

Universidade Presbiteriana Mackenzie

APLICAÇÃO DE REALIDADE AUMENTADA PARA LOCOMOÇÃO NO CAMPUS HIGIENÓPOLIS DA UNIVERSIDADE PRESBITERIANA MACKENZIE

Andrey Gustavo Fonseca Vieira Fava (IC) e Daniela Vieira Cunha (Orientadora) Apolo:PIVIC
Mackenzle

RESUMO

Cada vez mais tecnologias estão disponíveis a um baixo custo e consequentemente acessíveis a um maior número de usuários. Essa disponibilidade de dispositivos facilita o acesso às ferramentas tal como a realidade aumentada (RA). A RA exige uso de dispositivos móveis que possuem um mínimo de processamento e memória, bem como câmera acoplada. Neste trabalho, a proposta inicial foi o estudo da RA para desenvolvimento de uma aplicação com duas finalidades principais: (1) calcular a melhor rota entre o local do usuário (origem) e o local escolhido no campus (destino) e por meio de RA mostrará as direções até o local solicitado; (2) informar as atividades que acontecem dentro de cada prédio no campus assim que o aluno se posiciona na frente dele. Os objetivos propostos inicialmente não foram cumpridos com sucesso na sua totalidade devido a falta de recursos de hardware disponíveis durante a pandemia. Aplicações em RA requerem máquinas com configurações de hardware de muito processamento e memória e não foi possível dispender de tal equipamento. A proposta da aplicação foi reformulada e adaptada aos recursos disponibilizados pela universidade. Assim, o objetivo adaptado foi desenvolver um aplicativo que apresenta informações de atividades que acontecem dentro de um prédio específico do campus de acordo com a real posição do aluno, conseguida por geolocalização bem como apresentar uma rota entre dois pontos dentro do campus. Utilizou-se o Android Studio e uma API do Google Maps juntamente com as linguagens de programação Java e Kotlin.

Palavras-chave: aplicativo, geolocalização, Google Maps.

ABSTRACT

More and more technologies are available at a low cost and consequently accessible to a greater number of users. This availability of devices facilitates access to tools such as augmented reality (AR). AR requires the use of mobile devices that have a minimum of processing and memory, as well as a camera attached. In this work, the initial proposal was

the study of AR for the development of an application with two main purposes: (1) calculating the best route between the user's location (origin) and the chosen location on the campus (destination) and through AR show directions to the requested location; (2) inform the activities

XV Jomada de Iniciação Científica e IX Mostra de Iniciação Tecnológica - 2020

that take place inside each building on the campus as soon as the student positions himself in front of him. The objectives initially proposed were not successfully met in their entirety due to

the lack of hardware resources available during the pandemic. RA applications require machines with hardware configurations with a lot of processing and memory and it was not possible to have such equipment. The application proposal was reformulated and adapted to the resources provided by the university. Thus, the adapted objective was to develop an application that presents information about activities that take place within a specific building on the campus according to the student's real position, achieved by geolocation as well as presenting a route between two points within the campus. Android Studio and a Google Maps API, and the programming languages Java and Kotlin were used.

Keywords: application, geolocation, Google Maps.

Universidade Presbiteriana Mackenzie

ANÁLISE EXPLORATÓRIA E VISUALIZAÇÃO DE PROCESSOS DE CLUSTERIZAÇÃO DE PONTOS GEORREFERENCIADOS: PROPOSTA DE UMA FERRAMENTA GRÁFICA DE ANÁLISE E VISUALIZAÇÃO

Krizhan Wesley Santos da Silva e Arnaldo Rabello de Aguiar Vallim Filho

Apoio: PIBIC Mackenzie

RESUMO

O presente projeto de pesquisa foi desenvolvido com o intuito de desenvolver soluções e ferramentas gráficas envolvendo análises exploratórias e visualização de resultados de um processo de clusterização. O estudo foi desenvolvido para uma aplicação em operações logísticas em que os clusters representam agrupamentos de pontos geográficos que demandam por serviços logísticos. Foi desenvolvido e implementado em uma ferramenta computacional um processo em que os dados de entrada são pontos com sua informação de localização geografia e já clusterizados. Esses pontos passam por um tratamento de

transformação de coordenadas e os clusters são apresentados na ferramenta gráfica desenvolvida na linguagem R, com o uso da biblioteca R-Shiny., tendo como fundo um mapa da região de estudo. Além disso, a ferramenta permite ainda algumas análises exploratórias dos dados. O resultado do estudo se mostrou satisfatório, tendo-se chegado à implementação de uma ferramenta computacional de visualização que apresenta com clareza os resultados do processo de clusterização em um contexto de operação logística, que era o objetivo da pesquisa.

Palavras-chave: Visualização, clusterização, ferramenta gráfica

ABSTRACT

This project was developed aiming at developing solutions and graphic tools involving exploratory analysis and visualization of results from a clustering process. The study was developed for an application in logistics operations where clusters represent groupings of geographical points that demand logistics services. The process was developed and implemented in a computational tool in which the input data are points with their geographic location information and already clustered. These points undergo a coordinate transformation treatment and the clusters are presented in the graphic tool developed in R language, using the R-Shiny. library, having as background a map of the study region. In addition, the tool also allows some exploratory analysis of the data. The result of the study proved to be satisfactory, having reached the implementation of a computational visualization tool that clearly presents the results of the clustering process in a context of logistic operation, which was the research objective.

Keywords: Visualization, clustering, graphic tool

Universidade Presbiteriana Mackenzie

APLICAÇÃO DE REALIDADE AUMENTADA PARA LOCOMOÇÃO NO CAMPUS HIGIENÓPOLIS
DA UNIVERSIDADE PRESBITERIANA MACKENZIE

Andrey Gustavo Fonseca Vieira Fava (IC) e Daniela Vieira Cunha (Orientadora) Apolo:PIVIC
Mackenzle

RESUMO

Cada vez mais tecnologias estão disponíveis a um baixo custo e consequentemente

acessíveis a um maior número de usuários. Essa disponibilidade de dispositivos facilita o acesso às ferramentas tal como a realidade aumentada (RA). A RA exige uso de dispositivos móveis que possuem um mínimo de processamento e memória, bem como câmera acoplada. Neste trabalho, a proposta inicial foi o estudo da RA para desenvolvimento de uma aplicação com duas finalidades principais: (1) calcular a melhor rota entre o local do usuário (origem) e o local escolhido no campus (destino) e por meio de RA mostrará as direções até o local solicitado; (2) informar as atividades que acontecem dentro de cada prédio no campus assim que o aluno se posiciona na frente dele. Os objetivos propostos inicialmente não foram cumpridos com sucesso na sua totalidade devido a falta de recursos de hardware disponíveis durante a pandemia. Aplicações em RA requerem máquinas com configurações de hardware de muito processamento e memória e não foi possível dispender de tal equipamento. A proposta da aplicação foi reformulada e adaptada aos recursos disponibilizados pela universidade. Assim, o objetivo adaptado foi desenvolver um aplicativo que apresenta informações de atividades que acontecem dentro de um prédio específico do campus de acordo com a real posição do aluno, conseguida por geolocalização bem como apresentar uma rota entre dois pontos dentro do campus. Utilizou-se o Android Studio e uma API do Google Maps juntamente com as linguagens de programação Java e Kotlin.

Palavras-chave: aplicativo, geolocalização, Google Maps.

ABSTRACT

More and more technologies are available at a low cost and consequently accessible to a greater number of users. This availability of devices facilitates access to tools such as augmented reality (AR). AR requires the use of mobile devices that have a minimum of processing and memory, as well as a camera attached. In this work, the initial proposal was the study of AR for the development of an application with two main purposes: (1) calculating the best route between the user's location (origin) and the chosen location on the campus (destination) and through AR show directions to the requested location; (2) inform the activities

XV Jomada de Iniciação Científica e IX Mostra de Iniciação Tecnológica - 2020

that take place inside each building on the campus as soon as the student positions himself in front of him. The objectives initially proposed were not successfully met in their entirety due to

the lack of hardware resources available during the pandemic. RA applications require

machines with hardware configurations with a lot of processing and memory and it was not possible to have such equipment. The application proposal was reformulated and adapted to the resources provided by the university. Thus, the adapted objective was to develop an application that presents information about activities that take place within a specific building on the campus according to the student's real position, achieved by geolocation as well as presenting a route between two points within the campus. Android Studio and a Google Maps API, and the programming languages Java and Kotlin were used.

Keywords: application, geolocation, Google Maps.

Universidade Presbiteriana Mackenzie

MODELO DE PREDIÇÃO DE INTERRUPÇÕES NA OPERAÇÃO DE TURBINAS DE USINAS
HIDRELÉTRICAS BASEADO EM SEUS CICLOS DE CARGA: Uma Abordagem Por Redes Neurais

Daniel Farina Moraes (IC) e Prof. Arnaldo Rabello de Aguiar Vallim Filho (Orientador)
Apolo:PIBIC Mackenzle

RESUMO

Turbinas hidrelétricas estão propensas a sofrerem danos, causando falhas de equipamento e trazendo prejuízo para a empresa. E o modo em que a maioria das hidrelétricas lidam com esse problema é realizando manutenções periódicas. Desse modo há um desperdício de dinheiro por causa de manutenções desnecessárias, e ainda assim há paradas não programadas por causa de falhas nos equipamentos. O objetivo deste estudo é prever as falhas de equipamento para que o número de manutenções desnecessárias diminua e o número de manutenções necessárias aumente, de modo a reduzir os custos de manutenção, paradas operacionais e a quantidade de falhas. Neste projeto foi desenvolvido uma ferramenta de monitoramento de variáveis relevantes dos ciclos de carga (tempo de carga de jatos de água nas pás) dessas turbinas. Esses dados dos ciclos de carga foram utilizados junto com os dados de alarmes, que indicam quando ocorrem falhas de equipamento, para alimentar uma rede neural artificial capaz de prever em um período de 24 horas se ocorrerá uma falha de equipamento ou não. Os resultados dos experimentos desenvolvidos se mostraram promissores e com a rede neural artificial implantada na operação de uma usina hidrelétrica, trará uma redução de custos significativa.

Palavras-chave: Redes Neurais Artificiais. Manutenção Preditiva. Aprendizado de Máquina.

ABSTRACT

Hydroelectric turbines are always subject to damage, causing equipment failures and raising the company's maintenance costs. The way that most hydroelectric power plants deal with this problem is to carry out periodic maintenance. However, in this way, there is a waste of money due to unnecessary maintenance, and even though they perform extra maintenance, sometimes failures still happen. The purpose of this study is to predict equipment failures so that the number of unnecessary maintenances decreases, and the number of necessary maintenances increases, in order to reduce costs, operational downtime and the number of future failures. In this project, a monitoring dashboard is developed to monitor relevant data of

XVI Jomada de Iniciação Científica e X Mostra de iniciação Tecnológica - 2020

the water jets cycles (duration of the water jets hitting the impulse blades) of those turbines. This data from the water jets cycles was used together with alarm data, which indicates when an equipment failure happens, to feed an artificial neural network capable of predicting, within a 24-hour period, whether an equipment failure will occur or not. The experiment results were promising and with the artificial neural network installed in a hydroelectric power plant's system, it will bring a significative cost reduction.

Keywords: Artificial Neural Networks. Predictive Maintenance. Machine Learning.

Universidade Presbiteriana Mackenzie

DISTRIBUIÇÃO GEOGRÁFICA E SITUAÇÃO DE ATENDIMENTO DOS EQUIPAMENTOS PET-CT NO BRASIL, CONSIDERADOS OS FATORES DISTÂNCIA E POPULAÇÃO: ESTUDO E PROPOSTA A PARTIR DOS DADOS DO DATASUS.

Felipe Marliere Cabral Ciabattari e Mario Olimpio de Menezes (Orientador) Apolo: PIVIC Mackenzie

RESUMO

As distâncias aos equipamentos de diagnóstico PET-CT precisam ser reduzidas para que pacientes tenham um rápido diagnóstico de doenças, e como consequência, maior chance de sobrevivência através da detecção e início de tratamento forma precoce.

A utilização do PET-CT não se limita à oncologia, com a aplicação não oncológica mais predominante na neurologia (para diagnósticos relacionados a demência e epilepsia) e na cardiologia (para aferir a viabilidade miocárdica).

Logo, este estudo vai abordar a análise de como estão distribuídos os equipamentos PET-CT no território brasileiro e o que deve ser feito para melhorar o acesso da população a este.

Em paralelo à distribuição dos equipamentos, também é necessário considerar a instalação de pontos de fabricação de radiofármaco.

Dados como Códigos de Endereço Postais (CEPs) das atuais localidades onde existe o aparelho PET-CT de diagnóstico em conjunto com Cadastro Nacional de Estabelecimentos de Saúde da seção de cadastro de equipamentos do DATASUS foram

introduzidos em um algoritmo desenvolvido na linguagem Python, configurado em um ambiente que utiliza as bibliotecas Folium, Pycep e Pandas. Com plotagem de dados em mapas para visualização e análise de resultados.

O Brasil possui uma quantidade insuficiente de PET-CTs para atender de forma adequada às demandas da população em relação aos exames específicos de diagnóstico, proporcionados somente por este equipamento.

Palavras-chave: PET-CT, Algoritmo de localização e distribuição, DATASUS.

ABSTRACT

Distances to PET-CT diagnostic equipment need to be reduced so that patients can have a rapid diagnosis of diseases, and as a consequence, have a greater chance of survival through early detection and initiation of treatment.

The use of PET-CT is not limited to oncology, with non-oncological application more prevalent in neurology (for diagnoses related to dementia and epilepsy) and in cardiology (to measure myocardial viability).

XVII Jornada de Iniciação Científica - 2021

Therefore, this study will address the analysis of how PET-CT equipment is distributed in the Brazilian territory and what should be done to improve the population's access to it.

In parallel to the distribution of equipment, it is also necessary to consider the installation of radiopharmaceutical manufacturing points.

Data such as Postal Address Codes (zip codes) of the current locations where the PET-CT diagnostic device exists in conjunction with the National Register of Health Establishments from the equipment registration section of DATASUS were introduced in an algorithm developed in Python language, configured in an environment that uses the Folium,

Pycep and Pandas libraries. With data plotting on maps for visualization and analysis of results.

Brazil has an insufficient amount of PET-CTs to adequately meet the demands of the

population in relation to specific diagnostic tests, provided only by this equipment.

Keywords: PET-CT, Localization and distribution algorithm, DATASUS.

Universidade Presbiteriana Mackenzie

UM ESTUDO ANALÍTICO SOBRE DADOS DE UMA CENTRAL DE ATENDIMENTO

André Philipe Andriotti de Moraes e Leandro Nunes de Castro

Apoio: PIBIETI Mackenzie

RESUMO

Em sua grande parte, as centrais de atendimento pouco ou não exploram os dados gerados durante as conversas, apesar de serem uma excelente base para melhorar o desempenho da empresa. Em decorrência disso, faz-se necessária a criação de uma plataforma computacional inteligente que objetiva a análise do humor do cliente e a identificação do assunto do contato, por meio de recursos analíticos como mineração de textos e processamento de linguagem natural. Para alcançar tais objetivos, foram construídas diversas aplicações na plataforma analítica SOMMA, da empresa AXONDATA Tecnologia Analítica LTDA, permitindo a classificação de sentimento por meio de processamento de texto, remoção de stopwords, geração de tokens, estruturação do documento em um Bag of Words e treinamento manual de um classificador. Não somente isso, mas também foi implementada uma aplicação focada na extração de palavras-chave, utilizando um componente responsável pela contagem de palavras em uma base de dados, o que possibilitou a geração de nuvens de palavras compostas pelas palavras mais frequentes no documento, trazendo um modelo mais visual dos resultados da aplicação. Como resultado, foi obtido um conjunto de informações sobre o atendimento, solucionando os objetivos propostos. Dessa forma, gerou-se um melhor aproveitamento dos dados proporcionados pela central de atendimento e maior satisfação dos clientes e atendentes.

Palavras-chave: Análise de dados. Mineração de textos. Central de atendimento.

ABSTRACT

For the most part, call centers use little or no data generated during conversations, despite

being an excellent basis for improving company performance. As a result, it is necessary to create an intelligent computing platform that aims to analyze the customer's mood and identify

the subject of the contact, through analytical resources such as text mining and natural language processing. To achieve these goals, several applications were built on the SOMMA analytical platform, from the company AXONDATA Tecnologia Analítica LTDA, allowing the classification of sentiment through text processing, stopword removal, token generation, document structuring in a Bag of Words and training of a classifier. Not only that, but an application

focused on keyword extraction was also implemented, using a component responsible for counting words in a database, which enabled the generation of word clouds composed of the XVII Jomada de Iniciação Científica - 2021

most frequent words in the document, bringing a more visual model of the application's results.

As a result, a set of information about the service is obtained, solving the proposed objectives. In this way, it generates a better use of the data provided by the call center and greater customer and attendant satisfaction.

Keywords: Data analysis. Text mining. Call center.

Universidade Presbiteriana Mackenzie

USO DE METAHEURÍSTICAS E FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL EM PROBLEMAS DE ROTEIRIZAÇÃO DE VEÍCULOS

Jin Uema e Arnaldo Rabello de Aguiar Vallim Filho

Apoio: PIVIC Mackenzie

RESUMO

Neste projeto, foi realizado uma pesquisa sobre o uso de uma metaheurística no problema de roteirização de veículos - PRV, que é um problema clássico de otimização combinatória. Foi utilizada uma ferramenta computacional normalmente utilizada em inteligência artificial/aprendizado de máquina, e que se mostrou muito eficaz para resolver problemas combinatórios de otimização. Com base na literatura, foi visto que as metaheurísticas têm

uma boa performance ao se tratar de problemas de solução NP-difícil, onde se enquadram problemas de otimização combinatória e aplicamos as suas propriedades em um modelo para realizar otimizações em rotas veiculares urbanas, especialmente voltadas para as distribuições de carga em múltiplos pontos de entrega. O modelo foi ajustado para trabalhar com frotas homogêneas e heterogêneas de veículos, podendo assim otimizar as rotas com base em sua capacidade e distância. Existe também um cálculo de penalidade caso o veículo ultrapasse sua capacidade máxima ou fique abaixo de seu nível de carga mínimo. Após isto, realizamos o treinamento do modelo e mensuramos os seus ganhos durante as interações do algoritmo e gravamos sua distância global, as distâncias de cada rota e trechos, seus pesos e pontos em cada rota. Ao final de sua execução, tivemos uma porcentagem de redução que varia entre 74% e 94% em média na distância das rotas em comparação com sua distância inicial e com isso podemos ajudar aplicações existentes nas áreas de logística de múltiplas paradas.

Palavras-chave: Metaheurísticas, Inteligência Artificial, Roteirização de Veículos, Otimização,
ABSTRACT

In this project, a research was carried out on the use of a metaheuristic in the vehicle routing problem - VRP, which is a classic combinatorial optimization problem. A computational tool commonly used in artificial intelligence/machine learning was used, which proved to be very effective in solving combinatorial optimization problems. Based on the literature, it was seen that metaheuristics have a good performance when dealing with NP-hard solution problems, where combinatorial optimization problems are included, and we apply their properties in a model to perform optimizations in urban vehicular routes, especially aimed at for load distributions at multiple delivery points. The model was adjusted to work with homogeneous and heterogeneous vehicle fleets, thus being able to optimize routes based on their capacity and distance. The objective function to be optimized, considers also a penalty if the vehicle

XVII Jornada de Iniciação Científica - 2021

exceeds its maximum capacity or falls below its minimum load level. After that, we train the model and measure its gains during the algorithm's interactions and record its global distance, the distances of each route and segments, its weights and points in each route. At the end of its execution, we had a percentage reduction that varies between 74% and 94% on average in the distance of the routes compared to its initial distance and with this we can help existing

applications in the areas of multi-stop logistics.

Keywords: Metaheuristics, Artificial Intelligence, Vehicle Routing, Optimization

Universidade Presbiteriana Mackenzie

UM ESTUDO DE ANÁLISE E PREVISÃO DE SÉRIES TEMPORAIS APLICADO ÀS CURVAS DE INDICES DE ISOLAMENTOS DEVIDOS À PANDEMIA DE COVID-19

Diego Clementino Pompeo e Amaldo Rabello de Aguiar Vallim Filho (Orientador) Apolo: PIBIC Mackenzie

RESUMO

Nesta pesquisa, foi feita uma análise e previsão das séries temporais do índice de isolamento devido a pandemia da COVID-19, durante um período de um ano, fevereiro de 2020 a fevereiro de 2021, das principais cidades das 16 regiões administrativas do estado de São Paulo, divididas em 4 sub-regiões: a Capital, a Região Metropolitana de São Paulo (RMSP), o Litoral e o Interior. Para que fosse possível a análise e previsão do índice de isolamento de cada uma dessas regiões, foram utilizados os dados disponibilizados pelo governo de Estado de São Paulo, que após a limpeza, a transformação e a integração, foi possível o desenvolvimento de quatro tipos de análise: cálculo de média móvel, curva de tendência e curva de sazonalidade (anual e semanal), e ainda a previsão para períodos futuros. As previsões foram feitas para até duas semanas à frente e com uma margem de erro estipulada dentro no próprio algoritmo desenvolvido. Todo o estudo foi desenvolvido com o apoio da linguagem de programação R, e suas bibliotecas, tornando assim, possível a compreensão e análise do comportamento dos índices de isolamento de cada município analisado.

Palavras-chave: Séries Temporais, COVID-19, Ciência de Dados

ABSTRACT

In this research, an analysis and forecast of the time series of the isolation index due to the COVID-19 pandemic was carried out, over a period of one year, from February 2020 to February 2021, in the main cities of the 16 administrative regions of the state of São Paulo, divided into 4 sub-regions: the Capital, the Metropolitan Region of São Paulo (RMSP), the Coast and the Interior. In order to analyze and predict the isolation index of each of these

regions, data provided by the São Paulo State government were used, which after cleaning, transformation and integration, it was possible to develop four types of analysis: calculation of moving average, trend curve and seasonality curve (annual and weekly), as well as forecasting for future periods. The predictions were made for up to two weeks ahead and with a margin of error stipulated within the developed algorithm itself. The entire study was developed with the

support of the R programming language, and its libraries, thus making it possible to understand

and analyze the behavior of the isolation indices of each analyzed municipality.

Keywords: Time Series, COVID-19, Data Science

Universidade Presbiteriana Mackenzie

REQUISITOS DE ACESSIBILIDADE PARA EXPERIÊNCIAS IMERSIVAS PARA DEFICIENTES VISUAIS EM OBRAS DE ARTE BIDIMENSIONAIS

Claudia Fiorentino Andrade (IC) e Maria Amelia Eliseo (Orientador)

Apoio: PIVIC Mackenzie

RESUMO

A inserção de pessoas com deficiência na sociedade é ampliada gradualmente com o passar dos anos, sendo incentivada pela formulação de tecnologias assistivas. No que diz respeito à integração cultural, mais especificamente na percepção visual, ainda é fundamental investigar novos meios de expressar o conteúdo por alternativas não ópticas e efetivar sua instalação em museus e galerias de arte. A partir disso, esse artigo busca analisar requisitos na implementação de uma experiência multissensorial acessível e imersiva para o usuário com deficiência visual de forma que seja capaz de vivenciar os sentimentos e as emoções transmitidos por uma obra de arte. Para isso, foram estudados, a partir da literatura, experimentos que mostram estímulos multissensoriais para entender as necessidades do deficiente visual e alternativas para a vivência de experiências visuais. Assim, foi possível extrair quais são as necessidades do deficiente visual e as formas de exploração sensorial além da visão aplicada para substituir a carência existente. Para validar

os requisitos analisados, foi construída uma representação 3D da tela "A Cuca" da pintora Tarsila do Amaral, para aplicar os requisitos levantados. O modelo produzido confere características táteis que permitem a compreensão dos elementos presentes no quadro, como fauna e flora, e a diferenciação das cores presentes em sua composição.

Palavras-chave: Acessibilidade. Deficiência visual. Arte.

ABSTRACT

The inclusion of people with disabilities in society is gradually expanded over the years, being encouraged by the formulation of assistive technologies. With regard to cultural integration, more specifically in visual perception, it is still essential to investigate new ways of expressing content through non-optical alternatives and effecting its installation in museums and art galleries. From this, this article seeks to analyze requirements in the implementation of an accessible and immersive multisensory experience for the visually impaired user so that he is able to experience the feelings and emotions transmitted by a work of art. For this, experiments were studied from the literature that show multisensory

Universidade Presbiteriana Mackenzie

INVESTIGANDO UMA SOLUÇÃO PARA PREVISÃO DE FALHAS EM TURBINAS HIDRÁULICAS POR MEIO DA REDE DEEP LEARNING TRANSFORMER

Bruno Nardelli Santiago (IC) e Orlando Bisacchi Coelho (Orientador)

Apolo: PIVIC Mackenzle

RESUMO

O mercado da energia elétrica é muito importante na economia mundial porque reflete em diversos aspectos da sociedade. Além disso, a manutenção preditiva tem se tornado cada vez mais relevante nos processos de produção. Ela tem a peculiaridade de acompanhar periodicamente os dados coletados de equipamentos por meio de monitoração ou inspeções, assim, podendo antecipar-se de eventos anômalos, como falhas. Este estudo teve como objetivo investigar experimentalmente se a arquitetura de rede Deep Learning LSTM (Long Short Term Memory) apresenta um desempenho sensivelmente inferior à alcançada pela solução desenvolvida com base na arquitetura Transformer. Assim, ao fazer

este estudo comparativo, entre as duas arquiteturas, esta pesquisa poderá auxiliar em uma futura organização em seu processo de tomada de decisão, sobre qual arquitetura usar em uma solução para previsão de falhas em turbinas hidráulicas.

Palavras-chave: Deep Learning. Transformer. Manutenção preditiva. LSTM.

ABSTRACT

The electric power market is very important in the world economy because it reflects on several aspects of society. In addition, predictive maintenance has become increasingly relevant in production processes. It has the peculiarity of periodically following up the data collected from equipment through monitoring or inspections, thus being able to anticipate anomalous events, such as failures. This study aimed to experimentally investigate whether the Deep Learning network architecture LSTM (Long Short Term Memory) presents a performance appreciably inferior to that achieved by the solution developed based on the Transformer architecture. Thus, by making this comparative study between the two architectures, this research can help a future organization in its decision making process, about which architecture to use in a solution to predict failures in hydraulic turbines.

Keywords: Deep Learning. Transformer. Predictive Maintenance. LSTM.

Universidade Presbiteriana Mackenzie

INVESTIGANDO UMA SOLUÇÃO PARA PREVISÃO DE FALHAS EM TURBINAS HIDRÁULICAS POR MEIO DE DETECÇÃO DE ANOMALIA, DEEP LEARNING

E A ARQUITETURA AUTOENCODER

Karen Assunção de Farias (IC) e Orlando Bisacchi Coelho (Orientador)

Apoio: PIBIC MackPesquisa

RESUMO

Esta pesquisa apresenta uma solução baseada em redes neurais da classe deep learning, utilizando a arquitetura Autoencoder, para solucionar o problema de manutenção preditiva de falhas em turbinas hidráulicas. Foi utilizada uma base de dados fornecida por uma operadora de energia do estado de São Paulo, contendo dados, coletados por meio de sensores, relativos ao funcionamento de um equipamento e os alarmes gerados pelo sistema supervisor a partir desses dados. Foram identificados quais desses eram mais influentes

para o contexto do problema. Em seguida, foi desenvolvido um modelo computacional capaz de prever falhas futuras dessa turbina, o que possibilita a realização de manutenções preditivas. Abordamos a questão sob o paradigma de Detecção de Anomalias. Nesse sentido, optamos por uma rede deep learning Autoencoder. Para implementá-la, usamos o TensorFlow 2, uma biblioteca deep learning que pode ser chamada a partir de scripts em Python.

Palavras-chave: Manutenção Preditiva, Detecção de Anomalia, Deep Learning, Autoencoder, Turbinas Hidráulicas.

ABSTRACT

This research presents a solution based on Deep Learning, using the Autoencoder architecture, to solve the problem of predictive maintenance for hydraulic turbines. A dataset

was provided by an energy operator in the state of São Paulo, containing data related to the operation of the equipment, collected through sensors, and the alarms generated from the data

by the plant's supervisory system. First it was identified which data were more influential in the

context of the problem. Then a computational model capable of predicting possible future failures of these turbines was developed, which makes it possible to carry out predictive maintenance. We approached the issue as an Anomaly Detection problem. In this sense, we chose a deep learning network of the Autoencoder class. To implement it, we used TensorFlow 2, a deep learning library that can be called from Python scripts.

Keywords: Predictive Maintenance, Anomaly Detection, Deep Learning, Autoencoder, Hydraulic Turbines.

UM ESTUDO SOBRE A INFLUÊNCIA DA SELEÇÃO DE ATRIBUTOS NO AGRUPAMENTO DE VAGAS EM APLICAÇÕES DE E-RECRUITMENT

Joel Joaquim de Souza Junior e Leandro Nunes de Castro

Apoio: PIBIC Mackenzie

RESUMO

Sistemas de recomendação, em geral, têm por finalidade recomendar itens ao usuário de forma eficiente e eficaz baseado em seu perfil. Um sistema de recrutamento online

(e-recruitment) tem por finalidade recomendar vagas para um candidato de acordo com o seu perfil podendo também agir na via inversa, recomendando candidatos mais qualificados para determinada vaga. Definir quais variáveis serão utilizadas impacta diretamente na qualidade da recomendação de modo que, ao utilizar as variáveis mais importantes, temos uma melhor assertividade no processo. Tendo isso em foco, este trabalho destina-se a realizar a seleção das variáveis mais importantes de uma base de dados de recrutamento online utilizando técnicas de seleção de atributos (características). Foram utilizados os algoritmos de Mitra, SUD e ACA para realizar a seleção dos atributos aplicados no agrupamento da base. As bases utilizadas foram derivadas da base original considerando três cenários distintos: a base contendo os atributos com características das vagas; a base contendo o bag of words do atributo de descrição das vagas; e a base resultante da união das duas anteriores. Os subconjuntos de variáveis selecionadas em cada um dos cenários acima tiveram seu desempenho avaliado na tarefa de agrupamento. Os resultados obtidos em cada cenário mostram um ganho de desempenho do agrupamento quando é feita a seleção de variáveis em comparação com as bases originais.

Palavras-chave: Sistema de Recomendação. Recrutamento Online. Seleção de Atributos.

ABSTRACT

Recommender systems aim to effectively recommend items to the user based on their profile. An online recruitment system recommends jobs for a candidate according to his profile and can also act in reverse, recommending more qualified candidates for a particular job. Defining which variables will be used impacts directly on the recommendation quality so that, when using the most important variables, we have a better assertiveness in the process. With this in mind, the goal of this work is to select the most important features of an online recruitment database using feature selection techniques. More specifically, we used the algorithms of Mitra, SUD and ACA to perform feature selection. The datasets used were derived from the original dataset assuming three distinct scenarios: the dataset containing the attributes related with the jobs characteristics; the dataset containing the bag of words of the description attribute

of the jobs; and the dataset resulting from the union of the two previous ones. The features subsets selected in each of the above scenarios had their performance evaluated in a clustering task. The results obtained in each scenario show a performance gain of the clustering process when feature selection is made over the original data.

Keywords: Recommender System. Online Recruitment. Feature Selection.

EXPERIMENTOS COM O PROBLEMA DE ROTEIRIZAÇÃO DE VEÍCULOS BASEADO EM CAIXEIRO VIAJANTE: UMA ABORDAGEM PELA METAHEURÍSTICA SIMULATED ANNEALING

Leonardo Nogueira Pizzi e Arnaldo Rabello de Aguiar Vallim Filho (Orientador)

Apoio: PIBIC Mackenzie

RESUMO

O objetivo desta pesquisa é trabalhar com o Problema do Caixeiro Viajante (PCV), voltado para roteirização de veículos, que é um conhecido problema logístico de otimização combinatória. O PCV é uma aplicação clássica de Pesquisa Operacional e de Inteligência Artificial por ser um problema combinatório de alta dificuldade (NP-Difícil) e também por apresentar um desafio técnico. Assim, é geralmente tratado por meio do uso de heurísticas e metaheurísticas. A pesquisa desta Iniciação Científica baseia-se especificamente na metaheurística Simulated Annealing, com o propósito de utilizar um tipo de objetivo de otimização para o PCV diferente do tradicional. O objetivo de otimização desta pesquisa foi o de minimização do custo logístico total da operação, incluindo os componentes de custo de transporte, o custo de estoques em trânsito e em armazém e o custo de vendas perdidas. Foi utilizada a metodologia da Agência Nacional de Transportes Terrestres, proposta devido à greve dos caminhoneiros que ocorreu no Brasil em Maio de 2018. Para isto foi feita a implementação de um algoritmo de otimização de rotas. Partindo dessa implementação, diversos experimentos foram desenvolvidos com a proposta de comparar os resultados aqui obtidos com os resultados do objetivo tradicional de minimização da distância total percorrida. Os resultados obtidos são apresentados e discutidos neste artigo.

Palavras-chave: Roteirização de Veículos. Simulated Annealing. Caixeiro Viajante.

ABSTRACT

The goal of this research is to work over the Traveling Salesman Problem (TSP), directed to vehicle routing. This well-known combinatorial optimization logistic problem is normally managed by means of heuristics and metaheuristics. The TSP is a classic application of Operational Research and Artificial Intelligence. It is a complex combinatorial problem (NP-hard), and presents a technical challenge. The research of this Scientific Initiation is based specifically on the Simulated Annealing metaheuristic. Its purpose is to use a type of improvement goal to the TSP that is different from the traditional one. The improvement goal used in this research was the cost minimization. It also includes the cost of the transportation, the cost of the stock in transit and the stock in the warehouse, and the cost of the lost sales. It was used the methodology of the National Terrestrial Transport Agency (ANTT), created due to the truck drivers' strike that occurred in Brazil in May 2018. This goal was achieved through the implementation in a route optimization algorithm. Many experiments were developed from this implementation in order to compare the results achieved in this research with the results of the traditional goal, which is minimizing the total distance traveled. The aim of this research is to acquire more knowledge about the behavior of routing operations, and also the behavior of the algorithms used to solve the presented problem.

Keywords: Vehicle Routing. Simulated Annealing. Traveling Salesman

UM ESTUDO SOBRE A PERCEPÇÃO DE SEGURANÇA QUE AS LEIS DIGITAIS
EXERCEM SOBRE USUÁRIOS DE SERVIÇOS E PRODUTOS DIGITAIS.

Nathan Novais Borges (IC); Fabio Silva Lopes (Orientador)

Apoio: PIBIC Mackenzie

RESUMO

Leis que abrangem o contexto digital em relação à proteção de Dados vêm se tornando cada

vez mais comuns, tais como a brasileira LGPD (Lei Geral de Proteção de Dados) e a GDPR (General Data Protection Regulation). A criação de regulamentações de manipulação e coleta de dados, que abrangem o contexto digital, são consequência da crescente expansão do mundo digital junto ao aumento exponencial na produção de dados. O principal objetivo de tais legislações é regulamentar o armazenamento, processamento e utilização de dados pessoais, com o intuito de proteger e garantir a segurança dos titulares desses dados. Porém muitos indivíduos possuem deficiências no que se refere ao conhecimento de Leis que abrangem o contexto digital e boas práticas sobre o comportamento e vivência dentro do mundo virtual, criando vulnerabilidades de segurança para tais pessoas. Entre os principais fatores que aceleraram a adoção de serviços computacionais por pessoas com menor familiaridade com a tecnologia, podemos citar a pandemia internacional de covid-19, que ao criar um cenário de mudança de comportamento da sociedade visando um maior distanciamento social, gerou tanto um aumento em interações através dos meios digitais quanto uma escalada no número de golpes digitais. Com base nesse cenário, esse estudo busca entender como os indivíduos, cada vez mais associados a um contexto digital, compreendem e entendem as Leis digitais, criadas para protegê-los.

Palavras-chave: Leis de Regulação de Dados, Segurança de Dados, Cidadão Digital

ABSTRACT

Laws covering the digital context in relation to Data protection have become increasingly common, such as the Brazilian LGPD (Lei Geral de Proteção de Dados) and GDPR (General Data Protection Regulation). The creation of data handling and collection regulations, which cover the digital context, are a consequence of the growing expansion of the digital world along with the exponential increase in data production. The main purpose of such legislation is to regulate the storage, processing and use of personal data in order to protect and ensure the safety of the holders of such data. However, many individuals have deficiencies regarding the knowledge of Laws that cover the digital context and good practices on behavior and experience within the virtual world, creating security vulnerabilities for such people. Among the main factors which have accelerated the adoption of computer services by people with less familiarity with technology, we may mention the international pandemic of covid-19, which, by

XVIII Jornada de Iniciação Científica - 2022

creating a scenario of change in society's behavior aiming at a greater social distance, has generated both an increase in interactions through digital means and an escalation in the number of digital scams. Based on this scenario, this study seeks to understand how individuals, increasingly associated to a digital context, understand and comprehend the Digital

Laws, created to protect them.

Keywords: Data Regulation Laws, Data Security, Digital Citizen

ANÁLISE DE SENSIBILIDADE DOS PARÂMETROS DO ALGORITMO K VIZINHOS INFORMATIVOS MAIS PRÓXIMOS PARA PROBLEMAS DE CLASSIFICAÇÃO DE DADOS

Vinícius Gomes Pajaro Grande e Leandro Augusto da Silva

Apoio: PIBIC Mackenzie

RESUMO

O problema da parametrização é muito comum nos algoritmos usados em Mineração de Dados, pois a melhor escolha para resolver um determinado problema não é robusta o suficiente para ser generalizada a outras classes de problemas. Em geral, o melhor parâmetro é definido em um processo empírico onde são realizados diversos testes utilizando métricas de avaliação afim de decidir qual é a melhor medida para resolver um problema específico. O LI-KNN (do inglês Locally Informative - K Nearest Neighbor) é um algoritmo interessante proposto para ser eficiente em situações em que a separação das classes não é bem definida. Porém os autores que propuseram o algoritmo definiram os melhores parâmetros sem uma apresentação da metodologia que pudesse ser reproduzível ou ao menos justificável. Neste artigo é utilizada uma metodologia de literatura proposta para investigar classes sobrepostas e a relação com a escolha de parametrização através de análises quantitativas e qualitativas e do uso de bases reais e artificiais afim de comparar e justificar melhor a escolha destes parâmetros e analisar a sua sensibilidade em situações diversas. Os resultados da melhor parametrização de LI-KNN são contrastados com algoritmos clássicos como Random Forest e SVM. Os resultados para conjuntos de dados sintéticos e conjuntos de dados reais mostram através de gráficos e tabelas a eficiência de LI-KNN em contraste com KNN, Random Forest

e SVM.

Palavras-chave: Classificação de dados. Vizinho mais próximo. Informatividade.

ABSTRACT

The parameterization problem is very common in the algorithms used in Data Mining, because the best choice for solving a particular problem is not robust enough to be generalized to other problem classes. In general, the best parameter is an empirical process which various tests are performed using evaluation metrics to decide what is the best measure to solve a specific problem. Locally Informative Nearest Neighbor or simply LI-KNN is an interesting algorithm proposed to be efficient in situations that the classes separations are not well defined. However, the authors that proposed the algorithm define the best parameters without

a presentation of a methodology that is possible to be reproduced or at least justifiable. In this paper is used a literature methodology proposed to investigate overlapped classes versus

XV Jornada de Iniciação Científica e IX Mostra de Iniciação Tecnológica - 2019

parametrization choice through quantitative and qualitative analyzes and the use of real and artificial databases in order to better compare and justify the choice of parameters and analyze

their sensitivity in various situations. The results of best parametrization of LI-KNN is contrasted with classical algorithms such as Random Forest and SVM. The results for synthetic datasets and real datasets show through graphs and tables the efficiency of LI-KNN in contrasting to KNN, Random Forest and SVM.

Keywords: Data classification. Nearest neighbor. Informativeness.