



MINISTÉRIO DA EDUCAÇÃO

UNIVERSIDADE FEDERAL DO PARANÁ

SETOR DE CIÊNCIAS SOCIAIS APLICADAS

Coordenação do Curso de ou Departamento de Ciência e Gestão da Informação

Ficha 2 (variável)

Disciplina: Mineração de Dados		Código: SIN195					
Natureza: (X) Obrigatória () Optativa		(X) Semestral () Anual () Modular					
Pré-requisito: não há		Co-requisito: não há		Modalidade: () Totalmente Presencial () Totalmente EAD (X) Parcialmente EAD: 20% (12h) *CH			
CH Total: 60 CH Semanal: 4 Prática como Componente Curricular (PCC): Atividade Curricular de Extensão (ACE):	Padrão (PD): 2	Laboratório (LB): 2	Campo (CP):	Estágio (ES):	Orientada (OR):	Prática Específica (PE):	Estágio de Formação Pedagógica (EFP):
Número de vagas: 30							
Período de oferta: 31/01/2022 - 07/05/2022, conforme Instrução Normativa PROGRAD N°02/2021 de 27/08/2021							
Exame final da disciplina: 09/05/2022							
Justificativa da oferta 20% EaD: Portaria 280/ PROGRAD, de 13/12/2017 que "Aprova Ajuste Curricular no Curso de Gestão da Informação, aplicável à Resolução 64/13-CEPE" (0690698).							
EMENTA							
Aspectos teóricos e práticos da descoberta de conhecimento em bancos de dados utilizando tarefas e técnicas de mineração de dados.							

PROGRAMA

Semana	Data	Carga horária			Conteúdo	Metodologia	Ambiente & Ferramentas	ENTREGAS
		Presencial	Remota	Total				
1	31/01/2022	4	1	5	Apresentação da disciplina, equipes, trabalhos, ferramentas	Aula expositiva e definição da AT1	Sala de aula, UFPR Virtual, DSA, Weka	30/jan
2	07/02/2022	4	0	4	Conceitos básicos	Atividades: AT1 - IGNITE	Sala de aula, UFPR Virtual, Office 365	
3	14/02/2022	4	0	4	KDD e Mineração de Dados; Tarefa de Classificação: ID3 recursivo	Aula expositiva: ID3: AT2	Sala de aula, UFPR Virtual, Office 365	
4	21/02/2022	4	1	5	Tarefa de classificação: C4.5	Aula expositiva: AT5	Sala de aula, UFPR Virtual, Office 366	01/mar
5	28/02/2022	0	4	4	Inteligência Artificial e Ciência de Dados	DSA: AT3 e AT4	Office 365, UFPR Virtual, DSA, Weka	15/mar
6	07/03/2022	4	0	4	Tarefa de Associação WEKA e Excel	AT6	Laboratório, Sala de aula, UFPR Virtual, Office 367, Weka	15/mar
7	14/03/2022	4	0	4	Primeira Avaliação Individual		Sala de aula	
8	21/03/2022	4	0	4	Trabalho Final	TF	Laboratório, Office 365, UFPR Virtual, Weka	
9	28/03/2022	4	0	4	Tarefa de Associação: Apriori	Aula expositiva: Apriori: AT7	Sala de aula, UFPR Virtual, Office 366	05/abr
10	04/04/2022	0	4	4	Tarefa de Associação WEKA e Excel	AT8	Office 365, UFPR Virtual, Weka	12/abr
11	11/04/2022	4	0	4	Tarefa de agrupamento: K-Means	Aula expositiva: AT9	Sala de aula, UFPR Virtual, Office 366	19/abr
12	18/04/2022	4	0	4	Segunda Avaliação Individual		Sala de aula	
13	25/04/2022	2	4	6	Trabalho Final	TF	Laboratório, Office 365, UFPR Virtual, Weka	
14	02/05/2022	4	0	4	Apresentações dos trabalhos	Aula expositiva	Sala de aula, Office 365, UFPR Virtual, Weka	

	48	12	60				
Exame final da disciplina: 09/05/2022							
OBJETIVO GERAL							
O aluno deverá conhecer e ser capaz de aplicar diferentes técnicas estatísticas e computacionais de exploração de dados na identificação de padrões relevantes e não óbvios em bancos de dados.							
OBJETIVOS ESPECÍFICOS							
O aluno deverá ser capaz de:							
1 discutir os conceitos básicos relacionados à mineração de dados;							
2 aplicar e explicar resultados das principais tarefas de mineração de dados;							
3 justificar a relevância desta disciplina;							
4 identificar e compreender aplicações práticas da mineração de dados.							
PROCEDIMENTOS DIDÁTICOS							
Aulas, teórico práticas, expositivo dialogadas em encontros PRESENCIAIS nos quais a docente faz exposição do tema do programa, discute com os alunos os métodos e aplicações e propõe exercícios e atividades em grupo ou individuais, além das atividades avaliativas.							
Serão utilizados como recursos: sala de aula no SA, laboratório de computadores, sala no ambiente virtual de aprendizagem – AVA, mais especificamente o UFPR Virtual; o Microsoft Office 365 disponível institucionalmente, o site da Data Science Academy como suporte ao conteúdo e o software Weka. Para as eventuais aulas não presenciais o material didático será elaborado observando utilização de linguagem dialógica, apresentação e formato próprios para o ensino à distância. Como suporte ao ensino será utilizada a infraestrutura da UFPR tais como: software (Office 365) e AVA. Em atendimento à Resolução 72/10-CEPE, serão observados os seguintes aspectos:							
<ol style="list-style-type: none"> 1. sistema de comunicação – encontros presenciais, AVA (UFPR Virtual), correios eletrônicos trocados entre coordenação e alunos e professor e alunos; 2. modelo de tutoria a distância e presencial – não está prevista a tutoria nesta disciplina. As dúvidas serão esclarecidas pela docente e, quando existir, pelos alunos do programa de Monitoria ou alunos de PPGs em estágios docência; 3. material didático específico – o material didático será elaborado observando utilização de linguagem dialógica, apresentação e formato próprios para o ensino presencial; 4. infraestrutura de suporte tecnológico, científico e instrumental à disciplina – o curso conta com Ambiente Virtual de Aprendizagem específico e pacote Office 365, incluindo sala no Teams para eventuais esclarecimentos e interação; 5. previsão de período de ambientação dos recursos tecnológicos a serem utilizados pelos discentes: os alunos recebem capacitação no Ambiente de Aprendizagem Virtual no primeiro semestre do Curso e esta é uma disciplina de sexto período. Assume-se que os alunos já tenham domínio no uso das tecnologias. 							
TODAS AS AVALIAÇÕES (PROVAS) SERÃO INDIVIDUAIS E PRESENCIAIS.							
As atividades podem ser presenciais ou remotas, conforme definição em sala de aula e rubricas.							
FORMAS DE AVALIAÇÃO E CONTROLE DE FREQUÊNCIA							
Algumas etapas avaliativas serão assíncronas e outras presenciais. Os critérios das atividades e avaliações estão definidos (ao final desta Ficha 2) e serão publicados no ambiente virtual, no primeiro dia de aula, para todos os alunos. As três notas principais para composição da média final (cálculo da média aritmética simples das três notas), são:							
<ol style="list-style-type: none"> a. primeira nota: média aritmética das atividades (AT) realizadas individualmente ou em equipe (conforme solicitado) com avaliação conforme Rubricas 1, 2 e 3; b. segunda nota: média aritmética das notas das avaliações (AV) individuais com avaliação conforme Rubrica 5; c. terceira nota: trabalho (TF) em dupla / individual sobre um dos temas disponíveis com avaliação conforme Rubrica 4 que considerará: <ul style="list-style-type: none"> • profundidade na abordagem do tema; • qualidade do levantamento bibliográfico; • contextualização do tema na área de interesse; • explicação do funcionamento do algoritmo em questão; • exemplificações práticas e contextualizadas do tema; • análises dos resultados obtidos; • criatividade da apresentação; • uso de normas para elaboração de trabalhos de pesquisa (normas ABNT); • desempenho individual e coletivo dos membros da equipe durante a apresentação. 							
$MédiaFinal = \frac{\sum_{i=1}^9 AT_i + \sum_{i=1}^2 AV_i + TF}{3}$							
As frequências no AVA serão calculadas conforme a entrega das atividades e avaliações propostas.							
As demais frequências serão controladas por chamada em sala de aula.							
O limite de ausências permitido é 25% da carga horária da disciplina.							
BIBLIOGRAFIA BÁSICA (mínimo 03 títulos)							
Castro, Leandro Nunes de; FERRARI, Daniel Gomes. Introdução à mineração de dados: conceitos básicos, algoritmos e aplicações. São Paulo: Editora Saraiva, 2016. 978-85-472-0100-5. Disponível em: https://integrada.minhabiblioteca.com.br/#/books/978-85-472-0100-5/ . Acesso em: 16 abr. 2021.							
NG, Andrew. Machine learning learning. Disponível em: https://www.deeplearning.ai/machine-learning-yearning/ . Acesso em: 22 jun. 2020.							
SILVA, Leandro Augusto da; PERES, Sarajane Marques; BOSCARIOLI, Clodis. Introdução à mineração de dados - com aplicações em R. Rio de Janeiro: Grupo GEN, 2016. 9788595155473. Disponível em: https://integrada.minhabiblioteca.com.br/#/books/9788595155473/ . Acesso em: 16 abr. 2021							
BIBLIOGRAFIA COMPLEMENTAR (mínimo 05 títulos)							
BURKOV, Andriy. The hundred-page machine learning book. Disponível em: http://themlbook.com/wiki/doku.php . Acesso em: 14 abr. 2021.							
DEISENROTH, Marc Peter; FAISAL, A. Aldo e ONG, Cheng Soon. Mathematics for machine learning. Disponível em: https://mml-book.github.io . Acesso em: 16abr. 2021.							

GOOGLE. **Machine learning crash course**. Disponível em: <https://developers.google.com/machine-learning/crash-course/>. Acesso em: 15 abr. 2021.

POOLE, David L.; MACKWORTH, Alan K. **Artificial intelligence: foundations of computational agentes**. 2ª ed. Disponível em: <https://artint.info/2e/html/ArtInt2e.html>. Acesso em: 22 mar. 2021.

SHALEV-SHWARTZ, Shai; BEN-DAVID, Shai. **Understanding machine learning**. Disponível em: <https://www.cse.huji.ac.il/~shais/UnderstandingMachineLearning/>. Acesso em: 22 mar. 2021.

SUTTON, Richard S.; BARTO, Andrew G. **Reinforcement learning: an introduction**. 2ª ed. Disponível em: <http://incompleteideas.net/book/RLbook2018.pdf>. Acesso em: 22 mar. 2021.

Vídeos

RAMESH, Raj. **What is artificial intelligence?** In 5 minutes. Disponível em: <https://www.youtube.com/watch?v=2ePf9rue1Ao&t=155s>. Acesso em: 16 abr. 2021.

Ventury city. **Amazon's city of the future**. Disponível em: <https://www.youtube.com/watch?v=ZjraUKJRIt8>. Acesso em: 15 abr. 2021.

Demais materiais serão disponibilizados no UFPR Virtual, conforme a necessidade

Docente da Disciplina:

Prof. Drª. Denise Fukumi Tsunoda - E-mail: dtsunoda@ufpr.br

Rubricas

RUBRICA 1				
Atividade (AT1): Criar e apresentar (pode ser por meio de narração) um IGNITE sobre o assunto sorteado no encontro síncrono. O conteúdo deve contemplar: histórico, aplicações, exemplos e ferramentas. Atenção: a apresentação pode ser um vídeo!				
Critérios de avaliação	Atende plenamente	Atende suficientemente	Atende parcialmente	Não atende
Envia o IGNITE na plataforma	Apresenta todos os 4 itens solicitados, devidamente explicados.	Apresenta todos os 4 itens solicitados, entretanto, não estão devidamente explicados.	Apresenta apenas 3 ou menos itens.	Envio não realizado
Pontos	50	35	20	0
Apresenta o IGNITE	Utilizando template ou apresentação temporizada com 1 ou 2 conteúdos textuais (os demais são tópicos ou ilustrações).	Utilizando template ou apresentação temporizada com 3 ou mais conteúdos textuais.	Sem apresentação temporizada.	Não apresenta.
Pontos	30	23	13	0
Apresentação narrada ou síncrona	Apresenta com domínio de conteúdo, com 1 ou 2 leituras de texto.	Apresenta com domínio de conteúdo, com 3 ou mais leituras de texto.	Apresenta com pouco domínio de conteúdo.	Não apresenta.
Pontos	10	5	2	0
Linguagem	A apresentação possui 1 ou 2 deslizes gramaticais (ortografia, pontuação, concordância e acentuação).	A apresentação possui 3 ou 4 deslizes gramaticais (ortografia, pontuação, concordância e acentuação).	A apresentação possui mais de 4 deslizes gramaticais (ortografia, pontuação, concordância e acentuação).	Não realiza envio.
Pontos	10	7	5	0
	100	70	40	0
Descontos por envio da atividade em atraso	Entrega no prazo	Até 2 dias de atraso	De 3 a 5 dias de atraso	Mais de 5 dias de atraso
Pontos	0	-5	-10	-15
RUBRICA 2				
Atividade (AT3 e AT4): realizar os cursos e enviar o certificado pelo AVA				

Critérios de avaliação	Atende plenamente	Atende suficientemente	Atende parcialmente	Não atende
Envia o certificado com a nota	Recebe entre 86 e 100 pontos	Recebe entre 70 e 85 pontos	Envia o print da conclusão (sem avaliação)	Não realiza envio
Pontos	100	80	50	0

RUBRICA 3

Atividade (AT2 e AT5-AT9): cada questão das atividades será avaliada considerando os critérios abaixo definidos.

Critérios de avaliação	Atende plenamente	Atende suficientemente	Atende parcialmente	Não atende
Questões: entrega	Entrega no prazo	Até 2 dias de atraso	De 3 a 5 dias de atraso	Acima de 5 dias de atraso ou sem entrega.
Pontos	20	15	5	0
Questões: acerto	Entrega a atividade com a resposta correta	Entrega a atividade com a resposta incorreta com o detalhamento da solução	Entrega a atividade com a resposta correta e sem o detalhamento da solução	Entrega a atividade sem registro de resposta
Pontos	30	20	10	5
Questões: detalhamento	Entrega a atividade com o passo-a-passo da solução, devidamente explicados	Entrega a atividade com o passo-a-passo da solução, entretanto, não estão devidamente explicados	Entrega a atividade com 2 ou mais passos ausentes	Entrega a atividade sem o detalhamento da solução
Pontos	50	35	25	5
	100	70	40	10

RUBRICA 4

TRABALHO FINAL

ATIVIDADE INDIVIDUAL OU DUPLAS

Data de entrega: até 28/04/2021 - 23h deverão ser enviados os itens abaixo COMPACTADOS EM UM UNICO ARQUIVO PELO AVA!

Definição:

- Buscar uma base de dados para realizar a atividade OU escolher uma das listadas e divulgadas em sala de aula;
- Adaptar (pré-processamento) a base de dados à entrada no Weka e realizar a análise utilizando UM método de cada categoria recebida;
- Redigir relatório detalhando o experimento (de 15 a 20 páginas de conteúdo) e comparando os resultados encontrados.

Critérios de avaliação	Atende plenamente	Atende suficientemente	Não parcialmente	Não atende
Introdução	A introdução define integralmente o problema e objetivos do trabalho	A introdução define parcialmente o problema e objetivos do trabalho	A introdução é superficial.	Não existe introdução.
Pontos	15	10	5	0
Descrição detalhada das etapas do utilizadas no trabalho, método, bases	As etapas estão explicitadas em detalhes e permitem reprodutibilidade de 80 a 100% do experimento.	As etapas estão explicitadas em detalhes e permitem reprodutibilidade de 50 a 79% do experimento.	As etapas estão explicitadas em detalhes e permitem reprodutibilidade de até 49% do experimento.	As etapas não estão explicitadas.

Pontos	30	25	20	0
Análise dos resultados	Apresenta os resultados de forma consistente e com 4 ou mais de contribuições relevantes.	Apresenta os resultados de forma consistente e com menos de 4 contribuições.	Apresenta os resultados de forma consistente, mas sem relevância.	Não apresenta resultados consistentes ou apenas prints das telas de resultados sem análises.
Pontos	20	15	10	0
Considerações finais	Redige de forma clara as considerações com ponderações pertinentes.	Apresenta as considerações, mas não estabelece comparações ou realiza ponderações.	Apresenta as considerações sem grande relevância.	Não apresenta as considerações finais.
Pontos	20	15	10	0
Linguagem	O documento possui 1 ou 2 deslizes gramaticais (ortografia, pontuação, concordância e acentuação).	O documento possui 3 ou 4 deslizes gramaticais (ortografia, pontuação, concordância e acentuação).	O documento possui mais de 4 deslizes gramaticais (ortografia, pontuação, concordância e acentuação).	Não realiza envio.
Pontos	5	3	1	0
Normas ABNT / formato de entrega (AVA) / Prazo	Observa, integralmente, os 3 itens	Observa, parcialmente, os 3 itens	Observa 2 dos 3 itens	Observa menos de 2 itens
Pontos	10	7	4	0
	100	75	50	0

RUBRICA 5

Avaliação (AV1 - AV2): cada questão será avaliada conforme critérios abaixo.				
Critérios de avaliação	Atende plenamente	Atende suficientemente	Atende parcialmente	Não atende
Questões: acerto	Entrega a atividade com a resposta correta com o detalhamento da solução	Entrega a atividade com a resposta incorreta com o detalhamento da solução	Entrega a atividade com a resposta correta e sem o detalhamento da solução	Entrega a atividade sem registro de resposta
Pontos	50	35	25	5
Questões: detalhamento	Entrega a atividade com o passo-a-passo da solução, devidamente explicados	Entrega a atividade com o passo-a-passo da solução, entretanto, não estão devidamente explicados	Entrega a atividade com 2 ou mais passos ausentes	Entrega a atividade sem o detalhamento de todos os passos da solução
Pontos	50	35	25	5
	100	70	50	10



Documento assinado eletronicamente por DENISE FUKUMI TSUNODA, PROFESSOR DO MAGISTERIO SUPERIOR, em 23/12/2021, às 08:42, conforme art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por JOSE MARCELO ALMEIDA PRADO CESTARI, CHEFE DO DEPARTAMENTO DE CIENCIA E GESTAO DA INFORMACAO - SA, em 03/01/2022, às 14:21, conforme art. 1º, III, "b", da Lei 11.419/2006.



Documento assinado eletronicamente por LUCIANO HEITOR GALLEGOS MARIN, VICE / SUPLENTE COORDENADOR DO CURSO DE GESTAO DA INFORMACAO, em 05/01/2022, às 10:44, conforme art. 1º, III, "b", da Lei 11.419/2006.



A autenticidade do documento pode ser conferida [aqui](#) informando o código verificador 4021913 e o código CRC 7E32EF99.

