

# Universidade de Pernambuco

## Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

### Proposta de Tese de Doutorado

**Área:** Inteligência Computacional

**Título:** Inteligência de enxames e teste estatísticos para tarefas de clusterização

**Orientador** – Carmelo José Albanez Bastos Filho ([carmelofilho@upe.br](mailto:carmelofilho@upe.br))

**Co-orientador** – Hugo Valadares Siqueira [hugosiqueira@utfpr.edu.br](mailto:hugosiqueira@utfpr.edu.br) UTFPR

#### Descrição

Nos dias atuais, com o avanço das tecnologias de sensoriamento, de transmissão de dados e armazenamento de dados, tornou-se possível a criação de grandes bancos de dados (*data lakes*) não necessariamente estruturados e que podem receber e descartar dados com altas taxas (*big data*). Tais dados podem ser analisados para obter insights que levam a decisões melhores e ações estratégicas de negócio ou reconhecimento e entendimento de um sistema de interesse. Entretanto, o tratamento e processamento destas informações não é uma tarefa trivial. Tal problema atraiu o interesse das comunidades científicas e de desenvolvedores para utilização de abordagens de aprendizado de máquina (*machine learning*), especialmente de aprendizado não supervisionado. Uma etapa fundamental para aprendizado não supervisionado é a etapa de descoberta de padrões a partir de técnicas de aglomeração (do inglês *clustering*), como uma maneira de agregá-las de forma coerente e identificar padrões. Neste sentido, técnicas de Inteligência de Enxame podem ser utilizadas para assessorar técnicas de agrupamento usuais em problemas reais, podendo até substituí-las em alguns casos [1-4]. Todavia, mesmo modelos robustos como os bio-inspirados são geralmente baseados em otimizar métricas baseadas em distância. Testes estatísticos têm mostrado que os agrupamentos formados nem sempre são distintos ao final da tarefa. Assim, este projeto tem como objetivo investigar e implementar novas ferramentas de análise de dados que envolvam a estimação da distribuição de probabilidade e que superem a simples separação em fronteiras hiperdimensionais, como utilizado atualmente nos algoritmos estado da arte.

#### Referências Bibliográficas

1. Figueiredo, E., Macedo, M. G. M., Siqueira, H., Santana JR., C. J., Gokhale, A., Bastos Filho, C. J. A., Swarm intelligence for clustering A systematic review with new perspectives on data mining. Engineering Applications of Artificial Intelligence, v. 82, p. 313-329, 2019.
2. Jose-Garcia, A., Gomez-Flores, W., 2016. Automatic clustering using nature-inspired metaheuristics: A survey. Appl. Soft Comput. 41, 192–213.
3. Kennedy, J., Eberhart, R., Swarm Intelligence, [S.l.: s.n.], ISBN 1-55860-595-9
4. Nanda, S.J., Panda, G., 2014. A survey on nature inspired metaheuristic algorithms for partitional clustering. Swarm Evol. Comput. 16, 1–18.