

Dados de identificação		
<i>Disciplina</i>	Classificação e Pesquisa de Dados	
<i>Oferecida para</i>	Bacharelado em Ciência da Computação, Engenharia da Computação e Bacharelado em Matemática -	
<i>Período Letivo</i>	2020/2	
<i>Professor Responsável</i>	Leandro Krug Wives	
<i>Sigla</i>	INF01124	
<i>Carga horária (horas)</i>	60	
<i>CH Autônoma (horas)</i>	6	
<i>CH Coletiva (horas)</i>	54	
<i>CH Individual (horas)</i>	0	
Súmula	Métodos de classificação de dados e de pesquisa em tabelas. Conceituação de arquivo; organizações básicas e manipulação (métodos básicos). Índices: indexado-sequencial; B-tree, B+, B*, e outros; hashing, ext. hash, e outros; arquivos invertidos; multilistas. Estruturas não-convencionais: noções: dados no espaço em disco; dados textuais em disco.	
Objetivos	Capacitar o aluno na seleção e análise de algoritmos de classificação de dados, de pesquisa de dados em tabelas e de técnicas de organização e compactação de arquivos.	
Conteúdo Programático		
Título	Conteúdo	Semana
Introdução à Classificação e Ordenação de Dados	Apresentação da disciplina e do plano de ensino; Introdução, conceitos e definições relacionados; Taxonomia das famílias/classes de Métodos de Classificação e suas características.	1
Introdução à Complexidade de Algoritmos	Introdução à Complexidade de Algoritmos para a Classificação e Pesquisa de Dados	1
Classificação por inserção	Classificação por inserção direta (insertion sort): com busca linear e com busca binária; Shell sort	2
Classificação por trocas	Bubble sort, Comb sort e Shakesort Quicksort tradicional, randomizado e variações	5 e 6
Classificação por seleção	Seleção direta (selection sort) Heap sort	3
Classificação por intercalação	Algoritmo de merge two-way, multi-way merge sort	4
Classificação em tempo linear	Bucket sort Radix sort MSD e Radix sort LSD Counting sort	7
Classificação por distribuição / tempo constante	Hashing: visão geral Funções de hashing (método da divisão, funções para strings...) Técnicas de tratamento de colisões (endereçamento aberto e endereçamento fechado) Hashing dinâmico (tabelas extensíveis)	8
Estruturas e Projetos de Arquivos de dados	Visão geral sobre arquivos (de dados e de índice) Tipos de arquivos de dados: seriais (heap), sequenciais, ISAM e hash Projeto e modelagem de arquivos Racionalização e normalização	9
Estruturas de Arquivos de Índice	Árvores B e variações Árvores digitais (binárias, de prefixo e variações - TRIE e PATRICIA) Arquivos Invertidos	10 e 11
Indexação e busca em Textos e Strings	Processo de Indexação Estruturas de armazenamento Processamento de consultas booleanas Busca e comparação de strings	12 e 13
Compressão de dados	Supressão de dados/caracteres repetidos Codificação de itens (dicionários estáticos ou dinâmicos de comprimento fixo ou variável, i.e., LZW, Huffman)	14
Seminário de tópicos avançados	Estruturas e algoritmos atuais de busca e comparação de dados Seminário de estudo e apresentação de trabalhos	15

<p>Metodologia</p> <p><i>Estratégias didáticas em atividades remotas</i></p> <p><i>Estratégias didáticas em atividades presenciais</i></p> <p><i>Recursos disponibilizados</i></p>	<p>As atividades ocorrerão de forma síncrona e assíncrona, conforme previsto no cronograma disponibilizado no Moodle da disciplina. As aulas serão gravadas e disponibilizadas no Moodle da disciplina. Estão previstas, também, atividades práticas, a serem realizadas pelos alunos, de forma síncrona, em seus computadores pessoais. Essas atividades deverão ser realizadas até a data prevista. Todas as atividades serão propostas, entregues e avaliadas no Moodle da disciplina, onde constará as instruções a serem seguidas para sua realização. Em caso de dúvidas, os alunos poderão contar com atendimento individualizado do professor, em horário a ser combinado e realizado de forma remota.</p> <p>Não estão previstas atividades presenciais</p> <p>As atividades previstas assim como as instruções para sua realização serão disponibilizadas no Moodle do INF. Eventuais componentes externos ao Moodle e necessários para a realização das atividades estarão indicados no próprio Moodle. Também serão disponibilizados no Moodle links para os livros e artigos online e gratuitos, indicados para leitura e estudo, visando auxiliar na realização das atividades propostas.</p>
<p><i>Recursos computacionais</i></p> <p><i>Informações sobre Direitos Autorais e de Imagem:</i></p>	<p>Para acompanhar as atividades previstas é necessário ter acesso regular à Internet. Atividades síncronas e atendimentos pessoais serão disponibilizados através do Microsoft Teams, e podem ser acompanhadas através de telefone ou computador, preferencialmente com microfone e câmera. Para a realização das atividades propostas será necessário ter acesso a computador.</p> <p>Todos os materiais disponibilizados são exclusivamente para fins didáticos, sendo vedada a sua utilização para qualquer outra finalidade, sob as penas legais. Todos os materiais de terceiros que venham a ser utilizados devem ser referenciados, indicando a autoria, sob pena de plágio. A liberdade de escolha de exposição da imagem e da voz não isenta o aluno de realizar as atividades originalmente propostas ou alternativas; Todas as gravações de atividades síncronas devem ser previamente informadas por parte dos professores. Somente poderão ser gravadas pelos alunos as atividades síncronas propostas mediante concordância prévia dos professores e colegas, sob as penas legais. É proibido disponibilizar, por quaisquer meios digitais ou físicos, os dados, a imagem e a voz de colegas e do professor, sem autorização específica para a finalidade pretendida. Os materiais disponibilizados no ambiente virtual possuem licença de uso e distribuição específica, sendo vedada a distribuição do material cuja a licença não permita ou sem a autorização prévia dos professores para o material de sua autoria.</p>
<p>Carga Horária</p> <p><i>Teórica</i></p> <p><i>Prática</i></p>	<p>52</p> <p>8</p>
<p>Experiências de Aprendizagem</p>	<p>O conteúdo programático previsto para cada semana será apresentado na forma de leituras, vídeos ou apresentações assíncronas. Estão previstos quatro encontros síncronos para realização de laboratórios práticos pelos alunos, com a orientação remota do professor. As atividades práticas serão avaliadas e retornadas aos estudantes. Além disso, estão previstos dois encontros síncronos para apresentação de trabalhos por parte dos alunos. Todas as atividades síncronas serão gravadas e disponibilizadas para referência posterior dos alunos.</p>

<p>Critérios de Avaliação</p>	<p>As avaliações serão realizadas de forma remota, distribuídas em formato digital no horário regular da disciplina, com uma janela de tempo de horas para resolução e entrega dos resultados digitalizados por email. Em caso de problemas técnicos os alunos podem informar o professor até um prazo de 72h; neste caso o aluno pode repetir a prova num horário a ser combinado com o professor. As apresentações dos trabalhos acontecem em seminários síncronos online.</p> <p>A avaliação é composta por 4 itens:</p> <ul style="list-style-type: none"> - Parte 1 (20%): Listas de exercícios e tarefas (LET); - Parte 2 (30%): Trabalho final (TF), a ser realizado em grupos - Parte 3 (25%): Prova 1 (P1) na metade do semestre - Parte 4 (25%): Prova 2 (P2) no final do semestre <p>$MG = 0,25*P1 + 0,25*P2 + 0,30TF + 0,20*LET$</p> <p>9,0 <= MG <= 10,0 conceito A (aprovado)</p> <p>7,5 <= MG < 9,0 conceito B (aprovado)</p> <p>6,0 <= MG < 7,5 conceito C (aprovado)</p> <p>MG < 6,0 conceito D.</p> <p>De acordo com a Resolução do CEPE sobre o ERE, durante o período em que perdurar o ERE, fica inaplicável a atribuição de conceito FF, prevista no Parágrafo 2º, do Artigo 44, da Resolução nº 11/2013 do CEPE.</p> <p>Para os estudantes matriculados até o final do período e que deixaram de participar da Atividade de Ensino, deverá ser atribuído o registro NI (Não Informado) no campo de conceito do sistema acadêmico.</p> <p>Para os casos previstos no Parágrafo 1º, a justificativa do registro NI deverá conter a referência ao período de excepcionalidade.</p> <p>Os casos de não informação de conceito durante o ERE, deverão ser resolvidos até o fim do segundo período letivo, após o fim da situação emergencial de saúde.</p>
<p>Atividades de Recuperação Previstas</p>	<p>O aluno que não atingir 60% de aproveitamento poderá realizar uma única prova versando sobre todo o conteúdo desenvolvido na disciplina. Caso a média recalculada levando em conta o novo valor obtido (e o peso da prova substituída) for um conceito final melhor ou igual a C, conforme a tabela apresentada anteriormente, este conceito será o conceito final do aluno. Caso contrário, o aluno receberá o conceito C se a nota da prova de recuperação for >= 6, e conceito D se a nota for < 6.</p>
<p>Bibliografia</p>	<p>Sem alterações</p> <p>Básica Essencial</p> <p>Knuth, Donald Ervin. The art of computer programming. Reading: Addison-Wesley, c1997-c1998. ISBN 0201896834 (v.1); 0201896842 (v.2); 0201896850 (v.3).</p> <p>Básica</p> <p>Azeredo, Paulo Alberto de. Métodos de classificação de dados e análise de suas complexidades. Rio de Janeiro: Campus, c1996. ISBN 8535200045.</p> <p>Cormen, Thomas H; Leiserson, Charles E.; Rivest, Ronald L; Stein, Clifford. Introduction to Algorithms. Cambridge, Massachusetts: The MIT Press, 2009. ISBN 978-0-262-03384-8. Elmasri Elmasri; Shamkant B. Navathe. Fundamentals of database systems. Boston: Pearson, 2016. ISBN 9780133970777.</p> <p>Complementar</p> <p>Nivio Ziviani. Projeto de Algoritmos com implementações em Pascal e C. Sao Paulo: Editora Pioneira, 1999. ISBN 9788522103904.</p> <p>Robert Sedgewick; Kevin Wayne. Algorithms. London: Addison-Wesley Professional, 2011. ISBN 032157351X.</p> <p>Steven S. Skiena. The Algorithm Design Manual. New York: Springer, 1997. ISBN 0387948600.</p> <p>Veloso, Paulo Augusto Silva; Santos, Clesio Saraiva dos; Azeredo, Paulo Alberto de; Furtado, Antonio Luz. Estruturas de dados. Rio de Janeiro: Campus, 1984. ISBN 8570013523.</p>