

IMPLEMENTAÇÃO DE INTELIGÊNCIA ARTIFICIAL UTILIZANDO REDES NEURAIS ARTIFICIAIS (RNA) NA CLASSIFICAÇÃO DE VIDROS

Marcelo Pontes Rodrigues¹; Marcelo de Oliveira Rego²

¹ Programa de Pós-Graduação em Tecnologias Aplicadas a Animais de Interesse Regional–UFPI;

² Programa de Doutorado em Ciência da Computação Associação UFMA/UFPI/DCCMAPI/CCET

¹marcelo.rodrigues@ufpi.edu.br; ²oliveira@ufdpar.edu.br;

Resumo

A pesquisa sugeriu uma abordagem para classificação de tipos de vidro utilizando o conjunto de dados "Glass Identification" disponível na UCI (University of California, Irvine)(German 1987). O objetivo é identificar a composição de vidro com base em variáveis químicas e físicas, diferenciando entre tipos comuns e especiais. Para isso, o dataset foi pré-processado e normalizado, e foram projetadas Redes Neurais Artificiais (RNA) com uma camada oculta, onde foram treinadas para distinguir entre as classes. Para isso, foram abordadas neste trabalho quatro hipóteses: (i) Hipótese 1, dados com as 7 classes e sem modificações; (ii) Hipótese 2, dados com as 7 classes e retirando a coluna "Ca"; (iii) Hipótese 3, problema reduzido a 3 classes e retirando a coluna "Ca"; (iv) Hipótese 4, divisão em Flutuado e Não Flutuado excluindo a classe "Outros"; e (v) Hipótese 5, divisão em Flutuado e Não Flutuado. Com isso, culminou em um modelo eficaz com a Hipótese 4, com Acurácia 86,4%. Os achados sugerem que RNAs são ferramentas promissoras para a análise composicional de materiais.

Palavras-chave: vidro, machine learning e Redes Neurais Artificiais.

INTRODUÇÃO

O objetivo desta pesquisa é desenvolver um modelo de classificação de vidros com base em sua composição química. Utilizou-se o conjunto de dados "Glass Identification" do repositório UCI Machine Learning(German 1987), que contém informações detalhadas sobre a composição de diferentes tipos de vidros, como os usados em janelas e em ambientes especiais. Originalmente criado para fornecer uma base confiável e padronizada para investigações forenses e estudos em engenharia de materiais, este conjunto de dados permite a classificação automatizada de amostras, o que pode auxiliar na identificação rápida e precisa de materiais, economizando tempo e recursos. Neste estudo, exploramos redes neurais artificiais (RNAs) para a classificação dos tipos de vidro, com base nas propriedades químicas e físicas do material. O uso de RNAs é especialmente adequado, pois essas redes são capazes de detectar padrões complexos e não lineares em dados multidimensionais (2), uma característica essencial para analisar a diversidade nas composições químicas deste dataset.

METODOLOGIA

Inicialmente foi realizado uma avaliação dos dados e a correlação das features para auxiliar nas opções a serem escolhidas no modelo de RNAs e definir as hipóteses, conforme Figura 1.

O estudo foi conduzido de forma abrangente, revisitando cada etapa várias vezes para avaliar, ajustar e validar as hipóteses propostas. Em cada iteração, o tamanho do conjunto de testes foi mantido constante em 30% para efeitos de comparação, e análises detalhadas foram realizadas para decidir sobre a aceitação ou rejeição das hipóteses. Foram abordadas neste trabalho cinco hipóteses:

- (i) Hipótese 1, dados com as 7 classes e sem modificações;
- (ii) Hipótese 2, dados com as 7 classes e retirando a coluna "Ca";
- (iii) Hipótese 3, problema reduzido a 3 classes e retirando a coluna "Ca";
- (iv) Hipótese 4, divisão em Flutuado e Não Flutuado excluindo a classe "Outros"; e
- (v) Hipótese 5, divisão em Flutuado e Não Flutuado

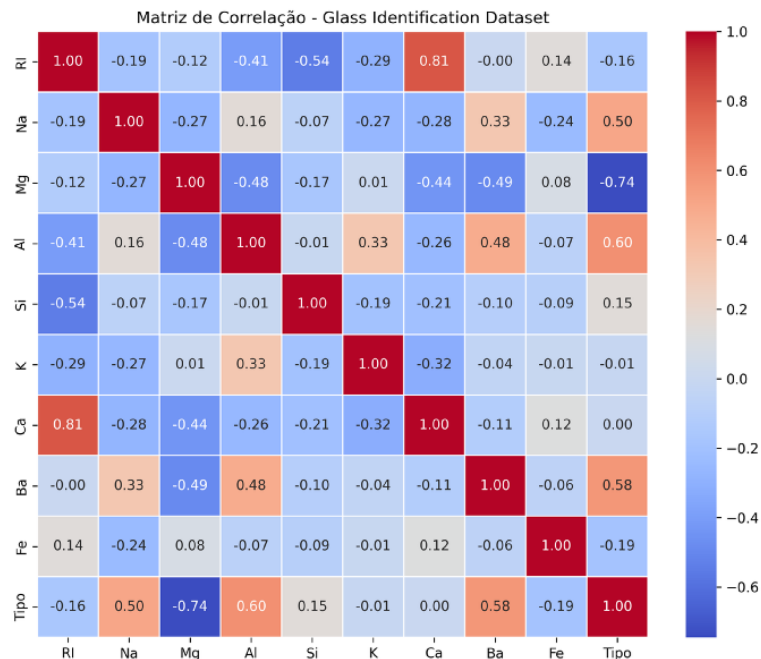


Figura 1: Matriz de Correlação

O estudo foi implementado na linguagem python e está disponível no repositório do GitHub (Pontes e Rego 2024)

RESULTADOS E DISCUSSÕES

Os resultados obtidos após a aplicação das diferentes hipóteses e ajustes de parâmetros demonstraram variações significativas nas métricas de desempenho. Percebemos que com 30% dos dados sendo destinados para o teste obtivemos as melhores métricas, portanto fizemos todos os testes subsequentes já com essa proporção.

Na Hipótese 1, dados com as 7 classes e sem modificações, foram utilizados 32 neurônios na camada escondida, 7 neurônios na camada de saída, taxa de aprendizagem de 0.1 e foi obtido acurácia de 72%.

Na Hipótese 2, dados com as 7 classes e retirando a coluna "Ca", foram utilizados 32 neurônios na camada escondida, 7 neurônios na camada de saída, taxa de aprendizagem de 0.1 e foi obtido acurácia de 71%.

Na Hipótese 3, problema reduzido a 3 classes e retirando a coluna "Ca", foram utilizados 32 neurônios na camada escondida, 3 neurônios na camada de saída, taxa de aprendizagem de 0.1 e foi obtido acurácia de 75%.

Na Hipótese 4, divisão em Flutuado e Não Flutuado excluindo a classe "Outros", foram utilizados 24 neurônios na camada escondida, 2 neurônios na camada de saída, taxa de aprendizagem de 0.06 e foi obtido acurácia de 86%.

Na Hipótese 5, divisão em Flutuado e Não Flutuado, foram utilizados 32 neurônios na camada escondida, 2 neurônios na camada de saída, taxa de aprendizagem de 0.1 e foi obtido acurácia de 72%.

CONCLUSÃO

O estudo confirmou a eficácia de abordagens específicas para otimizar a precisão e robustez do modelo de classificação. A Hipótese 4, divisão em Flutuado e Não Flutuado excluindo a classe "Outros" foi que apresentou melhor acurácia 86,4%. Esses resultados indicam que a simplificação do conjunto de variáveis pode ser uma estratégia eficaz para evitar redundâncias e melhorar o desempenho do modelo.

Esses insights estabelecem uma base robusta para futuras pesquisas, especialmente no aprimoramento de modelos preditivos em cenários onde a correlação entre variáveis desempenha um papel significativo.

Referências

GERMAN, B. *Glass Identification*. 1987. UCI Machine Learning Repository. DOI: <https://doi.org/10.24432/C5WW2P>.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.

PONTES, M.; REGO, M. *Repositório de Inteligência Artificial utilizando Árvore de Decisão no Diagnóstico Câncer de Mama*. 2024. Acessado em: 14 de novembro de 2024. Disponível em: <https://github.com/infopontes/inteligencia_artificial/>.