

Universidade Federal do Sul e Sudeste do Pará  
Instituto de Geociências e Engenharias  
Faculdade de Sistemas de Informação  
Disciplina: Inteligência Artificial    Turma: SI-2023  
Professor: Adam D. F. dos Santos  
Data de abertura: 28/10/2025  
Data de entrega: 16/12/2025  
Data de apresentação: 18/12/2025

## Trabalho Final (10 pontos)

### Instruções iniciais

- O envio do trabalho será exclusivamente via SIGAA na forma de arquivo compactado
- O arquivo compactado deve ser composto pelo menos por uma apresentação (pdf ou pptx, preferencialmente), códigos utilizados e base de dados utilizada (outros arquivos julgados necessários podem ser inclusos)
- Caso a base de dados exceda o tamanho máximo permitido para envio via SIGAA, envie apenas um link para ela
- O trabalho só será avaliado se apresentado na data para tal
- Todos os discentes envolvidos devem **obrigatoriamente** apresentar o trabalho
- Somente serão admitidos grupos de no máximo 4 (quatro) discentes
- Apenas **um** discente do grupo deve enviar o trabalho pelo SIGAA, incluindo os nomes de todos os integrantes do grupo
- **TRABALHOS IGUAIS OU COM SEMELHANÇA INDISCUTÍVEL SERÃO DESCONSIDERADOS**

Pede-se o desenvolvimento de um projeto de *machine learning* (aprendizado de máquina) ou de um projeto de *data science* (ciência de dados), seguindo as orientações apresentadas a seguir.

- Para projetos de *machine learning*, deve-se considerar a aplicação de pelo menos 3 (três) algoritmos de um dos paradigmas (classificação, clusterização, detecção de anomalias, redução de dimensionalidade etc.) estudados na disciplina em uma base de dados com pelo menos complexidade média, levando em consideração a comparação de resultados obtidos por cada algoritmo quando aplicado no problema escolhido;
- Para projetos de *data science*, deve-se considerar a descoberta de conhecimento e conclusões em uma base de dados com pelo menos complexidade média, levando em consideração pelo menos as etapas de pré-processamento (transformação dos dados, criação/exclusão de variáveis, etc.), consultas/seleções, agrupamentos e visualização gráfica;
- Em ambos os casos deve-se gerar uma apresentação que sintetize o domínio escolhido, a base de dados, os algoritmos utilizados, os resultados obtidos, as conclusões derivadas e as referências utilizadas;
- A apresentação mencionada no item anterior conjuntamente com os códigos e a base de dados utilizados formam o entregável mínimo deste trabalho final;
- Indica-se, como fonte primária de consulta de domínio/base de dados, o site: <https://www.kaggle.com/datasets>;
- Obviamente, os domínios e as bases de dados podem ser identificados em fontes diferentes daquela sugerida, assim como os algoritmos, bibliotecas/pacotes, linguagens de programação e softwares não devem se limitar apenas ao que foi apresentado na disciplina.