A. Kübler, K.-R. Müller, An introduction to brain-computer interfacing, Toward brain-computer interfacing (2007) 1–25.

[161] K.-R. Müller, M. Krauledat, G. Dornhege, G. Curio, B. Blankertz, Machine learning techniques for brain-computer interfaces.

[162] P. Sajda, A. Gerson, K.-R. Muller, B. Blankertz, L. Parra, A data analysis competition to evaluate machine learning algorithms for use in brain-computer interfaces, IEEE Transactions on neural systems and rehabilitation engineering 11 (2) (2003) 184–185.

[163] B. Y. Gravesteijn, D. Nieboer, A. Ercole, H. F. Lingsma, D. Nelson, B. Van Calster, E. W. Steyerberg, C. Åkerlund, K. Amrein, N. Andelic, et al., Machine learning algorithms performed no better than regression models for prognostication in traumatic brain injury, Journal of clinical epidemiology 122 (2020) 95–107.

[164] Z. Lv, L. Qiao, Q. Wang, F. Piccialli, Advanced machine-learning methods for brain-computer interfacing, IEEE/ACM Transactions on Computational Biology and Bioinformatics.

[165] M. A. Khan, I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, A. Rehman, S. A. C. Bukhari, Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists, Diagnostics 10 (8) (2020) 565.

[166] A. E. Flanders, L. M. Prevedello, G. Shih, S. S. Halabi, J. Kalpathy-Cramer, R. Ball, J. T. Mongan, A. Stein, F. C. Kitamura, M. P. Lungren, et al., Construction of a machine learning dataset through collaboration: the rsna 2019 brain ct hemorrhage challenge, Radiology: Artificial Intelligence 2 (3) (2020) e190211.

[167] S. Lin, H. Jia, L. Abualigah, M. Altalhi, Enhanced slime mould algorithm for multilevel thresholding image segmentation using entropy measures, Entropy 23 (12) (2021) 1700.

[168] E. H. Houssein, K. Hussain, L. Abualigah, M. Abd Elaziz, W. Alomoush, G. Dhiman, Y. Djenouri, E. Cuevas, An improved opposition-based marine predators algorithm for global optimization and multilevel thresholding image segmentation, Knowledge-Based Systems 229 (2021) 107348.

[169] L. Abualigah, A. Diabat, P. Sumari, A. H. Gandomi, A novel evolutionary arithmetic optimization algorithm for multilevel thresholding segmentation of covid-19 ct images, Processes 9 (7) (2021) 1155.

[170] B. J. Erickson, P. Korfiatis, Z. Akkus, T. L. Kline, Machine learning for medical imaging, Radiographics 37 (2) (2017) 505–515.

[171] D. Shen, G. Wu, D. Zhang, K. Suzuki, F. Wang, P. Yan, Machine learning in medical imaging., Comp. Med. Imag. and Graph. 41 (2015) 1–2.

[172] K. Suzuki, P. Yan, F. Wang, D. Shen, Machine learning in medical imaging, International journal of biomedical imaging 2012.

[173] P. Yan, K. Suzuki, F. Wang, D. Shen, Machine learning in medical imaging (2013).

[174] M. L. Giger, Machine learning in medical imaging, Journal of the American College of Radiology 15 (3) (2018) 512–520.

[175] J. Zhang, K.-K. Ma, M.-H. Er, V. Chong, Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine, in: International Workshop on Advanced Image Technology (IWAIT'04), 2004, pp. 207–211.

[176] I. El-Naqa, Y. Yang, M. N. Wernick, N. P. Galatsanos, R. M. Nishikawa, A support vector machine approach for detection of microcalcifications, IEEE transactions on medical imaging 21 (12) (2002) 1552–1563.

[177] S. Lemm, B. Blankertz, T. Dickhaus, K.-R. Müller, Introduction to machine learning for brain imaging, Neuroimage 56 (2) (2011) 387–399.

[178] S. Wang, R. M. Summers, Machine learning and radiology, Medical image analysis 16 (5) (2012) 933–951.

[179] J.-G. Lee, S. Jun, Y.-W. Cho, H. Lee, G. B. Kim, J. B. Seo, N. Kim, Deep learning in medical imaging: general overview, Korean journal of radiology 18 (4) (2017) 570–584.

[180] J. Schmidhuber, Deep learning in neural networks: An overview, Neural networks 61 (2015) 85–117.

[181] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, C. I. Sánchez, A survey on deep learning in medical image analysis, Medical image analysis 42 (2017) 60–88.

[182] D. Shen, G. Wu, H.-I. Suk, Deep learning in medical image analysis, Annual review of biomedical engineering 19 (2017) 221–248.

[183] N. K. Batmanghelich, B. Taskar, C. Davatzikos, Generative-discriminative basis learning for medical imaging, IEEE transactions on medical imaging 31 (1) (2012) 51–69.

[184] K. Suzuki, Pixel-based machine learning in medical imaging, Journal of Biomedical Imaging 2012 (2012) 1.

[185] Y. Wang, Y. Fan, P. Bhatt, C. Davatzikos, High-dimensional pattern regression using machine learning: from medical images to continuous clinical variables, Neuroimage 50 (4) (2010) 1519–1535.

[186] L. Oakden-Rayner, J. Dunnmon, G. Carneiro, C. Ré, Hidden stratification causes clinically meaningful failures in machine learning for medical imaging, in: Proceedings of the ACM conference on health, inference, and learning, 2020, pp. 151–159.

[187] M. J. Willemink, W. A. Koszek, C. Hardell, J. Wu, D. Fleischmann, H. Harvey, L. R. Folio, R. M. Summers, D. L. Rubin, M. P. Lungren, Preparing medical imaging data for machine learning, Radiology 295 (1) (2020) 4–15.

[188] A. A. Ewees, L. Abualigah, D. Yousri, A. T. Sahlol, M. A. Al-qaness, S. Alshathri, M. A. Elaziz, Modified artificial ecosystem-based optimization for multilevel thresholding image segmentation, Mathematics 9 (19) (2021) 2363.

[189] R. Rehouma, M. Buchert, Y.-P. P. Chen, Machine learning for medical imaging-based covid-19 detection and diagnosis, International Journal of Intelligent Systems.

[190] A. Barragán-Montero, U. Javaid, G. Valdés, D. Nguyen, P. Desbordes, B. Macq, S. Willems, L. Vandewinckele, M. Holmström, F. Löfman, et al., Artificial intelligence and machine learning for medical imaging: A technology review, Physica Medica 83 (2021) 242–256.

[191] G. A. Kaissis, M. R. Makowski, D. Rückert, R. F. Braren, Secure, privacy-preserving and federated machine learning in medical imaging, Nature Machine Intelligence 2 (6) (2020) 305–311.

[192] D. Ravì, C. Wong, B. Lo, G. Yang, A deep learning approach to on-node sensor data analytics for mobile or wearable devices, IEEE Journal of Biomedical and Health Informatics 21 (1) (2017) 56–64.

[193] N. Y. Hammerla, S. Halloran, T. Ploetz, Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables, arXiv e-prints.

[194] B. M. Eskofier, S. I. Lee, J. Daneault, F. N. Golabchi, G. Ferreira-Carvalho, G. Vergara-Diaz, S. Sapienza, G. Costante, J. Klucken, T. Kautz, P. Bonato, Recent machine learning advancements in sensor-based mobility analysis: Deep learning for parkinson's disease assessment, in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016, pp. 655–658.

[195] Y. Kwon, K. Kang, C. Bae, Unsupervised learning for human activity recognition using smartphone sensors, Expert Systems with Applications 41 (14) (2014) 6067 – 6074.

[196] J.-Y. Yang, J.-S. Wang, Y.-P. Chen, Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers, Pattern recognition letters 29 (16) (2008) 2213–2220.

[197] J. Mantyjarvi, J. Himberg, T. Seppanen, Recognizing human motion with multiple acceleration sensors, in: 2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236), Vol. 2, IEEE, 2001, pp. 747–752.

[198] C. Zhu, W. Sheng, Human daily activity recognition in robot-assisted living using multi-sensor fusion, in: 2009 IEEE International Conference on Robotics and Automation, IEEE, 2009, pp. 2154–2159.

[199] C. Zhu, W. Sheng, Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 41 (3) (2011) 569–573.

[200] W. Jiang, Z. Yin, Human activity recognition using wearable sensors by deep convolutional neural networks, in: Proceedings of the 23rd ACM international conference on Multimedia, ACM, 2806333, 2015, pp. 1307–1310.

[201] F. Ordóñez, D. Roggen, Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition, Sensors 16 (1) (2016) 115.

[202] E. P. Stuntebeck, J. S. Davis II, G. D. Abowd, M. Blount, Healthsense: classification of health-related sensor data through user-assisted machine learning, in: Proceedings of the 9th workshop on Mobile computing systems and applications, ACM, 2008, pp. 1–5.

[203] K. H. Lee, N. Verma, A low-power processor with configurable embedded machine-learning accelerators for high-order and adaptive analysis of medical-sensor signals, IEEE Journal of Solid-State Circuits 48 (7) (2013) 1625–1637.

[204] A. Mannini, D. Trojaniello, A. Cereatti, A. Sabatini, A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and huntingtons disease patients, Sensors 16 (1) (2016) 134.

[205] O. D. Lara, M. A. Labrador, A survey on human activity recognition using wearable sensors, IEEE Communications Surveys & Tutorials 15 (3) (2013) 1192–1209.

[206] S. Shan, T. Yuan, A wearable pre-impact fall detector using feature selection and support vector machine, in: IEEE 10th International Conference On Signal Processing Proceedings, IEEE, 2010, pp. 1686–1689.

[207] F.-T. Sun, C. Kuo, H.-T. Cheng, S. Buthpitiya, P. Collins, M. Griss, Activity-aware mental stress detection using physiological sensors, in: International conference on Mobile computing, applications, and services, Springer, 2010, pp. 282–301.

[208] K. Yatani, K. N. Truong, Bodyscope: a wearable acoustic sensor for activity recognition, in: Proceedings of the 2012 ACM Conference on Ubiquitous Computing, ACM, 2012, pp. 341–350.

[209] J. Suutala, S. Pirttikangas, J. Röning, Discriminative temporal smoothing for activity recognition from wearable sensors, in: International Symposium on Ubiquitious Computing Systems, Springer, 2007, pp. 182–195.

[210] T. Zhang, J. Wang, L. Xu, P. Liu, Fall detection by wearable sensor and one-class svm algorithm, in: Intelligent computing in signal processing and pattern recognition, Springer, 2006, pp. 858–863.

[211] J. P. Varkey, D. Pompili, T. A. Walls, Human motion recognition using a wireless sensor-based wearable system, Personal and Ubiquitous Computing 16 (7) (2012) 897–910.

[212] S. Patel, K. Lorincz, R. Hughes, N. Huggins, J. Growdon, D. Standaert, M. Akay, J. Dy, M. Welsh, P. Bonato, Monitoring motor fluctuations in patients with parkinson's disease using wearable sensors, IEEE transactions on information technology in biomedicine 13 (6) (2009) 864–873.

[213] L. Clifton, D. A. Clifton, M. A. Pimentel, P. J. Watkinson, L. Tarassenko, Predictive monitoring of mobile patients by combining clinical observations with data from wearable sensors, IEEE journal of biomedical and health informatics 18 (3) (2014) 722–730.

[214] T. Huynh, U. Blanke, B. Schiele, Scalable recognition of daily activities with wearable sensors, in: International Symposium on Location-and Context-Awareness, Springer, 2007, pp. 50–67.

[215] A. Fleury, M. Vacher, N. Noury, Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results, IEEE transactions on information technology in biomedicine 14 (2) (2010) 274–283.

[216] T. Huynh, B. Schiele, Towards less supervision in activity recognition from wearable sensors, in: 2006 10th IEEE International Symposium on Wearable Computers, Citeseer, 2006, pp. 3–10.

[217] M. Stikic, D. Larlus, S. Ebert, B. Schiele, Weakly supervised recognition of daily life activities with wearable sensors, IEEE transactions on pattern analysis and machine intelligence 33 (12) (2011) 2521–2537.

[218] C. S. Tucker, I. Behoora, H. B. Nembhard, M. Lewis, N. W. Sterling, X. Huang, Machine learning classification of medication adherence in patients with movement disorders using non-wearable sensors, Computers in Biology and Medicine 66 (2015) 120 – 134.

[219] A. Sano, A. J. Phillips, Z. Y. Amy, A. W. McHill, S. Taylor, N. Jaques, C. A. Czeisler, E. B. Klerman, R. W. Picard, Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones, in: 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN), IEEE, 2015, pp. 1–6.

[220] A. Sano, R. W. Picard, Stress recognition using wearable sensors and mobile phones, in: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, IEEE, 2013, pp. 671–676.

[221] O. Ojetola, E. I. Gaura, J. Brusey, Fall detection with wearable sensors–safe (smart fall detection), in: 2011 Seventh International Conference on Intelligent Environments, 2011, pp. 318–321.

[222] E. M. Tapia, S. S. Intille, W. Haskell, K. Larson, J. Wright, A. King, R. Friedman, Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor, in: 2007 11th IEEE international symposium on wearable computers, IEEE, 2007, pp. 37–40.

[223] C. K. Williams, C. E. Rasmussen, Gaussian processes for machine learning, Vol. 2, MIT press Cambridge, MA, 2006.

[224] L. Clifton, D. A. Clifton, M. A. F. Pimentel, P. J. Watkinson, L. Tarassenko, Gaussian processes for personalized e-health monitoring with wearable sensors, IEEE Transactions on Biomedical Engineering 60 (1) (2013) 193–197.

[225] L. L. Duan, X. Wang, J. P. Clancy, R. D. Szczesniak, Joint hierarchical gaussian process model with application to personalized prediction in medical monitoring, Stat 7 (1) (2018) e178.

[226] L.-F. Cheng, B. Dumitrascu, G. Darnell, C. Chivers, M. Draugelis, K. Li, B. E. Engelhardt, Sparse multi-output gaussian processes for online medical time series prediction, BMC medical informatics and decision making 20 (1) (2020) 1–23.

[227] J.-R. Fouefack, B. Borotikar, M. Lüthi, T. S. Douglas, V. Burdin, T. E. Mutsvangwa, Dynamic multi feature-class gaussian process models, arXiv preprint arXiv:2112.04495.

[228] L. Abualigah, A. Diabat, Improved multi-core arithmetic optimization algorithm-based ensemble mutation for multidisciplinary applications, Journal of Intelligent Manufacturing (2022) 1–42.

[229] L. Abualigah, A. Diabat, Chaotic binary group search optimizer for feature selection, Expert Systems with Applications 192 (2022) 116368.

[230] M. Zhang, A. A. Sawchuk, A feature selection-based framework for human activity recognition using wearable multimodal sensors, in: Proceedings of the 6th International Conference on Body Area Networks, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011, pp. 92–98.

[231] J. Lester, T. Choudhury, N. Kern, G. Borriello, B. Hannaford, A hybrid discriminative/generative approach for modeling human activities, in: Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence, 2005.

[232] S. Pirttikangas, K. Fujinami, T. Nakajima, Feature selection and activity recognition from wearable sensors, in: Ubiquitous Computing Systems, Springer Berlin Heidelberg, 2006, pp. 516–527.

[233] L. Wang, T. Gu, X. Tao, H. Chen, J. Lu, Recognizing multi-user activities using wearable sensors in a smart home, Pervasive and Mobile Computing 7 (3) (2011) 287 – 298.

[234] W. Ugulino, D. Cardador, K. Vega, E. Velloso, R. Milidiú, H. Fuks, Wearable computing: Accelerometers' data classification of body postures and movements, in: Advances in Artificial Intelligence - SBIA 2012, Springer Berlin Heidelberg, 2012, pp. 52–61.

[235] R. Brugarolas, D. Roberts, B. Sherman, A. Bozkurt, Machine learning based posture estimation for a wireless canine machine interface, in: 2013 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems, IEEE, 2013, pp. 10–12.

[236] B. T. Cole, S. H. Roy, C. J. De Luca, S. H. Nawab, Dynamical learning and tracking of tremor and dyskinesia from wearable sensors, IEEE Transactions on Neural Systems and Rehabilitation Engineering 22 (5) (2014) 982–991.

[237] A. Mannini, A. M. Sabatini, Machine learning methods for classifying human physical activity from on-body accelerometers, Sensors 10 (2) (2010) 1154–1175.

[238] O. Aziz, S. N. Robinovitch, An analysis of the accuracy of wearable sensors for classifying the causes of falls in humans, IEEE transactions on neural systems and rehabilitation engineering 19 (6) (2011) 670–676.

[239] X. Ma, H. Wang, B. Xue, M. Zhou, B. Ji, Y. Li, Depth-based human fall detection via shape features and improved extreme learning machine, IEEE journal of biomedical and health informatics 18 (6) (2014) 1915–1922.

[240] A. E. Ezugwu, A. M. Ikotun, O. O. Oyelade, L. Abualigah, J. O. Agushaka, C. I. Eke, A. A. Akinyelu, A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects, Engineering Applications of Artificial Intelligence 110 (2022) 104743.

[241] K. H. Almotairi, L. Abualigah, Hybrid reptile search algorithm and remora optimization algorithm for optimization tasks and data clustering, Symmetry 14 (3) (2022) 458.

[242] D. Wu, H. Jia, L. Abualigah, Z. Xing, R. Zheng, H. Wang, M. Altalhi, Enhance teaching-learning-based optimization for tsallis-entropy-based feature selection classification approach, Processes 10 (2) (2022) 360.

[243] M. M. Rahman, P. Bhattacharya, B. C. Desai, A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback, IEEE transactions on Information Technology in Biomedicine 11 (1) (2007) 58–69.

[244] Z. Wang, M. Jiang, Y. Hu, H. Li, An incremental learning method based on probabilistic neural networks and adjustable fuzzy clustering for human activity recognition by using wearable sensors, IEEE Transactions on Information Technology in Biomedicine 16 (4) (2012) 691–699.

[245] R. Akhavian, A. H. Behzadan, Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers, Advanced Engineering Informatics 29 (4) (2015) 867–877.

[246] A. Özdemir, B. Barshan, Detecting falls with wearable sensors using machine learning techniques, Sensors 14 (6) (2014) 10691–10708.

[247] M. Zhang, A. A. Sawchuk, Human daily activity recognition with sparse representation using wearable sensors, IEEE journal of Biomedical and Health Informatics 17 (3) (2013) 553–560.

[248] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, Y. Amirat, Physical human activity recognition using wearable sensors, Sensors 15 (12) (2015) 31314–31338.

[249] B. Barshan, M. C. Yüksek, Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units, The Computer Journal 57 (11) (2014) 1649–1667.

[250] O. N. Oyelade, A. E. Ezugwu, T. I. Mohamed, L. Abualigah, Ebola optimization search algorithm: A new nature-inspired metaheuristic algorithm with application in medical image classification problem, IEEE Access.

[251] J. O. Agushaka, A. E. Ezugwu, L. Abualigah, Dwarf mongoose optimization algorithm, Computer Methods in Applied Mechanics and Engineering 391 (2022) 114570.

[252] L. Abualigah, M. Abd Elaziz, P. Sumari, Z. W. Geem, A. H. Gandomi, Reptile search algorithm (rsa): A nature-inspired meta-heuristic optimizer, Expert Systems with Applications 191 (2022) 116158.

[253] L. Abualigah, D. Yousri, M. Abd Elaziz, A. A. Ewees, M. A. Al-Qaness, A. H. Gandomi, Aquila optimizer: a novel meta-heuristic optimization algorithm, Computers & Industrial Engineering 157 (2021) 107250.

[254] L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz, A. H. Gandomi, The arithmetic optimization algorithm, Computer methods in applied mechanics and engineering 376 (2021) 113609.

[255] M. R. Boutell, J. Luo, X. Shen, C. M. Brown, Learning multi-label scene classification, Pattern recognition 37 (9) (2004) 1757–1771.

[256] X.-w. Chen, J. C. Jeong, Enhanced recursive feature elimination, in: Sixth International Conference on Machine Learning and Applications (ICMLA 2007), IEEE, 2007, pp. 429–435.

[257] M. Shehab, A. T. Khader, M. Laouchedi, O. A. Alomari, Hybridizing cuckoo search algorithm with bat algorithm for global numerical optimization, The Journal of Supercomputing (2018) 1–28.

[258] Y. Nojima, H. Ishibuchi, Designing fuzzy ensemble classifiers by evolutionary multiobjective optimization with an entropy-based diversity criterion, in: 2006 Sixth International Conference on Hybrid Intelligent Systems (HIS'06), IEEE, 2006, pp. 59–59.

[259] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, International journal of computer vision 115 (3) (2015) 211–252.

[260] Z. Huang, A fast clustering algorithm to cluster very large categorical data sets in data mining., DMKD 3 (8) (1997) 34–39.

[261] S. Balaji, K. Srivatsa, Unsupervised learning in large datasets for intelligent decision making, International Journal of Scientific and Research Publications 2 (1) (2012) 2250–3153.