RESEARCH ARTICLE OPEN ACCESS

Role of Big Data in Revolutionizing Health Management Systems

Nithin Narayan Koranchirath
Subject Matter Expert (SME), Leading Heath Insurance Company, Richmond, United States

ABSTRACT:

Big data is crucial and popular elements for innovation and predictive analytics in healthcare, propelling the digital healthcare transformation. Organizations are already developing an intelligent big data analytics platform based on data integration principles. This platform serves as the new cornerstone for the organization to improve population health management, value-based care, and address emerging healthcare challenges. Utilizing this new data platform for community and population health offers several benefits, including improved healthcare outcomes, enhanced clinical operations, reduced care costs, and the generation of accurate medical information. The authors have implemented multiple advanced analytics framework that leverage the large, standardized datasets integrated into the platform to enhance the effectiveness of public health interventions, improve diagnoses, and provide clinical decision support. The data integrated into the platform are sourced from Electronic Health Records, Laboratory Information Systems, Hospital Information Systems and Radiology Information Systems as well as data generated by public health platforms, mobile data, social media, and clinical web portals. This substantial volume of data is integrated using big data techniques for storage, retrieval, processing, and transformation. This paper outlines the design of a digital health platform within a healthcare organization, aiming to integrate operational, clinical, and business data repositories with advanced analytics to enhance the decision-making process for population health management.

KEYWORDS: decision support systems; population health management; big data; advanced analytics; personalized patient care.

INTRODUCTION

World over several initiatives now focused on developing digital healthcare platforms with collaborative access tools. These tools are specifically designed to facilitate the exchange and sharing of information and knowledge throughout the healthcare process [1]. The primary goal of these frameworks and architectures is to optimize the quality and efficiency of patient care, while also ensuring appropriate attention to patient conditions and risks.

For example, the implementation of digital health services aims to improve efficiency and enhance security in patient care. Another organization is working on the development of a healthcare cyber-physical system assisted by big data. This system comprises a data integration layer, a data management layer, and a data analytics service layer, all aimed at enhancing the functioning of the healthcare system.

Medical organizations are increasingly investing in the development of healthcare platforms that integrate data, applications, business processes, and user interfaces to gain knowledge and insights for clinical decisions, drug recommendation systems, and better disease diagnoses. Various big data applications in healthcare, such as healthcare monitoring and analytical platforms, have been developed to detect errors in medical prescriptions and clinical errors, as well as to apply advanced analytics frameworks for predictive analytics based on large volumes of clinical data [2]. Platforms for genomics data analytics and healthcare knowledge systems are also being implemented to generate predictions based on DNA molecular changes and mutations, supporting clinical decision-making and diagnosis.

These initiatives reflect a growing emphasis on decision support systems, population health management, big data, advanced analytics, and personalized patient care in the development of digital healthcare platforms [3].

The healthcare landscape is undergoing significant changes due to big data and advanced analytics implementations. These advancements are impacting disease research and increasing the complexity of population health management [4]. The traditional fee-for-service approach is being replaced by the value-based care model, leading to a shift in healthcare goals for the future.

Population health management involves aggregating patient health data from various sources, analyzing and transforming it into actionable insights to drive informed decisions aimed at improving clinical and financial outcomes.

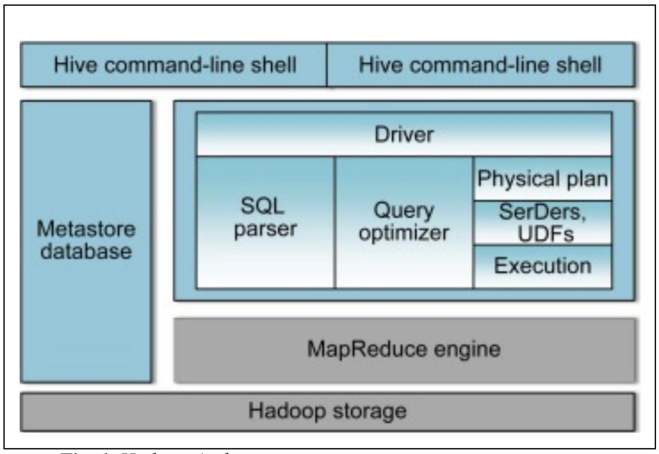

Fig. 1. Hadoop Architecture

Big data technologies allow for the consolidation of structured and unstructured data from disparate sources into a central repository for examination and analysis [5]. Advanced analytics play a crucial role in uncovering insights from complex datasets, including identifying unseen patterns, making new predictions, and analyzing health data trends. Advanced analytics is being applied across various clinical domains to analyze numerous clinical parameters, resulting in more effective and efficient analytics for improving medical care outcomes and quality.

The implementation of this platform demonstrates the significant potential of using big data to personalize medical treatment, improve patient lives, deliver better medical care, and reduce operational waste. Other potential uses for applying big data in healthcare include allowing doctors to assess a patient's risk of dependency before prescribing treatment, enabling the development of customized treatment plans, using psychosocial and clinical medical data to accurately identify and diagnose chronic diseases, gaining insights into performance, improving clinical care, and restructuring care delivery to enhance patient outcomes using the platform's analytics to enhance the quality of care and patient experience at a reduced cost, in line with the organization's core objectives etc.

Big Data Health Platform

Implementing big data platform is a complex process that involves integrating health data from various sources, developing multiple technological and functional components, and establishing robust IT management processes to ensure interoperability with other systems. The digital health information platform encompasses a wide range of data, including patient information, EHR, diagnostic reports, prescriptions, medical images, pharmacy records, research data, operational data, financial data, and human resources data.

This platform represents a significant innovation as it pioneers the design and construction of a comprehensive digital health platform for healthcare organizations, with a primary focus on putting the patient at the center and consolidating all relevant information based on a standardized enterprise data repository [6]. It enables quick and intuitive access to information as needed, while concealing technical complexities and providing tools for longitudinal process management and decision support for healthcare professionals.

The platform stands out due to the development of a medical portal featuring a 360-degree view of the patient, using data from the enterprise data repository to generate real-time early warning scores, patient surveillance, open APIs for hospital integration, prediction of health risk patterns, identification of high-risk markers, and assessment of co-morbidities.

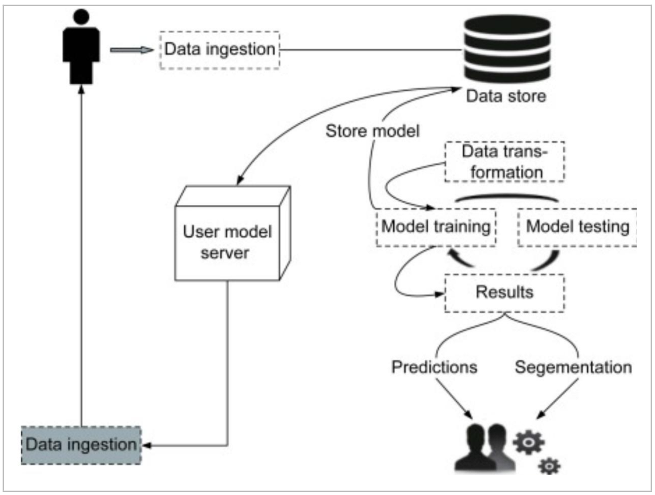

***Fig. 2**. Big Data Health platform overview*

Furthermore, the platform incorporates advanced technologies such as big data, advanced analytics to support decision-making processes and enhance population health management, ultimately contributing to the delivery of personalized patient care.

ARCHITECTURE

The initial approach involved the creation of a big data processing pipeline using a Big Data Hadoop architecture to support both real-time and batch analytics. This architecture encompasses various methods for ingesting data, tailored to the source and required timing for generating insights. Furthermore, this approach facilitates contributions from professionals with diverse skill sets to collectively build the platform.

The architecture comprises a data curation layer, a diverse real-time data layer, and a data analytics layer [6]. The PaaS layer oversees persistent storage and has the capability to scale horizontally. The real-time layer is responsible for processing streaming data, executing dynamic computation, importing data into repositories, and leveraging prediction analytics.

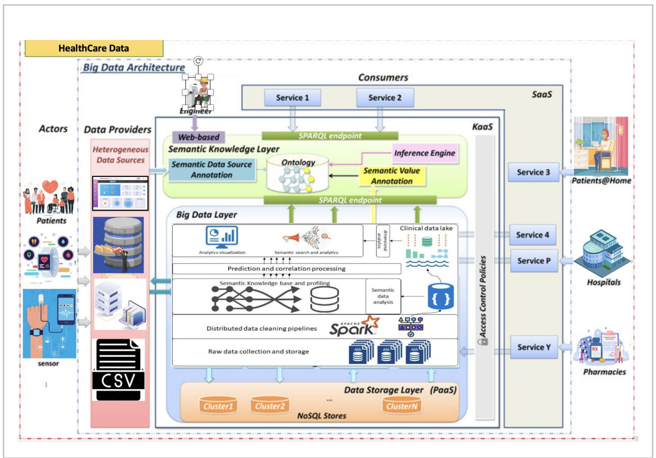

***Fig. 3**. Big Data Health platform overview*

From an infrastructure standpoint, the platform offers the flexibility to be deployed in local data processing centers through the utilization of virtualization techniques, in memory processing, and load distribution systems. The infrastructure design was developed to provide a versatile set of resources that can be scaled up based on specific workload requirements [7]. The infrastructure deployment heavily relied on automation to ensure seamless operations.

This platform integrates decision support systems, population health management, big data, advanced analytics and personalized patient care.

DATA STORAGE LAYER

A comprehensive evaluation is currently underway to establish a staging repository for the big data analytics platform. Big data enables the consolidation of data of various sizes, types, and ingestion speeds into a single location for the storage of diverse data types. The suggested staging area encompasses essential features such as security, scalability, reliability, and availability. Data processing can occur directly from this staging area, or the data can be ingested into a data storage layer for historical loading, preparation, and serving for decision support systems, population health management, and the requirements of advanced analytics applications [8]. The data warehouse repository features a scale-out architecture and a massively parallel processing (MPP) engine.

Data analytics have been developed to cover clinical, social, and healthcare program domains. Data analytics validates and processes incoming data, effectively segregating the data processing and administration from the data source. These data analytics can be expanded to include additional attributes specific to the implementation, enabling them to subscribe to particular types of messages using the mapping and filtering options provided by the data processing pipelines. Once these subscriptions are established, the results will be loaded with all pertinent messages for the subscribed entities and stored in the big data.

Regarding data storage, the data is loaded into a data warehouse with a daily refresh. This healthcare data repository features a highly normalized data model tailored to facilitate swift and efficient querying and analysis, addressing the demands of personalized patient care.

ADVANCED ANALYTICS LAYER

The platform's advanced analytics layer is capable of processing real-time data essential for operations and providing comprehensive delivery services. It incorporates in-memory processing, data encryption, and compliance with international security standards such as HIPAA.

Within the analytic data component, there are BI analytics designed for tactical, operational, and strategic decision-making, as well as a set of prediction analytics. These insights assist healthcare providers in identifying and monitoring chronic diseases.

The advanced analytics component is utilized to develop, test, utilize, and deploy predictive analytic as required by the organization. The platform offers self-service dashboards and visualizations that leverage data to facilitate the decision-making process. The advanced analytics application layer is a vital aspect of this platform.

After the data has been integrated, aggregated, and standardized within the system, the platform provides a tool to manage knowledge through the business intelligence interface, delivering data analysis, design, and advanced analytics [9]. It also facilitates the development and management of results-based care indicators and population health management. Furthermore, the platform includes a tool for clinicians, researchers, and scientists to extract valuable information from the mined data.

Advanced analytics frameworks can be customized in preconfigured data domains, enabling the storage of results for future utilization. Data researchers and scientists can develop advanced tools to extract information and value from the data stored in the solution, leveraging the design and validation component. This platform succinctly showcases the implemented predictive analytics.

Key Performance Indicator

Healthcare organizations face growing pressure to reduce care costs, with population health management teams at the forefront. Big data insights are pivotal in standardizing production processes and enhancing overall healthcare efficiency. For instance, big data applications can monitor performance and identify patients not receiving appropriate care as per guidelines. This entails using graphs to identify care gaps, assess risk profiles, pinpoint prescription gaps for high-risk populations, and track progress in meeting guideline targets for conditions such as high blood pressure and high cholesterol.

Advanced analytic tools offer valuable insights into care and cost variation, as well as improvement opportunities and performance tracking. Performance dashboards can showcase individual and combined indicators, facilitating benchmarking among peers and entities for a specified reporting period. These reports enable the tracking of performance benchmarks over time and can be shared with clinical teams, along with specific metrics linked to incentives for closing care gaps.

Decision support systems can provide customized alerts for patients receiving suboptimal care, empowering administrative staff to establish recall and reminder alerts for management consultations [10]. These tools aid in optimizing population health management by addressing care gaps and serving as incentives to limit avoidable variation in established cardiovascular care processes. The use of big data and personalized patient care is critical in meeting the evolving demands of healthcare management.

Benefits Of Big Data Platform

In the realm of limited public health resources, it is crucial to go beyond simply collecting and presenting data and instead use data to generate concepts and insights. This shift is essential for developing and evaluating effective practices and policies at local, state, and federal levels to enhance outcomes [11]. For example, data can be used to tailor clinical care for underserved patients and collaborate with local welfare workers to address underlying health determinants.

While previous attempts to extract knowledge from data mining have led to new discoveries, their practical application in clinics has been limited. New data science tools have the potential to reveal strong and hidden connections between social factors within extensive datasets. For instance, advanced analytics can assist researchers in predicting future trends using data from sources such as the Behavioral Risk Factor Surveillance System and the American Community Survey.

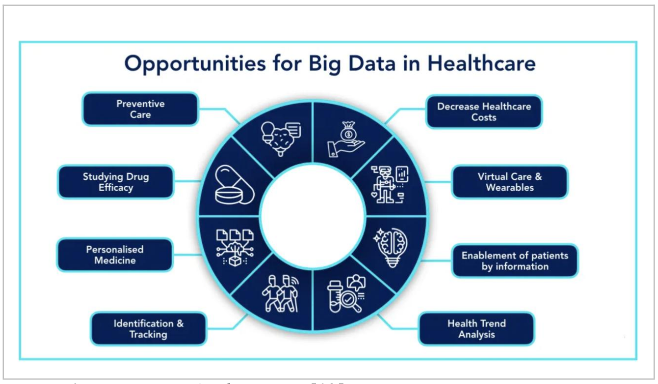

Fig 4 Big Data Applications [12]

The emergence of big data on a population level enables the reevaluation and adjustment of existing associations, such as cardiovascular risk factors and outcomes. Recent analyses have integrated neighborhood and individual clinical factors to enhance the depth of patient risk profiles for cardiovascular disease and heart failure.

Additionally, traditional statistical analysis tools can be repurposed in population health with big data.

Despite its potential, the utilization of big data in clinical practice is still in its early stages. Current clinical practice could support the integration of individual-level clinical data with community data, potentially by incorporating new codes to capture patients at risk of non-medical issues like food and housing. This approach could enable doctors to focus on patient-centered care and empower public health practitioners to devise strategies based on social determinants, ultimately strengthening the patient-physician relationship and providing more personalized care. This shift aligns with the growing emphasis on decision support systems, population health management, big data, advanced analytics, and personalized patient care.

Limitations Of Big Data Platform

The use of big data comes with certain limitations, particularly in securely storing large volumes of data. This is especially relevant when handling extensive datasets such as EMR, claims, and other sources of big data. Factors like missing data, coding errors, and the over or under-representation of specific medical conditions can all affect the reliability of findings derived from big data. Addressing these concerns necessitates the use of appropriate methodological tools and expertise in statistical analysis.

Observational data is inherently limited by selection and information bias, confounding, and other study design constraints. These limitations must be considered to ensure the accuracy and generalizability of estimates [12]. Operational challenges, such as resistance from healthcare providers, potential errors in systems, and limitations in data interoperability, also need to be taken into account.

Moreover, while mobile and wearable devices offer new opportunities for data sharing and patient engagement, accessibility to these technologies may be restricted, and privacy issues related to wearable technology can create further barriers.

Despite these limitations, the widespread application of big data, supported by innovative tools, presents exciting opportunities for evidence-based care management at a population level. This includes clinical risk reduction, targeted medical interventions, and enhanced patient satisfaction, demonstrating the potential for personalized patient care and population health management facilitated by big data decision support systems.

Conclusion

This discussion introduces a cutting-edge, secure healthcare platform that is revolutionizing the healthcare industry through the provision of advanced information to patients and care teams. The platform's implementation has resulted in reduced healthcare costs.

The proposed digital health platform facilitates the resolution of population health challenges, offers a comprehensive understanding of patient health, and uncovers hidden patterns that may be overlooked by traditional data analytics. By utilizing unified patient-generated data, financial data, and socioeconomic data, the organization can identify patterns and groups of patients with similar health behaviors. The analysis of clinical and non-clinical data enables more precise predictions of patient health. Additionally, the platform simplifies the identification of health insights and the execution of actions based on treatment history for both individuals and patient groups.

Every organization strives to enhance quality outcomes, provider engagement and recruitment, and its economic health. To accomplish these goals, the organization emphasizes clinician engagement and organizational alignment, ensuring widespread access to meaningful, actionable data, and leveraging the healthcare analytics platform to guide decisions and drive improvements [13]. Integration of machine learning into the organization represents one of the most critical and life-saving technologies ever implemented. The platform's potential to improve and expedite clinical, workflow, and financial outcomes is virtually limitless.

Further enhancements are vital for the platform to continuously maximize all the benefits for the entire organization. Tools for conducting the knowledge discovery process will be integrated into the ecosystem. The organization intends to begin implementing prescriptive analytics models to assist in making informed decisions in population health management. The architecture team will explore the feasibility of incorporating Map/Reduce-based computations for processing data with high scalability and executing low latency and high concurrency analytical queries on top of Hadoop clusters.

References

[1] WHO, "Global Health and Aging," 2011. [Online]. Available: http://www.who.int/ageing/publications/globalhealth.pdf.

[2] N. Krishna Chythanya and L. Rajamani, "Neural Network Approach for Reusable Component Handling," in Proceedings - 7th IEEE International Advanced Computing Conference, IACC 2017, art. no. 7976764, pp. 75-79. [Online]. Available: https://doi.org/10.1109/iacc.2017.0030

[3] M. Chen, S. Mao, and Y. Liu, " Mobile Networks and Applications, vol. 19, no. 2, 2014.

[4] S. Shahrivari, "Beyond batch processing: towards real-time and streaming big data," Computer Journal, vol. 3, no. 4, pp. 117–129, 2014.

[5] J. Dean and S. Ghemawat, "MapReduce: simplified data processing," Communications of the ACM, vol. 51, no. 1, pp. 107–113, 2008.

[6] N. Tatbul, "Streaming data integration: Challenges and opportunities," in Proceedings of the 2010 IEEE 26th International Conference on Data Engineering Workshops, ICDEW 2010, pp. 155–158, USA, March 2010.

[7] D. Singh and C. K. Reddy, "A survey on platforms for big data analytics," Journal of Big Data, vol. 2, no. 1, 2015.

[8] B. Dhanalaxmi, G. Apparao Naidu, and K. Anuradha, "Adaptive PSO based association rule mining technique for software defect classification using ANN," in Procedia Computer Science, 2015, vol. 46, pp. 432-442. [Online]. Available: https://doi.org/10.1016/j.procs.2015.02.041

[9] L. Neumeyer, B. Robbins, A. Nair, and A. Kesari, "S4: Distributed stream computing platform," in Proceedings of the 10th IEEE International Conference on Data Mining Workshops (ICDMW ’10), pp. 170–177, Sydney, Australia, December 2010.

[10] K. Padmavathi and K. S. R. Krishna, "Myocardial infarction detection using magnitude squared coherence and Support Vector Machine," in International Conference on Medical Imaging, m-Health and Emerging Communication Systems, Med Com 2014, art. no. 7006037, pp. 382-385. [Online]. Available: https://doi.org/10.1109/medcom.2014.7006037.

[11] M. Zaharia, R. S. Xin, P. Wendell et al., "Apache spark: A unified engine for big data processing", vol. 59, no. 11, 2016.

[12] Mona Ciotta, " medudoc education GmbH," 2020. [Online]. Available: https://medium.com/medudoc/how-big-data-is-improving-healthcare-d88f2b9269c0.

[13] T. Huang, L. Lan, X. Fang, P. An, J. Min, and F. Wang, "Promises and challenges of big data computing in health sciences", vol. 2, no. 1, pp. 2–11, 2015.