57º SBPO
SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
Gramado • RS
DE 5 A 9 DE OUTUBRO DE 2025

COMPARAÇÃO DE MODELOS PREDITIVOS PARA UM PROBLEMA APLICADO DE CARREGAMENTO DE PALETES DO DISTRIBUIDOR

Matheus Aguilar de Oliveira e André Gustavo dos Santos
Universidade Federal de Viçosa (UFV)
Av. P H Rolfs, s/n - Campus Universitário, Viçosa - MG
matheus.a.aguilar@ufv.br; andresantos@ufv.br

Thiago Alves de Queiroz
Universidade Federal de Catalão (UFCAT)
Av. Dr. Lamartine Pinto de Avelar, 1120, Vila Chaud, Catalão - GO
taq@ufcat.edu.br

Marco Taccini e Manuel Iori
University of Modena and Reggio Emilia (UNIMORE)
Via Giovanni Amendola, 2, 42122 Reggio Emilia, Itália
marco.taccini@unimore.it; manuel.iori@unimore.it

RESUMO

Este artigo tem como motivação as operações de uma indústria italiana de produção de cerâmica, que lida diariamente com o problema de carregamento de paletes do distribuidor. A indústria está interessada em uma abordagem que seja rápida e precisa ao lidar com o problema. É proposta uma abordagem híbrida que combina heurísticas construtivas com diferentes modelos de aprendizado de máquina: XGBoost, LightGBM e Random Forest. Usando dados da indústria italiana, os modelos são treinados, validados e comparados com o *software* PackVol. Os modelos propostos mostraram desempenho superior em termos de precisão e tempo de execução, indicando que a abordagem híbrida é eficaz mesmo com diferentes algoritmos de aprendizado de máquina.

PALAVRAS CHAVE. Problema de carregamento de paletes, Aprendizado de máquina, Heurística, Indústria de cerâmica.

IA – PO e IA, POI – PO na Indústria, L&T – Logística e Transportes.

ABSTRACT

This paper is motivated by the operations of an Italian ceramic production company that deals daily with the distributor's pallet loading problem. The industry is interested in a fast and accurate approach to this problem. We propose a hybrid approach, combining constructive heuristics with different machine learning models: XGBoost, LightGBM, and Random Forest. Using data from the Italian industry, the models are trained, validated, and compared with the software PackVol. The proposed models showed superior accuracy and computing time performance, indicating that the hybrid approach is effective even with different machine learning models.

KEYWORDS. Pallet Loading Problem, Machine learning, Heuristic, Ceramic industry.

IA – OR and AI, POI - OR in Industry, L&T - Logistics and Transportation.

https://proceedings.science/p/212248?lang=pt-br

1. Introdução e definição do problema

O mercado global de produção cerâmica cresceu consideravelmente na última década, com a produção mundial atingindo quase 17 bilhões de metros quadrados segundo a ACIMAC Research Department [2023], sendo o italiano um dos maiores. Indústrias da área enfrentam desafios logísticos relacionados ao transporte e carregamento de paletes, nos quais caixas de diferentes tamanhos e pesos precisam ser organizadas para envio aos clientes. Esse processo exige respeitar restrições como limite de peso, estabilidade e orientação das caixas, buscando sempre minimizar o número de paletes utilizados.

O problema estudado neste trabalho surge de uma indústria cerâmica localizada na região da Emília-Romanha, que lida diariamente com o chamado *Distributor's Pallet Loading Problem* (DPLP) [Silva et al., 2016]. Na literatura, o DPLP é considerado uma variante dos problemas de corte e empacotamento, amplamente estudados na pesquisa operacional. Diversos trabalhos exploraram abordagens como heurísticas [Gunawardena et al., 2021], algoritmos exatos [Iori et al., 2021], metaheurísticas [Calzavara et al., 2021] e técnicas de aprendizado de máquina [Magnani et al., 2025], muitas vezes aplicados em contextos industriais reais com restrições práticas complexas [Gzara et al., 2020].

No contexto da empresa italiana estudada, os pedidos chegam por meio de um sistema de e-commerce, e a empresa precisa informar ao cliente, já no momento da compra, a quantidade estimada de paletes necessários. Essa informação é essencial para que os clientes organizem o transporte de forma adequada. As cerâmicas são armazenadas em paletes homogêneos, posicionados no armazém por um sistema automatizado, e depois transferidas para paletes de envio conforme cada pedido. O carregamento é feito por operadores em baias com espaço limitado, que recebem automaticamente até quatro paletes de produção por vez. Como não controlam quais paletes chegam, os operadores precisam organizar rapidamente os paletes de envio para liberar espaço, respeitando restrições de peso, estabilidade, orientações permitidas e fragilidade das caixas.

Devido à dinamicidade e às incertezas do sistema automatizado, torna-se inviável adotar soluções exatas ou heurísticas tradicionais para atender ao problema. Assim, a indústria necessita de uma abordagem que forneça estimativas rápidas e precisas da quantidade de paletes necessários por pedido. Para isso, este trabalho propõe uma solução híbrida que integra heurísticas construtivas e técnicas de aprendizado de máquina treinadas com dados históricos da indústria para gerar previsões precisas em tempo real. Alguns resultados dessa abordagem, limitados ao modelo XGBoost, foram publicados em Taccini et al. [2025].

2. Abordagem Híbrida Proposta

A abordagem híbrida usa dados históricos da indústria e de limitantes válidos gerados por heurísticas construtivas, que são organizados para a entrada dos modelos de aprendizado. Diferente do trabalho anterior em Taccini et al. [2025], que se concentrou unicamente no modelo XGBoost, é feita a sua extensão para a inclusão dos modelos *Light Gradient-Boosting Machine* (LGBM) e *random forest* (RF), com objetivo de generalizar a abordagem.

Para alimentar os modelos, foi utilizado o mesmo conjunto de colunas definido em Taccini et al. [2025]: as quantidades de cada tipo de cerâmica, peso e volume totais e médios do pedido, além de 3 heurísticas que estimavam um limite inferior, superior e valor aproximado de paletes para um pedido. Os hiperparâmetros dos modelos foram calibrados através de uma busca em grade (*grid search*) considerando valores obtidos de análises empíricas, valores padrão das bibliotecas que implementam tais modelos e com base na literatura [Zouinina, 2024; Ahmed, 2022; Sharma, 2024; Saxena, 2024]. A Tabela 1 contém os melhores valores obtidos após a calibragem.

3. Experimentos computacionais

A abordagem híbrida foi implementada em Python 3.10, utilizando a biblioteca *scikit-learn*. Para os experimentos, empregou-se um computador com processador Intel Core i5, 8 GB de

Tabela 1: Conjunto de valores dos hiperparâmetros obtidos durante a calibragem pela busca em grade.

XGBoost		LGBM		RF	
learning_rate	0.1	learning_rate	0.1	max_depth	50
gamma	0.1	max_depth	∞	min_samples_split	20
max_depth	10	num_leaves	1023	min_samples_leaf	1
min_child_weight	3	min_data_in_leaf	50	max_features	6
subsample	0.5	feature_fraction	1.0	bootstrap	False
colsample_bytree	0.5				

Tabela 2: Resultados dos modelos em termos de MSE e MAE.

Modelo	MSE	MAE
XGBoost	0,728	0,329
LGBM	0,726	0,326
RF	0,779	0,332

(a) Conjunto de pedidos de treinamento

Modelo	MSE	MAE
XGBoost	0,500	0,433
LGBM	0,467	0,400
RF	0,400	0,400
PackVol	2,133	0,933

(b) Conjunto de novos pedidos (validação)

memória e sistema operacional *Linux Mint 21 Cinnamon*. Os modelos de aprendizado de máquina foram treinados a partir de aproximadamente 167 mil pedidos históricos da indústria de 2022 e 2023, considerando apenas os pedidos heterogêneos, ou seja, aqueles com mais de um tipo de cerâmica. Os dados foram divididos em 80% para treinamento e 20% para validação.

Durante a fase de treinamento, os três modelos de aprendizado — XGBoost, LGBM e Random Forest — apresentaram desempenhos semelhantes em termos de erro médio quadrático (MSE) e erro absoluto médio (MAE). O modelo LGBM obteve ligeira vantagem sobre os demais (conforme apresentado na tabela 2a).

Na etapa de validação, os modelos foram avaliados com 30 pedidos reais recebidos em 2024, todos distintos dos utilizados no treinamento e contendo de 2 a 6 tipos de cerâmica. Os resultados mostraram que todos os modelos superaram o software PackVol [PackVol, 2024], utilizado pela indústria, apresentando MSE até cinco vezes menor e MAE duas vezes menor. Entre os três, LGBM e Random Forest se destacaram com resultados muito próximos e ligeiramente melhores que o XGBoost (conforme apresentado na tabela 2b). Além disso, enquanto o PackVol levou cerca de 1 segundo para resolver cada instância, os modelos de aprendizado processaram todas as 30 em menos de 0,5 segundos.

4. Conclusão e Trabalhos Futuros

Neste artigo é proposta uma abordagem híbrida para a previsão do número de paletes no contexto do DPLP em uma indústria de cerâmicas. Mesmo considerando diferentes modelos de aprendizado de máquina, esta abordagem se mostrou mais precisa e rápida do que a abordagem (PackVol) atualmente usada pela indústria, avançando as conclusões apontadas em Taccini et al. [2025], que apresentava apenas o modelo XGBoost como parte da abordagem híbrida.

Trabalhos futuros poderiam avaliar outros modelos de aprendizado de máquina, especialmente fora da família dos métodos baseados em árvore de decisão, e outras heurísticas para os limitantes, bem como estender a previsão para pedidos que envolvem diferentes tipos de paletes.

Referências

ACIMAC Research Department (2023). World production and consumption of ceramic tiles. https://www.mecs.org/it/shop/cat/16-mercato-ceramico/prod/111-produzione-e-consumo-mondiale-di-piastrelle-ceramiche. Acessado: 2025-05-17.

Ahmed, A. M. A. (2022). Lightgbm cheat sheet. https://awadrahman.medium.com/lightgbm-cheat-sheet-ddde43f9eaf1. Acessado: 2025-05-04.

Calzavara, G., Iori, M., Locatelli, M., Moreira, M. C. O., e Silveira, T. (2021). Mathematical models and heuristic algorithms for pallet building problems with practical constraints. *Annals of Operations Research*. ISSN 1572-9338.

Gunawardena, K., Wijayanayake, A., e Kavirathna, C. (2021). Solve manufacturer's pallet loading problem (mplp) with practical warehouse constraints. In *2021 6th International Conference on Information Technology Research (ICITR)*, p. 1–6.

Gzara, F., Elhedhli, S., e Yildiz, B. C. (2020). The pallet loading problem: Three-dimensional bin packing with practical constraints. *European Journal of Operational Research*, 287(3):1062–1074.

Iori, M., de Lima, V. L., Martello, S., Miyazawa, F. K., e Monaci, M. (2021). Exact solution techniques for two-dimensional cutting and packing. *European Journal of Operational Research*, 289(2):399–415. ISSN 0377-2217.

Magnani, M., Franchini, G., Dell'Amico, M., e Zanni, L. (2025). A machine learning approach to speed up the solution of the distributor's pallet loading problem. In Sergeyev, Y. D., Kvasov, D. E., e Astorino, A., editors, *Numerical Computations: Theory and Algorithms*, p. 108–120, Cham. Springer Nature Switzerland. ISBN 978-3-031-81241-5.

PackVol (2024). Optimization software for load planning. https://www.packvol.com/index.html. Acessado: 2024-12-10.

Saxena, S. (2024). A beginner's guide to random forest hyperparameter tuning. https://www.analyticsvidhya.com/blog/2020/03/beginners-guide-random-forest-hyperparameter-tuning. Acessado: 2025-05-04.

Sharma, K. (2024). Mastering random forest hyperparameter tuning for enhanced machine learning models. https://medium.com/@kalpit.sharma/mastering-random-forest-hyperparameter-tuning-for-enhanced-machine-learning-models-2d1a8c6c426f. Acessado: 2025-05-04.

Silva, E., Oliveira, J. F., e Wäscher, G. (2016). The pallet loading problem: a review of solution methods and computational experiments. *International Transactions in Operational Research*, 23(1-2):147–172.

Taccini, M., Aguilar de Oliveira, M., dos Santos, A. G., Alves de Queiroz, T., e Iori, M. (2025). A hybrid approach for pallet loading in ceramic tile industry. In *Proceedings of the International Conference on Optimization and Learning (OLA2025)*, p. 1–12.

Zouinina, S. (2024). A deep dive into lightgbm: How to choose and tune parameters. https://medium.com/@sarahzouinina/a-deep-dive-into-lightgbm-how-to-choose-and-tune-parameters-7c584945842e. Acessado: 2025-05-17.

https://proceedings.science/p/212248?lang=pt-br