Bulletin of Environment, Pharmacology and Life Sciences
Bull. Env. Pharmacol. Life Sci., Special Issue [1]2022 : 300-305

Online ISSN 2277-1808
Journal's URL:http://www.bepls.com
CODEN: BEPLAD
REVIEW ARTICLE **OPEN ACCESS**

Upcoming future of Biotechnology: Machine learning and Artificial intelligence

Simran Srivastava, Arsh Singh and Amit Gupta*
Department of Biotechnology, Graphic Era Deemed to be University, Dehradun
***Corresponding author's** Email: dr.amitgupta.bt@geu.ac.inamitvsbt@gmail.com

ABSTRACT
In this modern era the emerging field of biotechnology has been an integration of many other important fields. One of them is Computational Sciences which is a vast field including data management Systems, artificial intelligence, computational statistics, machine learning and many more. Biotechnology is an interdisciplinary area in which the use of the advanced biological techniques and novel fields such as nano-technology, gene editing technology, CRISPR, etc. computational techniques are very robust and can be used in multiple ways like discovery of genetic biomarkers, production of protein chips for enhanced and robust diagnosis, omics studies, machine learning based tools. However, this requires great knowledge of biology, data analysis and coding skills. we here are going to discuss about the integrated fields and their therapeutic advances and their uses. In this paper, we discuss about contribution of machine learning and artificial intelligence in upcoming future of biotechnological field.
Key words: *Biotechnology; artificial intelligence; machine learning*

Received 19.02.2022 Revised 11.03.2022 Accepted 02.04.2022

INTRODUCTION

Biotechnology is "the implementation of organisms, cells, portions of organisms, and molecular replicas for products and services through the integration of biological sciences and applied sciences." In 1919, Karl Ereky first utilized the title *biotechnology* which means the act of making goods out of raw resources with the assistance of life forms [1].Biotechnology is an integrative area of study comprising not just biology, but other areas like physics, nanotechnology, engineering, mathematics, computational biology, chemistry and so on. It is the fusion of several innovations combined with viable cells for the manufacture of a certain product or to modify it. Its aid and implications range from areas like agriculture to industry (chemical, food, textiles, pharmaceutical, bioproducts etc.), medicinal, nutritional, ecological sustainability, animal sciences etc. rendering it one of the biggest emerging fields[2]. Biotechnology has provided treatments and potential cause for the majority of ailments and autoimmune conditions; gene therapy has aided progression in the disciplines of medicine, agricultural sectors, and private sector; and its implementation in atmospheric sciences has resulted in waste minimization and the development of sustainable energy sources[3]. Laboratory management system, product-specific software, and bio-imaging software are among the several kinds of software accessible. Biotechnology has unquestionably risen to the top due to the rapid rate of new discoveries and technologies.

MACHINE LEARNING IN BIOTECHNOLOGY

Machine learning (ML) is the analysis of computer algorithms that adapt and develop on their own with experience and data [4].It is considered to be a type of artificial intelligence(AI) technology. ML algorithms create a model based on training data to make predictions or choices despite having to be specifically designed to do so[5].ML algorithms are widely used in a number of applications, including healthcare, web filtering, language processing, and computer vision, where developing traditional algorithms to do the required tasks is hard or impossible. The major applications of ML in biotechnology [4-6]are-

Identifying the coding areas of genes-Next generation sequencing has revolutionized genomics by speedily analysing a genome in a short period of time. In order to discover gene coding are as in a genome, dynamic zone ML is used. ML-based gene prediction methods would be more precise than traditional homolog-based sequence analyses.

Prediction of structure- The aid of ML in structure prediction has increased efficacy from 70% to over 80%. Using training sets to find new or unique pharmacological targets from many journal articles and scanning secondary databases, the application of ML in text mining is extremely promising.
Neural networks- Deep learning is a relatively modern ML subfield that is an evolution of neural networks. The term "deep" in deep learning represents the amount of layers that data is converted through as a result, deep learning is analogous to a multi-layer neural network. These multi-layer nodes attempt to simulate how the human mind solves problems. In contrast, ML already use neural networks and to undertake analysis, neural network-based machine learning algorithms require enhanced or meaningful insights from raw data sets. However, as the amount of genome sequencing data grew, it became more difficult to analyse useful data and subsequently undertake analysis. The information is filtered by multiple layers in a neural network, which interact with each other and allow the output to be refined.
Deep learning algorithms gather traits from big data sets, such as a set of photos or genomes, and use those features to build a model. Algorithms can even use the produced model to analyse different data sets once the model has been developed. Deep learning algorithms are now being used by scientists to do cellular image classification, genome assessment, drug identification, and determining how image and genome data are linked with electronic clinical records. Deep learning is now a popular topic in computational biology. Deep learning is being used on slightly elevated biological data in order to gain a better understanding of a large data set.

ML BASED TOOLS USED IN BIOLOGY
Cell Profiler: Until recently, biological image analysis software only assessed a single parameter from a series of photos. Anne Carpenter, a computational biologist from MIT and Harvard, developed Cell Profiler in 2005 for the quantitative determination of individual parameters like fluorescent cell number in the microscope field. However, CellProfiler can presently generate thousands of characteristics using deep learning approaches[7].
DeepVariant: Deep learning applications are widely employed in genetic data mining technologies. DeepVariant, a deep learning tool developed by Verily Life Science and Google, predicts a frequent form of genetic mutation more precisely than conventional techniques[7, 8].
Atomwise: Another domain wherein deep learning plays a key role in drug discovery. Atom wise, a biotech firm based in San Francisco, has developed an algorithm that aids in the conversion of molecules into 3D pixels. This representation aids in the accurate accounting of the 3D structure of proteins and micro entities. The system may then suggest small molecules that may bind with a specific protein using these features [7, 8].

BIG DATA IN BIOTECHNOLOGY
In recent years, the possibilities of applying technology for evaluating enormous volumes of data in Big Data database systems in healthcare have grown. And, thanks to the use of Big Data databases for research, all other implementations of biotechnology, such as agriculture and genetic engineering, may be developed even better and more efficient. Big Data is similar to a virtual library in that it stores and analyses a vast amount of data and information. And pursuing it in a specific way can aid biotech in improving and achieving even greater success in a shorter period of time with less effort.The major applications of big data in biotechnology are-
Genomics- The Human Genomic Project(HGP) took a very long time of research reported worldwide and providing assistance to explored more than 20,000 genes and alsogetting the sequencing of all genome bases (3 billion). This effort costs billions of dollars around the world, yet today's biotechnology companies use a Big Data database that can analyse complete genomes for just a few thousand dollars. The genomics industry assists several data organisations that utilise tools and concepts to perform massive and complex computational activities in order to evaluate genetic, medicinal, and biological data. These businesses frequently collaborate with computer hardware technology firms to increase usage performance and Big Data analysis outcomes[9].
Agriculture- In the domain of agriculture, big data has a lot of potential and stored big data acquired from GPS technology, and many GPS-enabled machines can assist farmers manage with evolving climatic conditions by accurately executing farming. Big data is also transforming the nature of the biotechnology industry with its participation to genetic analysis in generating genetically modified organisms(GMOs). All these modified crops can be updated with data from Big Data to increase crop productivity, adapt to altering environments, and produce healthy plants[10].
Pharma Automation- Almost every pharmaceutical firm acquires millions of molecules before deciding which ones are suitable for pre-clinical testing. It takes a significant quantity of time and resources to get an effective drug development. As a result, there are numerous software solutions that aid inefficiencies

and reduce drug discovery time. Big Data based modelling makes use of massive amounts of data and storage, such as terabytes(TBs) of data and knowledge about various chemicals and their properties. As a result, it functions as a virtual library with information on billions of compounds, allowing researchers to pinpoint the compounds that are most likely to succeed. The trial parameters and anticipated outcomes are compared to the specific disease and chemical patterns in these predictive modelling algorithms. Automation in the pharmaceutical industry lowers risks, saves a lot of money, and speeds up the research-to-market process [9, 10].

Crowdsourcing- Crowdsourcing is a model for obtaining products and services by individuals or groups. These services, which comprise ideas and funds, cater to a vast, open, and frequently changing community of web users. It distributes work among participants in order to achieve a collective product. Hence it is frequently utilised outsourcing workers and entrepreneurial initiatives. Some pharmaceutical firms have developed online multiplayer platforms that include disease descriptions, research tasks, and medical puzzle solving. Patients drive research using online questionnaires, allowing them to execute their respective research and studies, submit their private medical data, and provide insights about their illness and symptoms to the sake of the entire medical community [11, 12].

Preventing drug-related fraud- Fake medications kill countless individuals and harm their health every day in developing countries. Patients and their families have lost faith in pharmaceutical businesses as a result of this tragic situation, and sales have suffered as a result. According to the World Health Organization(WHO), fake forms of anti-malaria and tuberculosis drugs kill 700,000 sick people in Africa each year, costing drug companies $75 billion. The problem is so serious that Sproxil has teamed up with IBM ($IBM) to help drug businesses examine Big Data sources to discover indications of forgery drug activity. With a system that allows consumers to short message codes from prescription bottles to verify if the meds are legitimate, Sproxil hopes to collect vast volumes of transactional data. According to Big Blue, drugmakers can tap a significant proportion of statistics on drug trades in real-time using IBM's visualisation technology and other algorithms. Prescription medicine fraud should, presumably, be detectable [13].

Genetic biomarkers are being discovered- Numerous genomic screening techniques are available to find DNA code variations and disease risk factor genetic biomarkers. Certain Big Data analytics and informatics tools can combine several data kinds to produce better results. The capacity to link vast data banks, morphological, and genomic data allows researchers to gain a better knowledge of illness determinants, symptoms, and progression. Researchers can evaluate sequence conformations and disease data among objective findings using software solutions. As a result, Big Data aids in the discovery of biomarkers as well as the identification and treatment of diseases [11-13].

NANOTECHNOLOGY IN BIOTECHNOLOGY

Nanotechnology is concerned with the study and management of materials at scales of 1 to 100 nanometers, where peculiar phenomena allow for fresh applications. Nanotechnology is defined as the image processing, modelling, measuring, designing, production, characterization, and implementation of structures, gadgets, and systems at the nanoscale level (macromolecular, molecular, and atomicscale) that results in arrangements, tools, and systems including at least one novel/superior attribute or property. Nanotechnology in healthcare and physiology refers to mechanisms and technologies that are so technologically advanced that they can associate with sub-cellular (i.e. molecule) stages of the body with extreme precision. By using a focused cell or tissue-specific clinical intervention, clinical effectiveness can be maximized with minimal adverse effects[14, 15].The major applications [16-18]of nanotechnology in biotechnology are-

Diagnosis applications- Many presently utilized/standard clinical check-ups disclose the existence of an ailment triggering organism or a molecule and identifying the capacity of antibody binding with ailment-associated target. Generally, such tests or experiments are executed through conjoining especially antibodies with some organic/inorganic dyes and envisioning the indications which is reported in the samples or specimens. This is mainly observed and confirmed through electronic microscopy or fluorescence microscopy. Dyes, on the other hand, frequently impair the accuracy and applicability of detection techniques. The use of semiconductor nanocrystals in nanobiotechnology provides a solution (also known as "quantum dots"). So, microscopic probes may be able to tolerate far more excitations and light discharges cycles than normal organic molecules, which degrade more quickly.

Proteins are more suggestive of functionality and play a vital role in determining the biological appearance of organisms in both healthy and sick states. As a result, proteomics is crucial in illness diagnostics and medicine, where medications to modify signalling pathways can be produced. Chemical classes, or micro modular peptide elements, can be added to protein chips to specifically attach to proteins with a specific structural or biochemical sequence [16].Agilent, Inc. and NanoInk, Inc. are two businesses that are currently active in this application field. Agilent creates microarrays using a non-contact ink-jet method that prints oligos and full cDNAs onto glass plate at the nanoscale. Dip-pen nanolithography (DPN) technique is used by NanoInk to construct structures on a nanoscale.

In normal physiological circumstances, sparse cells are indeed relatively uncommon and physiologically diverse from their surrounding cells (e.g. cancer cells, lymphocytes, foetal cells and HIV-infected T cells). They play an important role in the identification and treatment of a variety of genetic disorders. Identifying and isolating these scarce cells, on the other hand, is difficult. Nanobiotechnology opens up new avenues for progress in this field. Nanosystems have been developed that can successfully select sparse cells from blood and tissues. The distinctive features of sparse cells, exhibited in variances in deformation, surface charges, and sensitivity for certain receptors and/or ligands, are used in this technology. Cells can be accurately separated based on spatial charge by putting wires into microchannels, for example. Biocompatible layers with accurate nanopores can also be used to sort them. These technologies are currently being used at Cornell University's nano-biotechnology centre to build impactful diagnostic equipment for the extraction and detection of several illnesses.

Therapeutic uses [19-22]

Nanoparticles as treatments can be administered to specific places, even those that are difficult to reach with conventional medications. If a treatment can be chemically linked to a nanoparticle, for example, it can be steered to the disease or infection site via radio or magnetic indicators. These nanodrugs can also be programmed to only "release" when certain molecules are present or external stimuli (like as infrared heat) are applied. The release of medications can be controlled far more accurately than ever before by encapsulating them in nanosized components (like as nanoshells, organic dendrimers, and hollow polymer capsules). Drugs can have a therapeutic burden (chemotherapy, gene therapy, or radiotherapy) as well as be used for imaging. Microparticles and nanoparticles can enhance immunisation, according to recent developments in encapsulation and the establishment of relevant animal models.

The inherent limitations of successful pharmaceutical processing and manufacturing, as well as the risk of reversion of a created mutant to the natural type, concern existing gene therapy systems. The pathogenicity of viral vectors used to carry genes is also a concern. To solve this problem, nanotechnological methods in human gene therapy have been investigated, and nanoparticle-based nonviral vectors for plasmid DNA transportation (typically 50-500 nm in size) have been described. As a result, the effective adoption of less immunogenic nanoscale gene carriers as a replacement for the controversial viral vectors appears to be helpful in human gene repair or replacement.

The area of dentistry will benefit from nanotechnology in the future. The employment of biotechnology, nanorobotics, and nanomaterials in nanodentistry will provide greater dental health. Lots of people who are currently receiving substandard dental treatment will benefit from this significant advancement in oral health science. In addition, nanodental procedures for extensive tooth repair may develop. Reconstructive tooth nanorobots could be utilised in minutes to selectively and precisely occlude specific tubules, allowing for a faster and more durable healing. Nanodentistry may have a considerable benefit in natural tooth care. Artificial materials with covalent bonds, such as sapphire, may be used to replace the uppermost enamel covering to improve the aesthetic and lifespan of teeth.

Because of its capacity to penetrate over lipid membranes and cellular membranes of the target, a liposome constituted of a lipid membrane can be employed in gene therapy. The employment of many types of liposomes in a regional delivery system has recently been proven to be beneficial. Liposomes can potentially aid in the delivery of specific medication. By connecting nanoparticle (such as polyethylene glycol) processed liposomes to a monoclonal antibody for human insulin reporter, Zhang et al exhibited broad reporter expression in the brains of rhesus monkeys. These successful studies illustrate the importance of nanometer-sized structures in the evolution of molecular medicine and the future of targeted therapy.

ARTIFICIAL INTELLIGENCE (AI) IN BIOTECHNOLOGY

AI and ML has turned out to be omnipresent in technology startups, fueled to a great extent by the expanding accessibility, measure of amount of information to be processed and less expensive, yet more powerful computers. In the course of recent years, AI and ML have been able to enrich the biotech domain duet analogous transformation of biotech information. A survey of pharma and life sciences experts showed that 44% were using AI in their R&D activities. The survey also revealed that AI majorly finds applications in preclinical phases of drug development. The major applications [23-26] of AI in biotechnology are-

Drug research and development, as well as clinical trials- The most intriguing implementation of Artificial Intelligence and Machine Learning has been drug discovery. Organizations are pursuing a structure-based approach to drug development, employing machine learning to find tiny compounds with therapeutic potential based on known target structures. The majority of AI use-cases and emerging technologies for clinical trials appear to be focused on three key applications: patient recruiting, clinical trial design, and trial optimization.

Prognostics- AI is being used to help with the treatment of malignant tumours. With Quest Diagnostics, IBM developed IBM Watson Genomics, which employs artificial intelligence to improve the accuracy of malignant growth detection. Pathology and the diagnosis of rare diseases are two more ML uses. According to a current study, ML is more accurate than cardiologists in identifying cardiac diseases.

Radiology and radiotherapy- AI has proven to be useful in shortening the planning procedure for radiation therapy to only minutes, saving doctors many days and increasing patient care. DeepMind Health is collaborating with University College London Hospital to develop machine learning calculations that will improve the precision of radiotherapy planning by distinguishing healthy tissues from cancerous ones.

Gene editing- Complicated activities, such as creating structures for gene editing, are handled by AI systems that act as assistants. Desktop Genetics has developed an AI-powered platform using CRISPR gene editing structures. From selection of RNA to data analysis, their gene editing software manages and controls the entire process.

Electronic Health Record(EHR)- Evidence-based therapy and clinical choice support systems built on the machine learning platform have the potential to enhance the power of an EHR system by assisting doctors in making smart clinical decisions based on a patient's preferences and clinical history. AI and digital technology can also be used to successfully handle medical records. For better patient care, the massive volume of data can be securely maintained, structured, and accessible.

Medication management- To track patients' medication regimens, mobile applications are being created. Patients' prescriptions are managed using a smartphone webcam connected to AI. Patients with chronic illnesses, as well as clinical trial participants, may benefit from them.

CONCLUSION

In this review, we looked and focused on some historical aspects of biotechnology and correlate with ML and AI. Finally, this software's related to ML and AI may help in order to automate, accelerate and scale the process and help scientists or medical specialists discover drugs systematically faster.

REFERENCES

1. Fari MG and Kralovánszky UP. (2006). The founding father of biotechnology: Karl Ereky Int J Hort Sci; 12: 9–12.
2. Steinberg FM, Raso J. (1999). Biotech pharmaceuticals and biotherapy: an overview. J Pharm Pharm Sci ; 1: 48–59.
3. Verma AS, Agrahari S, Rastogi S, Singh A. (2011). Biotechnology in the realm of history. J Pharm Bioallied Sci ; 3: 321–3.
4. Tian H, Ning Z, Zong Z, Liu J, Hu C, Ying H, Li H. (2002). Application of Machine Learning Algorithms to Predict Lymph Node Metastasis in Early Gastric Cancer. Front Med (Lausanne); 8:759013.
5. Raina CK. (2016). A review on machine learning techniques. International Journal on Recent and Innovation Trends in Computing and Communication; 4(3): 395-399.
6. Jordan MI and Mitchell TM. (2015). Machine learning: Trends, perspectives, and prospects. Science; 349(6245): 255-260.
7. Praveena M and Jaiganesh V. (2017). A literature review on supervised machine learning algorithms and boosting process. International Journal of Computer Applications; 169(8): 32-35.
8. Angermueller C, Pärnamaa T, Parts L and Stegle O. (2016). Deep learning for computational biology. Molecular systems biology; 12(7): 878.
9. Provost F and Fawcett T. (2013). Data science and its relationship to big data and data-driven decision making. Big Data ; 1:51–9.
10. Antle J, Jones J and Rosenzweig C. (2017). Next Generation agricultural system data, models and knowledge products: introduction. Agric Sys ; 155: 186–190.
11. Howe J. (2006). The rise of crowd sourcing. Wired Magazine;14:1–4.
12. Plenge RM, *et al*. (2013). Crowdsourcing genetic prediction of clinical utility in the Rheumatoid Arthritis Responder Challenge. Nat. Genet. 45:468–469.
13. George SL. (1997). Perspectives on scientific misconduct and fraud in clinical trials. Chance. 10(4):3–5.
14. Emerich DF, Thanos CG. (2003). Nanotechnology and medicine. Expert Opin Biol Ther, 3: 655-663.
15. Sahoo KS, Labhasetwar V: (2003). Nanotech approaches to drug delivery and imaging. DDT. 8 (24): 1112-1120.
16. Vasir JK, Reddy MK, Labhasetwar V: (2005). Nanosystems in drug targeting: opportunities and challenges. Curr Nanosci ; 1: 47-64.
17. Bethi B and Sonawane SH. (2018). Nanomaterials and its application for clean environment. In: Nanomaterials for green energy ; 385–409.
18. Bodalo-Santoyo A, Gómez-Carrasco JL, Gomez-Gomez E, Maximo-Martin F and Hidalgo-Montesinos AM. (2003). Application of reverse osmosis to reduce pollutants present in industrial wastewater. Desalination; 155(2):101–108.

19. Chiang HM, Xia Q, Zou X, Wang C, Wang S, Miller BJ, Howard PC, Yin JJ, Beland FA, Yu H, Fu PP. (2012). Nanoscale ZnO induces cytotoxicity and DNA damage in human cell lines and rat primary neuronal cells. J Nanosci Nanotechnol ; 12(3):2126–2135.
20. Cheng SH, Lee CH, Chen MC, Souris JS, Tseng FG, et al. (2010). Tri-functionalization of mesoporous silica nanoparticles for comprehensive cancer theranostics-the trio of imaging, targeting and therapy. Journal of Materials Chemistry ; 20: 6149-6157.
21. Menaa B, Menaa F, Aiolfi-Guimaraes C, Sharts O. (2010). Silica-based nanoporous sol-gel glasses: from bioencapsulation to protein folding studies. International. Journal of Nanotechnology ; 7: 1.
22. Menaa F, Menaa B, Sharts O. (2011). Development of carbon-fluorine spectroscopy for pharmaceutical and biomedical applications. Faraday Discussions ; 149: 269-278.
23. Tripathi MK, Nath A, Singh TP, Ethayathulla AS, Kaur P. (2021). Evolving scenario of big data and Artificial Intelligence (AI) in drug discovery. Mol Divers ; 25(3):1439-1460.
24. Kashyap K, Siddiqi MI. (2021). Recent trends in artificial intelligence-driven identification and development of anti-neurodegenerative therapeutic agents. Mol Divers. 25(3):1517-1539.
25. Zhu H. (2020). Big Data and Artificial Intelligence Modelling for Drug Discovery. Annu Rev PharmacolToxicol; 60:573-589.
26. Istepanian RSH, Al-Anzi T.M (2018). Health 2.0: New perspectives on mobile health, machine learning and big data analytics.Methods ; 151:34-40.

CITATION OF THIS ARTICLE

Simran Srivastava, Arsh Singh and Amit Gupta. Upcoming future of Biotechnology: Machine learning and Artificial intelligence . Bull. Env.Pharmacol. Life Sci., Spl Issue [1] 2022 : 300-305