



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



How to describe the trend, accuracy and uncertainty of predictions in Artificial Neural Networks?

CLAUDIONEI PEREIRA DA CUNHA FILHO¹, STEFANY DIOVANA CUNHA COSTA²,
VINICIUS FRANCISCO ROFATTO³, MARCELO TOMIO MATSUOKA⁴

RESUMO

No desenvolvimento de redes neurais artificiais, como em qualquer algoritmo de *machine learning*, é comum separar os dados relativos ao problema em um conjunto de treinamento e um conjunto de testes. Os dados de treinamento são utilizados para estimar os parâmetros inerentes à rede, e os dados de teste são empregados para verificar a performance da rede sob condições reais de utilização. Entretanto, esse método apresenta algumas limitações: (i) os resultados produzidos são baseados em uma divisão não controlada; (ii) a divisão inadequada do conjunto de dados pode ter um efeito adverso no desempenho do modelo; (iii) não é adequado em situações em que os dados são insuficientes ou esparsos. Portanto, como confiar nas predições fornecidas por uma única rede neural artificial (RNA) em particular? Devido à aleatoriedade relacionada à própria rede (arquitetura, inicialização e procedimento de aprendizado), geralmente não há a melhor escolha. Considerando essa questão, fornecemos uma abordagem que captura a incerteza relacionada à própria rede. Consequentemente, a predição por intervalo está disponível em vez da clássica predição pontual. O método proposto foi aplicado em um problema de interpolação de dados para geração de um Modelo Digital do Terreno (MDT). O método caracteriza espacialmente os pontos de maior ou menor tendência, acurácia e incerteza. Apesar de restringirmos ao modelo de rede neural, o método proposto também pode ser estendido a outras ferramentas estatísticas modernas, como *Krigagem*, Colocação por Mínimos Quadrados, entre outros métodos preditivos.

Palavras-chave: Redes Neurais. Reamostragem. Incerteza.

1 INTRODUÇÃO

Uma das tarefas mais importantes no desenvolvimento de uma rede neural diz respeito à forma com que os dados disponíveis (amostra) são divididos. Normalmente, os dados coletados são separados em duas categorias: dados de treinamento, que serão utilizados para o treinamento da rede e dados de teste, que serão utilizados para verificar seu desempenho sob condições reais de utilização (HAYKIN, 2001). Além dessa divisão, é comum usar também uma subdivisão do conjunto de treinamento, conhecido como validação. Esse conjunto é frequentemente utilizado para verificar a eficiência da rede quanto a sua capacidade de generalização durante o treinamento, e pode ser empregado

¹ Universidade Federal de Uberlândia, claudioneidacunha@gmail.com

² Universidade Federal de Uberlândia, stefanydiovana.sd@gmail.com

³ Universidade Federal de Uberlândia, vfrofatto@gmail.com

⁴ Universidade Federal de Uberlândia, tomiomatsuoka@gmail.com



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



como critério de parada no processo de treinamento.

Depois de determinados estes conjuntos, eles são, geralmente, colocados em ordem aleatória para prevenção de tendências associadas à ordem de apresentação dos dados (HAYKIN, 2001). Este método de divisão é também conhecido como *Hold-out* (YADAV; SHUKLA, 2016). Também é comum encontrar o termo validação cruzada. Entretanto, existem limitações significativas para sua aplicação (ZIGGAH et. al., 2019): (i) os resultados produzidos são baseados em uma divisão não controlada; (ii) a divisão inadequada do conjunto de dados pode ter um efeito adverso no desempenho do modelo; (iii) é inadequado em situações de baixa densidade de dados (dados insuficientes e/ou esparsos).

Como alternativa, o método conhecido como *K-fold* tem sido recomendado (BURMAN, 1989; REITERMANOVÁ, 2010). Neste caso, os dados são separados em K subconjuntos (*K-fold*) de tamanho aproximadamente igual, de modo que cada subconjunto estará apenas uma única vez no conjunto de teste. Embora *K-fold* apresente vantagens, ainda não está claro como escolher os subconjuntos, e em alguns casos os subconjuntos não são de tamanhos iguais, o que não garante uma versão balanceada de validação cruzada. Além disso, tanto o *K-fold* quanto o *Hold-out* fornecem previsões pontuais. Aqui, surge uma questão importante: como confiar nas previsões fornecidas por uma única rede neural artificial (RNA) em particular? Devido à aleatoriedade relacionada à própria rede (arquitetura, inicialização e procedimento de aprendizado), geralmente não há a melhor escolha (PAN, 1998).

À luz desta questão, apresentamos uma abordagem que consiste em realizar sucessivas réplicas do método *Hold-out*. Com efeito, a previsão em cada ponto-teste pode ser não somente de forma pontual (por exemplo, baseado em estatísticas, como a média, a mediana, etc.), mas também por intervalo (intervalos de confiança).

2 MATERIAL E MÉTODOS

Uma rede neural Perceptron Multi-Camadas (MLP), treinada com o algoritmo back propagation, foi desenvolvida usando as coordenadas no sistema UTM (E, N) como atributos de entrada e a componente altimétrica (no caso, a altitude geométrica) como de saída (Figura 1). O método *Levenberg Marquardt* foi empregado para estimar os parâmetros da rede, ou seja, os pesos e vieses (MARQUARDT, 1963). A funcionalidade de linha de comando da caixa de ferramentas Rede Neural do Matlab (R2019b) foi usada para treinar e



II Simpósio Regional de Agrimensura e Cartografia

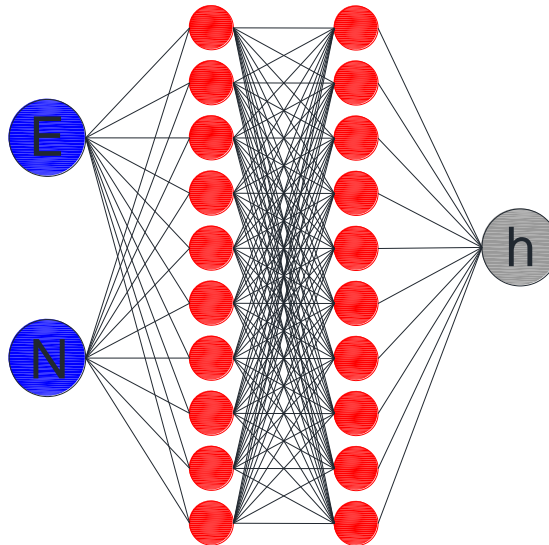
“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



validar a rede neural. Após vários testes, a arquitetura de rede neural ótima encontrada foi definida por duas camadas ocultas, com cada camada composta por dez neurônios. Como os dados de entrada são compostos apenas por coordenadas UTM (E, N), a rede neural foi denominada rede neural espacial (Figura 1).

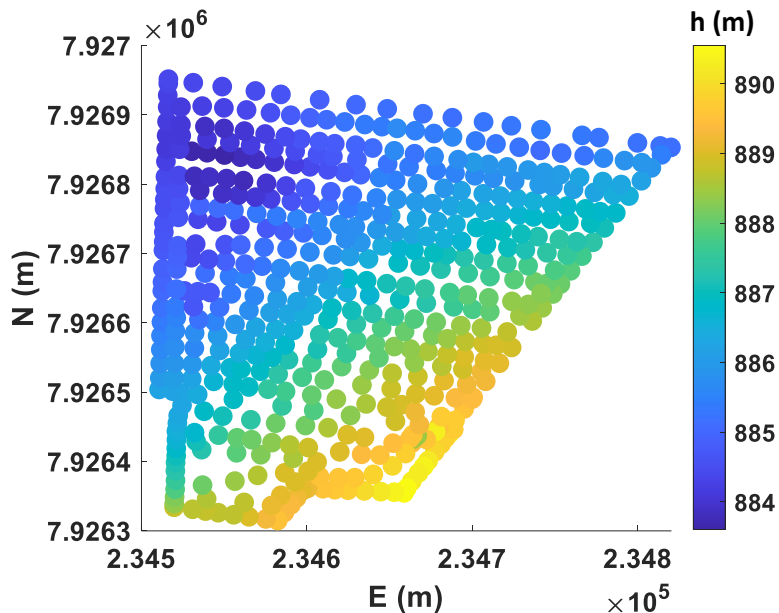
Figura 1 – Arquitetura da rede neural usada para estimativa da altitude



Os dados relativos ao problema referem-se à um levantamento realizado na região localizado em Monte Carmelo, Minas Gerais (Latitude: 18° 44' 5" Sul, Longitude: 47° 29' 47" Oeste, Fuso 23 S, Meridiano Central 45°). A aquisição das coordenadas (E, N, h) foi realizada por meio do método relativo cinemático em tempo real, conhecido como RTK (*Real Time Kinematic*). Um total de 511 pontos foram obtidos em campo (Figura 2).



Figura 2 – Dados de campo (E, N, h)



Após diversos testes, a melhor divisão aleatória e uniforme dos dados foi encontrada, como sendo 80% (409 pontos) para treinamento, 10% validação (51 pontos) e 10% para teste (51 pontos). Como já mencionei, esse método de divisão de dados é conhecido como *Hold-out*. Aqui, entretanto, o *Hold-out* foi estendido para sua versão denominada de *Hold-out trials*, que consiste em realizar sucessivas réplicas aleatórias do método *Hold-out*. Assim, cada predição será repetida certo número de vezes. Neste experimento, escolhemos 1000 realizações do *Hold-out*, o que proporciona aproximadamente 100 predições para cada ponto amostral.

3 RESULTADOS E DISCUSSÃO

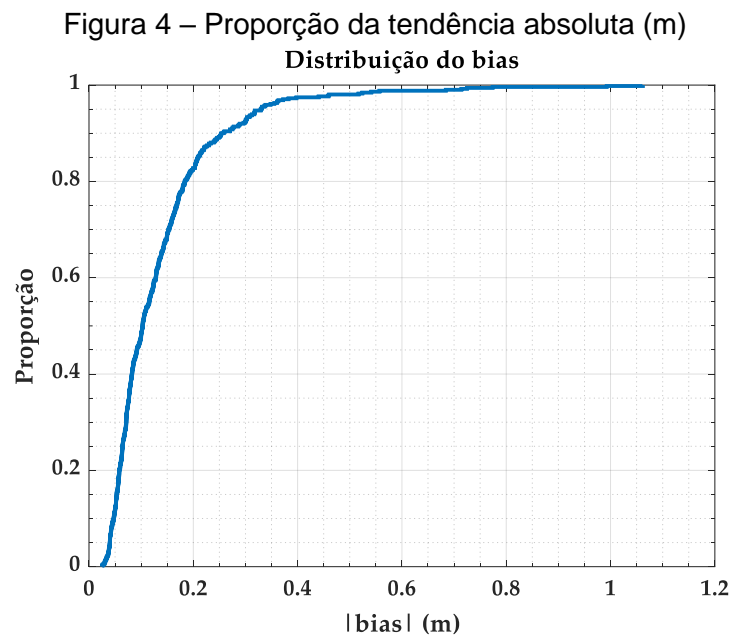
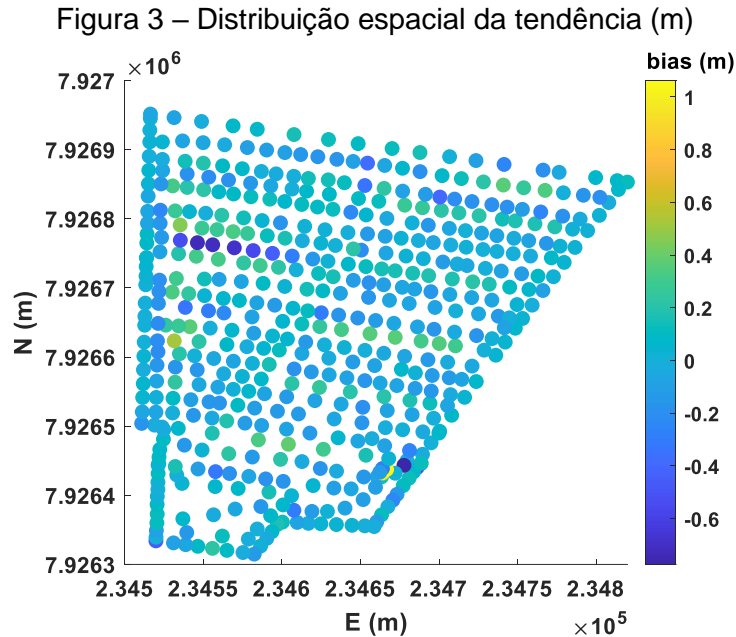
O método proposto permite descrever a distribuição espacial da tendência (bias), ou seja, a diferença entre a média da predição e o valor observado em campo para cada ponto individual (Figura 3). Nota-se que 85% dos pontos (434 pontos) ficaram com tendência absoluta (em módulo, |bias|) menor ou igual a 20 cm, e na média foi da ordem de 10 cm (Figura 4).



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



Nota-se que algumas variações abruptas da elevação dentro de uma região não foram capturadas pela RNA. Isso é verificado pelas tendências isoladas em uma dada região (Figura 5). Essas variações abruptas podem ser entendidas como outliers (ROFATTO, et. al., 2020), uma vez que o algoritmo de treinamento *Levenberg-Marquardt* (LM) não os considera nos dados. Portanto, um algoritmo de detecção de outliers ou algum estimador robusto, que leve em consideração essas anomalias, deverá ser empregado em conjunto com LM.

Figura 5 – Exemplo de uma variação abrupta da elevação no terreno e sua tendência



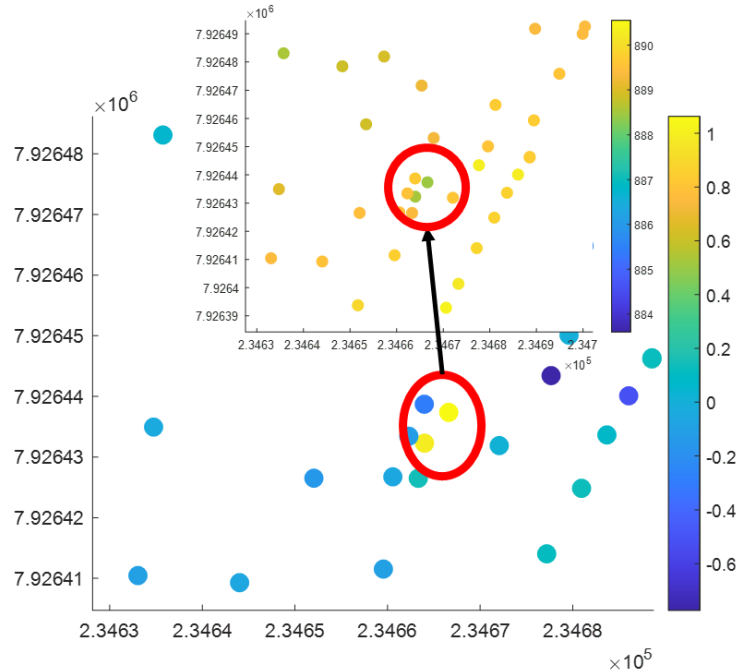
II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



expressa na predição (m)



Em termos de incerteza, nota-se que 90% dos pontos possuem desvio-padrão (1σ) menor ou igual a 11 cm, e na média da ordem de 6 cm (Figura 6 e 7). No geral, notamos também que as regiões de maior variabilidade em termos de elevação são aquelas em que as predições também apresentaram maiores incertezas (Figura 6 e 2). A incerteza caracteriza cada predição como sendo uma variável aleatória e, portanto, a estimativa não é apenas pontual (valor médio), mas valor médio \pm desvio-padrão (predição intervalar).



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



Figura 6 – Distribuição espacial do desvio-padrão (σ)

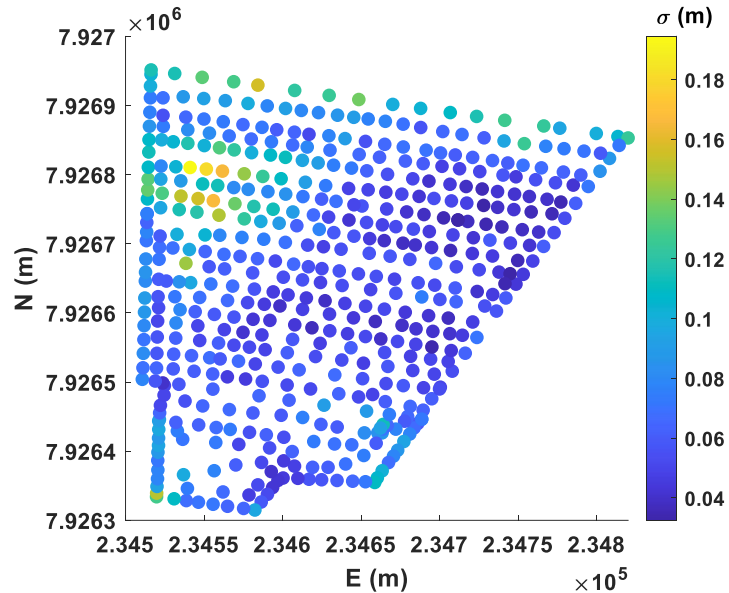
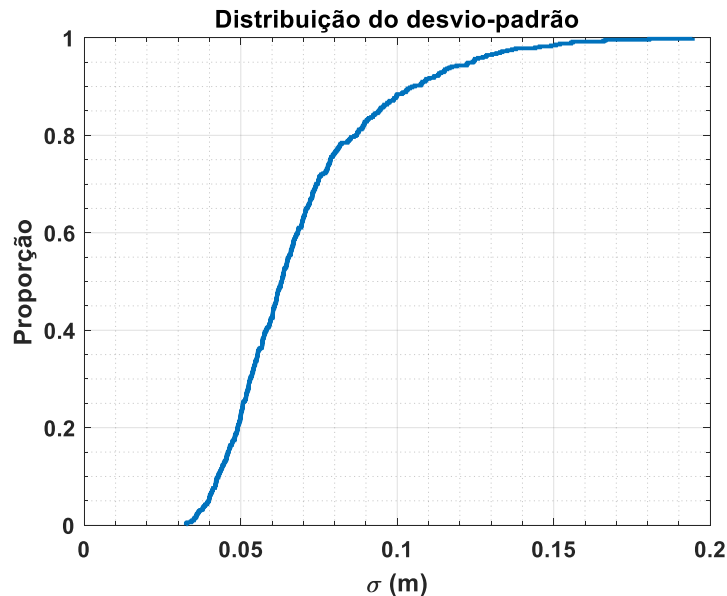


Figura 7 – Proporção da incerteza expressa pelo desvio-padrão (σ)



Em termos de acurácia, nota-se que 80% dos pontos possuem RMSE (*Root Mean Square Error* – Raiz do Erro Médio Quadrático) menor ou igual a 20 cm, e na média da ordem de 12 cm. Esses resultados também sugerem que diferentes redes neurais poderão ser desenvolvidas para cada região particular do estudo.



Figura 8 – Distribuição espacial do RMSE

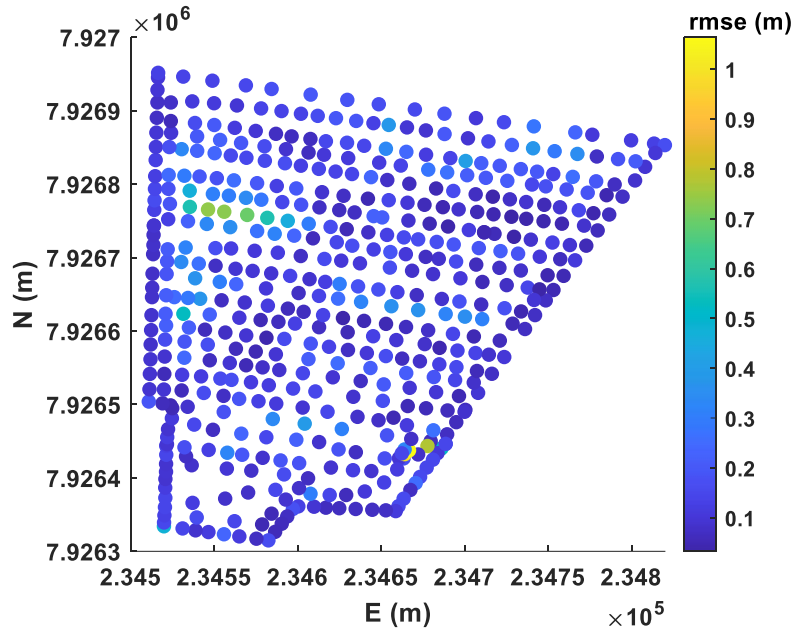
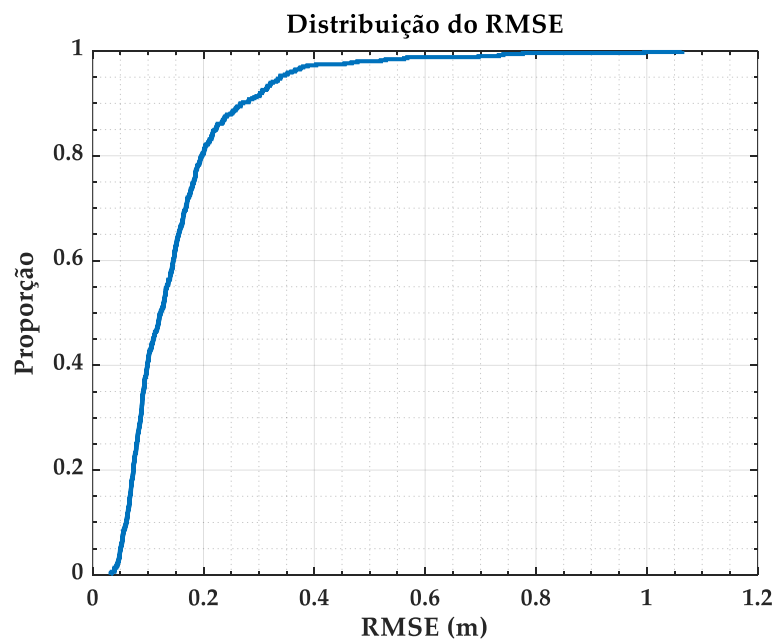


Figura 9 – Proporção da acurácia expressa pelo RMSE



4 CONCLUSÕES

Aqui, apresentamos um método de reamostragem em redes neurais, denominado de *Hould-out trials*. Os produtos gerados pelo método permitem descrever espacialmente a



II Simpósio Regional de Agrimensura e Cartografia

“Ampliando os horizontes e discutindo o futuro da geoinformação e do cadastro territorial do Brasil”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo
22 a 24 de novembro de 2021



tendência, a incerteza e acurácia das predições. A tendência e acurácia necessitam de valores observados em campo para as análises, o que por muitas vezes estão indisponíveis em aplicações práticas do modelo previamente construído. Entretanto, a incerteza expressa pelo desvio-padrão não precisa de valores de referência. Assim, a incerteza revela os pontos de maior ou menor confiança e, portanto, pode auxiliar nas tomadas de decisão.

Outros métodos de reamostragem vêm sendo integrados em rede neurais pelos autores, tais como *Jackknife* e *Bootstrap*. Apesar de restringirmos ao modelo de rede neural, o método proposto também pode ser estendido a outras ferramentas estatísticas modernas, como *Krigagem*, Colocação por Mínimos Quadrados, entre outros métodos preditivos.

REFERÊNCIAS

HAYKIN, S. **Redes neurais: princípios e prática**. Porto Alegre: Bookman, 2001.

YADAV, S.; SHUKLA, S. **Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification**, 2016 *IEEE 6th International Conference on Advanced Computing (IACC)*, 2016, pp. 78-83, doi: 10.1109/IACC.2016.25.

ZIGGAH, Y.Y.; YOUJIAN, H.; TIERRA, A.R.; LAARI, P. B. 2019. **Coordinate transformation between global and local datums based on artificial neural network with k-fold cross-validation: A case study, Ghana**. *Earth Sciences Research Journal*, v.23, n.1: p. 67-77.

BURMAN, P. 1989. **A comparative study of ordinary cross-validation, v-fold cross-validation and their repeated learning-testing methods**. *Biometrika*, v.76, n.3: 503-514.

REITERMANOVÁ, Z. 2010. **Data splitting**. *WDS'10 Proceedings of contributed papers, Part I: Mathematics and Computer Sciences*, p. 31-36. eds. Charles University, Faculty of Mathematics and Physics, Prague, CZ.

PAN, L. 1998. **Resampling in neural networks with application to financial time series**. *Ph.D. Thesis, The University of Guelph, ON, CA*.

MARQUARDT, D. W., 1963. **An algorithm for least-squares estimation of nonlinear parameters**. *Journal of the society for industrial and applied mathematics*, 11 (2), 431–441. Available from: <https://doi.org/10.1137/0111030>.

ROFATTO, V. F.; MATSUOKA, M. T.; KLEIN, I.; VERONEZ, M. R.; SILVEIRA JUNIOR, L. G. d. **A Monte Carlo-Based Outlier Diagnosis Method for Sensitivity Analysis**. *Remote Sens.* 2020, 12, 860. <https://doi.org/10.3390/rs12050860>