



Plano de Ensino

1. IDENTIFICAÇÃO

Curso: Ciência da Computação

Componente curricular: Tópicos Especiais em Banco de Dados (optativo)

Fase: Décima

Ano/semestre: 2016.1

Número da turma: 13295

Número de créditos: 4

Carga horária - Hora aula: 72

Carga horária - Hora relógio: 60

Professor: Guilherme Dal Bianco

Atendimento ao Aluno: Segundas e Terças-feiras das 18h00 às 19h00

2. OBJETIVO GERAL DO CURSO

O curso tem por objetivo a formação integral de novos cientistas e profissionais da computação, os quais deverão possuir conhecimentos técnicos e científicos e serem capazes de aplicar estes conhecimentos, de forma inovadora e transformadora, nas diferentes áreas de conhecimento da Computação. Adicionalmente, os egressos do curso deverão ser capazes de adaptar-se às constantes mudanças tecnológicas e sociais, e ter uma formação ao mesmo tempo cidadã, interdisciplinar e profissional

3. EMENTA

Mineração de Dados. Aprendizado de Máquina. Map Reduce. Aplicação de mineração de dados em grandes bases de dados.

4. OBJETIVOS

4.1 GERAL

- Entender o processo de mineração de dados e seus algoritmos, principalmente aqueles da área de aprendizado de máquina. Aplicar tais conceitos em grandes volumes de dados.

4.2 ESPECÍFICOS

Apresentar aos estudantes técnicas de mineração de dados utilizando aprendizado de máquina.

- Capacitar os estudantes a utilizarem algoritmos de aprendizado de máquina para minerar dados.
- Capacitar os estudante na utilização do modelo de programação MapReduce
- Capacitar os estudantes no uso da técnica MapReduce para minerar grandes bases de dados



5. CRONOGRAMA E CONTEÚDOS PROGRAMÁTICOS

| | | |
|----|-------|--|
| 1 | 29/02 | Apresentação da disciplina. Tarefas da mineração de dados e definição dos seminários |
| 2 | 1/03 | Introdução ao Modelo MapReduce |
| 3 | 7/03 | Aula ao framework hadoop |
| 4 | 8/03 | Exercícios utilizando o MapReduce |
| 5 | 14/03 | Multiplicação de matrizes utilizando hadoop I |
| 6 | 15/03 | Multiplicação de matrizes utilizando hadoop II |
| 7 | 21/03 | Atividade prática. Descrição dos seminários |
| 8 | 22/03 | Introdução ao PageRank |
| 9 | 28/03 | Atividade sobre PageRank no MapReduce |
| 10 | 29/03 | Desenvolvimento da atividade prática |
| 11 | 04/04 | Apresentação seminários I |
| 12 | 05/04 | Apresentação seminários II |
| 13 | 11/04 | Regras de associação. Métricas de Avaliação (Precisão, recall, F1) |
| 14 | 12/04 | Introdução a ferramenta Weka |
| 15 | 18/04 | Classificação de dados: conceitos básicos e árvores de decisão |
| 16 | 19/04 | Aula prática sobre a ferramenta Weka para classificação de dados. Exercícios. |
| 17 | 25/04 | ERBD 2016 - Trabalho prático |
| 18 | 26/04 | Trabalho em aula (comitê de árvores de decisão) |
| 19 | 02/05 | Apresentação do trabalho |
| 20 | 03/05 | Introdução ao Agrupamento de dados |
| 21 | 09/05 | Algoritmo k-means |
| 22 | 10/05 | Aula Prática: agrupamento de dados |
| 23 | 16/05 | Prova G1. |
| 24 | 17/05 | Entrega das notas e elaboração das questões. |



| | | |
|----|-------|--|
| 25 | 23/05 | Análise de sentimento |
| 26 | 24/05 | Técnicas para análise de sentimento |
| 27 | 30/05 | Atividade prática sobre análise de sentimento |
| 28 | 31/05 | Atividade prática sobre análise de sentimento |
| 29 | 6/06 | Descrição do projeto final I |
| 30 | 7/06 | Descrição do projeto final |
| 31 | 13/6 | Desenvolvimento Projeto final |
| 32 | 14/6 | Seminários de Apresentação do projeto final II |
| 33 | 20/6 | Prova G2 |
| 34 | 21/6 | Entrega das notas. Resolução da prova. |
| 35 | 27/6 | Recuperação G2 |
| 36 | 28/6 | Entrega das notas |

6. PROCEDIMENTOS METODOLÓGICOS

Cada conteúdo da disciplina será organizado em três momentos: exposição de conceitos, implementação de algoritmos e desenvolvimento de atividade avaliativa. Dependendo dos conteúdos da semana, haverá a exposição pelo professor dos conceitos, estruturas de dados e algoritmos relacionados ao tema seguida da avaliação de complexidade e da implementação dos algoritmos. Em alguns casos, a implementação ficará como tarefa extraclasse. O desenvolvimento das atividades avaliativas são momentos para esclarecimento de dúvidas e socialização de resultados. Recomenda-se que todas as atividades propostas sejam trabalhadas pelos estudantes até o início da aula seguinte para que as mesmas possam ser discutidas com a turma. A premissa das atividades avaliativas da disciplina é permitir a prática dos conteúdos no maior tempo possível para que se detectem eventuais dificuldades dos estudantes para atingir os objetivos da disciplina.

7. AVALIAÇÃO DO PROCESSO ENSINO-APRENDIZAGEM

A avaliação tem caráter integrador, uma vez que existe um crescimento gradativo na mobilização dos conhecimentos, habilidades e atitudes desenvolvidos. O resultado é expresso sob a forma de nota que varia de 0 (zero) a 10 (dez), com intervalos de 0,1 (um décimo). É exigido, no mínimo, a média 6,0 (seis) para fins de aprovação na unidade curricular.

O resultado final do aproveitamento é expresso através da média aritmética dos graus G1 e G2, conforme equação (1).

$$Nf = (G1 + G2)/2 \quad (1)$$

Os graus G1 e G2 são compostos pela seguinte equação (2):

$$GX = 0,7*PX + 0,3*TX, \text{ sendo: } PX \text{ a nota da avaliação teórica do Grau 1 ou } 2 \text{ (0 a 10).}$$

$$TX \text{ é a média dos trabalhos do Grau 1 ou 2 (0 a 10).}$$

Em caso de plágio as seguintes regras serão aplicadas.



Prova:

-O aluno recebe nota zero na prova onde o plágio foi detectado, além disso, por demonstrar prática não aceitável o caso será levado ao conhecimento do colegiado;
Trabalhos:

-É permitido usar conteúdo da internet, livros, colegas, etc., contanto que uma citação seja feita. A nota do trabalho será proporcional ao conteúdo original;
-Caso seja detectado plágio o aluno recebe zero no trabalho em questão, além disso, por demonstrar prática não aceitável o caso será levado ao conhecimento do colegiado;

7.1 RECUPERAÇÃO: NOVAS OPORTUNIDADES DE APRENDIZAGEM E AVALIAÇÃO

Tendo em vista que o objetivo do processo de ensino aprendizagem é permitir verificar se, ao final do período letivo, os estudantes possuem as competências e habilidades mínimas necessárias relacionadas aos conteúdos da disciplina, então a recuperação será realizada da seguinte forma: uma prova de substituição de grau contemplando o conteúdo abordado para cada uma das avaliações G1 e G2. Sendo sua nota de grau definida pela equação (4).

$$\text{Grau 1 ou 2} = (\text{NProva ou NRec}) * 0.7 + \text{NTrabalho} * 0.3 \quad (4)$$

Somente será permitida a substituição de um dos graus quando a média for inferior a 7,0. A substituição de grau deve ser realizada na data estabelecida no Calendário Acadêmico. Somente será substituída a nota da prova, se mantendo a nota nos trabalhos. Ao realizar a substituição de grau o aluno atingirá, no máximo, a média final com nota 7,0 (sete).

8. REFERÊNCIAS

8.1 BÁSICA

MEIRA Jr, W., ZAKI, M. J. Data Mining and Analysis: Fundamental Concepts and Algorithms. New York.

Cambridge University Press, 2014.

MITCHELL, T. Machine Learning. New York, McGrawHill Press, 1997

WHITE, TOM. Hadoop: The definitive guide. " O'Reilly Media, Inc.", 2012.

8.2 COMPLEMENTAR

DE AMO S.: Curso Introductorio de Mineração de Dados Compilação de Notas de Aulas, 2006, 201 páginas.

Disponível em www.deamo.prof.ufu.br

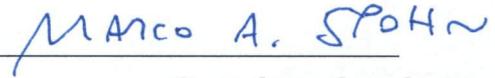
MICHALSKI, R. S., BRATKO, I., KUBAT, M. Machine Learning and Data Mining. Chichester, John Wiley & Sons Ltd, 1998.

TAN, P.N., STEINBACH, M., KUMAR, V. Introduction to Data Mining. New York, AddisonWesley, 2006.



Universidade Federal da Fronteira Sul


Professor



Coordenador do curso

MARCO AURÉLIO SPOHN
Siape nº.1521671
Coord. do Curso de Ciência da Computação
Universidade Federal da Fronteira Sul-UFFS
Campus Chapecó-SC

