RESEARCH ARTICLE OPEN ACCESS

Role of Cloud Computing in Bioinformatics

[1]Sunny Sharma, [2]Karandeep Kaur, [3]Amritpal Singh
[1, 2, 3] Department of Computer Science, Guru Nanak Dev University, Amritsar, Punjab

----------**********----------

Abstract:

The evolution to network and computational paradigm has gone through a amazing phase of expansion and development. The growth curve was indeed very steep in many major domains. The advent of Cloud computing & Machine learning has enhanced the implementation in application area like Bioinformatics. With huge application-domain scope Cloud computing has emerged as a special area of interest for many bioinformatics researchers. Research is being done on different aspects of Cloud computing with bioinformatics for identifying areas of improvement and their respective remedies for living beings. Specially the cloud computing are acting very helpful for identifying H1N1 virus in human. H1N1 is an infectious virus which, when spread affects a large volume of the population. It spreads very easily and has a high death rate. Similarly cloud computing doing good job for detection of Hypertension, Diabetics, Cancer and Heart patient with software as a service, so the development of healthcare support systems using cloud computing is emerging as an effective solution with the benefits of better quality of service, reduced costs. This paper, provide an effective review towards cloud computing important effort in a field of bioinformatics.

Keywords — **Cloud computing, Bioinformatics, Data Analytics, Machine Learning**

----------**********----------

I. INTRODUCTION

Cloud computing (CC) is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources e.g., networks, servers, storage, applications, and services that can be rapidly provisioned and released with minimal management effort or service provider interaction. In CC various service providers (SPs) namely; Amazon, Google, Salesforce etc. provides different kind of services across the world, even different companies (Microsoft, Google, IBM 2, Yahoo) in different locations are deploying very high computational data centers (DCs). Recent DCs have high end servers for hosting applications [1]. Cloud computing for healthcare support services makes possible many extra benefits like ease in swapping records between hospitals, in management of centralized medical records. It not only relives user from keeping whole records side by side for doctors to take important decision by looking previous detail of the patient. So we can say cloud computing helps in accessibility, availability, scalability, cost effectiveness and storage capability are some of the major benefits. These are the benefits which encourage any healthcare departments to shift towards medical-health clouds system.

II. INTRODUCTION TO BIOINFORMATICS

In 1970 P. Hogeweg and B. Hesper innovate the term bioinformatics which to alludes to the investigation of data handling or information processing in biotic frameworks. Bioinformatics is conceptualizing biology regarding atoms or molecules (in the feeling of Physical science) and applying "informatics methods" got from computer science mathematics, and statistics) to comprehend and sort out the information connected with these molecules, on a substantial scale.

The essential objective of bioinformatics is to build

the comprehension of natural biological procedures. It is the exploration of merging subatomic science or molecular biology with computer technology. Bioinformatics is the mixture of different fields like computer science, Biology, Biochemistry, statistics, arithmetic and considerably more. To concentrate on biological data distinctive ideas of computer science, statistics, biology and designing are joined together.

Biological data + Computer Calculations= Bioinformatics

Bioinformatics is new science to oversee of natural data. To assemble, store, break down, analyze, incorporate natural biological and hereditary genetics data, computer science concepts are utilized. Figure 1 describes the integration of Bioinformatics with multidisciplinary fields.

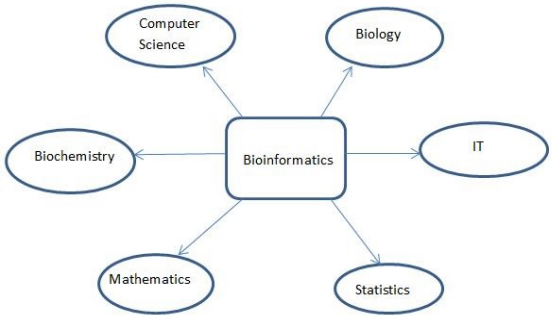

Figure 1: Bio-informatics multidisciplinary fields

Aims of bioinformatics are:

- To arrange information that permits scientists to effortlessly make and get to data
- To create tools that encourages the investigation and administration of information.
- To utilize biological information to investigate and decipher the outcomes in a biologically important way.

Bioinformatics incorporates arithmetic, measurements and Computer science innovation to take care of complex biological issues. These atomic issues can't be settled by some other fields. A few utilizations of bioinformatics are Sequence Analysis, Genome Annotation, Comparative Genomics, Wellbeing and medication disclosure, drug disclosure. Area of Computer Science assumes an imperative part in bioinformatics. The blend of Biological information and Computer Calculations is Bioinformatics. To extricate the information encoded in biological data progressed computational innovations, calculations and tools required. Essential issues in bioinformatics like protein structure forecast, numerous arrangements of sequences, etc. are inalienably non deterministic polynomial time. To tackle these sorts of issues computerized reasoning especially Artificial Intelligence (AI) methodologies are utilized, Researchers have utilized AI systems like Artificial Neural Networks (ANN), Fuzzy Logic, Genetic Algorithms, and Support Vector Machines (SVM) to take care of issues in bioinformatics. Manufactured Neural Networks is one of the AI procedures generally being used in view of its capacity to catch and speak to complex info and yield connections among information.

III. DATA MINING ROLE IN BIOINFORMATICS

Extracting or "mining" information from a lot of information is alluded to as Data mining. Data Mining (DM) is the investigation of discovering new intriguing examples and relationship in tremendous measure of data. Data mining is additionally called Knowledge Discovery in Databases (KDD). It is characterized as "the procedure of finding important new connections, designs from a lot of information put away in Warehouses". In bioinformatics mining natural information removes helpful learning from monstrous datasets assembled in biology, and in different zones related life sciences. The importent task of data mining is the essential objective of information mining is separating important new examples from information or data. The distinctive errands performed by information mining are Classification, Estimation, Prediction, Association guidelines, Clustering, Description & visualization

IV. SOFTWARE AND TOOLS

Some simple programming software tools utilized for bioinformatics is command line as well as GUI. Few open-source programming bundles ranges

from BioPerl, Bioconductor, Biopython, BioRuby, BioJava, Bi-oclipse, .NET-Bio, UGENE & EMBOSS etc. Few present in market. there are few web administrations are present in bioinformatics e.g. To run an application on one computer the important algorithms on the planet to utilize calculations, information and figuring assets on servers in different parts of the world, SOAP-and REST-based interfaces are being created and used.
But the cloud computing architecture plays a vital role here. With cloud computing technique people can be watched online on regular interval of time and with the help of cloud computing technique the instant diagnosis is possible especially for H1N1, Heart Diabetic, and Hypertension patient.

V. LITERATURE SURVEY

In 2001 J. Su. et. al. [3] found medical data discovery using three basic techniques named Bayessian Network, Decision Trees and Back Propagation Neural Networks. According to this approach mined medical information is not only classified for research perspectives but also made available to the physician to improve his practices. Six steps of knowledge discovery include data collection firstly followed by data filtration, data enhancement, data encoding, data mining and at last knowledge representation. High correlation parameter is chosen for construction of Bayessian networks. While learning rate of 0.5 with learning time of 1000 nanoseconds is taken for back propagation technique. After generating decision tree the results of three algorithms were discussed by considering average of high accuracy of data sets. In 2006, S. Yoon et. al. [9] [10] found usefulness of biochips for acquisition of biological data with high throughput. The main advantages of micro fluidic lab on a chip include ease of use, speed of analysis, low sample and reagent consumption, and high reproducibility due to standardization and automation. Without effective data-analysis methods, however, the merit of acquiring massive data through biochips will be marginal.

In 2014, Abbas et. al. [2] stated importance cloud computing in healthcare services in addition to other business and social domains. Cloud computing with healthcare support services reduces maintenance costs. Author consider certain issues such as twenty four hours availability of Medical health data and equipments to the use for Medical-health clouds were also highlighted. In 2014, Sood et. Al. [4] proposed a proactive resource provisioning methodology and highlight effective use of artificial neural network (ANN). In 2014, Wang et al. [10] proposed a hybrid model of mobile cloud computing for medical data monitoring, which efficiently diagnosis large scale data. They proposed self health management monitoring system and express how it is most convenient to user. In 2014, Xu et al. [11] author developed a emergency medical services support system and provides flexibility to the user with furnished semantic data model to store & access ubiquitous Internet of things data.

VI. CONCLUSION

Revolutionary change in the industry working has been seen in the past few years. The concept of cloud computing in various application areas is bringing new inventions like in bioinformatics. The cloud as a service to the patient is like a God's blessing. The patient data is processed over the cloud throughout the year, twenty four hours, seven days a week & monitoring the patient records. Whenever there is any abnormality in the records the message is send to the patient, to their family members as well as to the doctors. This kind of system plays important role in critical disease like H1N1 where the person affecting from virus can affect some other persons present in public. For nation's government and healthcare departments easily transmitted diseases are one of the major concerns, with the increase in information technologies like cloud computing, it is possible to control many infections in most effective and efficient way. This paper highlights the important and effective use to cloud computing in a field of bioinformatics.

REFERENCES

[1] Ilango Sriram, Ali Khajeh-Hosseini, "Research Agenda in Cloud Technologies" submitted to the 1st ACM Symposium on Cloud Computing, SOCC 2010, arXiv:1001.3259.

[2] A. Abbas, S.U. Khan, A review on the state-of-the-art privacy-preserving approaches in the e-Health clouds, IEEE J. Biomed. Health Inf. 18 (4) (2014) 1431–1441.

[3] Jenn-Lung Su, Guo-Zhen Wu, I-Pin Cha. (2001, October). *THE APPROACH OF DATA MINING METHODS FOR MEDICAL DATABASE*. Paper presented at IEEE, Istanbul, Turkey.

[4] R. Sandhu, S.K. Sood, G. Kaur, An intelligent system for predicting and preventing MERS-CoV infection outbreak, J. Supercomput. (2015), http://dx. doi.org/10.1007/s11227-015-14740.

[5] Wang, C., Xiao, Z., Liu, Y., Xu, Y., Zhou, A., & Zhang, K. (2013). SentiView: Sentiment Analysis and Visualization for Internet Popular Topics. *IEEE Transactions on Human-Machine Systems*, *43*(6), 620-630. doi:10.1109/thms.2013.2285047.

[6] Jeonghee Yi Wayne Niblack. (2005). Sentiment Mining in WebFountain. San Jose, CA.

[7] Yang, C. C., & Dorbin Ng, T. (2011). Analyzing and Visualizing Web Opinion Development and Social Interactions With Density-Based Clustering. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *41*(6), 1144-1155. doi:10.1109/tsmca.2011.2113334.

[8] Jin Xingxing, Cai Yingkun, *Xie Kunqing, Ma Xiujun, Sun Yuxiang, Cai Cuo. (2006, January). *A Novel Method to Integrate Spatial Data Mining and Geographic Information System*. Paper presented at IEEE, Beijing, China.

[9] Wang, C., Xiao, Z., Liu, Y., Xu, Y., Zhou, A., & Zhang, K. (2013). SentiView: Sentiment Analysis and Visualization for Internet Popular Topics. *IEEE Transactions on Human-Machine Systems*, *43*(6), 620-630. doi:10.1109/thms.2013.2285047.

[10] X. Wang, Q. Gui, B. Liu, Z. Jin, Y. Chen, Enabling smart personalized healthcare: a hybrid mobile-cloud approach for ECG telemonitoring, IEEE J. Biomed. Health Inf. 18 (3) (2014) 739–745.

[11] B. Xu, L. Da Xu, H. Cai, C. Xie, J. Hu, F. Bu, Ubiquitous data accessing method in iot-based information system for emergency medical services, IEEE Trans. Ind. Inf. 10 (2) (2014) 1578–1586.

[12] Yang, C. C., & Dorbin Ng, T. (2011). Analyzing and Visualizing Web Opinion Development and Social Interactions With Density-Based Clustering. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *41*(6), 1144-1155. doi:10.1109/tsmca.2011.2113334.

[13] S. Sharma, " A Review on Efficacy of Artificial Neural Networks in Medical & Business Areas ", International Journal of Recent Trends in Engineering & Research, Volume 02, Issue 04, April – 2016

[14] S. Sharma, " Cervical Cancer stage prediction using Decision Tree approach of Machine Learning ", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 4, April 2016.