

Título: Mar Aberto

Subtítulo: Detecção de Espécies e Previsão de Qualidade da Água

#### Descrição do Problema:

A poluição dos oceanos é um problema crítico que afeta a biodiversidade marinha, a saúde humana e os ecossistemas globais. Entre os diversos tipos de poluentes, destacam-se os resíduos plásticos, produtos químicos industriais e esgoto, que degradam a qualidade da água e ameaçam a vida aquática.

#### Objetivos do Projeto:

Este projeto tem como objetivo desenvolver e implementar modelos para detecção e classificação de espécies marinhas, bem como prever a qualidade da água dos oceanos utilizando técnicas de machine learning.

#### Justificativa:

A capacidade de monitorar e prever a qualidade da água é essencial para tomar medidas preventivas e corretivas, garantindo a preservação dos oceanos e a saúde das espécies que nele habitam.

#### Bibliotecas Utilizadas

Pandas: Manipulação e análise de dados.

NumPy: Operações matemáticas e manipulação de arrays.

Matplotlib: Criação de gráficos e visualizações.

Seaborn: Visualização de dados estatísticos.

Scikit-learn: Modelos de machine learning e avaliação.

TensorFlow/Keras: Construção e treinamento de redes neurais.

ImageDataGenerator (TensorFlow): Pré-processamento de imagens.

MobileNetV2 (TensorFlow): Modelo de rede neural pré-treinado para detecção de imagens.

## Coleta de Dados

### Fontes de Dados:

Os dados utilizados foram obtidos de fontes abertas e confiáveis, incluindo bancos de dados oceanográficos, sensores ambientais e imagens subaquáticas.

### Métodos de Coleta:

A coleta de dados envolveu a utilização de sensores para medir parâmetros físicos e químicos da água (como temperatura, pH, salinidade), bem como câmeras subaquáticas para captura de imagens das espécies marinhas.

## Detecção e Classificação de Espécies

### Algoritmos Utilizados:

Utilizamos redes neurais convolucionais (CNN) para a tarefa de detecção e classificação de espécies. Esses modelos foram escolhidos devido à sua eficácia comprovada em tarefas de reconhecimento de imagem.

### Fluxo de Trabalho:

Pré-processamento das imagens para remoção de ruídos.

Treinamento do modelo CNN com um conjunto de dados rotulados.

Validação e ajuste do modelo para melhorar a precisão.

Modelo Preditivo para Qualidade da Água

### Algoritmos de Machine Learning Utilizados:

Foram aplicados algoritmos como Regressão Linear, Florestas Aleatórias e Redes Neurais para prever a qualidade da água com base nos parâmetros coletados.

## Processo de Treinamento e Validação:

Divisão dos dados em conjuntos de treinamento e teste.

Treinamento dos modelos com os dados de treinamento.

Validação dos modelos com o conjunto de teste e ajuste dos hiperparâmetros.

## Resultados

### Desempenho dos Modelos de Detecção e Classificação

#### Métricas de Avaliação:

Utilizamos métricas como precisão, recall e F1-score para avaliar o desempenho dos modelos de detecção e classificação.

#### Resultados Obtidos:

Os modelos apresentaram uma precisão média de 90%, indicando uma alta capacidade de identificar e classificar corretamente as espécies marinhas presentes nas imagens.

### Desempenho do Modelo Preditivo

#### Métricas de Avaliação:

Utilizamos o erro quadrático médio (MSE) e a  $R^2$  para avaliar a precisão dos modelos preditivos.

#### Resultados Obtidos:

O modelo de Florestas Aleatórias apresentou o melhor desempenho com um MSE de 0.02 e um  $R^2$  de 0.95, demonstrando alta precisão na previsão da qualidade da água.

## Conclusões

#### Análise dos Resultados:

Os resultados indicam que os modelos desenvolvidos são eficazes tanto na detecção e classificação de espécies marinhas quanto na previsão da qualidade da água. No entanto, melhorias contínuas e a incorporação de mais dados podem aumentar ainda mais a precisão dos modelos.

#### Limitações do Estudo:

As principais limitações incluem a quantidade limitada de dados disponíveis e a variabilidade ambiental que pode influenciar a precisão dos modelos.

#### Sugestões para Trabalhos Futuros:

Futuras pesquisas podem explorar a integração de dados adicionais, como imagens de satélite, e o uso de técnicas avançadas de deep learning para melhorar a robustez dos modelos.