Course Code: DA461 **Course Name: Bioinformatics** **Credits: 2-0-2-6**

Pre-requisite: None

Syllabus: Primer on molecular biology: Structure and function of nucleic acids and protein; genes and gene expression; the central dogma of biology; Sequencing, structure determination, and bioinformatics databases. Sequence alignment: Pairwise and multiple sequence alignment; Dynamic Programming; Global alignment; Local alignment; Scoring matrices and gap penalty; Carrillo–Lipman algorithm; Feng–Doolittle algorithm; CLUSTAL; Sequence logo. Pattern detection: Gibbs sampler; Hidden Markov model; Profile HMM Alignment. Phylogenetic analysis: Molecular evolution, homolog, orthologs, paralogs; Rooted and unrooted phylogenetic tree; Maximum parsimony method; Distance-based method; Genome assembly and next-generation sequencing: Shortest superstring approach; Overlap graph approach; de Bruijn graph approach; NGS read mapping; RNA-seq read mapping; Peak calling method. The lectures will focus on the well-established algorithms in these topics, and the laboratory exercises will supplement those lectures with programming assignments and mini projects.

Textbooks:

- Neil C. Jones, and Pavel A. Pevzner, An Introduction to Bioinformatics Algorithms, 1st Edition, ANE Books, 2009.
- Richard Durbin, Sean R. Eddy, Anders Krogh, and Graeme Mitchison, Biological Sequence Analysis, Cambridge University Press, 1st Edition, 1998.
- Wing-Kin Sung, Algorithms For Next-Generation Sequencing, 1st Edition, CRC Press, 2020.

References:

- Phillip Compeau and Pavel Pevzner, Bioinformatics Algorithms: An Active Learning Approach, Vol. I, 2nd Edition, Active Learning Publisher, 2015.
- Mourad Elloumi and Albert Y. Zomaya, Algorithms In Computational Molecular Biology, 1st Edition, Wiley, 2011.