

Healthcare Analytics: A Comprehensive Review

Alshymaa Alghamdi

College of Computer and Information Systems
Umm Al-Qura University
Mecca, Saudi Arabia
s44180219@st.uqu.edu.sa

Tahani Alsubait

College of Computer and Information Systems
Umm Al-Qura University
Mecca, Saudi Arabia
tmsubait@uqu.edu.sa

Abdullah Baz

College of Computer and Information Systems
Umm Al-Qura University
Mecca, Saudi Arabia
aobaz01@uqu.edu.sa

Hosam Alhakami

College of Computer and Information Systems
Umm Al-Qura University
Mecca, Saudi Arabia
hhhakam@uqu.edu.sa

Abstract-Big data have attracted significant attention in recent years, as their hidden potentials that can improve human life, especially when applied in healthcare. Big data is a reasonable collection of useful information allowing new breakthroughs or understandings. This paper reviews the use and effectiveness of data analytics in healthcare, examining secondary data sources such as books, journals, and other reputable publications between 2000 and 2020, utilizing a very strict strategy in keywords. Large scale data have been proven of great importance in healthcare, and therefore there is a need for advanced forms of data analytics, such as diagnostic data and descriptive analysis, for improving healthcare outcomes. The utilization of large-scale data can form the backbone of predictive analytics which is the baseline for future individual outcome prediction.

Keywords-healthcare; big data; healthcare analytics; descriptive analytics; diagnostic analytics

I. INTRODUCTION

Data have become a topic of particular interest because of their high but often hidden potential [1]. Their importance has been proven in almost all aspects of life. Data give the ability to study the past and predict the future while developing the present. Although profit is not the primary motivator, healthcare organizations should consider acquiring knowledge for available techniques, tools, and infrastructure to leverage big data [2]. In most sectors data grow exponentially. Healthcare large scale data describe the large data volume [3] that overburdens a healthcare institution on a daily basis. Existing logical methods can be applied to the tremendous amount of existing patient-related well-being and clinical information, offering more profound comprehension and improving healthcare [4]. In real world scenarios, individual information could help each doctor in deciding the most appropriate treatment for a given patient. Healthcare has taken too long to utilize big data compared to other sectors. Healthcare big data activities include collecting, analyzing, holding of consumer, patient, physical, and clinical data that

are too large and complex to be understood by the traditional data processing methods. An example of this big data processing is the use of complex and sophisticated sensors in healthcare. As a result, big data are analyzed and processed by machine learning algorithms and specialized data scientists [5]. Healthcare facilities require new methods to work out their recurrent challenges, especially concerning patients' diagnosis and treatment information. As a solution to these challenges that are related to data volume, velocity, variety, and veracity, healthcare facilities and organizations have to incorporate new technologies capable of data collection, storage, and information analysis to produce actionable insights [3]. Different situations call for the application of different data strategies. The idea is to use healthcare analytics to create actionable information that can be utilized for improving patient care and the efficiency of processes. The paper at hand examines previous studies and evaluates the utilization of descriptive, diagnostic, and predictive healthcare analytics.

II. HEALTHCARE ANALYTICS

The utilization of data analytics is helpful to healthcare professionals, especially in terms of predicting, diagnosing, and treating diseases, improving the quality of services and minimizing healthcare costs [6]. It is estimated that the U.S. healthcare system saves more than \$450 billion annually through data mining [7]. Many in-depth studies have examined the application of big data in mental health and other areas such as the methodological challenges of data mining perspectives [8]. This large volume of data is generated from fast-growing industries attracting clinicians' and scientists' attention [9], emerging the need for advanced forms of data analytics. Healthcare professionals could significantly benefit from these techniques, reducing considerable effort, resources, and time, smoothening the running. These methods could also satisfy professionals by offering them more valuable knowledge and information [10]. Big data analytics includes analysis, observation, influence, and prediction.

Corresponding author: Alshymaa Alghamdi

A. Observation

Noting and taking records is essential for generating critical values from large scale data, and a new scientific field called "data science" has evolved [11]. A critical factor is data visualization in a quite friendly way which is vital for the development of society. Management of these big data becomes challenging due to the organization and reorganization of the same data [12]. Therefore highly advanced techniques are often employed for efficient and cost-effective data management. Healthcare services can be offered with analytics tools through proper storage [13].

B. Influence

Data collection is important for every company which depends in future predictions [14]. More information causes more deliveries. Gigabytes or terabytes are usually collected structurally in health monitoring toolkits, including sensors, structure, and data acquisition. Large data volumes are often caused by either the number of sensors or the monitoring data structure size [15]. The scope is given by structure, while sensors provide resolution.

C. Prediction

Prediction becomes more useful when knowledge can be transformed into action. Prediction and intervention need to be integrated into the same system in order to increase system's efficiency. This is a process of making future predictions based on historical data [16]. This analysis improves healthcare by improving decisions and personalizing care for each individual. Healthcare organizations use predictive analysis drawn from new data streams directly from patients [17].

D. Analysis

Analyzing big data has a workflow similar to what is described in [8]. A data warehouse is the storage of massive amounts of data produced from various sources. It uses analytic pipelines for smarter and affordable healthcare options. Modern techniques have evolved to aid in this, increasing the resolution in which biological events are observed and recorded. Therefore such big data analytics from medical healthcare systems can help providing new strategies [18].

III. DESCRIPTIVE HEALTH ANALYTICS

Healthcare analytics are very important in medicine, as they offer insights into data, including hospital patient records, costs, and diagnoses, among others. Descriptive analytics is a sub-branch of healthcare analytics [19], aiming to analyze raw data and come up with viable solutions to help patients. Big data are firstly turned into actionable insights, mostly relying on public health and pharmacy among other factors [9]. Descriptive analytics include various methods, entailing stages or processes whose outcomes enable easy manipulation and utilization to figure out solutions in real life. Descriptive analytics use control data to categorize, characterize, aggregate, and classify data into valuable information for the healthcare professionals, in understanding and analyzing decisions, performance, and outcomes [20]. Besides, data from descriptive analytics are prepared using tables and graphs showing the average length of stay in hospitals [4], the rate of discharge, and the rates of hospital occupancy among other

indicators [21]. Descriptive analytics is usually the easiest level of data to understand. Moreover, control data are utilized in categorization, characterization, aggregation, and classification of the data, transforming them into valuable information for healthcare professionals in the process of understanding and analyzing decisions, performance, and outcomes. Data visualization can answer specific questions related to patient care patterns, which ultimately highlight a broad perspective of evidence-based clinical practice [22]. As a result, real-time or near real-time data management has to be enabled in the process. Using this method, literature predicts that previously unnoticed observed patterns in terms of hospital readmission are enabled to enhance a balance between cost and capacity [14]. Descriptive analytics help because they connect previous decisions and their implications. It answers the question: "what happened?" giving a good basis of decision making.

IV. CLASSIFICATION OF DESCRIPTIVE ANALYTICS

A. Data Classification

Classification falls under supervised machine learning methods. Artificial neural networks consist one of the most common underlying structures exploited [23]. Training and validation sets are used whereas adequate output for given inputs is generated. This is the process of organizing healthcare data into relevant groups and categories that are easier to understand and analyze. This step is important as it allows easy data retrieval when needed. In healthcare applications the lack of data classification could result on much duplication complicating the tracking down of certain events and occurrences related to patients and health facilities. Data classification process in healthcare widely revolves around content, context, and users. Context refers to the location where it is applicable. User refers to the person who is using the data being classified. People in charge of the classification process have to take good care of the classification criteria, as wrong criteria can lead to wrong classifications. The categories should include information about the type of data included. Moreover, security and privacy considerations should be examined with rules for data retrieval, transmission, storage, and possible risks associated with security policy violations [24].

B. Data Mining Algorithms Algorithms

Heuristic-based data mining algorithms in hospitals can be helpful in the stipulation of the relationship between different symptoms and a particular disease. The same algorithms can be used to stipulate the relationships between cause and effects of human behaviors and diseases [24]. One of their main roles is to ensure the standardization of healthcare practices. Medical algorithms are evidence-based and data-driven [8].

C. Methods/Tools

The methods applied are regression, association, clustering, data warehousing, and sequential pattern mining. Data can be analyzed electronically or manually using traditional methods such as pen and paper through graphical methods [25]. This information is used by healthcare providers to shed light on facts such as occupancy rates and financial state of hospitals, number of in- and out-patients, discharge, and length of stay among other entities or variables of study. Other data analysis

methods are utilized including statistical analysis, which features descriptive, inferential, and diagnostic analysis among others. In healthcare, analysis of data is performed via statistical analysis using appropriate tools [7]. Data from hospital financial records, number of patients, and staff can be inferred through statistical analysis of collected data. Current data analysis tools include SPSS, Matlab, SQL, Java, Weka, Rapidminer, R analytics suite, and Python Scikit-learn [7, 20].

D. Problem Domain

Problem domain analytics involve the process of creating a model that describes the problem to be solved. In healthcare, this deals with modeling solutions and curative measures for different diseases [26]. Once there is a healthcare problem domain analysis for a given facility, this can be applied to multiple implementations at different periods as technology advances. Problem domain analysis is fully independent of the platform where the solution is implemented [4].

E. Data Visualization

Data visualization is the process of representing data graphically. The visual elements used in this stage are charts, graphs, and maps. These elements along with data visualization tools enable one to see and understand patterns, trends, and outliers of data. Data visualization is important for users in order to understand the enormous amounts of daily generated data quickly and effectively. In healthcare, where data are vital and the volumes of circulating data are huge, data visualization tools and technologies are essential to analyze information in order to enable data driven decisions. Efficient use of data visualization techniques in healthcare enhances usability, hence the provision of quality care [27]. This stage should be taken seriously into consideration, as the advantages or benefits of effective data visualization are many, and they lead to better decision-making, hence improvement of the health sector. Data visualization is incorporated in healthcare in an effort to answer specific questions or patterns of care, providing a broad view of evidence-based clinical practice. The activity of data analysis gives sufficient space for proper data management, use and study of operational content, and helps in capturing visual data and information including patient records. Henceforth, data previously not observed and patterns not studied before are brought into light via data visualization [28]. Differences brought up in patient records, financial records, and healthcare provision are captured and recorded through visualization tools. It is hard to imagine a healthcare institution that does not visualize data for its decision making [29].

V. DIAGNOSTIC HEALTH ANALYTICS

Diagnostic health analytics is the second major health analytics method. This method is used in the analysis of healthcare organizations in order to transform them to fully data-driven institutions, offering competitive advantages. Diagnostic analysis answers the "why" in a phenomenon [30], aiming at finding the causes of events. Diagnostic analytics explain "why it happened" [30-31], such as "why did a patient get worse even with the best care". These answers enable healthcare providers to figure out problems and solutions. Diagnostic health analytics take a critical look on data, trying to understand why events and behaviors occur [31].

A. Data Discovery, Exploration, and Analysis

Exploration, discovery, and analysis of data are major processes involved in diagnostic health analytics. Discovery is bringing to light concealed structures and patterns [27]. This process is the act of collecting and examining data from different sources to gain insight from hidden patterns, trends, and outliers. It is the first step to fully exploit a health organization's data and inform on critical clinical decisions [14]. Through data discovery, data are gathered, combined, and analyzed in a sequence of steps. Exploration of data is the initial step in data analysis where healthcare providers explore an extensive data set in an unstructured manner to uncover original patterns and trends, characteristics, and points of interest [32]. During clinical research or undertaking, even during decision making by healthcare providers, patient or organizational records are analyzed to get an insight of previous patterns and behaviors. This process is not meant to reveal any patterns in detail, but rather to create a broad picture of important trends, as data exploration is extracting data without knowing what it is looking for [33]. Data exploration uses manual methods and automated tools, such as visualization tools. This process is vital in healthcare as it enables healthcare providers to cut down large sets of data to manageable sizes and focus on the efforts to optimize their analysis. Analysis of data follows just after the discovery and exploration of data to emerge viable solutions to a healthcare organization's problems [20].

VI. QUALITATIVE AND QUANTITATIVE DIAGNOSTIC ANALYTICS

These are the two major classifications of diagnostic analytics. Qualitative refers to interpreting meanings or processes, while quantitative aspect deals with numerals and statistical calculations of analysis [41]. Qualitative diagnostic analytics focuses on meanings rather than quantifiable phenomena. Qualitative diagnostic analytics puts more emphasis on the impact of the analysis on patients and healthcare providers as opposed to presumptions of the possibility of value-free inquiry [20]. Therefore, qualitative diagnostic analytics aims at gaining an understanding of underlying motivations and reasons, while uncovering dominant trends in thoughts and opinions. Quantitative diagnostic analytics, on the other hand, aim at quantifying data and generalizing results from a sample to the population of the study group [42].

A. Focus of Qualitative and Quantitative Methods

Qualitative research is widely used in research. It does not use statistics and mathematical operations to process data. Qualitative research focuses on using research methods to address phenomena by analyzing experiments [34-35]. Qualitative research focuses on the use of words or notes [36]. However, quantitative research uses methods, statistical processes, mathematics, and numerical data processing in a systematic experimental research of phenomena [37-38].

B. Data Collection

During the qualitative research, both primary and secondary data collection methods are used. Data could be collected from secondary sources such as social media platforms. LinkedIn

was used in [39], where authors looked for healthcare professionals, in order to get relevant data from prominent practitioners in the medical field. The research focused on big data analysis on LinkedIn as well as the management and surveillance of the diseases, while primary data were collected from interviews. People with relevant sources of information were interviewed in order to extract first-hand information. Some questions were developed before the study, while other questions emerged during the study in order to adequately cover all relevant areas.

C. Objective

The primary objective of health analytics is to provide insights in predicting and detecting the outcome of a disease, such as the prediction of heart-related diseases, HIV, and cancer [39]. Moreover, it provides opportunities to health professionals for improvement in their field of operation, through the appropriate patient management service and their diagnosis. Patients are served in the right manner, following the most suitable procedures.

TABLE I. COMPARISON OF QUALITATIVE AND QUANTITATIVE METHODS

Feature	Qualitative method	Quantitative method	References
Focus	Examining the breadth and depth of the phenomena, words and observations	Quantity (how much, number)	Focus on qualitative method: [34-35-36] Focus on quantitative method: [37-38]
Objective	Descriptive, exploration	Descriptive, prediction	Focus on objectives: [39]
Analysis	Inductive (by the researcher)	Deductive (statistical)	Focus on analysis: [40]
Data collection	Interviews, observation	Questionnaire, test	Focus on data collection: [39]
Findings	Comprehensive	precise, numerical	Focus on findings: [39]

VII. CLASSIFICATION OF DIAGNOSTIC ANALYTICS

A. Data Classification

Diagnostic data discovery aims at making noisy data clear for classification, free of inconsistencies, and easy to understand. Data discovery in healthcare is beneficial as it enables healthcare providers to gather actionable insights for better decision making, saves time, and makes data reusable. Data discovery consists of five steps: connecting and blending, cleansing and preparation, sharing, analysis and generation of insights, and visualization [24]. Data classification in diagnostic analytics aims on understanding the causes of events and behaviors.

B. Algorithms

Algorithms aim at answering why an event happened. Due to their changing nature, algorithms change very fast. Typically, algorithms utilize machine learning methods.

C. Methods/Tools

Statistical methods are used in diagnostic analytics, and data can be analyzed in the same manner as in descriptive analytics. Data can be analyzed electronically or manually using traditional methods [25]. This information is used by healthcare providers to shed light on entities or variables of study. In healthcare, data analysis is performed via statistical analysis using appropriate techniques [32].

D. Analysis

Healthcare analytics provide the most reliable information for a given patient. It also provides a given healthcare unit the right track about how its resources are being used [40], boosting accountability. The system is also used in detecting people who are most vulnerable to be infected by a particular infection. As a result, healthcare professionals are able to specify the most appropriate procedures to prevent the spread of a specific contagious disease. Moreover healthcare analytics play a vital role in communication process within a particular healthcare center.

E. Findings

Big data strategies under healthcare analytics have major contribution to the research for new drugs [39]. The future of the health sector relies on big data strategies, since they enhance effectiveness, improve relations among the various stakeholders in a given healthcare unit, and help in reducing operational costs. Some studies utilizing both qualitative and quantitative methods are shown in Table I.

D. Problem Domain

In diagnostic analytics, the problem domain analytics has to be elaborated more than in descriptive analytics. It also involves the process of creating a diagnosis model to describe the problems that are needed to be solved. In diagnosis, this deals with modeling solutions and curative measures for different diseases [32]. Once there is a healthcare problem domain analysis for a given facility, this can be applied to multiple implementations at different periods and as the technology advances. Problem domain analysis is wholly independent of the platform where the solution is implemented. Apart from descriptive analytics and diagnostics, there are predictive analytics and prescriptive analytics. Predictive analytics is used to predict the future values for identification of unknown events, while prescriptive analytics is used for providing the best solution [19].

VIII. PREDICTIVE HEALTH ANALYTICS

Predictive analytics involve mining information out of raw data in order to predict future trends [43]. Data mining and predicting is increasing progressively in the medical field. Support systems based on different prediction frameworks have become key tools in disease diagnosis in order to ensure prevention and the quality of prediction for a possible disease [44]. Disease prediction involves the extraction of hidden information and forecasting of the course of a disease development from medical data. Many studies have been conducted to develop frameworks for predicting diseases based on machine learning techniques [44]. Predictive health analytics is therefore the process of getting this information in

the health sector and processing it for exploiting in future predictions. Health is one of the key areas that predictive analytics has been utilized [43]. Predictive Analytics applies known parameters to produce a model that can predict results for a different set of data. Modeling gives this product as a prediction representing a probability of the variable being predicted based on estimates from input variables.

Most of the predictive modeling techniques in use include: decision trees, regression and neural networks. It has been noted that predictive analytics has been very useful in forecasting the pandemic of COVID-19 [43]. Healthcare institutions apply predictive analytics in various ways, such as in allocating resources based on previous trends, scheduling staff, gauging patient-risk to manage readmission costs, and manage pharmaceutical requirements. More than half of healthcare managers used predictive analytics in their organizations, and 57% of them believed that they were beneficial for their institutions [45]. In fact, 26% of them predicted saving over 25% of their total budget, as a result of applying such methods.

IX. DISCUSSION

This section discusses and compares descriptive and diagnostic health analytics. Diagnostic health analysis mostly answers the question "why something happened". Therefore, directed analysis and extensive exploration are required with the help of other tools, such as the use of visualization techniques that help uncover the genesis of problems, in order to help users to determine the extent to which these problems affect them [9]. For instance, it is possible to track down an increase in waiting times to provide several healthcare services, showing several factors that relate to patient, provider, or related organizational factors [8]. Descriptive analysis usually gives an overview of what happened. It often deals with the discovery and exploration of new information in a given dataset. It connects previous decisions and their implications providing the best basis for decision making. It is a branch of health analysis aiming to analyze data and come up with relevant solutions that can help in proper decision making about a given patient. Descriptive analysis quantifies events and transforms them into readable forms. In the health sector, descriptive analysis relies on clinical health support, public health, and pharmacy mental health among others, during the decision making process. On the other hand, diagnostic analysis enables any given healthcare unit to rely on data during the process of decision making, gaining competitive advantages over competing organizations. The diagnostic analysis is very essential since it gives detailed information about why certain things happened, for instance why a given patient situation worsened. This shows that diagnostic analysis explains the cause of events and behaviors, involving the discovery and exploration of any new data available. Both descriptive and diagnostic analyses are outstanding and very essential in the decision making process. Descriptive analysis relies on new information while the diagnostic analysis relies on already existing data. Consequently, the most reliable conclusion emerges from combining these two methods.

X. CONCLUSION

In general, healthcare analytics gives a deeper inspection of data in medical institutions. Generation and access of data has been made easier these days, since they are global and mobile. It is evident that big data offer the ability to capture and manage visual data and records. An effective application of health analytics and large scale data improves health results. The recommended descriptive and diagnostic analysis methods could be used by healthcare sectors in collecting and analyzing, patient and clinical data offering outstanding results and being essential in the decision-making process. Predictive analysis uses data in order to predict future trends. Descriptive analysis forms the easiest level of data by categorizing, characterizing, aggregating, and classifying them. Utilization of techniques such as data visualization in medical fields improves usability, which in turn enhances healthcare quality. Instead of replacing health professionals in order to advance medical facilities, the use of large scale data could form the backbone of predictive analytics, which is also the base line for future individual outcome prediction. Medical institutions should incorporate big data systems for improving health prediction and results.

REFERENCES

- [1] Z. A. Al-Sai and L. M. Abualigah, "Big data and E-government: A review," in *2017 8th International Conference on Information Technology (ICIT)*, May 2017, pp. 580–587, <https://doi.org/10.1109/ICITECH.2017.8080062>.
- [2] S. Lavalle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big Data, Analytics and the Path From Insights to Value," *MIT Sloan Management Review*, vol. 52, no. 2, pp. 21–32, Feb. 2011.
- [3] M. S. Hossain and G. Muhammad, "Healthcare Big Data Voice Pathology Assessment Framework," *IEEE Access*, vol. 4, pp. 7806–7815, 2016, <https://doi.org/10.1109/ACCESS.2016.2626316>.
- [4] N. Mehta and A. Pandit, "Concurrence of big data analytics and healthcare: A systematic review," *International Journal of Medical Informatics*, vol. 114, pp. 57–65, Jun. 2018, <https://doi.org/10.1016/j.ijmedinf.2018.03.013>.
- [5] V. Agarwal, "Research on Data Preprocessing and Categorization Technique for Smartphone Review Analysis," *International Journal of Computer Applications*, vol. 131, pp. 30–36, 2015, <https://doi.org/10.5120/ijca2015907309>.
- [6] S. B. Merriam, *Qualitative Research and Case Study Applications in Education. Revised and Expanded from "Case Study Research in Education."* San Francisco, CA, USA: Jossey-Bass Publishers, 1998.
- [7] G. M. Kakandikar and V. M. Nandedkar, "Big Data in Healthcare: Technical Challenges and Opportunities," in *Big Data Analytics in Healthcare*, A. J. Kulkarni, P. Siarry, P. K. Singh, A. Abraham, M. Zhang, A. Zomaya, and F. Baki, Eds. Cham, Switzerland: Springer International Publishing, 2020, pp. 67–81.
- [8] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: management, analysis and future prospects," *Journal of Big Data*, vol. 6, no. 1, p. 54, Jun. 2019, <https://doi.org/10.1186/s40537-019-0217-0>.
- [9] M. S. Islam, M. M. Hasan, X. Wang, H. D. Germack, and M. Noor-E-Alam, "A Systematic Review on Healthcare Analytics: Application and Theoretical Perspective of Data Mining," *Healthcare*, vol. 6, no. 2, p. 54, Jun. 2018, <https://doi.org/10.3390/healthcare6020054>.
- [10] M. Alkhatib, A. Talaei-Khoei, and A. Ghapanchi, "Analysis of Research in Healthcare Data Analytics," presented at the Australasian Conference on Information Systems, Sydney, Australia, Dec. 2015.
- [11] G. Cai, "Big Data Analytics in Structural Health Monitoring," Ph.D. dissertation, Vanderbilt University, Nashville, TN, USA, 2017.
- [12] C. Lee *et al.*, "Big Healthcare Data Analytics: Challenges and Applications," in *Handbook of Large-Scale Distributed Computing in*

- Smart Healthcare*, S. U. Khan, A. Y. Zomaya, and A. Abbas, Eds. Cham, Switzerland: Springer International Publishing, 2017, pp. 11–41.
- [13] Y. Wang, L. Kung, W. Y. C. Wang, and C. G. Cegielski, "An integrated big data analytics-enabled transformation model: Application to health care," *Information & Management*, vol. 55, no. 1, pp. 64–79, Jan. 2018, <https://doi.org/10.1016/j.im.2017.04.001>.
- [14] Y. Wang and N. Hajli, "Exploring the path to big data analytics success in healthcare," *Journal of Business Research*, vol. 70, pp. 287–299, Jan. 2017, <https://doi.org/10.1016/j.jbusres.2016.08.002>.
- [15] B. Nithya and V. Ilango, "Predictive analytics in health care using machine learning tools and techniques," in *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, Jun. 2017, pp. 492–499, <https://doi.org/10.1109/ICCONS.2017.8250771>.
- [16] Y.-K. Lin *et al.*, "Healthcare Predictive Analytics for Risk Profiling in Chronic Care: A Bayesian Multitask Learning Approach," *MIS Quarterly*, vol. 41, no. 2, pp. 473–495, Feb. 2017, <https://doi.org/10.25300/MISQ/2017/41.2.07>.
- [17] N. D. Shah, E. W. Steyerberg, and D. M. Kent, "Big Data and Predictive Analytics: Recalibrating Expectations," *JAMA*, vol. 320, no. 1, pp. 27–28, Jul. 2018, <https://doi.org/10.1001/jama.2018.5602>.
- [18] M. M. Malik, S. Abdallah, and M. Ala'raj, "Data mining and predictive analytics applications for the delivery of healthcare services: a systematic literature review," *Annals of Operations Research*, vol. 270, no. 1, pp. 287–312, Nov. 2018, <https://doi.org/10.1007/s10479-016-2393-z>.
- [19] R. Kunnivil and N. S. Murthy, "Healthcare Data Utilization for the Betterment of Mankind - An Overview of Big Data Concept in Healthcare," *International Journal of Healthcare Education & Medical Informatics*, vol. 5, no. 2, pp. 14–17, Aug. 2018, <https://doi.org/10.24321/2455.9199.201807>.
- [20] H. B. Patel and S. Gandhi, "A Review on Big Data Analytics in Healthcare Using Machine Learning Approaches," in *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, May 2018, pp. 84–90, <https://doi.org/10.1109/ICOEI.2018.8553788>.
- [21] G. Rao *et al.*, "Identifying, Analyzing, and Visualizing Diagnostic Paths for Patients with Nonspecific Abdominal Pain," *Applied Clinical Informatics*, vol. 9, no. 4, pp. 905–913, Oct. 2018, <https://doi.org/10.1055/s-0038-1676338>.
- [22] M. Khalifa and I. Zabani, "Utilizing health analytics in improving the performance of healthcare services: A case study on a tertiary care hospital," *Journal of Infection and Public Health*, vol. 9, no. 6, pp. 757–765, Nov. 2016, <https://doi.org/10.1016/j.jiph.2016.08.016>.
- [23] A. Cornuéjols and L. Miclet, *Apprentissage artificiel: concepts et algorithmes*. Paris: Eyrolles, 2003.
- [24] C.-W. Tsai, M.-C. Chiang, A. Ksentini, and M. Chen, "Metaheuristic Algorithms for Healthcare: Open Issues and Challenges," *Computers & Electrical Engineering*, vol. 53, pp. 421–434, Jul. 2016, <https://doi.org/10.1016/j.compeleceng.2016.03.005>.
- [25] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Turkey, *Graphical Methods for Data Analysis*. Boca Raton, FL, USA: CRC Press, 2017.
- [26] R. Pastorino *et al.*, "Benefits and challenges of Big Data in healthcare: an overview of the European initiatives," *European Journal of Public Health*, vol. 29, no. Supplement_3, pp. 23–27, Oct. 2019, <https://doi.org/10.1093/eurpub/ckz168>.
- [27] S. Stieglitz, M. Mirbabaie, B. Ross, and C. Neuberger, "Social media analytics – Challenges in topic discovery, data collection, and data preparation," *International Journal of Information Management*, vol. 39, pp. 156–168, Apr. 2018, <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>.
- [28] P. K. D. Pramanik, S. Pal, and M. Mukhopadhyay, in *Healthcare Big Data: A Comprehensive Overview*, Hershey, PA, USA: IGI Global, 2019.
- [29] A. Khan, H. Mukhtar, H. F. Ahmad, M. A. Gondal, and Q. M. Ilyas, "Improving Usability through Enhanced Visualization in Healthcare," in *2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)*, Bangkok, Thailand, Mar. 2017, pp. 39–44, <https://doi.org/10.1109/ISADS.2017.36>.
- [30] A. Banerjee, T. Bandyopadhyay, and P. Acharya, "Data Analytics: Hyped Up Aspirations or True Potential?," *Vikalpa: The Journal for Decision Makers*, vol. 38, no. 4, pp. 1–12, Oct. 2013, <https://doi.org/10.1177/0256090920130401>.
- [31] M. I. Dascalu, C. N. Bodea, E. L. Trifan, E. Lazarou, and M. Nitu, "Learning analytics in an e-testing application: Premises and conceptual modelling," in *Proceedings of the 14th International Scientific Conference "eLearning and Software for Education"*, Bucharest, Romania, Apr. 2018, vol. 2, pp. 239–246.
- [32] Stefan Thaler and Vlado Menkovski, "The Role of Deep Learning in Improving Healthcare," in *Data Science for Healthcare*, Cham, Switzerland: Springer, 2019, pp. 75–116.
- [33] S. Idreos, O. Papaemmanouil, and S. Chaudhuri, "Overview of Data Exploration Techniques," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, New York, NY, USA, May 2015, pp. 277–281, <https://doi.org/10.1145/2723372.2731084>.
- [34] Sharan B. Merriam and Elizabeth J. Tisdell, *Qualitative Research: A Guide to Design and Implementation*. Hoboken, NJ, USA: John Wiley & Sons, 2015.
- [35] M. Hennink, I. Hutter, and A. Bailey, *Qualitative Research Methods*. Newbury Park, CA, USA: SAGE, 2020.
- [36] D. Amaratunga, D. Baldry, M. Sarshar, and R. Newton, "Quantitative and qualitative research in the built environment: application of 'mixed' research approach," *Work Study*, vol. 51, no. 1, pp. 17–31, Jan. 2002, <https://doi.org/10.1108/00438020210415488>.
- [37] A. Bhattacharjee, *Social science research: principles, methods, and practices*. Global Text Project, 2012.
- [38] R. Kumar, *Research Methodology: A Step-by-Step Guide for Beginners*. Newbury Park, CA, USA: SAGE, 2018.
- [39] N. Das, S. Rautaray, and M. Pandey, "Big Data Analytics for Medical Applications," *International Journal of Modern Education and Computer Science*, vol. 10, Feb. 2018, <https://doi.org/10.5815/ijmecs.2018.02.04>.
- [40] B. Saenyi, "Opportunities and challenges of Big Data Analytics in healthcare: An exploratory study on the adoption of big data analytics in the Management of Sickle Cell Anaemia," M.S. thesis, Jönköping University, Jönköping, Sweden, 2018.
- [41] C. Hammond, "Interpreting qualitative data (4th Ed.) by D. Silverman," *British Journal of Psychology*, vol. 103, no. 4, pp. 576–579, 2012, <https://doi.org/10.1111/j.2044-8295.2012.02131.x>.
- [42] Al-Khasawneh, Mohammad, "An Exploration of Consumer Response Towards Sponsored Search Advertising (SSA) from a Consumer Behaviour Perspective," Ph.D. dissertation, Griffith University, Brisbane, Australia, 2010.
- [43] S. F. Ardabili *et al.*, "COVID-19 Outbreak Prediction with Machine Learning," *Algorithms*, vol. 13, no. 10, p. 249, Oct. 2020, <https://doi.org/10.3390/a13100249>.
- [44] B. Trstenjak, D. Donko, and Z. Avdagic, "Adaptable Web Prediction Framework for Disease Prediction Based on the Hybrid Case Based Reasoning Model," *Engineering, Technology & Applied Science Research*, vol. 6, no. 6, pp. 1212–1216, Dec. 2016, <https://doi.org/10.48084/etasr.753>.
- [45] P. Galetsi and K. Katsaliaki, "Big data analytics in health: an overview and bibliometric study of research activity," *Health Information & Libraries Journal*, vol. 37, no. 1, pp. 5–25, 2020, <https://doi.org/10.1111/hir.12286>.