**PROJETO FINAL**

**DISCIPLINA:** Introdução ao Machine Learning

**CURSO:** Mestrado em Administração Pública

**ÁREA DE CONCENTRAÇÃO**: Ciência de Dados e Inteligência Artificial

**DOCENTE RESPONSÁVEL:** Roberta Moreira Wichmann

**E-MAIL:** [roberta.wichmann@idp.edu.br](mailto:roberta.wichmann@idp.edu.br)

**ANO E BIMESTRE DE REFERÊNCIA:** 2025/2

**ENTREGA PARA A PROFESSORA:** dia31/10/2025

Template Relatório – Atividade 2 (**Adicionar o Título Do Projeto**)

Nome do Aluno(a): (**Adicionar o nome completo**)

**1. Introdução**

Nesta seção, o discente deverá contextualizar o conjunto de dados utilizado, relembrando brevemente sua origem e os objetivos da análise. Também é importante descrever de forma geral as etapas que serão abordadas no relatório (Fonte: time news roman, tamanho 12, alinhamento justificado e espaçamento 1,5).

**2. Metodologia**

Descrever a sequência de passos adotados no desenvolvimento do trabalho, destacando os principais procedimentos:

● Análise descritiva preliminar.

● Divisão treino e teste.

● Pré-processamento e transformações.

● Construção e escolha do modelo.

● Métricas utilizadas.

● Otimização de hiperparâmetros.

● Avaliação final do modelo.

**3. Resultados e Discussões**

Nesta seção, o discente deve apresentar os resultados obtidos em cada etapa do trabalho e interpretá-los de forma crítica.

● Resultados da descritiva preliminar: destacar padrões relevantes, outliers encontrados, correlações importantes, insights extraídos dos gráficos.

● Resultados dos Modelos Testados: apresentar tabelas ou gráficos com métricas de desempenho (ex.: acurácia, precisão, recall, F1, AUC-ROC para classificação; RMSE, MAE, R² para regressão).

● Discussão Crítica: comparar os modelos entre si, apontar qual teve melhor desempenho e justificar o motivo (ex.: “O Random Forest obteve melhor resultado pois conseguiu capturar relações não lineares que a regressão linear não modelou bem”).

● Limitações: mencionar eventuais problemas, como poucos dados disponíveis, variáveis pouco informativas ou necessidade de mais pré-processamento.

**4. Conclusão**

Aqui o discente deve resumir o trabalho, destacando:

● Principais descobertas do EDA.

● Importância do pré-processamento na qualidade do modelo.

● Modelo que apresentou o melhor desempenho e por quê.

● Potenciais melhorias (ex.: coletar mais dados, incluir novas variáveis, testar algoritmos mais sofisticados, aplicar técnicas de balanceamento de classes).

**5. Referências (Se tiver)**

Nesta seção devem ser listados todos os materiais utilizados para embasar o trabalho, como:

● Livros (ex.: James, G. et al. An Introduction to Statistical Learning).

● Artigos científicos.

● Documentação oficial de bibliotecas Python (scikit-learn, pandas, matplotlib etc.).

● Tutoriais, sites e materiais de aula.

**Deixar na formatação segundo a ABNT.**

**FIM!**