International Journal of Management, IT & Engineering
Vol. 11 Issue 12 December 2021,
ISSN: 2249-0558 Impact Factor: 7.119
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gate as well as in Cabell's Directories of Publishing Opportunities, U.S.A

BIG DATA ANALYTICS AND MACHINE LEARNING: PERSONALIZED, PREDICTIVE HEALTH AND BOOST EXACTITUDE MEDICINE RESEARCH

Mukesh Kumar Saini
Research Scholar OPJS University

Dr. Jaibir Singh
Asst. Professor

ABSTRACT

Healthcare data is enabling physician in building predictive models and better patient profiles for more effective anticipation, diagnosis and treatment of various diseases. In addition, partnerships and collaborations between healthcare communities and researchers have resulted into the development of data pools, which can be used for establishing better personalized healthcare models. The increased use of artificial intelligence and machine learning is shifting the paradigm of medical research and treatment.

These advanced technologies are providing researchers or medical practitioner real-time access to every white paper and clinical case study conducted on a genetic disorder and also help in detecting fraud in healthcare quickly and efficiently, with the use of statistical tools and algorithms that can help a faster development of more accurately targeted vaccines with a cost effective ways for discovering more clinically relevant ways to analyses the disease and treat the patients in an effective manner.

The objective of proposed approach to develop such mechanism to read patients clinical records, radiology data and provide the insight on the disease and allows researchers or medical practitioner to not only understand the full scope of a medical condition, but further shorten the amount of time it takes to develop a cure, treatment options, helping, and healing patients in need of healthcare.

KEYWORDS: Artificial Intelligence, Machine Learning, Health and Health Care, human performance, software engineering, big data

INTRODUCTION

Big data analytics helps the healthcare industry improve the accuracy of personalized medication and risk analysis. It also help standardize terminology across predictive analytics, medicine wastages and automated reporting for patient data. There is huge amount of data generated in health care industry and it will continue to increase. This includes electronics health records data, imaging data, patient generated data, sensor data, and other forms of data that is difficult to process.

Artificial Intelligence, Mobile, social and Internet of Things (IoT) are increasing data complexity and providing new forms and source of data. Big data analytics is used together with advance analytics methods on very large, diverse data sets that include structured, semi-structured and un-structured data, from different sources, and in different sizes from terabytes to zettabytes [1]. Analyzing these new forms of big data will enable analysts, researches, medical practitioner and business users to make more accurate and faster decision using data that was previously in accessible to them. By leveraging advanced analytics techniques such as text analytics, machine learning, predictive analytics, data mining, statistics and natural language processing businesses will accurately analyze previously untapped data sources gaining new insights and understandings, which will results in improved and more accurate decisions.

Big data analytics in healthcare is evolving into a promising field for proving insight from very large data sets for improving outcomes while reducing costs. As per McKinsey estimates that big data analytics can enable more than $300 billion in savings per year in U.S. healthcare, two thirds of that through reductions of approximately 8% in national healthcare expenditures[2], [7].

Big Data Analytics Advantages: Big data analytics has provided a way to healthcare organizations to develop actionable insights and decision support system, organize their future vision, boost up the outcomes and reduce time to value. This strategy is also helpful to provide meaningful information to the healthcare organizations regarding their management, planning and the measurements. The evaluated results can further help enhance the decision making capacity of the top management [2]-[3].

- Detecting diseases at earlier stages when patient can be treated more easily and effectively

- Managing specific individual and population health and detecting health care fraud a lot more quickly and efficiently.
- Numerous questions can be addressed with big data analytics
- Certain developments or outcomes may be predicted and/or estimated based on vast amounts of historical data, such as
  - Length of stay
  - Patients who will choose elective surgery
  - Patients who likely will not benefit from surgery
  - Complications
  - Patients at risk for medical complications
  - Patients at risk for sepsis, MRSA, C. difficile, or other
  - Hospital-acquired illness
  - Disease progression
  - Patients at risk for advancement in disease states;
  - Causal factors of illness/disease progression and possible comorbid conditions (EMC Consulting).

Big Data analytics and Machine Learning in healthcare can be used to raise the standards in following fields.

Public Health: By evaluating the disease patterns and recording disease occurrences, public health issues can be improved with analytics strategy. Enormous amount of data can help to determine the requirement, offer required services, predict and prevent the future crises to benefit the patients population [4]-[5].

Electronic Medical Record or EMR: An EMR contains the standard (structured and unstructured) medical data that can be evaluated with the data analytic approach to predict patients at risk and provide him effective care.

Patient Profile Analytics: Advanced and predictive analytics can be applied to patients' profile for identifying individuals who could benefit from proactive approach. This may include lifestyle changes.

Genomic Analytics: The data analytic approach can be effectively included in genomic analytics to make this approach a part of regular medical care decision process.

Fraud Analysis: Healthcare fraud is a serious financial ditch on the healthcare systems in many authorities. It characterizes a severe drain on the effectiveness of providing healthcare to those in need. Data analytics approach helps analyze greater range of claim requests to curtail down fraud cases, an effective analysis can help to reduce fraud, waste and abuse.

Safety Monitoring: Data analytics can be used to analyze real time large volumes of brisk data in hospitals. The strategy may help in the safety monitoring and negative event prediction.

MRI and PET (Positron emission tomography): it can provide assessment of both cerebral metabolism and amyloid. Computer procedures, algorithm and techniques are being adapted for medical purposes to recognize specific elements of images in order to offer early diagnose of disease or flag up the warning signs to physicians.

Video streaming Statistics: It is used for the purposes of understanding the consumption patterns (behavioral and optimizing viewing experience).

Patient care in radiology: Big data in radiology is more about decision support than anything else and plays an important role in defining the way radiologists use clinical decision support systems to assist them in reading images.

Virtual care and wearable health care technologies. Helping providers drive virtual care initiatives to increase quality of care and provide patients with more access, but there is the question, how secure is the ecosystem in which more and more personal health information is being exposed?

Project Data Sphere. SAS is playing a important role in the development of industry wide pharmaceutical data transparency. The objective is a secure, globally reachable data and analysis environment where multiple enterprises can share anonymized clinical trial information. This would help scientists learn from experiments and research more quickly, and thus speed improvements in healthcare.

Disease progression models- Most recently disease advancement was studied on the organ and system level. Using vital organ, labs and invasive procedures, the patient disease state is determined as the function of the organ or the system.

BIG DATA ANALYTICS AND MACHINE LEARNING: DISCOVERY FOR PERSONALIZED HEALTHCARE

Big data technologies make it possible in a short time to analyze a large collection of data from millions of patients, identify clusters and correlations, and develop predictive models using statistical or machine learning modeling techniques. A very high throughput technologies make viable for exploration is the normalization of digital medical images to conventional space-time reference systems, using elastic registration methods, followed by the treatment of the quantities expressed by each three dimensional value in the image as self-determining data quanta. The voxel values of the scan then become another medical dataset, potentially to be correlated with average blood pressure, body weight, age, or any other clinical information.

By using statistical modeling and machine learning techniques, we could obtain a good predictors, valid for the range of the datasets examined; if a database contains outcome observables for a subset of patients, we will be able to compute automatically the correctness of such a predictor. Normally the result of this process would be a prospective clinical tool with known correctness and in some cases the result would provide a predictive accuracy sufficient for clinical purposes, in others a higher accuracy might be anticipated. In some cases there is need for an explanatory theory, which answers the “how” question, and which may be used in a wider context than that a statistical model normally is. As an alternative step, one could use the correlation linked by the empirical modeling to unfold possible mechanistic theories. Given that the available mechanistic knowledge is relatively deficient, in numerous cases we will be suitable to express a mathematical model only for a part of the process to be modeled; various “grey-box” modeling methods have been developed in the last few years that allow one to combine partial mechanistic knowledge with phenomenological modeling [14].

The final step is where biomechanics physiology, biochemistry, and biophysics mechanistic models are used. These models contain a large amount of validated knowledge, and require only a relatively small amount of patient- pacific data to be properly identified. In numerous cases these automatous models are exceptionally expensive in terms of computational cost; so input-output sets of these models may also be stored in a data repository in order to recognize reduced order models (also referred as “surrogate” models and “meta-models”) that accurately replace a computationally expensive model with a cheaper/faster simulation. Experimental design approaches are used to choose the input and output parameters or variables with which to run the mechanistic model in order to produce the meta-model’s state space description of the input-output relations which is frequently replaced with a piecewise partial least-squares regression approximation. Another approach is to use Nonlinear Auto-Regressive Moving Average model with eXogenous inputs in the framework of nonlinear systems identification.

It is fascinating to note that no real model is ever completely white-box. In all scenarios, some phenomenological modeling is required to define the interaction of the portion of reality under exploration with the rest of the universe. If we accept that a model describes a process at a positive characteristic space-time scale, everything that happens at any scale smaller or larger than that must also be accounted for phenomenologically. Therefore, it is possible to imagine a difficult process being modeled as an instrumentation of sub-models, each predicting a part of the process (for instance at different scales), and we can anticipate that, while initially all sub-models will be phenomenological, more and more will progressively include some automatous knowledge.The awareness of a progressive increase of the descriptive content of a hyper model is not fundamentally new and other domains of science already pursued the strategy described here.

PREDICTIVE ANALYTICS AND BIG DATA FOR PERSONALIZED HEALTHCARE

Health care is the act of taking preventative and necessary medical procedures for maintaining and restoration of patients health and prevention of diseases. Predictive Analytics in health care combines technology and statistical methods to search through huge amount of information, and analyze it to predict individual patient’s results. It can include historical treatment as well as the latest medical research and experimental data.

By using predictive analytics enhance the results for both the patients and the businesses participating in the rigid health care market. The below mentioned summarizes indicate how predictive analytics is aiding different segments of the health care domain.

- *Increase the accuracy of diagnoses:* Medical practitioner can use predictive algorithms to assist them more accurate diagnoses as an example, once patients come to the ER with chest pain, it's typically difficult to know whether or not the patient should be hospitalized. If the doctors were able to answers questions on the patient and his condition into a system with a tested and accurate predictive algorithm that would assess the chance that the patient may be sent home safely, then their own clinical judgments would be power-assisted. The prediction wouldn't replace their judgments however rather would assist.
- *Preventive medicine and public health:* Healthcare facilities can take a more pre-emptive approach to treatment. For instance, by exactly predicting which patients will develop chronic situations, or which ones will respond best to certain types of medications, health care enterprises can emphasis not only on treating present conditions, also on preventing recurrences.

- *Insight to physicians for individual patients:* Evidence-based medicine (EBM) is a step within the right direction and provides more help than straightforward hunches for physicians. However, what works best for the middle of a traditional distribution of individuals might not work best for an individual patient seeking treatment. Predictive Analytics can facilitate doctors decide the precise treatments for those individuals. It's wasteful and potentially dangerous to give treatments that are not required or that won't work specifically for an individual. Improved diagnoses and more targeted treatment procedure will naturally lead to increases in good outcomes and fewer resources used, including the physician's time.

- *Optimized Healthcare Costs:* Analytics models built to detect the risk of patients based on their age, lineage, diet habits, previous medical history and which diseases they are most vulnerable to and yields right protection, prevention treatment plans for individual patients populations, optimizing the health care and insurance costs.

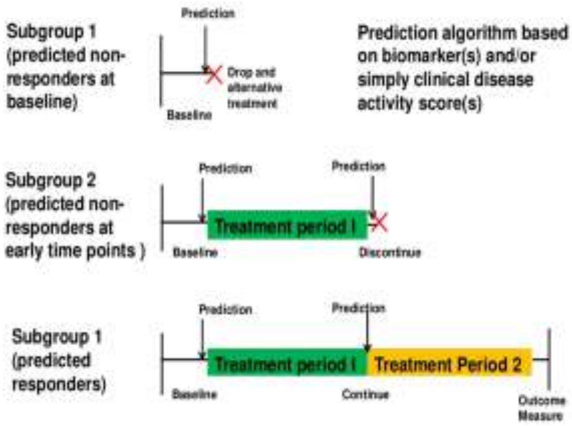
- *Improve patient outcomes and lower costs*: The use of sophisticated mathematical models, algorithms, powerful servers, storage, networking, and optimized analytics code, the solution gives healthcare organization the ability to analyze extensive amounts of structured Or Un structured data in a scalable and cost-effective way. Predictive analytics built on this foundation can enable data-driven insights across a variety of clinical, operational, and financial use cases, leading to improve access to higher quality care at lower cost.

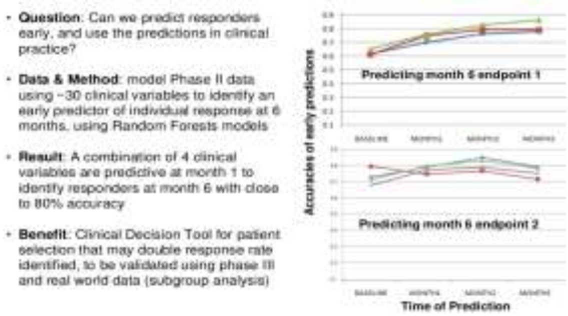
- Predictive analytics Potential benefit of better outcomes to patients: Health care organizations can obtain clinical facts from huge amounts of patient information to understand patient histories and identify future outcomes. By carefully analyzing which treatments and procedures work best, providers can make intellectual decisions about treatment plans, minimizing obstacles and patient readmissions

Approach and vision of big data and predicative analytics for personalized healthcare: Personalized Medicine based on the recognition that unprecedented types of information will be obtainable from genetic, genomic, proteomic, imaging technologies , etc. which help further known diseases into new categories and managing a patient's health based on the individual patient's specific characteristics.

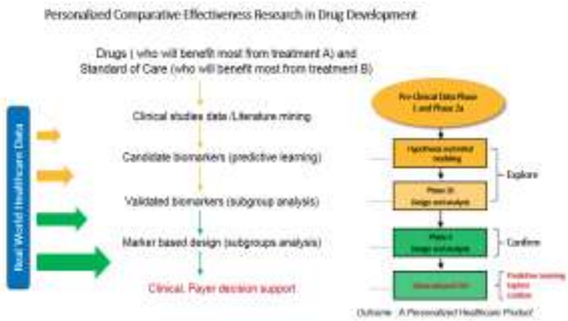
Comparative effectiveness research, an analysis of comparative effectiveness is simple a rigorous evaluation of the impact of different treatment options that are available for treating a given medical condition [16]-[17].

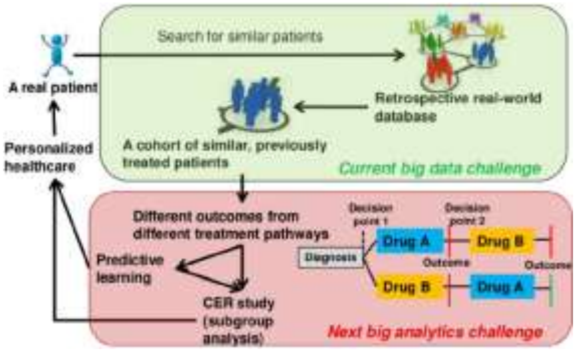
How's predictive learning identify responders early in treatment course? Predictive Learning Analytics is unlike from other metrics because it concentrations on the individual learner, instead of the learning program as a whole. This makes Predictive Analytics uniquely helpful in tackling the problem of ineffective learning. Predictive Learning Analytics allows you to conclude who did and did not learn the material, and who is most, or even least, possible to apply the things they learned to their jobs.





Personalized Comparative effectiveness research, will answer to what patient subgroup, what disease stage, what treatment pathway, and where in the treatment pathway, a comparative effectiveness evidence is applicable.





BIG DATA ANALYTICS: WHAT CHALLENGES CAN BE ADDRESS BEFORE BIG DATA TECHNOLOGIES CAN BE EFFECTIVELY AND EXTENSIVELY USED IN HEALTHCARE?

There is a lot of potential in delivering more targeted and actionable, wide-reaching, and cost-efficient healthcare by exploiting current big data trends and technologies. Though, it has also been shown that the healthcare industry has some very specific characteristics and challenges that needs a targeted effort and research in order to realize the full potential [16]-[21]:

- Big data access, availability and quality: there is a huge amount of existing data distributed in several repositories and new data generated daily by billions of connected devices or self-generated by people. It is then necessary to find more appropriate and effective ways to leverage these data in line with privacy and ethical principles, to access them, to understand the purposes for their acceptance and their quality in order to enhance and optimize care processes, diseases diagnosis, and personalized care in general the healthcare system.
- Patients and healthcare professionals profiting from big data: here is a need to develop strategy that allow for humans and machines to collaborate more closely on exploiting big data for a better health. This includes assurances on the reliability of information, a focus on generating actionable advice, and improving the interactivity and understandability of big data processing and analytics. The necessities of different target groups, researchers, healthcare providers, care-givers, patients and general population may need different focus.
- Multi-modal data analytics: there is a need for technologies, which can handle, analyses and exploit the set of very diverse, interlinked and complex data that

already exists in the healthcare world to improve the healthcare quality and decrease costs.

- Healthcare knowledge: Next to the big and dissimilar healthcare data sets, there is already a big amount of medical and healthcare knowledge. This knowledge exists in research papers and widely available in internet, but also in the heads of healthcare providers. In fields of epidemiology, wearable sensors, a completely different knowledge on the real world, organizations and how people live their lives is very valuable to understand patients and the healthcare system in general. New strategies are required that bring together big data and knowledge, such that knowledge can be used to make better sense of data, and data can be used to generate more meaningful learning.

Integrity and privacy in a big data world: Additional practical approaches are needed to effectively balance the benefit and threats of more and more detailed and complex data being available. With the respect to an increasing amount of complexity and automation in clinical trial data processing and decision support, in particular in the light of the move towards personal health assistant on mobile devices, a targeted focus on the ethical difficulties connected with these new technologies seems advisable.

CONCLUSION

The most important challenges in clinical practice and medicine research include the need to develop and apply novel tools for the effective integration, analysis and interpretation of complex medical data with the aim to identify testable hypothesis, and build accurate models.

Big data Analytics gives a great boost to leverage the benefits of chaotic environment in healthcare, latest computing and behavioral research can lead to transformative changes in delivering high quality and personalized health care, diagnosing the acute diseases and predicting their re-occurrence is vital in health care industry and by identifying the recurrence of diseases in advance and providing health advice.

The use of sophisticated mathematical models, algorithms, powerful servers, storage, networking, and optimized analytics code, the solution gives healthcare organization the ability to analyze extensive amounts of structured Or Un structured data in a scalable and cost-effective way. Predictive analytics built on this foundation can enable data-driven insights across a variety of clinical, operational, and financial use cases, leading to improve access to higher quality care at lower cost. There is undoubtedly a significant research space examining the ultimate methods and technologies for big data analytics, it is vital to recognize that it is also necessary to fund domain based targeted research and experiments that allows dedicated solutions to be developed for precise applications.

REFERENCES

[1].Song TM. Efficient utilization of big data on health and welfare. Health Welf Policy Forum. 2012;(193):68–76.

[2].O. Terzo, P. Ruiu, E. Bucci, and F. Xhafa, "Data as a service (DaaS) for sharing and

processing of large data collections in the cloud," in Proc. Int.
Conf. Complex Intell. Softw. Intensive Syst., 2013, pp. 475–480. https://docplayer.net

[3].B. Wixom, T. Ariyachandra, D. Douglas, M. Goul, B. Gupta, L. Iyer, U. Kulkarni, J. G. Mooney, G. Phillips-Wren, and O. Turetken, "The current state of business intelligence in academia: The arrival of big data," Commun. Assoc. Inform. Syst., vol. 34, no. 1, p. 1, 2014.

[4].Practical Predictive Analytics and Decisioning Systems for Medicine Linda Miner Pat BoldingJoseph Hilbe Mitchell GoldsteinThomas Hill Robert Nisbet Nephi WaltonGary Miner

[5].[Predictive Analytics for Personalized Health Care Zhaohui (John) Cai, MD, PhD Biomedical Informatics Director, AstraZeneca Big Data and Analytics for Pharma Philadelphia, PA June 12, 2013]
https://ieeexplore.ieee.org/Xplore/home.jsp

[6].[https://www.netsec.news/definition-hipaa-covered-entity/][Ohm, Paul (August 2010). "Broken Promises of Privacy:
Responding to the Surprising Failure of Anonymization". UCLA Law Review. 57 (6): 1701–1777.]

[7].https://www.intel.com/content/dam/www/public/us/en/documents/brief/hls-predictive-analytics-business-brief.pdf]

[8].[Personalized Medicine Coalition, Washington, DC, USA - Edward Abrahams, 1 Geoffrey S. Ginsburg2 and Mike Silver3]
https://www.researchgate.net/
www.ijitjournal.org

[9].W. B. Nelson, Accelerated Testing: Statistical Models, Test Plans, and Data Analysis. New York, NY, USA: Wiley, 2009. [40] C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY, USA: Springer, 2006.

[10].J. Ide, R. Chen, D. Shen, and E. H. Herskovits, "Robust brain registration using adaptive probabilistic atlas," Med. Image Comput. Comput. Assist. Interv., vol. 11, no. Pt 2, pp. 1041–1049, 2008. https://docksci.com/

[11].M. Rusu, B. N. Bloch, C. C. Jaffe, N.M. Rofsky, E.M. Genega, E. Feleppa, R. E. Lenkinski, and A. Madabhushi, "Statistical 3D prostate imaging atlas construction via anatomically constrained registration," Proc. SPIE, vol. 8669, 2013. www.forbes.com

[12].S. A. Billings, Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains. New York, NY, USA: Wiley, 2013.

[13].Kayyali B, Knott D, Van Kuiken S. [2016-11-11]. The big-data revolution in US health care: accelerating value and innovation. McKinsey & Company. 2013

[14].Youssef AE. A framework for secure healthcare systems based on big data analytics in mobile cloud computing environments.Int J Ambient Syst Appl. 2014

[15].Schilsky RL, Michels DL, Kearbey AH, Yu PP, Hudis CA. Building a rapid learning health care system for oncology: the regulatory framework of CancerLinQ. J Clin Oncol. 2014 Aug

[16].US Department of Health and Human Services Genomics Data Sharing. 2014. https://gds.nih.gov/17summary_dbGaP_statistics.html webcite. www.slideshare.net www.ijese.org

[17].Feldman B, Martin E, Skotnes T. [2016-11-09]. Big data in healthcare hype and hope. GHDonline. 2012 Oct. https://www.ghdonline.org/uploads/big-data-in-healthcare_B_Kaplan_2012.pdf webcite.

[18].Chawla NV, Davis DA. Bringing big data to personalized healthcare: a patient-centered framework. J Gen Intern Med. 2013 Sep;28(Suppl 3):S660–5. doi: 10.1007/s11606-013-2455-8.

[19].US Department of Health and Human Services [2016-11-09]. State Health Information Exchange Cooperative Agreement Program. US Printing Office. HealthIT. 2011. https://www.healthit.gov/policy-researchers-implementers/state-health-informationexchange webcite.

[20].Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA. 2013 Apr 3;309(13):1351–2. doi:10.1001/jama.2013.393. [PubMed] [Cross Ref] www.builtinla.com

[21].Shu-Hsien Liao, Pei-Hui Chu, Pei-Yuan Hsiao, "Data mining techniques and applications - A decade review from 2000 to 2011", Expert Systems with Applications (2012), 11303-11311, Elsevier.