BT 302 Bioinformatics 2-0-2-6

Description/Preamble

This course is designed to introduce the basic concepts and methods employed in Bioinformatics through hands on learning experiences. Topics include sequence-based homology detection, gene prediction, phylogenetic analysis, etc. with clear emphasis on the understanding and utilization of concepts and algorithms. The objective is to help the students rapidly reach the frontier of Bioinformatics and use standard tools to solve problems in modern biology.

Industrial Relevance

The advent of genomic era heralds an ever-expanding deluge of biomedical information concerning human health and wellbeing. Computational approaches contribute immensely to glean these large volumes of biological data for devising suitable therapeutic interventions. Global healthcare is a trillion-dollar industry. However, cost of developing a new drug from concept to market is projected to be around 35000 crore rupees (5 billion US dollars) with a success rate in clinical phase approval remain at a mere 12 %. Such 'drug failures' are key contributors to development costs. Computational methodologies have therefore become a crucial component for many drug discovery programs, especially in the early stages of the pipeline to minimize this exorbitant cost and verify transition probabilities from phase to phase using *in silico* prediction tools. Undergraduate students in IIT are good in their quantitative skills and engineering practices. A good training in computational methods can make them more employable and complement current industry strategy aimed at accelerating and economizing discovery and development process. India already has a matured information technology platform in place, which can potentially supplement computational drug discovery initiatives, but not exploited effectively till date. Development of new solutions demand trained personnel who have a good understanding of the operational principles in biology along with computer aided design to effectively supplement industrial research.

Objective

To provide a broad understanding of computational techniques and resources available to biological scientists. Student should be able to supplement their future research on both experimental and computational projects with this training.

Pre-requisites

This course is designed for advanced undergraduates with reasonably strong backgrounds in molecular biology, mathematics, and computer science, but not necessarily all.

Syllabus

Introduction to biological databases: collection, organization, storage and retrieval of data; Concept of homology and definition of associated terms; Pairwise sequence alignment: dynamic programming algorithm, global (Needleman-Wunsch) and local (Smith-Waterman) alignments; BLAST Scoring matrices (PAM and BLOSUM families), gap penalty, statistical significance of alignment; Multiple sequence alignment: progressive alignment, iterative alignment, Sum-of-pairs method, CLUSTAL W; Pattern recognition in protein and DNA sequences, Hidden Markov Model, Profile construction and searching, PSI-BLAST; Big data analysis: Introduction to Next-generation sequencing analysis, RNA-seq, CHIP-seq, Introduction to phylogeny: maximum parsimony method, distance method (neighbor-joining), maximum-likelihood method; Gene prediction in prokaryotes and eukaryotes, homology and ab-initio methods; Genome analysis and annotation; comparative genomics, cluster of orthologous groups.

Text Books/References

1. Marketa Zvelebil, Jeremy O. Baum. Understanding Bioinformatics. Garland Science, 2007.

2. Mount, David W. Bioinformatics: Sequence and Genome Analysis (2nd Ed.). CSH press, 2005
3. Bourne, Philip E. Structural Bioinformatics. (2nd Ed.). Wiley, 2009.
4. Richard Durbin, Sean R. Eddy, Anders Krogh, Graeme Mitchison. Biological Sequence Analysis. Cambridge University Press, 1998.
5. Kochan, Stephen G; Wood, Patrick. UNIX shell programming (3rd Ed.). SAMS, 2003.

Evaluation & Grading

Evaluation will be based on assignments, lab exercises and small projects along with mid semester and end semester examinations. Grading from AS to DD and F, as per standard grading system followed by institute.