UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

LÚCIO DA CRUZ DE MORAES

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

Cornélio procópio

2025

**LÚCIO DA CRUZ DE MORAES**

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

**Optimization of School Classroom Assignment Process Using Genetic Algorithms and Linear Regression: A Theoretical Approach Based on Hybrid Modeling and Literature Review**

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia de Software da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Rodrigo Henrique Cunha Palácios

**CORNÉLIO EM PROCÓPIO**

**2025**

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

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia de Software da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Rodrigo Henrique Cunha Palácios

Data de aprovação: A Preencher

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Nome completo e por extenso do Membro 1 (de acordo com o Currículo Lattes) – A Preencher

Titulação (Especialização, Mestrado, Doutorado)

Nome completo e por extenso da instituição a qual possui vínculo

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Nome completo e por extenso do Membro 2 (de acordo com o Currículo Lattes) – A Preencher

Titulação (Especialização, Mestrado, Doutorado)

Nome completo e por extenso da instituição a qual possui vínculo

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Nome completo e por extenso do Membro 3 (de acordo com o Currículo Lattes) – A Preencher

Titulação (Especialização, Mestrado, Doutorado)

Nome completo e por extenso da instituição a qual possui vínculo

**CORNÉLIO EM PROCÓPIO**

**2025**

Dedico este trabalho à minha família, pelos momentos de ausência e, em especial, à minha namorada, que sempre me deu suporte e foi uma grande revisora dessa monografia.agradecimentos

À minha família, especialmente aos meus pais, pelo apoio financeiro e emocional que tornaram possível a realização deste curso e desta pesquisa. Sem vocês, nada disso seria possível.

À minha namorada, companheira de todas as horas, que não apenas me apoiou emocionalmente, mas também contribuiu significativamente como revisora deste trabalho, oferecendo sugestões valiosas e correções importantes.

Ao meu orientador, Professor Rodrigo Henrique Cunha Palácios, pela paciência, dedicação e conhecimento compartilhado ao longo do desenvolvimento desta monografia.

Aos professores do curso de Engenharia de Software, que contribuíram para minha formação acadêmica e profissional.

Aos colegas de turma, pelas trocas de experiências e momentos de aprendizado compartilhados.

Resumo

No decorrer da monografia será sugerido uma abordagem teórica, para otimização do ensalamento escolar. Através da combinação de algoritmos genéticos (AG) e regressão linear (RL). O processo de ensalamento é considerado um não polinomial, portanto um problema complexo e de difícil resolução (NP-difícil), envolvendo múltiplas restrições rígidas e flexíveis que impactam diretamente na qualidade da gestão educacional. A proposta é usar a regressão linear para prever e visualizar, com base em dados históricos e a demanda por alunos, salas e recursos; e, com essas informações, alimentar os algoritmos genéticos, para encontrar alocações otimizadas que minimizem conflitos de horário, superlotação, subutilização de salas e outros problemas que possam ser causados pela má otimização dos recursos. A metodologia implantada inclui uma revisão sistemática da literatura, a modelagem conceitual do sistema híbrido, definição de entradas e saídas, sugestão de parametrização da função de fitness e a simulação de cenários fictícios para ilustrar a aplicação do modelo. Espera-se que esse modelo reduza significativamente o tempo gasto com planejamento manual, aumente a eficiência no uso de recursos e seja adaptável a diferentes contextos — desde escolas públicas até universidades. A proposta também enfoca na transparência das decisões e a possibilidade de planejamento estratégico com base nas previsões.

Palavras-chave: Genetic algorithms, Classroom assignment, Timetabling problem, Combinatorial optimization, Academic scheduling, Linear regression, Hybrid algorithms;

LISTA DE Equações

[Equação 1 – Equação Geral da Regressão Linear](#_Toc195178567) 19

[Equação 2 – Taxa de Ocupação Média de Salas de Aula 24](#_Toc195178568)

LISTA DE ABREVIATURAS E SIGLAS

|  |  |
| --- | --- |
| IA | Inteligência Artifical |
| NP-difícil | Problemas Polinomiais difíceis |
| SGE | Sistema de Gestão Escolar |
| PSO | Particle Swarm Optimization |
| UTFPR | Universidade Tecnológica Federal do Paraná |
| AG | Algoritmos Genéticos |
| RL  STP | Regressão Linear  School Timetabling Problem |
| CAP | Classroom Assignment Problem |
|  |  |

**SUMÁRIO**

|  |  |
| --- | --- |
| 1. Introdução ...................................................................................................... | 13 |
| 1.1 Cronograma Objetivo Geral ...................................................................... | 14 |
| 1.2 Objetivos.................................................................................................... | 14 |
| 1.2.1 Objetivo Geral................................................................................... | 14 |
| 1.2.2 Objetivos Específicos....................................................................... | 15 |
| 2. Fundamentação Teórica ................................................................................ | 16 |
| 2.1 Problemas do Ensalamento Escolar ......................................................... | 16 |
| 2.2 Algoritmos Genéticos ................................................................................ | 17 |
| 2.3 Regressão Linear ..................................................................................... | 19 |
| 2.4 Modelagem Híbrida .................................................................................. | 21 |
| 3. Desenvolvimento ........................................................................................... | 25 |
| 3.1 Levantamento Bibliográfico ...................................................................... | 25 |
| 3.2 Construção do Arcabouço Teórico ........................................................... | 26 |
| 3.3 Proposição da Modelagem Conceitual ..................................................... | 26 |
| 3.4 Modelagem Abstrata do Algoritmo Híbrido ............................................... | 26 |
| 3.5 Estudo de Caso Fictício ............................................................................ | 26 |
| 3.6 Proposição de Métricas Teóricas de Avaliação ........................................ | 27 |
| 4. Resultados Esperados (Baseados na Literatura) .......................................... | 29 |
| 5. Discussão ...................................................................................................... | 32 |
| 5.1 Vantagens do Modelo Híbrido ................................................................. | 32 |
| 5.2 Limitações e Obstáculos .......................................................................... | 33 |
| 5.3 Questões Éticas e Educacionais ............................................................. | 33 |
| 6. Aplicações Práticas em Situações Hipotéticas .............................................. | 35 |
| 6.1 Possibilidades de Aplicação em Diferentes Contextos Educacionais ...... | 35 |
| 7. Conclusão ...................................................................................................... | 37 |
| 8. Referências Bibliográficas ............................................................................. | 39 |
| Anexo A – Lei nº 9.610, de 19 de fevereiro de 1998 ......................................... | 41 |

1. Introdução

O ensalamento escolar, procedimento delicado na gestão de instituições de ensino e aprendizagem, consiste em um problema complexo que envolve aspectos logísticos e pedagógicos, conforme demonstrado por Sena, Liberalino e Lima Júnior (2022) em sua abordagem híbrida para alocação de horários universitários. Afetando assim, imediatamente a qualidade e consistência do ensino, como no bem-estar de discentes e docentes. A proposta envolve a 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 manualmente (Elloumi et al. 2014) ou semiautomático, quase sempre baseado em planilhas digitais e um conjunto de normas estipuladas pela experiência dos superiores.

Todavia, a ascendente dinamização dos sistemas de educação – caracteriza-se por uma dimensão maior de alunos, transferências recorrentes entre as 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. Segundo Abduljabbar e Abdullah (2021), o problema de alocação de horários apresenta uma dificuldade combinatória do problema transcende a capacitação de uma liderança encontrar as melhores soluções. Enquanto verifica o monitoramento da instituição, isso resulta em salas superlotadas ou subutilizadas, divergências nos horários e desorganização na carga de trabalho de docentes e, em última instância, aproveitamento indevido de recursos físicos e humanos. Nesse dilema, o aprofundamento de métodos tecnológicos avançados, como a IA e pesquisa operacional (Ochoa e Chicano, 2015), na otimização do processo de ensalamento, não é apenas uma tendência.

Durante esta monografia será proposto 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. Analisar-se-á 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-se com um ambiente educacional mais propício ao aprendizado.

## Cronograma

|  |  |  |  |
| --- | --- | --- | --- |
| Junho | |  | | --- | |  |  |  | | --- | | Revisão do TCC 1 com orientador Definição dos dados a serem utilizados | |
| Julho | Coleta e organização de dados Construção do modelo preditivo (Regressão Linear) |
| Agosto | Desenvolvimento do algoritmo genético  Testes iniciais |
| Setembro | Integração da regressão linear com o algoritmo genético Testes de validação e ajuste de parâmetros |
| Outubro | Documentação dos resultados alcançado  Geração dos Gráficos e Tabelas |
| Novembro | Revisão do Documento de TCC |

**1.2 Objetivos**

**1.2.1 Objetivo Geral**

Propor e analisar teoricamente um modelo híbrido que utilize Algoritmos Genético e Regressão Linear para otimização de processos de ensalamento escolar, a fim de gerar uma alocação mais adaptável e justa dos recursos educacionais.

**1.2.2 Objetivos Específicos**

1.2.2.1 Realizar uma revisão bibliográfica de artigos e textos que tratam de métodos de otimização de ensalamento escolar, com foco principal em algoritmos genéticos e regressão linear;

1.2.2.2 Desenvolver um modelo híbrido em que a capacidade preditiva da RL é interaja junto do poder de otimização e busca dos AGs.

1.2.2.3 Analisar teoricamente a capacidade do sistema híbrido superar as abordagens convencionais e delimitar suas limitações.

1.2.2.4 Demonstrar como modelo proposto pode maximizar a eficiência no uso de recursos, otimizar a distribuição de resultados e mudar para um ambiente pedagógico mais propício ao aprendizado.

1.2.2.5 Avaliar teoricamente a viabilidade do modelo em resolver problemas reais.

2. Fundamentação teórica

## 2.1 Problemas 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, deste modo 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 variado 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:
  + - 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, logo 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:
  + - Preferência dos estudantes, evidenciada no estudo de Mahlous e Mahlous (2023).
    - 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;
    - Favorecer 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, 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. (ALVES; OLIVEIRA; ROCHA NETO, 2017)

## 2.2 Algoritmos Genéticos

Os Algoritmos Genéticos (AGs) constituem técnicas de busca e otimização inspiradas em processos biológicos, formalmente introduzidas por John Holland (1975).

Fundamentam-se na simulação de mecanismos da evolução natural e da genética, como seleção natural, cruzamento genético e mutação. Essas estratégias são especialmente indicadas para lidar com problemas complexos, como evidenciado por Goldberg (1989), em sua obra Genetic Algorithms in Search, Optimization, and Machine Learning, na qual categorizou esses algoritmos como adequados para espaços de busca extensos, não lineares e com múltiplos ótimos locais — situações em que abordagens tradicionais de otimização tendem a 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 ao qual sugere 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 um indivíduo (ou cromossomo) pode ser representado como um vetor, no qual cada gene (ou elemento do vetor) 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, o que contribui 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 se encontre 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, obtendo soluções de alta qualidade mesmo na ausência de um conhecimento profundo sobre a estrutura matemática do problema.



## 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 a maneira como a variável dependente varia em função das variáveis independentes, assim prevendo o seu comportamento e fornecendo, dessa forma, um resultado que pode ser utilizado em cenários futuros.

* No contexto educacional, a regressão linear — especialmente em sua forma múltipla — tem se mostrado uma ferramenta eficaz para atividades de análise e previsão, como demonstrado por Riandari et al. (2022) em estudos sobre a estimativa do número de estudantes com base em variáveis acadêmicas e institucionais.
* **Previsão de demanda:** Estimar o número futuro de matrículas por série, curso ou unidade escolar, com base em dados históricos, 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, a previsão do 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 conforme a equação 1:

* Conforme expressa a Equação 1, em que 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: recursos físicos e humanos,por exemplo o aumento na demanda por laboratórios de informática).
* 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 outras variáveis constantes.
* 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. A resultante da soma das vantagens é maior que a das desvantagens.

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, adaptando-se 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 sobre a 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, (AGBOLADE et al., 2024), e indústria, em que previsões alimentam algoritmos de otimização para planejamento de rotas, estoques ou escalonamentos (TALBI, 2009). 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.

1. Desenvolvimento

No decorrer desta monografia será desenvolvida 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:

* A obtenção das referências foi realizada por meio de uma busca criteriosa em bases de dados científicas de alta relevância — incluindo IEEE Xplore, ResearchGate, Scopus, Web of Science, ACM Digital Library e SciELO — complementada pela análise de livros técnicos e literatura acadêmica disponível 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", “scheduling 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, com foco direcionado principalmente a livros.
* 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.

## Modelagem Abstrata do Algoritmo Híbrido:

* Concepção de um modelo conceitual robusto integrando Algoritmos Genéticos e Regressão Linear 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.

## 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.

## 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, para se calculá-la utilizasse a fórmula descrita na Equação 2:

Equação 2:

Em que:

**Total de alunos alocados:** número total de estudantes que foram efetivamente distribuídos nas salas de aula.

**Somatório da capacidade das salas utilizadas nos horários de aula:** soma da lotação máxima permitida em todas as salas usadas no período analisado.

* 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.

1. 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. A otimização proporcionada por algoritmos genéticos no tratamento de múltiplas e complexas restrições do ensalamento, como demonstrado por Mahlous e Mahlous (2023) ao incluir preferências estudantis, pode levar a uma redução de episódios de superlotação e a um uso mais eficiente dos espaços, 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 de o 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, 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.

1. Discussões

Com base na modelagem proposta, foi elaborado um estudo de caso fictício para simular o funcionamento do modelo híbrido em um cenário educacional realista. Essa simulação não tem caráter experimental, mas busca ilustrar o potencial do sistema na resolução do problema de ensalamento escolar.

O cenário envolve uma instituição com 200 alunos distribuídos em diferentes disciplinas, cinco salas com capacidades variáveis (entre 30 e 50 estudantes), turnos matutino, vespertino e noturno, bem como professores com disponibilidade limitada e exigências pedagógicas diversas. A Regressão Linear foi aplicada teoricamente para prever a demanda por turmas e espaços, e os Algoritmos Genéticos foram configurados para considerar as principais restrições rígidas e flexíveis identificadas na fundamentação teórica.

Entre os resultados da simulação estão: uma alocação que respeita todas as restrições rígidas (sem conflitos de horário, sobrecarga de salas ou alocação simultânea de professores); uma taxa média de ocupação próxima a 90%; e um nível parcial de atendimento às preferências dos docentes, configurado como uma penalização leve na função de avaliação.

Esses resultados indicam que o modelo é capaz de produzir soluções teoricamente viáveis e adaptadas às exigências institucionais, demonstrando coerência entre a modelagem proposta e os objetivos de otimização do ensalamento. Embora não tenham sido utilizados dados reais, a aplicação conceitual sugere que o sistema possui potencial para redução de tempo de planejamento, melhoria na eficiência dos recursos e maior justiça na distribuição de turmas e espaços.

1. APLICAÇÕES PRÁTICAS EM SITUAÇÕES HIPOTÉTICAS

A transição de um modelo teórico para a aplicação prática em instituições de ensino envolve diversas variáveis técnicas, administrativas e sociais. Esta seção discute os principais fatores que condicionam ou dificultam a adoção do modelo híbrido proposto.

O primeiro ponto a ser considerado diz respeito à qualidade dos dados. A etapa de Regressão Linear depende de séries históricas consistentes, atualizadas e devidamente estruturadas. Muitas instituições, sobretudo da rede pública, enfrentam dificuldades na coleta e manutenção de dados acadêmicos e operacionais, o que representa um desafio inicial importante.

Além disso, a etapa de otimização via Algoritmos Genéticos requer capacidade computacional mínima e, sobretudo, conhecimento técnico especializado para configurar os parâmetros do algoritmo — como taxas de mutação, critérios de parada e estrutura da função de avaliação. Isso pode exigir capacitação das equipes ou parceria com setores de tecnologia da informação.

Outro fator relevante é a resistência institucional à inovação. A adoção de sistemas automatizados no lugar de processos manuais tradicionais pode gerar desconfiança por parte de gestores e professores. Por isso, a implementação do modelo deve ser acompanhada de ações de sensibilização, explicação dos critérios utilizados na alocação e oferta de ferramentas de auditoria e ajuste manual.

No aspecto técnico, é necessário desenvolver uma interface de integração com os Sistemas de Gestão Escolar (SGE), permitindo o intercâmbio de dados como matrículas, turmas, professores e infraestrutura. Essa etapa é crítica para a automatização completa e exige compatibilidade com as plataformas existentes.

Apesar dos desafios, o modelo apresenta alto potencial de aplicabilidade em redes de ensino com maior maturidade tecnológica e disponibilidade de dados, como universidades, institutos federais ou escolas privadas com sistemas informatizados robustos. A implantação piloto em contextos controlados pode fornecer os dados empíricos necessários para validar e calibrar o sistema antes de uma adoção em larga escala.

1. 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 (Ochoa e Chicano 2015).
* **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.

1. Referências Bibliográficas

ABDULJABBAR, I. A.; ABDULLAH, S. M. An evolutionary algorithm for solving academic courses timetable scheduling problem. *Baghdad Science Journal*, [S.l.], v. 19, n. 2, p. 399–408, 2021. DOI: <https://doi.org/10.21123/bsj.2022.19.2.0399>.

AGBOLADE, S. J.; AYINLA; ODENIYI, L. A.; AKINOLA, S. O. Optimisation of university examination timetable using hybridised genetic and greedy algorithms: a case study of Computer Science Department, University of Ibadan. *International Journal of Computer*, [S.l.], v. 51, n. 1, 2024. Disponível em: [https://ijcjournal.org/index.php/internationaljournalofcomputer/article/view/2788.](https://ijcjournal.org/index.php/internationaljournalofcomputer/article/view/2788.%20) Acesso em: 4 jun. 2025.

ALVES, S.; OLIVEIRA, S.; ROCHA NETO, A. R. A recursive genetic algorithm-based approach for educational timetabling problems. In: MELLO, R. F.; CARVALHO, A. C. P. L. F. (org.). *Designing with Computational Intelligence*. Cham: Springer, 2017. p. 161–175. (Studies in Computational Intelligence, v. 664). DOI: <https://doi.org/10.1007/978-3-319-44735-3_9>.

CHEN, S.-H.; HSIEH, S.-H. Using genetic algorithm methods to solve course scheduling problems. *Expert Systems with Applications*, [S.l.], v. 25, n. 1, p. 39–50, 2003. DOI: <https://doi.org/10.1016/S0957-4174(03)00015-9>.

CHEN, X. et al. Design and application of an improved genetic algorithm to a class scheduling system. *International Journal of Emerging Technologies in Learning*, [S.l.], v. 16, n. 1, p. 4–16, 2021. DOI: <https://doi.org/10.3991/ijet.v16i01.18225>.

DAVIS, L. *Handbook of Genetic Algorithms*. [S.l.]: Van Nostrand Reinhold, 1991.

ELLOUMI, A. et al. The classroom assignment problem: complexity, size reduction and heuristics. *Applied Soft Computing*, [S.l.], v. 14, part C, p. 677–686, 2014. DOI: <https://doi.org/10.1016/j.asoc.2013.09.003>.

GAREY, M. R.; JOHNSON, D. S. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. [S.l.]: W.H. Freeman, 1979.

GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. [S.l.]: Addison-Wesley, 1989.

HOLLAND, J. H. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press, 1975.

MAHLOUS, A. R.; MAHLOUS, H. Student timetabling genetic algorithm accounting for student preferences. *PeerJ Computer Science*, [S.l.], v. 9, e1200, 2023. DOI: <https://doi.org/10.7717/peerj-cs.1200>.

MITCHELL, M. *An Introduction to Genetic Algorithms*. Cambridge: MIT Press, 1996.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. *Introduction to Linear Regression Analysis*. [S.l.]: John Wiley & Sons, 2012.

OCHOA, G.; CHICANO, F. (eds.). *Evolutionary Computation in Combinatorial Optimization: 15th European Conference, EvoCOP 2015, Copenhagen, Denmark, April 8–10, 2015, Proceedings*. [S.l.]: Springer, 2015.

RIANDARI, F.; SIHOTANG, H. T.; HUSAIN, H. Forecasting the number of students in multiple linear regressions. *Matrik: Jurnal Manajemen Teknik Informatika dan Rekayasa Komputer*, [S.l.], v. 21, n. 2, p. 249–256, mar. 2022. DOI: <https://doi.org/10.30812/matrik.v21i2.1348>.

SENA, D. C.; LIBERALINO, C. H. P.; LIMA JÚNIOR, F. C. Avaliação de um algoritmo híbrido na geração de soluções para o problema de horários de cursos universitários. In: *Seminário Integrado de Software e Hardware – SEMISH*, 2022, evento online. *Anais [...].* Porto Alegre: SBC, 2022. DOI: <https://doi.org/10.5753/semish.2022.223191>. Acesso em: 2 jun. 2025.

SIVANANDAM, S. N.; DEEPA, S. N. *Introduction to Genetic Algorithms*. [S.l.]: Springer, 2008.

TALBI, E.-G. *Metaheuristics: From Design to Implementation*. Hoboken, NJ: John Wiley & Sons, 2009. ISBN: 9780470278581. DOI: <https://doi.org/10.1002/9780470496916>.