



RESOLUÇÃO DE PROBLEMA DE FORMAÇÃO DE GRUPOS UTILIZANDO UM ALGORITMO EVOLUCIONÁRIO

LAÍS GONÇALVES VIANA – lais.viana@ufv.br
UNIVERSIDADE FEDERAL DE VIÇOSA - UFV - RIO PARANAÍBA

HELEN VITALINE DE CASTRO SANTOS – helen.castro@ufv.br
UNIVERSIDADE FEDERAL DE VIÇOSA - UFV - RIO PARANAÍBA

THIAGO HENRIQUE NOGUEIRA – thiagoh.nogueira@ufv.br
UNIVERSIDADE FEDERAL DE VIÇOSA - UFV - RIO PARANAÍBA

RAIANE RIBEIRO MACHADO GOMES – raianemachado@ufv.br
UNIVERSIDADE FEDERAL DE VIÇOSA - UFV - RIO PARANAÍBA

LARISSA SOUSA CAMPOS – larissa.sousa@ufv.br
UNIVERSIDADE FEDERAL DE VIÇOSA - UFV - RIO PARANAÍBA

ÁREA: 3. PESQUISA OPERACIONAL

SUBÁREA: 3.1 – MODELAGEM, SIMULAÇÃO E OTIMIZAÇÃO

RESUMO: O PROBLEMA DE FORMAÇÃO DE GRUPOS (GFP) É FUNDAMENTAL EM ÁREAS COMO LOGÍSTICA, EDUCAÇÃO E GESTÃO DE RECURSOS HUMANOS, ONDE É NECESSÁRIO OTIMIZAR A DIVISÃO DE ELEMENTOS EM SUBGRUPOS CONFORME MÉTRICAS ESPECÍFICAS. ALGORITMOS GENÉTICOS (AGS) SURGEM COMO UMA SOLUÇÃO EFICAZ PARA ESTE PROBLEMA, CAPAZES DE LIDAR COM OBJETIVOS DIVERSOS, COMO A FORMAÇÃO DE GRUPOS HOMOGÊNEOS OU HETEROGÊNEOS, E DE ATENDER A RESTRIÇÕES ESPECÍFICAS, COMO O EQUILÍBRIO NA CARGA DE TRABALHO OU NAS HABILIDADES DOS MEMBROS. EM AGS, CADA INDIVÍDUO DA POPULAÇÃO REPRESENTA UMA SOLUÇÃO POSSÍVEL, AVALIADA POR CRITÉRIOS DE APTIDÃO, COM OS MAIS APTOS CONTRIBUINDO PARA AS PRÓXIMAS GERAÇÕES. ESTUDOS ANTERIORES COMPROVAM A EFICÁCIA DOS AGS NA RESOLUÇÃO DO GFP EM DIFERENTES CONTEXTOS, COMO NA MANUFATURA E NA FORMAÇÃO DE GRUPOS ESCOLARES. ESSA ABORDAGEM OFERECE FLEXIBILIDADE E EFICIÊNCIA, TORNANDO-SE UMA FERRAMENTA VALIOSA PARA ENFRENTAR DESAFIOS PRÁTICOS NA FORMAÇÃO DE GRUPOS.

PALAVRAS-CHAVES: PROBLEMA DE FORMAÇÃO DE GRUPOS; ALGORITMOS GENÉRICOS; OTIMIZAÇÃO; EFICIÊNCIA.

1. INTRODUÇÃO

O *Group Formation Problem* (GFP), ou Problema de Formação de Grupos, é uma questão fundamental em diversas áreas, como logística, educação, organização de eventos e gestão de recursos humanos. O desafio consiste em agrupar elementos em subconjuntos otimizados, de acordo com critérios específicos, para maximizar ou minimizar determinadas métricas, como eficiência, diversidade, equidade ou satisfação dos participantes. Esse problema é conhecido por sua complexidade, especialmente em cenários com muitos elementos e restrições a serem consideradas (KING; NAKORNCHA, 1982).

Para enfrentar o *Group Formation Problem*, uma das abordagens mais promissoras é o uso de Algoritmos Genéticos (AGs). Algoritmos Genéticos são métodos inspirados no processo evolutivo da seleção natural, que se baseiam na evolução de populações de soluções candidatas para encontrar aproximações às melhores soluções para um problema. (LAMBORA; GUPTA; CHOPRA, 2019).

Através da aplicação de Algoritmos Genéticos ao *Group Formation Problem*, é possível alcançar diferentes objetivos, como a formação de grupos homogêneos ou heterogêneos, a minimização do desvio padrão das características dos grupos ou a maximização da diversidade entre eles. Além disso, os AGs podem lidar com restrições e preferências específicas, como a necessidade de balancear a carga de trabalho, a compatibilidade de habilidades entre membros dos grupos ou a consideração de limites de tamanho para os subconjuntos. (KATOCH; CHAUHAN; KUMAR, 2021).

Nesta abordagem, cada indivíduo da população representa uma possível solução para o *Group Formation Problem*, e sua avaliação é feita com base em critérios de aptidão que refletem o quão bem a solução atende aos objetivos e restrições estabelecidos. Ao longo das gerações, os indivíduos mais aptos têm maior probabilidade de contribuir com seus genes para a próxima geração, enquanto os menos aptos são eliminados gradualmente (SMITH, 1978). Dessa forma, os Algoritmos Genéticos conduzem a uma busca eficiente pelo espaço de soluções, resultando em grupos otimizados ou aproximadamente otimizados, dependendo da complexidade do problema (MIRJALILI; MIRJALILI, 2019).

Em suma, o uso de Algoritmos Genéticos para abordar o *Group Formation Problem* é uma abordagem poderosa e flexível, capaz de encontrar soluções eficientes e adaptáveis para a formação de grupos em diversos contextos. Sua capacidade de explorar amplamente o

espaço de soluções e lidar com restrições torna-os uma ferramenta valiosa para enfrentar desafios práticos de agrupamento de elementos.

2. REFERENCIAL TEÓRICO

Brown e Sumichrast (2001) testaram um algoritmo genético de agrupamento (GGA) para resolver o problema de formação de células de máquinas e peças (MPCF) na manufatura. O MPCF aborda a formação de famílias de peças com base nos requisitos de processamento dos componentes e a identificação de grupos de máquinas com base em sua capacidade de processar famílias específicas de peças. O GGA foi avaliado quanto à qualidade de suas soluções usando medidas de eficiência e eficácia.

Ani (2010) utilizou a abordagem de algoritmos genéticos para resolver um problema de formação de grupos. Uma avaliação dos indivíduos foi utilizada para categorizar o conhecimento em programação de cada indivíduo e foram criados times de trabalho que possuísem habilidades semelhantes em programação. O estudo mostrou que o método utilizado é capaz de produzir grupos equilibrados, nos quais cada grupo é composto por estudantes com habilidades de programação boas, moderadas e fracas.

Moreno, Ovalle e Vicari (2012) propõe um método baseado em algoritmo genético para formar grupos inter-homogêneos e intra-heterogêneos, considerando que a formação de grupos é um dos principais processos na aprendizagem colaborativa. Esse método permite considerar várias características dos alunos, transformando o problema de agrupamento em uma otimização multiobjetivo. Para validar o método foram consideradas três características: estimativa dos níveis de conhecimento, habilidades de comunicação e habilidades de liderança dos estudantes. Os resultados do experimento validaram o método, tanto do ponto de vista computacional, medindo o desempenho do algoritmo, quanto do ponto de vista pedagógico, comparando os resultados dos alunos com duas estratégias tradicionais de formação de grupos: aleatória e auto-organizada.

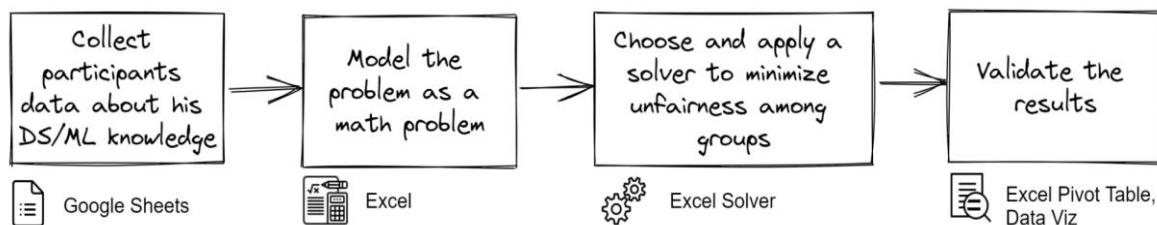
3. DESENVOLVIMENTO

O trabalho foi desenvolvido para formar grupos equilibrados para uma competição dentro de evento de programação que conta com participantes de níveis variados. O objetivo principal do evento é promover a aprendizagem no campo da ciência de dados, além de oferecer prêmios para as equipes vencedoras. No entanto, surge uma pergunta crucial: como os organizadores podem garantir que os participantes maximizem sua aprendizagem durante o desafio?

Além disso, há uma preocupação com a composição dos grupos, especialmente a presença de novatos ou especialistas em algumas equipes. O cenário ideal envolveria grupos equilibrados, onde indivíduos com diferentes estilos de aprendizagem e níveis de expertise se complementam de forma eficaz. Para alcançar isso, foram implementadas estratégias para avaliar os níveis de habilidade e estilos de aprendizagem dos participantes antes de formar as equipes, garantindo uma mistura de níveis de experiência e expertise. Dessa forma, o evento pode oferecer um ambiente de aprendizagem inclusivo, permitindo que os indivíduos aprendam uns com os outros e, em última análise, aprimorem seus conhecimentos e habilidades em ciência de dados.

Foram registrados no evento centenas de participantes de forma que organizar os grupos manualmente seria inviável e pouco eficaz, uma vez que o objetivo central é formar grupos com mesmo nível de conhecimento. O ponto de partida foi, portanto coletar informações sobre o nível de conhecimento de cada participante sobre *Data Science* e *Machine Learning*, para em seguida ser feita a modelagem matemática do problema, a aplicação do *solver* e por fim a validação dos dados, conforme apresentado na figura 1.

FIGURA 1 - Fluxograma simplificado da proposta de trabalho



Fonte: Autores (2023)

O primeiro passo do processo para resolver esse problema foi coletar dados que pudessem ajudar a formar grupos equilibrados. Nesse sentido, foi feito um *brainstorming* para pensar em quais dados poderiam ajudar a conhecer a expertise dos participantes. Foi feito o refinamento das perguntas e possíveis respostas, e criado um formulário usando o *Google Forms* para coletar esses dados. O foco foi em perguntas como: "Qual é o seu nível de conhecimento sobre problemas de regressão?" com algumas opções fechadas para escolher: "Baixo", "Médio" e "Alto". Respostas como essas podem ser facilmente convertidas em números e colocadas em uma faixa (0: Baixo, 1: Médio, 2: Alto) e depois normalizadas para uma escala (por exemplo, [0, 1]).

As respostas das questões 1 a 8 foram coletadas no *Google Forms*, codificadas e normalizadas para números usando um codificador e um ajuste de escala para normalização. A coluna "*participant_id*" é um identificador único, a coluna "*group*" indica o grupo em que o participante está e a coluna "*score*" é uma combinação linear de pontos com base nas respostas de cada participante para cada pergunta, sendo o peso das perguntas de 1 a 8 respectivamente 0,1; 0,1; 0,125; 0,15; 0,1; 0,1; 0,125 e 0,2.

FIGURA 2 - Exemplo de dados coletados sobre os participantes

group	participant_id	question 1	question 2	question 3	question 4	question 5	question 6	question 7	question 8	score
2	1	1.00	0.50	0.50	0.67	1.00	0.00	0.00	0.00	0.4125
8	2	0.00	0.00	0.50	0.33	0.00	0.00	0.00	0.00	0.1125
1	3	1.00	0.50	0.50	0.67	1.00	1.00	1.00	0.25	0.6875
1	4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0000
2	5	0.00	0.50	0.00	0.00	0.00	1.00	0.50	1.00	0.4125

Fonte: Autor

O segundo passo foi transformar esse problema em um problema matemático. A função objetivo era produzir grupos que tivessem a mesma quantidade de expertise conjunta. Ou seja, se fosse possível atribuir uma pontuação de expertise a cada participante, seria possível calcular a pontuação de expertise do grupo ao qual ele pertencia. Para gerar uma pontuação de grupo, somou-se as pontuações de cada participante. Em uma situação ideal, cada grupo teria a mesma pontuação, então a soma das diferenças absolutas entre cada grupo e os outros grupos seria zero. Portanto, a função objetivo pode ser reescrita como minimizar as diferenças absolutas entre cada grupo.

O foco era na atribuição de grupos para cada participante (variáveis de decisão). Se o grupo do participante de id 8 fosse trocado com o participante de id 6, haveria uma nova formação para o grupo 2 e uma nova pontuação para o grupo 2. Desta forma, com a troca de indivíduos entre grupos é possível testar novas combinações que possibilitem ao algoritmo de otimização decidir os grupos para cada participante de forma que as diferenças nas pontuações dos grupos fossem zero. Portanto, o algoritmo de otimização automaticamente alteraria a coluna "grupo" tentando alcançar nosso objetivo principal.

FIGURA 3 - Exemplo de troca de participante em um grupo e modificação do score

group	participant_id	question 1	question 2	question 3	question 4	question 5	question 6	question 7	question 8	score
2	1	1.00	0.50	0.50	0.67	1.00	0.00	0.00	0.00	0.4125
2	5	0.00	0.50	0.00	0.00	0.00	1.00	0.50	1.00	0.4125
2	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0000
2	9	0.00	0.50	0.00	0.33	1.00	0.00	0.00	0.00	0.2000
2	10	0.00	1.00	0.50	0.33	1.00	0.00	0.00	0.00	0.3125
2	11	0.25	0.50	0.50	0.67	1.00	0.00	0.00	0.00	0.3375
										group 2 score 1.6750
group	participant_id	question 1	question 2	question 3	question 4	question 5	question 6	question 7	question 8	score
2	1	1.00	0.50	0.50	0.67	1.00	0.00	0.00	0.00	0.4125
2	5	0.00	0.50	0.00	0.00	0.00	1.00	0.50	1.00	0.4125
2	6	0.75	1.00	0.50	1.00	1.00	1.00	1.00	0.50	0.8125
2	9	0.00	0.50	0.00	0.33	1.00	0.00	0.00	0.00	0.2000
2	10	0.00	1.00	0.50	0.33	1.00	0.00	0.00	0.00	0.3125
2	11	0.25	0.50	0.50	0.67	1.00	0.00	0.00	0.00	0.3375
										group 2 score 2.4875

Fonte: Autor

Por fim, as restrições foram definidas para o sistema, objetivando a geração de soluções viáveis e adequadas à população:

- Todos os indivíduos devem ser associados a um grupo apenas obrigatoriamente;
- Os grupos devem ter 6 participantes;
- O número de grupos é 24.

A partir do modelo estruturado foi utilizado o *solver* do Excel, objetivando a agilidade na geração de soluções viáveis. Para tal foi estabelecida uma solução viável inicial como ponto de partida, observando todas as restrições listadas acima. A entrada de dados no solver é apresentada na figura 3, sendo que o método de solução empregado foi o algoritmo evolucionário, por se tratar de um algoritmo probabilístico que fornece um mecanismo de busca paralela e adaptativa, adequado ao problema apresentado.

FIGURA 4 - Exemplo de dados coletados sobre os participantes

Set Objective: ↑

To: ☐ Max ☒ Min ☐ Value Of:

By Changing Variable Cells: ↑

Subject to the Constraints:

\$A\$3:\$A\$169 <= \$T\$5
 \$A\$3:\$A\$169 = integer
 \$A\$3:\$A\$169 >= 1
 \$P\$4:\$P\$31 <= 6
 \$P\$4:\$P\$31 >= 6

☒ Make Unconstrained Variables Non-Negative

Select a Solving Method: Evolutionary Options

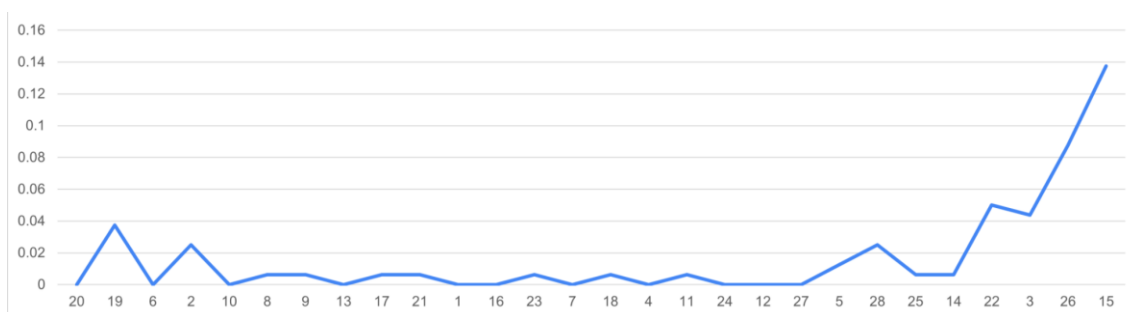
Solving Method
 Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.

Fonte: Autor

4. RESULTADOS E DISCUSSÃO

Após algumas execuções, o *Excel Evolutionary Solver* conseguiu reduzir a função objetivo em cerca de dez vezes, mas ao visualizar os resultados, ficou claro que alguns grupos estavam desequilibrados figura 5. Os grupos foram ordenados em ordem decrescente de pontuação e foram calculadas as diferenças entre eles, para cada grupo. Por exemplo, o grupo 19 tem uma pontuação de 1,7, e o grupo 20 tem uma pontuação de 1,7375; portanto, dado que o grupo 20 é o predecessor do grupo 19, a diferença entre eles é de 0,0375.

FIGURA 5 - Diferença entre os grupos a partir do grupo predecessor



Fonte: Autor

Conforme apresentado na figura 5, é visível que grupos 22, 3, 26 e principalmente o 15 não parecem estar equilibrados em comparação com os outros, De forma complementar foram então foram feitas algumas trocas manuais de participantes entre os grupos, para tornar melhor a configuração dos grupos. A implementação do modelo em outras plataformas além

do *Excel Solver* e utilizando outros algoritmos é imprescindível para a melhoria contínua dos resultados.

5. CONCLUSÃO

A metodologia utilizada para seleção dos grupos é mais eficaz e eficiente que a seleção manual ou aleatória dos grupos, quando considerada a necessidade de criar grupos equilibrados entre si. O fator "justiça" na escolha dos grupos contribui com o desenvolvimento da competição e impulsiona a troca de aprendizado dos participantes. De forma análoga os membros dos grupos estão mais propensos a sentir que todos os grupos têm chance de serem vitoriosos, reduzindo a frustração ao longo do desenvolvimento do torneio.

Como contribuição futura se objetiva a implementação do modelo matemático em *Python* e utilizando conceitos de *Machine Learning* em inteligências artificiais, para se ter uma base de comparação dos modelos e consequentemente a melhoria do sistema. A evolução do questionário para perguntas mais aplicadas e a avaliação das respostas por uma inteligência artificial também é um recurso a ser explorado, em busca de se alcançar ainda mais assertividade no entendimento das capacidades de cada indivíduo.

REFERÊNCIAS

- ANI, Zhamri Che et al. A method for group formation using genetic algorithm. **International Journal on Computer Science and Engineering**, v. 2, n. 9, p. 3060-3064, 2010.
- BROWN, Evelyn C.; SUMICHRIST, Robert T. CF-GGA: a grouping genetic algorithm for the cell formation problem. **International Journal of Production Research**, v. 39, n. 16, p. 3651-3669, 2001.
- KATOCH, Sourabh; CHAUHAN, Sumit Singh; KUMAR, Vijay. A review on genetic algorithm: past, present, and future. **Multimedia tools and applications**, v. 80, p. 8091-8126, 2021.
- KING, John R.; NAKORNCHAI, Vias. Machine-component group formation in group technology: review and extension. **The international journal of production research**, v. 20, n. 2, p. 117-133, 1982.
- LAMBORA, Annu; GUPTA, Kunal; CHOPRA, Kriti. Genetic algorithm-A literature review. In: **2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)**. IEEE, 2019. p. 380-384.
- MIRJALILI, Seyedali; MIRJALILI, Seyedali. Genetic algorithm. **Evolutionary Algorithms and Neural Networks: Theory and Applications**, p. 43-55, 2019.

MORENO, Julián; OVALLE, Demetrio A.; VICARI, Rosa M. A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. **Computers & Education**, v. 58, n. 1, p. 560-569, 2012.

SMITH, J. Maynard. Optimization theory in evolution. **Annual review of ecology and systematics**, v. 9, n. 1, p. 31-56, 1978.