***Machine learning* aplicado na informatização de dados de colheita mecanizada de madeira**

**Fernando Reis Andrade1, Rafaele Almeida Munis1, Danilo Simões2**

1 Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agronômicas, Botucatu, São Paulo; 2 Universidade Estadual Paulista (Unesp), Campus Experimental de Itapeva, Itapeva, São Paulo. (fernandofr\_sp@hotmail.com).

**RESUMO:** A gestão florestal das operações de colheita mecanizada de madeira a partir dos dados reais de campo permite aplicação de técnicas de aprendizado de máquina que auxiliam as tomadas de decisões. Deste modo, foi avaliado se a aplicação do algoritmo de *support* *vector* *regression* propicia aos gestores florestais a informatização dos dados necessária para a predição da produtividade de um *feller-buncher* empregado na derrubada de árvores. A modelagem dos dados deu-se por meio do algoritmo de *machine learning* de *support vector regression*, parâmetro de *Kernel* linear, com dados particionados em treinamento e teste. Foram estimadas as importâncias relativas das variáveis preditoras pautadas no erro médio quadrático e absoluto. A performance do modelo resultou em um coeficiente de determinação de 79,31%. As variáveis relacionadas, volume de madeira colhida em metros cúbicos e horas de máquina disponíveis impactam na predição da produtividade do *feller buncher*.

**Palavras-chave:** *feller-buncher,* operações florestais, *support vector regression*

1. **INTRODUÇÃO**

Por demandar de aportes econômicos consideráveis às indústrias de base florestal, a gestão das operações de colheita mecanizada de madeira deve estar sempre alinhada com a produtividade aumentada e redução de custos. Logo, gestores florestais precisam empreender ações que alcancem com satisfatoriedade tais objetivos.

Assim, suportados pela enorme quantidade de dados gerados em campo, técnicas de mineração de dados aliadas ao *machine learning*, permitem a estes gestores florestais a extração do conhecimento e tomadas de decisões inteligentes, resultando em um gerenciamento ideal de recursos e serviços (SUNHARE *et al*., 2020). Contudo, existem vários parâmetros dos algoritmos de treinamento que precisam ser selecionados para os diferentes métodos de aplicação (ESCRIG *et al*., 2020).

Neste sentido, algoritmos de *machine learning*, como o de *support vector* *regression*, são formulados para problemas de otimização (TAVARES JÚNIOR *et al*., 2020) e comumente utilizados para aumentar a precisão de previsões, por meio da obtenção de vetores de suporte (KOR; ALTUN, 2020), o que justifica sua implementação no gerenciamento florestal de operações de colheita mecanizada de madeira com vista à predição de variáveis.

Portanto, foi avaliado se a aplicação do algoritmo de *support vector regression* propicia aos gestores florestais a informatização dos dados necessária para a predição da produtividade de um *feller-buncher* empregado na derrubada de árvores.

1. **MATERIAL E MÉTODOS**

*Dataset*

Os dados utilizados para aplicação do algoritmo de *support vector regression* foram oriundos de uma base de dados de operações de colheita mecanizada de madeira de *Eucalyptus* em florestas plantadas com espaçamento de 3,30 m x 1,80 m, em relevo suave ondulado e latossolo.

O *feller-buncher* da marca *Tigercat* e modelo L870C possuía 5.372,00 horas de uso acumulada, rodantes em esteiras, com um cabeçote da marca *Tigercat*, modelo ST5702-26 e 26 polegadas de diâmetro.

*Modelagem dos dados*

A predição da produtividade do *feller-buncher* (PR) deu-se a partir das variáveis: horas de máquina disponíveis (HM) e volume de madeira colhida (VM) em metros cúbicos, somando-se a isso cinco variáveis que ocasionaram interrupções operacionais (IO), foram elas: Abastecimento e Lubrificação (IO1); Aguardando Mecânico Interno (IO2); Manutenção do Implemento (IO3); Manutenção da Máquina Base (IO4); Análise de Falhas e Ajustes (IO5).

A modelagem dos dados por *support vector regression* ocorreu com o auxílio do *software* R (R CORE TEAM, 2019) e aplicação da função *svm*, inclusa no pacote *e1071*, parâmetro de *Kernel* linear, tipo *eps-regression*. Destarte, os dados foram particionados em 70% para treinamento e 30% para teste, com validação cruzada com a técnica *k-fold*.

Foram estimadas as importâncias relativas das variáveis preditoras por meio da permutação de uma variável e a manutenção das demais fixas, para o cálculo do erro médio quadrático e erro médio absoluto. Ademais, o coeficiente de determinação ajustado (R² ajustado) foi utilizado como estatística de qualidade do ajuste do modelo.

1. **RESULTADOS E DISCUSSÃO**

Ao verificar quais das variáveis relacionadas para construção do modelo preditor da produtividade do *feller-buncher*, percebeu-se que apenas duas foram capazes de minimizar o erro médio quadrático em relação ao modelo expandido composto por todas as variáveis, isto é, quando mantidas apenas as horas de máquina disponíveis e o volume de madeira colhida em metros cúbicos o erro médio quadrático resultante da estimativa da produtividade foi o menor observado (Tabela 1).

Tabela 1. Estatísticas de precisão da produtividade do *feller-buncher* operando em florestas plantadas com *Eucalyptus*.

|  |  |  |
| --- | --- | --- |
| Variáveis Preditoras | Erro Médio Quadrático | Erro Médio Absoluto |
| (m³ h-1) | (m³ h-1) |
| Volume de Madeira Colhida (m³) – VM | 38,36 | 28,71 |
| Horas de Máquina Disponíveis (h) – HM | 24,05 | 19,89 |
| Manutenção da Máquina Base (h) – IO4 | 15,73 | 11,19 |
| Manutenção do Implemento (h) – IO3 | 15,26 | 9,48 |
| Abastecimento e Lubrificação (h) – IO1 | 15,23 | 8,05 |
| Análise de Falhas e Ajustes (h) – IO5 | 14,96 | 8,44 |
| Aguardando Mecânico Interno (h) – IO2 | 14,92 | 9,14 |
| Todas | 16,75 | 10,33 |
| VM + HM | 15,92 | 11,45 |

Assim, Shansara e Scüürmann (2020) ressaltam que o *machine learning*, por meio da adoção do algoritmo de *support vector regression* fornece resultados validados e precisos para expandir o conhecimento das características que compõe e impactam os processos.

Exemplo disto, a assertividade na determinação da produtividade do *feller-buncher* na operação de derrubadas das árvores na colheita mecanizada de madeira, uma vez que quando avaliado o poder preditivo do algoritmo, com o modelo final composto pelas variáveis VM e HM, foi obtido R² ajustado de 74,26, entretanto, R² ajustado de 79,31 para o conjunto de dados teste.

Logo, o poder preditivo do modelo ajustado possibilitou a inferência da produtividade do *feller-buncher* que, por conseguinte, torna possível identificar regras de previsão importantes para gestores florestais de operações de colheita mecanizada de madeira que devem enfrentar situações específicas e concretas (ROSSIT *et al*., 2019).

1. **CONCLUSÕES**

O modelo composto pelas variáveis volume de madeira colhida em metros cúbicos e horas de máquina disponíveis propicia a predição da produtividade do *feller-buncher*, dado o coeficiente de determinação ajustado obtido pela utilização do algoritmo de *support vector regression*.

As variáveis que despendem de tempo por interrupção operacional, não impactam na predição da produtividade do *feller-buncher*.

1. **AGRADECIMENTOS**

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

1. **REFERÊNCIAS**

ESCRIGA, J.; WOOLLEY, E.; SIMEONE, A.; WATSONA, N. J. Monitoring the cleaning of food fouling in pipes using ultrasonic measurements and machine learning. **Food Control**, v. 116, n. 107309, 2020.

KOR, K.; ALTUN, G. Is Support Vector Regression method suitable for predicting rate of penetration? **Journal of Petroleum Science and Engineering**, v. 194, n. 107542, 2020.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2019. Disponível em: <https://www.R-project.org/> Acesso em: 15/09/2020.

ROSSIT, D. A.; OLIVEIRA, A.; CÉSPEDES, V. V.; BROZ, D. A Big Data approach to forestry harvesting productivity. **Computers and Electronics in Agriculture**, v. 161, p. 29-52, 2019.

SHAMSARA, J.; SCHÜÜRMANN, G. A machine learning approach to discriminate MR1 binders: The importance of the phenol and carbonyl fragments. **Journal of Molecular Structure**, v. 1217, n. 128459, 2020.

SUNHARE, P.; CHOWDHARY, R. R.; CHATTOPADHYAY, M. K. Internet of Things and Data Mining: An Applications Oriented Survey. Journal of King Saud University - **Computer and Information Sciences**. In press, 2020.

TAVARES JÚNIOR, I. S.; TORRES, C. M. M. E.; LEITE, H. G.; CASTRO, N. L. M.; SOARES, C. P. B.; CASTRO, R. V. O.; FARIAS, A. A. Machine learning: Modeling increment in diameter of individual trees on Atlantic Forest fragments. **Ecological Indicators**, v. 117, n. 106685, 2020.