Integration of AI, ML and Blockchain Technology: Transforming Codes into Life

Hari Krishna Garg

Department of Biotechnology,
Institute for Excellence in Higher Education, Bhopal, India

Abstract

Data-driven approaches using AI, ML, and blockchain have revolutionized biotechnology and genomics. The present paper delves into the exciting possibilities that arise from the intersection of AI, ML, and blockchain technology with genomics, healthcare, and biotechnology. The use of these advanced technologies allows for the creation of radical applications that leverage machine learning, Big Data analytics, natural language processing, decision support, and reasoning under uncertainty. Such applications provide unprecedented avenues for improving human health and well-being.

By assimilating AI with biotechnology, researchers can develop cutting-edge applications that enable genomic sharing, next-generation sequencing, gene editing, clinical workflow optimization, risk prediction, diagnosis, and precision medicine. The potential applications of AI, ML, and blockchain in these areas are truly transformative, and have the power to revolutionize the future of healthcare. The survey showcases the significant impact of these technologies in improving patient outcomes, reducing costs, and increasing the efficiency of healthcare delivery. With the help of AI, ML, and blockchain, one can realize a future where healthcare is more personalized, effective, and accessible to everyone.

Keywords: Artificial Intelligence, Biotechnology, Blockchain, Deep Learning, Digital Transformation, Machine Learning.

Introduction

Big data has become an essential aspect of modern society, with its significance spread across various industries and fields. In 2001, Gartner introduced the 3Vs of data: Volume, Velocity, and Variety. Since then, the field of data analytics has expanded on this concept by adding two more Vs - Value and Veracity. Here, volume refers to the massive amount of data, which is often complex and heterogeneous. Traditional database technology cannot handle this volume of data, leading to the need for advanced analytics to extract insights. Velocity refers to the speed at which new data is generated and moves around. Variety refers to the different types of structured, semi-structured, and unstructured data available, such as social media conversations and voice recordings. Veracity refers to the certainty, accuracy, relevance, and predictive value of the data, while value refers to the conversion of data into business insights.

The genomics and healthcare industry is one of the sectors that have been impacted significantly by big data and artificial intelligence (AI). Big data analytics and AI have become omnipresent across the entire healthcare spectrum, including payers, providers, policy-makers/ government, patients, and product manufacturers. Healthcare fraud and abuse account for up to 10% of global healthcare expenditure, and AI-based tools can help mitigate this problem in payer programs [Joudaki, et al., 2015]. Medical coding errors and incorrect claims also account for substantial losses, and the reliable identification of these errors can save payers, providers, and governments significant amounts of money and time [Davenport and Kalakota, 2019].

AI is also used for evidence-based clinical decision support, detection of adverse events, and predicting patients at risk for readmission. Healthcare policymakers and government can use AI-based tools to control and predict infections and outbreaks.

With the advent of the global pandemic coronavirus disease 2019 (COVID- 19) in early 2020, AI models could be used to predict at-risk populations and provide additional risk information to clinicians caring for at-risk patients [Vaishya et al., 2020]. The big data analytics for patients and biotechnology/ healthcare products is a crucial aspect of healthcare and the future of healthcare will depend significantly on these technologies - AI, ML and Blockchain to provide efficient and effective care to patients.

Artificial Intelligence Vs Machine Learning

Artificial Intelligence (AI) and Machine Learning (ML) are two terms that are often used interchangeably, but they represent different concepts. AI refers to the broad vision of generating computers and software that can perform tasks that require human intelligence. On the other hand, ML is a subfield of AI that involves training computers to perform tasks without explicit instructions using patterns and insights from data. Deep Learning (DL) is a subset of ML that uses artificial neural networks with many layers to learn and make decisions. It is particularly useful for tasks that involve analyzing large amounts of data [Garg, 2023a].

The AI field was initiated in 1956 when a group of computer scientists met at Dartmouth College in Hanover, New Hampshire. The group had ambitious goals to create machines that could simulate every aspect of human intelligence. Since then, AI has gone through many ups and downs, including an AI winter in the 1980s. However, the rise of statistical data-driven ML helped to revive the field of AI. Today, AI is experiencing a resurgence, and the latest natural language technology developed by OpenAI, called ChatGPT, is proof of what AI can do.

There are three major classes of ML: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning aims to predict a classification or label of data points using a given set of labeled training examples. In contrast, unsupervised learning aims to learn inherent patterns within the data. Reinforcement learning is based on rewarding desired behavior and punishing undesired behavior of software agents.

DL models are more flexible than standard ML methods and can model more complex relationships between inputs and outputs [LeCun et.al, 2015; Zou et.al, 2019]. Different types of neural networks have been developed for specific tasks such as convolutional neural networks, which capture spatial dependencies, and recurrent neuronal networks, which handle sequential or time-series data.

Thus, AI aims to provide the theoretical fundamentals for ML to develop software that can learn autonomously from previous experience. To reach a level of usable intelligence, we need to learn from prior data, extract knowledge, generalize, fight the curse of dimensionality, and disentangle the underlying explanatory factors of the data. The grand goal is to create software that can learn automatically without human intervention.

AI in Genomics and Healthcare

Artificial Intelligence (AI) has revolutionized many domains, and healthcare is no exception. It involves the use of technology to create software that mimics human-like critical thinking [Ramesh et al., 2004]. AI uses techniques such as fuzzy expert systems and artificial neural networks [Hessler and Baringhaus, 2018; Mintz and Brodie, 2019] to provide personalized experiences where predictions are backed by mathematical data points. The field of AI in healthcare can be divided into two subunits, virtual and physical. The virtual aspect of AI involves electronic healthcare records [Esteva et al., 2019] and neural networks guiding patient treatments [McDonnell et al., 2021], while the physical aspect involves robots assisting in surgeries and AI-generated prosthetics for the disabled.

Over the past decade, AI has seen remarkable growth and acceptance in genomics and biotechnology. It provides rich opportunities for designing intelligent products, creating novel services, and generating new business models. The use of AI in medicine can introduce social and ethical challenges to security, privacy, and human rights.

AI technologies in medicine exist in many forms, from the purely virtual to cyber-physical. AI technologies have enabled many image-based detection and diagnostic systems in healthcare to perform as well or better than clinicians. AI-enabled clinical decision-support systems may reduce diagnostic errors, augment intelligence to support decision making, and assist clinicians with EHR data extraction and documentation tasks.

Emerging computational improvements in natural language processing, pattern identification, efficient search, prediction, and bias-free reasoning will lead to further capabilities in AI that address currently intractable problems [Biamonte et al., 2017]. However, the advances in the computational capability of AI have prompted concerns that AI technologies will eventually replace physicians.

Therefore, the term augmented intelligence [Ashby, 1957] may be a more apt description of the future interplay among data, computation, and healthcare providers, and perhaps a better definition for the abbreviation AI in healthcare.

Insights into the Blueprint of Life

The human genome is the foundation for the expression of human traits, consisting of unique biological DNA that makes each individual distinct. The advent of genomics has revolutionized the field of molecular biology, enabling scientists to map the structure and function of genomes. Each human genome contains 20,000 to 25,000 genes, with every gene comprising a few hundred to 2 million DNA bases [International Human Genome Sequencing Consortium, 2001]. The mapping of the human genome in 2003 opened up numerous possibilities for using genomics in the medical field.

Gene expression occurs through transcription and translation, with RNA splicing in between, resulting in diversity in protein coding. However, errors in splicing or mutations can cause a range of diseases [Fletcher et al., 2013]. While protein-coding DNA accounts for only a small percentage of the genome, a significant portion is transcribed into non-protein-coding RNAs (ncRNAs) that regulate gene expression and transcription initiation and termination [Mattick and Makunin, 2006]. Genomic sequencing has revolutionized the way researchers read the genetic blueprint, but affordability and data management are two significant challenges that must be addressed.

Challenges with Gene Sequencing

The ability to sequence DNA has provided researchers with unprecedented opportunities to understand human biology and develop new therapies for diseases. However, the cost of using genome sequencing in routine clinical care remains a significant challenge. At present, the cost of sequencing a single genome in a single laboratory is around $1000 [Schwarze et al., 2020]. This cost can be prohibitively expensive for many people, limiting access to potentially life-saving genetic information. To make genomic sequencing more affordable, researchers are developing new technologies that could reduce the cost of sequencing and improve the accuracy of results.

1. Data Management and Privacy

Another significant challenge in genomics is data management. The collection, sharing, ownership, and storage of genomic data are all complex and time-consuming processes that require special attention to detail, precision, and privacy. Genomic data contains highly personal information about an individual's past, present, and future generations. Therefore, researchers must take special care to ensure that this information is recorded and managed securely to prevent potential misuse or breaches of privacy.

2. Potential Misuses of Genomic Data

The potential misuse of genomic data is a significant concern in the field of genomics. This information could be used to develop harmful medicines or even commit crimes, highlighting the importance of managing and securing genomic data carefully. Researchers must be mindful of the potential consequences of any data breaches or misuses, and must take steps to minimize the risk of these occurrences.

Transforming Genomics through ML

In the area of genome sequencing, machine learning can be used to identify patterns within high volume genetic data sets. These patterns are then used to create computer models that can help predict an individual's probability of developing certain diseases or inform the design of potential therapies. This is particularly useful in the field of precision medicine, where treatments are tailored to an individual's unique genetic makeup. By analyzing large data sets, machine learning algorithms can identify subtle differences in genetic patterns that may be associated with increased disease risk or specific treatment responses.

Advancements in genomics continue to offer insights into human health and disease. For instance, researchers have employed genomics to identify genetic variations that contribute to various diseases, such as cancer, diabetes, and Alzheimer's disease. By comparing the genomes of healthy individuals to those with specific diseases, researchers can identify genetic differences and develop targeted treatments.

Even companies offering genomic sequencing services to individual consumers are using machine learning algorithms to gain a greater understanding of how an individual's genes may impact their health. By analyzing genetic data, companies can predict an individual's likelihood of developing certain conditions, such as weight gain, and provide personalized advice on diet and exercise to help individuals manage their weight.

Machine learning is also being used to predict pharmaceutical properties of drug targets and drug candidates. By analyzing large data sets on the molecular properties of potential drugs, machine learning algorithms can predict their effectiveness in treating specific diseases. This has the potential to greatly accelerate the drug discovery process, ultimately leading to more effective treatments.

Another important area of application is in the analysis of multimodal data from genomics and other omics fields, combined with clinical data [National Research Council, 2011]. By integrating large data sets from multiple sources, machine learning algorithms can generate new diagnostic and predictive models for diseases, including their underlying genetic causes. This has the potential to greatly improve disease diagnosis and treatment, leading to better patient outcomes.

Pharmacogenomics is another promising application of genomics, which helps doctors assign medication and corresponding dosage based on the patient’s genetic markers. This technique has enabled specialists to provide more personalized care and improve patient outcomes. CRISPR is another revolutionary technology that has made it possible to treat chronic diseases like HIV [Xiao et al., 2019], β-thalassemia [Frangoul et al., 2021], cancers [Chen et al., 2019], leukemia [Tzelepis et al., 2016], and sickle cell anemia [Frangoul et al., 2021].

Despite ongoing debates on the ethics of genetic testing without a clear cure, the availability of genetic information through next-generation sequencing and direct-to-consumer testing makes personalized prevention and management of serious diseases a reality.

1. Next Generation Sequencing

Next Generation Sequencing (NGS) technology has revolutionized genome sequencing and emerged as the leading method. Compared to classic Sanger sequencing that took over a decade to complete the human genome, NGS allows researchers to sequence a whole human genome in just one day. Illumina sequencing is currently the most popular technology due to its cost, speed, and accuracy [Liu et al., 2012]. However, long-read sequencing technologies like those created by Oxford Nanopore [Green and Sambrook, 2018] and Pacific Biosciences [Rhoads and Au, 2015] generate longer reads that are thousands of base pairs long, but lower in quality than short-read sequencing.

NGS data has the potential to supplement other genomic sequencing methods and improve the effectiveness of precision medicine by better identifying disease risk and actionable genetic mutations in cancer patients. This technology can aid in the development of drugs targeting tumors and matching patients to therapy methods. Companies like Deep Genomics are using machine learning algorithms to interpret genetic variation by identifying patterns in large genetic datasets and translating them into computer models.

DNA sequencing data is stored in the FASTQ format, which consists of four corresponding lines of text for each sequence. FASTA is another commonly used text-based file format for storing reference genomes. Algorithms map sequencing reads to reference genomes, and these results are stored in Sequence Alignment Map (SAM) or its binary equivalent (BAM) file formats [Hoogstrate et al., 2021]. While SAM files are readable by humans, BAM files are used to compress the data due to the large file sizes. Finally, variant call format (VCF) files describe the sequence variations, insertions, and deletions found in samples along with rich annotations [Zhang, 2016; Morash et al., 2018].

Although the efficacy of NGS data in precision medicine remains controversial due to experimental design, the technology's potential for development is immense. The improved methods for analyzing sequenced data can help in the development of precision medicine. NGS technology, combined with machine learning, can also help identify and interpret genetic variation and its effects on crucial cellular processes.

2. Variant Discovery

Variant discovery is a critical step in understanding the genetic basis of various diseases. Whole genome sequencing (WGS) is a technique that involves sequencing an individual's entire genome, including both protein-coding and non-protein-coding regions, while the entire exon sequencing (WES) focuses solely on the protein-coding regions [Petersen et al., 2017]. By using variant calling, researchers can identify various types of variants, providing valuable insights into disease diagnosis and prevention.

There are three main types of pipelines used for WGS and WES: cloud-computing, centralized, and standalone [Ahmed et al., 2021]. Cloud-computing pipelines are utilized in environments with on-demand compute resources provided by external vendors. On the other hand, centralized pipelines are used in local computers, while standalone pipelines are mainly used in high-performance computing environments. These pipelines have been designed to effectively collect and process data from WGS or WES, allowing researchers or medical professionals to recognize the links between genetic variants and diseases.

3. Gene Editing

Gene editing involves making targeted changes to DNA at the cellular or organism level. CRISPR is a gene editing technology that has made this process faster and less expensive. However, selecting the appropriate target sequence for CRISPR can be a challenging task. Luckily, the use of machine learning has the potential to significantly reduce the time, cost, and effort required to identify the right target sequence. Continued research and development in this area could revolutionize the field of gene editing.

At the intersection of AI and CRISPR, London-based software company Desktop Genetics has emerged. The company works with experimental or reference data uploaded to Google Cloud, which is then processed and formatted before being sent to their bioinformatics and machine learning teams. By analyzing this data, they can design and conduct CRISPR experiments, develop new models, and generate FASTQ data that feeds back into the workflow.

Recently, the company published two significant findings from their research. Firstly, they found that an increased amount of training data improves the accuracy of the algorithm's ability to predict CRISPR activity. Secondly, they discovered that the model's accuracy decreases when applied to a different species, such as humans versus mice. Although these findings may not be surprising, they highlight the importance of ongoing research to continue improving processes and push the boundaries of how machine learning can impact CRISPR.

4. Clinical Workflow

In today's world, where technology has penetrated every aspect of our lives, it is no surprise that the genomics and healthcare industries are also reaping the benefits of technological advancements. With the help of artificial intelligence (AI) and machine learning, the healthcare sector is trying to revolutionize the way it functions.

One of the challenges that the healthcare sector faces is the availability of patient data to the various members of the healthcare team serving a patient. However, this challenge has sparked an interest in using machine learning to improve the efficiency of the clinical workflow process.

Intel, a major tech company, has created an Analytics Toolkit that integrates machine learning capabilities to evaluate factors like a patient's risk of developing multiple cancers. The algorithm utilized in the toolkit was created with four primary components, including a centralized genomic data database linked to clinical and patient data, electronic health record (EHR) access for all clinicians and genetic counselors, integration of all data from genetic tests into EHRs, and access to operational Clinical Decision Support tools (CDS). Examples of clinical decision support include family health histories, screenings, and past clinical data.

It has been reported that a sample workflow for a patient can be screened in just 3 to 5 minutes with the workflow model developed using machine learning. This has contributed to improved data accessibility. Despite the regulatory issues and complex sales cycles, many of the major players in artificial intelligence are recognizing the significant economic value of AI in healthcare.

5. Direct-to-Consumer Genomics

The market for predictive genetic testing and consumer genomics is set to expand dramatically, and is expected to touch $5 billion by 2025. This growth is fueled by the increasing awareness of how genomic testing can aid in identifying one's risk of developing certain illnesses. Proper guidance can make these tests a valuable tool in preventative healthcare, despite concerns regarding regulation and the need for health professionals to interpret results for patients.

Direct-to-consumer genomics is a rapidly expanding industry, especially as people become more conscious of their lifestyle and dietary habits. Personalized analyses of an individual's genetic makeup, taking into account factors like genotype, sex, age, and self-identified primary ancestry, can help determine how one's genetic material may impact their weight. However, there are still concerns about the regulation of these tests and the necessity of professional interpretation of results.

6. Clinical Genomics

Clinical genomics is a rapidly developing field that leverages sequencing techniques to identify genes associated with diseases. The approach can detect abnormalities in patients, predict their susceptibility to certain diseases, and facilitate the development of treatments for rare diseases. However, the usefulness of genomics data depends on how it is organized and assimilated.

One essential tool for organizing and assimilating genomic and phenotypic data is gene-disease databases. Despite the existence of approximately 18,000 gene-disease databases [Huang et al., 2018] only a few are approved by the American College of Medical Genetic and Genomics (ACMG). One significant challenge with these databases is their lack of standardization, which may lead to outdated or irrelevant information about diseases.

To address this challenge, researchers have developed IOS applications such as PAS-Gen and PROMIS-APP-SUITE. These applications provide a centralized database for genomic and disease information [Stenson et al., 2017], making it more accessible and practical for researchers and healthcare professionals. By providing standardized and up-to-date information, these apps can accelerate medical discoveries and aid in the development of treatments for various genetic diseases.

7. Precision Medicines

The integration of machine learning into genomics has brought about significant advancements in precision medicine. Machine learning algorithms have revolutionized the analysis of vast amounts of genomic data, enabling the identification of genetic mutations and patterns that are linked to various diseases and disorders. Such insights are used to develop patient-specific treatment plans, thereby improving outcomes and lowering healthcare expenses.

Precision medicine is an approach to patient care that takes into account an individual's unique genetics, behaviors, and environment. Its goal is to create tailored treatment interventions instead of a one-size-fits-all approach. For example, matching a patient in need of a blood transfusion to a donor with the same blood type can significantly reduce the risk of complications.

Despite the potential benefits of precision medicine, a significant obstacle to its widespread implementation is the high cost of collecting and analyzing patient data. Machine learning techniques are useful in reducing these costs by swiftly and effectively analyzing vast amounts of data. Furthermore, as the cost of genome sequencing continues to decline, genomics is becoming more accessible and affordable.

By leveraging machine learning techniques, genomics firms and researchers can hasten the pace of discovery and create more personalized treatment plans for patients. As the field of genomics progresses, we can anticipate exciting advancements in precision medicine and other aspects of healthcare. Overall, the integration of machine learning into genomics has the potential to significantly enhance patient outcomes and lower healthcare costs.

8. Diagnostics

The use of artificial intelligence (AI) in medical biotechnology has great potential to revolutionize the field. However, implementing AI algorithms in in vitro diagnostics (IVD) companies presents significant challenges, particularly related to ethical and legal issues. Despite these obstacles, AI can be utilized in several ways to improve medical biotechnology.

Drug target identification: One way in which AI can be utilized in medical biotechnology is drug target identification. By analyzing genomic data and protein-protein interaction data, AI can identify potential therapeutic targets for the treatment of diseases. Machine learning algorithms can identify patterns and correlations that may not be apparent to humans.

Drug screening: Another application of AI in medical biotechnology is drug screening. AI can analyze data on the activity of potential drugs against different targets and identify those most likely to be effective. Machine learning algorithms can predict the likelihood of a particular drug being effective based on its characteristics and the characteristics of the target.

Image screening: AI can also be utilized in medical image screening. By analyzing CT scans and MRI images, AI can identify abnormalities and diagnose diseases. Deep learning algorithms can automatically segment and classify structures in medical images.

Predictive Modeling: AI can be used for predictive modeling. By analyzing data from electronic health records and wearable devices, machine learning algorithms can make predictions about an individual's health. This includes predicting the likelihood of an individual developing a particular disease or the likelihood of a particular treatment being effective.

9. Cardiovascular Disease

The field of cardiovascular medicine has a rich history of employing predictive modeling to evaluate patient risk. Recent advancements have enabled the prediction of heart failure and other cardiac events in asymptomatic individuals. When utilized in conjunction with personalized prevention strategies, these predictive models hold the potential to have a positive impact on disease incidence and its effects. The intricacy of diseases such as cardiovascular disease necessitates the integration of various factors, including gender, genetics, lifestyle, and environmental factors. As a result, it is critical to consider the heterogeneity of the data, and artificial intelligence (AI) approaches have exhibited promise in identifying intricate connections among a vast number of factors. A Vanderbilt study showcased the early successes of merging electronic health record (EHR) and genetic data, yielding favorable outcomes in cardiovascular disease prediction [Zhao, et al., 2019]. AI-powered recognition of phenotype features via EHR or images and correlating those features with genetic variants may enable more rapid genetic disease diagnosis [Gurovich et al., 2019].

Future Prospects

The field of genomics is rapidly advancing, and machine learning is expected to have a significant impact in several areas. One of these areas is the development of patient-specific pharmaceutical drugs. Machine learning models are being used to determine stable doses of drugs including those commonly administered to patients following solid organ transplants to prevent acute rejection of the new organ. Pharmacogenomics is an emerging field that uses genetics to understand how individuals respond to drugs, and machine learning is expected to play a crucial role in this field.

Another area where machine learning is expected to have a significant impact is in newborn genetic screening. As this practice becomes more widespread, data collected at birth will be integrated into individuals' electronic health records. Non-invasive screening capabilities for diseases such as Down Syndrome may be available to women during pregnancy.

Roadblocks

Managing, analyzing, and storing the large amounts of data generated by healthcare and genomics industries is a daunting task. Current data management systems face various challenges, such as data sharing, analysis cost, data ownership, privacy, and security. Researchers have developed different solutions to tackle these problems, including Data Cloud Architecture, Data Commons, and Data Ecosystem. However, these solutions still have scalability and flexibility issues.

Recently, blockchain technology has emerged as a promising solution to address these challenges. With its decentralized, distributed, and immutable nature, blockchain can provide secure and transparent data management solutions. Moreover, blockchain can reduce the analysis cost of genomics and healthcare applications by enabling faster and more efficient transactions compared to traditional processes.

Blockchain technology also offers pseudo-anonymity to ensure personal data security and privacy. Individuals can modify their data access permissions and use encryption methods, such as symmetric encryption, to secure their data further [Garg, 2023b]. Thus, blockchain technology has enormous potential to be a valuable cornerstone in building a Data Ecosystem for healthcare applications.

Blockchain as a Way-out

Blockchain technology can revolutionize the healthcare industry by enabling secure, transparent, and efficient sharing of electronic health records (EHRs) and genetic test results. The use of blockchain technology has led to the development of several platforms such as Coral Health, Patientory, Medicalchain, and GemOS, all built on Ethereum and Hyperledger protocols.

Coral Health is a data sharing platform that creates a secure and accessible healthcare ecosystem through a precision medicine program. The system uses SMART and FHIR protocols to connect mobile devices of patients and other environments hosting their medical data [Coral Health, 2023]. EncrypGen and Gene-Chain are other platforms that use blockchain technology to de-identify genomic data and enable safe, traceable, and unhackable transactions of genomic data.

Health Nexus is an open-source blockchain protocol that offers a more efficient, trustworthy, and secure path for data to travel in the healthcare community. Medicalchain is an EHR sharing project that employs a dual blockchain structure with Hyperledger controlling access to health records and Ethereum underlies all the applications and services. MedRec and Opal are other encrypted platforms that use blockchain technology to manage authentication, confidentiality, accountability, and data sharing of sensitive healthcare information.

Nebula Genomics is a platform that leverages blockchain technology to empower individuals to own their personal genomic data, lower sequencing costs, and enhance data privacy. The platform uses an open protocol to enable data buyers to efficiently aggregate standardized data from many individuals and genomic databanks [Nebula Genomics, 2023]. Zenome is another platform that focuses on genetic data sharing and has an Ethereum-based ZNA token [Kulemin, Popov & Gorbachev, 2017].

These platforms offer a scalable and flexible approach to patient-centric healthcare and personalized medicine and present an exciting opportunity for innovation in data transfer for the healthcare community.

Discussion

In recent years, Artificial Intelligence (AI) has played a significant role in the biotechnology industry, particularly in fields such as drug discovery, drug safety, proteomics, pharmacology, and pharmacogenetics. These fields require the storage, filtering, analysis, and sharing of large amounts of data, and AI software solutions have provided support to increase speed and reduce manual errors [McAlister et al., 2017; Ginsburg and Phillips, 2018]. The adoption of new technologies and processes to improve efficiency, accuracy, and speed through digital transformation can further accelerate the development and use of AI in biotechnology.

In healthcare, the successful adoption of AI is dependent on three key principles: data and security, analytics and insights, and shared expertise. Shared expertise refers to the complementary relationship between AI systems and human professionals. Moreover, precision medicine, which aims to personalize care for every individual, is providing an equal or even greater influence than AI on the direction of healthcare. Precision medicine requires access to massive amounts of data, and the convergence of AI and precision medicine can accelerate the goals of personalized care and tightly couple AI to healthcare providers for the foreseeable future.

The future of biotechnology and healthcare is dependent on several key areas, including genomics, AI, big data, and blockchain. Precision medicine is one of the target areas in this field, with numerous advantages such as more accurate diagnoses, easy access to medical data, and a better understanding of diseases and their causes [Hasin et al., 2017]. However, implementing precision medicine can be challenging due to the lack of a system to compare multi-omics patient data and identify appropriate approaches to use with different types of medical data [Picard et al., 2021]. Ethical and logistical issues also need to be considered in clinical genomics, big data, and pharmacogenomics implementation.

To address these challenges, multiple approaches need to be integrated into precision medicine. By combining information from different fields, researchers can gain a more comprehensive understanding of a medical case and select an appropriate treatment method. For example, the integration of AI and genomics has led to significant developments in disease analysis and prediction, resulting in faster decision-making.

However, the integration of multi-omics datasets is crucial in capturing the complexity of each omics approach. More benchmark studies are needed to determine the best machine-learning strategy to implement. Multi-omics integrative models can help in understanding disease abnormalities that are not always possible with only genomic or other single-omics analysis.

The field of precision medicine has seen remarkable growth with the advent of advancements in AI, healthcare, clinical genomics, and pharmacogenomics. These developments have generated an enormous amount of data, which can provide insights into personalized treatment options. However, incomplete or inaccurate healthcare-specific information in open access clinical data and claims data presents difficulties in determining how patient-specific treatment is appropriate or effective. The same limitation exists in genomic databases, where data cannot be easily transferred from one database to another. These limitations complicate the process of cross-referencing data between different databases, which can prove to be an obstacle to efficient patient treatment.

Pharmacogenomics focuses on an individual's reaction to specific treatments and medications rather than the disease itself. By correlating a patient's genomic makeup and their reaction to treatments, it allows for more precise and personalized prescription of treatment. However, this field is still developing and has not been utilized reliably.

The large influx of data generated in precision medicine presents an issue, as no reliable or standardized means of analysis has been developed. The use of AI and ML techniques alleviates this issue by allowing for efficient data management and the ability to recognize patterns in complex datasets. These techniques can predict pharmaceutical properties of drug targets and drug candidates, which is especially beneficial in clinical settings.

Conclusion

The integration of artificial intelligence (AI), machine learning (ML), and blockchain technology in the fields of genomics, healthcare, and biotechnology has the potential to modernize these industries. AI can improve diagnosis accuracy, aid in clinical decision-making, and provide personalized patient experiences. However, ethical and social considerations must be taken into account.

ML is increasingly important in genomics research and may become an even more crucial tool in unlocking the secrets of the genome. In the healthcare sector, the integration of AI and ML can revolutionize patient care by improving access to data and using it more efficiently. Additionally, the development of precision medicine requires a combination of different approaches.

Blockchain technology can enhance security and transparency in the storage and sharing of patient data. However, energy and computation efficiency should be considered when implementing this technology. Despite the challenges that need to be addressed, the potential benefits of these technologies in improving patient care and disease prevention cannot be overlooked.

The impact of big data analytics and AI on the healthcare industry is enormous. AI-based tools can help mitigate healthcare fraud, medical coding errors, and improve patient care. Healthcare policymakers and government also use AI-based tools to control and predict infections and outbreaks. With the advent of COVID-19, AI models can predict at-risk populations and provide additional risk information to clinicians caring for at-risk patients. In our journey towards a progressively technology-driven future, it is crucial that we place utmost importance on the well-being of every individual inhabiting our planet.

References

- Ashby WR (1957). An Introduction to Cybernetics. Chapman & Hall Ltd., London, UK.
- Biamonte J, Wittek P, Pancotti N (2017). Quantum machine learning. Nature, 549: 195-202.
- Chen M, Mao A, Xu M, Weng Q, Mao J, Ji J (2019). CRISPR-Cas9 for Cancer Therapy: Opportunities and Challenges. Cancer Letter, 447: 48–55.
- Coral Health (2023). https://www.coralhealth.com/
- Davenport T & Kalakota R (2019). The potential for artificial intelligence in healthcare. Future Healthcare Journal, 6: 94–98.
- Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K (2019). A Guide to Deep Learning in Healthcare. Nature Medicine, 25 (1): 24-29.
- Fletcher S, Meloni PL, Johnsen RD, Wong BL, Muntoni F, Wilton SD (2013). Antisense Suppression of Donor Splice Site Mutations in the Dystrophin Gene Transcript. Molecular Genetics & Genomic Medicine, 1 (3): 162-173.
- Frangoul H, Altshuler D, Cappellini MD, Chen YS, Domm J, Eustace BK, Foell J, de la Fuente J, Grupp S, Handgretinger R, Ho TW, Kattamis A, Kernytsky A, Lekstrom-Himes J, Li AM, Locatelli F, Mapara MY, de Montalembert M, Rondelli D, Sharma A, Sheth S, Soni S, Steinberg MH, Wall D, Yen A, Corbacioglu S (2021). CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β-Thalassemia. N. English Journal of Medicine, 384 (3): 252–260.
- Garg R (2023a). Virtual Worlds, Real Lives: Impact of AI, ML and Blockchain. Taylor & Francis, Routledge, Oxfordshire, UK.
- Garg R (2023b). Blockchain for Real World Applications. John Wiley & Sons Inc., New Jersey, US.
- Gartner (2001). https://blogs.gartner.com/
- Green MR, Sambrook J (2018). The Basic Polymerase Chain Reaction (PCR). Cold Spring Harbor Protocols, 2018 (5).
- Gurovich Y, Hanani Y, Bar O (2019). Identifying facial phenotypes of genetic disorders using deep learning. Nature Medicine, 25: 60–64.
- Hessler G, Baringhaus KH (2018). Artificial Intelligence in Drug Design. Molecules 23 (10), 2520.
- Mintz Y, Brodie R (2019). Introduction to Artificial Intelligence in Medicine. Minimally Invasive Therapy & Allied Technologies, 28 (2): 73–81.
- International Human Genome Sequencing Consortium (2001). Initial Sequencing and Analysis of the Human Genome. Nature, 412 (6846): 565.
- Joudaki H, Rashidian A, Minaei-Bidgoli B, Mahmoodi M, Geraili B, Nasiri M, Arab M (2015). Improving fraud and abuse detection in general physician claims: a data mining study. International Journal of Health Policy Management, 5: 165-172.
- LeCun Y, Bengio Y, Hinton G (2015). Deep learning. Nature, 521(7553): 436- 444.
- Liu L, Li Y, Li S, Hu N, He Y, Pong R et al. (2012). Comparison of Next-Generation Sequencing Systems. Journal of Biomedical Biotechnology, 2012: 1-11.
- Mattick JS, Makunin IV (2006). Non-coding RNA. Human Molecular Genetics. 15 (Suppl. L-1), R17-R29.
- McDonnell JM, Evans SR, McCarthy L, Temperley H, Waters C, Ahern D, et al (2021). The Diagnostic and Prognostic Value of Artificial Intelligence and Artificial Neural Networks in Spinal Surgery. Bone & Joint Journal, 103-B (9): 1442-1448.

- Mintz Y, Brodie R (2019). Introduction to Artificial Intelligence in Medicine. Minimally Invasive Therapy & Allied Technologies, 28 (2): 73–81.
- National Research Council (2011). Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease (The National Academies Press, Washington, DC).
- Nebula Genomics (2023). https://nebula.org/whole-genome-sequencing-dna-test/
- Ramesh A, Kambhampati C, Monson J, Drew P (2004). Artificial Intelligence in Medicine. Annals of Royal College of Surgeons of England, 86 (5), 334–338.
- Rhoads A, Au KF (2015). PacBio Sequencing and its Applications. Genomics, proteomics Bioinformatics, 13 (5): 278-289.
- Schwarze K, Buchanan J, Fermont JM (2020). The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. Genet Med 22, 85–94.
- Tzelepis K, Koike-Yusa H, De Braekeleer E, Li Y, Metzakopian E, Dovey OM, Mupo A, Grinkevich V, Li M, Mazan M, Gozdecka M, Ohnishi S, Cooper J, Patel M, McKerrell T, Chen B, Domingues AF, Gallipoli P, Teichmann S, Ponstingl H, McDermott U, Saez-Rodriguez J, Huntly BJP, Iorio F, Pina C, Vassiliou GS, Yusa K (2016). A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. Cell Reports, 17 (4): 1193-1205.
- Vaishya R, Javaid M, Khan IH, Haleem A (2020). Artificial intelligence (AI) applications for COVID- 19 pandemic. Diabetes & Metabolism Syndrome. 14, 337-339.
- Xiao Q, Guo D, Chen S (2019). Application of CRISPR/Cas9-based Gene Editing in HIV-1/AIDS Therapy. Frontiers in Cellular and Infection Microbiology, 9: 69.
- Zhao J, Feng Q, Wu P, Lupu RA, Wilke RA, Wells QS, Denny JC, Wei WQ (2019). Learning from longitudinal data in electronic health record and genetic data to improve cardiovascular event prediction. Scientific Reports, 9: 717.
- Zou J, Huss M, Abid A, Mohammadi P, Torkamani A, Telenti A (2019). A primer on deep learning in genomics. Nature Genetics, 51(1): 12- 18.