

Serviço Nacional de Aprendizagem Comercial - Faculdade Senac - Jaraguá do Sul/SC

Curso: Programador de Sistemas

Disciplina: Implementar Banco de Dados

Prof: Cesar Ricardo Velasque Trindade

Data: 18 de março de 2025

Atividade: O naufrágio do Titanic e o desafio de predição no Kaggle

Nome dos alunos:

---



## 1 O naufrágio do Titanic e o desafio de predição no Kaggle.

O RMS Titanic foi um dos transatlânticos mais famosos da história, marcado tanto por sua grandiosidade quanto por sua tragédia. Construído para ser um dos navios mais luxuosos e seguros de sua época, o Titanic partiu de Southampton, na Inglaterra, em 10 de abril de 1912, com destino a Nova York, nos Estados Unidos. Entretanto, na noite de 14 de abril de 1912, colidiu com um iceberg e afundou nas primeiras horas do dia seguinte, resultando na morte de mais de 1.500 passageiros e tripulantes.

A tragédia do Titanic impactou profundamente a sociedade da época, levando a mudanças significativas nas leis de segurança marítima. Anos mais tarde, sua história foi imortalizada no cinema, com destaque para o filme "Titanic" de 1997, dirigido por James Cameron. A produção foi um enorme sucesso de bilheteria, conquistando 11 Oscars e se tornando um dos filmes mais assistidos de todos os tempos.

Paralelamente, o naufrágio do Titanic tornou-se uma referência em análise de dados. Na plataforma Kaggle, existe um desafio de aprendizado de máquina baseado nos dados do desastre. O objetivo desse desafio é construir um modelo preditivo capaz de determinar quais passageiros teriam sobrevivido ao naufrágio com base em um conjunto de variáveis.

O dataset do Titanic disponibilizado no Kaggle contém diversas informações sobre os passageiros, incluindo:

- **PassengerId:** Identificação do passageiro.
- **Survived:** Indica se o passageiro sobreviveu (1) ou não (0).
- **Pclass:** Classe da cabine do passageiro (1<sup>a</sup>, 2<sup>a</sup> ou 3<sup>a</sup> classe).
- **Name:** Nome completo do passageiro.
- **Sex:** Sexo do passageiro.
- **Age:** Idade do passageiro.
- **SibSp:** Número de irmãos ou cônjuges a bordo.
- **Parch:** Número de pais ou filhos a bordo.
- **Ticket:** Número do bilhete.
- **Fare:** Valor da tarifa paga pelo passageiro.
- **Cabin:** Número da cabine.
- **Embarked:** Porto de embarque (C = Cherbourg, Q = Queenstown, S = Southampton).

Com esses dados, cientistas de dados podem criar modelos de aprendizado de máquina utilizando técnicas como regressão logística, árvores de decisão e redes neurais para prever a sobrevivência dos passageiros. Esse desafio é um excelente ponto de partida para iniciantes no campo da ciência de dados, pois permite explorar conceitos fundamentais de análise exploratória, limpeza de dados e modelagem preditiva.

Assim, o naufrágio do Titanic, além de ser um dos eventos mais marcantes da história marítima e cinematográfica, também serve como uma valiosa fonte de estudo na área de ciência de dados e aprendizado de máquina.