Estimativa do teor de boro no solo por meio de técnicas de aprendizado de máquina

Rafael Schultz Drzevieski (PIBIC/CNPq), Alaine Margarete Guimarães (Orientadora), e-mail: alainemg@uepg.br

Universidade Estadual de Ponta Grossa/ Departamento de Informática

Ciência da Computação: Metodologia e Técnicas de Computação.

Palavras-Chave: *Computação Aplicada à Agricultura, Aprendizado de máquina, Boro, Espectroscopia NIR, Modelo de predição.*

Resumo

O intuito deste projeto foi de gerar um modelo preditivo, através de técnicas de aprendizado de máquina, capaz de predizer o teor de boro presente em amostras de solo utilizando dados de espectroscopia de refletância difusa no espectro NIR.

Introdução

O setor agropecuário é um dos pilares da economia brasileira, sendo indispensável tanto no mercado interno como externo. Diante disso, é imprescindível o desenvolvimento de novas técnicas que mantenham esta área otimizada. Uma de tais tecnologias é a espectroscopia de reflectância difusa no espectro do infravermelho próximo (NIR), capaz de auxiliar na análise dos elementos físicos e químicos presentes no solo. Porém, para que o uso deste método seja possível, é necessário a determinação de um modelo numérico capaz de estimar o teor do elemento presente em uma amostra, com base na resposta espectral obtida em um espectrofotômetro. Entre os elementos está o boro (B) que por ser um micronutriente no solo, é de difícil estimativa por espectroscopia. O objetivo deste trabalho foi criar um modelo de estimativa para o B em amostras de solo utilizando espectroscopia NIR e técnicas de aprendizado de máquina.

Materiais e Métodos

O estudo do trabalho de Proença (2012) e de Zhu et al. (2022) foi fundamental para a compreensão do tema. Para a construção do modelo de estimativa nesse subprojeto, optou-se pela utilização da ferramenta computacional WEKA (FRANK, 2016), que contempla vários algoritmos para aprendizado de máquina. A base de dados espectral estudada caracteriza-se pela alta dimensionalidade, contendo um total de duzentos e setenta e um atributos, dos quais duzentos e setenta e cinco compõem os dados espectrais.

Resultados e Discussão

Devido à alta complexidade da base de dados, a análise dos valores individuais de B não gerou resultados com coeficiente de correlação suficiente para ser utilizado em aplicações práticas. Devido a isto, optou-se pela discretização em classes binárias de igual frequência, para manter o balanceamento das classes. A primeira classe contempla os valores de 0 a 0,165mg/dm³ e a segunda classe variando de 0,165mg/dm³ ao valor máximo. Os algoritmos utilizados nesta abordagem foram BayesNet, Jrip e Roughset, sendo que os três foram capazes de gerar modelos com taxa de acerto próximos de 74%.

Conclusão ou Conclusões

O trabalho atingiu resultados com modelos preditivos apresentando acurácia em torno de 74%. Apesar das limitações observadas, principalmente relacionadas à alta dimensionalidade da base de dados, os resultados indicam potencial para o uso de técnicas de aprendizado de máquina na estimativa do teor de boro no solo.

Agradecimentos

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, pela bolsa concedida.

FRANK, E; HALL, M.; WITTEN, I. **The WEKA Workbench**. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

PROENÇA, C. **Redes Neurais Artificiais para predição dos teores de matéria orgânica e argila do solo na Região dos Campos Gerais utilizando espectroscopia de reflectância difusa**. 2012. Dissertação (Mestrado em Computação Aplicada) – Universidade Estadual de Ponta Grossa, Ponta Grossa, 2012.

ZHU, J.; JIN, X.; LI, S.; HAN, Y.; ZHENG, W. **Prediction of Soil Available Boron Content in Visible-Near-Infrared Hyperspectral Based on Different Preprocessing Transformations and Characteristic Wavelengths Modeling**. Computational Intelligence and Neuroscience, v. 2022, n. 1, 2022.