

A MINERAÇÃO DE DADOS PARA SELEÇÃO DE HEURÍSTICAS NO PROBLEMA DE EMPACOTAMENTO BIDIMENSIONAL RETANGULAR

Alvaro Neuenfeldt Júnior

Universidade Federal de Santa Maria
Av. Roraima 1000, prédio 7, sala 302, 97105900, Santa Maria/Brasil
alvaroj.eng@gmail.com

RESUMO

O problema de empacotamento bidimensional retangular em faixas consiste em posicionar um conjunto de pequenos retângulos em uma faixa de largura fixa e comprimento virtualmente infinito, minimizando a comprimento necessário para posicionar todos os retângulos na faixa. O objetivo geral deste projeto é ajustar modelos de classificação de algoritmos para selecionar, com acurácia, a melhor opção de heurística de melhoria, de acordo com as características de cada instância do problema. A metodologia de pesquisa está baseada no uso de técnicas supervisionadas de mineração de dados para o ajuste de modelos de classificação de algoritmos. O enfoque está na seleção das heurísticas de melhoria que possuem um maior potencial de encontrar soluções de qualidade para o problema, somente utilizando como informação as características das instâncias. Após a condução da pesquisa foi possível observar que a técnicas supervisionadas de mineração de dados *support vector machine with polynomial kernel* foi a mais eficiente no que tange a busca pela melhor opção de heurística de melhoria para o contexto do problema. Ainda, notou-se que as características das instâncias são capazes de remeter informações importantes para a resolução do problema.

Palavra-chave: Empacotamento bidimensional retangular; Pesquisa Operacional; Mineração de dados; Heurísticas.

ABSTRACT

The rectangular two-dimensional strip packing problem is to position a set of small rectangles within a fixed width and virtually infinite length range, while minimizing the length required to position all rectangles. The overall objective of this project is to fit algorithm classification models capable of accurately selecting the best improvement heuristic option according to the characteristics of each instance of the rectangular two-dimensional strip packing problem. The research methodology is based on the use of supervised data mining techniques for the adjustment of algorithm classification models. The focus is on the selection of heuristics that have the greatest potential to find quality solutions to the problem, using only as information the characteristics of the instances used. After conducting the research, it was possible to observe that the supervised data mining techniques support vector machine with polynomial kernel is the most efficient in the search for the best improvement heuristic option for the problem context. Still, it was noted that the characteristics of the instances are able to send important information to solve the problem.

Keywords: Rectangular Two-dimensional strip packing; Operational research; Data mining; Heuristics.

Como Citar:

NEUENFELDT JÚNIOR, Alvaro. A mineração de dados para seleção de heurísticas no problema de empacotamento bidimensional retangular. *In: SIMPÓSIO DE PESQUISA OPERACIONAL E LOGÍSTICA DA MARINHA*, 19., 2019, Rio de Janeiro, RJ. **Anais** [...]. Rio de Janeiro: Centro de Análises de Sistemas Navais, 2019.

1. INTRODUÇÃO

A crescente competitividade entre as indústrias exige que cada vez mais investimentos devam ser direcionados para a otimização dos processos de produto. As indústrias que envolvem diretamente operações de corte ou empacotamento de materiais não são exceção. Independente do contexto, tanto os problemas de corte como os de empacotamento possuem uma mesma estrutura lógica, apesar de serem intuitivamente diferenciáveis. Ambos envolvem a pesquisa pelo melhor padrão na disposição dos itens pequenos (ou peças pequenas) em objetos grandes (superfícies grandes), de modo a satisfazer um conjunto de restrições específicas do problema abordado. Porém, o processo de pesquisa é fortemente influenciado pelo aumento do número de itens a serem processados, ou por grandes variações nos formatos dos itens ou do objeto. O que afeta diretamente a eficiência da resolução de problemas deste, pois exigem maiores capacidades computacionais para gerar soluções.

O problema de empacotamento bidimensional retangular em faixas, do inglês *two-dimensional strip-packing problem* (2D-SPP), consiste em posicionar um conjunto de pequenos retângulos em uma faixa de largura fixa e comprimento virtualmente infinito, minimizando a comprimento necessário para posicionar todos os retângulos. Os retângulos devem ser posicionados sem qualquer sobreposição entre eles e completamente dentro da faixa. De acordo com a tipologia proposta em [1], esta descrição se encaixa na definição de problemas de corte e empacotamento como um problema de dimensão aberta retangular bidimensional. No 2D-SPP abordado nesta pesquisa, os retângulos devem ser posicionados ortogonalmente dentro da faixa e podem girar 90 graus. Aplicações típicas 2D-SPP são encontradas no contexto industrial, incluindo o corte de metais, têxteis ou rolos de papel. A Figura 1 mostra uma representação simples do 2D-SPP.

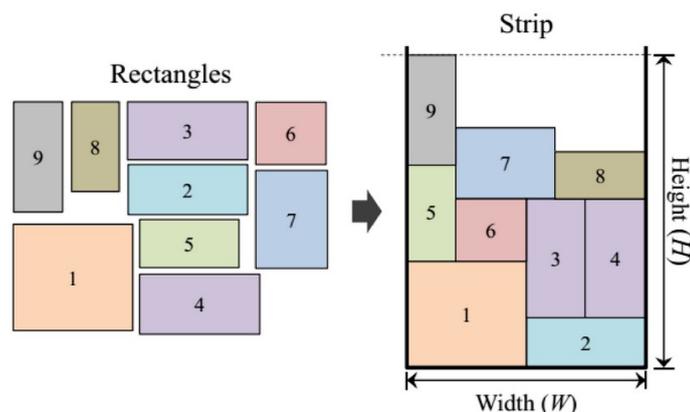


Figura 1 – Representação do 2D-SPP.

Assim, um dos principais desafios é encontrar o melhor custo-benefício entre o tempo de processo computacional a ser dispensado e o nível de qualidade esperado das soluções, para que o foco de otimizar uma função objetivo seja alcançado. O problema pode estar proposto para a maximização de lucros ou da quantidade de peças utilizadas, bem como

da minimização dos custos ou da quantidade de material desperdiçado.

Um conjunto de problemas industriais relacionados com o SPP podem ser encontrados, como por exemplo no empacotamento de dados em diversas áreas tecnológicas, automação das operações de corte em metais, vidros, papéis, tecidos, madeira, mármore ou plásticos, escalonamento multidimensional de recursos limitados, dentre outros.

Nos cortes bidimensionais de tecidos, o rolo (objeto) é dividido em pequenas partes. Já no corte de vidros ocorre a substituição dos rolos de materiais por placas, capazes de processar principalmente itens com formatos muito variados. Aplicações do SPP surgem também em áreas tecnológicas. Nos circuitos integrados do tipo *Very-Large-Scale Integration* (VLSI), uma das etapas mais importantes é o empacotamento de um conjunto de módulos (dados) dentro do domínio retangular (*chip*). O objetivo é minimizar a área utilizada, o que é crucial para obter reduções nos custos da produção de *chips* por *wafer*.

Portanto, é notória a necessidade da utilização de técnicas para coordenar mais eficientemente o modo como o posicionamento dos itens no objeto deve ocorrer, para a obtenção de soluções de melhor qualidade. Fato que torna relevante a condução de investigações sobre o 2D-SPP. Em específico, o desenvolvimento de algoritmos mais eficientes para reduzir os desperdícios no momento de encontrar o melhor arranjo dos retângulos dentro da faixa afeta diretamente na redução dos custos operacionais e de matéria-prima na indústria, além de impactar significativamente na redução de impactos ambientais, decorrentes do menor desperdício de matéria-prima quando o arranjo para o corte das faixas possuir uma opção melhor ajustada.

O 2D-SPP é um problema NP-difícil e pode ser resolvido por métodos exatos ou abordagens heurísticas [2,3]. Os métodos exatos são baseados em modelos de programação matemática e são capazes de obter soluções ótimas ou, pelo menos, determinar a qualidade da solução. No entanto, esses métodos não conseguem lidar com a complexidade e o tamanho dos aplicativos do mundo real. Em vez disso, abordagens heurísticas têm sido usadas nessas situações, dados os baixos tempos computacionais necessários quando comparados aos métodos exatos. No entanto, as abordagens heurísticas não garantem a otimização e não fornecem informações sobre a qualidade da solução [4,5].

As heurísticas são formadas pela combinação entre heurísticas construtivas (para posicionar por etapas os retângulos na faixa) e heurísticas de melhoria (para melhorar as soluções completas até que um critério de parada seja atingido). O comportamento intrínseco de cada combinação de heurísticas construtivas e de melhoramento é utilizado para definir quando as soluções possuem maior ou menor qualidade em relação as características das instâncias do problema.

O objetivo geral deste projeto é ajustar modelos de classificação de algoritmos capazes de selecionar, com acurácia, a melhor opção de heurística de melhoria de acordo com as características de cada instância do problema de empacotamento bidimensional retangular em faixas.

Soluções de qualidade são aquelas que o arranjo dos retângulos posicionados admite menores quantidades de espaços vazios dentro da faixa. Por se tratar de um problema em que a quantidade de retângulos a ser empacotada/cortada é fixa, desperdícios elevados acarretam em maiores quantidades de sobras ou retalhos na matéria-prima (faixas), demandando o uso de uma quantidade maior de material para produzir a quantidade de retângulos pré-estabelecida.

2. METODOLOGIA

A metodologia de pesquisa está baseada no uso de técnicas supervisionadas de mineração de dados para o ajuste de modelos de classificação de algoritmos. O enfoque está

na seleção das heurísticas que possuem um maior potencial de encontrar soluções de qualidade para o problema, somente utilizando como informação as características das instâncias utilizadas.

As instâncias são uma forma de representar problemas encontrados na realidade industrial do corte de retângulas a partir de faixas bidimensionais, contendo informações a respeito do número de retângulos, dimensões dos retângulos e largura da faixa. Por se tratar de um problema com dimensão aberta, o comprimento da faixa não é informado a priori. Assim, valores aproximados, como os *bounds* [2] ou estimativas [6,7] podem ser utilizadas para a verificação aproximada da quantidade de desperdício de espaços obtidas no arranjo dos retângulos na faixa. Esta medida é fundamental justamente para permitir, no ajuste do modelo de classificação, a mensuração de qual heurística é capaz de reduzir os desperdícios de espaços com maior eficiência, com base nas características das instâncias.

Um dos pontos fundamentais para o desenvolvimento do projeto é o conhecimento mais profundo sobre o problema. E isto pode ser obtido através de uma revisão detalhada da literatura sobre os principais métodos de seleção de algoritmos existentes, aplicações comuns do 2D-SPP e principais heurísticas utilizadas para a resolução do problema retangular que permite rotação de 90 graus [8].

Quanto aos parâmetros de entrada e saída, um set de instâncias (definido como *Dataset*) deverá ser utilizado como base de dados para desenvolver a análise e obter modelos de classificação. Em complemento, nesta etapa ocorre a definição das heurísticas construtivas e de melhoria, bem como da forma com que o arranjo entre as heurísticas estará disposto. Por ser uma metodologia que envolve o uso da mineração de dados supervisionada, com 19 variáveis explicativas relacionadas a geometria dos retângulos e da faixa e a características do problema [9]. Além disso, uma variável dependente foi definida, o gap entre a menor altura calculada nas heurísticas de melhoria e um valor de referência característico de cada instância, no caso o *lower bound* da área [2], para vincular a performance de cada heurística.

A estrutura baseada em mineração de dados será desenvolvida em torno dos conjuntos de dados de processamento usados para ajustar e validar modelos de classificação, que podem ser aplicados para selecionar a opção heurística de aposta para o problema. Subetapas de transferência de dados de entrada (variáveis e conjuntos de dados) para ajustar ou usar no modelo de classificação devem ser construídas, em conjunto com etapas de descoberta do conhecimento sobre o problema para a mineração dos dados de entrada.

O estudo experimental é uma etapa de fundamental importância pois possibilita o ajuste de alguns dos parâmetros de entrada mais relevantes para o desenvolvimento deste tipo de projeto, tais como: definição das métricas de performance das heurísticas; número de iterações; definição da melhor opção de método *multi-label* de transformação; seleção da técnica de mineração de dados supervisionada a ser usada para ajustar o modelo de classificação. As conclusões obtidas nesta etapa são fundamentais para que a análise de classificação seja viabilizada, com o uso da base de dados *Dataset*.

Por fim, se a validação do modelo de classificação estiver coerente com a expectativa proposta, ocorre a compilação dos resultados e, posteriormente, o encerramento do projeto. Senão, duas opções são possíveis. A primeira é verificar no estudo experimental se os parâmetros de entrada estão definidos de maneira coerente. A segunda opção é verificar se algum parâmetro da mineração de dados foi ajustado de forma precipitada, e isto estar influenciando na falta de coerência dos modelos de classificação ajustados.

3. RESULTADOS PRELIMINARES

A presente seção descreve a aplicação da análise de classificação para o 2D-SPP. Primeiramente, um panorama geral sobre as técnicas supervisionadas de mineração de dados *support vector machine with polynomial kernel* é apresentado. A seguir, os procedimentos para a análise de classificação são descritos, bem como da verificação do nível de acurácia dos modelos de classificação ajustados. Uma verificação detalhada da capacidade de prever com precisão a melhor opção de heurística de melhoria conforme as características das instâncias é proposta.

O *support vector machine with polynomial kernel* (SVM) [10] é uma técnica de aprendizado de máquina supervisionada que tenta encontrar padrões (ou clusters) no conjunto de dados de entrada e usa essas informações para classificar novos conjuntos de dados com base em suas características. O objetivo é maximizar as distâncias entre o ponto de dados mais próximo de todas as classes e o hiperplano, uma curva capaz de definir a fronteira que melhor segrega as classes. Uma característica atrativa é a convexidade do problema de otimização tratado para ajustar o modelo de classificação, que resulta na existência de apenas um mínimo ótimo global, evitando restringir as soluções encontradas com a técnica a mínimos locais. Além disso, o SVM ignora possíveis outliers de uma classe específica, maximizando o hiperplano de acordo com as características predominantes da classe, o que torna o SVM muito robusto para problemas em que outliers são comumente encontrados. Tais fatos estão relacionados as características da formulação dos algoritmos SVM com o uso de funções kernel não-lineares.

A heurística construtiva *bottom-left* (BL), e a sua variante *bottom-left-fill* (BLF), foram escolhidas para encontrar soluções completas viáveis. A variável dependente utilizada para descrever a melhor classe, neste estudo denominadas como heurísticas de melhoria, pode assumir seis valores diferentes: *Completely random local search* (CR); *dynamic random* (DR); *random weight procedure* (RW); *tabu search* (TS); *simulated annealing* (SA); e *genetic algorithm* (GA). Cada instância deve ser caracterizada por uma ou mais classes (heurísticas de melhoria) que possuem o menor *gap*.

A primeira etapa da análise de classificação compreende a definição do tamanho da amostra utilizada para o BL e BLF. O tamanho da amostra original proposto para *Dataset* é de 15666 instâncias. Porém, instâncias com mais de uma heurística de melhoria capaz de representar uma solução de qualidade para a variável dependente foram excluídas, restando um total de 13750 instâncias.

Para compreender melhor o comportamento dos resultados, um diagrama de *Venn* foi desenhado para visualizar as interseções existentes para o desempenho das heurísticas de melhoria para as 15666 instâncias de *Dataset* (Figura 2), usando o *InteractiVenn* (Heberle2015).

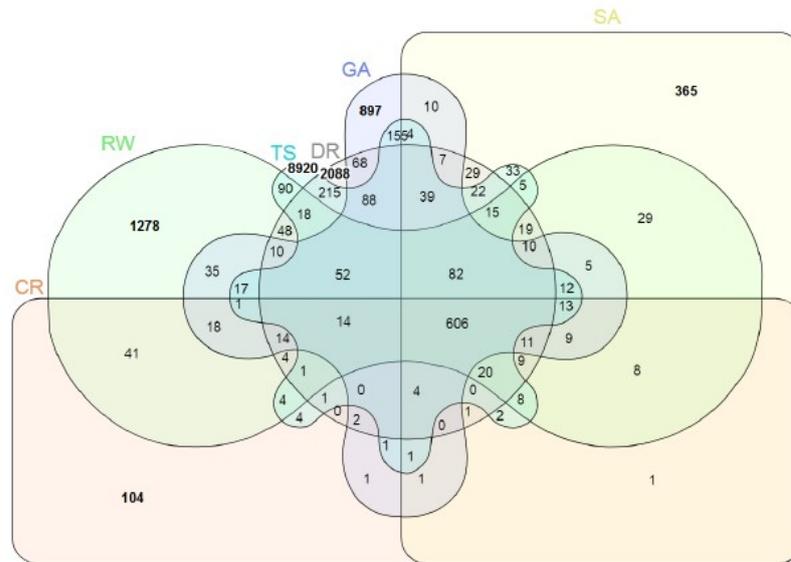


Figura 2 – Diagrama de Venn.

Um total de 2014 instâncias *multi-label* com mais de uma heurística de melhoria representativa foram encontradas, constituindo um grupo de instâncias que possui o maior nível de complexidade para serem explicadas, ou até mesmo determinar um padrão do comportamento das características para cada classe.

O diagrama de *Venn* é útil neste estudo para identificar as interseções entre heurísticas de melhoria com maior número de ocorrências, a fim de definir possíveis motivos para que as ocorrências mais frequentes terem ocorrido. Em complemento, as instâncias pertencentes a uma determinada interseção podem ser úteis para identificar características do problema que tornam a busca por padrões na classificação das heurísticas de melhoria difusas ou complexas, a ponto de as técnicas supervisionadas de mineração de dados não serem capazes de encontrar padrões.

A próxima etapa foi dividir cada novo conjunto de dados em duas partes: Conjunto de dados de treinamento (80%) e conjunto de dados de teste (20%). O conjunto de dados de treinamento é usado para ajustar modelos de classificação. O conjunto de dados de teste é usado para verificar o nível de precisão do melhor modelo de classificação selecionado para cada heurística de melhoria.

Para melhorar a qualidade da análise de classificação, o processo de validação cruzado cinco vezes (*5-fold cross-validation*) foi utilizado, dobrando a amostra de treinamento em cinco partes diferentes. O conjunto de dados de treinamento pode ser dividido em mais dobras com um pequeno número de observações se o número de dobras aumentar, o desafio é encontrar o melhor número de amostras dobradas para evitar encaixar os modelos de sobreposição. As instâncias do problema foram inseridas aleatoriamente em apenas uma dobra. Um total de dez modelos de classificação foram gerados, dois para cada amostra dobrada, e o modelo de classificação com a maior precisão entre as classes previstas e as classes calculadas foi usado para selecionar o modelo de classificação ideal. A Tabela 1 apresenta uma pequena descrição sobre os parâmetros de entrada definidos para o SVM no pacote “*caret*” [11] usando o software *RStudio*.

Tabela 1: Parâmetros de entrada.

Técnica	Parâmetro	Descrição	BL	BLF
SVM	<i>degree</i>	Grau polinomial	2	2
	<i>scale</i>	Escala	0.01	0.01
	<i>C</i>	Custo	1	1

A Tabela 2 mostra a acurácia das predições realizadas com os modelos de classificação ajustados. Em complemento, estão descritas as heurísticas de melhoria que se apresentaram como mais dominantes e a margem de erro das predições.

Tabela 2: Nível de acurácia usando o conjunto de dados de treinamento.

Heurística construtiva	Acurácia	Prevalência	95% intervalo de confiança
BL	0.66	TS; DR; RW	-0,0183, +0,0178
BLF	0.65	TS; DR; RW	-0,0182, +0,0178

Em média, a acurácia dos modelos de classificação ajustados é 58%, aceitável para o nível de complexidade e das características do problema. Diferenças de acurácia foram verificadas devido as diferenças intrínsecas existentes entre as heurísticas de melhoria escolhidas o teste do framework proposto. Destaca-se que a acurácia obtida é fortemente impactada pela estratégia *multi-label* selecionada para desenvolver a análise de classificação. Na prática, a estratégia *multi-label* se mostrou eficiente para ajustar os modelos do 2D-SPP, confirmando as expectativas realizadas na etapa de testes apresentada anteriormente. Para todas as estratégias, a margem de erro das predições com base no intervalo de confiança equivalente a 95% está entre 0,017 e 0,019, tanto positivamente quanto negativamente.

Para o BL e BLF os resultados são mais uniformes. A escolha do TS como melhor opção de heurística de melhoria foi verificada para mais de 60% das instâncias de teste utilizadas. O fato do BL ser fortemente dependente da ordem de entrada dos retângulos, além de ser a heurística de melhoria que não permite mudanças na ordem ao decorrer do processo de posicionamento dos retângulos na faixa evidencia a tendência pela maior uniformidade dos resultados verificados.

Apesar do uso de diferentes técnicas supervisionadas de mineração de dados, a acurácia e a prevalência de BF e BFF é similar. A busca pela melhor combinação entre o espaço na faixa e as dimensões dos retângulos é simples, sendo a ordenação inicial dos retângulos um fator de razoável relevância para aumentar a eficiência na busca por soluções de qualidade. Tal fato é preponderante para que heurísticas de melhoria como DR, RW e TS se sobressaíam em relação as demais, visto que a mudança no ordenamento de entrada dos retângulos não é completamente aleatória como em CR. Por outro lado, não possui uma estrutura tão rígida como a verificada para GA e SA.

Sobre a performance das heurísticas de melhoria, a estratégia que possui a performance mais efetiva da vasta maioria das instâncias está diretamente relacionada ao TS. De modo geral, TS e DR possuem a maior capacidade de obter soluções de melhor qualidade, pois foram capazes de realizar uma busca eficiente no espaço das possíveis soluções viáveis para o 2D-SPP. DR e RW são similares por permitirem uma busca semi-aleatória no espaço de soluções, armazenando informações relevantes de iterações passadas para a construção de um novo ordenamento de entrada dos retângulos. Por outro lado, não são rígidas pois, a cada iteração, permitem que partes do ordenamento sejam desenvolvidos de maneira aleatória. Em oposição, GA e SA possuem uma estrutura mais rígida para a mudança no ordenamento dos retângulos, sendo altamente sensíveis aos valores dos parâmetros de entrada. Como valores genéricos foram considerados, fato que é uma premissa

da aplicação das heurísticas de melhoria neste estudo, uma série de instâncias não atingiram um desempenho esperado, visto a grande amplitude adotada no gerador de instâncias para cobrir a maior parte das características do problema. Finalmente, CR não guarda qualquer informação das iterações passadas, explorando o espaço das soluções de forma muito ampla e aleatória, o que torna a busca por soluções de qualidade pouco intuitiva.

Para visualizar em detalhes a performance dos modelos de classificação ajustados, a Tabela 3 mostra a matriz de confusão entre os valores observados e preditos da variável dependente para BL e BLF.

Tabela 3: Matriz de confusão.

BL: Pre- dicted	BL: Observed						BLF: Pre- dicted	BLF: Observed					
	CR	SA	DR	RW	GA	TS		CR	SA	DR	RW	GA	TS
CR	0	0	0	0	0	0	CR	0	0	0	0	0	0
SA	0	0	0	0	0	0	SA	0	0	0	0	0	0
DR	0	0	2	1	0	0	DR	0	0	0	0	0	0
RW	2	0	2	12	4	5	RW	1	1	6	10	4	5
GA	0	0	0	0	0	0	GA	0	0	0	0	0	0
TS	24	22	470	263	178	1763	TS	19	72	411	245	175	1779

O modelo de classificação ajustado para BL e BLF possui uma baixa precisão para prever heurísticas de melhoria diferentes de TS. Uma das causas é a alta prevalência de TS (0,643 para BL e 0,654 para BLF), viés que se refletiu fortemente na sensibilidade do modelo em se adequar a situações em que outra heurística de melhoria é a melhor opção. Em contrapartida, o modelo é praticamente perfeito para identificar quando uma heurística de melhoria não é a mais recomendada de ser utilizada para uma determinada instância, devido ao alto valor da especificidade obtido. Tal fato é importante pois, mesmo o modelo não sendo preciso para prever qual heurística de melhoria é melhor, desconsiderando a TS, pode fornecer informações importantes para o descarte de heurísticas de melhoria que, necessariamente, não são as mais adequadas para as características de uma determinada instância.

Algumas observações gerais foram realizadas após observar os resultados da etapa de testes da análise de classificação. Destaca-se que a semelhança no padrão dos resultados quando do uso ou não da estratégia “Fill” para posicionar os retângulos dentro da faixa.

Por se tratar do uso de técnicas supervisionadas de mineração de dados para análise de classificação, é natural esperar que os modelos de classificação sejam capazes de prever com maior acurácia situações em que a prevalência de uma heurística de melhoria se sobressai sobre outras. Uma tendência é que a proporção da prevalência conduza a maiores valores na sensibilidade do modelo em prever a heurística de melhoria que mais se aproxima as características da instância. Outro comportamento interessante observado é capacidade do modelo manter proporcional a prevalência determinante.

De modo geral, um maior nível de acurácia não foi obtido devido a impossibilidade de superar algumas dificuldades que são intrínsecas do problema. A principal dificuldade verificada foi traduzir de maneira mais precisa a grande quantidade de variações geométricas no arranjo entre os retângulos dentro da faixa, com base nas diferentes possibilidades de dimensões que os retângulos e a faixa podem assumir. Apesar disso, pode-se afirmar que as variáveis explicativas utilizadas para mensurar as características das instâncias do problema fornecem dados relevantes e consistentes sobre os retângulos e a faixa, e que a limitação apresentada é resultante da complexidade do 2D-SPP. A pior consequência prática deste efeito pode ser a identificação de uma mesma heurística de melhoria para instâncias com

características diferentes. Porém, tal fato não dificultou para que um bom nível de acurácia, com base nos padrões no modelo de classificação fossem encontrados.

4. CONCLUSÃO

Após a condução da pesquisa foi possível observar que a técnicas supervisionadas de mineração de dados *support vector machine with polynomial kernel* foi eficiente no que tange a busca pela melhor opção de heurística de melhoria para o contexto do 2D-SPP. Ainda, notou-se que as características das instâncias são capazes de remeter informações importantes para a resolução do problema.

Apesar de serem similares, as heurísticas construtivas escolhidas para fazer parte do presente estudo forneceram informações relevantes sobre a forma com que modelos de classificação supervisionados podem ser utilizados no contexto dos problemas de corte e empacotamento. Assim, a abordagem proposta é inovativa no sentido de ser uma abordagem diferente das tradicionais adotadas para resolver problemas deste tipo para problemas de corte e empacotamento. Em desenvolvimento está a aplicação da metodologia em outras heurísticas construtivas notórias na literatura, como a *Best-Fit* e a *Fast-Heuristic*. A longo prazo a expectativa é aplicar os passos aplicados neste estudo para a prospecção de cenários em outros problemas de corte e empacotamento como, por exemplo, o *bin packing problem* ou o *knapsack problem*.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Wäscher, G., Haußner, H., & Schumann, H. (2007). An improved typology of cutting and packing problems. *European journal of operational research*, 183(3), 1109-1130.
- [2] Martello, S., Monaci, M., & Vigo, D. (2003). An exact approach to the strip-packing problem. *INFORMS Journal on Computing*, 15(3), 310-319.
- [3] Alvarez-Valdés, R., Parreño, F., & Tamarit, J. M. (2008). Reactive GRASP for the strip-packing problem. *Computers & Operations Research*, 35(4), 1065-1083.
- [4] Hopper, E. B. C. H., & Turton, B. C. (2001). An empirical investigation of meta-heuristic and heuristic algorithms for a 2D packing problem. *European Journal of Operational Research*, 128(1), 34-57.
- [5] Ntene, N., & van Vuuren, J. H. (2009). A survey and comparison of guillotine heuristics for the 2D oriented offline strip packing problem. *Discrete Optimization*, 6(2), 174-188.
- [6] Neuenfeldt Júnior, A. (2017). *The Two-Dimensional Rectangular Strip Packing Problem*. PhD Thesis.
- [7] Júnior, A. N., Silva, E., Gomes, A. M., Soares, C., & Oliveira, J. F. (2019). Data mining based framework to assess solution quality for the rectangular 2D strip-packing problem. *Expert Systems with Applications*, 118, 365-380.
- [8] Oliveira, J. F., Neuenfeldt Júnior, A., Silva, E., & Carravilla, M. A. (2016). A survey on heuristics for the two-dimensional rectangular strip packing problem. *Pesquisa Operacional*, 36(2), 197-226.

- [9] Júnior, A. N., Silva, E., Gomes, A. M., & Oliveira, J. F. (2017, June). The Two-Dimensional Strip Packing Problem: What Matters?. In Congress of APDIO, the Portuguese Operational Research Society (pp. 151-164). Springer, Cham.
- [10] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- [11] Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of statistical software*, 28(5), 1-26.