Instituto Federal do Rio Grande do Norte
Campus Natal - Central
Diretoria de Gestão e Tecnologia da Informação
Tecnologia em Análise e Desenvolvimento de Sistemas

Explorando Regiões de Denso Interesse e Outliers em um ambiente de dados espaço-temporal

Francisco Bento da Silva Júnior

Natal-RN
Dezembro, 2018

Francisco Bento da Silva Júnior

Explorando Regiões de Denso Interesse e Outliers em um ambiente de dados espaço-temporal

Trabalho de conclusão de curso de graduação do curso de Tecnologia e Análise em Desenvolvimento de Sistemas da Diretoria de Gestão e Tecnologia de Informação do Instituto Federal do Rio Grande do Norte como requisito parcial para a obtenção do grau de Tecnologo em Análise e Desenvolvimento de Sistemas.

Banco de Dados:
Análise de Dados

Orientador

Dr. Plácido Antonio de Souza Neto

TADS – Curso de Tecnologia em Análise e Desenvolvimento de Sistemas
DIATINF – Diretoria Acadêmica de gestão e Tecnologia da Informação
CNAT – Campus Natal - Central
IFRN – Instituto Federal do Rio Grande do Norte

Natal-RN

Dezembro, 2018

Trabalho de Conclusão de Curso de Graduação sob o título *Título* apresentada por Nome completo do autor e aceita pelo Diretoria de Gestão e Tecnologia da Informação do Instituto Federal do Rio Grande do Norte, sendo aprovada por todos os membros da banca examinadora abaixo especificada:

Nome completo do orientador e titulação
Presidente
DIATINF – Diretoria Acadêmica de Gestão e Tecnologia da Informação
IFRN – Instituto Federal do Rio Grande do Norte

Nome completo do examinador e titulação
Examinador
Diretoria/Departamento
Instituto

Nome completo do examinador e titulação
Examinador
Diretoria/Departamento
Universidade

Natal-RN, data da defesa (dia, mês e ano).

Homenagem que o autor presta a uma ou mais pessoas.

Agradecimentos

Agradecimentos dirigidos àqueles que contribuíram de maneira relevante à elaboração do trabalho, sejam eles pessoas ou mesmo organizações.

Citação

Autor

Explorando Regiões de Denso Interesse e Outliers em um ambiente de dados espaço-temporal

Autor: Francisco Bento da Silva Júnior
Orientador(a): Titulação e nome do(a) orientador(a)

Resumo

O resumo deve apresentar de forma concisa os pontos relevantes de um texto, fornecendo uma visão rápida e clara do conteúdo e das conclusões do trabalho. O texto, redigido na forma impessoal do verbo, é constituído de uma sequência de frases concisas e objetivas e não de uma simples enumeração de tópicos, não ultrapassando 500 palavras, seguido, logo abaixo, das palavras representativas do conteúdo do trabalho, isto é, palavras-chave e/ou descritores. Por fim, deve-se evitar, na redação do resumo, o uso de parágrafos (em geral resumos são escritos em parágrafo único), bem como de fórmulas, diagramas e símbolos, optando-se, quando necessário, pela transcrição na forma extensa, além de não incluir citações bibliográficas.

Palavras-chave: Palavra-chave 1, Palavra-chave 2, Palavra-chave 3.

Explorando Regiões de Denso Interesse e Outliers em um ambiente de dados espaço-temporal

Author: Francisco Bento da Silva Júnior
Supervisor: Titulação e nome do(a) orientador(a)

Abstract

O resumo em língua estrangeira (em inglês *Abstract*, em espanhol *Resumen*, em francês *Résumé*) é uma versão do resumo escrito na língua vernícula para idioma de divulgação internacional. Ele deve apresentar as mesmas características do anterior (incluindo as mesmas palavras, isto é, seu conteúdo não deve diferir do resumo anterior), bem como ser seguido das palavras representativas do conteúdo do trabalho, isto é, palavras-chave e/ou descritores, na língua estrangeira. Embora a especificação abaixo considere o inglês como língua estrangeira (o mais comum), não fica impedido a adoção de outras linguas (a exemplo de espanhol ou francês) para redação do resumo em língua estrangeira.

Keywords: Keyword 1, Keyword 2, Keyword 3.

Lista de figuras

Lista de tabelas

Lista de abreviaturas e siglas

Sumário

1 Introdução

Nos últimos dez anos, a busca por termos como *big data*, análise de dados e visualização de dados tem aumentado notoriamente. Existem muitas razões para esse fenomeno, um deles é que com o avanço do poder computacional nós agora lidamos com imensas quantidades de dados, que crescem diariamente, de diversas fontes, em vários formatos e num incrível curto espaço de tempo. Dessa forma, novos desafios vêm surgindo na área de análise de dados: Como processar essas imensas quantidades rápida e eficientemente? Como visualizar esse montante de dados? Como *limpar* o conjunto de dados sem perder pontos importantes?

No que se refere ao campo da análise de dados e o analista está lidando com grandes *datasets*, é muito comum encontrar vários pontos com atributos muito distantes do resto de seu conjunto. Isto acontece porque quanto maior é o dataset, mais facilmente se pode encontrar pontos atípicos que serão mais distantes da distribuição normal. Esse comportamente é importante ser estudado pelo analista para que seja descoberto mais informações sobre o dataset em si e com isso possa ser tomado decisões mais precisas e proposto melhores afirmações. Geralmente, essa preocupação sobre dados anômalos não era tão relevante para a maioria das pesquisas, mas isso vem mudando a partir de que informações importantes podem ser descobertas com essa análise de pontos incomuns.

Esse tipo específico de dado com essas características é chamado de Outliers e é muito importante que os atuais analistas de dados deêm mais atenção para esses dados, pois informações importantes podem estar escondidas entre esses conjuntos particulares.

Por exemplo, se pegarmos um conjunto de dados isolado sobre os fluxo de taxis de Nova Iorque de 2011 e analisarmos a frequência de corridas no ano inteiro, irão aparecer muitos pontos fora dessa curva média e isso indicaria um comportamento anômalo nessa coleção.

Geralmente, a primeira tarefa a se fazer nessa situação era remover esses pontos irregulares e continuar o processamento com o resto do conjunto, mas se pegarmos outro

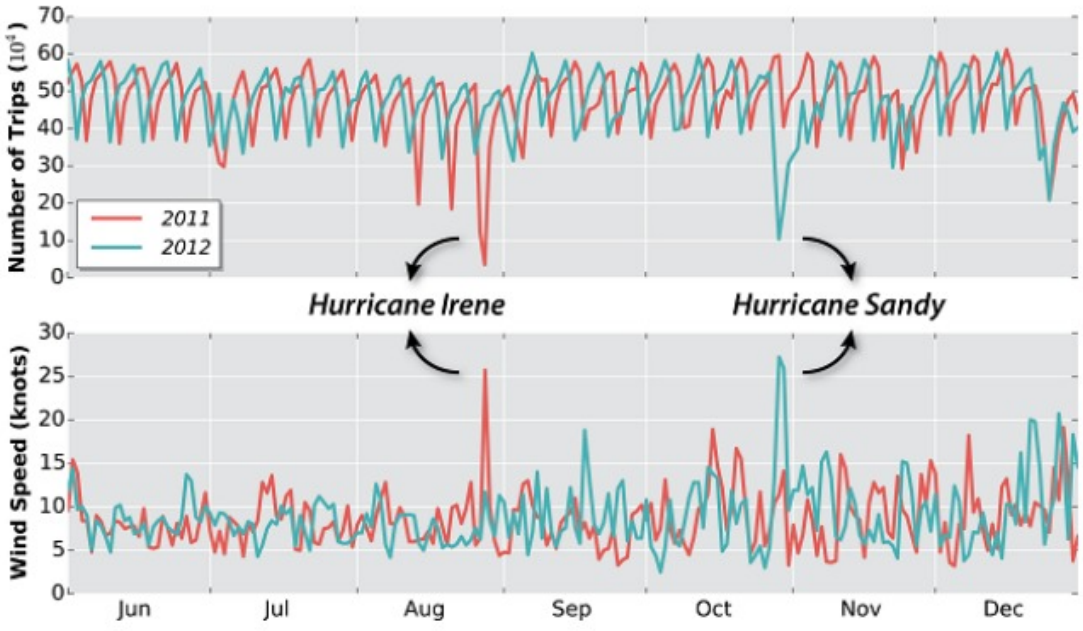

Figura 1: Figura tirada do (FREIRE et al., 2016) mostrando a relação entre o número de corridas de táxis e a velocidade do vento

conjunto de dados isolado sobre a velocidade do vendo na região de Nova Iorque e no mesmo intervalo de tempo, nós iremos perceber algum picos de alta velocidade indicando furacões no mesmo momento e na mesma região da queda das corridas de táxi, como apresentado na Figura 1. Análises como essa provam a importancia de detectar, estudar e interpretar esse outliers para acrescentar o conhecimento obtido de um dataset.

1.1 Contextualização

Hoje em dia nós estamos mais e mais conectados com múltiplas aplicações que acessam um imenso montante dos nossos dados existentes e ainda gera mais dados para melhorar suas análises sobre nós por diveros motivos. Ferramentas como Google Maps, Uber e Waze possuem muitos dados espaciais em tempo real sobre o nosso comportamento em relação ao trafégo (carros, transporte público, taxis, etc.), local de trabalho, locais de viagem frequentes, etc.

Quando tratamos usuários comuns, é muito comum que ele se perca frente à tanta massa de dados espaciais e isso vai prejudicar sua possível análise sobre o conjunto, mesmo a mais simples. Esse problema comum ainda não tem uma solução definitiva, então pesquisas atuais tentam indicar possíveis estratégias para mitigar esse problema e se aproximar de uma solução funcional. Essas abordagens são baseada em: agrupar um grande conjunto

de dados por um ou mais atríbutos específicos e resumir esses grupos baseado nesses atributos para conseguir simples *insights* sobre esses conjuntos, filtrar o dataset para reduzir os dados visíveis e focar em dados específicos para uma análise mais precisa (mas não vasta), e muitas outras estratégias para reduzir a complexidade da análise.

Junto desses problemas mais comuns, existe um importante que acontece antes do primeiro passo da análise que é: *O que fazer quando partes do dataset parecem irregular ou com dados corrompidos?*. Existem técnicas que ajudam na limpeza dessas partes de forma que não prejudique a análise, mas estudos recentes demonstram a importância desses dados *"anormais"* e o quanto um análista pode aprender estudando mais precisamente esse conjunto (FREIRE et al., 2016).

Nesse ambiente complexo de análise de dados espaciais com bastante variáveis e possibilidades, um usuário pode facilmente falhar numa dessas tarefas e comprometer seriamente o resultado de suas análises. Combinando todos esses detalhes, nós sugerimos uma abordagem que leve em consideração o *feedback* do usuário (capturando o movimento do mouse) e, baseado nesse feedback, nós nos tornaremos apto a analisar o interesse do usuário e dentro disso nós podemos detectar, estudar e propor ações para serem tomadas quando um dado considerado um outlier apareça nessa região de interesse do usuário.

1.2 Objetivos

Nesta seção estão definidos os objetivos gerais e específicos do trabalho.

1.2.1 Objetivos Gerais

- Introduzir o problema da análise e visualização em grandes conjuntos de dados espaço temporal atualmente.
- Explicar nossa abordagem proposta para detecção de outliers espaciais em grandes datasets utilizando o conceito de IDR e capturando o feedback do usuário.
- Apresentar nossos resultados utilizando a proposta para detecção de outliers no nosso ambiente espaço-temporal e os benefícios desse experimento.

1.2.2 Specific Objectives

- Analyze the latest researches in the field of outlier detection in spatial-temporal datasets.
- Present our proposed tool for spatial-temporal data analysis and visualization.
- Compare the presented researches showing the pros and cons of each work.
- Describe the concept of IDR used in our tool to mapping the user preference in a spatial-temporal environment.
- Summarize the most known existing outlier detection algorithms for generic and spatial data.
- Display our chosen outlier detection algorithm and explain the reasons for this choice.
- Apply our IDR concept and our chosen outlier detection algorithm in a spatial-temporal data environment.
- Present the results of our application and indicate our future work.

1.3 Work Organization

The document is organized as follows. Section 2 summaries the existing researches in data analysis and visualization field comparing with our proposed tool. Section 3 describes the concepts of IDRs (Interesting Dense Region) and Outliers with the existing algorithms for their detection and our chosen algorithm to detect outliers in our platform. Section 4 explains how we apply the IDRs and outliers detection in the GeoGuide tool. Section 5 presents two applications using distinct datasets to demonstrate the applicability of it to real-world problems and the advantages of this approach. It shows and discusses the outlier detection results. Finally, a conclusion and some directions for future works are given in Section 6.

2 Background

In this chapter we will present the existing works about data analysis and outlier detection with algorithms and strategies of how to use this approach to improve the analysis of data and increase the amount and the quality of the information that we can retrieve from datasets.

2.1 Related Work

2.1.1 GeoGuide

Pursuing improve the spatial data analysis and the guidance approach for this kind of data, the GeoGuide (OMIDVAR-TEHRANI et al., 2017) is a interactive framework that aims to highlight to the analyst a subset of k interesting spatial points, based on the analyst *implicit* (e.g. mouse tracking) and *explicit* (e.g. points clicked) feedbacks, that he may not seem because of the huge amount of information on his screen. This framework considers two metrics to give the subset highlight. The first one is the **relevance** of each point to the point selected by the analyst considering the attributes of those points. The second one is the geographically **diversity** to expand the analyst area in chase of possible new interesting regions.

All this process can be used in generic spatial datasets, as long as each point have two characteristics: geographical attributes (i.e. latitude and longitude) and metadata attributes about its own domain. For instance, the Airbnb[1] platform has open datasets about the available home-stays to rent and each one has the geographical attributes and *price, hostname, availability* as their metadata attributes that are specific for each dataset type as presented in Figure 2. Using this approach, the GeoGuide is the first efficient interactive highlighting framework for spatial data, combining analyst feedbacks with relevance and diversity metrics to display to the analyst a set of interesting points

[1] *http://www.airbnb.com*

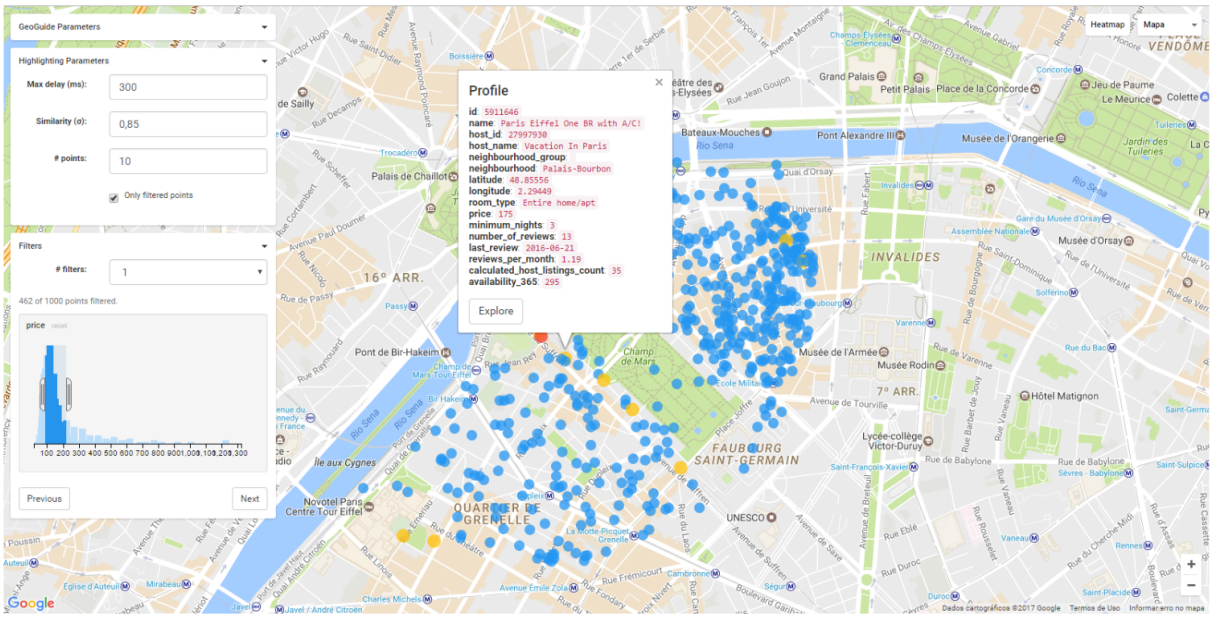

Figura 2: GeoGuide Image on Airbnb Dataset - Paris City

that he may be not focused during his analysis.

2.1.2 Outliers

Outliers in the statistics area are when one finds "aberrant values" in a given series of data, that is when one finds an atypical value or with a great distance from the normal distribution in that set. For example, when a researcher wants to monitor the temperature of his CPU during a certain time interval and it has been realized that the temperature range is between 34 °C and 45 °C degrees being 48 °C the maximum temperature and 27 °C the minimum temperature and in the middle of this sample are some punctual registers of 0 °C, this can be characterized as an outlier and, most likely, will be understood as a malfunction of the equipment that performed the collection of these CPU temperatures.

However, there are several ways to interpret an Outlier (not only as a collection error), but also as: data that belong to a different population of the sample, a damaged data, areas in which a certain theory is not valid or even, when the sample is too large, it is normal to have some small amounts of outliers in that group. In cases where it is proven that it is not the fault of a collection equipment malfunction or that it was not a human mistake, it is extremely important to know the why of that outlier and try to understand it, because it is not interesting for a research simply remove it from the sample or re-signify it by assigning a new value. This change may compromise the validity of the research, and if this is done, it is extremely important to document and record those changes.

As the information technologies improve and continuously increase their computational power, a great variety of algorithms for outlier detection has been surging and applied in different contexts with diverse characteristics and the choice for one of those algorithms is based on the problem domain. Next section will present some of those algorithms with a brief explanation about each one.

2.2 Algorithms

2.2.1 Z-Score

Z-Score is one of the simplest methods for detecting outliers in a dataset. It is a parametric method and takes into account only one attribute per execution. It is also necessary the input of a threshold (usually is a value between 2.5 to 3.5) to be able to define if a given data can be considered an outlier or not. This method is suitable for small datasets that follow the Gaussian distribution.

2.2.2 DBSCAN

It is a density-based spatial clustering algorithm (ESTER et al., 1996) that can be applied in datasets that cannot be presumed what their distribution. It accepts multidimensional datasets (with 3 or more dimensions). However, you need a parameter (MinPts) that defines how many minimum points are needed to form a cluster. Thus, if the size of the set change, this parameter will need to be updated, otherwise the DBSCAN can become inefficient.

2.2.3 Isolation Forests

It is an algorithm of detection of outliers (LIU; TING; ZHOU, 2008) that uses a concept of machine learning that is the decision tree. It is one-dimensional (only takes one attribute at a time) and is required few parameters (this facilitates the configuration and use of the algorithm). No need to climb your values and a very robust algorithm for large datasets.

2.2.4 FDC

It is a depth-based algorithm (JOHNSON; KWOK; NG, 1998) approach for detection of outliers in 2D datasets based on the concept of ISODEPTH algorithm. The FDC computes

the first k 2D depth contours (the points that can be considered outliers) by restricting to a small part of the complete dataset. In this way, it is more efficient by not having to calculate in the complete dataset and thus scaling more easily for large datasets of two dimensions.

2.2.5 HOD

It is a distance-based outlier detection method (XU et al., 2016) that emerges to overcome the statistics-based concept because in the vast majority of datasets the probability distribution is not known. In this way, the method search for outliers based on their distance from their neighbors and if that point has a distance greater than a predefined parameter, then that point is considered an outlier. However, if there is a cluster of outliers in the dataset, this can affect its detection by distance-based algorithms, with this comes the concept of HOD (Hidden Outlier Detection) algorithms that aim to find outliers even when they are grouped and in enough quantity to form a cluster.

2.2.6 ORCA

It is a distance-based algorithm (BAY; SCHWABACHER, 2003) that optimizes a simple nested loop algorithm (which are logarithmic algorithms and extremely inefficient when dealing with large datasets) by removing possible non-outliers during their execution. This way, instead of processing the complete dataset by calculating all possible distances, it removes unnecessary calculations that would be executed if a non-outlier point were taken to the end. From him, new researches have emerged further refining this concept.

2.2.7 Linearization

It is a distance-based algorithm (ANGIULLI; PIZZUTI, 2002) that detects outliers by calculating the sum of the distances of a point in relation to its neighbor, calling it weight, and setting an outlier as the points with the greatest weights in the dataset. In this way, it is an efficient algorithm and it is linearly scaled both in the number of points and in the number of dimensions. To calculate these outliers more efficiently the algorithm uses the concept of the Hilbert space-filling curve.

2.2.8 RBRP

It is an algorithm for high-performance multidimensional datasets that is based on distances between the points to be able to define what the outliers are (GHOTING; PARTHASARATHY; OTEY, 2006). Its difference to the other distance-based algorithms is that it is more efficient for datasets with multiple dimensions and in comparisons with others, its scalability is approximately linear for the number of dimensions and logarithmic for the number of points in the dataset.

2.2.9 LOF

It is a density-based algorithm that adds a new concept in the search for outliers: the Local Outlier Factor (LOF) (BREUNIG et al., 2000), which is a degree of propensity to be an outlier so that the process of outlier definition is not more binary, but something gradual. With this, the approach is not to define whether a point is an outlier or not, but rather the "how outlier" that point is in that dataset. The outlier factor is local in the sense that only a neighborhood of that point is taken into account to define its factor.

2.2.10 ABOD

It is an angle-based algorithm (KRIEGEL; HUBERT; ZIMEK, 2008) for detection of outliers that is focused on high-dimensional datasets, different from other distance-based algorithms that end up damaged when one has many dimensions. Your approach is based on the calculation of a degree of angle between the different vectors of a point with its neighbors. With this, more centralized points within the cluster will have this degree calculated with a high value, the points more on the edge of the clusters will have this degree a little smaller and the possible outliers will have that degree with a very small value, since they will generally be far from the cluster in a particular direction.

Based on the algorithms presented, we organize each one according the answer of proposed questions about outlier detection algorithms in general. The proposed questions are: **I** *Is it parametric?*; **II** *Which is the approach?*; **III** *Is it scalable in performance terms?*; **IV** *Is it multidimensional scalable?* and **V** *Does it receive any argument?*. The result is presented in the Table 1.

Each question has a specific reason to be in the Table 1. The question **I** is about

Algorithms	**I**	**II**	**III**	**IV**	**V**
Z-Score	Yes	Model Based	No	No	Yes
DBSCAN	No	Density Based	No	Yes	Yes
Isolation Forests	No	Depth Based	No	No	No
FDC	No	Depth Based	Yes	No	No
Hidden Outlier Detection	No	Distance Based	Yes	Yes	Yes
ORCA	No	Distance Based	No	Yes	Yes
Linearization	No	Distance Based	Yes	Yes	Yes
RBRP	No	Distance Based	Yes	Yes	Yes
LOF	No	Density Based	No	Yes	Yes
ABOD	No	High Dimensional	No	Yes	No

Tabela 1: Comparison of the presented Outlier Detection Algorithms

the probability distribution of the dataset. If the algorithm is parametric, so we can assume a probability distribution based on a fixed set of parameters. The question **II** serves to classify each algorithm based on his approach for detect a data as an outlier or not, the options are: *Model Based*, *Density Based*, *Depth Based*, *Distance Based*, *High Dimensional*. The question **III** is relative to the performance of each algorithm, if the algorithm execution time is not compromised as the input data increase, it means the algorithm is scalable in perfomance terms. The question **IV** is about the performance of each algorithm too, but in this case is related to the perfomance when the dimension of the data increase. At least, the question **V** indicates if the algorithm require any input argument to process the dataset, this is important because if it requires many arguments, this can be a accuraccy problem when the dataset increase.

3 IDRs and Outliers

3.1 IDRs

TODO: brief description about the next subsections

3.1.1 User feedback

TODO: describe about user feedback and its benefit, *implicit and explicit* feedback with examples: gaze track, mouse click, mouse tracking, mouse track polygons, polygons intersections. use some image to demonstrate the idea

3.1.2 Interesting Dense Regions

TODO: describe about Interesting Dense Regions and its usability and advantages based on mouse track polygons. Use the figure from the IDR paper with description.

3.1.3 Algorithm

TODO: Show the used algorithm for IDR detection and explain how it works. present its complexity with pros and cons

3.2 Outliers

TODO: brief description about the next subsections

3.2.1 Atypical spatial data

TODO: explain about the existing process to capture spatial data and its different sources. show the existing spatial datasets with brief descriptions about each one with

references. summaries the problems of cleaning datasets and the bad analysis with *noise points*

3.2.2 Outliers in GeoGuide

TODO: explain the utility of detect outliers in GeoGuide and its advantages. describe a practical usage with real datasets examples

3.2.3 Algorithm

TODO: show the selected outlier detection algorithm. use image and explain step by step. explain about its complexity. present its pros and cons.

4 Considerações finais

As considerações finais formam a parte final (fechamento) do texto, sendo dito de forma resumida (1) o que foi desenvolvido no presente trabalho e quais os resultados do mesmo, (2) o que se pôde concluir após o desenvolvimento bem como as principais contribuições do trabalho, e (3) perspectivas para o desenvolvimento de trabalhos futuros, como listado nos exemplos de seção abaixo. O texto referente às considerações finais do autor deve salientar a extensão e os resultados da contribuição do trabalho e os argumentos utilizados estar baseados em dados comprovados e fundamentados nos resultados e na discussão do texto, contendo deduções lógicas correspondentes aos objetivos do trabalho, propostos inicialmente.

4.1 Principais contribuições

Texto.

4.2 Limitações

Texto.

4.3 Trabalhos futuros

Texto.

Referências

ANGIULLI, F.; PIZZUTI, C. Fast outlier detection in high dimensional spaces. In: ELOMAA, T.; MANNILA, H.; TOIVONEN, H. (Ed.). *Principles of Data Mining and Knowledge Discovery*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. p. 15–27. ISBN 978-3-540-45681-0.

BAY, S. D.; SCHWABACHER, M. Mining distance-based outliers in near linear time with randomization and a simple pruning rule. In: *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2003. (KDD '03), p. 29–38. ISBN 1-58113-737-0. Disponível em: <http://doi.acm.org/10.1145/956750.956758>.

BREUNIG, M. M. et al. Lof: Identifying density-based local outliers. *SIGMOD Rec.*, ACM, New York, NY, USA, v. 29, n. 2, p. 93–104, maio 2000. ISSN 0163-5808. Disponível em: <http://doi.acm.org/10.1145/335191.335388>.

ESTER, M. et al. A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. AAAI Press, 1996. (KDD'96), p. 226–231. Disponível em: <http://dl.acm.org/citation.cfm?id=3001460.3001507>.

FREIRE, J. et al. Exploring what not to clean in urban data: A study using new york city taxi trips. *IEEE Data Eng. Bull.*, v. 39, n. 2, p. 63–77, 2016. Disponível em: <http://sites.computer.org/debull/A16june/p63.pdf>.

GHOTING, A.; PARTHASARATHY, S.; OTEY, M. E. Fast mining of distance-based outliers in high-dimensional datasets. In: *Proceedings of the 2006 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, 2006. p. 609–613. Disponível em: <https://doi.org/10.1137/1.9781611972764.70>.

JOHNSON, T.; KWOK, I.; NG, R. Fast computation of 2-dimensional depth contours. In: *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining*. AAAI Press, 1998. (KDD'98), p. 224–228. Disponível em: <http://dl.acm.org/citation.cfm?id=3000292.3000332>.

KRIEGEL, H.-P.; HUBERT, M. S.; ZIMEK, A. Angle-based outlier detection in high-dimensional data. In: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: ACM, 2008. (KDD '08), p. 444–452. ISBN 978-1-60558-193-4. Disponível em: <http://doi.acm.org/10.1145/1401890.1401946>.

LIU, F. T.; TING, K. M.; ZHOU, Z. Isolation forest. In: *2008 Eighth IEEE International Conference on Data Mining*. [S.l.: s.n.], 2008. p. 413–422. ISSN 1550-4786.

OMIDVAR-TEHRANI, B. et al. Geoguide: An interactive guidance approach for spatial data. In: *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Exeter, United Kingdom, June 21-23, 2017*. [s.n.], 2017. p. 1112–1117. Disponível em: <https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2017.170>.

XU, H. et al. Index based hidden outlier detection in metric space. *Scientific Programming*, Hindawi Limited, v. 2016, p. 1–14, 2016. Disponível em: <https://doi.org/10.1155/2016/8048246>.

APÊNDICE A – Primeiro apêndice

Os apêndices são textos ou documentos elaborados pelo autor, a fim de complementar sua argumentação, sem prejuízo da unidade nuclear do trabalho.

ANEXO A – Primeiro anexo

Os anexos são textos ou documentos não elaborado pelo autor, que servem de fundamentação, comprovação e ilustração.