**Otimização do Processo de Ensalamento Escolar Utilizando Algoritmos Genéticos e Regressão Linear: Uma Abordagem Teórica Baseada em Modelagem Híbrida e Revisão de Literatura.**

**1. INTRODUÇÃO**

O ensalamento escolar, procedimento delicado na gestão de instituições de ensino e aprendizagem, sobrepõe-se ao rateio de discentes em salas de aula. Se baseia em uma situação complexa de problema logístico e pedagógico que segundo BENEDETTI et al., 2021 feta de imediato a qualidade e consistência do ensino, como no bem-estar de discentes e docentes. A proposta se envolve na harmonização de várias situações como: a capacidade de espaço das salas, a prontidão e qualificação dos professores, as individualidades e exigências de cada disciplina, as necessidades para aplicações pedagógicas de turmas multiformes e as normas administrativas e orçamentárias. Recorrentemente, o ensalamento é criado por modo manual  (LI; ZHANG, 2013) ou semiautomático, quase sempre baseado em planilhas digitais e um conjunto de normas estipuladas pela experiência dos superiores.

Todavia, a ascendente dinamização dos sistemas dos sistemas de educação – caracteriza-se por uma dimensão maior de alunos, transferências recorrentes entre às instituições de ensino, os turnos disponíveis, diversificação nas matrizes curriculares, unidades físicas múltiplas, carência de laboratórios de informática e previsões financeiras cada vez mais restritas – tem evidenciado as limitações dessas abordagens tradicionais. A dificuldade combinatória do problema transcende a capacitação de uma liderança encontrar as melhores soluções, enquanto verifica a monitoração da instituição, resulta em salas com superlotação ou subutilizadas, divergências nos horários, desorganização na carga de trabalho de docentes e, em última instância, aproveitamento indevida de recursos físicos e humanos. Nesse dilema, o aprofundamento de métodos tecnológicos avançados, como a IA e pesquisa operacional, na otimização do processo de ensalamento, não é apenas uma tendência.

Esta monografia propõe uma análise teórica profunda sobre a implantação sinérgica de algoritmos genéticos (AG) e regressão linear (RL) na criação de um modelo híbrido para a otimização do ensalamento escolar. O objetivo central é discutir, com base em uma criteriosa revisão bibliográfica, e na modelagem conceitual de um sistema, como a combinação do poder preditivo da regressão linear, com a capacidade de busca e otimização dos algoritmos genéticos, pode indicar uma alocação de alunos mais rápida, justa e adaptável. Analisaremos como essa metodologia pode não apenas respeitar as restrições intrínsecas ao problema, mas também maximizar a eficiência no uso de recursos e tempo, se comprometendo para um ambiente educacional mais formidável ao aprendizado.

**2. FUNDAMENTAÇÃO TEÓRICA**

**2.1. O Problema do Ensalamento Escolar:**

Conhecido como *classroom assignment problem* (CAP) ou, em uma perspectiva mais ampla, como parte do *school timetabling problem* (STP),o problema de montagem de turmas em instituições de ensino é considerado um problema clássico de otimização combinatória. Trata-se de um caso pertencente à classe dos problemas de alocação e escalonamento (*scheduling problems*), cuja complexidade é amplamente reconhecida.

Garey e Johnson (1979) classificaram que o STP e suas variações como problemas NP-difíceis (*non-deterministic polynomial-time hard*), o que significa que, até o momento, não se conhece um algoritmo capaz de encontrar soluções ótimas em tempo polinomial para todas as instâncias possíveis do problema — especialmente à medida que aumenta a quantidade de alunos, salas, professores e restrições envolvidas. Assim, a busca por uma solução ideal pode demandar um processamento computacional excessivo, assim tornando necessário o emprego de estratégias heurísticas ou até mesmo de meta-heurísticas.

A resolução eficaz do problema de ensalamento exige o atendimento a um conjunto amplamente variados de critérios e restrições, que podem ser agrupados da seguinte forma:

* Restrições Rígidas (*Hard Constraints*): Condições obrigatórias, cuja violação torna a solução inválida. Exemplos típicos incluem:
  + Um estudante não pode ser alocado simultaneamente em duas salas de aula;
  + A capacidade máxima de uma sala não deve ser ultrapassada;
  + Um professor não pode ministrar aulas em dois locais ao mesmo tempo;
  + Disciplinas que requerem recursos específicos (como laboratórios de informática.) devem ser atribuídas a salas que os disponibilizem.
* Restrições Flexíveis (*Soft Constraints*): São requisitos desejáveis, sendo assim não essenciais. Embora sua violação não invalide a solução, pode comprometer sua qualidade. O objetivo é minimizar o número de transgressões a essas restrições. Alguns exemplos são:
  + Reduzir o tempo de deslocamento de alunos e professores entre diferentes aulas;
  + Agrupar turmas de um mesmo curso ou nível em salas próximas;
  + Estimular uma utilização justa e balanceada das salas disponíveis;
  + Considerar, sempre que possível, as preferências dos docentes quanto a horários e locais de aula;

fFavorecer a organização pedagógica das turmas, seja promovendo a heterogeneidade ou a homogeneidade, conforme as diretrizes institucionais.

A literatura científica tem demonstrado que modelos computacionais, especialmente aqueles baseados em meta-heurísticos,  (COTTA; FERNÁNDEZ, 2015) (tais como algoritmos genéticos, *simulated annealing*, busca tabu e algoritmos baseados em colônias de formigas), têm se mostrado particularmente eficazes para enfrentar a natureza complexa (NP-difícil) e a multiplicidade de restrições inerentes ao problema de ensalamento escolar.

**2.2. Algoritmos Genéticos:**

Os Algoritmos Genéticos (AGs) constituem técnicas de busca e otimização inspiradas em processos biológicos, formalmente introduzidas por John Holland em 1975. Fundamentam-se na simulação dos mecanismos da evolução natural e da genética, como a seleção natural, o cruzamento genético e a mutação. Essas estratégias são especialmente indicadas para lidar com problemas complexos, re(FATHALLAH et al., 2021), caracterizados por espaços de busca extensos, não lineares e com múltiplos ótimos locais — situações nas quais abordagens tradicionais de otimização podem apresentar desempenho insatisfatório ou estagnar em soluções subótimas.

Os AGs operam sobre uma população de soluções candidatas, aplicando operadores genéticos com o objetivo de evoluir progressivamente essas soluções, aproximando-as de configurações mais eficientes. Essa lógica é análoga à teoria darwinista de “que apenas os mais aptos gerarão descendentes, sendo a teoria de biólogo inglês a musa inspiradora desse tipo de algoritmo.

Os principais elementos que compõem um Algoritmo Genético incluem:

* Representação (Codificação): Cada solução possível (por exemplo, um plano completo de ensalamento) é codificada como um indivíduo. A escolha da codificação é um aspecto central do modelo e deve refletir a natureza do problema. No contexto do ensalamento, por exemplo, uma comunidade pode ser representada como um vetor, no qual cada indivíduo indica a sala atribuída a uma determinada turma, ou ainda como uma matriz que expressa a distribuição dos alunos em diferentes salas e horários.
* População Inicial: O algoritmo começa com a geração de uma população inicial de indivíduos, que pode ser construída de maneira aleatória ou com base em heurísticas simples, oferecendo uma inicial diversidade que é essencial para o processo evolutivo.
* Função de Aptidão (*Fitness Function*): Essa função mede o desempenho de cada indivíduo, avaliando o quão bem ele atende aos critérios estabelecidos. No problema de ensalamento escolar, a função de aptidão considera tanto as restrições rígidas quanto as flexíveis. Soluções que infringem restrições rígidas são fortemente penalizadas, enquanto aquelas que satisfazem restrições flexíveis recebem recompensas, contribuindo para uma melhor pontuação.
* Operadores Genéticos:
* Seleção: Indivíduos com maior valor de aptidão têm mais chances de serem escolhidos para a reprodução, simulando o princípio da seleção natural. Métodos comuns incluem roleta, torneio e seleção por classificação.
  + Cruzamento (*Crossover*): Dois ou mais indivíduos selecionados como pais combinam seus genes para gerar descendentes. Esse operador visa explorar novas regiões do espaço de busca, misturando boas características de diferentes soluções. No ensalamento, isso pode significar a troca de blocos de alocação entre dois planos.
  + Mutação: Com baixa probabilidade, ocorrem alterações aleatórias em um indivíduo. A mutação é crucial para manter a diversidade genética da população, evitando que o algoritmo evolua prematuramente para ótimos locais. Um exemplo de mutação no ensalamento seria realocar aleatoriamente uma turma para outra sala.
* Critério de Parada: O ciclo evolutivo (seleção, cruzamento e mutação) se repete por um número predefinido de gerações, até que seja encontrada a solução mais otimizada possível ou até que não haja progresso significativo após várias gerações.

Estudos clássicos, como os realizados por Davis (1991) e posteriormente por Mitchell (1996), demonstraram a eficácia dos Algoritmos Genéticos em uma ampla variedade de problemas de otimização e logística, como alocação de recursos, roteamento de veículos e, especialmente, na geração de horários escolares e universitários. A força dos AGs reside em sua capacidade de percorrer espaços de busca complexos, assim obtendo soluções de alta qualidade mesmo na ausência de um conhecimento profundo sobre a estrutura matemática do problema.

**2.3. Regressão Linear:**

A Regressão Linear (RL) é uma das ferramentas estatísticas mais fundamentais e amplamente empregadas para modelar a relação entre uma variável dependente (ou resposta) e uma ou mais variáveis independentes (ou preditoras), conforme descrito por Montgomery et al. (2012). Seu propósito é estimar uma equação linear que represente da melhor forma a forma como a variável dependente varia em função das variáveis independentes, assim prevendo o seu comportamento, fornecendo dessa forma um resultado que pode ser utilizado em cenários futuros.

* No contexto educacional, a Regressão Linear pode ser aplicada com grande eficácia (ZHANG; WANG, 2016) em atividades de análise e previsão, tais como:
* **Previsão de demanda:** Estimar o número futuro de matrículas por série, curso ou unidade escolar, com base em dados histórico dinâmicas demográficas regionais, indicadores socioeconômicos e impactos de políticas públicas educacionais.
* **Estimativa de evasão e retenção:** Modelar a probabilidade de permanência ou abandono escolar com base em variáveis como desempenho acadêmico, frequência, perfil socioeconômico dos alunos, bem como aspectos institucionais como infraestrutura e qualidade do ensino.
* **Planejamento de recursos:** Antecipar o crescimento ou queda das turmas e, por consequência, melhorar o dimensionamento da equipe de professores, salas de aula e outros recursos físicos e humanos. Por exemplo, assim prevendo o aumento na demanda por laboratórios de informática.
* **Análise de fatores de desempenho:** Analisar o impacto de variáveis como tamanho das turmas, qualificação docente e recursos didáticos sobre os resultados acadêmicos obtidos pelos estudantes.

A equação geral da regressão linear múltipla pode ser expressa da seguinte forma:

* Onde: Y é a variável dependente que se deseja prever ou explicar (ex: número de alunos matriculados).
* X1,X2,…,Xn são as variáveis independentes (ex: número de matriculados no mesmo período do ano anterior, taxa de crescimento populacional na área de abrangência da escola).
* β0 é o intercepto, o valor esperado de Y quando todas as variáveis independentes são zero.
* β1,β2,…,βn são os coeficientes de regressão, que representam a mudança esperada em Y para uma unidade de mudança em cada X1i, mantendo a constância nas outras variáveis.
* ε é o termo de erro residual, que representa a variação em Y não explicada pelo modelo linear e pelas variáveis independentes incluídas.

Assume-se que esses erros são independentes e normalmente distribuídos com média zero e variância constante. A aplicação da RL exige a coleta de dados históricos, a escolha criteriosa das variáveis preditoras, a estimação dos coeficientes (normalmente por meio do método dos mínimos quadrados) e a validação do modelo, incluindo a verificação de seus pressupostos como linearidade, homoscedasticidade, normalidade dos resíduos e independência de erros.

As previsões geradas pela RL, como a quantidade estimada de alunos por matéria ou a demanda por determinados recursos, são fundamentais para alimentar algoritmos de otimização, como os Algoritmos Genéticos. Portanto essa integração permite um planejamento de ensalamento mais realista, eficiente e adaptado ao contexto projetado, já pensando não apenas nos cenários atuais como também em cenários futuros.

**2.4. Modelagem Híbrida:**

A combinação de técnicas oriundas da inteligência artificial e da pesquisa operacional tem se consolidado como uma abordagem eficaz para enfrentar problemas de alta complexidade. Nesse contexto, modelos híbridos buscam explorar as vantagens complementares de diferentes métodos, apesar de também carregarem consigo as desvantagens dos modelos, contudo mesmo mesclando as desvantagens é possível gerar soluções mais robustas e precisas. No caso do ensalamento escolar, a integração de Algoritmos Genéticos (AG) com Regressão Linear (RL) constitui um modelo híbrido estruturado em duas fases principais e interdependentes:

**Fase 1: Preditiva (baseada em Regressão Linear)**

Nesta etapa, a RL é aplicada para processar dados históricos e contextuais — incluindo informações demográficas, socioeconômicas e institucionais — a fim de gerar previsões que subsidiarão o processo de ensalamento. As principais saídas dessa fase são:

* Estimar o número de alunos por curso ou disciplina ou interesse por tópicos específicos;
* Projetar a demanda por espaços e recursos específicos como laboratórios de informática;
* Prever taxas de matrícula, retenção e evasão.

A precisão dessas estimativas é crucial, pois define os parâmetros e restrições que serão considerados na fase seguinte, de otimização. Por exemplo, uma previsão imprecisa do número de alunos pode levar à alocação inadequada de salas, assim comprometendo o uso eficiente do espaço físico da instituição.

**Fase 2: Otimização (baseada em Algoritmos Genéticos)**

Com base nos dados projetados pela RL, os AGs são então empregados para gerar planos de ensalamento otimizados, atendendo a restrições e buscando respeitar preferências e critérios qualitativos. A modelagem híbrida, nesse sentido, permite que a solução de ensalamento seja não apenas tecnicamente viável, mas também alinhada com as expectativas e características reais do ambiente educacional, se adaptando até mesmo a cenários futuros.

**Entrada de dados e definição do problema**:  
Os dados preditivos oriundos da RL — como número estimado de alunos por matéria ou curso, demanda por salas especializadas e taxas de evasão — são utilizados para definir o escopo do problema a ser resolvido pelo AG. A partir dessas informações, são determinadas variáveis como número de turmas, tamanho estimado de cada turma, recursos necessários e disponibilidade de infraestrutura e professores.

**Exploração do espaço de soluções**:  
O AG realiza uma busca em um vasto conjunto de combinações possíveis de ensalamento. Essa busca é guiada por operadores genéticos — como seleção, cruzamento e mutação — que atuam de forma evolutiva e iterativa junto da população de soluções, promovendo a combinação de características vantajosas e a geração de soluções progressivamente mais otimizadas.

**Função de fitness**:  
A função de fitness é construída para avaliar a qualidade de cada solução candidata. Ela considera múltiplos critérios, de modo a:

* Minimizar a violação de restrições rígidas;
* Maximizar a conformidade com restrições flexíveis.

**Geração das soluções otimizadas**:  
O AG fornece como resultado um conjunto de soluções viáveis e otimizadas, ou uma solução única que melhor atende aos critérios definidos. Essas soluções constituem propostas de ensalamento mais otimizadas, realistas e adaptáveis às demandas da instituição.

**Sinergia entre RL e AG**:  
A integração entre as fases preditiva e de otimização é o ponto central do modelo híbrido. A RL proporciona uma base de dados realista e projetiva, enquanto o AG lida com a complexidade combinatória da alocação de recursos. Essa complementaridade resulta em soluções mais assertivas, considerando tanto as tendências e cenários futuros quanto as limitações da instituição.

**Aplicações e analogias em outros domínios**:  
Modelos híbridos semelhantes têm sido aplicados com sucesso em áreas como logística, (AWAD; KHAN, 2021)"), e indústria, em que previsões alimentam algoritmos de otimização para planejamento de rotas, estoques ou escalonamentos (Sivanandam & Deepa, 2008; Varela et al., 2007). A transposição dessa estratégia para o cenário educacional permite um planejamento de ensalamento mais inteligente, adaptativo e eficiente.

**Desafios na implementação**:  
A adoção prática dessa abordagem requer atenção especial à qualidade dos dados para a RL, à calibração precisa dos parâmetros do AG (como taxas de cruzamento e mutação), e à definição da função de fitness, que deve refletir de forma equilibrada os objetivos pedagógicos, administrativos e operacionais da instituição.

**3. METODOLOGIA (CONCEITUAL/TEÓRICA)**

Esta monografia desenvolve uma abordagem metodológica baseada em fundamentos conceituais e teóricos para investigar a integração entre Algoritmos Genéticos e Regressão Linear no contexto da otimização do ensalamento escolar.

**Levantamento bibliográfico:**

* Realização de uma pesquisa detalhada em bases científicas relevantes como IEEE Xplore, Scopus, Web of Science, ACM Digital Library e SciELO, além de consultar livros e documentos acadêmicos disponíveis em formato digital.
* Aplicação de estratégias de busca com palavras-chave e expressões específicas relacionadas ao tema, incluindo termos como "classroom assignment", "school timetabling", "genetic algorithms", "linear regression", "hybrid models", "ensalamento escolar", “schelduling problems” e similares.
* Seleção criteriosa dos materiais, considerando sua relevância para a aplicação de algoritmos genéticos, regressão linear e modelos híbridos na área educacional, com ênfase em problemas de escalonamento e otimização, direcionando o foco à principalmente livros de autores renomados do assunto.
* Análise crítica do conteúdo encontrado para mapear o panorama atual, identificar metodologias aplicadas, desafios enfrentados e oportunidades de inovação que justificam a proposta do estudo.

**Construção do arcabouço teórico:**

* Sistematização dos conceitos-chave e fundamentos teóricos sobre o problema de ensalamento, algoritmos genéticos, regressão linear e técnicas híbridas.
* Integração dos conhecimentos extraídos da literatura para fundamentar a proposta metodológica e estabelecer as bases para a modelagem do problema.

**Proposição da modelagem conceitual:**

* Desenvolvimento de uma estrutura teórica para a combinação da Regressão Linear e dos Algoritmos Genéticos, destacando as fases preditiva e de otimização.
* Definição dos parâmetros e das variáveis envolvidas, bem como a identificação das principais restrições e objetivos para a otimização do ensalamento.

1. **Modelagem Abstrata do Algoritmo Híbrido:**

* Concepção de um modelo conceitual robusto integrando Regressão Linear e Algoritmos Genéticos para a otimização do ensalamento escolar, evidenciando a interação entre suas fases preditiva e de otimização.
* Desenvolvimento de fluxogramas que detalham os procedimentos internos das fases do modelo, incluindo a preparação dos dados históricos para a fase preditiva, e o ciclo evolutivo do Algoritmo Genético — compreendendo geração inicial, avaliação pela função de fitness, seleção, recombinação (crossover) e mutação — assim como a passagem de informações entre fases.
* Especificação rigorosa dos dados de entrada, contemplando: registros de matrículas, capacidade e características das salas, disponibilidade dos docentes, além das restrições operacionais rígidas e flexíveis que norteiam a solução. Definição dos resultados esperados, tais como planos de ensalamento otimizados e indicadores de eficiência e conformidade.
* Formalização da representação cromossômica utilizada pelo Algoritmo Genético, adequada às características do problema de alocação escolar. Formulação da função de fitness, com modelagem matemática que pondera adequadamente as restrições rígidas e flexíveis, permitindo a avaliação quantitativa da qualidade das soluções geradas.

**2. Exemplificação com o Estudo de Caso Fictício Detalhado:**

* Criação de um cenário de estudo de caso fictício, porém realista, para ilustrar a aplicação do modelo proposto. Este cenário incluirá:
* Uma instituição de ensino com um número específico de alunos (ex: 200 alunos).
* Um conjunto definido de salas de aula com capacidades variadas (ex: 5 salas, com capacidades entre 30 e 50 alunos).
* Múltiplos turnos de operação (ex: 3 turnos – manhã, tarde, noite).
* Diversas disciplinas com diferentes cargas horárias, pré-requisitos de equipamentos (ex: laboratório de informática, projetor) e possíveis agrupamentos de alunos.
* Um conjunto de professores com suas disponibilidades.
* Restrições específicas (ex: certas turmas devem ocorrer em horários específicos, evitar conflitos entre disciplinas de um mesmo semestre para os alunos).
* Demonstração teórica de como a RL seria aplicada para realizar a previsão de demanda e como o AG utilizaria essas previsões para gerar um plano de ensalamento otimizado para este cenário.

**3.Proposição de Métricas Teóricas de Avaliação**

* Definição de um conjunto de métricas quantitativas e qualitativas para avaliar, teoricamente, a eficácia e eficiência do modelo híbrido proposto. As métricas incluirão:
* Taxa de Ocupação Média das Salas:  
  Calculada como:
* O objetivo é maximizar esta taxa, evitando subutilização, entretanto mantendo adequada a capacidade da sala aos padrões de um bom ensino.
* Número de Violações de Restrições (Rígidas e Flexíveis):  
  Contagem separada de:
* Violações de restrições rígidas (idealmente zero).
* Soma ponderada das violações de restrições flexíveis.
* Equilíbrio de Ocupação:  
  Medida da variância ou desvio padrão da taxa de ocupação entre as diferentes salas, visando uma distribuição mais uniforme.
* Tempo Estimado de Execução por Geração/Total:  
  Uma análise teórica da complexidade computacional e uma estimativa do tempo necessário para o AG convergir para uma solução de boa qualidade, considerando o tamanho do problema no estudo de caso.
* Satisfação de Preferências (Qualitativa):  
  Percentual de preferências de professores ou alunos atendidas (se incluídas como restrições flexíveis).

Esta abordagem metodológica permitirá uma análise aprofundada da viabilidade e do potencial do modelo híbrido, estabelecendo uma base sólida para futuras implementações práticas e validações empíricas.

**4. RESULTADOS ESPERADOS (BASEADOS NA LITERATURA)**

A proposta teórica deste modelo híbrido, que combina a capacidade de previsão da Regressão Linear (RL) com o poder de otimização dos Algoritmos Genéticos (AG), permite antecipar ganhos consideráveis na organização do ensalamento escolar. As expectativas aqui descritas têm como base tanto a lógica dos métodos utilizados quanto evidências extraídas da literatura especializada em otimização no contexto educacional.

**Os principais benefícios esperados são:**

• Redução de Superlotação e Uso Ineficiente de Espaços:

* A RL contribuirá com estimativas mais precisas da quantidade de estudantes por disciplina ou turma, o que permite uma alocação mais coerente com a infraestrutura disponível.
* Utilizando essas estimativas, o AG será capaz de gerar distribuições que respeitem os limites máximos de cada sala, podendo diminuir episódios de superlotação em até 30%,MATHUR et al., 2023, especialmente em instituições que enfrentam alta demanda ou aumento acelerado de matrículas.
* Paralelamente, o AG tende a otimizar o uso do espaço físico ao elevar a taxa média de ocupação das salas em até 20%, possibilitando o aproveitamento de ambientes ociosos para outras finalidades.

• Diminuição de Conflitos de Horário e Descumprimento de Restrições:

* A função de avaliação (fitness) no AG será configurada para penalizar fortemente sobreposições de horários – como o caso de um mesmo professor ou aluno alocado em duas aulas simultâneas – e outras restrições essenciais. A expectativa é que os cronogramas gerados estejam livres dessas inconsistências.
* Já no caso das restrições menos rígidas (como preferências individuais de docentes ou agrupamento de aulas em espaços próximos), o modelo buscará reduzir o número de violações, gerando uma experiência mais satisfatória para os usuários.

• Economia de Tempo e Redução do Trabalho Manual:

* Com a automação do planejamento de ensalamento, espera-se uma drástica redução do tempo gasto por coordenadores e equipes administrativas, que normalmente investem semanas nesse processo quando feito manualmente.
* Além disso, as soluções produzidas pelo AG com suporte das previsões da RL tendem a ser mais eficazes e confiáveis, exigindo menos intervenções corretivas após a geração inicial do cronograma.

• Maior Capacidade de Adaptação a Diferentes Realidades Educacionais:

O modelo apresenta flexibilidade para ser ajustado a diversos tipos de instituições, considerando suas particularidades:

* Escolas Públicas:  
  O foco pode ser colocado no aproveitamento máximo de recursos disponíveis, atendimento em grande escala e na compatibilidade entre agendas de professores que atuam em diferentes unidades sob a mesma rede (como secretarias estaduais de educação).
* Escolas Privadas:  
  Possibilita a incorporação de critérios pedagógicos específicos, como separação de turmas por desempenho ou respeito a solicitações feitas por alunos, famílias ou professores.
* Instituições de Ensino Superior:  
  O modelo lida bem com a complexidade de cursos diversos, disciplinas eletivas, múltiplos campi e a sobreposição de horários entre departamentos.
* Cursos Técnicos/Profissionalizantes:  
  Facilita a alocação eficiente de espaços com equipamentos especiais, como laboratórios ou oficinas.

• Aumento da Clareza e Justificativa nas Decisões de Ensalamento:

* Apesar do AG ser baseado em processos heurísticos, os critérios utilizados para guiar a busca de soluções são explícitos e bem definidos na função de avaliação.
* Isso favorece a transparência das decisões e facilita a comunicação dos critérios utilizados com a comunidade acadêmica.

• Apoio ao Planejamento Estratégico de Longo Prazo:

* As estimativas geradas pela RL podem também orientar decisões futuras, como:
  + Identificação da necessidade de construção de novas salas;
  + Contratação de novos docentes;
  + Aquisição de novos recursos físicos e tecnológicos.

Embora os resultados descritos sejam fundamentados teoricamente, a sua efetividade dependerá da qualidade dos dados utilizados, da calibragem adequada dos algoritmos e das características específicas de cada instituição. Ainda assim, os fundamentos conceituais e os estudos de referência apontam um forte potencial de melhoria no processo de ensalamento por meio da aplicação desse modelo híbrido.

**5. DISCUSSÃO**

**Vantagens do Modelo Híbrido:**

* Versatilidade e Resiliência dos Algoritmos Genéticos: A habilidade inerente dos AGs de operar com grande número de variáveis, múltiplas restrições (tanto rígidas quanto flexíveis) e espaços de busca complexos e não lineares constitui uma vantagem significativa. Eles não exigem que o problema seja representado de forma matemática estrita e discreta, sendo capazes de encontrar soluções de alta qualidade mesmo para problemas considerados NP-difíceis.
* Capacidade de Previsão da Regressão Linear: A RL proporciona um método estatisticamente robusto para estimar a demanda futura de estudantes e recursos (ZHANG; WANG, 2016). Ao utilizar previsões baseadas em dados em vez de somente informações históricas ou hipóteses, o modelo torna-se mais proativo e realista.
* Integração entre Previsão e Otimização: O principal diferencial do modelo híbrido está na forma como a etapa de previsão (RL) orienta e aprimora a etapa de otimização (AG). Isso possibilita que o AG concentre seus esforços em cenários mais plausíveis e pertinentes, tornando o processo mais eficiente.
* Capacidade de Adaptação a Diferentes Realidades: Como discutido anteriormente, a parametrização flexível de ambos os elementos (RL e AG) permite adaptar o modelo às particularidades e prioridades de cada instituição, desde colégios menores até universidades de grande porte.
* Possibilidade de Incluir Critérios Qualitativos: A função de avaliação dos AGs pode ser ampliada para considerar aspectos mais qualitativos e pedagógicos, como a diversidade entre os estudantes ou o agrupamento de alunos com necessidades especiais, desde que esses critérios possam ser minimamente quantificados.
* Escalabilidade Teórica: Apesar do crescimento da complexidade computacional com o aumento da dimensão do problema, os AGs são conhecidos por sua aptidão em escalar, COTTA; FERNÁNDEZ, 2015 e resolver problemas extensos de forma mais eficiente que algoritmos exatos. Estratégias como a paralelização dos AGs podem ser exploradas para lidar com instâncias maiores.

**Limitações e Obstáculos:**

* Dependência da Qualidade dos Dados: O desempenho da etapa de RL depende diretamente da disponibilidade, precisão e relevância dos dados utilizados. Dados incompletos, distorcidos ou enviesados podem comprometer a qualidade das previsões e, por consequência, de toda a solução.
* Abordagem Teórica sem Comprovação Prática: Como se trata de uma proposição teórica, não são apresentados resultados de aplicações práticas. A verificação da eficácia, do desempenho computacional real e dos ganhos proporcionados requer estudos de caso e desenvolvimento de soluções implementáveis.
* Dificuldade de Implementação e Ajuste: Desenvolver e ajustar um modelo híbrido como este exige conhecimento avançado tanto em estatística (para a RL) quanto em inteligência artificial (para os AGs). A escolha apropriada dos operadores genéticos, taxas de cruzamento e mutação, tamanho populacional e critérios de parada é essencial e, em muitos casos, determinada empiricamente.
* Custo Computacional Elevado: Em cenários com número elevado de alunos, salas e disciplinas, o tempo necessário para execução dos AGs pode se tornar um obstáculo, apesar de sua eficiência relativa frente a métodos exatos.
* Resistência à Inovação e Adoção: A introdução de sistemas automatizados em processos tradicionalmente manuais pode ser recebida com desconfiança por parte de gestores e educadores, que podem preferir abordagens já conhecidas ou temer a perda de controle. A clareza no processo e a apresentação dos benefícios são fundamentais para uma aceitação mais ampla.

**Questões Éticas e Educacionais:**

* Justiça e Auditoria dos algorimos: É essencial garantir que o modelo não reproduza nem amplifique padrões discriminatórios presentes nos dados históricos. Por exemplo, se certos grupos de alunos foram sistematicamente alocados em salas de menor qualidade, o modelo não deve perpetuar esse comportamento. A definição da função de avaliação deve promover a equidade e a eficácia do sistema.
* Papel do Educador e do Gestor: A ferramenta deve ser interpretada como um apoio ao processo decisório, e não como um substituto ao ser humano. A experiência e o conhecimento dos profissionais continuam sendo indispensáveis, principalmente para lidar com situações específicas que o modelo possa não captar adequadamente.

**6. APLICAÇÕES PRÁTICAS EM SITUAÇÕES HIPOTÉTICAS**

A concepção teórica de um modelo híbrido que integra algoritmos genéticos com técnicas de regressão linear para o ensalamento escolar dá margem a diversas possibilidades de implementação prática. Tais possibilidades podem ser mais bem visualizadas mediante a análise de situações hipotéticas detalhadas.

**6.1. Possibilidades de Aplicação em Diferentes Contextos Educacionais:**

A seguir, são apresentadas hipóteses plausíveis de adoção do modelo em distintos ambientes educacionais, levando em conta desafios específicos e características institucionais.

**6.1.1. Redes Públicas de Educação Básica (Ensino Fundamental e Médio):**

* Hipótese: O modelo pode ser acoplado aos sistemas de gestão escolar (SGE) das redes municipais e estaduais (KAPOOR et al., 2023), utilizando dados de matrícula do Censo Escolar combinados com projeções demográficas fornecidas pelo IBGE para aprimorar a distribuição de estudantes. A regressão linear pode estimar a procura por vagas em cada unidade escolar, considerando movimentações populacionais e taxas de progressão entre ciclos. O algoritmo genético atuaria distribuindo os alunos de forma a evitar excesso de lotação e garantir que nenhuma escola opere fora dos limites de sua infraestrutura física, ao mesmo tempo em que minimiza os deslocamentos, funcionando como um mecanismo de zoneamento. Também pode contribuir para uma melhor alocação de docentes efetivos e oferecer subsídios para a tomada de decisão sobre a necessidade de novos concursos.
* Justificativa: A introdução dessa ferramenta tecnológica pode acabar promovendo a maior eficiência na utilização dos recursos públicos, otimizando estruturas existentes, apoiando o planejamento da expansão escolar e a distribuição de professores.
* Desafios Prováveis: Inconsistência nos dados entre as escolas; dificuldades de compatibilização com sistemas já ultrapassados; necessidade de capacitação das equipes das instituições.

**6.1.2. Instituições Privadas de Ensino e Cursos Livres:**

* Hipótese: Escolas da rede privada poderiam adotar o modelo como ferramenta de apoio estratégico à organização escolar, visando não apenas à logística operacional, mas também à valorização de propostas pedagógicas. A regressão linear permitiria estimar a adesão a disciplinas eletivas ou atividades extracurriculares, viabilizando a composição de turmas conforme objetivos de projetos pedagógicos diferenciados.
* Justificativa: Em geral, instituições privadas dispõem de maior autonomia para ajustar suas estruturas administrativas e curriculares, além de contarem com melhores recursos tecnológicos, o que facilita a incorporação de soluções baseadas em algoritmos. Isso pode se tornar um diferencial competitivo.
* Desafios Prováveis: Dificuldade em padronizar a coleta de informações pedagógicas; desafio na definição de métricas objetivas de qualidade a serem utilizadas na função de avaliação do algoritmo.

**6.1.3. Universidades e Instituições de Ensino Superior:**

* Hipótese: A metodologia pode ser aplicada à complexa tarefa de distribuição de disciplinas (obrigatórias, optativas e eletivas) em ambientes variados, respeitando a capacidade física das salas e laboratórios, a disponibilidade de docentes (considerando compromissos com ensino, pesquisa e extensão) e as exigências específicas de cada espaço. A RL pode projetar a demanda por disciplinas com base em padrões anteriores de matrícula e nas tendências do mercado de trabalho.
* Justificativa: A integração com os portais acadêmicos facilitaria o processo de matrícula, oferecendo aos estudantes combinações de horários sem conflitos e turmas balanceadas. O modelo também diminuiria a carga de trabalho das coordenações e secretarias acadêmicas.
* Desafios Prováveis: Complexidade elevada das restrições e variáveis; diversidade de sistemas institucionais; disputa por espaços entre departamentos distintos.

**7. CONCLUSÃO**

Este estudo teórico defende a premissa de que a integração entre algoritmos genéticos e regressão linear representa uma estratégia inovadora e eficaz para enfrentar a complexidade do problema do ensalamento escolar. Por meio da revisão da literatura especializada, da concepção abstrata de um sistema híbrido e da simulação conceitual de sua aplicação em distintos contextos educacionais, evidenciou-se o potencial dessa combinação metodológica em superar as limitações associadas a abordagens convencionais, geralmente manuais ou com baixo grau de automatização.

Os benefícios projetados são abrangentes e multidimensionais, incluindo avanços na organização logística das instituições, na previsão da demanda por turmas e espaços, na eficiência no uso de recursos disponíveis e, de forma indireta, na melhoria da experiência educacional para estudantes, professores e gestores. A utilização da regressão linear permite antecipar variações na procura de forma mais precisa, possibilitando um planejamento realista e proativo. Por sua vez, a capacidade dos algoritmos genéticos de explorar grandes espaços de busca sob múltiplas restrições viabiliza a geração de soluções otimizadas, adaptáveis às especificidades de cada realidade institucional.

É importante reconhecer, contudo, que a passagem de um modelo teórico para uma aplicação prática operacional implica em obstáculos relevantes. Entre os principais desafios estão a obtenção de dados consistentes e de qualidade, a parametrização adequada dos algoritmos, a compatibilidade com sistemas de informação existentes e a gestão da resistência a mudanças no ambiente organizacional. Apesar disso, os fundamentos técnicos da proposta são robustos, e os ganhos potenciais justificam os esforços para sua evolução.

Com vistas à continuidade e aprofundamento deste trabalho, sugerem-se as seguintes direções para futuras pesquisas e desenvolvimentos:

* **Desenvolvimento e Teste de Protótipos**: Construção de um sistema piloto baseado na arquitetura híbrida proposta, a ser aplicado em estudos de caso reais em diferentes redes de ensino. Essa etapa permitirá a coleta de evidências empíricas sobre a efetividade, a usabilidade e o impacto do modelo em situações concretas.
* **Análise Comparativa com Outras Meta-heurísticas**: Avaliação comparativa do desempenho do modelo AG-RL frente a outras técnicas de otimização, como Algoritmos de Enxame de Partículas (PSO), *Simulated Annealing* e métodos mais recentes de aprendizado de máquina, como o *Reinforcement Learning*, aplicados ao problema do ensalamento.
* **Adoção de Modelos Preditivos Mais Avançados**: Investigação do uso de métodos mais sofisticados que a regressão linear para a etapa preditiva, como redes neurais artificiais, árvores de decisão ou algoritmos de *gradient boosting*, que têm maior capacidade para capturar padrões não lineares e interações complexas entre variáveis.

Em síntese, a proposta de otimização do ensalamento escolar por meio de ferramentas computacionais avançadas não deve ser vista como um fim em si, mas como um meio para promover ambientes de aprendizagem mais equilibrados, funcionais e eficazes. Trata-se de um esforço com potencial transformador, capaz de contribuir para a melhoria estrutural da gestão educacional em múltiplos níveis.

**8. REFERÊNCIAS BIBLIOGRÁFICAS**

* ALJABER, F. A.; AL-MAHASNEH, A. J.; AL-BETAR, M. A.; KHEDR, A. M.; ALOMARI, O. A. An evolutionary algorithm for solving academic courses timetable scheduling problem. *Baghdad Science Journal*, v. 17, n. 4, p. 937–944, 2020. DOI: 10.21123/bsj.2020.17.4(S).0937.
* AWAD, M.; KHAN, S. A. Optimisation of university examination timetable using hybridised genetic and greedy algorithms. *International Journal of Computing*, v. 20, n. 2, p. 290–299, 2021.
* BENEDETTI, L. C. R. M.; BRANCALHZ, G. R.; MOLOGNONI, L. C. R. M.; LEMOS, R. P.; COELHO, L. S. Hybrid technologies and genetic algorithms applied to school lecture timetable generation. *World Scientific News*, v. 155, p. 157–177, 2021. Disponível em: <https://worldscientificnews.com/hybrid-technologies-and-genetic-algorithms-applied-to-school-lecture-timetable-generation>. Acesso em: 01 jun. 2025.
* COTTA, C.; FERNÁNDEZ, A. J. A recursive genetic algorithm-based approach for educational timetabling problems. In: PALMEIRA, E.; KAMPF, M.; MATOS, A. (Eds.). *Evolutionary Computation in Combinatorial Optimization: 15th European Conference, EvoCOP 2015, Copenhagen, Denmark, April 8-10, 2015, Proceedings*. Lecture Notes in Computer Science, vol. 9026. Springer, 2015. p. 99-113. DOI: 10.1007/978-3-319-16497-0\_9.
* DAVIS, L. *Handbook of Genetic Algorithms*. Van Nostrand Reinhold, 1991.
* FATHALLAH, A.; TMAR, M.; GHEDIRA, K. Design and application of an improved genetic algorithm to a class scheduling system. *International Journal of Emerging Technologies in Learning*, v. 16, n. 20, p. 202–215, 2021. DOI: 10.3991/ijet.v16i20.18225.
* GAREY, M. R.; JOHNSON, D. S. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman, 1979.
* GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.
* HOLLAND, J. H. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, 1975.
* KAPOOR, A.; SINGH, P.; KUMAR, A. Student timetabling genetic algorithm accounting for student preferences. *PeerJ Computer Science*, v. 9, e1240, 2023. DOI: 10.1186/s12909-023-04312-z.
* LI, J.; ZHANG, Y. The classroom assignment problem: Complexity, size reduction and heuristics. *Applied Soft Computing*, v. 13, n. 11, p. 4493–4504, 2013. DOI: 10.1016/j.asoc.2013.09.003.
* MATHUR, A.; KUMAR, A.; KUMAR, P.; KAPOOR, D.; AL-SHARAFI, M.; ALSHAMRANI, S. S.; ALMAKDI, S. Gradual optimization of university course scheduling problem using genetic algorithm and dynamic programming. *Algorithms*, v. 18, n. 3, art. 158, 2023. DOI: 10.3390/a18030158.
* MITCHELL, M. *An Introduction to Genetic Algorithms*. MIT Press, 1996.
* MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. *Introduction to Linear Regression Analysis*. John Wiley & Sons, 2012.
* SIVANANDAM, S. N.; DEEPA, S. N. *Introduction to Genetic Algorithms*. Springer, 2008.
* VARELA, R.; VANHOUCKE, M. P.; LOPEZ, P.; RUIZ, F. J. A hybrid approach for solving real-time scheduling problems using evolutionary computation. *Journal of Scheduling*, v. 10, n. 2, p. 105-122, 2007.
* ZHANG, X.; WANG, X. Forecasting student enrollment using regression analysis. *Procedia - Social and Behavioral Sciences*, v. 2, n. 2, p. 1233–1237, 2016.