

	INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ Campus Canindé - Graduação Tecnológica em Análise e Desenvolvimento de Sistemas	Nota:
	Disciplina: Inteligência Artificial	
	Professor: Raphael Torres Santos Carvalho Data: 25/04/2022	Etapa: 1

1. INTRODUÇÃO

Inteligência Computacional Aplicada é uma área de conhecimento da computação que vem ganhando bastante destaque nos últimos anos em virtude de ser composto de diversas ferramentas que permitem resolver problemas que não são possíveis de serem solucionados pela programação tradicional.

Regressão é a ferramenta de inteligência computacional aplicada que estuda as relações entre variáveis numéricas. Tem como principal objetivo verificar como certas variáveis de interesse influenciam uma variável resposta Y e criar um modelo matemático capaz de prever valores de Y com base em novos valores de variáveis preditoras X. É usado para diversas análises preditivas, como previsão de demandas, preços etc.

Com esta avaliação da N1, os alunos serão capazes de desenvolver scripts na linguagem R para resolução de problemas utilizando algoritmos de regressão linear e não linear.

2. REQUISITOS E CRITÉRIOS DE AVALIAÇÃO

Para cada questão apresentada na próxima seção, o aluno deverá elaborar um script (código) na linguagem R e realizar os procedimentos solicitados.

Ao final, o aluno deverá preparar um relatório contendo o código elaborado na linguagem R, os gráficos solicitados e responder às perguntas solicitadas em cada questão.

Cada questão será avaliada de forma **percentual, atribuindo-se os percentuais de 0%, 25%, 50%, 75% e 100% sobre a pontuação indicada para a questão/item.**

O relatório e todos os códigos scripts em R elaborados deverão ser enviados pela atividade postada na Sala de Aula do Google da disciplina dentro dos prazos previstos neste documento. **Caso seja observada a cópia – no todo ou em parte do teor do trabalho, seja de outros alunos ou da Internet, a(s) equipe(s) envolvida(s) terá(ão) sua nota anulada.**

3. ENTREGA DO TRABALHO

O trabalho é **INDIVIDUAL** e cada discente deverá preparar os artefatos solicitados em cada questão e enviá-los, devidamente compactados no formato ZIP ou RAR, na atividade “1ª Avaliação da N1” postada na sala de aula do Google da disciplina. Qualquer falha na compactação dos arquivos, bem como atraso nas entregas fora é de responsabilidade do aluno.

A data limite para entrega do trabalho **será dia 04/05/2022**. Não será aceito a entrega de nenhum arquivo ou do trabalho após o dia 04/05/2022, resultando em **nota zero** no item ou em toda a avaliação.

4. QUESTÕES

QUESTÃO 1 (3,0 pontos): Utilizando Regressão Linear Simples e o dataset denominado “aerogerador.data”, implemente um script em R para treinar o modelo que se ajusta melhor aos dados. Para isso, o script deverá dividir os dados em dois conjuntos (treinamento com 80% e teste com 20%) selecionados aleatoriamente. Em seguida, faça o que se pede:

- Plote o gráfico de dispersão dos dados de treinamento e o gráfico da reta obtida pelo modelo de regressão linear em cores diferentes.
- Indique a equação da reta obtida com os dados de treinamento;
- Qual o valor do R-quadrado ajustado para o modelo linear obtido com os dados de treinamento? Qual o significado desse valor?
- Teste o modelo obtido e plote o gráfico de dispersão dos dados de teste sobre o gráfico da reta obtida pelo modelo de regressão linear em cores diferentes.
- Com os dados de testes, compare os valores previstos pelo modelo com os valores reais dos dados de teste. O que é possível inferir? Esta reta é adequada para representar os dados do aerogerador?

QUESTÃO 2 (3,0 pontos): Utilizando Regressão Não-Linear Polinomial e o dataset denominado “aerogerador.data”, implemente um script em R para treinar o modelo que se ajusta melhor aos dados. Para isso, o script deverá dividir os dados em dois conjuntos (treinamento com 80% e teste com 20%) selecionados aleatoriamente. Em seguida, faça o que se pede:

- Plote o gráfico de dispersão dos dados de treinamento e o gráfico da reta obtida pelo modelo de regressão linear em cores diferentes.
- Implemente um modelo de regressão não-linear polinomial de 2º grau e indique a equação da obtida com os dados de treinamento pelo método da regressão polinomial;
- Qual o valor do R-quadrado ajustado para o modelo polinomial obtido com os dados de treinamento? Qual o significado desse valor?
- Plote o gráfico de dispersão dos dados de teste sobre o gráfico do modelo polinomial obtido pelo modelo de regressão polinomial em cores diferentes.
- Comparando os resultados obtidos com os dados de teste com os dados reais, o que se consegue inferir? A curva obtida é adequada para representar os dados do aerogerador?

QUESTÃO 3 (2,0 pontos): Repita o realizado na questão 2, mas altere o modelo para usar um modelo de Regressão Não-Linear Polinomial de 3º e compare o resultado obtido com o modelo de 2º grau. Plote os gráficos e faça a análise dos dados obtidos, conforme executado na questão 2.

QUESTÃO 4 (2,0 pontos): Utilizando Regressão Linear Múltipla e o dataset denominado “mtcars” disponível na library “datasets”, implemente um script em R para treinar o modelo que se ajusta melhor aos dados. Este conjunto de dados “mtcars” apresenta dados do consumo de combustível (mpg) de 32 modelos de automóveis em função de 10 características. Utilizando as características de potência do motor (hp) e de peso do veículo (wt), encontre a equação que estima o consumo de combustível (mpg). Em seguida, faça o que se pede:

- Plote o gráfico de dispersão dos dados dos 32 veículos em função da potência do motor (hp), do peso do veículo (wt) e do consumo de combustível (mpg).
- Indique a equação da superfície obtida com o modelo de regressão linear múltipla. c) Plote

o gráfico de dispersão dos dados sobre o gráfico da superfície obtida pelo modelo de regressão linear múltipla em cores diferentes.

d) Sobre o resultado obtido, o que se consegue entender sobre a relação entre potência do motor (hp) e de peso do veículo (wt) com o consumo de combustível?