Bioinformatics –An Information Explosion Arena: An Overview

M.Selvi
Assistant Technical Officer
University of Madras,Chennai
Selvinandu2909@gmail.com

Abstract

Recent years have seen an explosive growth in biological information. Biological information is becoming increasingly important to the modern research community. The task of managing the information is important yet the task presenting this information to the researchers is of even greater importance .One of the major challenges for the academicians and researchers working in the field of biological resources. This field of science is growing at an exponential pace with immense possibilities of innovation. Research in Bioinformatics includes the methods of development for storage, retrieval and analysis of data

Keywords

Bioinformatics, Biological data management.

Electronic access

The journal is available at www.jalis.in

Journal of Advances in Library and Information Science
ISSN: 2277-2219 Vol. 1. No.4. 2012. pp. 192-196

Introduction

Over the past three centuries the rapid developments in genomic and other molecular research technologies and developments in information technologies have combined to produce a tremendous amount of information related to molecular biology. Bioinformatics is the name given to these mathematical and computing approaches used to glean understanding of biological processes." The tremendous increase of information leads to the globalization of Information". We are drowning in the sea of Information. According to Varian [1] "When you look at the challenge we face how do we manage all this information? In Biological field information explosion leads to lot of changes where in the case of Biotechnology or Nanotechnology.

From the 17th Century the biologist started dealing with problems of information management. Developments in laboratory technology facilitated collection of data at a rate that was faster than data interpretation Biologist facing difficulties in analyzing and interpreting this data overload.Bioinformaics has a role in developing statistical methods for analysis of large data sets an in developing methods of information management for the New types of data being generated.

Bioinformatics:

The term bioinformatics was coined by Paulien Hogeweg and Ben Hesper in 1978 for the study of informatics processes in biotic systems. Its primary use since at least the late 1980s has been in genomics and genetics, particularly in those areas of genomics involving large-scale DNA sequencing.
Bioinformatics, as the name suggest, It is an interdisciplinary field involving majorly biology, computer science and information technology. Chemistry and Mathematics also plays an important role. Bioinformatics deals the creation and advancement of databases, algorithms, computational and statistical techniques and theory to solve formal and practical problems arising from the management, analysis and interpretation of biological data.

Definition:

According to T.K.Attwood and D.J.Parry-Smith the term Bioinformatics is used to encompass almost all computer applications in biological sciences…

The NCBI [6] defines as: "Bioinformatics is the field of science in which biology, computer science, and information technology merge into a single discipline. There are three important sub-disciplines within bioinformatics: the development of new algorithms and statistics with which to assess relationships among members of large data sets; the analysis and interpretation of various types of data including nucleotide and amino acid sequences, protein domains, and protein structures; and the development and implementation of tools that enable efficient access and management of different types of information."

BISTIC definition Committee defines "Bioinformatics and computational biology are rooted in life sciences as well as computer and information sciences and technologies. Both of these interdisciplinary approaches draw from specific disciplines such as mathematics, physics, computer science and engineering, biology, and behavioral science. Bioinformatics and computational biology each maintain close interactions with life sciences to realize their full potential. Bioinformatics applies principles of information sciences and technologies to make the vast, diverse, and complex life sciences data more understandable and useful. Computational biology uses mathematical and computational approaches to address theoretical and experimental questions in biology. Although bioinformatics and computational biology are distinct, there is also significant overlap and activity at their interface.

The NIH Biomedical Information Science and Technology Initiative Consortium agreed on the following definition of bioinformatics recognizing that "*Bioinformatics:* Research, development, or application of computational tools and approaches for expanding the use of biological, medical, behavioral or health data, including those to acquire, store, organize, archive, analyze, or visualize such data.

History

The history of biology in general, B.C. and before the discovery of genetic inheritance by G. Mendel in 1865, is extremely sketch and inaccurate. This was the start of Bioinformatics history. Gregor Mendel is known as the "Father of Genetics". He did experiment on the cross-fertilization of different colors of the same species. He carefully recorded the data and analyzed the data. Mendel illustrated that the inheritance of traits could be more easily explained if it was controlled by factors passed down from generation to generation.

The understanding of genetics has advanced remarkably in the last thirty years. In 1972, Paul berg made the first recombinant DNA molecule using ligase. In that same year, Stanley Cohen, Annie Chang and Herbert Boyer produced the first recombinant DNA organism. In 1973, two important things happened in the field of genomics. The advancement of computing in 1960-70s resulted in the basic methodology of bioinformatics. However, it is the 1990s when the INTERNET arrived when the full fledged bioinformatics field was born.

Objectives:

- To provide the background information which will help demystify computer usage
- To provide an introduction to the resources available to the biologist
- To encourage the analysis an dissemination of biological knowledge

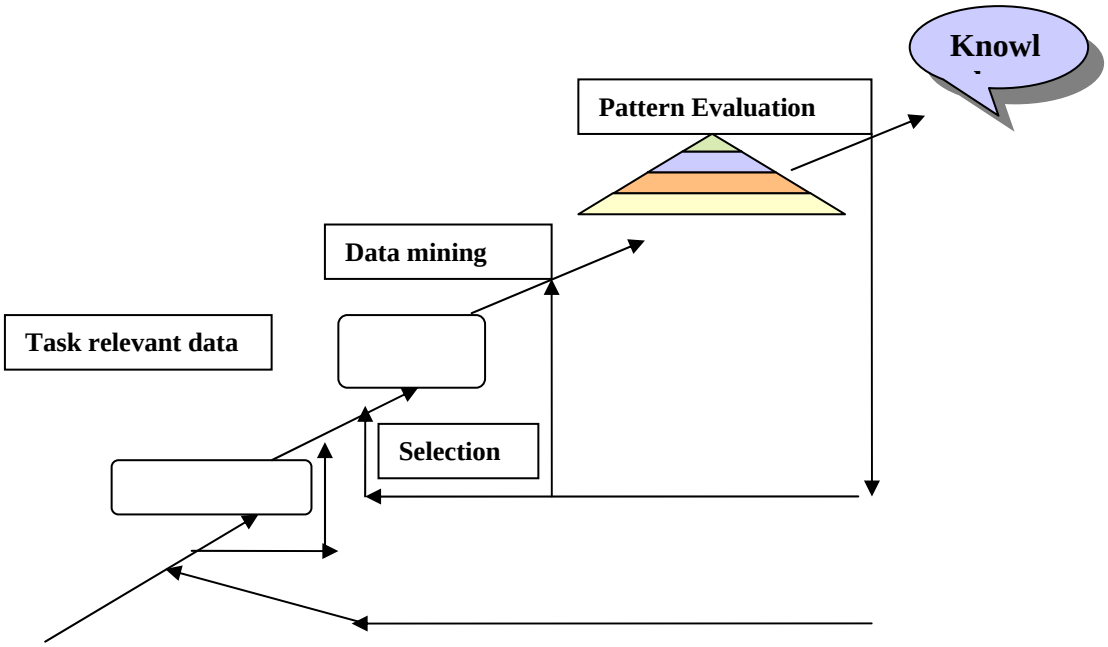
Applications of Bioinformatics

1. Molecular medicine
2. Preventative medicine
3. Gene therapy
4. Drug development
5. Microbial genome applications
6. Waste cleanup
7. Climate change Studies
8. Alternative energy sources
9. Biotechnology
10. Antibiotic resistance
11. Forensic analysis of microbes
12. Bio-weapon creation
13. Evolutionary studies
14. Crop improvement
15. Insect resistance
16. Improve nutritional quality
17. Development of Drought resistance varieties
18. Vetnary Sciences

Data mining of Biological Databases:

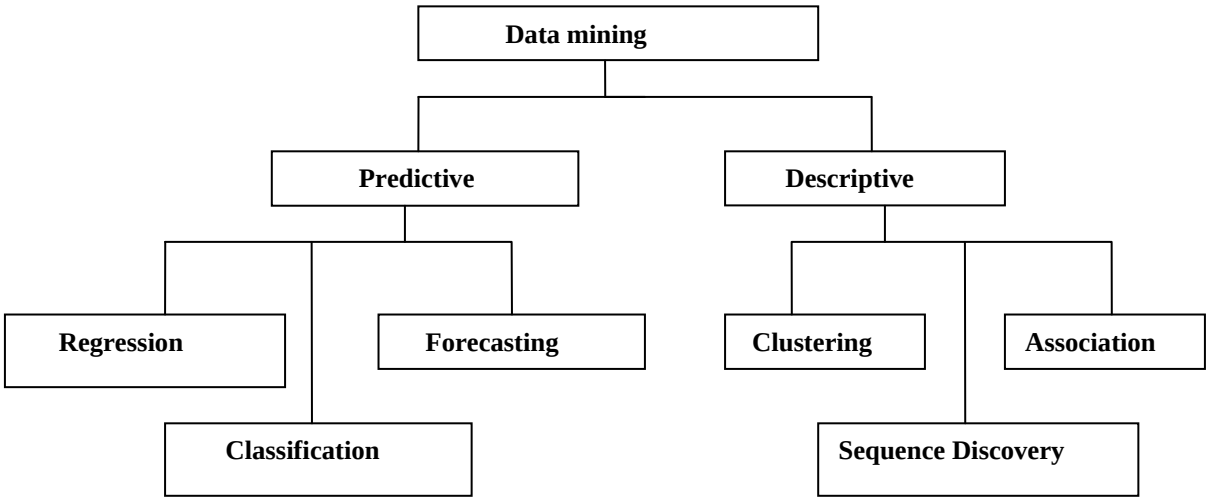
One of the most important applications of computers in modern biology is the data mining method to biological databases. Data mining is a process of knowledge discovery in data bases and may include knowledge extraction, data analysis, data archaeology, data dredging etc.

Steps in knowledge discovery [8]



These techniques grouped in to predictive and non-predictive. Predictive techniques use the past data to learn. Descriptive techniques provide a statistical view of the data to help in the analysis

Classification of various data mining tools



.

Challenges in Information Integration

Integration is a major challenge facing by the researchers and institutions that wish to explore these deposits of information. Data integration is an ongoing active area in the commercial world. Over the past decades, enormous efforts and progress have been made in many data integration system. However the following fundamental features are desirable for robust data integration system[9]

Bioinformatics scenario in India:

During the 7th five year plan - In 1986 Department of Biotechnology launched Biotechnology Information systems (BTIS). It is also a nationwide network with 10 Distributed Information Centers (DICs) and 48 Sub Distributed Information Centers (SDICs). Its mission is to establish India as a leader in Bioinformatics
During the 8th five year plan BTISnet also established. All these DICs and SubDICs were connected through satellites and terrestial links under two major network service providers - NICNET and ERNET.

These DICs and SubDICs have been given the tasks of providing discipline oriented information to all institutions interested in the relevant fields [10]. A list of various disciplines with respective DICs responsible for the same is given below.
1. Genetic Engineering: Jawaharlal Nehru University, Delhi. Indian Institute of Science, Bangalore, Madurai Kamarajar University, Madurai. Bose Institute, Calcutta
2. Protein Modelling and Protein Engineering: Institute Microbial Technology, Chandigarh.
3. Plants tissue culture, Photosynthesis and plant microbiology: Indian Agricultural Research Institute, New Delhi.
4. Animal cell culture and virology: University of Pune, Pune.
5. Oncogenes, reproduction physiology, Cell transformation, Nucleic acid and protein sequences: centre for cellular and molecular Biology, Hyderabad.
6. Immunology: National Institute of Immunology, New Delhi
7. Neuro Informatics: National Brain Research Centre, Delhi
8. Teraflops supercomputing facility: Super computing facility for Bioinformatics and Computational Biology, at IIT Delhi
9. Biotechnology and related fields: Department of Biotechnology, Government of India New Delhi
During the 10^{th} five year plan the union government also planned to set NBI under DBT which will be on the lines of the NCB under the National Institute of Health of USA.
In the 11^{th} five year plan, Strengthening Bioinformatics R&D Infrastructure in terms of more super computing facilities.

Private Organisations:

Indian based companies such as Biocon India, Informatics Pvt Ltd, Advance Biochemical Lab working as contract research organizations for companies abroad as well as companies involved in customized software development work of India. There are pure information technology companies heading towards bioinformatics by collaborations with government organizations. Such companies like TCS, Nicholas, Piramal, Satyam, Centre for DNA Finger printing and Diagnosis (CDFD) Centre for Biochemical Technology (CBT) Center for Cellular and Molecular Biology (CCMB) etc.Almost all Indian pharmaceutical companies are in to bioinformatics research and development. Some of them are Ranbaxy, Dr.Reddy's Lab, Dabur, Smith Kleim Beecham, Pfizer, Cipla, Zydus, cadila, Wockhardt, Astra Zeneca and East Indian Pharmaceuticals, Pharma and Pharma related companies in India have the unique and exciting opportunity of

Academic Scenario in India:

India is gearing to improve the academic scenario in the field of Bioinformatics. DBT has started one year Advance diploma in Bioinformatics in five Indian Universities – Jawaharlal University, Calcutta University, PuneUniveristy, Madurai Kamaraj University and Pondicherry University. The NBI is proposed to regulate the research in Bioinformatics. It have the separate wings to develop various activities like database development, human resources, genomics, proteomics and services. IIT (Delhi) conducted certificate courses in supercomputing facility and plans to offer M.Tech in Bioinformatics. Bioinformatics Institute pf India (BII) the only institute that providing a distance learning programme on bioinformatics in the country

Bioinformatics Careers:

Career opportunities are available in various streams
- Database Programmer
- Bioinformatics software Developer
- Research scientist / Associate
- Computational Biologist
- Network analysist
- Biostatistician
- Cheminformatician
- Pharmacogenomic Research Scientist
- Pharmacogenetician
- Gene analyst

Conclusion:

In the age of information explosion, when everybody is talking of information overload, satisfying users' information needs in a simple, personalised and efficient way is becoming a challenge for library and information centers. Without Bioinformatics new research in most fields of medicine and biology would become to standstill. Now a days bioinformatics has a vital role in scientific research.

References:

[1].Lyman,P.,Varian,H.,*etal*.(2003)Howmuchinformation?http://www.sims.berkeley.edu/research/projects/how-much-info-2003/

[2.] Bioinfomatics and molecular Evolution by Higgs paul.G and Attwood Teresa.K. – 10

[3].Hogeweg, P. (1978). "Simulating the growth of cellular forms". *Simulation* 31 (3): 90–96.

[4].. Hogeweg, P. (2011). Searls, David B.. ed. "The Roots of Bioinformatics in Theoretical Biology". PLoS *Computational Biology* 7 (3):

[5]. T.K.Attwood and DJ.Parry-Smith "Introduction to BioInformatics" Published by Prentice Hall in 1999

[6]. The National Center for Biotechnology Information (NCBI 2001)

[7]. BISTIC Definition Committee and released on July 17, 2000

[8]. Bioinformatics – Methods and applications Genomics, Proteomics and Drug discovery – 3rd Ed by S.C.Rastogi, N.Mendiratta, P.Rastogi

[9]. Bioinformatics – Managing Scientific data by Lacroix Zoe and Critchlow.

[10]. Bioinformatics –Databases,tools and Algorithms by Bosu Orpita and Thukral Simminder p.13-17