

Instituto de Informática
Departamento de Informática Aplicada

Dados de identificação

Disciplina: CLASSIFICAÇÃO E PESQUISA DE DADOS

Período Letivo: 2016/2

Período de Início de Validade: 2016/2

Professor Responsável pelo Plano de Ensino: LEANDRO KRUG WIVES

Sigla: INF01124

Créditos: 4

Carga Horária: 60

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.

Currículos

Currículos	Etapa Aconselhada	Natureza
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO	3	Obrigatória
ENGENHARIA DE COMPUTAÇÃO	3	Obrigatória
BACHARELADO EM MATEMÁTICA - ÊNFASE MATEMÁTICA APLIC COMPUTACIONAL	8	Alternativa

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 de arquivos e compactação de arquivos.

Conteúdo Programático

Semana: 1 a 15

Título: Classificação e Pesquisa de Dados

Conteúdo: 1 - Métodos de Classificação de Dados.

- 1.1 - Introdução.
- 1.2 - Formas de apresentação do resultado de uma classificação.
- 1.3 - Famílias de métodos de classificação.
- 1.4 - Método da Inserção Direta.
- 1.5 - Método dos Incrementos Decrescentes (Shellsort).
- 1.6 - Método da Bolha (Bubblesort) e suas variações .
- 1.7 - Método da Partição e Troca (Quicksort).
- 1.8 - Método da Seleção Direta.
- 1.9 - Método da Seleção em Árvore (Heapsort).
- 1.10 - Método da Distribuição de Chave.
- 1.11 - Método da Intercalação (Mergesort).
- 1.12 - Método de Cálculo de Endereços (Hashsort).
- 1.13 - Intercalação de arquivos classificados.

2 - Introdução à Análise da Complexidade de Algoritmos.

- 2.1 - Complexidade dos algoritmos de classificação.
- 2.2 - Complexidade dos algoritmos sobre pesquisa de tabelas.

3 - Pesquisa de Dados em Tabelas.

- 3.1 - Introdução.
- 3.2 - Pesquisa Sequencial.
- 3.3 - Pesquisa Binária.
- 3.4 - Cálculo de Endereços.
- 3.4.1 - Introdução.

- 3.4.2 - Tratamento de Colisões.
 - 3.4.2.1 - Endereçamento Aberto.
 - 3.4.2.1.1 - com Busca Linear.
 - 3.4.2.1.2 - com Realeatorização.
 - 3.4.2.1.3 - Desempenho.
 - 3.4.2.2 - Encadeamento.
 - 3.4.2.2.1 - Com Tabela de Descritores.
 - 3.4.2.2.2 - Com Encadeamento Direto.
 - em Área Separada.
 - com Fusão de Listas.
 - 3.4.2.3 - Alocação em Blocos.
- 3.5 - Hashing Dinâmico.
 - 3.5.1 - Introdução.
 - 3.5.2 - Hashing Estendível.
 - 3.5.3 - Realocação de tabelas.
- 3.6 - Funções de Aleatorização.
 - 3.6.1 - Compressão de Chaves Alfanuméricas.
 - 3.6.2 - Método da Divisão.
 - 3.6.3 - Funções Dependentes da Distribuição.
- 3.7 - Pesquisa em árvore.
 - 3.7.1 - Árvore binária de pesquisa.
 - 3.7.2 - Árvores digitais.
 - 3.7.3 - Árvores-B.
- 4 - Organização de Arquivos.
 - 4.1 - Arquivos seqüenciais.
 - 4.2 - Arquivos indexados.
 - 4.3 - Arquivos invertidos.
 - 4.4 - Arquivos de texto.
- 5 - Compressão de Arquivos.
 - 5.1 - Introdução.
 - 5.2 - Racionalização do projeto de arquivos.
 - 5.3 - Supressão de caracteres repetidos.
 - 5.4 - Codificação de itens.
 - 5.5 - Compressão de seqüências.
 - 5.6 - Códigos de comprimento variável.
 - Huffman.
 - 5.7 – Compressão de Dados Multimídia: Compressão de Imagens

Metodologia

A disciplina será desenvolvida através de aulas expositivas, com explicação dos conteúdos e discussão de exercícios. Espera-se que os alunos tenham uma participação ativa nas aulas, realizando anotações, procurando compreender os conteúdos e, através de perguntas, resolvam suas dúvidas. Além disso, serão utilizadas listas de exercícios, trabalhos e atividades para auto-estudo e avaliação em sala de aula ou no ambiente Moodle (extra classe).

O material apresentado em sala de aula será disponibilizado no ambiente Moodle Institucional ([HTTP://moodleinstitucional.ufrgs.br](http://moodleinstitucional.ufrgs.br)). Nele serão disponibilizados exercícios e atividades extraclasse, notas, especificações de trabalho e demais materiais que sejam necessários. Tal ambiente pode ser acessado por todos os alunos matriculados na disciplina, bastando para tal utilizar o seu número de matrícula (Cartão UFRGS) e sua senha pessoal do Portal do Aluno.

Espera-se que aluno utilize e verifique constantemente o material disponível no ambiente citado, antes e após as aulas. Esse material deve ser utilizado como auxílio ao estudo e como complementação aos apontamentos realizados em aula. Toda comunicação oficial do professor com os alunos será feita APENAS através do Moodle. Solicita-se, portanto, que os alunos mantenham seus e-mails atualizados no sistema da UFRGS.

As 60 horas previstas para atividades teóricas e práticas indicadas neste Plano de Ensino incluem 30 encontros de 100 minutos de duração (2 períodos de 50 minutos por encontro, 2 encontros por semana, durante 15 semanas), num total de 3.000 minutos, e mais 10 horas (600 minutos) de atividades autônomas, realizadas sem contato direto com o professor, correspondentes a exercícios e trabalhos extraclasse, conforme Resolução 11/2013 do CEPE/UFRGS, Artigos 36 a 38. O Professor poderá se valer de aulas presenciais ou à distância (utilização de recursos da EAD), assim como do apoio de Professores Assistentes (Alunos de Pós-Graduação) em Atividades Didáticas.

Carga Horária

Teórica: 60

Prática: 0

Experiências de Aprendizagem

Ao longo do semestre, serão realizadas as seguintes atividades:

- i. Duas provas: P1, na metade do semestre, e P2, como prova final. P1 corresponderá a 25% da nota final; P2 a 35% da nota final;
- ii. Listas de exercícios e tarefas (LET) serão disponibilizados em aula e/ou no ambiente Moodle, e serão contabilizados. A soma de todos os exercícios e tarefas corresponde a 15% da nota final;
- iii. Um trabalho final (TF), a ser realizado em grupos de até dois estudantes, representando 20% da nota final;

Critérios de avaliação

O aluno será avaliado com base no desempenho nas provas, exercícios, trabalhos, atividades e no trabalho final, bem como por sua participação em aula. Conforme regulamento da Universidade, a frequência às aulas é obrigatória.

Ao longo do semestre, serão considerados para a avaliação:

- i. Duas provas: P1, na metade do semestre, e P2, como prova final. P1 corresponderá a 25% da nota final; P2 a 35% da nota final;
- ii. Listas de exercícios e tarefas (LET) serão disponibilizados em aula e/ou no ambiente Moodle, e serão contabilizados. A soma de todos os exercícios e tarefas corresponde a 15% da nota final;
- iii. Um trabalho final (TF), a ser realizado em grupos de até dois estudantes, representando 20% da nota final;
- iv. Além disso, será atribuída nota pela participação (NP) em sala de aula, calculada em função da percentagem de aulas assistidas, o que representará 5% da nota final;
- v. A falta de mais do que 25% das aulas caracteriza o conceito final FF.

A média geral (MG) será obtida por meio da seguinte fórmula:

$$MG = 0,25 \cdot P1 + 0,35 \cdot P2 + 0,2 \cdot TF + 0,15 \cdot LET + 0,05 \cdot NP$$

As provas, trabalhos, exercícios, atividades e projeto final serão avaliados com nota entre 0.0 e 10.0.

A conversão da MG para conceitos é feita por meio da seguinte tabela:

9,0 \leq MG = 10,0 : conceito A (aprovado).

7,5 \leq MG < 9,0 : conceito B (aprovado).

6,0 \leq MG < 7,5 : conceito C (aprovado).

4,0 \leq MG < 6,0 : sem conceito (recuperação), podendo passar para conceito C em caso de aprovação ou conceito D em caso de reprovação na recuperação.

0,0 \leq MG < 4,0 : conceito D (reprovado).

Faltas > 25% : conceito FF (reprovado).

Atenção:

1 - Somente serão calculadas as médias gerais daqueles alunos que tiverem, ao longo do semestre, obtido um índice de frequência igual ou superior a 75% das aulas previstas. Aos que não satisfizerem esse requisito, será atribuído o conceito FF (Falta de Frequência).

2 - Pedidos de revisão de notas só serão avaliados se o aluno tiver pelo menos 75% de presença.

3 - O TF será desenvolvido em diversas etapas, a serem estabelecidas no início das aulas e disponibilizadas no Moodle. Todos os itens do trabalho são importantes e a não entrega ou não participação em qualquer uma das atividades (relatórios, apresentações parciais ou finais) comprometem o seu desenvolvimento, o que influencia na nota do trabalho.

Atividades de Recuperação Previstas

- Os alunos cujas médias gerais forem inferiores a 6,0 (seis) e que satisfizerem a condição 1 acima (ver critérios de avaliação) poderão prestar prova de recuperação, a qual versará sobre toda a matéria da disciplina. Essa prova substitui a menor nota obtida nas provas (mesmo que o aluno obtenha um valor inferior na recuperação). A média é então recalculada levando em conta o novo valor obtido (e o peso da prova substituída), gerando o conceito final conforme a tabela apresentada anteriormente. A prova de recuperação avalia o conteúdo de todas as unidades e sua data está prevista no cronograma disponibilizado pelo professor, no Moodle.

- Não é previsto nenhum tipo de recuperação para os itens LET, TF e NP (ver critérios de avaliação), exceto para os casos previstos na legislação (ver a seguir).

- A recuperação das provas ou atividades individuais será realizada somente para os casos previstos na legislação (saúde, parto, serviço militar, convocação judicial, luto, etc.), devidamente comprovados, através de processo aberto na Junta Médica da UFRGS ou nos respectivos órgãos competentes, conforme o caso. Tendo o direito a recuperação, o professor estipulará a data, o horário e o local de sua realização.

Prazo para Divulgação dos Resultados das Avaliações

As notas das provas e trabalhos serão disponibilizadas em até 20 dias depois de sua realização.

Bibliografia

Básica Essencial

Robert Sedgewick; Kevin Wayne. Algorithms. Addison-Wesley Professional, 2011. ISBN 978-0-321-57351-3. Disponível em: <http://algs4cs.princeton.edu/home/>

Básica

Azeredo, Paulo Alberto de. Métodos de classificação de dados e análise de suas complexidades. Rio de Janeiro: Campus, c1996. ISBN 8535200045.

Cormen, Thomas H.; Charles Eric Leiserson; Ronald Rivest; Clifford Stein. Algoritmos : teoria e prática. Campus, 2012. ISBN 9788535236996. Disponível em: <http://www.submarinocombr.com/produto/111052655/algoritmos-teoria-e-pratica>

Ramez Elmasri; Shamkant Navathe. Fundamentals of Database Systems. Pearson, 2015. ISBN 9780133970777. Disponível em: <https://www.pearsonhighered.com/program/Elmasri-Fundamentals-of-Database-Systems-7th-Edition/PGM189052.html>

Complementar

Knuth, Donald Ervin. The Art of Computer Programming. Addison-Wesley, 2011. ISBN 978-0321751041. Disponível em: <https://www.amazon.com.br/Computer-Programming-Volumes-1-4a-Boxed/dp/0321751043>

Santos, Clesio Saraiva dos; Azeredo, Paulo Alberto de. Tabelas :organização e pesquisa. Porto Alegre: Sagra Luzzatto, 2001. ISBN 8524106484.

Szwarcfiter, Jayme Luiz; Markenzon, Lilian. Estruturas de dados e seus algoritmos. Rio de Janeiro: Livros Tecnicos e Cientificos, 1994. ISBN 8521610149.

Veloso, Paulo Augusto Silva; Santos, Clesio Saraiva dos; Azeredo, Paulo Alberto de; Furtado, Antonio Luz. Estruturas de dados. Rio de Janeiro: Campus, c1984. ISBN 8570013523.

Outras Referências

Não existem outras referências para este plano de ensino.

Observações

Nenhuma observação incluída.