Comparação entre aplicação de modelo linear e Algoritmo Genético em agrupamento de pacientes com agravamento de COVID-19

Arthur Dantas Mangussi[a]

Patricia Shizue Matsumura Ueda[a]

Lucas de Abreu Julião[a]

Victor Castro Nacif de Faria[a]

Ana Carolina Lorena[a]
[a]Programa de Pós-Graduação em Pesquisa Operacional UNIFESP/ITA
Instituto Tecnológico de Aeronáutica
{arthuradm,victorvcnf,aclorena}@ita.br

RESUMO

Uma das formas de descrever a diferença entre observações de um conjunto de dados é por meio do agrupamento das observações que compõem o conjunto. Neste caso, as observações são organizadas em um ou mais grupos de acordo com a similaridade de suas características. Este trabalho faz a comparação de desempenho de um modelo linear inteiro (modelo exato) e da meta-heurística Algoritmo Genético (AG) para identificar agrupamentos de pacientes hospitalares no município de São José dos Campos, no período de pico da pandemia. O conjunto de dados inicial apresentava desbalanceamento com relação ao agravamento do quadro de COVID-19. Este atributo foi considerado apenas para o balanceamento da base de dados, com o objetivo de selecionar, na mesma proporção, as amostras para o experimento computacional. Tais amostras, com tamanho de 10, 20, 30 e 50 pontos, foram sorteadas aleatoriamente do conjunto de dados balanceado, totalizando 12 amostras. Uma amostra adicional com 100 pontos, a partir da mesma metodologia, também foi gerada para análise. A comparação entre os algoritmos é feita tanto pelos valores de minimização da distância euclidiana entre as unidades amostrais (variância intra-grupos) quanto no tempo de processamento para obtenção da solução pelos dois algoritmos. Para comparação dos resultados entre o modelo exato e o AG, criou-se uma métrica, denominada GAP. O AG atinge o valor ótimo do modelo exato para as 6 primeiras instâncias com 10 e 20 pontos. Entretanto, quando aumenta-se o número de pontos, o AG não encontra a resposta ótima para o problema proposto, apresentando um GAP médio de 14,84%. Para 100 vértices, o modelo exato não obtém a solução ótima com tempo de processamento inferior a 10 horas, o qual foi utilizado como critério de parada do modelo. Contudo, o tempo total de processamento, considerando 10 repetições do AG, foi, em média, 11 minutos. De forma geral, entende-se que a aplicação de Algoritmo Genético é eficiente em termos de tempo e qualidade da solução.

PALAVRAS CHAVE. Otimização, Agrupamento, Algoritmo Genético.