DIJEMSS: Dinasti International Journal of Education Management and Social Science
E-ISSN: 2686-6331
P-ISSN: 2686-6358
DINASTI PUBLISHER
https://dinastipub.org/DIJEMSS dinasti.info@gmail.com +62 811 7404 455
DOI: https://doi.org/10.38035/dijemss.v6i3

The Role of Artificial Intelligence in Optimizing Electronic Health Records for Early Detection of Disease

Yuda Syahidin[1], Ade Irma Suryani[2]
[1]Health Information Management, Piksi Ganesha Politechnic, Bandung, Indonesia, mail4yudasy@gmail.com
[2]Health Information, Piksi Ganesha Politechnic, Bandung, Indonesia, adeirmasuryani20@gmail.com

Corresponding Author: mail4yudasy@gmail.com[1]

Abstract: The Role of Artificial Intelligence in Optimizing Electronic Health Records for Early Detection of Disease is a critical focus area, as AI technologies enhance the ability to analyze vast datasets within EHRs, facilitating timely identification of health risks and improving patient outcomes. By leveraging AI, healthcare providers can streamline data management processes and support more accurate and efficient predictions, ultimately leading to better disease management and resource allocation. This paper emphasizes the importance of integrating artificial intelligence into EHR systems to maximize their potential for early disease detection and improve overall healthcare delivery.

Keywords: Predictive Analytics, Electronic Health Record, Machine Learning

INTRODUCTION

The rapid development of the internet, electronic medical information has become popular in all cities and around the world, such as electronic health records (EHR) to replace paper medical records, online appointments, and online reports, so that the volume of data growth is growing and many have encouraged research. intensive in the development of technological devices to store, manage, and analyze data in the health sector (Hong dkk, 2019). Although a significant amount of health data remains underutilized, Unstructured Data Analysis is needed as the next innovation in health data science. The scalability and efficiency of analyzing medical record data are still insufficient for generating comprehensive Electronic Health Reports, and decision-making systems in healthcare are limited due to the presence of unstructured data.

The use of Electronic Health Records (EHR) makes it possible to analyze large amounts of medical data (Shickel dkk, 2017). Recently, deep learning techniques can play an important role in managing the enormous medical data that has been generated every day (Ismail dkk, 2020; Xu dkk, 2020)]. Moreover, deep learning achieves success in various fields by effectively building deep hierarchical features.

The Role of Artificial Intelligence in Optimizing Electronic Health Records for Early Detection of Disease is becoming increasingly significant, as AI technologies enhance the capacity of predictive analytics to process and interpret vast amounts of health data. By integrating AI into Electronic Health Records (EHR), healthcare providers can more effectively identify patterns and trends that facilitate early disease detection. This integration not only streamlines the analysis of EHR data but also improves the accuracy of predictions related to patient outcomes. Consequently, the application of artificial intelligence within EHR systems holds the potential to transform patient care by fostering timely interventions and enabling more personalized healthcare solutions (Jensen dkk, 2012), as well as in predictive analysis to determine the likelihood of predicting a patient's readmission to the hospital.

Electronic Health Record (EHR) has become a widely used data source for clinical risk prediction, which offers unique opportunities and challenges for analysis. One of the main challenges in predicting outcomes with EHR data is that the performance of predictive algorithms is highly dependent on how the data is represented and the selection of relevant features. Another difficulty in feature selection is the time-consuming process of analyzing, selecting and evaluating raw EHR data, which often involves a trial-and-error process. This process is further complicated by the presence of potentially thousands of predictive variables, especially when clinical notes from doctors, nurses, and other providers are vague (Goldstein dkk, 2017; Weiskopf dkk, 2013). Therefore, it is imperative to develop predictive models that can perform pattern recognition, statistical analysis, database management, and visualization to overcome the challenges of extracting information from large healthcare databases.

In this paper, we will discuss the predictive analytics model by previous research, as well as the model that will be examined based on previous research opportunities.

Electronic Health Records

An Electronic Health Record (EHR) is a digital system that securely stores a person's health information, providing instant access to authorized users. EHRs include details such as patient diagnoses, medications, vital signs, treatment plans, progress notes, radiological images, and test results. There are two types of EHR data: structured and unstructured. Written or spoken notes about the patient's health that are based on the clinical context are known as unstructured data, and they are very useful for clinical documentation. However, they present a problem for computer analysis because to their disorganized nature, multiple typos and spelling mistakes, and usage of acronyms, abbreviations, and peculiarities. Administrative and supplementary data can be used to categorize organized RKE data. Administrative data either don't change during the clinical encounter (like demographic information) or keep changing over time (like procedures and diagnoses). Extra information can be continuous (like blood pressure and respiration) or discrete (like physiological measurements, drugs, and lab tests)..

EHR data includes a wide range of information, from structured data like drug prescriptions, which consist of dates and dosages recorded through standardized prescribing systems, to unstructured data such as clinical narratives that explain the medical reasoning behind the records. The connection between structured and unstructured data is illustrated in Figure 1 below.

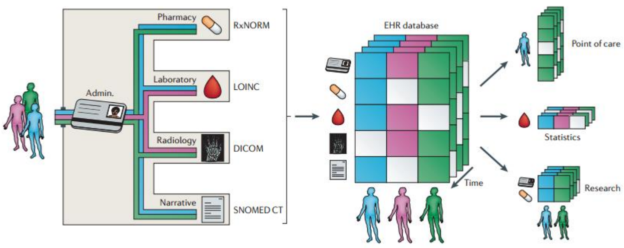


Figure 1. Electronic health record content (Jensen dkk, 2012)

A patient's Electronic Health Record (EHR) serves as a repository of information about their health in a computer-readable format. When integrated with the healthcare system, it generates different types of patient-related data. This data is stored in a database and can be accessed in formats tailored to the needs and permissions of specific user groups. For instance, a doctor might retrieve EHR data for a specific patient, statistical summaries of all lab procedures, or data extractions for research purposes.

The computerized collection of patient health data is referred to as an electronic health record (EHR) or electronic medical record (EMR). EHRs fall into one of the following functional categories: Basic EHRs without clinical records (i), basic EHRs with clinical records (ii), and comprehensive systems (iii) are the three options. Even in their most basic configuration, EHRs offer a wealth of data to researchers. Data can encompass a wide range of information and can be shared across networks, as previously mentioned. EHR is mostly intended for use in internal hospital administration duties, and several systems are available in various organizational configurations.(Liang dkk, 2019).

METHOD

Predictive Analytics involves various methods that use both current and historical data to forecast future outcomes. It aims to identify relationships within data and reveal patterns. To carry out Predictive Analytics, a formula or set of rules, referred to as an analytical model, is applied to generate a score or code. This modeling process uses mathematical techniques to uncover significant relationships between variables (Leventhal, 2018). Each model used is a simplification of the existing reality, which can help in understanding problems and making predictions.

There are several models used for Predictive Analytics, including:

A. *Decision Tree Analysis.*
   Decision trees are good predictive models and have many advantages, including being easy to understand and can be presented in very simple or detailed ways. The Decision Tree will start at the root of the tree, which represents all the data, and continue to break each category into two separate categories based on the optimal method, to be able to guess the best characteristics to identify each of the two categories (splitting a node). Then the algorithm will continue to divide the data until it can no longer divide, or stop in another way based on the control parameters (Leventhal, 2018).
B. *Random Forest*
   Random Forest is a method derived from Decision Trees. A decision tree, also known as a decision-making tree, is a flowchart resembling a tree structure. It consists of a root node for data collection, inner nodes that contain questions about the data, and leaf nodes that are used for problem-solving and decision-making. The decision tree classifies a data sample with an unknown class into existing categories. This approach helps prevent overfitting of the dataset while maximizing accuracy [20].
C. *Regression Analysis*
   Regression analysis is a statistical process used to analyze the relationship between variables, especially to predict one or more target variables (Hao dkk, 2019).
D. *Supervised Neural Network*
   Supervised Neural Networks are a good solution when the relationship between the target and predictor variables is complex and unknown (Sidey-Gibbons & Sidey-Gibbons, 2019).
   Neural Network processing basically mimics biological neurons, as seen in Figure 2 (M. Li dkk, 2021).

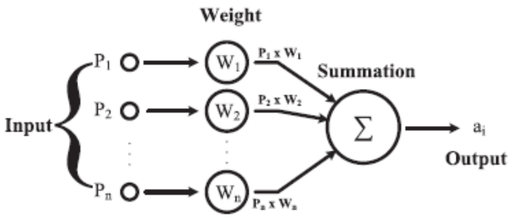


Figure 2. Artificial Neural Network Satu Layer (Sharma & Shah, 2021)

RESULT AND DISCUSSION

The Machine Learning method is most commonly used to build medical database systems from EHRs for patients who have undergone health checks (Hauskrecht dkk, 2013), Machine learning methods are still used by previous studies in making predictions that tend to use one or a combination of several algorithms (Sohn Dr. dkk, 2013). The following is previous research using machine learning techniques.

1) S. Mohan [14]

Making choices and predictions from the vast amounts of data produced by the healthcare sector has been made easier with the help of machine learning. In 2019, S. Mohan (Mohan dkk, 2019) suggests a novel approach that seeks to identify noteworthy characteristics through the use of machine learning algorithms, leading to an improvement in the prediction accuracy of cardiovascular disease. A variety of features and established categorization methods are combined to introduce predictive models. Heart disease prediction model using the linear Hybrid Random Forest Model (HRFLM). like picture 2 below.

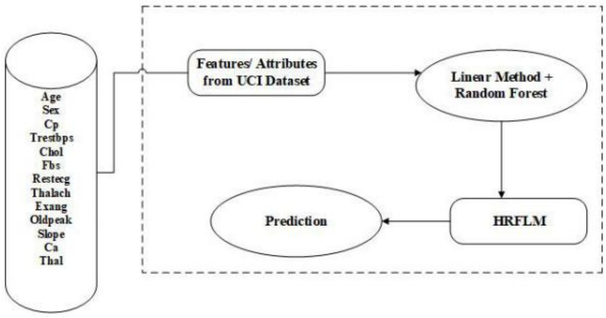


Figure 3. HRFLM Model (Mohan dkk, 2019)

2) K. Yu and X. Xie (Wang dkk, 2021)

This research (Meng dkk, 2020) allows medical professionals to identify high-risk individuals and take prompt action. For patient readmissions to hospitals, this method makes use of predictive analytics based on electronic health records (EHR). A framework for hospital readmissions is suggested in light of this. Hospitals' inpatient administration data from the national health care dataset is used in this method. The framework model for hospital readmission is depicted in Figure 3 below..

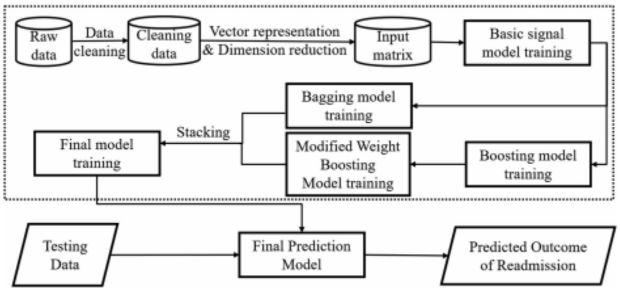


Figure 4. Model Framework Hospital Readmision (Nguyen dkk, 2017)

3) Y. Sun et al (Sun & Zhang, 2019)

Research by Y.Sun (Sun & Zhang, 2019) provides a collection of machine learning models to identify individuals with diabetic retinopathy (DR) using electronic health record data and develop a treatment plan. The approach utilized in the Machine Learning model with the five algorithm strategies, as displayed in Figure 4 below.

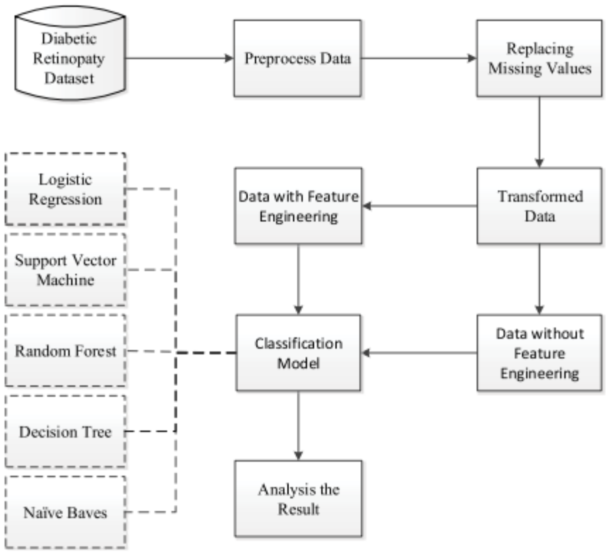


Figure 5. Model Framework Diabetic Retinopathy (Sun & Zhang, 2019)

4) R. Ghorbani (Ghorbani dkk, 2020)

The novel hybrid model combines an ensemble technique that combines Stacking and Boosting with a Genetic Algorithm for feature selection. Using this technique, the best subset of pertinent information is found to improve the creation of prediction models. In particular, feature selection strategies can lower the dataset's dimensionality by eliminating noise-filled or inconsequential information, improving the accuracy of prediction models.

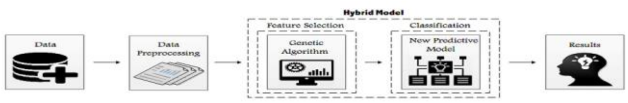


Figure 6. Model Framework New Hybrid (Ghorbani dkk, 2020)

5) Deep Learning

Over the past several years, there has been an increase in the application of deep learning to analyze EHR data; further expansion in this area is made possible by the availability of more EHR data, advancements in deep learning, and creative ways to integrate these two trends. The best possible Deep Learning architecture to research effective patient representation to forecast emergency department admission and heart failure patients (Guller & Guller, 2015). A directed graph can be used to depict a neural network (NN), in which one or more hidden layers process the output of the preceding layer after the input layer receives a signal vector. An strategy called deep learning models seeks to create end-to-end systems that can learn from unprocessed data and carry out certain activities on their own without human oversight. In comparison to a neural network, a deep neural network has more layers and nodes in each layer. The quantity of parameters that must be adjusted to improve neural network performance. Furthermore, deep learning cannot be learnt in the absence of sufficient data and a strong computer.

Various kinds of deep neural networks exist, including: (i) Convolutional Neural Networks (CNNs), in which the structure of the visual brain serves as an inspiration for the pattern of connectivity between layers; (ii) Recurrent Neural Network (RNN), in which nodes' connections create a directed graph that runs along a series and [34] (iii) Information travels in a single direction—forward—from input nodes to output nodes via hidden nodes, if any, and back again in a feedforward neural network (FFN). The network is free of loops and cycles. An Artificial Neural Network (ANN) with a few interconnected nodes (neurons) stacked in multiple layers is the foundation upon which deep learning algorithms are constructed. Layers that are not part of the input or output layers are known as hidden nodes (Wanyan dkk, 2020).

1) Miotto, R et al (Miotto dkk, 2016)

This model (Miotto dkk, 2016) suggests using an unsupervised Deep Learning method called the Deep Patient model to extract hierarchical characteristics and patterns from EHR data. This study demonstrates that hierarchical characteristics based on EHR datasets can be obtained by unsupervised Deep Learning.

a model that uses deep learning methods to depict the predicted patient status outcomes using EHR data. By estimating the probability that the patient is presently developing further ailments, this depiction seeks to forecast the patient's health status. They search the dataset in Figure 5 below for a hierarchical representation using autoencoder algorithms.

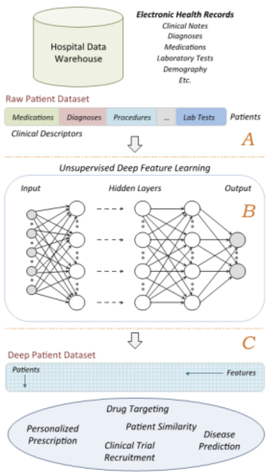


Figure 7. Model Deep Patient (Miotto dkk, 2016)

2) M. Jamshidi et al. (Jamshidi dkk, 2020)

The compilation of data is the initial phase. Medical data, including clinical reports, notes, photographs, and other data formats that can be transformed into machine-readable data, make up the data being discussed here. Human intervention is a component of the machine learning process that involves looking into and analyzing data to extract information with characteristics, patterns, and structure. similar to figure 8 down below.

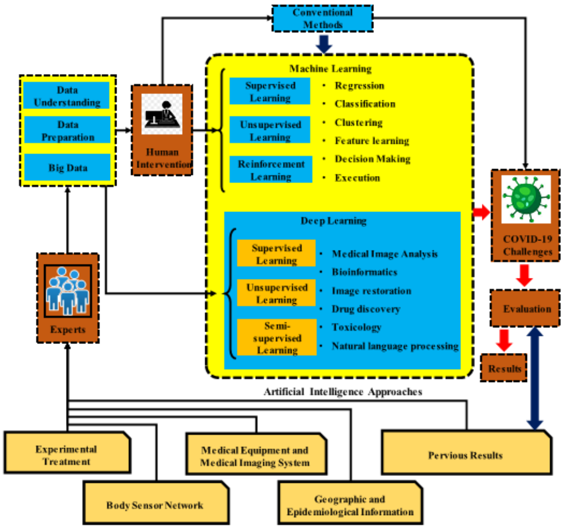
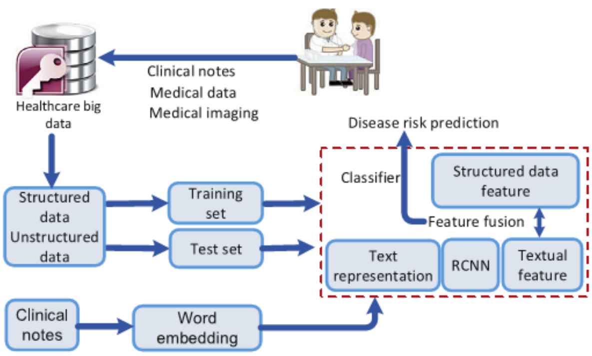


Figure 8. Deep Learning for Diagnosis and Treatment (L. F. Li dkk, 2020)

3) Yixue Hao (Hao dkk, 2019)

He presented the MD-RCNN model, which uses multi-modal data to predict disease risk, in the research that was done. In order to derive a highly non-linear relationship between structured and unstructured data, the multimodal data-based recurrent convolutional neural network (MD-RCNN) model extracts structured and unstructured characteristics. as displayed here in Figure 9.



Gambar 9. Model MD-RCNN (Hao dkk, 2019)

4) J. Baek et al (Çelik dkk, 2018)

Selected context variables provide the basis for the model's operation. High-relationship context information is then retrieved and used as DNN-context input for learning. Following processing, the data is separated into two sets: a training set and a test set. The process of data feature extraction involves entering the data into the DNN, extracting the data features, and displaying them in the hidden layer. In order to assess how good the model is, the output data is categorized. As seen in Figure 10 below, this model is contrasted with other models.

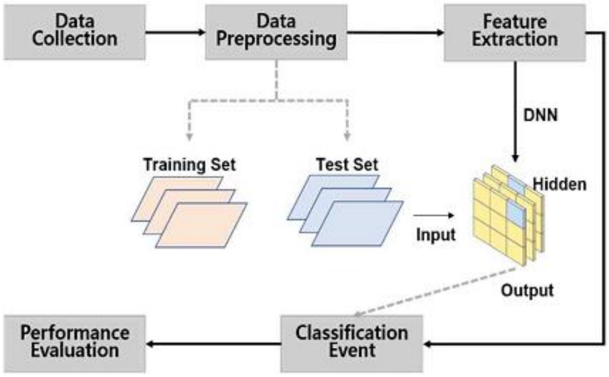


Figure 10. Model Context DNN (Baek & Chung, 2020)

CONCLUSION

Based on the review paper that has been described in the Predictive Analytics model from previous research, the techniques employed are visible. Table 1 below provides an explanation of earlier studies on predictive analytics in the context of electronic health records.

TABLE I. COMPARISON OF MODEL PREDICTIVE ANALYTICS IN ELECTRONIC HEALTH RECORDS

Paper	Method	Focus	Performance
S. Mohan (Mohan dkk, 2019)	Hybrid Random Forest with model linier (HRFLM).	Heart Disease	Make predictions of heart disease up to an accuracy level of 88.7%
K. Yu and X. Xie (Yu & Xie, 2020)	The joint ensemble-learning model, Modified Weight Boosting-Stacking (MWBS)	Return of the patient to the hospital	Make predictions about patient arrivals up to an accuracy level of 89.1%
Y. Sun (Sun & Zhang, 2019)	Classification Model: Logistic Regression (LR), Support Vector Machines (SVM), Random Forests (RF), Decision	Diabetic retinopathy (DR) – Eye disorder due to diabetes	Predict eye disorders using 5 (five) classification algorithms and produce predictions with RF up to 92%

	Trees (DT), Naïve Bayes (NB)		
R. Ghorbani (Ghorbani dkk, 2020)	Genetic Algorithm	Mortality	This model achieves 98.20% by 88.47% with the AUC test
Miotto, R et al (Miotto dkk, 2016)	Unsupervised deep feature learning	78 kinds of disease (disease)	Generate predictive results of more than one kind of prediction
P. Nguyen et al (Sabokrou dkk, 2018)	Multilayered architecture based on CNN	Clinical Notes, Medical Codes	Hospital re-admission prediction
Yixue Hao (Hao dkk, 2019)	MD-RCNN algorithm	Clinical notes, Medical data, Medical Imaging	Melakukan prediksi terhadap resiko penyakit berdasarkan data kesehatan yang terstruktur maupun tidak terstruktru
J. Baek et al [36], 2020	Deep Neural Network dan Multiple Regression	Depression	Nilai yang diprediksi adalah nilai antara 0 dan 1. Risiko depresi memiliki empat tahap, yaitu 'baik', 'tidak buruk', 'berisiko', dan 'sangat berisiko'

Based on the explanation in previous studies, the challenges in doing Predictive Analytics include unstructured and structured data. Therefore, one of the main challenges is to integrate and harmonize data that has differences. In addition, the heterogeneous nature of data types including numeric data, date time objects, free text (clinical outcomes) poses a significant challenge to obtain important information in EHRs and because EHRs data is in the form of structured and unstructured, classification and clustering models are needed to determine predictive analytics in the field of Electronic Health Records.

REFERENCES

Baek, J. W., & Chung, K. (2020). Context Deep Neural Network Model for Predicting Depression Risk Using Multiple Regression. IEEE Access, 8, 18171–18181. https://doi.org/10.1109/ACCESS.2020.2968393

Çelik, A., Yaman, H., Turan, S., Kara, A., Kara, F., Zhu, B., Qu, X., Tao, Y., Zhu, Z., Dhokia, V., Nassehi, A., Newman, S. T., Zheng, L., Neville, A., Gledhill, A., Johnston, D., Zhang, H., Xu, J. J., Wang, G., … Dutta, D. (2018). FUNDAMENTALS OF MACHINE LEARNING FOR PREDICTIVE DATA ANALYTICS. In Journal of Materials Processing Technology (Vol. 1, Nomor 1).

Futoma, J., Morris, J., & Lucas, J. (2015). A comparison of models for predicting early hospital readmissions. Journal of Biomedical Informatics, 56, 229–238. https://doi.org/10.1016/j.jbi.2015.05.016

Ghorbani, R., Ghousi, R., Makui, A., & Atashi, A. (2020). A New Hybrid Predictive Model to Predict the Early Mortality Risk in Intensive Care Units on a Highly Imbalanced Dataset. IEEE Access, 8, 141066–141079. https://doi.org/10.1109/ACCESS.2020.3013320

Goldstein, B. A., Navar, A. M., Pencina, M. J., & Ioannidis, J. P. A. (2017). Opportunities and challenges in developing risk prediction models with electronic health records data: A systematic review. Journal of the American Medical Informatics Association, 24(1), 198–208. https://doi.org/10.1093/jamia/ocw042

Guller, M., & Guller, M. (2015). Big Data Technology Landscape. Big Data Analytics with Spark, Ml, 1–15. https://doi.org/10.1007/978-1-4842-0964-6_1

Hao, Y., Usama, M., Yang, J., Hossain, M. S., & Ghoneim, A. (2019). Recurrent convolutional neural network based multimodal disease risk prediction. Future Generation Computer Systems, 92, 76–83. https://doi.org/10.1016/j.future.2018.09.031

Hauskrecht, M., Batal, I., Valko, M., Visweswaran, S., Cooper, G. F., & Clermont, G. (2013). Outlier detection for patient monitoring and alerting. Journal of Biomedical Informatics, 46(1), 47–55. https://doi.org/10.1016/j.jbi.2012.08.004

Hong, W., Xiong, Z., Zheng, N., & Weng, Y. (2019). A Medical-History-Based Potential Disease Prediction Algorithm. IEEE Access, 7, 131094–131101. https://doi.org/10.1109/ACCESS.2019.2940644

Ismail, W. N., Hassan, M. M., Alsalamah, H. A., & Fortino, G. (2020). CNN-based health model for regular health factors analysis in internet-of-medical things environment. IEEE Access, 8, 52541–52549. https://doi.org/10.1109/ACCESS.2020.2980938

Jamshidi, M., Lalbakhsh, A., Talla, J., Peroutka, Z., Hadjilooei, F., Lalbakhsh, P., Jamshidi, M., Spada, L. La, Mirmozafari, M., Dehghani, M., Sabet, A., Roshani, S., Roshani, S., Bayat-Makou, N., Mohamadzade, B., Malek, Z., Jamshidi, A., Kiani, S., Hashemi-Dezaki, H., & Mohyuddin, W. (2020). Artificial Intelligence and COVID-19: Deep Learning Approaches for Diagnosis and Treatment. IEEE Access, 8(December 2019), 109581–109595. https://doi.org/10.1109/ACCESS.2020.3001973

Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records: Towards better research applications and clinical care. Nature Reviews Genetics, 13(6), 395–405. https://doi.org/10.1038/nrg3208

Leventhal, B. (2018). Predictive Analytics for marketers. Kogan Page.

Li, L. F., Wang, X., Hu, W. J., Xiong, N. N., Du, Y. X., & Li, B. S. (2020). Deep Learning in Skin Disease Image Recognition: A Review. IEEE Access, 8, 208264–208280. https://doi.org/10.1109/ACCESS.2020.3037258

Li, M., Liu, Y., Liu, X., Sun, Q., You, X., Yang, H., Luan, Z., Gan, L., Yang, G., & Qian, D. (2021). The Deep Learning Compiler: A Comprehensive Survey. IEEE Transactions on

Parallel and Distributed Systems, 32(3), 708–727. https://doi.org/10.1109/TPDS.2020.3030548

Liang, H., Tsui, B. Y., Ni, H., Valentim, C. C. S., Baxter, S. L., Liu, G., Cai, W., Kermany, D. S., Sun, X., Chen, J., He, L., Zhu, J., Tian, P., Shao, H., Zheng, L., Hou, R., Hewett, S., Li, G., Liang, P., … Xia, H. (2019). Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nature Medicine, 25(3), 433–438. https://doi.org/10.1038/s41591-018-0335-9

Meng, L., Dong, D., Li, L., Niu, M., Bai, Y., Wang, M., Qiu, X., Zha, Y., & Tian, J. (2020). A Deep Learning Prognosis Model Help Alert for COVID-19 Patients at High-Risk of Death: A Multi-Center Study. IEEE Journal of Biomedical and Health Informatics, 24(12), 3576–3584. https://doi.org/10.1109/JBHI.2020.3034296

Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. Scientific Reports, 6. https://doi.org/10.1038/srep26094

Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. IEEE Access, 7, 81542–81554. https://doi.org/10.1109/ACCESS.2019.2923707

Nguyen, P., Tran, T., Wickramasinghe, N., & Venkatesh, S. (2017). Deepr: A Convolutional Net for Medical Records. IEEE Journal of Biomedical and Health Informatics, 21(1), 22–30. https://doi.org/10.1109/JBHI.2016.2633963

Sabokrou, M., Fayyaz, M., Fathy, M., Moayed, Z., & Klette, R. (2018). Deep-anomaly: Fully convolutional neural network for fast anomaly detection in crowded scenes. Computer Vision and Image Understanding, 172(January), 88–97. https://doi.org/10.1016/j.cviu.2018.02.006

Sharma, T., & Shah, M. (2021). A comprehensive review of machine learning techniques on diabetes detection. Visual Computing for Industry, Biomedicine, and Art, 4(1). https://doi.org/10.1186/s42492-021-00097-7

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2017). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. arXiv, 22(5), 1589–1604.

Sidey-Gibbons, J. A. M., & Sidey-Gibbons, C. J. (2019). Machine learning in medicine: a practical introduction. BMC Medical Research Methodology, 19(1), 1–18. https://doi.org/10.1186/s12874-019-0681-4

Sohn Dr., S., Wagholikar, K. B., Li, D., Jonnalagadda, S. R., Tao, C., Elayavilli, R. K., & Liu, H. (2013). Comprehensive temporal information detection from clinical text: Medical events, time, and TLINK identification. Journal of the American Medical Informatics Association, 20(5), 836–842. https://doi.org/10.1136/amiajnl-2013-001622

Sun, Y., & Zhang, D. (2019). Diagnosis and Analysis of Diabetic Retinopathy Based on Electronic Health Records. IEEE Access, 7, 86115–86120. https://doi.org/10.1109/ACCESS.2019.2918625

Wang, Q., Wang, S., Wei, B., Chen, W., & Zhang, Y. (2021). Weighted K-NN Classification Method of Bearings Fault Diagnosis with Multi-Dimensional Sensitive Features. IEEE Access, 9, 45428–45440. https://doi.org/10.1109/ACCESS.2021.3066489

Wanyan, T., Vaid, A., De Freitas, J. K., Somani, S., Miotto, R., Nadkarni, G. N., Azad, A., Ding, Y., & Glicksberg, B. S. (2020). Relational Learning Improves Prediction of Mortality in COVID-19 in the Intensive Care Unit. IEEE Transactions on Big Data, December 2020. https://doi.org/10.1109/TBDATA.2020.3048644

Weiskopf, N. G., Hripcsak, G., Swaminathan, S., & Weng, C. (2013). Defining and measuring completeness of electronic health records for secondary use. Journal of Biomedical Informatics, 46(5), 830–836. https://doi.org/10.1016/j.jbi.2013.06.010

Xu, D., Hu, P. J. H., Huang, T. S., Fang, X., & Hsu, C. C. (2020). A deep learning–based, unsupervised method to impute missing values in electronic health records for improved patient management. Journal of Biomedical Informatics, 111(October). https://doi.org/10.1016/j.jbi.2020.103576

Yu, K., & Xie, X. (2020). Predicting Hospital Readmission: A Joint Ensemble-Learning Model. IEEE Journal of Biomedical and Health Informatics, 24(2), 447–456. https://doi.org/10.1109/JBHI.2019.2938995