**IV SEMESTER**

**PAPER – Bioinformatics and Biostatistics**

**UNIT - I**

**TOPIC – Definition, Scope of Bioinformatics**

**Bioinformatics vs Computational Biology**

**Branches of bioinformatics**

**SOURCE – INTERNET**

**NAME OF INSTRUCTOR – G.N.V. SATISH**

## Definition, nature and scope of bioinformatics

Bioinformatics is a hybrid branch of biology and information technology. It encompasses the knowledge of computer science, statistics, mathematics, chemistry, biochemistry and physics.

It is science of developing computer databases and algorithms for purpose of speeding up and enhancing biological research.

The term Bioinformatics was first came into use in the 1970s, the branch bioinformatics mainly deals with management, analysis and storage of DNA, RNA and protein sequence data etc. research in Bioinformatics include methods development for, storage, retrieval and analysis of the data. the national center for Biotechnology information (NCBI), 2001 defines Bioinformatics as "Bioinformatics is the field of science in which Biology, computer science and information technology merge into a single discipline.

### Scope of Bioinformatics

Bioinformatics process data on gene sequence that is performed by biotechnologies and makes the information accessible for the research, Bioinformatics molecular Biology and computer Science, the main source is to fetch all relevant data and processes into useful information.

Bioinformatics has a very wide scope that includes,

→ Management and analysis of a wide set of biomedical data.

→ It is specially used in human gene sequence where large set of data are being handled

→ Bioinformatics play a major role in the research in development of the biomedical field.

→ Bioinformatics uses computational coding for several applications that involves finding gene and protein functions and sequences, developing evolutionary relationship and analysing the three-dimensional structure of protein.

→ Research works based on genetic diseases and microbial disease entirely depends on Bioinformatics where the derived information can be willing to produce personalized medicines.

### Bioinformatics versus computational Biology:

- Bioinformatics refers to the study of large set of biological statistics and scientific study results.
- It is the field that involves the development and application of computational methods to analyse and interpret large biological data sets.
- This includes everything from DNA sequencing data to protein structure information and even clinical data from patient samples.

- Bioinformatics is used in a wide range of biological research applications, from studying the genetics of disease to designing new drugs.

**Computational Biology:**

- Computational biology is a broader field that encompasses both bioinformatics and other areas of research that involve computational modelling and simulation of biological systems.
- Computational biology uses mathematical models and computer simulations to study complex biological systems and processes such as protein folding, gene regulation and cellular signalling pathways.
- Computational biology is an extremely powerful tool for understanding biological systems as it allows researchers to simulate and manipulate systems that are too complex to study experimentally.

**Computational Biology versus Bioinformatics**

- While the two fields of computational biology and bioinformatics are closely related, there are some key differences between the two
- Bioinformatics is primarily concerned with the development and applications of computational methods for analyzing and interpreting large biological data sets.
- Computational biology on the other hand, involves the development and applications of mathematical models, and computer simulations to study complex biological systems and processes.
- The key differences between computational biology and bioinformatics are:

| **Aspect** | **Bioinformatics** | **Computational Biology** |
|---|---|---|
| Primary focus | Development and applications of computational tools and algorithms for managing, processing and interpreting large scale biological data | Development and applications of methods, mathematical models and computer simulations to study biological systems or phenomenon. |
| Goal | Extract meaningful, obtainable insights from existing biological data such as DNA sequences or protein structures | Understand functions and biological processes by creating and testing hypotheses through simulations |
| Data usage | primarily deals with analysing raw, large data sets (e.g., genomic, proteomic data) | uses analysed data to inform, validate and build complex models for simulating hypothetical scenarios |

| | | |
|---|---|---|
| Key tools and techniques | Sequence alignment (BLAST), gene finding, genome assembling, statistical analysis and machine learning algorithms | molecular dynamic tools, e.g., protein folding or model simple systems |
| Application | Genomics, pharmacogenomics, drug discovery, personalized medicine and database creation, maintenance (NCBI) | Systems biology, synthetic biology, model of evolution and cell division and simulation protein folding or cellular signalling pathways |

**Branches of Bioinformatics**

The science of bioinformatics can be divided into several branches based on experimental material used for the study.

**Bioinformatics is broadly divided into two groups:**

1. Animal bioinformatics
2. Plant bioinformatics

**Animal bioinformatics**:

- It deals with the computer aided study of genomics, proteomics and metabolomics in various animal species.
- It includes study of gene mapping, gene sequencing and animal gene resources etc.

**Plant bioinformatics:**

- It deals with computer aided study of plant species. It includes gene mapping, gene sequencing and plant data base etc.
- It can be further divided into agricultural bioinformatics, medicinal plants bioinformatics and forest plant bioinformatics.
- A living cell is a system, with cellular components interacting with each other and these interactions determine the fate of the cell, characterisation of these three components i.e., the genome, the transcriptome and the proteins and the associated development of analytical methods lead to the establishment of the three closely related branches of bioinformatics - Genomics, Proteomics, Transcriptomics.

**Genomics:**

- Genomics involves extensive analysis of nucleic acids through molecular biological techniques.
- Genomics is a science that attempts to describe a living organism in terms of the sequence of its genome.

**Proteomics:**

- It represents the earliest attempt to identify a major subclass of cellular components, the proteins and their interactions.
- It has been coined from the word 'proteome' which is the complete protein compliment of a system.
- Proteomics involves the sequencing of amino acids in a protein, determining its three dimensional structure and relating it to the function of the protein.
- Proteomics also focuses on identifying when and where proteins are expressed in a cell, so as to establish their physiological roles in an organism.

**Transcriptomics:**

- It focuses the expression level of genes, using techniques like DNA microarrays.
- Transcriptomics has been coined from the word 'transcriptome' which is the set of all m-RNA molecules in one (or) a population of biological cells for a given set of environmental circumstances.

**Functional genomics:** It assigns functional relevance to genomic information. It is the study of genes their resulting proteins and the role played by the proteins.

**cheminformatics:**

- Drug design through bioinformatics is one of the most actively pursued areas of research.
- cheminformatics deals with compounds (low molecular weight). The products of secondary metabolism, often called natural products which shows bioreactivity.
- cheminformatics involves organisation of chemical data in a logical form to facilitate the process of understanding chemical properties, their relationship to structures and making inferences.
- chemical structures are the input to identify similar compounds for screening for biological activity.
- It also helps to assess the property of new compounds by comparison with the known compounds.