International Journal of Advanced Science and Technology
Vol. 29, No. 08, (2020), pp. 2076- 2083

An Effective Analysis of Personalized Medicine using Convolutional Neural Networks

Anbarasi R [1]
anbukavi1011@gmail.com

Dr. Krishnaveni R[2]
krishnaveni@hindustanuniv.ac.in

Department of Computer Science Engineering
Hindustan Institute of Technology and Science
Chennai, Tamil Nadu

***Abstract* — The high throughput technology causes the advanced a vast development in biological technology as a holistic science to achieve more precise modeling of complex diseases. It provides sufficient support in the medical field. Nowadays, people moving from two-tiered health systems to two-tiered personalized medicine. In the proposed model, The implemented the biomedical big data in personalized medicine. This review provides an update of important developments in the analysis of big data and facilitating the design of clinical procedures in Personalized Medicine.The dataset contains helpful information which are required for the further research improvement. The final product gives a decent expectation results for additional improvement about the quality. It contains results about the qualities, variety, definite content and forecast score about the dataset.**

Index: Big data, Personalized Medicine

Introduction

The development of rich context, diversity of the data being generated in health care is driven in the development of big data in health care. The volume, varieties of the data generates numerous challenges for the researchers. Tackling these Challenges would pave the way for more intelligent healthcare systems which help to provide early detection and personalized treatments. To identify the risk factor and diagnosing the medical result which helps to diagnose illnesses and conditions in individual patients. Also, data derived from electronic health records (EHRs), social media, the web, and other sources provides healthcare organizations and government agencies with up-to-the-minute information on infectious disease threats or outbreaks. The electronic patient health record (EHR) is consist of socio-demo-graphics, medical conditions,

ISSN: 2005-4238 IJAST

genetics, and treatments of the patients. A standard pattern is required for interpretation of results which support to organize the datain efficient manner and achieve the set goal based on the threshold value. Using the genome sequences and the extraction of the accurate clinical records from the digitalized clinical record which efficiently embrace the personalized medicines. The result shows a higher throughput value and increases the quality and quantity of the prescribed medicines. The outstanding improvement in automated collection of massive data volumes is exemplified by community movements such as the Global Alliance for Genomics and Health (GA4GH, www.ga4gh.org), research infrastructures like ELIXIR [2] and Big Data to Knowledge (BD2K) [3], and international initiatives such as the International Cancer Genome Consortium (ICGC), the International Human Epigenome Consortium (IHEC), and the International Rare Disease Consortium (IRDiRC), among others. Patients will see a shift from population-based healthcare to personalized medicine that includes targeted diagnostics and treatment based on each patient's history, ancestry and genetic profile. By applying big data analytics in personalized medicine improve the overall quality and efficiency of treating disease. It creates a premise in the personalized medicine programs that will significantly improve patient care. The transfer of these medical big data through the internet may affect with the insecurity. To generate, maintain, transfer and analyze large-scale information securely in biomedicine and to integrate data with other data sets, such as clinical data from patients is a tedious process. But due to advanced technologies, storing and analyzing the data will be less expensive. Transferring the data from one place to another place similar to mailing or transfer files using external drives, the security, and privacy of the data from individuals are also a concern before and during data transfer. Possible solutions to these issues include the use of better security systems with advanced encryption and de-identification algorithms, such as those used by banks in the financial sector to secure their clients' privacy. To accumulate all the life science data, it is important to improvise the computational infrastructure. By collaborating the existing and proposed work the personalized medicine and represent the new and unique form of data.

Architecture diagram

The architecture diagram consists of the following major faces such Patients, Omics profile, data integration, analysis & data interpretation, Patient specific interaction and clinical care innovation.

ISSN: 2005-4238 IJAST

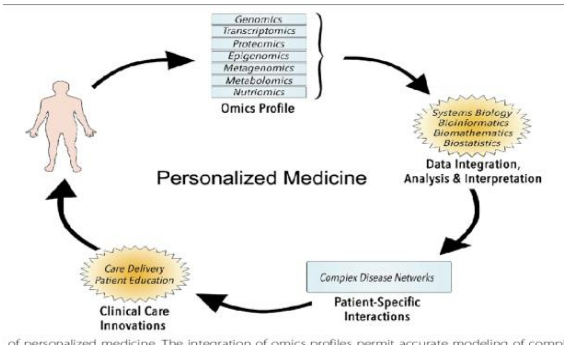


Fig 1.Personalized medicine

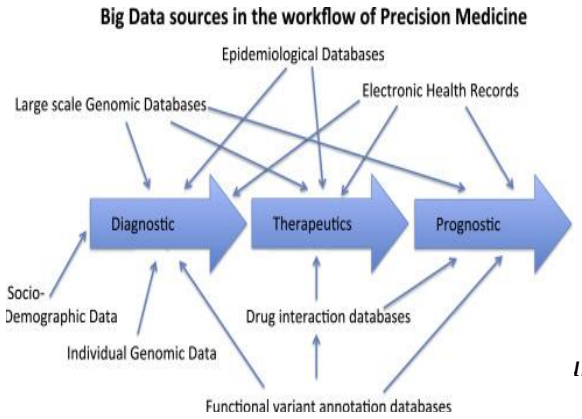
Various patients' details are collected based on various biological factors which is processed and carries out to the next section which is so called the omics profile face. The omics profile consist of different section such as the genomics, proteomics, transcriptomics, metabolomics, epigenomics, metagenomics and the nuticmics. Once the biological data are collected and categorized as per the Omics profile it is carried out to the data integration section where the scatted biological data of the patients are collected, combined, analyzed together to produce a useful report. Once the three faces are achieved the data is processed to the fourth face called the patient specific interaction where the complex disease network obtained. Once all the faces are successfully carried out the patients are provide with the care delivery and patient's education tips.

Datatype

The medical information is available in the digital format such as structured and unstructured content of EHR. The symptoms description is aligned in the unstructured content whereas ICD codes are represented in structure content. The medical big data is to act as valuable resources for researchers and model development. The experimental platforms such as proteome profiling (mass spectrometer), interactome profiling, transcriptome sequencing (RNA-seq and ribosome profile) are making the biomedical information is accessible and faster[6][7]. With access to genotypic and phenotypic data on tens of millions of patients, researchers could transform the face of healthcare. The ability to survey the entire genome enables scientists to identify novel pathways involved in the disease and develop therapies that are more likely to be relevant to humans. But because these variants are often rare, or have small effect sizes, large datasets are required to make valid inferences about the role of these variants in disease. As such, it is critical

ISSN: 2005-4238 IJAST

to pool large datasets. The workflow of the big data in precision medicine is illustrated in Fig1. The data are classified under three different categories such as Diagnostic, Therapeutics and prognostic.



ırces of precision medicine

The wearable devices and implanted devices provide efficient health data generated by the devices. It leads to an increase in the relevant big data type in personalized medicine [12] and illustrated in Fig 1. To maintain and improve the data quality, the mobile app data of the patient is tracked and thus enhance the health care services. The biometric measurement of the patient is monitored in real-time and updated in the cloud. It prompts the development of new areas such as stream computing and analyzing real-time flows of data. To enhance the precision medicine research it utilized the cloud-based platform for a store and efficiently retrieves the data.

Data Management

The big data is applied in a distinct area such as basic research in cancer, neurodegeneration, and diabetes and cardiovascular pathologies. It is necessary to develop research and data on those areas which highly appreciable by the social and government agencies which are the basis of all large scale biomedical projects. More than 2.6 billion euros is invested in personalized medicine research by European commission through FP7 and Horizon 2020 programs [13]. This biomedical big data constituent of highly distributed acquisition, format heterogeneity, and content sensitivity. Based on blockchain technology [11] patient details are collected using smart card technology but still require investigation. The utilization of the medical research relay on the effective framework and it has to state the variation among the disease and stages. The framework must be in A large amount of precision medicine research relies on the use of

ISSN: 2005-4238 IJAST

predictive frameworks to better understand disease states and how they vary at the patient level. By integrating all the workflows and databases in a single digital ecosystem, it would be easy for researchers to apply the appropriate analytics that will generate the complex predictive models necessary to inform personalized treatment choices. This personalized treatment choice led to some remarkable developments in healthcare technology as well as treatments that have improved patient outcomes. However, the full potential of precision medicine is yet to be realized, and future progress will be accelerated through the more effective harnessing of genotypic and phenotypic data. This type of data is shared globally develop insights into genetic factors that contribute to disease development, as well as large-scale population sequencing projects will enhance a broader understanding of genetic variation. With access to genotypic and phenotypic data on tens of millions of patients, researchers could transform the face of healthcare. Each sensor's data from the patient is stored in the hub and integrated with molecular profiling data and medical records. Based on the data, predictive modeling is generated which evaluate a patient's risk of developing a disease and to identify appropriate preventative measures or plan individualized treatments.Machine learning methods can be effectively applied to deliver integrative solutions for multi-view data to explain an event or predict an outcome. For instance, generalized linear models (GLMs) comply with a broad model formulation where the outcome is linearly related to factors and covariates through a link function that permits the estimation of the model parameters typically by maximum likelihood or Bayesian techniques.

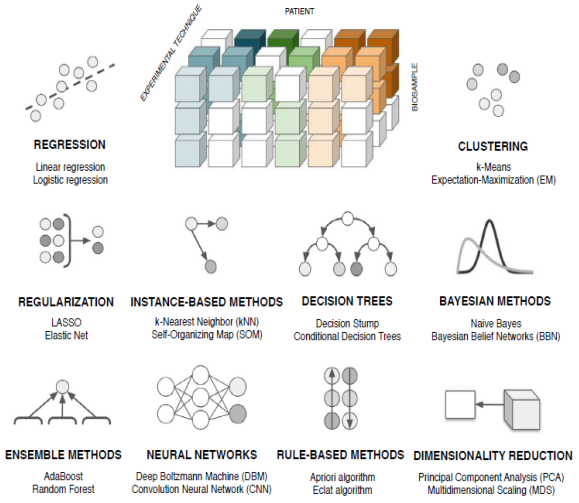


Fig 3.Machine learning methods

ISSN: 2005-4238 IJAST

As features of biological data often exhibit some form of structure, such as groups of genes with similar functions, structural regularization methods such as Sparse Group LASSO (Least Absolute Shrinkage and Selection Operator) is a common approach for supervised multi-view feature selection. Along with GLMs, common machine learning models for multi-view data are Bayesian models such as the naive Bayes classifier, ensemble-learning models such as random forest, neural networks and more recently deep learning in Fig 3. Machine learning algorithms for multi-view data analysis. Bio samples from several experimental techniques (e.g. genomic, proteomic, and metabolomics data) can be used to identify associations within and between multiple sets of patients, and generate integrative models for patient stratification.

Result and Discussion

The dataset contains useful data which are required for the further research development. The end result gives a good prediction results for further development about the gene. It contains results about the genes, variation, detailed text and prediction score about the dataset.

Prediction Accuracy : 75.0

ID	Gene	Variation	Text	Class[Label]	Predicted
1	PTPRT	A1022E	most cancers characterized increased stat activation phosphorylated stat levels associated reduced survival molecular mechanisms underlying aberrant stat phosphorylationactivation human malignancies been elusive findings provide mechanistic basis tumorspecific stat hyperactivation head neck squamous cell carcinoma hnscc demonstrate receptorlike protein tyrosine phosphatases	4.0	4
2	PTPRT	Y412F	the receptor protein tyrosine phosphatase ptprtptp the frequently mutated tyrosine phosphatase human cancer ptp mediates homophilic cellcell aggregation its extracellular region	4.0	4
3	DICER1	Deletion	mesenchymal cell populations contribute microenvironments regulating stem cells the growth malignant cells osteolineage cells participate the hematopoietic stem cell niche	1.0	1
4	CBL	C396R	oncogenic mutations the monomeric casitas bilineage lymphoma cbl gene been found many tumors	4.0	4

Sample complex disease network report

Conclusion

Recent developments in many bioinformatics topics, such as cancer diagnosis, precision medicine, and health-informatics systems have a keen need for integrative machine learning models to incorporate all available data. In this paper, we analyze different machine learning concepts. These methods, particularly multi-view matrix factorizations and multi-modal deep learning, will revolutionize the way of using information and play a key role in integrative bioinformatics. At present, a large number of multi-omics, imaging, medical devices, and EHR data are available from large-scale cohort and population studies, revealing subtle differences in

ISSN: 2005-4238 IJAST

human genetics and allowing Personalized Medicine interventions, while engaging infrastructural and research management innovation and sustainability. Challenges in big data analytics are pointing to the development of effective applications in areas where finding connections and insights can be difficult due to data abundance and the complexity of biological systems. In the future, deep learning and cognitive computing play a vital role in the data-driven analysis. Thus, this improvement causes great innovation in health care and personalized medicines.

REFERENCES

[1]. Tsiouris KM, Pezoulas VC, Zervakis M, Konitsiotis S, Koutsouris DD, Fotiadis DI: A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. Comput Biol Med 2018, 99:24-37.

[2]. Bakkar N, Kovalik T, Lorenzini I, Spangler S, Lacoste A, Sponaugle K, Ferrante P, Argentinis E, Sattler R, Bowser R: Artificial intelligence in neurodegenerative disease research: use of IBM Watson to identify additional RNA-binding proteins altered in amyotrophic lateral sclerosis. Acta Neuropathol 2018, 135:227-247.

[3]. Hernandez D, Greenwald T: IBM has a Watson dilemma. Wall Street J 2018. Retrieved from http://online.wsj.com.

[4]. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S: Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017, 542:115-118.

[5]. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, Liu PJ, Liu X, Marcus J, Sun M et al.: Scalable and accurate deep learning with electronic health records. NPJ Digit Med 2018, 1:1609.

[6]. Patel NM, Michelini VV, Snell JM, Balu S, Hoyle AP, Parker JS, Hayward MC, Eberhard DA, Salazar AH, McNeillie P et al.: Enhancing next-generation sequencing-guided cancer care through cognitive computing. Oncologist 2018, 23:179-185.

[7]. Zhou J, Theesfeld CL, Yao K, Chen KM, Wong AK, Troyanskaya OG: Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. Nat Genet 2018, 50:1171-1179.

ISSN: 2005-4238 IJAST

[8]. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow P-M, Zietz M, Hoffman MM et al.: Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 2018, 15.

[9]. Ratner A, Bach SH, Ehrenberg HR, Fries JA, Wu S, Re´ C: Snorkel: rapid training data creation with weak supervision. Proc VLDB Endowment 2017, 11:269-282.

[10]. Gupta S, Kar AK, Baabdullah A, Al-Khowaiter WAA: Big data with cognitive computing: a review for the future. Int J Inf Manage 2018, 42:78-89.

[11]. Kiyomoto S, Rahman MS, Basu A: On blockchain-based anonymized dataset distribution platform. 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA) 2017:85-92.

[12]. Shen D, Wu G, H-I Suk: Deep learning in medical image analysis. Annu Rev Biomed Eng 2017, 19:221-248.

[13]. Nimmesgern E, Norstedt I, Draghia-Akli R: Enabling personalized medicine in Europe by the European commission's funding activities. Pers Med 2017, 14:355-365.

[14]. Barcelona Supercomputing Center (BSC), C/Jordi Girona 29, 08034, Barcelona, Spain. Electronic address: davide.cirillo@bsc.es.ICREA, Pg. Lluís Company's 23, 08010, Barcelona, Spain

[15]. www.aetna.com

[16]. Davide cirillo,Alfonso Valencia:Big data analytics for personalized medicine.current opinion in biotechnology 58,161-167,2019

[17]. www.frontiersin.org

[18]. www.reuters.com

[19]. www.delltechnologies.com

[20]. www.easi.com

ISSN: 2005-4238 IJAST