

# Previsão de Qualidade de Água

Bruno Ferreira de Lima 112389

*Eng.Computação.*

*Fundação Hermínio Ometto*

Araras, Brasil

brunoferreira48@alunos.fho.edu.br

Marina de Souza Pina Oliveira 111838

*Eng.Computação.*

*Fundação Hermínio Ometto*

Araras, Brasil

marina.oliveira@alunos.fho.edu.br

**Abstract**—Create an artificial intelligence that calculates and analyzes data about drinking water for humans.

**Index Terms**—

## I. INTRODUÇÃO

A água é um recurso vital para a vida, essencial não apenas para o consumo humano, mas também para a agricultura, a indústria e a manutenção dos ecossistemas. No entanto, a crescente poluição causada por atividades humanas, como o descarte inadequado de resíduos e o uso excessivo de produtos químicos, tem comprometido a qualidade da água em diversas regiões. Essa situação representa um sério risco à saúde das populações e ao equilíbrio ambiental, tornando urgente o desenvolvimento de métodos eficazes para monitorar e prever a qualidade da água.

Nesse cenário, o uso de técnicas avançadas de aprendizado de máquina, como o Support Vector Machine (SVM), surge como uma solução promissora. O SVM é capaz de analisar grandes volumes de dados complexos e identificar padrões que ajudam a antecipar mudanças na qualidade da água, possibilitando ações preventivas e mais eficientes.

Este trabalho busca explorar o potencial do SVM na previsão da qualidade da água avaliando parâmetros como PH, índice de turbidez e índice de coliformes fecais.

## II. TRABALHOS RELACIONADOS

O trabalho fundamental para este projeto aborda a aplicação de diversas técnicas de aprendizado de máquina com o objetivo de analisar o índice de qualidade da água, considerando múltiplos indicadores de qualidade. A pesquisa avalia o desempenho de cada modelo e os compara utilizando variáveis de controle, como a taxa de acerto.

O principal objetivo deste estudo consiste no desenvolvimento de um modelo preditivo com base na técnica de Support Vector Machine (SVM). A escolha deste algoritmo foi fundamentada na análise comparativa do artigo de referência, que demonstrou a sua eficácia na classificação de águas em níveis de "regular" e "ótima" [1].

As variáveis selecionadas para a análise da qualidade da água são baseadas nos parâmetros que compõem o Índice de Qualidade das Águas (IQA), conforme definido pela Agência Nacional de Águas (ANA). Os indicadores utilizados incluem oxigênio dissolvido, coliformes termotolerantes, potencial hidrogeniônico (pH), temperatura da água, nitrogênio total,

fósforo total, turbidez e sólidos totais [2], por meio desses índices o modelo SVM será capaz de classificar corretamente a qualidade da água.

## III. METODOLOGIA

### III. METODOLOGIA

Este trabalho foi desenvolvido com o objetivo de construir um modelo preditivo capaz de avaliar a qualidade da água com base em parâmetros físico-químicos e biológicos. O processo foi dividido em quatro etapas principais: coleta e preparação dos dados, tratamento e normalização, treinamento do modelo SVM, e avaliação dos resultados.

#### A. Coleta e preparação dos dados

Os dados utilizados para o treinamento do modelo foram obtidos de bases públicas de monitoramento ambiental disponibilizadas pela Agência Nacional de Águas (ANA) e complementadas por conjuntos de dados abertos referentes ao Índice de Qualidade da Água (IQA). As variáveis consideradas incluem os principais parâmetros de classificação: oxigênio dissolvido, coliformes termotolerantes, pH, temperatura da água, nitrogênio total, fósforo total, turbidez e sólidos totais. Cada registro representa uma amostra de um corpo d'água analisado em determinado local e data, acompanhado do rótulo de classificação de qualidade, categorizado em níveis como ótima, boa, regular e ruim.

#### B. Tratamento e normalização dos dados

Antes do treinamento, os dados passaram por uma etapa de limpeza, removendo registros incompletos e valores inconsistentes. Em seguida, todas as variáveis numéricas foram normalizadas para o intervalo [0,1], garantindo que nenhuma característica dominasse o processo de aprendizado devido à diferença de escala. Também foi realizada uma análise exploratória de dados (EDA), com o objetivo de compreender a distribuição das variáveis, identificar correlações e verificar a relevância de cada parâmetro para o modelo.

#### C. Treinamento do modelo SVM

O algoritmo Support Vector Machine (SVM) foi escolhido pela sua eficiência na classificação de dados não lineares e pela capacidade de generalização em conjuntos de dados complexos. Para este trabalho, utilizou-se a implementação disponível na biblioteca scikit-learn da linguagem Python, com experimentos realizados nos modos linear e RBF (Radial Basis

Function) para comparar o desempenho de diferentes funções de kernel. Os dados foram divididos em 70

#### D. Avaliação e métricas de desempenho

A etapa final consistiu na avaliação do modelo por meio de métricas amplamente utilizadas em problemas de classificação, tais como acurácia, precisão, revocação (recall) e F1-score. Adicionalmente, foi gerada uma matriz de confusão para análise visual dos acertos e erros em cada classe. Os resultados foram interpretados considerando a correspondência entre as previsões e os valores reais do índice de qualidade da água, permitindo validar a eficácia do modelo SVM como ferramenta de apoio ao monitoramento ambiental.

## IV. RESULTADOS

## V. CONCLUSÃO

## REFERENCES

- [1] NASCIMENTO, Mário Elias Carvalho do. Previsão da qualidade da água utilizando técnicas de aprendizado de máquina. 2020. [Trabalho de Conclusão de Curso (Graduação em Engenharia Ambiental)] – Universidade Federal do Paraná, Curitiba, 2020. Disponível em: <https://acervodigital.ufpr.br/xmlui/bitstream/handle/1884/89304/R> Acesso em: 20 Setembro. 2025.
- [2] AGÊNCIA NACIONAL DE ÁGUAS (ANA). Índice de Qualidade das Águas (IQA). Brasília, [2024?]. Disponível em: <https://www.ana.gov.br/portalpnqa/indicadores-indice-aguas.aspx>. Acesso em: 29 set. 2025.