

Review of Data Analytic Techniques for Health Data

Dr. Vaishali Pawar¹ and Radhika Chandwadkar²

¹⁻²MVPS's KBT COE, Nashik, Maharashtra, India

Email: hod.comp@kbtcoe.org, chandwadkar.radhika@kbtcoe.org

Abstract—Recent years have experienced a tremendous explosion in digital healthcare data. The field of data analytics has also emerged with advance technologies of machine learning and deep learning. It has boosted the data driven medical research based on data analytics. Deep learning has gained importance in the field of machine learning due to its techniques of feature representation and pattern recognition. Health data is obtained in various forms, from varied sources, contexts, technologies, and their nature can affect the proper analysis. These variations and differentiations needs to be managed for meaningful analysis. Such health data needs aggregation and integration for the effective application of data analytic techniques. In this review we survey the challenges in data analytics, sources, types of medical data and the techniques that can be applied for the effective diagnosis. Various machine learning techniques and deep learning models are also reviewed for their effective and efficient application in the health care domain.

Index Terms— Deep learning, Big data, Electronic Health Record (EHR), Machine learning.

I. INTRODUCTION

Healthcare is a data-intensive field. Digital clinical data are increasing rapidly in recent years, with the massive accumulation of data in medical institutions. This leads to generation of big data and provides opportunities for healthcare research. The key attributes of volume, velocity, and variety make big data a demanding field for the latest innovations. Data in the health care domains is collected from different platforms and sources that too in varied forms which are widely known. The Electronic Health Record system (EHR) stores the data associated with each patient including demographic information, diagnoses, laboratory tests and results, prescriptions, radiological images, clinical notes, and more. The patient data contained in EHR systems is used for tasks as medical concept extraction, patient trajectory modelling, disease inference, clinical decision support systems and more[1]. The existing hospital informatics systems is not adequate for performing data analysis. Hence there is a need to develop an integrated analytical system of medical data that will aggregate data and provide insight to the physicians based on the analysis of the data in real time. Hence it is essential to study the relationship of the big data and the diseases using a data driven approach and identify the risk factors related to diseases.

A. Motivation

In the previous years the health data sources were limited to classical testing equipment and techniques such as Electrocardiogram (ECG) mammography, Magnetic Resonance Imaging (MRI), ultrasound, CT scanners, and much other testing equipment. All the analytical tasks needed to be performed manually by the physician. The traditional software applications were used to perform tasks like patient admission in hospital, billing of patient where as the diagnostic tasks were performed by the physicians by probable conclusions. So with this need

computational intelligence has to play an active role. It is now very difficult to aggregate all health data in a unified manner due to its varied nature and description. Also the interpretation of health data and the process of deriving inferences from it require additional analytic solutions for a specific dataset. Hence an integrated healthcare solution is essential that will consolidate data from EHR, additional data sources that are in use today and apply special analytics for quick and accurate diagnosis of the patient health. The usual algorithms of decision tree and K-means are effective only in deciding whether a patient should be admitted in hospital and lack critical analysis. In-depth analysis thus requires advance machine learning and deep learning techniques. The choice of a technique for a given health instance requires a thorough analysis. In this review, we provide options to choose from, depending on the health predicament.

II. RELATED WORK

Over the last few years, most analyses of medical data have used traditional statistical methods. The study of health care data previously focused on the data mining techniques. W. Raghupathi and V. Raghupathi, presented the pioneer review in the field of big data analytics in health care[2]. It covered potentials of big data analytics but lacks technical aspects of big data analytics in health care. M. Herland presented a holistic view of health care analytics by focussing on data mining approaches [3]. Later the study of deep explanation of data mining algorithms to mine health data provided a concise analysis of how each data mining technique is applied to various diseases.[4].In the further studies with advances in machine learning techniques statistical techniques such as logistic regression, support vector machines (SVM), and random forests were applied to health data.[5]. These methods have achieved some reliable results across many subfields of medical studies. In recent years, deep learning has been under vigorous development and has achieved great success in many competitions, [6] due to its superior performance in learning the feature representation from raw data in an end-to-end way. Hidden layers in the neural network architecture are responsible for capturing the nonlinear relationships; hence for complex analysis more number of layers can be added. In combination with the large volume of data and Graphical Processing Units (GPU), deep learning makes it possible to explore clinical data in the direction of precision medicine

III. PROBLEM DOMAIN

Medical field has very large amount of data that requires processing for accurate prediction. This medical data exists in various forms hence the processing of these different forms of data is a major challenge. The paper studies the recent machine learning and deep learning techniques applicable to the medical data.

ELECTRONIC HEALTH RECORD (EHR)

Electronic health record system was basically designed for managing the hospital administration. Medical data has various terminologies and vocabularies which need to be mapped with standard conventions for uniformity across the EHR of different hospital for appropriate integration of data. United Medical Language System (UMLS) and the Systemized Nomenclature of Medicine - Clinical Terms (SNOMED CT) are the standard medical resources that specifies the mapping of medical terms.EHR in the hospital is mainly used to store the patient's information like clinical notes of diagnosis, prescribed medication, lab reports and medical investigation reports.EHR usage at hospitals and clinics has the potential to improve patient care by minimizing errors, increasing efficiency, and improving care coordination .They also provide a rich source of data for researchers.EHR systems vary in terms of functionality, and are categorized into basic EHR without clinical notes, basic EHR with clinical notes, and comprehensive systems. One of the challenges in working with EHR data is the heterogeneous nature by which it is represented, with data types like: **(1) numerical quantities** like BMI, **(2) datetimeobjects** like birth date or admission date **(3) categorical values** like the medical terms indication of patient disease that uses codes from controlled vocabularies like ICD-10 diagnoses **(4) natural languagefree-text** such as progress notes or discharge summaries. Additionally, these data types can be ordered chronologically to form the basis for **(5) derived time series** such as pre operative vital sign signals or multimodal patient history [7]

The goals of designing an integrated health analytics is to provide correct, accurate and timely recommendation for practical diagnosis. The system should also be able to predict possible

Medical aggravations that might occur. Kuo *et al.* [8] considered that the challenges in big data are as per stages of big data pipeline so four distinct stages are considered. Each stage has its own challenges to overcome. The following challenges at each stage are considered:

A. Data Aggregation

The challenge of data aggregation is due to the high volume and variety of data from various sources like data ware houses. The solution can be the use of high-speed file transfer technologies. Another solution that is presented to deal with data aggregation challenge is data compression.

B. Data storage and maintenance

Storage of voluminous health big data is one of the challenges. Growth rate of the health data is exponential and hence the traditional database management systems are incapable to manage it. Solution for the problem is the usage of database systems that are non SQL like MongoDB, Cassandra and Hadoop Distributed File System (HDFS).

Cloud computing is also a powerful solution for the storage of vast amount of data. Wan and Sankaranarayanan[9] proposed a cloud-based EHR that can help all stakeholders to access the EHR by making use of three cloud computing components: Software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). The system allows mobile access for all stakeholders

C. Data integration and interoperability

The health big data are heterogeneous. They are obtained from many sources with divergent forms and structure hence making interoperability is a big challenge. Even the EHR which is the most structured of all data source can present a challenge, especially when more than one EHRs are involved. A patient may be treated at different hospitals, which may leave his health data scattered across various EHRs. Interoperability problems between hospitals still pose a big barrier for systems integration.

D. Data Analysis

Depending on the complexity of a health problem the traditional SQL based querying time increases exponentially as the number of records increases. The hardware and software needed to analyze data need to be so robust and expensive to get a meaning from the huge dataset. In terms of hardware Supercomputers and cloud computing are the most widely used. The motive of data analysis is to apply a predictive model for the prediction of complications.

Another big challenge to be considered is confidentiality as not many hospitals permit their records to be accessed by others. To solve the compatibility problems, various studies have been conducted and one of the most trusted technologies is the Blockchain [10]. The Blockchain technology solves the challenge of personalization and ubiquitous access to records which are the two traits that are required in the healthcare industry. Blockchain ensures data integrity by allowing an immutable way to update records and prevents the tampering with them once they have been logged.

IV. PROBLEM DEFINITION

Health big data has number of challenges as already stated to overcome these challenges ,the paper aims to study different data analytic techniques applicable for the medical data .

CLINICAL DATA

Characteristics of clinical data

An Electronic Health Record (EHR) of a patient includes many kinds of clinical data and plays an important role in medical research. In recent years, intelligent wearable devices have arisen to collect personal daily health data (such as blood sugar, blood pressure, and heart rate), which further accelerates data growth. We are witnessing clinical data growing at an unprecedented rate, accompanied by a growing variety of data sources (structured relational forms, unstructured images or text, semi-structured records). At the same time, if some real-time data is produced it also needs to be retrieved and analyzed immediately for timely decision making. It is obvious that clinical data has the “4Vs” characteristics of big data: big volume, large data variety, high velocity of data generation/update, and high veracity leading to big value.

Different types of clinical data

Clinical data is quite heterogeneous. When a patient undergoes the medical treatment then the information related to the patient's health status and treatment leads to generation of health data. It has varied forms from images to lab reports. Following are the types of data: [11]

A. Medical images

With the development of medical imaging techniques, there is much medical image data being generated by X-rays, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Optical Coherence Tomography (OCT), microscopy image, and Positron Emission Tomography (PET). These medical images are the direct sources of raw data which reflects patient's health condition.

B. Clinical notes

Text data that is in the form of patient diagnosis, physician's clinical notes regarding patient health status, lab reports, medication details and discharge summaries, in case of hospitalization, forms another category of health data.

C. Other types

Apart from medical images and clinical notes, the other types of data that are used for analysis are - Physiological measurements, socio-economic information, bill payment details and insurance related information. This data is structured and is hence stored in two-dimensional relational tables. The diverse lab items have different units and different reference ranges of normal values (considering age and gender). Hence normalization of the data is essential before performing the analysis of lab test results

V. INNOVATIVE CONCEPTS

DATA ANALYTIC TECHNIQUES

The major categories of machine learning technique are supervised and unsupervised learning. In supervised learning the target result is known and the learning process is guided to achieve the target. Logistic regression and support vector machine are the examples of supervised learning. Where as in unsupervised learning the process of learning is not guided, it is self-learning. Clustering and density estimation are the examples of unsupervised learning. Machine learning techniques are vital for predicting disease occurrences or their complications. Though all machine learning algorithms and practices can be applied to healthcare problems, each illness and its complications are best described by a single algorithm or a combination of some of them. Hence, it is necessary to inspect the algorithms closely and apply them where they are appropriate. The following algorithms are widely used in health informatics. We briefly described them as well as their specific use in diseases diagnosis.

A. K nearest neighbour algorithm

K-nearest neighbour (KNN) [11] is one of the simplest and most classical machine learning techniques. This algorithm is effective in classifying a disease by matching it with the already known cases and the resulted complications. For example -to determine if intestinal cells are cancerous, we need to classify intestinal cells in five possible classes of tumor cells: adenocarcinoma, sarcoma, carcinoid tumor, gastrointestinal tumor, and lymphoma. To classify cells under test, we can use Principal Components Analysis, a statistical feature that has more variance on data, and plots the known cell classes.

B. Support Vector Machine (SVM)

SVM [12] is an effective classifications algorithm that is used to classify data. SVM splits data using a hyperplane, which is a plane that separates the nearest points that are known as support vectors. Choosing the hyperplane is a constrained optimization problem as we must optimally choose the margin between classes that is wide enough. A hyperplane is a plane that is chosen to divide the classes, and sometimes, the data cannot be separated linearly. If such problems arise we can combine the SVM with additional kernel techniques like Radial Basis Functions. Brown *et al.* [13] introduced a method to classify genes using gene expression data by applying SVM techniques. The purpose of the experiment was to check the nature of the genes, whether active or silent. Khedheret *et al.* [14] has proposed a method for early diagnosis of Alzheimer's disease using SVMs on segmented MRI images. Principal Component Analysis was used to perform feature selection and SVM was used to classify if a certain patient is suffering from cognitive problems of old age or is developing the Alzheimer's disease.

C. Neural networks

Neural networks [15] are algorithms are a better choice to be applied on unstructured data for analyzing information. Health data analytics needs algorithms which receives inputs from varied sources and extract meaningful information. Medical unstructured data comprise doctors' notes, radiology scans, MRI images, microscopy, CT scans, ultrasound images, and so on. Interpreting these data using neural networks is an

excellent technique. Neural networks are a series of neuron-like layers of computation that apply a chain of computing algorithms to the input data computation cell to produce outputs. A Convolution Neural Network (CNN) is comprised of many layers that are capable of transforming inputs by applying convolution filters to produce a clear output. Neural networks are currently used for medical imaging, such as brain lesions analysis, fetal imaging, and cardiac analysis. To understand neural networks in the context of healthcare, consider its application in treating or predicting if a given breast tumor is malignant or benign.

Another key area of the neural network application is matching medical text reports with medical images, which is critical as we can interpret one using another. Kooiet *al.* [16] compared the traditional CAD-based mammography system, which relies on manual features, and a CNN-based mammogram. Another interesting application of neural networks and brain activity scans is the prediction of survivability using MRI images.

D. K-Means clustering techniques

The K-means clustering technique is one of the most popular unsupervised learning techniques [17]. It is a clustering algorithm that is used in classifying points on a Euclidian plane into categories. It is an iterative technique that takes several n points and classifies them into k possible clusters around k centroids. Gash and Eisen [18] used fuzzy k-means clustering and identified overlapping clusters of yeast genes. Fuzzy k-means allows a given point to belong to every cluster; however, with varying degree of membership. K-means algorithm can be applied to perform medical images segmentation. The method combined k-means with an improved watershed algorithm with k-means taking 2D MRIs as inputs and producing clustered images. A hybrid of k-means and SVMs can be applied for breast cancer diagnosis, in which, k-means can be used for identifying various hidden patterns of benign and malignant tumors.

E. Ensemble Learning

Medical data is very complex and hence single algorithm is not sufficient for generating informative results. Hence the combination of several learning techniques into a more effective and robust learning technique better than the constituent algorithms is done.

An ensemble [19] is a collection of various learning algorithm working together in parallel or in sequence to produce better results. It is mainly used in classification problems like sentiment classification. Ensemble learning has shown to provide very efficient and favourable outcome compared to individual learning systems as its usually perceived as a process of consulting multiple experts before deciding. The starting point of an ensemble process is feeding the training data to a base algorithm then voting is used to make an accurate decision from classification outcomes of each classifier in the ensemble. The most popular ensemble learning techniques are; Bagging and Boosting. During the bagging process we take the training data and divide it into bags (or subsets). We then apply an individual classifier on each subset and finally, we apply voting to produce a final prediction. Boosting is rather dependent on refining wrong predictions. We divide the training set as we do for bagging, we apply a learning algorithm to a subset, we test the classifier and check the wrong predictions. We take these cases and we combine them with a new subset and we apply another classifier and so on until we finish all the subsets. Boosting produces extremely refined decisions.

F. Markov Decision Process (MDP)

Markov Decision Process (MDP) [20] is a powerful stochastic control algorithm used in decision support. Even though the physician is presented with appropriate algorithms for a certain health case, there are some stochastic problems that are dependent on critical operations that need a decision support system. Some decisions such as a transplant or even the diagnosis itself will need mathematical models for accurate decisions to be made.

At each stage of the treatment, a patient's health state is analyzed as a Markov state and death or other complications are considered as terminal nodes in this MDP. The events that influence a patient's health are modeled as transitions between nodes. By building a Markov state diagram, critical decisions will be made and decisions which lead to catastrophic results will be omitted. Hence, MDP can be incorporated into a health analytic solution for decision making purposes.

VI. DEEP LEARNING MODELS

The central idea in deep learning is that of *representation*. Deep learning techniques learn optimal features directly from the data itself, without any human guidance. Deep learning models allow the discovery of hidden relationships in the data. Complex data representation is possible due to its expression as compositions of other, simpler representation in deep learning.

The vast majority of deep learning algorithms and architectures are built upon the framework of the *artificial neural Network* (ANN). ANNs are composed of a number of interconnected nodes (neurons). Neurons not contained in the input or output layers are referred to as hidden units. Every hidden unit stores a set of weights W which are updated as the model is trained. Optimization of the weight values in ANN is achieved by minimizing a loss function.

In recent years, deep learning methods have been highly effective big data analyses. Deep learning involves a wide variety of approaches and has great scope in the field of medical research. In this section, we briefly introduce the most common deep learning models that have been used to analyze medical data.

A. Multilayer Perceptron (MLP)

A multilayer perceptron is a type of ANN composed of multiple hidden layers, where every neuron in a layer is fully connected to every other neuron in next layer. There also exist few hidden layers in the network. Data flow in the network is only in one direction. During the training phase of ANN the weights in hidden layer are adjusted and optimized which helps to understand the association between input and output. If the number of hidden layers is increased then the data representation will be in abstract manner.

B. Deep neural networks

A Deep Neural Network (DNN) can be defined as Artificial Neural Network (ANN) with multiple hidden layers. Each hidden layer can be seen as a process of feature extraction. Features from the lower layers are composed in the higher layers so it can better model the complex data with few units. Due to this DNN can model complex non-linear relationships using the training data. To solve different problems, multiple variants and extensions have been developed[21].

C. Convolution neural networks

Convolution neural networks (CNN) have gain much importance in the field of image processing. CNN are better option of analysis when the input data has a *spatial* structure as the filters are smaller than the input due to less parameters. They support parameter *sharing* as every filter is applied across the entire input. In a CNN, a *convolution layer* is a number of convolution filters, all receiving the same input from the previous layer, which ideally learn to extract different lower level features. After the convolutions to aggregate the extracted features, a *pooling* or *subsampling* layer is applied. (CNNs) are also known as shift invariant and space invariant ANNs, which are based on their shared-weights convolution kernels and translation invariance characteristics [22]. They are widely used in computer vision and NLP. CNNs have greatly promoted medical image research for imaging, diagnosis, image segmentation, and Electrocardiograph (ECG) monitoring. CNNs and their variants can capture overall structural features from feature-rich images.

D. Recurrent neural networks

A Recurrent Neural Network (RNN) is a kind of deep feed-forward ANN. The connections between the units of an RNN form a directed graph along a sequence. RNNs are an appropriate choice when data is *sequentially* ordered (such as time series data or natural language). RNNs are better option while dealing with long-range temporal dependencies. Standard RNNs are made up of interconnected hidden units, such that each unit in a gated RNN is replaced by a special cell that contains an internal recurrence loop and a system of gates that controls the flow of information. RNNs have taken a central place in sequence data processing despite the gradient vanishing problem generated by long input sequences. A variant of RNN called a Long Short-Term Memory (LSTM) has been proposed to solve the problem by introducing gate mechanism and memory cell[23]. In addition, Gated Recurrent Unit (GRU) is a simplified version of LSTM. As compared with the plain model, these two variants achieve a better performance for long sequence processing. RNN is mainly used to solve sequence related problems, such as sentence classification, speech recognition, and machine translation.

E. Restricted Boltzmann machines

Restricted Boltzmann Machine (RBM) is a type of ANN that is generative and stochastic in nature. It can learn the probability distribution over a set of inputs. It is an unsupervised learning model and designed for feature extraction in unlabeled training data. Feature extraction by an RBM is helpful for the original classification problem. RBMs are viewed as generative models that model the underlying process by which the data was generated. All units are fully connected in standard Boltzmann machine (BM), where as in RBM there are no connections between any two units either visible or hidden. For training the algorithm performs Gibbs sampling to compute weight updates and achieves stochastic optimization. RBMs can be hierarchically stacked to form a deep belief network (DBN) for supervised learning tasks.

F. Deep auto encoders

One of the deep learning models exemplifying the notion of unsupervised representation learning is the auto encoder (AE). They were the early tool to pertain supervised deep learning models and were used for scarce labeled data that retain the usefulness of unsupervised tasks. Auto encoders are designed to *encode* the input into a lower dimensional space z . The encoded representation is then *decoded* by reconstructing an approximated representation $\tilde{x}$ of the input x [24]. Once an AE is trained, a single input is fed through the network, with the innermost hidden layer activations serving as the input's encoded representation. It is used to transform the input data in a format that stores the important derived dimensions. In this manner, they are similar to standard dimensionality reduction techniques like principal component analysis (PCA) and singular value decomposition (SVD), but with a significant advantage for complex problems due to nonlinear transformations via each hidden layer's activation functions. Deep AE networks can be constructed and trained in a greedy fashion by a process called stacking. Many variants of AEs have been introduced, including de noising auto encoders (DAE), sparse auto encoders (SAE), and variation auto encoders (VAE)

The intention of deep learning is representation, which means representing the features of input data effectively by multilayer neural network training. The deep learning methods mentioned above have made great progress in the study of clinical data. Moreover, the combination of models will bring more opportunities for effective analysis.

VII. CONCLUSION

The paper provides a review of various data analytic techniques from the field of machine learning and deep learning. This provides a platform to select a technique for the health care application as health care analysis requires a combination of various technologies. A successful analysis requires incorporation of varied medical data from different sources and the selection of the appropriate technology. For example - effective diagnosis using image data requires medical images to be closely analysed hence deep learning algorithms are better. Convolution Neural Networks (CNN) applies optimization algorithms like the Gradient Descent provides an effective analysis of spots on medical images. For data mining single algorithm may not provide solution hence ensemble learning, that includes many machine learning algorithms can provide a better analysis. The review of the techniques provides an insight to the researchers regarding the opportunities and challenges involved in developing a precise, accurate and robust application for the health care analysis.

REFERENCES

- [1] Gaspard Harerimana, Beakcheol Jang, JongWook Kim, and Hung Kookpark, "Health Big Data Analytics: A Technology Survey", IEEE ACCESS, Digital Object Identifier 10.1109/ACCESS.2018.2878254
- [2] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: Promise and potential," *Health Inf. Sci. Syst.*, vol. 2, no. 1, p. 3, 2014.
- [3] M. Herland, T. M. Khoshgoftaar, and R. Wald, "A review of data mining using big data in health informatics," *J. Big Data*, vol. 1, no. 1, p. 2, 2014
- [4] Fang, S. Pouyanfar, Y. Yang, S.-C. Chen, and S. S. Iyengar, "Computational health informatics in the big data age: A survey," *ACM ComputSurv.*, vol. 49, no. 1, p. 12, 2016.
- [5] N. Barakat, A. P. Bradley, and M. N. H. Barakat, "Intelligible support vector machines for diagnosis of diabetes mellitus," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 1114_1120, Jul. 2010.
- [6] Benjamin Shickel, Patrick James Tighe, Azra Bihorac, and Parisa Rashidi, "Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis", IEEE JOURNAL OF BIOMEDICAL AND HEALTHINFORMATICS, VOL. 22, NO. 5, SEPTEMBER 2018
- [7] M.-H. Kuo, T. Sahama, A. W. Kushniruk, E. M. Borycki, and D. K. Grunwell, "Health big data analytics: Current perspectives, challenges and potential solutions," *Int. J. Big Data Intell.*, vol. 1, nos. 1_2, pp. 114_126, 2014.
- [8] T. Wan and S. Sankaranarayanan, "Development of a Health Information System in the Mobile Cloud Environment," in *Proc. IEEE Int. Conf. Embedded Ubiquitous Comput. High-Perform. Comput. Commun. (HPCC_EUC)*, Nov. 2013, pp. 2187_2193.
- [9] M. Swan, *Blockchain: Blueprint for a New Economy*. Newton, MA, USA: O'Reilly Media, 2015.
- [10] Ying Yu, Min Li, Liangliang Liu, Yaohang Li, and Jianxin Wang, "Clinical Big Data and Deep Learning: Applications, Challenges, and Future Outlooks", BIG DATA MINING AND ANALYTICS, ISSN 2096-0654 06/08 pp288-305, Volume 2, Number 4, December 2019
- [11] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 4, pp. 580_585, Jul./Aug. 1985.
- [12] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18_28, Jul./Aug. 2008.

- [13] M. P. Brown *et al.*, "Knowledge-based analysis of microarray gene expression data by using support vector machines," *Proc. Nat. Acad. Sci. USA*, vol. 97, no. 1, pp. 262_267, 2000.
- [14] L. Khedheret *al.*, "Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images," *Neurocomputing*, vol. 151, pp. 139_150, Mar. 2015
- [15] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85_117, Jan. 2015.
- [16] T. Kooiet *al.*, "Large scale deep learning for computer aided detection of mammographic lesions," *Med. Image Anal.*, vol. 35, pp. 303_312, Jan. 2017.
- [17] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881_892, Jul. 2002
- [18] P. Gasch and M. B. Eisen, "Exploring the conditional coregulation of yeast gene expression through fuzzy k-means clustering," *Genome Biol.*, vol. 3, no. 11, p. research0059-1, 2002
- [19] T. G. Dietterich, "Ensemble methods in machine learning," in *Proc. Int. Workshop Multiple Classifier Syst.*, 2000, pp. 1_15.
- [20] A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts, "Modeling medical treatment using Markov decision processes," in *Operations Research and Health Care*. Boston, MA, USA: Springer, 2005, pp. 593_612.
- [21] Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [22] Y. LeCun and Y. Bengio, Convolutional networks for images, speech, and time series, in *Handbook of Brain Theory and Neural Networks*, M. A. Arbib, ed. Cambridge, MA, USA: MIT Press, 1995.
- [23] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [24] G. E. Hinton and R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science*, vol. 313, no. 5786, pp. 504–507, 2006.