

Planejamento da construção de aeroportos na cidade de Belém utilizando o método de clusterização *k-means*

Eduardo Gil S. Cardoso¹, Gabriela S. Maximino¹, Igor Matheus S. Moreira¹

¹ Instituto de Ciências Exatas e Naturais – Faculdade de Computação
Universidade Federal do Pará – Belém, PA – Brasil

{gabriela.maximino,igor.moreira}@icen.ufpa.br

{eduardo.gil.s.cardoso}@gmail.com

Abstract. *This article demonstrates the application of k -means algorithm to the clustering problem involving the construction of three airports in the city of Belém, Pará. This report is part of the deliverable associated to the task proposed by professor Reginaldo Cordeiro dos Santos Filho for the Artificial Intelligence course, taught under the Computer Science Bachelor's degree program at the Federal University of Pará.*

Resumo. *Este artigo demonstra a aplicação do algoritmo k -means para o problema de clusterização envolvendo a construção de três aeroportos na cidade de Belém, Pará. Este trabalho é parte do entregável relativo à tarefa proposta pelo Prof. Dr. Reginaldo Cordeiro dos Santos Filho para a disciplina de Inteligência Artificial, ministrada sob o curso de Bacharelado em Ciência da Computação na Universidade Federal do Pará.*

1. Introdução

A cidade de Belém, capital paraense, é o segundo município mais populoso da região Norte do Brasil [IBGE 2019]. Possuindo 71 bairros divididos em distritos administrativos [Wikipédia 2020], a cidade possui apenas um aeroporto, o Aeroporto Internacional Val-de-Cans. Contudo, em vista da grande demanda de viagens no estado, fez-se necessária a construção de mais três aeroportos para atender de forma igualitária os bairros da cidade, considerando - devido à maior demografia - apenas cinco dos oito distritos: de Belém; do Entroncamento; do Guamá; do Benguí; da Sacramenta. Para realizar a construção, entretanto, é necessário saber a melhor localização para os aeroportos.

Para esse problema, alguns critérios foram considerados: cada novo aeroporto deve atender a pelo menos um bairro e ao máximo de bairros próximos a ele; a distância total dos bairros para os aeroportos deve ser mínima no intuito de evitar o desperdício de recursos; a localização do aeroporto Val-de-Cans também deve ser considerada. Diante desse cenário, nota-se que o problema pode ser resolvido por meio da clusterização - uma técnica de aprendizado não-supervisionado que realiza o agrupamento de dados de acordo com características em comum. Nesse caso, os bairros que possuem as menores distâncias em relação a certo aeroporto serão agrupados em um *cluster*, de forma que o centróide desse seja a localização do aeroporto a ser construído.

Considerando as informações levantadas, escolheu-se o algoritmo *k-means* para realizar a clusterização. Esse método foi considerado ideal devido à utilização de

centróides, cujos valores são atualizados no decorrer da execução, ao redor dos quais as amostras são agrupadas baseadas na distância delas do *cluster*. Além disso, o fato do valor de k (4) ser conhecido e os dados serem do tipo real também fizeram com que a escolha convergisse para esse método.

Diante do exposto, as seções subsequentes estão divididas da seguinte forma: a seção 2 apresenta uma descrição da base de dados construída; a seção 3 descreve o processo de realização do trabalho; a seção 4 apresenta os resultados da aplicação do algoritmo; por fim, a seção 5 sintetiza o trabalho e apresenta as considerações finais.

2. Descrição da base de dados

A base de dados construída é composta pelas coordenadas geográficas do ponto central de cada bairro da cidade de Belém. Como resultado, obteve-se um banco de dados contendo 71 instâncias e 2 *features*, que são 'Bairro' e 'Coordenadas'. As coordenadas estão no formato coord-x, coord-y.

[foto da base?]

3. Metodologia do trabalho

O trabalho foi dividido em três etapas: construção da base de dados; implementação do algoritmo e realização de testes; escrita do artigo.

A construção da base de dados ficou a cargo de um estudante. Para essa etapa, inicialmente verificou-se a lista de bairros da cidade de Belém no site da prefeitura¹. Em seguida, buscou-se a coordenada de cada bairro utilizando a ferramenta Nominatin do projeto *OpenStreetMap*, a qual fornece a delimitação e o ponto central dos bairros. Ao todo, 71 pontos centrais foram coletados e armazenados em uma planilha no formato .csv.

A implementação do algoritmo [...]

Para a escrita do artigo, dois membros do grupo foram necessários, enquanto o terceiro ficou responsável pela revisão geral do trabalho.

4. Resultados

5. Conclusão

Referências

IBGE (2019). https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/7d410669a4ae85faf4e8c3a0a0c649c7.pdf. (Acessada em 23/12/2020).

Wikipédia (2020). Lista dos bairros de Belém (Pará). [https://pt.wikipedia.org/wiki/Lista_de_bairros_de_Belm_\(Par\)](https://pt.wikipedia.org/wiki/Lista_de_bairros_de_Belm_(Par)). (Acessada em 23/12/2020).