# AMNE26 INFORMATION CODING & TECHNIQUES

## UNIT-1 INTRODUCTION
1.1 Biological data in digital symbol sequences, genomes, proteins and proteomes, biological sequences, molecular function and structure.
1.2 Biological Databases: Sequence databases, mapping databases, information retrieval, genomic databases.
1.3 Machine Learning Foundations: The probabilistic framework and examples.

## UNIT-2 MACHINE LEARNING ALGORITHMS
2.1 Introduction, dynamic programming, gradient descent, EM/GEM algorithms,
2.2 MarkovChain Monte Carlo methods, simulated annealing, evolutionary and genetic algorithms, learning algorithms.
2.3 Neural Network: Theory and Applications.
2.4 Hidden Markov Models: Theory and applications

## UNIT-3 PROBABILISTIC GRAPHICAL MODELS IN BIOINFORMATICS
3.1 Markov Models and DNA symmetries, gene finders, hybrid models and neural network parameterization of graphical models, single model case, bidirectional recurrent neural networks for protein secondary structure prediction.
3.2 Probabilistic models of evolution: phylogenetic trees.

## UNIT-IV STOCHASTIC GRAMMARS AND LINGUISTICS
4.1 Introduction, formal grammars, Chomsky hierarchy, applications of grammars, learning algorithms, applications of SCFGs.
4.2 Microarrays and gene expression: Introduction, Probabilistic modelling of array data, clustering, gene regulation.

**References Books:**
1. TK Attwood & DJ Parry-Smith," Introduction to Bioinformatics", Pearson Education
2. Edward Keedwell and Ajit Narayanan, "Intelligent Bioinformatics" John Wiley & Sons, Ltd.
3. A Tramontano, "Introduction to Bioinformatics", Chapman & Hall/CRC.