



Ministério da Educação
UNIVERSIDADE FEDERAL DO PARANÁ
Setor de Ciências Exatas
Departamento de Estatística

Ficha 2 (variável)

Disciplina: Machine Learning					Código: CE042	
Natureza: () Obrigatória (x) Optativa		(x) Semestral () Anual () Modular				
Pré-requisito: não há		Co-requisito: não há		Modalidade: Presencial		
CH Total: 60	Padrão	Laboratório	Campo (CP):	Estágio	Orientada	Prática Específica
CH semanal: 4	(PD): 60	(LB): 0	0	(ES): 0	(OR): 0	(PE): 0
Início da oferta: 02/02/22						
Fim da oferta: 06/05/22						
EMENTA (Unidade Didática)						
Métodos de reamostragem, aprendizado não supervisionado: análise de componentes principais, k-means, dbscan. Aprendizado supervisionado: regressão logística, ridge, lasso e elastic net, bagging, boosting, random forest e support vector machines.						
PROGRAMA						
<ol style="list-style-type: none"> 1. Introdução <ul style="list-style-type: none"> • O que é Machine Learning; • Definição dos objetivos; • Dados faltantes e criação de características; • Dados desbalanceados e variáveis correlacionadas; • Como as máquinas aprendem. 2. Métodos de reamostragem <ul style="list-style-type: none"> • Tipos de função custo; • Amostras de treinamento, validação e teste; • Validação cruzada K-fold; • Validação cruzada certo e errado. 3. Aprendizado não supervisionado <ul style="list-style-type: none"> • Análise de componentes principais; • K-means e cluster hierárquico; • DBSCAN. 4. Classificação <ul style="list-style-type: none"> • Regressão logística; • Análise de discriminante linear e quadrática; • Comparação de modelos, Naive Bayes e tipos de erro; 5. Regularização <ul style="list-style-type: none"> • Problemas de “small n, large p” e Regressão penalizada; • Penalização Ridge, Lasso, Elastic net e Horseshoe; • Seleção dos parâmetros de tuning e Regressão com componentes principais. 6. Métodos baseados em árvores <ul style="list-style-type: none"> • Árvores de decisão; • Bootstrap Aggregation (Bagging) e Random Forest; • Boosting. 7. Support vector machines <ul style="list-style-type: none"> • Maximal Margin Classifier e Support Vector Classifier; • Expansão das características e truque do Kernel. 8. Gradiente descendente <ul style="list-style-type: none"> • Gradiente descendente batch e estocástico; • Gradiente boosting. 						

OBJETIVO GERAL

Proporcionar ao discente o conhecimento sobre o aprendizado de máquina em um contexto supervisionado e não supervisionado.

OBJETIVO ESPECÍFICO

O aluno deve demonstrar que compreende os conceitos de aprendizagem de máquina, suas aplicações e limitações.

FORMAS DE AVALIAÇÃO

Os alunos serão avaliados em três dimensões:

1. Desempenho na prova escrita (peso: 40%) (0 a 100).
2. Desempenho na atividade prática (peso: 30%) (0 a 100).
3. Desempenho no trabalho final (peso: 30%) (0 a 100).

Serão apresentados aos alunos na página da disciplina:

- Calendário das avaliações, com as datas, horários e objetivos;
- Tipo de avaliação que será realizada;
- Sistema de aprovação (médias das avaliações, frequência e avaliações, etc.)

Critérios para aprovação:

- Frequência de pelo menos 75% e *Nota* igual ou acima de 70 → Aprovação sem Exame Final.
- Frequência de pelo menos 75% e *Nota* entre 40 e 70 → Exame Final.
- Média entre: *Nota* e Exame Final igual ou acima de 50 → Aprovação.
- *Nota* inferior a 40 ou presença inferior a 75% → Reprovação.
- Média entre: *Nota* e Exame Final inferior a 50 → Reprovação.

BIBLIOGRAFIA BÁSICA (mínimo 03 títulos)

- James, G., Witten, D., Hastie, T. e Tibshirani, An Introduction to Statistical Learning, 2013, Unofficial Solutions;
- Hastie, T., Tibshirani, R. e Friedman, J., [The Elements of Statistical Learning](#), 2009;
- <http://cursos.leg.ufpr.br/ML4all/>
- <http://www.rizbicki.ufscar.br/ame/>

BIBLIOGRAFIA COMPLEMENTAR (mínimo 05 títulos)

1. Tan, Steinbach, and Kumar, Introduction to Data Mining, Addison-Wesley, 2005.
2. Lantz, B., Machine Learning with R, Packt Publishing, 2013;
3. **Pattern Recognition and Machine Learning;**
<https://www.microsoft.com/en-us/research/publication/pattern-recognition-machine-learning/>
4. <https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/understanding-machine-learning-theory-algorithms.pdf>
5. http://martin.zinkevich.org/rules_of_ml/rules_of_ml.pdf

Professor da Disciplina: Eduardo Vargas Ferreira

Contato do professor da disciplina (e-mail): e.ferreira@ufpr.br

Chefe de Departamento ou Unidade equivalente: PAULO JUSTINIANO RIBEIRO JUNIOR