

CARGA HORÁRIA TOTAL : 88h

EMENTA

A disciplina explora o escopo e a natureza multidisciplinar da ciência de dados com foco na solução de problemas usando dados em várias áreas. Estuda como as ferramentas analíticas podem ser usadas para descobrir padrões e significado nos dados. Desenvolve a mentalidade exploradora de estruturas de ciência de dados, podendo ser aplicadas a qualquer setor, empresa ou organização.

COMPETÊNCIAS

I – ANALISAR E RESOLVER PROBLEMAS

VII – PENSAMENTO VOLTADO A CIÊNCIA EXATAS E NATURAIS

VIII – PENSAMENTO LÓGICO

XIII - DOMÍNIO DA TECNOLOGIA DA INFORMAMAÇÃO E COMUNICAÇÃO

OBJETIVOS DE APRENDIZAGEM

- Compreender a relação entre a estatística, a ciência da computação e a ciência dos dados.
- Conhecer sobre uma breve história da evolução da estatística, ciência da computação e ciência dos dados.
- Saber das principais linguagens de programação usadas hoje na estatística e na ciência dos dados.
- Conhecer sobre as fases do desenvolvimento de modelos preditivos.
- Entender o estudo de caso proposto e formar paralelos com situações similares.
- Conhecer sobre o propósito do desenvolvimento de um modelo preditivo quantitativo.
- Conhecer sobre análise descritiva de dados e suas principais técnicas.
- Entender sobre dados estruturados e não estruturados, dados retangulares e data-frames.
- Conhecer sumários estatísticos e principais técnica gráficas de visualização de dados.
- Conhecer sobre modelos de regressão linear simples e múltipla.
- Entender o princípio de funcionamento da técnica de regressão linear.
- Saber reconhecer situações similares onde aplicar regressão linear como modelo preditivo.
- Conhecer sobre os paradigmas de aprendizagem supervisionada e não supervisionada.
- Entender o princípio de funcionamento dessas duas diferentes abordagens e seus históricos.
- Saber reconhecer a diferença entre os dois principais tipos de aprendizagem supervisionada: regressão e classificação.
- Conhecer sobre modelos de regressão logística simples e múltipla.
- Entender o princípio de funcionamento da regressão logística e outros classificadores.
- Conhecer o estudo de caso exemplo de predição de inadimplência com cartões de crédito.
- Conhecer sobre desenvolvimento de modelos de regressão logística simples.
- Consolidar o entendimento sobre o princípio de funcionamento da técnica de regressão logística.
- Realizar a análise descritiva dos dados do estudo de caso em questão: predição de inadimplência com cartões.
- Conhecer sobre modelos de regressão logística múltipla.
- Entender o princípio de funcionamento da técnica de regressão logística múltipla.
- Saber reconhecer situações similares onde aplicar regressão logística múltipla como modelo preditivo qualitativo.
- Adquirir uma visão geral sobre os principais métodos de aprendizagem supervisionada.
- Ter uma noção inicial sobre o princípio de funcionamento de cada dos principais métodos de aprendizagem supervisionada.
- Saber reconhecer situações onde aplicar tais métodos, na ciência ou nos negócios.
- Conhecer sobre modelos de árvores de decisão para regressão e classificação.
- Entender o princípio de funcionamento das árvores de regressão e de classificação.
- Saber reconhecer situações onde aplicar árvores de decisão e suas principais vantagens.
- Conhecer sobre modelos preditivos com base em árvores de decisão para classificação.
- Conhecer sobre o potencial de aplicação dessa técnica tomando um estudo de caso como exemplo.
- Saber reconhecer situações similares onde aplicar árvores de decisão para classificação.
- Adquirir uma visão geral sobre métodos de aprendizagem não supervisionada.
- Entender o princípio de funcionamento das técnicas de aprendizagem não supervisionada.
- Saber reconhecer situações similares onde aplicar os métodos de aprendizagem não supervisionada.
- Conhecer sobre os principais algoritmos de agrupamento (clusterização).
- Entender o princípio de funcionamento dos algoritmos de k-médias e de clusterização hierárquica.
- Saber reconhecer situações similares onde aplicar algoritmos de agrupamento.

- Conhecer sobre agrupamento hierárquico e suas aplicações.
- Entender o princípio de funcionamento da técnica de agrupamento hierárquico.
- Saber reconhecer situações similares onde aplicar agrupamento hierárquico.

CRONOGRAMA DE AULA	
<p>Unidade 1</p> <ul style="list-style-type: none"> • Estatística, Machine Learning e Ciência dos Dados • Predição com Regressão Linear - Estudo de Caso • Análise Descritiva dos Dados • Predição com Modelos de Regressão Linear 	<p>Objetivos de Aprendizagem</p> <ul style="list-style-type: none"> ▪ Compreender a relação entre a estatística, a ciência da computação e a ciência dos dados. ▪ Conhecer sobre uma breve história da evolução da estatística, ciência da computação e ciência dos dados. ▪ Saber das principais linguagens de programação usadas hoje na estatística e na ciência dos dados. ▪ Conhecer sobre as fases do desenvolvimento de modelos preditivos. ▪ Entender o estudo de caso proposto e formar paralelos com situações similares. ▪ Conhecer sobre o propósito do desenvolvimento de um modelo preditivo quantitativo. ▪ Conhecer sobre análise descritiva de dados e suas principais técnicas. ▪ Entender sobre dados estruturados e não estruturados, dados retangulares e data-frames. ▪ Conhecer sumários estatísticos e principais técnica gráficas de visualização de dados. ▪ Predições com modelos de regressão linear: Predição de valor de imóveis com regressão linear simples e múltipla.
	<p>Estratégias de Ensino</p> <p>Utilização de material referencial em diferentes formatos: vídeos, textos de referência conceitual, atividades de pesquisa, estudos de caso, infografias interativas, entre outros.</p> <p>Sequência sugerida:</p> <ul style="list-style-type: none"> ✓ Explorar a seção “Inspire-se” que contextualiza o tema da unidade e traz informações de tendências e inovações na respectiva área de conhecimento, aplicação prática ou estudos de caso, depoimentos ou entrevistas com profissionais qualificados do mercado de trabalho. ✓ Conhecer e entender os conceitos básicos da unidade apresentados na seção “Explore”. Neste material são apresentados os aspectos teóricos, exemplos práticos e conteúdos complementares que ampliam o conhecimento sobre as temáticas da unidade. Explorar os vídeos e infografias interativas.
	<p>Atividade</p> <p>Atividade não pontuada disponível na seção “Pratique e Compartilhe”.</p> <ul style="list-style-type: none"> ✓ Estudos de caso, resoluções, proposta de pesquisa ou produção criativa que integram atividades práticas aos conceitos teóricos básicos da unidade. ✓ As respostas e resultados da atividade proposta devem ser postados no fórum disponível na sessão “Compartilhe”. ✓ Após a postagem será disponibilizado feedback com modelo de resposta).
	<p>Avaliação Formativa</p> <p>Realizar a “Atividade Avaliativa” que constitui o recurso de avaliação pontuada da unidade. A pontuação desta atividade fará parte da nota final na N1 (ver item “Avaliação” deste plano”).</p>
	<p>Objetivos de Aprendizagem</p> <ul style="list-style-type: none"> ▪ Conhecer sobre os paradigmas de aprendizagem supervisionada e não supervisionada. ▪ Entender o princípio de funcionamento dessas duas diferentes abordagens e seus históricos. ▪ Saber reconhecer a diferença entre os dois principais tipos de aprendizagem supervisionada: regressão e classificação. ▪ Conhecer sobre modelos de regressão logística simples e múltipla.
<p>Unidade 2</p> <ul style="list-style-type: none"> • Tipos de Aprendizagem de Máquina • Estudo de Caso - Predição de Inadimplência • Análise Descritiva dos Dados 	

<ul style="list-style-type: none"> Predições com Modelos de Regressão Logística 	<div> <ul style="list-style-type: none"> Conhecer sobre desenvolvimento de modelos de regressão logística simples. Consolidar o entendimento sobre o princípio de funcionamento da técnica de regressão logística. Realizar a análise descritiva dos dados do estudo de caso em questão: predição de inadimplência com cartões. Entender o princípio de funcionamento da regressão logística e outros classificadores. Conhecer o estudo de caso exemplo de predição de inadimplência com cartões de crédito Conhecer sobre modelos de regressão logística múltipla. Entender o princípio de funcionamento da técnica de regressão logística múltipla. Saber reconhecer situações similares onde aplicar regressão logística múltipla como modelo preditivo qualitativo. </div> <div> <p>Estratégias de Ensino</p> <p>Utilização de material referencial em diferentes formatos: vídeos, textos de referência conceitual, atividades de pesquisa, estudos de caso, infografias interativas, entre outros.</p> <p>Sequência sugerida:</p> <ul style="list-style-type: none"> ✓ Explorar a seção “Inspire-se” que contextualiza o tema da unidade e traz informações de tendências e inovações na respectiva área de conhecimento, aplicação prática ou estudos de caso, depoimentos ou entrevistas com profissionais qualificados do mercado de trabalho. ✓ Conhecer e entender os conceitos básicos da unidade apresentados na seção “Explore”. Neste material são apresentados os aspectos teóricos, exemplos práticos e conteúdos complementares que ampliam o conhecimento sobre as temáticas da unidade. Explorar os vídeos e infografias interativas. </div> <div> <p>Atividade</p> <p>Atividade não pontuada disponível na seção “Pratique e Compartilhe”.</p> <ul style="list-style-type: none"> ✓ Estudos de caso, resoluções, proposta de pesquisa ou produção criativa que integram atividades práticas aos conceitos teóricos básicos da unidade. ✓ As respostas e resultados da atividade proposta devem ser postados no fórum disponível na sessão “Compartilhe”. ✓ Após a postagem será disponibilizado feedback com modelo de resposta. </div> <div> <p>Avaliação Formativa</p> <p>Realizar a “Atividade Avaliativa” que constitui o recurso de avaliação pontuada da unidade. A pontuação desta atividade fará parte da nota final na N1 (ver item “Avaliação” deste plano).</p> </div>
<p>Unidade 3</p> <ul style="list-style-type: none"> Visão Geral Sobre Aprendizagem Supervisionada Estudo de Caso: Volume de Vendas de um Produto de Varejo Análise Descritiva dos Dados Predições com Árvore de Decisão para Classificação 	<div> <p>Objetivos de Aprendizagem</p> <ul style="list-style-type: none"> Adquirir uma visão geral sobre os principais métodos de aprendizagem supervisionada. Ter uma noção inicial sobre o princípio de funcionamento de cada dos principais métodos de aprendizagem supervisionada. Saber reconhecer situações onde aplicar tais métodos, na ciência ou nos negócios. Classificação com árvores de decisão – Estudo de caso: Árvores de decisão para classificação; Como árvores de decisão para classificação podem ajudar a otimizar as vendas de um produto de varejo. Conhecer sobre modelos preditivos com base em árvores de decisão para classificação. Conhecer sobre o potencial de aplicação dessa técnica tomando um estudo de caso como exemplo. Saber reconhecer situações similares onde aplicar árvores de decisão para classificação. </div>

	<ul style="list-style-type: none"> ▪ Conhecer sobre modelos de árvore de decisão para classificação. ▪ Entender o princípio de funcionamento da técnica de árvore de decisão para classificação. ▪ Saber reconhecer situações similares onde aplicar árvores de decisão para classificação.
	Estratégias de Ensino
	<p>Utilização de material referencial em diferentes formatos: vídeos, textos de referência conceitual, atividades de pesquisa, estudos de caso, infografias interativas, entre outros.</p> <p>Sequência sugerida:</p> <ul style="list-style-type: none"> ✓ Explorar a seção “Inspire-se” que contextualiza o tema da unidade e traz informações de tendências e inovações na respectiva área de conhecimento, aplicação prática ou estudos de caso, depoimentos ou entrevistas com profissionais qualificados do mercado de trabalho. ✓ Conhecer e entender os conceitos básicos da unidade apresentados na seção “Explore”. Neste material são apresentados os aspectos teóricos, exemplos práticos e conteúdos complementares que ampliam o conhecimento sobre as temáticas da unidade. Explorar os vídeos e infografias interativas.
	Atividade
	<p>Atividade não pontuada disponível na seção “Pratique e Compartilhe”.</p> <ul style="list-style-type: none"> ✓ Estudos de caso, resoluções, proposta de pesquisa ou produção criativa que integram atividades práticas aos conceitos teóricos básicos da unidade. ✓ As respostas e resultados da atividade proposta devem ser postados no fórum disponível na sessão “Compartilhe”. ✓ Após a postagem será disponibilizado feedback com modelo de resposta.
<p>Unidade 4</p> <ul style="list-style-type: none"> • Análise Exploratória • Estudo de Caso - Violência Urbana • Análise Descritiva dos Dados • Agrupamento ou Análise de Cluster 	Avaliação Formativa
	Realizar a “Atividade Avaliativa” que constitui o recurso de avaliação pontuada da unidade. A pontuação desta atividade fará parte da nota final na N1 (ver item “Avaliação” deste plano)
	Objetivos de Aprendizagem
	<ul style="list-style-type: none"> ▪ Adquirir uma visão geral sobre métodos de aprendizagem não supervisionada. ▪ Entender o princípio de funcionamento das técnicas de aprendizagem não supervisionada. ▪ Saber reconhecer situações similares onde aplicar os métodos de aprendizagem não supervisionada. ▪ Conhecer sobre os principais algoritmos de agrupamento (clusterização). ▪ Entender o princípio de funcionamento dos algoritmos de k-médias e de clusterização hierárquica. ▪ Saber reconhecer situações similares onde aplicar algoritmos de agrupamento. ▪ Conhecer sobre agrupamento hierárquico e suas aplicações. ▪ Entender o princípio de funcionamento da técnica de agrupamento hierárquico. ▪ Saber reconhecer situações similares onde aplicar agrupamento hierárquico. ▪ Revisão geral do aprendizado sobre estatística aplicada à ciência dos dados. ▪ Consolidar uma visão geral sobre a estatística aplicada à ciência dos dados. ▪ Saber reconhecer situações da vida onde aplicar à estatística e a ciência dos dados.

	Estratégias de Ensino
	Utilização de material referencial em diferentes formatos: vídeos, textos de referência conceitual, atividades de pesquisa, estudos de caso, infografias interativas, entre outros. Sequência sugerida: <ul style="list-style-type: none"> ✓ Explorar a seção “Inspire-se” que contextualiza o tema da unidade e traz informações de tendências e inovações na respectiva área de conhecimento, aplicação prática ou estudos de caso, depoimentos ou entrevistas com profissionais qualificados do mercado de trabalho. ✓ Conhecer e entender os conceitos básicos da unidade apresentados na seção “Explore”. Neste material são apresentados os aspectos teóricos, exemplos práticos e conteúdos complementares que ampliam o conhecimento sobre as temáticas da unidade. Explorar os vídeos e infografias interativas.
	Atividade
	Atividade não pontuada disponível na seção “Pratique e Compartilhe” . <ul style="list-style-type: none"> ✓ Estudos de caso, resoluções, proposta de pesquisa ou produção criativa que integram atividades práticas aos conceitos teóricos básicos da unidade. ✓ As respostas e resultados da atividade proposta devem ser postados no fórum disponível na sessão “Compartilhe”. ✓ Após a postagem será disponibilizado feedback com modelo de resposta.
	Avaliação Formativa
	Realizar a “Atividade Avaliativa” que constitui o recurso de avaliação pontuada da unidade. A pontuação desta atividade fará parte da nota final na N1 (ver item “Avaliação” deste plano”).
Prova Presencial	Avaliação em formato de prova presencial constituída de atividades múltipla escolha contemplando as quatro unidades da disciplina (ver item “Avaliação” deste plano”).

AVALIAÇÃO

A Nota Final (NF) da disciplina considera os seguintes elementos e valores:

NOTA N1				NOTA N2
UNIDADE 1	UNIDADE 2	UNIDADE 3	UNIDADE 4	PROVA PRESENCIAL A5
Atividade Avaliativa A1	Atividade Avaliativa A2	Atividade Avaliativa A3	Atividade Avaliativa A4	Contendo Questões Objetivas e/ou Dissertativas, individual.
Avaliação Individual com nota de 0 a 10	Avaliação Individual com nota de 0 a 10	Avaliação Individual com nota de 0 a 10	Avaliação Individual com nota de 0 a 10	

Média Final (MF) é calculada com a seguinte média ponderada das duas notas, N1 e N2 e pesos, respectivamente, de 40% e 60%, resultante da seguinte equação:

$$MF = (N1*0,4) + (N2*0,6)$$

Para aprovação, a Nota Final da disciplina deverá ser igual ou superior a 6,0 (seis), além da necessária frequência mínima de 75%, que corresponde a realização de, no mínimo, três das quatro Atividades Avaliativas da N1

O estudante que não atingir a média final 6,0 (seis), poderá realizar uma Prova Substitutiva (A6), cuja nota substituirá a nota da N2 (A5) obtida, caso seja maior.

BIBLIOGRAFIA BÁSICA

Gupta, C. Bhisham Estatística e probabilidade com aplicações para engenheiros e cientistas / Bhisham C. Gupta, Irwin Guttman ; tradução Ana Maria Lima de Farias, Vera Regina Lima de Farias e Flores. - 1. ed. - Rio de Janeiro : LTC, 2017. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521632931/cfi/6/10!/4/14/8/2@0:38.1>

Estatística Básica: a arte de trabalhar com dados / João Ismael D. Pinheiro ... [et al.]. – 2. ed. – Rio de Janeiro: Elsevier, 2015. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595156449/cfi/6/6!/4/22@0:73.8>

SILVA, Leandro Augusto da; PERES, Sarajane Marques; BOSCARIOLI, Clodis. Introdução à mineração de dados: com aplicações em R. Rio de Janeiro: GEN LTC, 2016. Livro digital. (1 recurso online). (SBC (Sociedade Brasileira de Computação). ISBN 9788595155473.

BIBLIOGRAFIA COMPLEMENTAR

CASTRO, Leandro Nunes de; FERRARI, Daniel Gomes. Introdução à mineração de dados: conceitos básicos, algoritmos e aplicações. São Paulo: Saraiva, 2016. Livro digital. (1 recurso online). ISBN 978-85-472-0100-5. Disponível em: <https://integrada.minhabiblioteca.com.br/books/978-85-472-0100-5>. Acesso em: 9 abr. 2020.

FÁVERO, Luiz Paulo; BELFIORE, Patrícia. Manual de análise de dados: estatística e modelagem multivariada com Excel, SPSS e Stata. Rio de Janeiro: GEN LTC, 2017. Livro digital. (1 recurso online). ISBN 9788595155602.

LOESCH, Cláudio. Métodos estatísticos multivariados. São Paulo: Saraiva, 2012. Livro digital. (1 recurso online). ISBN 9788502146105. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788502146105>. Acesso em: 12 jun. 2020.

VIEIRA, Sonia. Estatística básica. 2. ed. rev. ampl. São Paulo: Cengage Learning, 2018. Livro digital. (1 recurso online). ISBN 9788522128082. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788522128082>. Acesso em: 12 jun. 2020.

ANDERSON, David R. et al. Estatística aplicada a administração e economia. 4. ed. São Paulo: Cengage Learning, 2019. Livro digital. (1 recurso online). ISBN 9788522128006. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788522128006>. Acesso em: 15 abr. 2020.