IOSR Journal of Research & Method in Education (IOSR-JRME)
e-ISSN: 2320–7388, p- ISSN: 2320-737x Volume 13, Issue 2 Ser. III (Mar. – Apr. 2023), 33-37
www.iosrjournals.org

Artificial intelligence and machine learning: its impact on clinical data management

Ankireddypalli Sreeja,[1] Lakshmi Prasanna Reddy[2], Noorush Shifa Nizami[3]
[1]*Corresponding Author, Pharm D, Student, ClinoSol Research Pvt Ltd*
[2]*Co-Author, Bachelor of physiotherapy, Student, ClinoSol Research Pvt Ltd*
[3]*Co-Author, Pharm D, Research Analyst*

Abstract: *Artificial Intelligence (AI) is defined as the study of intelligent agents, which are devices that perceive their environment and take actions to maximize their chance of success at some goal. A subset of AI is Machine learning (ML), which learns on a trial & error basis and improves its performance on the results. Deep learning, a subset of ML, uses neural networks, which are one of the most beautiful programming paradigms ever invented. Clinical drug development has remained relatively unchanged for the last 30 years. The most common applications of artificial intelligence (AI) in drug treatment have to do with matching patients to their optimal drug or combination of drugs, predicting drug-target or drug-drug interactions, and optimizing treatment protocols. AI can also identify patients who might respond better to a certain medication and predict possible complications, thus optimizing patient selection, augmenting the speed of discovery, and reducing the cost burden.*
Keywords: *Intelligent agents, trial & error basis, deep learning.*

Date of Submission: 22-03-2023 Date of Acceptance: 05-04-2023

I. Introduction:

Artificial Intelligence (AI) is defined as the study of intelligent agents, which are devices that perceive their environment and take actions to maximize their chance of success at some goal. A subset of AI is Machine learning (ML), which learns on a trial & error basis and improves its performance on the results. ML has continued to show promising results in optimizing processes and resource allocation and preparing new data using versatile methods. Deep learning, a subset of ML, uses neural networks, which are one of the most beautiful programming paradigms ever invented. Neural networks specifically and artificial intelligence in general have a vast eventuality to exponentially optimize health care exploration, especially in the epidemic where time is of the substance. Clinical data, epidemiological data, and inheritable data have to be all reused together to come up with the stylish possible forestallment, opinion, and operation of complaint, and decide it to public health measures. To attack similar variations, there's a high need for precise and intertwined command and control systems, and ML brings a result to all these complex problems. [1] The healthcare sector is witnessing a paradigm shift due to the recent growing trend of mass digitization of its information. utmost hospitals in the world are exercising electronic health records (EHRs) for recording all patient-affiliated data and, due to this digitization, a lot of structured and unshaped data is generated. Also, technological advancement has allowed data to be continuously generated from medical bias and wearable Internet of effects (IoT) bias. This has led to an exponential growth of healthcare data over a short period and redounded in what's known as big data. Despite having growth, there's a significant gap in the knowledge within the neurosurgical community regarding big data and its counter-accusations. Although there are many excellent reviews about big data in the neurosurgical literature, utmost of them concentrate on neurosurgical registries, thereby creating a false print in the mind of the anthology that big data solely comprises big registries. Many go further, using technical languages like artificial intelligence (AI), machine literacy, artificial neural network (ANN), deep literacy, and natural language processing without easily defining these generalities, therefore confining the knowledge and use of this field to only a sprinkle. Despite having access to this vast quantum of data, croakers and surgeons warrant the time, tools, and moxie to take full advantage of what it has to offer. At the other extreme, some interpreters believe that the application of big data will give results to all healthcare-related problems and exploration questions. To this end, report on the capabilities as well as the limitations and challenges of big data from the time of data accession to the final visualization of the outgrowth. [2]

Big Data

Big data is a term used to describe exceedingly large data sets, characterized by their complexity and variety. Big data is generally represented by the 4 V's volume, variety, haste, and veracity. [3]

Why Do We Need Similar Big Data?

Knowing the relationship between big data, machine literacy, and AI is a prerequisite to understanding the significance of data and registries in neurosurgery. utmost of the big data operations is directly or laterally affiliated with the field of data analytics, AI, and robotics. Then, we give a brief sapience into AI and robotics, followed by further general operations of big data in the medical field. AI and Robotics In general, AI refers to a subfield of data wisdom and computer wisdom, in which computers perform tasks that would generally bear mortal intelligence. The term artificial intelligence frequently conjures up images of robots, though this is generally not the case; robotics is just one element of AI. It's giving machines the capability to smell, reason, engage, and learn. Robotics and stir, natural language processing, voice recognition, and computer vision are types of AI. AI combines and uses machine literacy algorithms, ANN, deep literacy, and other data logical ways to gain intelligent capabilities. lately, the US FDA has cleared a deep- literacy algorithm- grounded AI calledViz.ai that analyzes images to descry implicit large- vessel occlusion strokes on CT angiography and alert the specialist to initiate exigency treatment, whereas in robotics, machines with mechanical corridor use the same styles of literacy to suppose and perform like humans. Clinical outgrowth Analysis Machine literacy can be used to ameliorate diagnostics as well as in the vaticination of patient issues and opinion. numerous studies have proven that prognostications deduced from machine literacy may be better than conventional statistical models of vaticination when applied to neurosurgical cases.[4].

The analysis and authorization of new pharmaceuticals is based upon the trust that clinical trials will intend to find the answers to the investigation problems by providing clinical data which further proves or disapproves a particular hypothesis. The type and quality of the clinical data plays an imperative task in the conclusion of the study performed. Therefore, this clinical data so obtained is appropriately managed to obtain the accurate results of the clinical study thus, a system Clinical Data Management (CDM) is needed for the authentication of the study. [5]

What is CDM?

Clinical Data Management forms a fundamental part in the clinical trial studies. CDM is implied in all the facets of operating computers, dispensation of the clinical data, managing the subject data and database systems to support the collection of the data. Clinical Data Management is precisely defined as the collection, integration and validation of the trial data. When the clinical trials are performed, the prime duty of the investigators is to collect the data of the patient's wellbeing after a specific interval of time Further, this data is given to the trial guarantor who quantifies and qualifies the given data by statistical means. When the blessing of new medicines is to be made by the nonsupervisory agencies it's reliant upon the clinical trial data presented. The trust on the clinical data is generally stuck to the quality practices and norms of the clinical trials performed. [6] Thus, the associations assure that the clinical trials performed and the data attained are in the hands of well good and trained staff and an inflow map representing the staff that's involved in the clinical data operation system. therefore, the crucial ideal of CDM is to offer high quality data by observing the crimes and missing data and keeping it as low as possible to congregate maximum data for analysis. colourful practices have been developed to insure that the data attained is complete, reused rightly and dependable. This has been fluently achieved by the use of operations of the software that presents royal discovery and stir of data disagreement and is used to maintain inspection trials. In clinical data operation, software is generally needed to address the electronic data prisoner, medication of the electronic FDA submission, acceleration of the clinical trial operation processes. [7] There are good clinical data operation practices that deal with data accession, sequestration, electronic data capturing, Case Report Form (CRF) printing, preservation of CRF, data storehouse, attestations and numerous further. Standard operating procedures are the processes that are followed to negotiate data operation conditioning and to support the responsibility of adhering the guidelines as per ICH GCP and 21 CFR part 11. Standard operating procedures or bribes are generally applied in medicinal processing and are also related for clinical studies. In clinical studies, the main focus is on recreating operation of innocent processes and their attestation thus, it supports the insulation of origins, goods and causes. Further, operations made are with respect to the precedence of patient treatments when the defined sources get employed according to estimation on urgency, staffing possibilities and ranking. The quality assurance platoon is responsible for covering that the study and test meet the bribe. bribes also act as a reference to new workers by answering questions without interposing administrators. [8]

Case report forms

A case report form can be electronic or a paper- grounded system and is generally shortened as CRF. It's extensively used tool by the guarantor for the collection of the data from the cases sharing in a clinical trial.[9]

Paper grounded systems

In the paper grounded systems the case report forms are filled manually at the point and are also posted to the company. The data that has been collected is farther transferred to the CDMS tool that is, Clinical Data Management System through data entry. For this matter, the most common system used is the double data entry

system wherein, two different data entry drivers enter the data in the system independently and both the entries made are compared by the system. In case, if there's any divergence in the entry the system sends cautions and the verification can be done manually. Also, a single data entry system is extensively used in which a single driver enters the data in the system. The data in the CDMS are further given for confirmation purposes also, during the data confirmation. The data explanation is done from colorful pots through paper forms, which contain the problem description and are also shoot to the investigator point and it responds by answering them through matters.

Electronic data system or EDC is a software system that's used to store the case's or a party's data that's unworried during clinical trials. Before transcribing into the system, the data is first recorded on the paper and also saved as electronic case report form(eCRF). EDC system software is used to conduct both complex and simple clinical trials in all aspects of exploration. Electronic data capture a slow and a steady pace medicinal companies are moving forward towards landing the data of the case record information from the source to the electronic system which has the function of submitting the data to the guarantors or the consumers. This shift of collection of data from paper to electronic system is appertained to as electronic data prisoner. Not all EDC systems are equal in nature and ramification and not every ramification offers a result to requirements of every association. The case report forms also contain the following data collection, crf shadowing, data entry, data confirmation, distinction operation and database locking. [11]

Clinical data operation softwares

All the associations similar as the medical, exploration, biotechnology and Medicinals are getting advantaged from clinical trial software. These software support all the aspects of a clinical trial starting from conscription to the submission of the study and archiving the study. These softwares can be used for small phase trials and also for large studies with thousands of campaigners. There are a variety of softwares that are available for the management of data and are thus called clinical data management systems or softwares. In the case of multicentric systems or trials these softwares play a vital role as they have the ability to handle large amounts of data. Most of these softwares in pharmaceutical industry are commercial but some of the softwares are also available as opensource tools. There are a variety of softwares that are used as CDM softwares in pharmacy some of them are ORACLE CLINICAL, MACRO, RAVE, CLINTRIAL, e-Clinical Suite, EZ-entry and so on. These softwares are more or less similar in their functionality and cannot offer a great advantage over one another. The drawback of these softwares is that they are expensive and the organizations need to acquire a refined Information Technology communications to function. These were the commercial tools that are available as softwares for pharmaceuticals. Among the opensource tools some of the software tools are Open Clinica, TrialIDB, open CDMS and PhOsCo. They are open source and are available as free of cost and are as good as their commercial counterparts. They can be easily downloaded from their websites as they are open source softwares. Some of the advantages offered by these clinical data management softwares are that they accelerate the timeline of the study and also control cost of the study, they have the ability of providing accurate predictions for the clinical trial, they enable a robust modelling of the program, facilitate screening, scheduling, recruiting, assist doctors, automatically adjust for foreign exchange rates. [12]

Key features of different clinical data softwares:

EZ –entry

EZ entry is a software program that is made from the modifications in the EpiData software system. The EpiData software system is the one in which the main functions were 1) use of simple syntax to set up e-CRFs 2) entry of data along with checkout theory 3) checking of the data after double input 4) export of data in various formats. The major drawback of EpiData was that it was a system with low security. EZ- entry is written in Microsoft Visual basic version 6.0 and the server used to manage the data set is Microsoft SQL Server 7.0., the main functions that are included in EZ entry are management of query, revision tracking, entry of the data, import and export of data. The two main functions of the EZ- entry are security protection and quality control. In the EZ-entry system the security depends upon the two aspects viz, user authentication and Revision tracking. In the user authentication system only, authorized user has the concession of accessing the database. In the revision tracking authorized user enters a new data entry or it makes revision in the original database wherein the system records the operations automatically. The quality of the data in EZ- entry system depends upon factors such as field value check, data entry alignment and query form. The software has various modules. [16]

Oracle

Clinical software Oracle software has been used by experts since 30 years as it provided information with a steadfast, protected and incorporated technology. Oracle software has been developed by the organization oracle itself and oracle spends more than US $30 million over the research and development annually. Various advantages offered by the oracle software include the effective team work, faster implementation, productive

marketing and higher return on investment. This software was developed on the basis of extensive experience of hundreds of organizations that conducted clinical trials. [14] More than 10,000 clinical trials have been conducted by using Oracle clinical by more than 200 organizations dealing with biotechnological and pharmaceutical products. Business Intelligence Return of results only at summary level.

SAS Clinical software

SAS Clinical Software is software that is run by the SAS solution for Life Sciences organization. It is one of the leading companies in the world. There are 44000 sites available in around 109 countries. There are no extra charges for technical support and skilled telephone and even for online technical support. Time to market new drugs is reduced by clinical data integration. Through this software system errors are eradicated and novel therapies for marketing are generated. SAS organization has designed the software on the basis that each drug that is to be marketed has inimitable needs. Thus this software addresses these needs in the form a portfolio which drives effectualness throughout the various stages of the drug lifecycle starting from discovery development, commercialization and beyond.

Cognos 8 Business Intelligence software

Cognos is one of the world's largest providers of business presentation softwares and is used by 24 of the world's top 30 pharmaceutical firms. Client list of Cognos exceeds 23000 in the present day and a broad range of health care and life sciences companies are included. Earlier, systems like EDC and CTMS were appointed which acted as point solutions and presented restricted decision support. Data management has been made simpler by the use of Cognos by accurately determining the data quality, performance of the personnel and partners and many more. This software delivers a variety of Business Intelligence capabilities, also provide service-oriented architecture (SOA). The BI capabilities may include.

1. Reporting of multiple languages is done in multiple export formats such as Excel, Pdf, Xml, Html, and CSV.
2. Analysis Resource allocation is improved through performance data. Reporting and analysis is carried out on performance information. Summarizing and processing of data is done within the application layer and results are submitted at a summary level.
3. Dashboards and scoreboards: The performance of the monitor adjacent to key milestones for example IRBs (Institutional Review Boards), site initiation, investigator selection and their availability at regional, national, territory and product levels are determined using dashboards and scoreboards.

Symmetric software Symmetric software has been developed by symmetric life sciences organization which is a cost-effective system for the purpose of clinical trial data collection and management which fully abides by the international standards and has a proper client base all over the world.

Akaza's Open Clinica software

It offers an open informatics solution which keeps a record of the needs of the organizations that are involved in research. Open Clinica was released by Akaza in 2005 which has been adopted by 50 leading research organizations all over the world. OpenClinica is an effective, low cost and flexible way to manage clinical research institutes and medical centres. It is an opensource software thus, its source code is freely available www.openclinica.org. It is an opensource software podium for electronic data capture in the CDM system. It has a modular architecture and is a collaborative model which offers flexibility for obtaining the quality of data.

Sigmasoft's DMSYS software

Sigmasoft is a global provider of software for data management which is used for clinical data management. High quality data management along with services and support is provided by this organization. Sigma soft was founded in the year 1998, since then the company has developed comprehensible and reasonable data management softwares. The clinical data management services offered befits both small scale and large scale clinical trials.

Progeny Clinical software

The organization Progeny Software, LLC came into existence in 1996 and gave the prime version of the software at the ASHG (American Society for Human Genetics) conference. This organization has customers across 60 countries to this day. Progeny Clinical software was developed ideally for keeping the track of the family history data. Since 1996, the software has provided the research institutes for managing the patient history data. [15]

Testing new biomedical treatments for safety and efficacy will also require new strategies, since it has been shown that existing therapies often only work for a small number of indicated individuals. The application of emerging digital technologies, such as next generation sequencing, though, have increased both our understanding of disease mechanisms in larger pool of patients and the potential for developing personalized therapies. For example, the majority of the new molecular entities approved by the U.S. FDA in recent years were designed to target specific aberrations implicated in disease initiation and maintenance—a hallmark of precision medicine.[13]

Broadly academic research labs, biotechnology corporations, and technology companies have been exploring the use of AI and ML in three key areas:

1. Machine-based learning to predict pharmaceutical properties of molecular compounds and targets for drug discovery.
2. Using pattern recognition and segmentation techniques on medical images (from, e.g., retinal scans, pathology slides and body surfaces, bones and internal organs) to enable faster diagnoses and tracking of disease progression and generative algorithms for computational augmentation of existing clinical and imaging data sets.
3. Developing deep-learning techniques on multimodal data sources such as combining genomic and clinical data to detect new predictive models. [17]

Clinical trial end points and drug compliance can also be determined using AI-based technologies. Medical Education, Decision-Making Processing and organizing medical information is sometimes beyond human capabilities and may warrant the use of intelligent machines. AI can be used in knowledge acquisition and knowledge representation. This allows the development of reasoning tools to help with the decision-making process. Many cognitive apprenticeship models have been developed for this purpose. Projects such as RadTutor, InforMed, SAFARI, the cardiac tutor, and GUIDON are a few examples.

II. Conclusion:

Clinical drug development has remained relatively unchanged for the last 30 years. This is due, in part, to uncertainties in regulatory requirements, risk aversion, and skepticism about rapidly emerging, yet largely unproven, technologies (such as machine learning, and wireless health monitoring devices and sensors), and the lack of relevant actionable biomedical data sources and advanced analytics to generate hypotheses that could motivate the development of innovative diagnostics and therapies. It is also important to understand that ML is designed to make predications or decisions from data without explicit programming as which they make discissions tomorrow that are better than today and also the ability of ML is to learn from real world evidence(training) and improve its performance(adaption). We still need to encourage it as it may reduce the time gap between trial phase which is nearly 15 years from phase 0 to 4 trails for a drug development. The most common applications of artificial intelligence (AI) in drug treatment have to do with matching patients to their optimal drug or combination of drugs, predicting drug-target or drug-drug interactions, and optimizing treatment protocols.AI can also identify patients who might respond better to a certain medication and predict possible complications, thus optimizing patient selection, augmenting the speed of discovery, and reducing the cost burden.

References:

[1]. Naseem M, Akhund R, Arshad H, Ibrahim MT. Exploring the Potential of Artificial Intelligence and Machine Learning to Combat COVID-19 and Existing Opportunities for LMIC: A Scoping Review. J Prim Care Community Health. 2020 Jan-Dec; 11:2150132720963634. doi: 10.1177/2150132720963634. PMID: 32996368; PMCID: PMC7533955.

[2]. Raju B, Jumah F, Ashraf O, Narayan V, Gupta G, Sun H, Hilden P, Nanda A. Big data, machine learning, and artificial intelligence: a field guide for neurosurgeons. J Neurosurg. 2020 Oct 2:1-11. doi: 10.3171/2020.5. JNS201288. Epub ahead of print. PMID: 33007750.

[3]. Kubben P, Dumontier M, Dekker A, editors. Fundamentals of Clinical Data Science [Internet]. Cham (CH): Springer; 2019. PMID: 31314217.

[4]. Ratner M. FDA backs clinician-free AI imaging diagnostic tools. Nat Biotechnol. 2018 Aug 6;36(8):673-674. doi: 10.1038/nbt0818-673a. PMID: 30080822.

[5]. McFadden E, LoPresti F, Bailey L, Clarke E, Wilkins P. Approaches to data management. Control Clin Trials 1995; 16:30–62.

[6]. Paul J, Seib R, Prescott T. The Internet and clinical trials: background, online resources, examples and issues. J Med Internet Res. 2005 Mar 16;7(1):e5. doi: 10.2196/jmir.7.1. e5. PMID: 15829477; PMCID: PMC1550630.

[7]. Verma K. Base of a research: good clinical practice in clinical trials. J Clin Trials 2013;3(1):1–5.

[8]. Guidelines for Good Clinical Practice. International Conference on Harmonisation (ICH). http://www.ich.org. [Accessed 5 April 2017].

[9]. Krishnankutty B, Bellary S, Kumar NBR, Moodahadu LS. Data Management in clinical research: an overview. Indian J Pharmacol 2012;44(2):168–72.

[10]. Lu Z, Su J. Clinical Data Management: current status, challenges, and future directions from industry perspectives. Open Access J Clin Trials 2010; 2:93–104.

[11]. Lu Z. Information technology in pharmacovigilance: benefits, challenges, and future directions from industry perspectives. Drug Healthc Patient Saf 2009; 1:35–45.

[12]. Argraves S, Brandt CA, Money R, Nadkarni P. Informatics tools to improve clinical research study implementation. Contemp Clin Trials 2006;27(2):112–22.

[13]. Gerritsen MG, Sartorius OE, vd Veen FM, Meester GT. Data management in multi-center clinical trials and the role of a nation-wide computer network: a 5 year evaluation. Proc Annu Symp Comput Appl Med Care 1993:659–62.

[14]. Almenoff JS, Pattishall EN, Gibbs TG, DuMouchel W, Evans SJ, Yuen N. Novel statistical tools for monitoring the safety of marketed drugs. Clin Pharmacol Ther 2007;82(2):157–66.

[15]. Litchfield J, Freeman J, Schou H, Elsley M, Fuller R, Chubb B. Is the future for clinical trials internet-based? A cluster randomized clinical trial. Clin Trials 2005; 2(1):72–9.

[16]. Gao QB, Kong Y, Fu Z, Lu J, Wu C, Jin ZC, et al. EZ-entry: a clinical data management system. Comp Biol Med 2008;38(9):1042–4.

[17]. Shah, P., Kendall, F., Khozin, S. *et al.* Artificial intelligence and machine learning in clinical development: a translational perspective. npj Digit. Med. 2, 69 (2019). https://doi.org/10.1038/s41746-019-0148-3.