

Comparação de Meta-heurísticas de Agrupamento Aplicados em Dados de Animais Bovinos

João V. Santana¹ Washington A. Oliveira² Carla T. L. S. Ghidini³
FCA/UNICAMP, Limeira, SP

Os avanços tecnológicos têm transformado o mundo em que vivemos, impactando todos os aspectos da nossa sociedade, desde a forma como nos comunicamos até a forma como realizamos tarefas diárias [4]. A inteligência artificial (IA) tem permitido que máquinas realizem tarefas antes exclusivas dos seres humanos. Algoritmos sofisticados e redes neurais artificiais são capazes de aprender com dados, tomar decisões com base neles e propor soluções para problemas complexos. Aprendizado de Máquina (AM) é uma área da IA em que algoritmos computacionais são projetados para emular a inteligência humana, aprendendo com o ambiente para melhorar gradualmente sua precisão [5]. O aprendizado não supervisionado é uma abordagem de AM, na qual não há rótulos ou classes conhecidas nos dados de entrada. Esse aprendizado desempenha um papel importante na análise exploratória de dados, ao segmentar o mercado, detectar anomalias, entre outras aplicações em que o objetivo é encontrar padrões e estruturas intrínsecas nos dados, permitindo, assim, uma melhor compreensão do banco de dados como um todo para identificar informações úteis [3]. Nesse processo de segmentação, obter grupos homogêneos (elementos com características similares nos dados), chamados de clusters, refere-se ao agrupamento por similaridade dos elementos por meio do uso de uma métrica particular de similaridade, facilitando a análise dos dados e permitindo o auxílio na tomada de decisões [1]. Ainda no âmbito do AM, as meta-heurísticas têm revelado sua eficácia na solução de problemas de otimização complexos e de grande porte, explorando o espaço de soluções de problemas desafiadores de forma eficiente. Essas técnicas são reconhecidas por sua versatilidade e pela facilidade de adaptação a diferentes problemas [2].

A classificação e o agrupamento de bovinos no pasto, com base em características, são importantes para a gestão eficaz do rebanho, permitindo a identificação de padrões de comportamento, necessidades de cuidados específicos e a otimização dos processos de produção. Este trabalho estudou e aplicou três meta-heurísticas — Algoritmo Genético (AG), Otimização por Enxame de Partículas (PSO) e Otimização por Colônia de Formigas (ACO) — para a tarefa de agrupamento de dados relacionados aos bovinos. O objetivo principal consistiu em avaliar o desempenho dessas abordagens na formação de agrupamentos de bovinos com base na similaridade dos animais. Para isso, foram realizadas análises comparativas dos resultados de cada meta-heurística, considerando diferentes métricas de distância, bem como a avaliação da qualidade dos agrupamentos obtidos por meio de métricas de validação de cluster e análise visual dos agrupamentos. Veja um exemplo ilustrativo na Figura 1 para os algoritmos AG, PSO e ACO.

Nos testes computacionais, as meta-heurísticas demonstraram sua capacidade de realizar com sucesso o agrupamento dos conjuntos de dados *benchmarks*, tendo sido possível confirmar a qualidade e coerência dos agrupamentos obtidos por meio das métricas de avaliação e a visualização dos resultados, bem como constatar que a distribuição e a densidade dos dados desempenham um papel crucial na eficiência das meta-heurísticas utilizadas, sendo a distribuição o fator de maior influência. A variabilidade dos resultados sugere que as meta-heurísticas reagem de forma distinta a depender da natureza dos dados, destacando a importância de estudos aprofundados nesse tópico.

¹victortorsantana@gmail.com

²waoliv@unicamp.br

³cghidini@unicamp.br

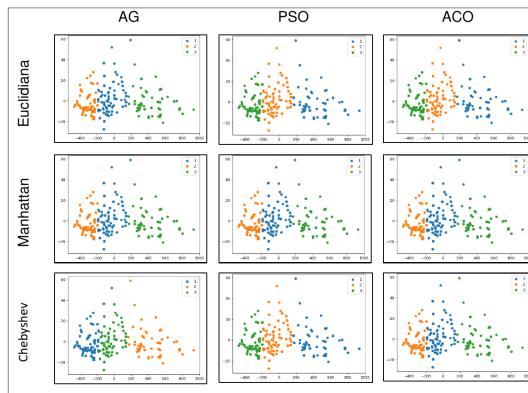


Figura 1: Exemplo de agrupamentos das meta-heurísticas. Fonte: própria.

Particularmente, o banco de dados dos bovinos estudado consiste de treze atributos, dos quais três são medidos diretamente no animal, quatro são por ultrassonografia e seis são atribuídas obtidas por meio de avaliação profissional. A determinação da melhor quantidade de clusters e da meta-heurística com melhor performance ocorreram por meio da análise dos índices de avaliação. No cenário final analisado, dois foi considerado como a melhor quantidade de clusters, a meta-heurística Otimização por Enxame de Partículas utilizando a métrica de distância de Manhattan foi mais promissora no particionamento dos dados, tendo conseguido identificar padrões e diferenças significativas entre os clusters, resultando em agrupamentos coesos e interpretáveis. As características dos dois grupos identificados podem ser utilizadas não apenas para a seleção de reprodutores, mas também para a gestão do rebanho em termos de nutrição, saúde e reprodução. Também abre caminho para a implementação de estratégias de manejo mais direcionadas e personalizadas, considerando as necessidades específicas de cada grupo de animais. Os resultados validaram a capacidade das estratégias de otimização adotadas de fornecer soluções adequadas para problemas de agrupamento, evidenciando seu potencial para aplicações práticas.

Agradecimentos

Os autores agradecem o suporte do CNPq (processo 309925/2021-5) e da FAPESP (processos 2013/07375-0, 2022/05803-3 e 2020/09838-0).

Referências

- [1] G. Caruso, S. A. Gattone, F. Fortuna e T. Di Battista. “Cluster analysis as a decision-making tool: a methodological review”. Em: **Decision Economics: In the Tradition of Herbert A. Simon’s Heritage: Distributed Computing and Artificial Intelligence**, 14th International Conference. Springer. 2018, pp. 48–55.
- [2] K. Hussain, M. N. Mohd Salleh, S. Cheng e Y. Shi. “Metaheuristic research: a comprehensive survey”. Em: **Artificial intelligence review** 52 (2019), pp. 2191–2233.
- [3] S. Solorio-Fernández, J. Carrasco-Ochoa e J. Martínez-Trinidad. “A review of unsupervised feature selection methods”. Em: **Artificial Intelligence Review** 53.2 (2020), pp. 907–948.
- [4] D. Susskind. **A world without work: Technology, automation and how we should respond**. Penguin UK, 2020.
- [5] Z. Ullah, F. Al-Turjman, L. Mostarda e R. Gagliardi. “Applications of artificial intelligence and machine learning in smart cities”. Em: **Computer Communications** 154 (2020), pp. 313–23.