Strauss Cunha Carvalho

Contatos:

e-mail: engsts@gmail.com

Linkedin: https://www.linkedin.com/in/strauss-cunha-carvalho-1a601a15/

Lattes: lattes.cnpg.br/4915404310911390

Formação: ano conclusão

1 - Instituto Tecnológico de Aeronáutica - ITA Mestrado

2014

Mestrado em Engenharia Eletrônica e Computação

Dissertação: MIDC - Um Método de Inferência Difusa para Classificação em Banco de Dados

Orientador: Dr. Adilson Marques da Cunha

2 - Universidade Federal do Rio Grande do Sul Pós-Graduação Lato Sensu 2023

Especialização em Ciência de Dados

3 - Universidade Federal Fluminense - UFF Pós-Graduação Lato Sensu 2008

Especialização em Criptografia e Segurança em Redes

4 - Centro Universitário de Barra Mansa - UBM Graduação 2005

Graduação em Engenharia da Computação

5 - Colégio Batista Americano - CIBA Técnico 2000

Técnico de Informática

6 - Escola Técnica Pandiá Calógeras – ETPC Curso aprendizagem industrial 1996

Eletromecânico

Experiência Profissional:

ano atuação

1 - Analista de Sistemas Serpro - Serviço Federal de Processamento de Dados 2006 - atualmente Atua como Desenvolvedor de Softwares utilizados por órgãos do governo federal. E, também, atuou como Administrador Regional de Redes da sétima região fiscal da RFB.

Papéis: Engenheiro de Dados, Cientista de Dados, Desenvolvedor Python, Desenvolvedor Java Spring, Database Administrator - Dba e Líder de Projetos.

- 1) Engenheiro de Dados Projetos: Data Lake da RFB e Data Lake corporativo.
- 2) Desenvolvedor Python/Natural/Adabas Projeto: Cadastro Nacional de Pessoas Jurídicas Cnpj;
- 3) Dba Projeto: Sistema Integrado de Inf. Econômico Fiscais Sief (módulo Cnpj);
- 4) Desenvolvedor Python Projeto: Sistema Interativo de Navegação no Linux Sinal;
- 5) Desenvolvedor Java Projeto: Piloto Cnpj; e
- 6) Administrador de Redes.

Tecnologias: Git, Python, PySpark, Java Spring Boot, Flask, Jupyter, Cloudera Hadoop, Kafka, Kudu, HBase, Grafana, Prometheus, Oracle e Natural/Adabas.

2 - Cientista de Dados 2rpnet

2017 - 2020

Atuou como Cientista de Dados em projetos aplicados ao setor varejista e bancário.

Papéis: Data Scientist e Desenvolvedor Python.

- 1) Data Scientist Projetos: a) elasticidade de preços de produtos de varejo; b) otimização de retorno financeiro na abertura e novas lojas físicas; e c) classificação de fraude em cartão de crédito;
- 2) Desenvolvedor Python Projetos: *Pipeline* de modelos de machine learning e descoberta de *Features Engineering*.

Tecnologias: Apache Zeppelin, Pyspark, Python, Apache Hadoop e Oracle database.

3 - Machine Learning Researcher FCMF - Fundação Cassimiro de Montenegro Filho 2015 - 2017

Regime: bolsista em curso de Doutorado

Atuou como Pesquisador Colaborador participando da pesquisa, desenvolvimento e aplicação, em ambiente de Produção, de dois Projetos de Machine Learning ligados à temática da detecção de fraudes em cartão de crédito e à concessão de crédito.

Tecnologias: Python (scikit-learn, statsmodels, pandas, numpy e seaborn), Jupyter, Microsoft SQL Server e Oracle database.

4 - Analista de Sistemas Instituto Biochimico

2005 - 2006

Regime: celetista 40h semanais

Desenvolvedor, no cargo de Analista de Sistemas Jr, lotado em projeto de desenvolvimento de um Enterprise Resource Planning - ERP interno.

Tecnologias: Microsoft .net e Oracle database.9i

5 - Programador SoftHouse

2001 - 2004

Desenvolvedor Visual Basic, C++ e SQL Server.

Desenvolvimento, implantação e sustentação de sistemas de informação comercial.

Idiomas:

1- Inglês

YES - Curso de Inglês (concluídos o níveis básico, intermediário e avançado)

1997 - 2000

Leitura: Boa, Escuta: Mediana, Escrita: Boa, Fala: Mediana

Publicação Científica:

1) Autor:

1.1 - A Classifier Evaluation for Payments Default Predictions in a Brazilian Retail Companydoi≥ https://doi.org/10.1007/978-3-319-77028-4 96 2018

This article presents an investigation about the performance of classification algorithms used for predicting payments' default. Classifiers used for modelling the data set include: Logistic Regression; Naive-Bayes; Decision Trees; Support Vector Machine; k-Nearest Neighbors; Random Forests; and Artificial Neural Networks. These classifiers were applied to both balanced and original data using the Weka data mining tool. Results from experiments revealed that Logistics Regression and Naive Bayes classifiers had the best performance for the chosen data set.

Publication: Springer International Publishing - Advances in Intelligent Systems and Computing. Switzerland. 2018, v. , p. 741-743. ISBN: 978-3-319-77028-4

1.2 - Investigating Attribute Assessment for Credit Granting on a Brazilian Retail Enterprise 2018 doi> https://doi.org/10.1007/978-3-319-77028-4 41

In this article, we investigate which features are required to enhance a credit scoring model for a Brazilian retail enterprise. In order to find attributes that can improve the performance of classifier algorithms for credit granting, a national and an international survey were carried out. A logistic regression classifier was used and the main result has improved the performance of data mining classifiers. The main contribution of this article was the verification that additional financial and behavioral data increase defaulting prediction performance on credit granting.

Publication: Springer International Publishing - Advances in Intelligent Systems and Computing. Switzerland. 2018, v. , p. 301-309. ISBN: 978-3-319-77028-4

1.3 - Applying a Method for Credit Limit Reviews in a Brazilian Retail Enterprise

This article is devoted to ensuring the reliability of information systems. An approach based on risk assessment and neutralization is used to increase the reliability of information systems. This approach allows for early risk assessment of the software development process and determines the most effective mitigation strategies.

Publication: Springer International Publishing - Advances in Intelligent Systems and Computing. Switzerland. 2018, v. , p. 731-733. ISBN: 978-3-319-77028-4

1.4 - Generating Flow Control Graph from NATURAL

2017

doi>https://doi.org/10.1007/978-3-319-54978-1 52

The work aims to generate White-box Test Cases from a mainframe NATURAL code fragment, using the Control Flow Graph Technique. Basically, it enables a code fragment analysis, generating its control flow represented by a Graph perspective. As a consequence, it provides the automatic generation of White-box Test Cases. Furthermore, this work contributes to attain a lower degree of difficult in the execution of White-box Tests within a Mainframe environment. Also, it brings a significant contribution related to the execution time, inherent to the software testing and provides automatic generation of test cases for empowerment of team decision, regarding which test cases are more relevant to be performed.

Publication: Springer International Publishing - Advances in Intelligent Systems and Computing. 1ed 2018, v. , p. 395-402. ISBN: 978-3-319-54978-1

1.5 – A Comparative Research between SCRUM and RUP Using Real Time Embedded Software Development

2013

doi> https://ieeexplore.ieee.org/document/6614397

This paper describes an academic experience performed by graduate students at the Brazilian Aeronautics Institute of Technology during the 2nd semester of 2011. It happened during a system prototype construction for Real Time Embedded System and Quality, Reliability, and Safety courses. The main idea behind this research was to conduct a comparative investigation between applications of two Software Development Processes: the Scrum and the Rational Unified Process, from system requirements' identification up to a final product delivery.

(Este artigo recebeu o prêmio de melhor Poster – The Best Poster Award)

Publication: IEEE - 2013 10th International Conference on Information Technology: New Generations Las Vegas, NV, USA. ISBN: 978-0-7695-4967-5

1.6 - Syntatic Patterns Recognition Applied To the Apf Analyses

2010

doi> https://doi.org/10.1109/ITNG.2010.84

The evolution of technology has enabled research in several areas, one is the software manufacturing cells planning. There are several techniques for the study of design patterns. This line of research and development, small scale production, has been offered some computer systems dedicated to of design patterns analysis. The aim of this paper is to present the software implementation that uses Grammar of design patterns to validate string of characters in valid design patterns.

Publication: IEEE - 2010 10th International Conference on Information Technology: New Generations

Las Vegas, NV, USA ISBN: 978-1-4244-6271-1

1.7 - Qualidade, Confiabilidade e Segurança nas disciplinas do processo unificado

2008

Este artigo descreve uma experiência acadêmica realizada por alunos de graduação e pós-graduação do Instituto Tecnológico de Aeronáutica (ITA), durante o segundo semestre de 2007. Essa experiência ocorreu durante a elaboração de um Protótipo de Sistema para as matérias Sistemas Embarcados de Tempo Real e Qualidade, Confiabilidade e Segurança (Safety) de Software. em um Estudo de Caso denominado Veículo Aéreo Não Tripulado - Estação de Controle - Satélite Artificial para Monitoramento Ecológico (VANT-EC-SAME).

Publicação: Simpósio Brasileiro de Engenharia de Software (SBES), 2008. UNICAMP, Campinas - SP.

2) Coautor:

O artigo propõe um método automatizado para detectar trincas e estimar o seu percentual, por quilômetro, em vídeos-registros, com o auxílio de técnicas de processamento de imagem e de aprendizado de máquina. Foram combinadas duas abordagens, uma que visa descartar regiões cuja probabilidade de ter trincas é baixa, e outra que se certifica de que as demais regiões têm trincas ou não. O modelo proposto foi validado em um conjunto de 138.909 imagens, obtendo-se acurácia de 0,90, sensibilidade de 0,89 e precisão de 0,75.

Publicação: Congresso de Pesquisa e Ensino em Transporte - ANPET, 33 edição. Novembro de 2019.

ISBN: 978-85-87893-22-2 Balneário Camboriú - SC

2.2 - Um Modelo de Controle Formal para o gerenciamento de riscos de processo em fábricas de software.

2009

O artigo apresenta um protótipo de um metamodelo que contempla as Áreas de Processo de risco do modelo de qualidade CMMi-DEV e que utiliza o formalismo das Redes de Petri para gerenciar risco de processo, em uma linha de produção de *software*.

Publicação: IX Simpósio Brasileiro de Segurança da Informação e Sistemas Computacionais (SBSEG), 2009. UNICAMP, Campinas - SP.

Prêmios:

1 - MIDCS - Um Método de Inferência Difusa para classificação de Sonegadores Fiscais

2014

Prêmio Criatividade e Inovação da Receita Federal do Brasil

Terceiro colocado

ESAF - Escola Superior de Administração Fazendária - Ministério da Fazenda - Receita Federal do Brasil Brasília - DF

2 - A Comparative Research between SCRUM and RUP Using Real Time Embedded Software Development

2014

10th International Conference on Information Technology - New Generations (ITNG), 2013 The Best Poster Award Las Vegas, Nevada, USA

3 - Um Modelo de Controle Formal para o Gerenciamento de Riscos de Projetos de Software 2014

Conserpro Terceiro colocado Florianópolis - SC

Localização:

Residência: São José dos Campos - SP