PCS8J001 ALGORITHMS FOR BIOINFORMATICS

UNIT I – **[10 HOURS]**

DYNAMIC PROGRAMMING ALGORITHMS (9 hours) Introduction to Algorithms, Dynamic Programming, Sequence Alignment: Edit distance, LCS. PAM and BLOSUM Scoring Matrices. Global alignments: Needleman Wunsch Algorithm, Local Alignments: Smith Waterman Algorithm, Gap Penalties.

UNIT II- **[8 HOURS]**

GRAPH ALGORITHMS (9 hours) Graph Algorithms, SBH and Eulerian Paths, De-novo Peptide Sequencing: Longest Paths and Space Efficient Alignment Algorithms. Fast LCS using Table Lookup.

UNIT III- **[10 HOURS]**

PATTERN MATCHING AND CLUSTERING (9 hours) Exact Pattern Matching: KMP Algorithm, Keyword Trees, Aho-Corasick Algorithm. Clustering Basics: Hierarchical Clustering, Multiple Sequence Alignment: CLUSTAL, Center-based Clustering, Clustering via Cliques.

UNIT IV- **[12 HOURS]**

EVOLUTIONARY TRESS AND PHYLOGENY (9 hours) Evolutionary Trees and Ultrametrics, Additive distance trees, Perfect Phylogeny Problem, Small Parsimony Problem, Nearest Neighbor Interchange. HIDDEN MARKOV MODELS, RANDOMIZED ALGORITHMS (9 hours) Hidden Markov Models: Basics, Forward and Backward (Viterbi) Algorithms, Randomized algorithms and their applications.

REFERENCES

1. Neil C. Jones and Pavel A. Pevzner, "An Introduction to Bioinformatics Algorithms", MIT Press, 2005.

2. Gusfields D, "Algorithms on strings, trees and sequences: Computer Science and Computational Biology", Cambridge University Press, 1997.

3. Steffen Schulze-Kremer, "Molecular Bioinformatics: Algorithms and Applications", Walter de Gruyter, 1996.

4. Gary Benson, Roderic Page (Eds.), "Algorithms in Bioinformatics", Springer International Edition, 2004.

5. Richard Durbin, Sean R. Eddy, Anders Krogh, Graeme Mitchison. "Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acid", Cambridge University Press, 1999.