

Campus: São José dos Campos		
Curso (s): Ciência da Computação, Engenharia de Computação		
Unidade Curricular (UC): Tópicos (Bioinformática Avançada)		
Unidade Curricular (UC): <i>[nome da UC em inglês] Advanced Bioinformatics</i>		
Unidade Curricular (UC): <i>[nome da UC em espanhol - opcional]</i>		
Código da UC: 7217		
Docente Responsável/Departamento: Reginaldo Massanobu Kuroshu / Departamento de Ciência e Tecnologia		Contato (e-mail): <i>[opcional]</i> rmkuroshu@unifesp.br
Docente (s) Colaborador/a (es/as)/Departamento (s):		Contato (e-mail): <i>[opcional]</i>
Ano letivo: 2024	Termo: 5º	Turno/Turma: Noturno
Nome do Grupo/Módulo/Eixo da UC (se houver):		Idioma predominante em que a UC será oferecida: (X) Português ( ) English ( ) Español ( ) Français ( ) Libras ( ) Outro:
UC: ( ) Fixa (X) Eletiva ( ) Optativa	Oferecida como: (X) Disciplina ( ) Módulo ( ) Estágio ( ) Outro:	Oferta da UC: (X) Semestral ( ) Anual
Ambiente Virtual de Aprendizagem: (X) Moodle (X) Classroom ( ) Outro: ( ) Não se aplica		
Pré-Requisito (s) - Indicar Código e Nome (s) da (s) UC: Algoritmos e Estruturas de Dados I		
Carga horária total (em horas): 72 hrs		
Carga horária teórica (em horas): 36 hrs	Carga horária prática (em horas): 36 hrs	Carga horária de extensão (em horas, se houver):
Se houver atividades de extensão, indicar código e nome do projeto ou programa vinculado na Pró-Reitoria de Extensão e Cultura (ProEC):		
Ementa: Introdução à linguagem Python e ao uso de bibliotecas de BioPython. Métodos de Alinhamento de Sequências simples (Needleman-Wunsch, Smith-Waterman, BLAST). Métodos de alinhamentos múltiplos. Modelos de Markov escondidos. Análise de Clusters. Análise de Componentes Principais. Métodos de sequenciamento, montagem e análise de genomas.		
Conteúdo programático: Introdução à Python, biblioteca Biopython. Genômica, métodos e dados de sequenciamento de DNA. Comparações de sequências: distância Hamming, edit distance, Needleman-wunsch, Smith-waterman, BLAST. Análise de genomas, homologia, repetições, evolução, alinhamento múltiplo. Montagem de genomas: OLC, grafos de bruijn. HMM. Clustering: agrupamento de expressão gênica, agrupamento hierárquico, k-means. Análise de Componentes Principais. Tópicos selecionados.		
Objetivos: Curso teórico-prático com o objetivo de introduzir conceitos e alguns dos principais problemas computacionais da bioinformática. Objetivos específicos: Capacitar o aluno a resolver problemas de biologia computacional por meio de programação e aplicação de ferramentas de bioinformática.		
Metodologia de ensino: Aulas expositivas; Laboratório de programação; Listas de exercícios; Atividades complementares a distância.		
Avaliação:		

Provas (P): duas provas P1 e P2 escritas ou práticas

$$P = 0,5 \times P1 + 0,5 \times P2$$

Exercícios (E)

Projeto (T)

$$\text{Nota final: } NF = 0,5 \times P + 0,25 \times E + 0,25 \times T$$

Bibliografia:

Básica:

1. N. C. Jones and P. A. Pevzner. An Introduction to Bioinformatics Algorithms, The MIT Press; 1 edition, 2004.
2. KINSER, J.M. Python for Bioinformatics. Jones & Bartlett, 2008.
3. MOUNT, D.W. Bioinformatics. 2nd Ed. CSHL Press, 2004.
4. LESK, A.M. Introdução à Bioinformática. 2ª Ed. Artmed, 2005.

Complementar:

1. GRAUR, D.; LI, W.H. Fundamentals of Molecular Evolution. 2nd Ed. Sinauer, 2000.
2. D. Gusfield. Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology. Cambridge, UK: Cambridge University Press, 1997
3. M. Waterman. Introduction to Computational Biology: Maps, Sequences, and Genomes, Boca Raton, FL: CRC Press, 1995.
4. R. Durbin, S. R. Eddy, A. Krogh, G. Mitchison: Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids, Cambridge University Press, 1998.

Cronograma: *[opcional]*