



## REAMOSTRAGEM EM REDES NEURAIS COM APLICAÇÃO A DADOS ESPACIAIS DA FAZENDA JULIANA

**Bruno Póvoa Rodrigues<sup>1</sup>, Vinícius Francisco Rofatto<sup>1</sup>, Marcelo Tomio Matsuoka<sup>1</sup>, Talita Teles Assunção<sup>1</sup>**

<sup>1</sup> Universidade Federal de Uberlândia, Monte Carmelo, Minas Gerais  
(brunopovoar@gmail.com).

**RESUMO:** No desenvolvimento de redes neurais artificiais (RNA) o conjunto de dados disponível é dividido em três categorias: treinamento, validação e teste. No entanto, surge um problema importante: como confiar na predição fornecida por uma única RNA? Devido à aleatoriedade relacionada à própria RNA (arquitetura, inicialização e procedimento de treinamento), geralmente, não existe melhor escolha. Para capturar a aleatoriedade intrínseca à RNA, uma abordagem baseada no método Jackknife de reamostragem estatística é apresentada. O Jackknife clássico consiste em remover uma observação do conjunto de dados disponíveis ( $n$ ) e usar as  $(n - 1)$  amostras restantes no processo de estimação. Este processo é repetido para cada observação individual. Ao final, terá  $n$  estimativas advindas de amostras diferentes. No caso de redes neurais, cada observação individual é selecionada para compor o conjunto de teste, enquanto o restante da amostra é destinado para o treinamento da rede. Nesse caso, o número de redes neurais é igual ao tamanho dos dados disponíveis. Entretanto, estendeu-se a ideia ao replicar esse procedimento por um inúmeras vezes. Logo, devido à característica aleatória da rede neural, as predições variam para um mesmo ponto amostral. Consequentemente, é possível descrever a distribuição de cada predição individual. Portanto, o método proposto fornece predições intervalares ao invés da tradicional predição pontual. O método proposto foi aplicado e testado utilizando dados cedidos pela Fazenda Juliana contendo potencial de hidrogênio, cálcio trocável e concentração de fósforo obtidos por meio da análise de 118 amostras de solos georreferenciados. Os resultados mostraram que a redução de 60% no conjunto de dados disponível oferece acurácia compatível em relação ao conjunto de dados completo e, portanto, reduz custos de amostragem em campo. Na avaliação externa, ou seja, na análise de dados que não participaram da reamostragem, sendo possível observar que a taxa de sucesso é maior quando usamos a predição intervalar em vez de usar a predição média. Embora esteja restrita a aplicação em redes neurais, o método proposto pode ser estendido a outras ferramentas estatísticas modernas, tais como Krigagem, e outros dados espaciais como gradientes ionosféricos, pressão atmosférica, e mais.

**PALAVRAS-CHAVE:** reamostragem, delete-1 jackknife, análise espacial



**AGRADECIMENTOS:** Os mais sinceros agradecimentos à Fazenda Juliana, por incentivar a ciência cedendo os dados mais que essenciais para que essa pesquisa fosse realizada, ao Alfredo P. de Oliveira Junior, Gerente Administrativo e Financeiro, que nos recebeu tão solícitamente.