



CURSO DE CIÊNCIA DE DADOS APLICADA AO PODER JUDICIÁRIO



MACHINE LEARNING

FICHA DO CURSO

Curso: Machine Learning**Modalidade:** EAD - autoinstrucional**Carga horária total:** 40 horas**Carga horária semanal:** 4 horas**Início da oferta:****Fim da oferta:****Pré-requisito:** Recomenda-se que o cursista tenha conhecimentos em fundamentos de estatística e em R intermediário.**Conteudistas:** Anderson Ara**1. Objetivos**

Proporcionar ao(à) aluno(a) conhecimento teórico e prático referente aos principais métodos de aprendizado de máquina não supervisionados e supervisionados. Habilitar o(a) aluno(a) a utilizar o software R e/ou Python para implementação de algoritmos de aprendizado de máquina.

Ao final do curso o participante deverá demonstrar capacidade para realizar tarefas, tais como: análise de agrupamento e redução de dimensionalidade, regressão e classificação.

2. Programa

Conceitos iniciais, Histórico e Aplicações. Análise de Agrupamentos: K-médias. Redução de Dimensionalidade: PCA e t-SNE. Classificação: Regressão Logística, KNN, Árvores de Decisão, Redes Neurais. Regressão: Regressão Linear, KNN, CART, Redes Neurais. Métodos ensemble: Florestas Aleatórias.

3. Procedimentos didáticos

A cada módulo (em um total de **10 módulos**) serão desenvolvidas as seguintes atividades pelos cursistas:

1. Acessar o material pré aula para preparar-se para a aula.
2. Assistir as vídeos-aulas programadas para o módulo
3. Fazer o estudo individual de materiais indicados como: leituras complementares, resolução de exercícios e acesso a vídeos adicionais.
4. Realizar as atividades de avaliação do módulo no formato de quiz.

Como suporte ao Ensino Remoto será utilizada a infraestrutura da plataforma moodle do Centro de Formação e Aperfeiçoamento de Servidores do Poder Judiciário (CEAJUD).

4. Procedimentos metodológicos

Conjunto de vídeo-aulas separadas por módulos de conhecimento acompanhado de avaliações para reforçar o aprendizado.

5. Distribuição da carga horária do cursista

A Tabela 1 a seguir apresenta distribuição da carga horária semanal total do cursista, de 4 horas por semana, nas atividades semanais previstas na vigência do curso.

Tabela 1 - Carga horária semanal do(a) cursista.

Atividade semanal	Carga horária (horas)	Fração (%)
Estudo individual de conteúdo pré-aula	00:30	12,5%
Assistir vídeo-aulas	02:00	50%
Estudo individual pós aula	00:30	12,5%
Atividades avaliativas	01:00	25%
Total	04:00	100%

6. Conteúdo programado

O curso está dividido em 10 módulos sendo que cada módulo configura uma semana de atividades. O conteúdo programado é apresentado na Tabela 2.

Tabela 2 - Cronograma detalhado do conteúdo das unidades didáticas.

Módulo	Conteúdo programado	Duração
1	Conceitos iniciais, Histórico e Aplicações	4 horas
2	Análise de agrupamentos: k-médias Definição, algoritmo, avaliação da escolha do número de grupos	4 horas
3	Redução de dimensionalidade: PCA e t-SNE Definição. Métodos de análise de componentes principais e incorporação de vizinho estocástico distribuído em T Diferenças entre os métodos	4 horas
4	Classificação: Regressão Logística Introdução à Classificação. Modelo de Regressão Logística Medidas de Desempenho. Métodos de Validação	4 horas
5	Classificação: KNN e Árvores de Decisão Algoritmos KNN, ID3 e C4.5. Tunagem	4 horas
6	Classificação: Redes Neurais Origem. Perceptron e Multilayer Perceptron. Gradiente Estocástico. Backpropagation	4 horas
7	Regressão: Regressão Linear Introdução à Regressão. Modelo de Regressão Múltipla e Regressão Polinomial	4 horas
8	Regressão: CART Árvores de Regressão e Classificação	4 horas
9	Regressão: KNN e Redes Neurais Adaptação KNN e Redes Neurais para regressão	4 horas

10 Métodos ensemble: Florestas Aleatórias Bagging. Florestas Aleatórias. Tunagem	4 horas
---	---------

Encerramento do curso

7. Desempenho no curso

O desempenho no curso será determinado pela nota em atividade avaliativa aplicada aos cursistas. Requer-se, no mínimo, 70% de aproveitamento para obtenção de certificado.

8. Referências bibliográficas

1. Faceli, K., Lorena, A. C., Gama, J., & Carvalho, A. C. P. L. F. (2011). **Inteligência artificial. Uma Abordagem de Aprendizado de Máquina**. Editora LTC.
2. Friedman, J., Hastie, T., Tibshirani, R. (2001). **The elements of statistical learning** (Vol. 1, pp. 337-387). New York: Springer series in statistics.
3. Hair, J. F., Anderson, R. E., Babin, B. J., & Black, W. C. (2010). **Multivariate data analysis: A global perspective** (Vol. 7).
4. Witten, I. H., Frank, E., Hall, M. A., Pal, C. J. (2016). **Data Mining: Practical machine learning tools and techniques**. Morgan Kaufmann.

9. Informações sobre os conteudistas

Anderson Ara

Graduado em Estatística (2009), Mestre em Estatística (2011), títulos obtidos pela Universidade Federal de São Carlos (UFSCar). Doutor em Estatística (2016) através dos Programas de Pós-graduação em Estatística (PPGEst-UFSCar) e Pós-graduação em Ciência da Computação (PPG-CC-UFSCar). Desde agosto de 2021 é Professor Adjunto da Universidade Federal do Paraná (UFPR), campus Curitiba/PR, Departamento de

Estatística (DEst) do Setor de Ciências Exatas. Foi Professor Adjunto da Universidade Federal da Bahia (2017-2021), campus Salvador/BA, Departamento de Estatística (DEst) do Instituto de Matemática e Estatística (IME) e Professor da Faculdade de Tecnologia SENAI-SP (2009-2015), campus São Carlos/SP. Docente da Especialização em Data Science & Big Data (DSBD-UFPR), MBA em Finanças Corporativas (UTFPR) e Ciência de Dados e Big Data (ECD-UFBA). Pesquisador do Programa de Pós-Graduação em Métodos Numéricos em Engenharia (PPGMNE) da UFPR pela linha de pesquisa Métodos Estatísticos. Pesquisador do Programa de Pós-Graduação em Matemática (PGMAT) com área de Concentração em Estatística do IME-UFBA. Em atual colaboração em pesquisas desenvolvidas na área de Ciência de Dados no Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS). Atua principalmente nas seguintes áreas: Aprendizado Estatístico de Máquina, Inferência Estatística e Métodos Computacionais. Tem orientado e publicado em periódicos nacionais e internacionais da área. Tem experiência no desenvolvimento de projetos multidisciplinares de pesquisa, extensão e de desenvolvimento tecnológico.

10. Requisitos Técnicos

Computador com acesso à internet. Permissão para instalar programas.

CURSO DE CIÊNCIA
DE DADOS APLICADA
AO PODER JUDICIÁRIO

