

Overview of AI and Machine Learning in Healthcare

Antonia Voutouri, Arina Kostina, Panagiotis Menelaou, Michaela Bratsa, and Spyridon Sachmpazidis

[avouto01@ucy.ac.cy, akosti02@ucy.ac.cy, pmenel01@ucy.ac.cy, mbrats01@ucy.ac.cy, ssachm02@ucy.ac.cy]

Department of Computer Science,

University of Cyprus,

Nicosia, Cyprus

Abstract—This paper examines Artificial Intelligence in healthcare, detailing its evolution from historical developments to current applications and future potential. It highlights AI’s impact on disease detection, health service delivery, and drug discovery, while also addressing future challenges and innovations in the field. Moreover, it describes the distinct approaches taken by each category of Machine learning: Supervised Learning, Unsupervised Learning, Semi-supervised Learning, and Reinforcement Learning. It offers an overview of two fundamental machine learning algorithms: Naive Bayes for classification and K-means for clustering. The paper underscores the key theoretical points, limitations, and practical applications associated with these algorithms. It also explores the significant advancements in medical imaging and electronic health records (EHR), highlighting the integration of artificial intelligence (AI) to enhance diagnostic precision and data management in healthcare and underscore the impact of technological progress on improving clinical outcomes and operational efficiencies in healthcare systems. Furthermore, this paper provides a perceptive examination of machine learning applications in important healthcare domains, with particular emphasis on breast cancer disclosure, thyroid condition diagnosis, and multiple sclerosis disease progression prediction. Symptoms, diagnoses, and the correlation of diseases with machine learning are also cited.

Index Terms—Future AI Healthcare, Artificial Intelligence in Healthcare, AI modern medicine, Past AI Healthcare, Present AI Healthcare, Future AI healthcare, Supervised Learning, Unsupervised Learning, Semi-supervised Learning, Reinforcement Learning, Naive Bayes, K-means, Classification algorithms, Clustering algorithms, Medical Imaging, Advancements in Medical Imaging, Electronic Health Records, Adoption of EHR, Digital Transformation in Healthcare, Overcoming challenges of EHR with AI, Thyroid Disorders, Hypothyroidism, Hyperthyroidism, Hashimoto’s Disease, Thyroid Cancer, Breast Cancer, Multiple Sclerosis.

I. INTRODUCTION

IN the modern era, where technology intersects seamlessly with daily life, Artificial Intelligence (AI) emerges as a revolutionary force in numerous fields, with healthcare at the forefront of this transformation. Projecting into the future, the already expansive role of AI in healthcare continues to evolve from theoretical underpinnings to real-world applications that promise to reshape medical practices. As healthcare challenges

intensify globally—exacerbated by aging populations, rising costs, and pandemics—the use of AI technologies holds great promise in enhancing disease detection, optimizing service delivery, and accelerating drug discovery.

This study offers insights into how early breakthroughs have paved the way for current advancements and future prospects by analyzing the historical evolution of AI in healthcare. It delves into the current state of AI applications in healthcare, discussing their impact on improving diagnostic accuracy, treatment efficacy, and operational efficiency. Furthermore, it projects into the future, speculating on how AI can develop further to meet the intricate ever-evolving problems facing the healthcare system of the future.

This comprehensive analysis aims not only to document the milestones of AI in healthcare but also to understand its limitations, address ethical considerations, and anticipate future trends. Through this exploration, the paper seeks to offer a nuanced perspective on AI’s potential to assist healthcare providers and revolutionize patient care, making healthcare more accessible, personalized, and effective.

II. AI IN HEALTHCARE

The Role of Artificial Intelligence in Modern Medicine

The healthcare industry is currently undergoing a profound transformation, prompted by escalating healthcare costs and a notable shortage of medical professionals. This shift is driven by the necessity to address multiple systemic challenges such as limited access to care, inefficiencies, and the impacts of an aging population[21], all magnified by global crises like the COVID-19 pandemic. These issues not only expose the vulnerabilities of existing healthcare systems but also highlight the urgent need for innovative solutions to enhance care delivery and improve administrative processes. The focus is shifting on tackling problems such as inequitable access, the absence of on-demand services, and a lack of transparency in pricing[35].

Artificial Intelligence (AI) is positioned as a pivotal tool in this transformation. It is being explored and implemented across various aspects of healthcare including disease detection, management of chronic conditions, delivery of

health services, and drug discovery. The potential of AI to significantly address these health challenges is immense. However, its practical effectiveness is often hampered by the quality of the data it relies on and its inherent lack of human traits, such as empathy and ethical judgment. Moreover, the deployment of AI in healthcare raises several ethical concerns, including the risk of inaccurate decisions, accountability for mistakes, challenges in validating AI outcomes, biases in training data, data protection issues, public trust, and the potential impacts on personal dignity and social isolation in care settings. A key challenge lies in ensuring that AI development and utilization are transparent, serve the public interest, and continue to foster innovation within the healthcare sector[51].

In the context of these ongoing developments, the necessity for AI in healthcare becomes evident. Modern healthcare systems are burdened by high costs and significant disparities in access to care. Traditionally, medical knowledge was confined to textbooks, journals, and expert opinions formed through a master-apprentice dynamic. Physicians would accumulate experience through direct patient care and outcome observation. However, this method has its limitations in assimilating and applying the vast array of available medical knowledge and experience to patient care. AI steps in as a crucial aid, merging extensive patient data to enhance and refine the effectiveness of medical treatments. When leveraged correctly, AI can uncover invaluable insights from this data, aiding clinical decisions and potentially transforming patient outcomes[1].

In the following paragraphs, we will explore the evolution of AI in healthcare, tracing its past developments, examining its current applications, and discussing its future potential to reshape healthcare practices.

The Past, Present and Future of AI in Healthcare

1) Historical Context of AI in Healthcare: The concept of Artificial Intelligence (AI) was formalized in 1956 during a workshop led by John McCarthy. This event set the stage for pioneering efforts by notable figures such as McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon, among others, who drove advancements in the field of AI. Alan Turing had earlier contributed foundational ideas with his article "Computing Machinery and Intelligence," proposing the potential for computers to simulate intelligence, including automatic learning capabilities. Subsequent developments by Newell and Simon in heuristic research and mathematical theorems, such as the Logic Theorist program, further expanded AI's scope. Innovations like Selfridge's face recognition technology and the mobile robot "Shakey," which utilized natural language processing, marked significant milestones in AI. By the late 1990s, technological advances and the availability of significant computing power allowed the creation of data-driven AI systems, leveraging vast datasets and sophisticated statistical techniques to solve complex problems[45].

In the healthcare sector, AI began its integration in the 1950s with the development of early computer-aided programs aimed at improving medical diagnoses. These initial applications leveraged AI to enhance clinical, diagnostic, and surgical processes by processing and analyzing vast amounts of medical data. This integration has significantly accelerated over the decades, thanks to advancements in computing capabilities and an increase in available digital data. Today, AI in healthcare not only continues to enhance various medical processes but also plays a crucial role in informed clinical decision-making and the discovery of new medical treatments. This historical overview highlights the significant, transformative impact that AI has had on medical practices from its inception to the sophisticated applications we encounter in modern healthcare[45], [35].

2) AI in Modern Medicine Today: Artificial intelligence (AI) is fundamentally described as computer software emulating human cognitive functions. Among various AI technologies, Artificial Neural Networks (ANN) have shown notable applications, especially in cardiology decision-making, proving more effective than traditional logistic regression models. For instance, ANNs have demonstrated predictive superiority in diagnosing myocardial ischemia in emergency patients with chest pain[46]. Another significant branch of AI, machine learning (ML), involves algorithms that learn from data to improve performance automatically. These algorithms are categorized into four types: supervised, unsupervised, semi-supervised, and reinforcement learning. In supervised learning, algorithms predict and stratify risks using methods like logistic regression and Bayesian networks[1], showing greater success in predicting outcomes compared to conventional clinical scores. Unsupervised learning, on the other hand, involves techniques like hierarchical clustering and principal component analysis, enhancing disease prediction accuracy in fields such as nuclear cardiology.

Deep learning, a more advanced form of ML, utilizes multiple layers of hidden neurons in neural networks to handle the increasing volume and complexity of big data. This method has been particularly effective in cardiac imaging, such as echocardiography and cardiac MRI, where it has improved the accuracy of endocardium tracking and left ventricle segmentation. Convolutional Neural Networks (CNN), a subtype of ANN derived from deep learning, have also shown efficacy in cardiac CT angiography for calculating coronary artery calcium using supervised learning, achieving more precise results than existing methods[51], [1].

These advancements in AI are dramatically reshaping healthcare today. AI is optimizing drug discovery by automating target identification and analyzing compounds for repurposing, thereby accelerating drug development and reducing redundancies. Major pharmaceutical companies are leveraging AI to enhance the discovery of treatments for diseases like cancer, signaling a shift towards more efficient and cost-effective drug development. AI is also revolutionizing

clinical trials by automating data management, enhancing trial accuracy, and shortening durations through predictive analytics on real-world data. In patient care, AI-powered systems and devices are improving the quality of life and precision in diagnostics across various medical settings. From enhancing maternal care through risk identification to enabling mobility with advanced robotics, AI's role is pivotal in modern healthcare[51].

As AI continues to evolve, it is clear that its integration into healthcare will lead to even more groundbreaking developments. The next section will explore the future potential of AI in healthcare, examining how it might further revolutionize the field and address the complex challenges of tomorrow's medical landscape.

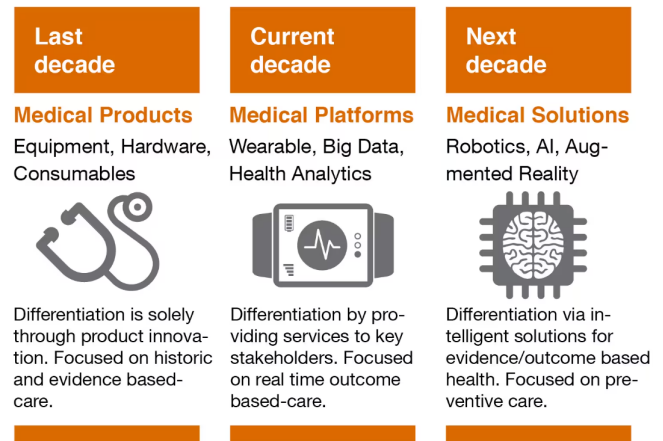
3) **Future Prospects of AI in Healthcare:** Artificial intelligence is already reshaping healthcare, yet the journey towards its full integration into clinical practice is fraught with complexities and challenges. One significant hurdle is the current regulatory environment, which lacks standardized criteria for assessing the safety and efficacy of AI algorithms. There is a clear need for legislation to address these issues before AI and machine learning can be fully embraced in clinical settings[21]. The FDA has made some strides by providing guidance on evaluating and implementing AI in general wellness products. Looking ahead, cognitive computing is expected to support clinicians by enhancing decision-making and improving patient outcome predictions. The vast data generated from routine medical procedures necessitates the integration of AI into everyday practice, a shift already beginning to take hold in specialties like pathology and radiology[51].

As healthcare becomes increasingly digitized, it is crucial for medical professionals to view AI not with fear but as an invaluable tool that deepens the understanding of patient data. Physicians must prepare for the AI era by acquiring the necessary skills to apply machine learning models and interpret their results accurately[21], [51].

Looking to the future, AI is set to revolutionize healthcare with its applications in disease detection, drug discovery, and chronic condition management. Yet, this transformation presents several ethical, governance, and technical challenges. Rapid developments in AI technology are testing existing regulatory frameworks that may need to be revised to maintain transparency, accountability, and public interest alignment. As AI systems grow more autonomous, the imperative for them to perform ethically in decision-making also intensifies, sparking ongoing philosophical debates about the moral capacity of machines and the determination of ethical standards[21], [51].

Moreover, the practical application of AI in healthcare is often limited by issues such as poor data quality, inconsistencies in digitization of medical records, and a lack of standardization in IT systems. These limitations highlight

AI's inability to fully replicate certain human qualities, such as compassion and contextual understanding, which are crucial in clinical settings. Along with concerns about the privacy and security of health data, these challenges outline the intricate landscape that AI must navigate to realize its potential in healthcare. It is essential for healthcare professionals and policymakers to acquire the skills needed to leverage AI technologies effectively, ensuring that AI fosters innovation while adhering to ethical standards and enhancing public welfare[21], [51].



Source: Frost & Sullivan, 'Transforming healthcare through artificial intelligence systems', 2016

Fig. 1. Healthcare Transformation: The Shift from Products to Intelligence. Source: PWC industries healthcare publications in AI robotics.

Conclusion: The Ongoing Evolution of AI in Healthcare and the Path Ahead

As we look to the future, the evolution of AI from its inception in the mid-20th century to its current applications in healthcare shows a trajectory of transformative growth[45]. AI has transitioned from enhancing basic medical diagnostics to playing a pivotal role in complex areas like drug discovery and personalized patient care. Looking ahead, AI promises even more revolutionary changes in healthcare[35]. However, the path forward requires navigating challenges such as refining regulatory standards, ensuring ethical use, and improving data quality. As AI becomes increasingly integrated into healthcare systems, it is imperative that we address these challenges to fully realize AI's potential to enhance patient outcomes and transform medical practices[21], [1], [46].

III. MACHINE LEARNING IS SPLIT INTO CATEGORIES

According to a leading expert in machine learning technologies, Arthur Samuel (1959), machine learning can be defined as the "field of study that gives computers the ability to learn without being explicitly programmed". Nowadays, several approaches attempt to tackle this problem, and numerous algorithms that implement machine learning have been developed. Their usage is essential in most fields of our modern society therefore a great deal of effort was invested in their research. Even though machine learning algorithms solve many different

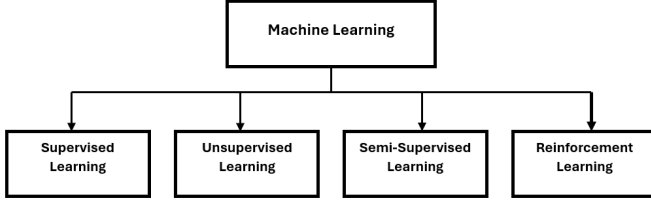


Fig. 2. Classification of Machine Learning Algorithms.

problems, they still share some characteristics that distinguish them into different groups. As seen in 2, machine learning can be categorized into 4 distinct categories: Supervised Learning, Unsupervised Learning, Semi-supervised Learning, and finally Reinforcement Learning[29].

A. Supervised Learning

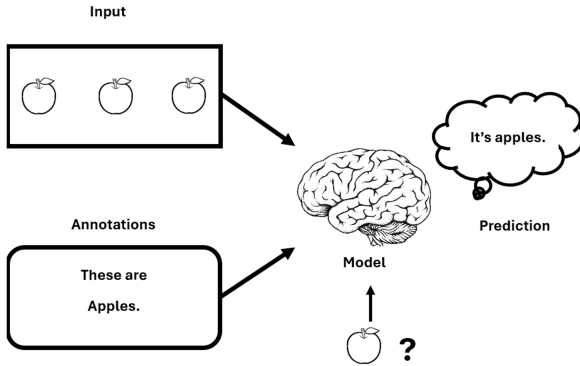


Fig. 3. Supervised learning model.

The first category of machine learning is Supervised Learning. All the algorithms that fall under this category follow the same key principles. The main principle behind supervised learning is the use of labeled data sets (which is data that has already been annotated with a label that meaningfully describes each element) in order to create a function that will be able to predict the outcome of future inputs that are unlabelled. These algorithms can also analyze their outputs and compare them with the intended results to finetune and rectify any potential errors in the model [12] [43]. A prime example of a Supervised Learning model can be seen in 3, where the labeled data (in this case apples) is given to the model for training purposes. After that, the model is asked to identify the given object (an apple) and the prediction of the machine is in this case correct. Essentially, supervised learning algorithms use the knowledge they gained in the past to predict the outcome [43]. Supervised Learning can be split into two categories: Classification and Regression. The former uses algorithms to classify the elements of a

dataset into distinct categories. For instance, some of the most common algorithms that use classification are the following: naïve Bayes, linear classifiers, k-nearest neighbor, random forest, support vector machines, and decision trees. The latter uses algorithms to explore the correlation of dependent and independent variables. For instance, some of the most common algorithms that use regression are the following: logistical regression, linear regression, and polynomial regression are all admired regression algorithms [19].

B. Unsupervised Learning

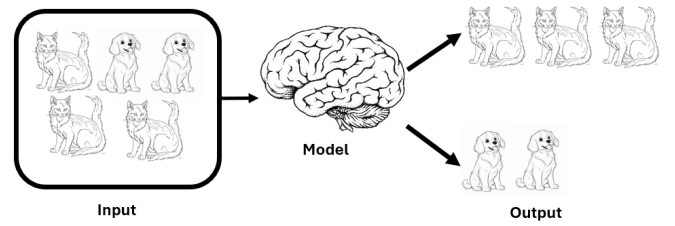


Fig. 4. Unsupervised learning model.

The second category of machine learning is Unsupervised Learning. The key difference between Unsupervised Learning and Supervised Learning is that it uses unlabelled data sets instead of labeled ones. Specifically, the model is given a data set consisting of unlabelled elements and it is up to the model to identify patterns in that data, in order to categorize it into groups. Naturally, this implies the usage of substantial amounts of data [31]. An example of Unsupervised Learning can be seen in 4, where unlabelled data (in this case images of cats and dogs) is given to the model. The model is then tasked with categorizing the data and using these algorithms it successfully distinguishes between the two types of animals [12]. Unsupervised Learning can be split into two categories: Clustering and Association. The former uses algorithms to group together elements that have the most correlation and separate them from those that have the least correlation. The latter uses algorithms to detect similar elements from big data sets. Some of the most common Unsupervised Learning algorithms include hierarchal clustering, k-means clustering, k-nearest neighbors, anomaly detection, apriori algorithm, neural networks, independent component analysis, principal component analysis and singular value decomposition. [19].

C. Semi-supervised Learning

The third category of machine learning is Semi-supervised Learning. This category is a combination of both Supervised Learning and Unsupervised Learning. It is characterized by the use of labelled and unlabelled data. Similarly to Supervised learning, using the provided data it aims to create a function to predict the output for future unknown test inputs [48] [19].

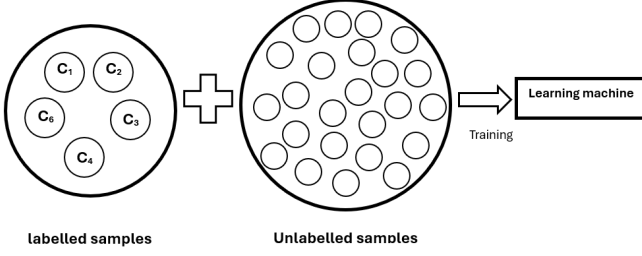


Fig. 5. Learning diagram for Semi-supervised learning.

As it can be seen in 5, the model is trained using labeled data and a much more significant quantity of unlabelled data, since unlabelled data is more prevalent in most fields of study [19]. Semi-supervised Learning can be categorized as Transductive or Inductive. The former uses the labeled data to train the model and then it takes the unlabelled data as input in order to predict a label for each element. The latter trains the model using both the labeled and unlabelled data as well as the unknown test data. As a result, it predicts the labels for the unlabelled data along with the unknown test data. These techniques can be used for classification, clustering dimension-reduction, and regression. Semi-supervised Learning depends on three assumptions: the smoothness assumption, the cluster assumption, and the manifold assumption. If the model has inadequate or inaccurate assumptions, then diminishing performance will be observed. Therefore, their accuracy is essential for a correct model [48].

D. Reinforcement Learning

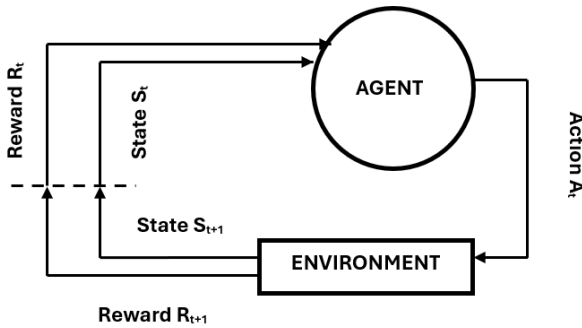


Fig. 6. Reinforcement learning framework.

Finally, the fourth category of machine learning is Reinforcement Learning. The key feature that separates this category is the usage of agents, the environment, and a feedback signal (reward) during the learning process. The philosophy behind these algorithms can be observed in 6, where the agent is continuously analyzing the environment (which is only a simulation of the real system since direct interaction with that

system is of high cost) and taking appropriate new actions depending on the result that previous action had on the state and the reward it earned. In the end, the experience acquired from this process is used to create the most favorable decision-making rule [37] [19]. Reinforcement Learning can be separated into two different categories: Positive Reinforcement and Negative Reinforcement. The former attempts everything to maximize the probability that the wanted behavior will reoccur. In contrast, the latter maximizes the same probability by avoiding unwanted behavior [19]. A Reinforcement Learning algorithm can be either Single Agent or Multi-Agent. As previously explained, for the single agent model in 6, the agent receives a reward based on the action he previously took. The value of the reward depends on the decision of the agent, positive means the decision was correct, and negative means it was incorrect. The objective of the agent is to maximize the reward across many several actions. The key difference in Multi-Agent models is, as the name suggests, the use of more than one agent. These agents make coordinated actions and together they aim to maximize the rewards they are receiving [37].

IV. MACHINE LEARNING ALGORITHMS

There is a huge amount of data generated from health monitoring systems, health service providers, and medical databases, which are too complex and voluminous to be processed and analyzed by traditional methods. Machine learning approaches offer the methodology and technology to transform these data into meaningful information for the automation of the decision making process in many areas of the medical field.[13]

With the fast improvement of the AI and ML field, there is a wide range of applications and sectors where these innovative techniques show great potential, and already are an indispensable tool that provides great assistance to medical professionals. There are many algorithms, each one of them with its own advantages and disadvantages, so choosing a machine learning algorithm mainly depends on the problem we are trying to solve. In this section, we are focusing on the most broadly used algorithms utilized in areas such as diagnosis-making, prediction, and treatment planning.

1) Classification Algorithms:

One of the most broadly used applications of machine learning in healthcare is the classification of data and digital imaging, for the purpose of detecting anomalies and stating the potential diagnosis. Machine learning classification algorithms are particularly good in this task as they predict which category an observation belongs to, based on the input features. In a classification task, the goal is to learn a mapping from the input features to the output classes, such that when new data is presented to the model, it can predict the class label for that data, based on the learned patterns from the training data. There are different types of classification algorithms that compute this mapping by learning the behavior of the data, ranging from linear classifiers like Logistic Regression, Naïve Bayes to Support Vector Machine, Tree-based Classifiers, and Neural Networks.[24]

Naïve Bayes

Naïve Bayes can be applied effectively for various tasks, including disease prediction, risk assessment, outcome prediction, cancer detection[20], and many more, and is one of the most effective classification algorithms. It is suited to use when the dimensionality of input is high[25] something that is very common in medical data because it often involves a large number of features or variables. These features can include patient demographics, medical history, genetic information, laboratory test results, image data, and more. Managing and analyzing such high-dimensional data is a significant challenge in medical research and diagnosis prediction tasks. That is why the Naive Bayes Classifier technique is often employed to handle this complexity effectively.

Theory

Naive Bayesian classifier calculates a set of probability by counting the frequency and combinations of values in a given data set. It assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. This is where the term "naive" comes from because it's a simplifying assumption that might not always hold true in real-world scenarios. However, the algorithm tends to learn quickly in a variety of controlled classification problems.[44] The Naive Bayes Classifier technique is based on the Bayes Theorem, which is stated below:

Let X be a data sample whose class label is not known and let H be some hypothesis, such that the data sample X may belong to a specified class C . Bayes theorem is used for calculating the posterior probability $P(C | X)$, from $P(C)$, $P(X)$, and $P(X | C)$. Where $P(C | X)$ is the posterior probability of the target class, which is the probability of the occurrence of event C when event X occurs. $P(C)$ is called the prior probability of class, which is the probability of the occurrence of C . $P(X | C)$ is the likelihood which is the probability of predictor of the given class, meaning that it is the probability of the occurrence of event X when event C occurs. $P(X)$ is the prior probability of the predictor of class, which is the probability of the occurrence of X . The formula is the following:

$$P(C|X) = \frac{P(X|C) * P(C)}{P(X)} \quad (1)$$

The Naive Bayes classifier does the following:

1. Let D be the training dataset associated with class labels. Each tuple, which is the data that is being assigned to a specific class, is represented by an n -dimensional element vector, $X = (x_1, x_2, x_3, \dots, x_n)$.
2. Suppose that we have m classes $C_1, C_2, C_3, \dots, C_m$. For an unknown tuple X , the classifier will predict that X belongs to the class with higher posterior probability, conditioned on X . i.e., the Naive Bayesian classifier assigns an unknown tuple X to the class C_i if and only if

$$P(C_i|X) > P(C_j|X) \quad (2)$$

For $1 \leq j \leq m$, and $i \neq j$, above posterior probabilities are computed using Bayes Theorem.[16]

The Naïve Bayesian classifier pays particular attention to the selection of features because this choice significantly impacts its performance, making it sensitive to feature selection. Features, also known as predictors, attributes, or independent variables, are the measurable properties or characteristics of the data that we provide to the machine learning model in order to train it. Therefore to improve the scalability, efficiency, and accuracy of the classifier it is important to have a good feature selection methodology that considers the domain and the algorithm characteristics.[5]

2) Clustering algorithms:

Clustering is an unsupervised Machine learning technique that is used to classify the unlabeled data set into groups with similar characteristics. If there are various data points then this can be defined as a way of grouping the unlabeled data point into various predefined classes. In medical data, there are often distinct subgroups or patterns that can be identified through clustering analysis. For example, in medical imaging data such as MRI or CT scans, different types of tissues or anomalies may form distinct clusters. Similarly, in genomic data, different genetic profiles or expression patterns may lead to the formation of clusters representing different patient subgroups or disease types. Some of the most broadly used algorithms are Hierarchical clustering, OPTICS, Gaussian clustering, and K- Mean Clustering[26]

K-means clustering

K-means is a numerical, unsupervised, non-deterministic, iterative method. It is simple and very fast, so in many practical applications, the method has proven to be a very effective way that can produce good clustering results. It is very suitable for producing globular clusters.[30]

Theory

Let x_1, x_2, \dots, x_p be our data points. Initially, we select how many clusters (k) we want to separate our data in. This is a crucial step because K-means requires you to specify the number of clusters upfront. Next, the algorithm randomly selects k points from your data as the initial cluster centers. These points are called centroids: $C_1, C_2, C_3, \dots, C_k$.

- 1) In the next phase each data point in the dataset is assigned to the nearest centroid based on some distance measure, usually the Euclidean distance. With the Euclidean equation the distance between a vector $x = (x_1, \dots, x_n)$ and a vector $y = (y_1, \dots, y_n)$ is calculated as follows:

$$D = ||x - y|| = \left[\sum_{i=1}^n (x_i - y_i)^2 \right]^{\frac{1}{2}} \quad (3)$$

- 2) After all data points have been assigned to clusters, the centroids are re-calculated. Each centroid is moved to the mean (average) of all the data points assigned to its cluster, therefore the new centroid is the center of its cluster:

$$C_j = \frac{1}{|S_j|} \sum_{x_i \in S_j} x_i \quad (4)$$

Where S_j is the set of data points assigned to centroid C_j .

The steps 1) and 2) are repeated iteratively and the goal is to minimize the criterion function. Supposing that the target object is x , and C_i is the centroid of cluster S_i , criterion function is defined as follows:

$$E = \sum_{i=1}^k \sum_{x \in S_i} \|x - C_i\|^2 \quad (5)$$

This process is repeated until the convergence. Convergence is typically defined by a condition where the centroids no longer change significantly between iterations or a maximum number of iterations is reached.

The k-means clustering algorithm always converges to local minimum. Once the algorithm converges, the process finishes and now each data point belongs to its cluster.[30][28] As the cluster centroids are initialized at random, the algorithm should be repeated with random initialization and the results should be combined, in order to obtain the best possible results.[22]

K-means algorithm is often used in combination with other algorithms to provide better accuracy in the specific task. K-means can effectively handle noise and outliers in the data.[42]. Also, the k-means clustering filtering strategy can effectively delete boundary samples, noise samples, and overlapping samples, and avoid the loss of important samples and the generation of new boundary samples.[49] K-means can be used as a preprocessing step for other machine learning algorithms. By classifying the data set through a certain number of clusters fixed apriori, we can reduce the dimensionality of the feature space.[41]

V. MEDICAL IMAGING

A. Evolution and Impact of Technology in Medical Imaging

Medical Imaging began in November 1895 with Wilhelm Conrad Roentgen's discovery of the X-ray[2]. As far as we are concerned, medical imaging is crucial to humanity regarding their well-being, from disease detection, lesion segmentation, diagnosis, treatment selection, response assessment, and clinical prediction. Computers entered the world of medical imaging in the early 1970s with the advent of computed tomography (CT) and then with magnetic resonance imaging (MRI)[2]. AI (Artificial Intelligence) with ML (Machine Learning) and DL (Deep Learning) methods can effectively assist clinical specialists in healthcare, to improve medical imaging understanding and in-depth analysis[38].

B. Medical Imaging consists of:

1) *X-rays* – W.C. Roentgen, 1895[7], Figure 7: **X-ray**, is one of the most commonly used techniques, where the image is produced by passing X-rays which are generated by a source through the body detecting the weakened X-rays on the other

side with a detector array[36]. The resulting image is a 2D projection with resolutions down to 100 microns. X-ray projection imaging has been ubiquitous in cardiovascular, mammography, musculoskeletal, and abdominal imaging applications among others.



Fig. 7. 29 y/o male with a thigh injury with an incidental finding on x-ray of enchondroma. Source: National Library of Medicine

2) *Nuclear medicine* - Cassen, 1951[7], Figure 8: **Nuclear medicine** relies on imaging gamma rays emitted from radioactive isotopes administered into the body. These isotopes emit radiation captured by an external camera to create images. Techniques such as Single Photon Emission Computed Tomography (SPECT - Kuhl, Edwards, 1963) and Positron Emission Tomography (PET – Ter-Pogossian, 1972)[36] compile these emissions into 3D images.

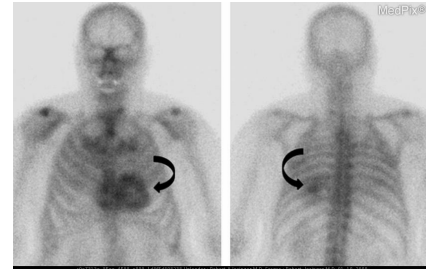


Fig. 8. Anterior and posterior planar images of the chest from a nuclear medicine bone scan (obtained 2 hours after injection of Tc-99m HDP) demonstrating abnormal radiopharmaceutical accumulation within the left ventricular walls (curved black arrows) of Nuclear Medicine Bone Scan Source: National Library of Medicine

3) *Ultrasound (US)* – Ian Donald, 1962[7], Figure 9: **Ultrasound imaging (US)**, which utilizes sound waves in the 1–10 MHz range to create images of internal tissues, works by measuring the echoes from pulses that scatter back from internal structures. Ultrasound is quick and capable of real-time imaging, such as monitoring blood flow in arteries. A key advantage is that it does not use ionizing radiation, making it safer for patients[36]. Ultrasound is widely used for real-time monitoring of heart activity and fetal development. Its application has grown to include 3D and 4D imaging, although, these methods may offer lower temporal resolution.

4) *Magnetic Resonance Imaging (MRI)* - Lauterbur, Mansfield, Hutchison, 1972[7], Figure 10: **Magnetic Resonance**

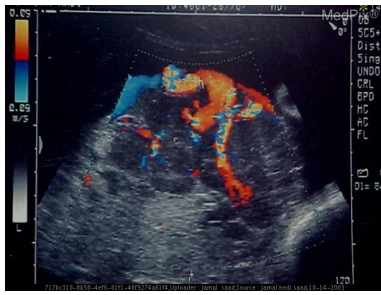


Fig. 9. A twenty-six-years-old female third gravida with a history of two prior cesarean sections was referred for a routine antenatal ultrasound scan at 20 weeks of gestation. Source: National Library of Medicine

Imaging (MRI) uses an external magnetic field and radio-frequency (RF) pulses to create detailed volumetric images, primarily of hydrogen nuclei, without the use of ionizing radiation. Widely used across various fields, MRI excels in musculoskeletal, cardiovascular, and neurological imaging due to its exceptional soft-tissue contrast[36].

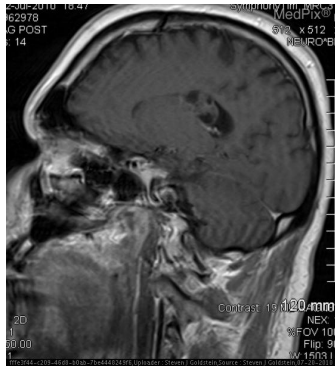


Fig. 10. 28-year-old man with a longstanding history of behavior problems and learning disability. The patient had an MRI scan done to evaluate a new onset (2 months) of headaches. Source: National Library of Medicine

5) *Computed Tomography (CT) - Hounsfield, 1973*[7], *Figure 11: Computed Tomography (CT)* like MRI, provides volumetric scans but constructs a 3D image from a series of 2D axial slices. Modern CT scanners, equipped with advanced solid-state detectors, have reached spatial resolutions as fine as 0.25 mm, and multiple detector rows allow broad spatial coverage with minimal slice thicknesses of 0.625 mm[36].

C. Advancements in Medical Imaging Through AI

Similar to human cognitive processes, algorithms analyze medical images to recognize patterns, having been trained with a large dataset of examinations and images[38]. These systems are capable of providing insights into the nature of abnormalities, primarily by offering conditional probabilities useful for Bayesian decision-making frameworks. Moreover, data obtained from radiomics (a rapidly evolving field of research concerned with the extraction of quantitative metrics within medical images) research, including aspects like intensity, shape, texture, and wavelength, can be extracted from medical images and utilized or incorporated into machine learning strategies. This information is crucial for predicting

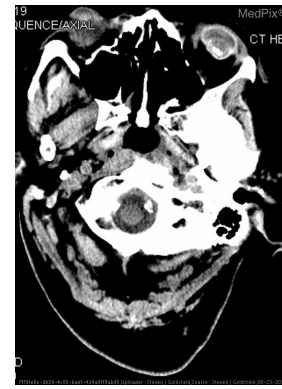


Fig. 11. 65-year-old man who had CT scan for stroke evaluation Source: National Library of Medicine

treatment outcomes, distinguishing between benign and malignant tumors, and evaluating the genetic aspects of various cancers[38].

VI. ELECTRONIC HEALTH RECORDS (EHR)

A. Origins of EHR

Ancient Egyptian hieroglyphic inscriptions from 1,600-3,000 BC indicate the use of medical records[8]. What is now most commonly referred to as the EHR started to enter clinical care not earlier than the 1960s[4]. A patient's electronic health record (EHR) can be viewed as a repository of information regarding their health status in a computer-readable form[18].

B. Challenges and Progress in the Adoption of Electronic Health Records

Initially, structured data entry methods were often superseded by free-text documents which were typically dictated by physicians. The adoption of Electronic Health Records (EHRs) has varied significantly across different regions and hasn't paralleled broader IT advancements[4]. Despite the global healthcare system not yet completely abandoning paper records, there has been a significant surge in the digitization of healthcare records over the past 5 to 10 years. Now, in some countries, nearly 90 percent of all healthcare records are maintained digitally thanks to the advancements in technology and the uprising of AI.

C. What does it take to achieve true digital transformation in healthcare?

EHR adoption is growing thanks to initiatives like the US 19 billion HITECH act10 in the United States and the 2 billion public-private partnership Innovative Medicines Initiative (IMI) 11 in the European Union. According to the Healthcare Information and Management Systems Society (HIMSS) Analytics, larger nations like the United States, Canada, and Germany fall behind smaller European countries such as Denmark and Sweden in achieving the highest level of paperless data sharing, storage, and decision support (Uwe Buddrus, HIMSS Analytics Europe, personal communication)[18]. Only a limited number of hospitals have implemented top-level

Electronic Health Record (EHR) systems that encompass all hospital operations, although, a significant portion have established IT infrastructures. These systems are capable of capturing valuable research data from various auxiliary clinical and administrative systems[18].

D. Overcoming Challenges in EHR Utilization with AI and Data Analytics

Due to high data volume, experts wishing to analyze EHR can face several obstacles such as scattered, heterogeneous data, as well as ethical and legal matters that limit access[4]. Currently, most clinical and basic research data are stored in separate and disparate systems, making it challenging for clinicians and researchers to access and share this information[4]. The crucial transformation needed to contribute to biomedical research and other critical fields like drug discovery is reliant on the availability of reliable and scalable EHRs for reuse. Innovative methods with the help of ML are being employed to derive meaning from these vast datasets. Data-driven approaches are used to identify statistical patterns in large, complex datasets, which are commonly found in EHR systems. In clinical settings, due to safety considerations, these models typically play a supportive role in treatment-related decision support as AI is not yet capable of accurate, error-free results[4]. This sets a goal for humanity to emphasize on digitalization as well as AI-driven methods to improve EHR analysis and support to clinical experts for an optimized and more practical healthcare system.

VII. APPLICATIONS OF MACHINE LEARNING IN HEALTHCARE

A. Diagnosis of Thyroid Disorders

1) What are Thyroid Disorders?:

*The **Thyroid** [34] is a small, butterfly-shaped gland in the front of the neck. Thyroid hormones control the way the body uses energy, affecting nearly every organ, including the heart. There are many types of thyroid disorders. Some of these are:[9]*

- I. **Hypothyroidism:** In this particular condition, the thyroid gland is unable to produce enough of some hormones, resulting in the slowing down of many body functions. The symptoms include, among others: increased levels of the hormone TSH, weight gain, slow heart rate, fatigue, and decreased metabolism.
- II. **Hyperthyroidism:** This condition is the exact opposite of hypothyroidism. The gland produces a much larger quantity of hormones than the body requires. This results in an acceleration of the metabolism. Some of the symptoms include decreased production of the hormone TSH, weight loss, sweating, increased heart rate, and elevated metabolism.
- III. **Hashimoto's disease:** Hashimoto's disease, also known as Hashimoto's thyroiditis (HT) [3], is a chronic, autoimmune disorder and the most common condition

affecting the thyroid gland. The most common symptoms [50] of HT include: neck pain, voice changes, discomfort in the throat, difficulty breathing, goiter, and fatigue.

IV. **Thyroid Cancer:** It is a disease in which cancerous cells appear in the gland tissues.[14] In the period from 2000 to 2020, according to the [European Thyroid Journal], there were 1387 cases in Europe, of which 1074 cases (77.43%) involved women. An individual may suspect they have thyroid cancer if they feel a hard lump in the front part of their neck, have a hoarse voice, have difficulty swallowing, and/or have shortness of breath, coughing, and weakness.[33]

2) Which factors can lead to Thyroid Disease onset?:

There are several factors that can lead someone to develop thyroid disease [32]. Women who have not undergone menopause are quite likely to develop thyroid cancer. Additionally, individuals suffering from other thyroid disorders, those consuming medications with high iodine content, as well as individuals undergoing radiation therapy to the head and/or neck, are also at risk of developing thyroid diseases.

3) How are they diagnosed?:

The diagnosis of such conditions is mainly performed through physical examination, ultrasound, and hormone and antibody tests via blood tests. However, in several instances, it can be quite a painful process if the patient needs to undergo fine-needle aspiration biopsies or even radioactive iodine uptake tests, where the patient needs to be isolated.

4) Contribution of machine learning to Thyroid Disorders:

[6] Traditional medical procedures are being revolutionized by machine learning, which is crucial in the identification of thyroid disorders. Machine learning algorithms can find patterns and connections in patient data that may be invisible to human clinicians by examining large databases of patient data, such as symptoms, medical histories, and test results. These algorithms help doctors make quick and precise diagnosis by predicting the likelihood of thyroid problems with outstanding accuracy. Furthermore, machine learning models have the capacity to learn from fresh data continuously, improving their diagnostic skills over time. This dynamic method improves patient outcomes in the management of thyroid disorders by enhancing diagnostic accuracy and enabling customized treatment recommendations for each patient.

B. Disclosure of Breast Cancer

1) What is exactly Breast Cancer?:

Type of cancer affecting the breast(s). Although it is a disease that is closely associated with women, it is worth noting that it also affects men.[47] It is important not to mistake all breast lumps for cancer. Most tumors that occur are benign such as fibroadenomas, phyllodes tumors and cysts. Types of breast

cancer [17] includes ductal and lobular carcinoma and sarcoma

2) Where does it come from?:

Factors [10] that can lead to breast cancer can be divided into 2 categories: immutable factors and modifiable factors that can serve as prevention. In the first category, age plays a significant role; individuals over 50 are more likely to develop breast cancer. Additionally, hereditary factors and family history, prior treatment with radiation therapy, and genetic mutations can also contribute to the development of breast cancer. In the second category, sedentary lifestyle, obesity, poor diet (especially postmenopausal for women), alcohol abuse, and smoking are included among other factors.

3) How are they diagnosed?:

The diagnosis [27] of breast cancer can be achieved through various methods. The most well-known is mammography, followed by ultrasound and biopsy. The latter is a rather painful procedure for the patient, as it involves the sampling of the tumor using a needle.

4) Machine learning for prediction and diagnosis of breast cancer:

Machine learning [40] is the most accessible way to diagnose breast cancer[15]. Some algorithms used are: Decision Tree, Naïve Bayes, Extreme Gradient Boosting, Logistic Regression, etc. Machine learning holds tremendous promise in the realm of breast cancer detection and treatment. Through the utilization of extensive data sets, machine learning algorithms are able to examine mammograms, genetic data, and patient medical history in order to detect patterns that may indicate breast cancer in its initial stages. Additionally, by helping radiologists analyze images more precisely, these algorithms can increase the accuracy of diagnoses. Additionally, by evaluating patient data to forecast the best courses of action and possible results, machine learning models can tailor treatment regimens, improving patient care and survival rates in the process. The field of breast cancer detection and therapy is undergoing a revolution thanks to the continuous research and innovation in machine learning, which is opening up new possibilities for early detection, accurate diagnosis, and customized treatment plans.[23]

C. Forecasting Disease Progression in Multiple Sclerosis

1) What is Multiple Sclerosis?: Multiple Sclerosis (MS)

[11] is a long-term autoimmune disease that affects the central nervous system. **Myelin** is a protein that surrounds and protects nerve fibers. The loss of myelin leads to the formation of scar tissue, known as *sclerosis*. Consequently, the nerve fibers are left unprotected and exposed. As a result, they cannot transmit electrical impulses to the brain normally, and they can be easily damaged or even die completely, leading to disability.

2) What causes it?:

Until today, the causes of being affected by the disease are not known. However, the most prevalent reasons include a

combination of environmental and hereditary factors as well as microbial, viral, and bacterial infections. According to the World Health Organization, data indicate that women are more susceptible to the disease. Also, regarding statistical data on the map of Europe, there are more than 700,000 people suffering from MS.

3) What are the symptoms?:

The most common symptoms of the disease include numbness of the limbs, difficulty, and instability in walking, double or blurred vision, as well as dizziness and vertigo.

4) How it is diagnosed?:

The diagnosis process of the disease is initially based on the patient's medical history and continues with specialized neurological examinations, magnetic resonance imaging (MRI) scans, and a more discomforting examination for the patient: the lumbar puncture for the analysis of cerebrospinal fluid.

5) How machine learning can help patients?:

[39] In the field of multiple sclerosis, machine learning assists in the diagnosis and treatment of the disease. Diagnosis is performed by examining magnetic resonance imaging (MRI) scans. Specialized machine learning algorithms (the description of these was given above - Section IV) take these scans as input, process them, and accordingly extract the results. The results yielded by these specific algorithms include, in addition to the initial diagnosis, the stage at which the patient is currently at, and they can also, in some way, "predict" the progression of the disease.

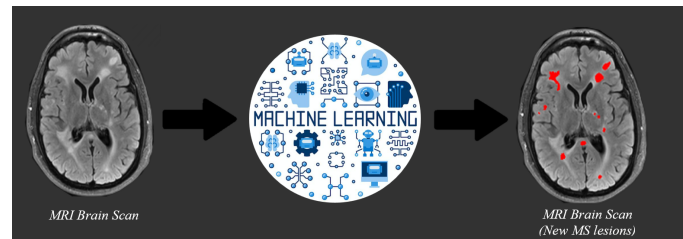


Fig. 12. Identification of New Multiple Sclerosis Lesions in the Brain Using Machine Learning Algorithms.

Source: <https://mostafasalem.netlify.app/publication/phdthesis/>

VIII. CONCLUSION

In conclusion, this paper has highlighted the role of Artificial Intelligence (AI) and machine learning in revolutionizing healthcare, from their early developments to their potential future impact. These technologies enhance patient care by improving the accuracy and personalization of diagnosis, treatment, and illness management. While AI and machine learning bring significant advancements, they also introduce challenges such as ethical concerns and the need for robust regulations. Despite these difficulties, the ongoing integration of these technologies into healthcare promises to make it more effective and patient-centered, provided that their development and implementation are carefully managed and overseen. In the

future, continued advancements in AI and machine learning are expected to further empower healthcare professionals by providing more comprehensive and real-time data analysis, potentially leading to breakthroughs in predictive medicine and patient care optimization.

REFERENCES

- [1] Natalie Banner, Alison Hall, Johan Ordish, Sobia Raza, Brent Mittelstadt, Ben Moody, and Reema Patel. Artificial intelligence (ai) in healthcare and research. 2018.
- [2] William G Bradley. History of medical imaging. *Proceedings of the American Philosophical Society*, 152(3):349–361, 2008.
- [3] P. Caturegli, A. De Remigis, and N.R. Rose. Hashimoto thyroiditis: Clinical and diagnostic criteria. *Autoimmunity Reviews*, 13(4):391–397, 2014. Diagnostic criteria in Autoimmune diseases.
- [4] Pascal Coorevits, Mats Sundgren, Gunnar O Klein, Anne Bahr, Brecht Claerhout, Christel Daniel, Martin Dugas, Danielle Dupont, Andreas Schmidt, Peter Singleton, et al. Electronic health records: new opportunities for clinical research. *Journal of internal medicine*, 274(6):547–560, 2013.
- [5] D’souza, Kevin Joy, Ansari, and Zahid. Big data science in building medical data classifier using naïve bayes model. pages 76–80, 2018.
- [6] Priyanka Duggal and Shipra Shukla. Prediction of thyroid disorders using advanced machine learning techniques. pages 670–675, 2020.
- [7] Alex Elliott. Medical imaging. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 546(1-2):1–13, 2005.
- [8] R Scott Evans. Electronic health records: then, now, and in the future. *Yearbook of medical informatics*, 25(S 01):S48–S61, 2016.
- [9] P. A. Farling. Thyroid disease. *BJA: British Journal of Anaesthesia*, 85(1):15–28, 07 2000.
- [10] Centers for Disease Control and Prevention. What are the risk factors? 2023.
- [11] Nikzad E. Ghasemi N, Razavi S. Multiple sclerosis: Pathogenesis, symptoms, diagnoses and cell-based therapy. 2016.
- [12] Gupta, Vratika, Mishra, Vinay Kumar, Singhal, Priyank, Kumar, and Amit. An overview of supervised machine learning algorithm. pages 87–92, 2022.
- [13] Haraty, Ramzi A., Dimishkieh, Mohamad, Masud, and Mehedi. An enhanced k-means clustering algorithm for pattern discovery in healthcare data. *Int. J. Distrib. Sen. Netw.*, 2015, jan 2015.
- [14] National Cancer Institute. Thyroid cancer screening (pdq®)—patient version. 2023.
- [15] Md. Milon Islam, Md. Rezwanul Haque, Hasib Iqbal, Mahmudul Hasan Md. Munirul Hasan, and Muhammad Nomani Kabir. Breast cancer prediction: A comparative study using machine learning techniques. 2020.
- [16] Jadhav, Sayali D, Channe, and HP. Comparative study of k-nn, naïve bayes and decision tree classification techniques. *International Journal of Science and Research (IJSR)*, 5(1):1842–1845, 2016.
- [17] Baltimore Jenni Sheng, MD Johns Hopkins University School of Medicine and MD This information is provided by Breastcancer.org. Types of breast cancer. 2024.
- [18] Peter B Jensen, Lars J Jensen, and Søren Brunak. Mining electronic health records: towards better research applications and clinical care. *Nature Reviews Genetics*, 13(6):395–405, 2012.
- [19] Khushi Kumari Jha, Roshan Jha, Ankita Kumari Jha, Md Al Mahedi Hassan, Saurav Kumar Yadav, and Tr Mahesh. A brief comparison on machine learning algorithms based on various applications: A comprehensive survey. pages 1–5, 2021.
- [20] Kamel, Hajer, Abdulah, Dhahir, Al-Tuwaijari, and Jamal M. Cancer classification using gaussian naïve bayes algorithm. pages 165–170, 2019.
- [21] Upendra K Kar. The future of health and healthcare in a world of artificial intelligence. 2018.
- [22] Khalid, Sara, Judge, Andrew, Pinedo-Villanueva, and Rafael. An unsupervised learning model for pattern recognition in routinely collected healthcare data. 2018.
- [23] Arslan Khalid, Arif Mehmood, Amerah Alabrah, Bader Fahad Alkhamees, Farhan Amin, Hussain AlSalman, and Gyu Sang Choi. Breast cancer detection and prevention using machine learning. *Diagnostics*, 13(19), 2023.
- [24] Khan, Hera, Srivastav, Ayush, Mishra, and Amit Kumar. Use of classification algorithms in health care. pages 31–54, 2020.
- [25] Kharya, Shweta, Soni, and Sunita. Weighted naïve bayes classifier: a predictive model for breast cancer detection. *International Journal of Computer Applications*, 133(9):32–37, 2016.
- [26] Kushwaha, Pradeep Kumar, Kumaresan, and M. Machine learning algorithm in healthcare system: A review. pages 478–481, 2021.
- [27] M.D. LECIA M. APANTAKU. Breast cancer diagnosis and screening. 2000.
- [28] Li, Youguo, Wu, and Haiyan. A clustering method based on k-means algorithm. *Physics Procedia*, 25:1104–1109, 2012.
- [29] Thomas Rincy N, Gupta, and Roopam. A survey on machine learning approaches and its techniques:. pages 1–6, 2020.
- [30] Na, Shi, Xumin, Liu, Yong, and Guan. Research on k-means clustering algorithm: An improved k-means clustering algorithm. pages 63–67, 2010.
- [31] Samreen Naeem, Aqib Ali, Sania Anam, and Munawar Ahmed. An unsupervised machine learning algorithms: Comprehensive review. *IJCDS Journal*, 13:911–921, 04 2023.
- [32] National Health Service (NHS). Causes of thyroid cancer. 2023.
- [33] National Health Service (NHS). Symptoms of thyroid cancer. 2023.
- [34] National Institute of Diabetes, Digestive, and Kidney Diseases (NIDDK). Hypothyroidism (underactive thyroid). 2021.
- [35] Omankwu, Obinnaya Chinecherem, Osodeke, Efe Charlse, Ubah, and Valetine Ifeanyi. Artificial intelligence a paradigm shift in healthcare sector past, present, future prospects. 5(3), 2023.
- [36] Andreas S. Panayides, Amir Amini, Nenad D. Filipovic, Ashish Sharma, Sotirios A. Tsiftaris, Alistair Young, David Foran, Nhan Do, Spyretta Golemati, Tahsin Kurc, Kun Huang, Konstantina S. Nikita, Ben P. Veasey, Michalis Zervakis, Joel H. Saltz, and Constantinos S. Pattichis. Ai in medical imaging informatics: Current challenges and future directions. *IEEE Journal of Biomedical and Health Informatics*, 24(7):1837–1857, 2020.
- [37] Pecioski, Damjan, Gavriloski, Viktor, Simona Domazetovska, Ignjatovska, and Anastasija. An overview of reinforcement learning techniques. pages 1–4, 2023.
- [38] Filippo Pesapane, Marina Codari, and Francesco Sardanelli. Artificial intelligence in medical imaging: threat or opportunity? radiologists again at the forefront of innovation in medicine. *European radiology experimental*, 2:1–10, 2018.
- [39] Mauro F. Pinto, Hugo Oliveira, Sónia Batista, Luís Cruz, Mafalda Pinto, Inês Correia, Pedro Martins, and César Teixeira. Prediction of disease progression and outcomes in multiple sclerosis with machine learning. 2020.
- [40] Rabiei, R.and Ayyoubzadeh, S. M., Sohrabei, S.and Esmaeili, M.and Atashi, and A. Prediction of breast cancer using machine learning approaches. *Journal of biomedical physics engineering*, 2022.
- [41] Santhanam, T, Padmavathi, and MS. Comparison of k-means clustering and statistical outliers in reducing medical datasets. pages 1–6, 2014.
- [42] Santhanam, T, Padmavathi, and MS. Application of k-means and genetic algorithms for dimension reduction by integrating svm for diabetes diagnosis. *Procedia Computer Science*, 47:76–83, 2015.
- [43] R. Saravanan and Pothula Sujatha. A state of art techniques on machine learning algorithms: A perspective of supervised learning approaches in data classification. pages 945–949, 2018.
- [44] Saritas, Mucahid Mustafa, Yasar, and Ali. Performance analysis of ann and naïve bayes classification algorithm for data classification. *International journal of intelligent systems and applications in engineering*, 7(2):88–91, 2019.
- [45] Silvana Secinaro, Davide Calandra, Aurelio Secinaro, Vivek Muthurangu, and Paolo Biancone. The role of artificial intelligence in healthcare: a structured literature review. (125), 2021.
- [46] Mohammed Yousef Shaheen. Applications of artificial intelligence (ai) in healthcare: A review. 2021.
- [47] American Cancer Society. What is breast cancer? 2021.
- [48] Xu, Wen, Tang, Jian, Xia, and Heng. A review of semi-supervised learning for industrial process regression modeling. pages 1359–1364, 2021.
- [49] Xu, Zhaozhao, Shen, Derong, Nie, Tiezheng, Kou, Yue, Yin, Nan, Han, and Xi. A cluster-based oversampling algorithm combining smote and k-means for imbalanced medical data. *Information Sciences*, 572:574–589, 2021.
- [50] Jiaojiao Yuan, Shuo Qi, Xufan Zhang, Hezheng Lai, Xinyi Li, Chen Xiaoheng, Zhe Li, Simiao Yao, and Zhiguo Ding. Local symptoms of hashimoto’s thyroiditis: A systematic review. *Frontiers in Endocrinology*, 13, 2023.
- [51] Ahmet İlker Tekkeşin. Artificial intelligence in healthcare: Past, present and future. 2019.