Applications and analytics of bioinformatics, healthcare informatics for modern healthcare system

Kshatrapal SINGH[1*], Ashish KUMAR[2], Manoj Kumar GUPTA[3]

[1] Krishna Engineering College, Ghaziabad, India

[2] I.T.S Engineering College, Greater Noida, India

[3] School of CSE, Shri Mata Vaishno Devi University, Katra, India

mekpsingh1@gmail.com, ashishcse29@gmail.com, manoj.cst@gmail.com

***Corresponding Author:** Kshatrapal SINGH
mekpsingh1@gmail.com

Abstract: Healthcare informatics emphasizes on the gathering, processing, assessment, and utilization of health records, information, and expertise. By connecting genetics, proteomics, and patient records, bioinformatics uses computational methods and tools to research and analyze huge biological datasets, completely understand illness, and grasp genetics and proteomics. Interpreting genomic and proteomic data for basic biological research, as well as for use in healthcare, drug discovery, and related fields, is the main focus. Large amounts of electronic data collection in the healthcare industry lead to the development of analytics. Data analytics is adept at enhancing healthcare, cutting costs, and preserving people. The explosion of data leads to the use of data analytics in the healthcare industry to explore understandings and make smart judgments. In order to create a better health service, this paper focuses on bioinformatics, healthcare informatics, and analytics. It examines the advantages, how each one helps to improve the health service, how bioinformatics, health informatics, and analytics connect, and lastly, what the coming times holds for healthcare informatics as well as analytics.

Keywords: Bioinformatics, Healthcare Informatics, Healthcare System, Analytics.

1. Introduction

A community must protect its citizens' mental and physical well-being in order to accomplish its dream and vision of sustainable progress. Healthcare is a crucial area of any economy. As a result, a country's ability to flourish depends on its residents' overall health, giving rise to the proverb "health is riches." So, it is highly likely that now the people will accept any innovation or method that is designed to improve the healthcare system (Flygare et al., 2016) (Ondov et al., 2016). Since the Human Genome Project's completion in 2003, there have been possibilities to obtain, analyze, and disseminate genetic data in more efficient, affordable, and trustworthy ways. Healthcare has been significantly impacted by this significant technical advancement in terms of novel prevention creation, better diagnosis, and early treatment for routine healthcare. It is predicted that these massive amounts of data from the bioinformatics as well as healthcare informatics areas, combined with analytics, will soon give preventive, predictive, and personalized healthcare solutions. It is anticipated that this will aid in the prosperity and progress of the country.

Experts now have the option to keep data, like DNA comparison, analysis, as well as interpretation, on machines in the format of databases due to bioinformatics. Even seasoned specialists cannot use biological data inside its unprocessed raw form. Therefore, the development of biological instruments to gather data from these databases to be used in the study has been made possible by bioinformatics. To understand the sequencing, structures, and activities of proteins, bioinformatics techniques are essential (Thankachan et al., 2017). A greater comprehension of an organism's biology results from having a thorough awareness of these three concepts.

Since the information in Big Data is so private, using it in healthcare comes with significant ethical and legal issues. The possibility of compromising one's privacy and autonomy, in addition to repercussions on the public's need for clarity, faith, and impartiality when using Big Data, are among the ethical and legal challenges. Critical technical and architectural concerns that could jeopardize a Big-Data-driven healthcare system include data diversity, data security, analytical processes in data analysis, and the lack of adequate infrastructures for file storage.

https://doi.org/10.33436/v32i4y202206

Although it is simple to maintain databases of basic biological data in systems, the real problem for experts is to retrieve the necessary information from the huge amount of data. Because of this, bioinformatics tools are essential for retrieving and interpreting data from the databases. The bioinformatics tools are made to be simple for biologists in using.

2. Literature review

In Situ Hybridization approach was used to map 579 human genes, which is when bioinformatics began. Ever since Marvin Carruthers as well as Leory Hood developed an automatic DNA sequencing technology in 1981, the field of bioinformatics experienced tremendous growth. The first comprehensive genome mapping was released in 1989 by the Human Genome Organization, a global group of scientists working on the Genome Sequencing.

The Project was started in 1990, and a sum of 1879 mankind genes were analyzed by the end of 1991. A topological mapping of the human genome was created in 1993 in France by the human genome-focused research facility Genethon. When Genethon released the full study on the Human Genetic Mapping after three years, the very first stage of the Human Genome project was complete.

After the Human Genome Experiment was finished in 2003, biologists, computer scientists, as well as mathematicians joined forces to create the field of bioinformatics. The bioinformatics field developed out of the necessity to evaluate genome data, gain a solid understanding of the genes associated with specific disorders, and reveal the complexity of such genes' biological activity. Thus, the administration, analytics, and analysis of data from biological data and components constitute the interdisciplinary field of bioinformatics.

The requirement to create huge databases, like EMBL, GenBank, as well as DNA, to collect and compare the Genomic DNA input produced by the other genome sequencing initiatives and the human genome significantly improved bioinformatics (Lu et al., 2017) (Wu et al., 2017). Nowadays, bioinformatics includes data from patient's preclinical and clinical studies, computational modeling of proteins, functional information about proteins and genes, and the metabolic functions of many organisms.

3. Bioinformatics

The necessity to organize the massive amounts of data produced by molecular biology techniques led to the development of bioinformatics. The technique exploits methods for locating the locations of protein-coding regions in DNA as well as methods for large-scale DNA sequencing (Refer Figure 1). Hybridization, image scanning, quantification, preprocessing, correction of background, normalization and summarization are required to perform for microarray. Whereas sequencing, base call and alignment to reference genome are needed to perform for RNA seq. Additionally, sequence repository construction took place concurrently.

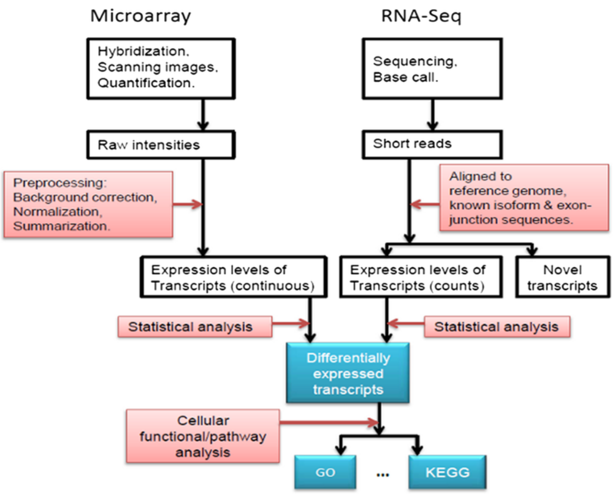

Figure 1. Bioinformatics

A new tool called Multiple Gene Expression Assessment is indeed producing a vast amount of fresh information. The idea is to find DNA chips with thousands of DNA nucleotide segments on them, with each person characterizing a gene to gene Encoding taken from an interest source. Numerous gene expression techniques are frequently employed to find subsets of genes. This enables monitored categorization, another term for the ability to distinguish among more than two biological circumstances. The categorization of samples and genes is accomplished by unsupervised classification, which seeks for groupings in the gene sample space. Improvements in computational models from computational techniques with their basis in computer science and mathematics were made possible by the study of gene expression information (Badea & Ţilivea, 2006). Bioinformatics tools for bio prospecting from meta-genomic sequence are shown in Figure 2. Sequencing, binning, assembly, gene prediction, gene annotation and data sharing is done with bioinformatics tools.

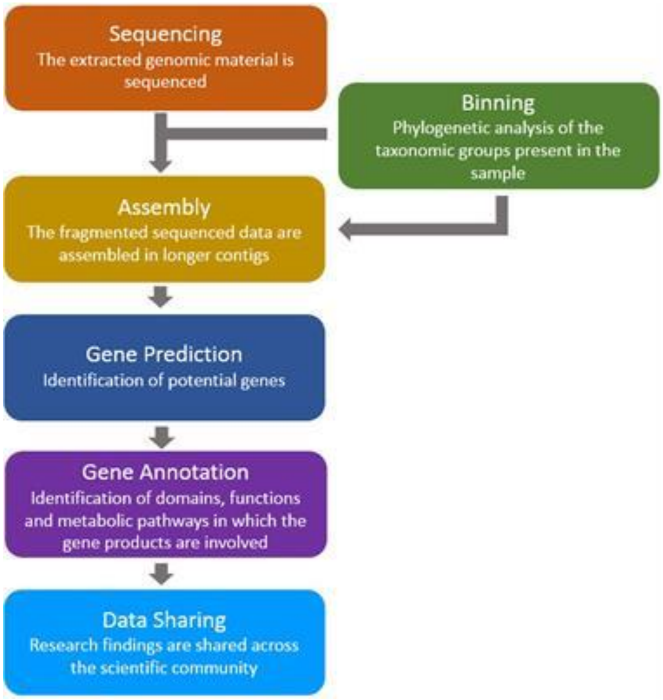

Figure 2. Bioinformatics tools

Additionally, significant achievements have been made in data mining techniques for genomes as well as proteomics. The accomplishments have enriched and increased the understanding of genes connected to a certain illness and the identification of new, highly efficient medications. Additionally, a lot of national and international bodies and associations support the inclusion of genetic data in electronic healthcare records (Navarro-Gomez et al., 2015). The organizations' endeavor is certain to have a significant impact on personalized medicine. The forefront concept of personalized medicine, which selects best treatment for the patients based on their genetic makeup, has been enhanced by advances in genetic investigation.

4. Healthcare informatics

Healthcare informatics focuses on making the greatest utilization of information through the application of technologies to modern clinical research, global health, and healthcare. Information seems to be more important than technique in this case. To improve people's lives, informatics combines computer and informatics disciplines with health and life sciences.

The basic explanation of healthcare informatics is the area of the industry that deals with giving the appropriate and meaningful information about a patient’s condition to the appropriate person, a health specialist, at the appropriate time, assisting them in declaring as well as taking intelligent decision regarding their patient's treatment (Wullianallur et al., 2014) (Tae-Min et al., 2013). Patients, healthcare practitioners, healthcare managers, Internet technologies companies, and administration of medical facilities collaborate and share information in the field of healthcare informatics.

Health informatics is described as the knowledge, skills as well as tools that allow information to be gathered, organized, utilized and distributed to enhance the healthcare service and upgrade health. The development of healthcare informatics resulted from the switch from a conventional healthcare delivery system to a system of care delivery for the digital age. The word "healthcare informatics" is a combination of several terms that have a similar purpose, including telehealth, medical informatics, as well as information management and technology.

Pre-clinical, clinical, and post-clinical medical fields, health manage and administration and information technology are all combined in healthcare informatics. Experts in health informatics are familiar with computer technology, how to use it to address problems in the real world, and how to manage complex processes for integrating new ideas into companies (IBM Institute for Business Value 2021). In addition to these raw skills, an expert in health informatics is familiar with the field from the perspectives of providing medical services as well as to manage, challenge, and demand medical organizations and procedures.

The crucial instruments for connecting data and information in the course of decision-making are made available by health informatics. Innovation and information management facilitate the gathering, analysis, and transmission of information from a variety of sources. The goal is to support service and business plans, improve organizational performance, aid in health care, and address business challenges. The word "eHealth," which is relatively new and is frequently used during European nations, refers to the use of electronic and digital innovations in healthcare practices, procedures, and communications.

Individuals who have health issues now have the ability to obtain healthcare in the convenience of their own homes thanks to telehealth, a combination of technology and healthcare services (Dale et al., 2019). An illustration of telehealth is a house Pod device placed in a patient house, which may be applied to track the patient's health state and transmit that data securely to a distant physician.

4.1. Need of healthcare informatics

Being one of the most important industries in modern civilization, healthcare is expanding quickly. The sector makes up a considerable portion of public spending, which is driven by ongoing calls for innovation. Information management, which is crucial for health practitioners, has significantly increased as a result of the explosion of medical and biological information. Experts in handling such data on behalf of healthcare professionals are known as health informatics.

The foundation of the healthcare sector is data. For the purpose of offering medical services, knowledge regarding clinical status must be current, reliable, and accurate. Information is also necessary for the efficient management of the numerous groups and people engaged in each healthcare treatment. Information is also required to manage the intricate business transactions involved in the sector (IHIT, 2021). Because of the complexity, intensity, and volume of health data, incorporating information technology for the advancement of efficiency and effectiveness with contextual medical facilities has presented significant challenges. As a result, highly qualified healthcare informatics professional drive and are depending on innovations in the healthcare sector.

Currently, the focus in healthcare is on reducing the amount of waste generated by implementing new mechanisms of health information technique (healthcare informatics). By doing so, inefficiency, duplication, as well as organizational costs are reduced, which encourages punctuality, transparency, and better outcomes in the delivery of healthcare.

4.2. Role of healthcare informatics into healthcare system

A field of study at the intersection of information systems and healthcare is known as healthcare informatics. It includes the tools, equipment, and techniques required to improve the gathering, archiving, retrieval, and application of information in healthcare and medicine. Figure 3 shows interconnection of various parameters for health outcomes (Raghupath, 2020). Computers, information and communication technologies, extensive medical terminologies, and diagnostic care are some of the technologies utilized in health informatics. A more accurate method of information storage and retrieval than asking a patient to remember their current drugs and reactions is made possible by health informatics (Dembosky, 2022). For such patient, it is the most crucial safety concern. An unfavorable pharmacological effect may result from incorrect or insufficient information provided by a perplexed patient. Consequently, it is crucial to provide precise information when it is required.

Health informatics enables combined care, in which several divisions, like the laboratories, surgery, radiography, accounting, or administration division, are linked to each other, thereby minimizing unnecessary duplication and accelerating processes. With the expanding requirements in healthcare, such moment feature is very important. It functions as a tool for making important decisions as well. The benefits of computerized standards include assisting patients and professionals in making wiser decisions. This makes it possible to maintain high-quality prescriptions and treatments (André, 2018). Furthermore, by spending less time on administrative duties, medical professionals can spend more time with patients.

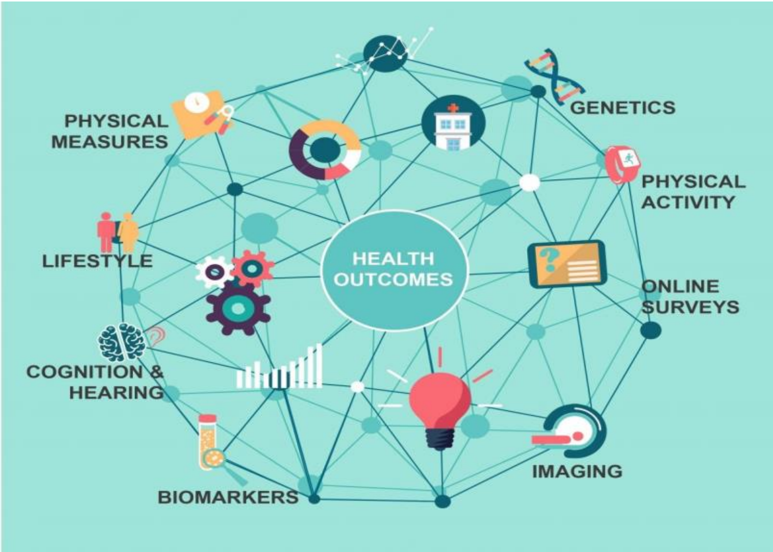

Figure 3. Health outcomes

The necessity to repeatedly record the same information is eliminated by health informatics. It reduces waste. Additionally, advances in health informatics guarantee effective, high-quality healthcare service. Consequently, it is an area that fuses information science, information systems, technologies, and healthcare. Resource planning, medical services standardization, research collaboration, and effective healthcare distribution all benefit from the knowledge of health informatics.

5. Healthcare analytics

Analytics is the methodical application of data and associated competitive intelligence created via applied analytical disciplines like statistical, experiential, numerical, accurate, intellectual, and other such as innovative approaches to support fact-based decision-making for organizing, management, evaluation, and understanding, as per IBM Institute for Business Value. Analytics might be predictive, informative, or both. Inside the three stages of data management i.e. data collection, data sharing, and data analytics healthcare has continued to improve. Phases of data gathering and sharing are what sparked the rapid adoption of electronic health records and healthcare data transfers. Analytics has a significant possibility of transforming healthcare into a fully data-driven industry.

Health records are predicted to rise dramatically and erratically in the next years. Similar changes are occurring in payment and delivery models, with pay-as-you-use services becoming increasingly important in the modern healthcare market. Healthcare organizations shouldn't be driven exclusively by revenue, but it's imperative that they get the facilities, procedures, and resources they need to effectively manage Big Data. Instead, they run the chance of losing a significant amount of money in sales and earnings. The health system as a whole has recently generated a sizable datasets as a result of bookkeeping, regulatory and compliance requirements, and patient care (Adams et al., 2020). The custom is to keep documents on paper, however the current trend is to quickly digitize each of these enormous datasets. This tendency is driven by the obligation and opportunity to raise the bar for healthcare provision while also guaranteeing cost containment. These enormous amounts of data, often familiar as "Big Data," ensure the assistance of numerous healthcare and medical activities, including health care delivery, illness monitoring, and clinical decision-making support.

5.1. Need of healthcare analytics

Analytics in health emerged as an output of the massive amount of health information being gathered. Digital patient bookkeeping, significant national programmes, as well as databases that connect patient facts across hospitals, areas, and other entities are some examples of this data. The ability to gather different data of patients, allow patients to access one‘s own information and extra health information, and eventually the ability to task across all these numerous information to track patient and population care outcomes, optimize and streamline the health system, and fully assess information systems are all requirements for healthcare analytics. The frightening issues that the healthcare sector is currently facing call for strong recommendations to expand the analytics jobs. Ineffective collection of information, sharing, and utilization are to blame for the lack of productivity.

In order to provide the best results, including novel treatments and technology, analytics should be used to improve understanding. Understanding via information may also help knowledgeable, educated, and informed customers take more responsibility for their own health (IBM 2021). As a result, analytics will increase productivity, facilitate finding and explorations, enhance service delivery as well as logistics, and provide a way to track and assess important corporate data. Analytics can also improve affordable healthcare, pay equity for performance, and sustainability in medical costs.

5.2. Role of healthcare analytics into healthcare system

Within the healthcare sector, there exists a great volume of data but little expertise and information (Anil, 2019) (Carlton, 2020). The use of analytics methodologies and tools in healthcare systems has a variety of advantages. These consist of:

- Method to healthcare management that takes wellbeing into account;
- Specialized medical analytics, collaboration, and health management tools;
- Enhanced real-time assistance and treatment judgments, user dedication;
- Searching for unidentified elements that affect quality, like hidden re-admission factors;
- Claims data analysis for frauds, scientific proof medicine that harmonizes clinical procedures, protocols, and recommendations, and so forth;
- Paying for better reimbursement changes operations, analytics, and reliability;
- Improved revenue planning through knowledge of regional markets.

6. Future scope

The development of genetic technology has the potential to speed up the transition to personalized and predicting medicine and better illness detection and treatment, particularly for complex diseases like cancer and genetic syndromes. The enormous amount of genetic data, computationally difficulties brought on by high throughput, and declines in next-generation sequence analysis have caused a shift away from such an exploration and toward clinical application with integral part of medical data. The interchange of innovation, data, and knowledge is the main goal of combining biology and health information. As a result, a database of illness, underlying reasons, trigger patterns, and potential treatments and cures will be made available. It will expedite access to healthcare and create a trustworthy and effective health service. The development of analytics came to the rescue in the healthcare sector when it came to the vast volume of data gathered electronically over disparate platforms. Analytics can therefore be used to produce insights that assist medical diagnoses and minimize costs, eliminate inefficiencies, improve outcomes, detect at-risk communities, forecast potential healthcare requirements for individuals, and empower patients.

7. Conclusion

When genetic technologies are being combined with health informatics data, it will significantly improve the present health system, and analytics will make it happen to spend resources extra effectively by making confirm that individuals who need to care the most get it. In order to maximize its potential for innovations, creativity, and improved patient care, academics must focus more closely on the combination of the three hot domains (bioinformatics, healthcare informatics, as well as analytics).

REFERENCES

1. Adams, J. R., Bakalar, M. D., Michael B., Karen K., Edgar L. M. and Neil S. (2020). *Healthcare and care delivery: Delivery models refined, competencies defined.* IBM Institute for Business Value.
2. Badea, L., Țilivea, D. (2006). Clustering and Meta-clustering Gene Expression Data with Positive Matrix. *Revista Română de Informatică şi Automatică (Romanian Journal of Information Technology and Automatic Control),* 16(4), p. 19-28.
3. André, W. K. (2018). *Forecasting in Health Care: Integrating Analytics with Electronic Health Records*, SAS Institute Inc.
4. Anil, J. (2019). Unlocking the power of big data to improve healthcare for everyone. *Explorys*.
5. Carlton, M. (2020). *How big data analytics reduced medical re-admissions*. IBM Corporation.
6. Dale, S., David, A. S. & Denis, P. (2019). *The Healthcare Analytics Adoption Model: A Framework and Roadmap*. White paper by Health Catalyst.
7. Dembosky, A. (2022). Data prescription for better healthcare. *Financial Times*, p. 19.
8. Flygare, S., Simmon, K., Miller, C., Qiao, Y., Kennedy, B., Di Sera, T., et al. (2016). Taxonomer: an interactive metagenomics analysis portal for universal pathogen detection and host mRNA expression profiling. *Genome Biol.*, 17:111.
9. IBM (2021). *IBM big data platform for healthcare*. Solution Brief.
10. IBM Institute for Business Value (2021). *The Value of Analytics in healthcare*. IBM Global Business Services, USA.
11. IHIT, (2021). *Transforming healthcare through Big data strategies for leveraging health care industry*. The Institute for Health Technology Transformation.
12. Lu, Y.Y., Tang, K., Ren, J., Fuhrman, J.A., Waterman, M.S., Sun, F. (2017). *CAFE: accelerated Alignment-Free sequence analysis*. Nucleic Acids Res., 45:7.
13. Navarro-Gomez, D., Leipzig, J., Shen, L., Lott, M., Stassen, A.P.M., Wallace, D.C., et al. (2015). Phy-Mer: a novel alignment-free and reference-independent mitochondrial haplogroup classifier. *Bioinformatics*, 31:1310–1312.
14. Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., et al. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.*, 17:132.

15. Raghupath, W. (2020). Data Mining in Health Care. In *Healthcare Informatics: Improving Efficiency and Productivity*, 211-223.
16. Tae-Min, K., Myung, S. K., Seong, K. M. & Yeun-Jun, C. (2013). Perspectives on Clinical Informatics: Integrating Large-Scale Clinical, Genomic, and Health Information for Clinical Care. *Genomics and Informatics*. Published online by Korea Genome Organization.
17. Thankachan, S. V., Chockalingam, S. P., Liu, Y., Krishnan, A., Aluru, S. (2017). A greedy alignment-free distance estimator for phylogenetic inference. *BMC Bioinformatics*. (2017); 18:238.
18. Wu, Q., Yu, Z.-G., Yang, J., (2017). DLTree: efficient and accurate phylogeny reconstruction using the dynamical language method. *Bioinformatics.*
19. Wullianallur, R. & Viju, R. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, 2:3.

Kshatrapal SINGH is B. Tech. (CSE), M. Tech (CSE) and Pursuing Ph.D. from Dr. A.P.J. Abdul Kalam Technical University, Lucknow (India). Currently, he is working as an Assistant Professor at Krishna Engineering College, Ghaziabad. His specialization is Bio-Computing, Bio-informatics, Compiler Design, Discrete Structure and Graph Theory.

Ashish KUMAR is B. Tech. (CSE), M. Tech (IIT-K) & Ph.D. from UPES, Dehradun (UK). Currently, he is working as Professor and Head, Department of Computer Science & Engineering at I.T.S. Engineering College, Greater Noida. His specialization is Networking, Web Technology, Java and Software Engineering.

Manoj Kumar GUPTA is B. Tech. (CSE), M. Tech (CSE) & Ph.D. from IIT, Roorkee. Currently, he is working as Associate Professor, Department of Computer Science at Shri Mata Vaishno Devi University, Katra (J&K). His specialization is Sequence Comparison, Networking, Object Oriented Technologies, Python etc.