



**Approved in 44<sup>th</sup> BoA Meeting (24-11-2021)**

<b>Course number</b>	<b>: BE502</b>
<b>Course Name</b>	<b>: Design and Analysis of Bioalgorithms</b>
<b>Credit Distribution</b>	<b>: 2-0-2-3</b>
<b>Intended for</b>	<b>: Core for Dual Degree Bio Engg. students Elective for other B.Tech and M.Tech students.</b>
<b>Prerequisite</b>	<b>: IC152 – Computing and Data Science IC136 - Understanding Biotechnology &amp; its applications IC260 – Signals and Systems</b>
<b>Mutually exclusive</b>	<b>: None</b>

---

### **1. Preamble:**

The course is meant to cover some important considerations in developing algorithms across various bio-informatics and bio-signal processing applications. Each module involves some well-established algorithms / methods targeting particular tasks, which will be compared and contrasted in terms of aspects such as their theoretical principles, quality of performance, complexity, speed, memory etc. After the introduction, 2 modules are dedicated to bioinformatics algorithms, and the other two modules cover bio-signal processing methods.

### **2. Course Modules with quantitative lecture hours:**

**Module 1: Introduction:** Background on genomics and proteomics, DNA-RNA gene, protein structures, Importance of Bio-algorithms, Bioinformatics algorithms and examples of algorithms designs, algorithm complexity and speed, Bio-signal processing algorithms and design considerations **(4 hours)**

**Module 2: Combinatorial pattern matching algorithms for genomes:** Hash Tables, Exact matching, Suffix Trees, Keyword Trees, Heuristic similarity search methods, approximate pattern matching, sequence similarity search, sequence alignment, BLAST, and motif finding **(9 hours)**

**Module 3: Graph-based and Clustering algorithms:** Graph terminologies and some basic algorithms, shortest superstring-based DNA sequencing, hybridization-based DNA sequencing, graphs for peptide sequencing, K-means clustering and Hierarchical clustering for gene expressions, graph-based clustering, phylogenetic tree

reconstruction (9 hours)

**Module 4: Bio-signal processing:** Pre-processing considerations in bio-signals (for EEG and FMRI), signal decomposition methods, graph-based and clustering algorithms for EEG and FMRI, feature extraction and their uses in diagnosis of diseases (6 hours)

**3. Text books:**

1. N.C. Jones and P.A. Pevzner. **An introduction to bioinformatics algorithms**, MIT Press, 2004
2. K. Najarian, R. Splinter. **Biomedical Signal and Image Processing**, 2<sup>nd</sup> Edition, CRC Press, 2012

**4. References:**

1. T. Cormen, C. Leiserson, R. Rivest, C. Stein. **Introduction to algorithms**, 3<sup>rd</sup> Edition, MIT Press, 2009

**5. Similarity Content declaration with existing courses:**

S. No.	Course Code	Similarity Content	Approx. % of Content
1.	BE5XX – Computational Biology	Some target tasks in the bio-informatics modules may overlap. But the proposed course (BE3XX) focuses more on the details and analysis of algorithms for such tasks	20%
2.	CS403 – Algorithms Design and Analysis	The principles of analysis of algorithms in the first three modules are similar. However, the algorithms covered in the proposed course are different and have an bio-informatics context.	20%

**6. Justification of new course proposal if cumulative similarity content is >30%:**