

Evandro Cleto

Cientista de Dados

Informações pessoais

E-mail
evandrocleto74@gmail.com.br

Telefone
+5511970690730

LinkedIn
<https://www.linkedin.com/in/evandro-cleto-51b49340/>

Endereço
Rua professor Joaquim de Oliveira, 122 -
Centro - São Roque/SP

GitHub
EvandroCleto

Website
<https://evandrocleto-portfolio-homeapp-6568et.streamlitapp.com/Achievements>

Softwares

SQL	★★★★★
Linguagem R	★★★★★
Machine Learning	★★★★★
Pacote Office	★★★★★
Python	★★★★★
Apache Spark	★★★★★
Azure Machine Learning	★★★★★
Estatística	★★★★★
Git/GitHub	★★★★★
Kanban	★★★★★
Linux	★★★★★
Power BI	★★★★★
Apache Hadoop	★★★★★
Deep Learning	★★★★★
PNL	★★★★★
Scrum	★★★★★

Idiomas

Português	★★★★★
Inglês	★★★★★
Francês	★★★★★

Experiência Profissional

2023-04
- 2023-05

Cientista de Dados

Projeto Acadêmico, São Roque/SP

Criação de projeto de Machine Learning para segurança do trabalho "Prevenção a Eficiência de Extintores de Incêndio" e com o objetivo da Classificação preditiva atingir 85% de acurácia.

O projeto foi realizado de forma independente por 4 semanas para carregar, analisar, limpar, pré-processar, realizar as análises estatísticas e criar e avaliar os modelos de Machine Learning, utilizando linguagem R e seus pacotes readxl(para importar a fonte de dados em Excel), car(para geração de qqPlots), e1071(para analisar Assimetria e Curtose), caTools(para divisão em dados em treino e teste), pROC(para visualizar, suavizar e comparar as características de curvas ROC) e multiROC(para calcular métricas de Specificity, Sensitivity e AUC)

Os algoritmos de Machine Learning utilizados neste projeto para criação dos modelos de Regressão foram: Regressão Logística(Benchmark), KNN(2 versões), Naive Bayes(2 versões) e SVM(2 versões)

Na avaliação dos modelos verificou-se que o modelo KNN V1 teve a maior acurácia(0.9312) e o maior AUC(0.9311), sendo escolhido para deploy pois obteve melhor desempenho em 2 das 3 métricas analisadas.

O Dataset utilizado no projeto foi obtido em <https://www.muratkoklu.com/datasets/vtdhnd07.php>

2023-03
- 2023-04

Cientista de Dados

Projeto Acadêmico, São Roque/SP

Realizado desenvolvimento do projeto de previsão de Consumo de Energia de Carros Elétricos, utilizando diversos modelos de veículos e com o objetivo de atingir uma precisão de 80% pelo modelo de Regressão.

Este trabalho foi realizado de forma independente por 3 semanas para carregar, analisar, limpar, pré-processar, realizar a análise estatística e criar e avaliar os modelos de Machine Learning.

Este projeto foi desenvolvido em linguagem R, utilizando as bibliotecas como readxl para carregar dados, dplyr para manipulação de data frames, ggplot2 para plot de gráficos e caret para criação dos modelos.

Na modelagem preditiva de Regressão, foram utilizados os algoritmos Regressão Linear(Benchmark), Random Forest e SVM(2 versões). O Random Forest também foi utilizado para Feature Selection.

Foram construídas 4 versões de modelos preditivos usando diferentes abordagens. Na avaliação, o modelo que apresentou melhor performance foi o modelo de Benchmark(usando Regressão Linear) com Rsquared = 0.7945 e Residual standard error = 0.1173 que praticamente atingiu métrica definida no início do projeto(80%).

O Dataset utilizado no projeto foi obtido em <https://data.mendeley.com/datasets/tb9yryptdn/2>

2018-03
- Atualmente

Consultor e Desenvolvedor ERP Totvs Protheus

Porto Seguro, São Paulo/SP(Home Office)

Consultoria, desenvolvimento, implantação e sustentação nos módulos Faturamento, Compras, Financeiro, Estoque, Fiscal, Gestão de Contrato e Gestão de Serviços no software ERP Totvs Protheus, bem como em integrações REST e gateway de pagamentos.

Formação acadêmica

2023-04
- 2023-04

Certificação DSCP , Data Scientist Certified Professional

Data Science Academy, Remoto

2021-03
- 2022-02

Formação, Cientista de Dados

Data Science Academy, Remoto

1995-01
- 2000-12

Bacharel, Administração de Empresas

Faculdades Metropolitanas Unidas, São Paulo/SP



Big Data Analytics com R e Microsoft Azure Machine Learning

Big Data Real-Time Analytics com Python e Spark

Engenharia de Dados com Hadoop e Spark

Machine Learning

Business Analytics

Visualização de dados e Design de Dashboards

Análise Estatística e Modelagem Preditiva de Séries Temporais

SQL Para Data Science

Microsoft Power BI Para Data Science

Introdução ao Sistema Operacional Linux