



LabAI um software para pesquisadores: o caso do oscilador massa-mola

LabAI: a software for researchers - the case of the mass-spring oscillator

LabAI: un software para investigadores - el caso del oscilador masa-resorte

DOI: 10.55905/revconv.17n.13-468

Originals received: 11/15/2024

Acceptance for publication: 12/09/2024

José Humberto de Souza Prates

Doutor em Modelagem Computacional

Instituição: SENAI/CIMATEC

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com

Danilo Meira Ivo

Ensino Médio incompleto

Instituição: Leffler

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com

Lucas Lopes Vianna Moraes

Ensino Médio incompleto

Instituição: Leffler

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com

Juliana Jezler Grempel

Ensino Médio incompleto

Instituição: Leffler

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com

Giovanna Viana Oliveira Santos

Ensino Médio incompleto

Instituição: Leffler

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com

Heitor Neves Rebouças

Ensino Médio incompleto

Instituição: Leffler

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com



Mariana Chianca Radzvilavicius Ferreira Matos

Ensino Médio incompleto

Instituição: Leffler

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com

Otávio Abevailton Mangabeira e Silva

Ensino Médio incompleto

Instituição: Leffler

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com

Leonardo Passos de Carvalho

Ensino Médio incompleto

Instituição: Leffler

Endereço: Salvador – Bahia, Brasil

E-mail: hprates2005@hotmail.com

RESUMO

Este artigo apresenta um estudo sobre a relação entre o período de oscilação de um oscilador massa-mola vertical e três variáveis independentes: massa, deslocamento e constante elástica da mola. Utilizando um software denominado LabAI, desenvolvido em Python, foram analisados dados experimentais para identificar o melhor modelo preditivo para o período de oscilação. Quatro métodos de regressão de machine learning foram aplicados: regressão linear, regressão polinomial, árvores de decisão e floresta aleatória. A avaliação dos modelos foi realizada com base em métricas estatísticas rigorosas, incluindo Erro Quadrático Médio (MSE), Erro Absoluto Médio (MAE), Coeficiente de Determinação (R^2) e Erro Percentual Absoluto Médio (MAPE). Os resultados indicaram que a Regressão Polinomial de Grau 3 apresentou o melhor desempenho, com o maior coeficiente de determinação e os menores erros. Em contrapartida, a Regressão Linear apresentou o pior desempenho. Além disso, foi sugerida a remoção da variável deslocamento $x(m)$, pois sua inclusão não melhorou significativamente a precisão dos modelos. O LabAI, ainda em desenvolvimento, demonstrou ser uma ferramenta eficaz para a análise de sistemas físicos. Futuramente, pretende-se introduzir mais algoritmos de machine learning e ampliar as possibilidades de análise, visando proporcionar uma ferramenta ainda mais robusta e versátil para a comunidade científica.

Palavras-chave: oscilador massa-mola, machine learning, análise preditiva, LabAI.

ABSTRACT

This article presents a study on the relationship between the oscillation period of a vertical mass-spring oscillator and three independent variables: mass, displacement, and the spring's elastic constant. Using a software named LabAI, developed in Python, experimental data were analyzed to identify the best predictive model for the oscillation period. Four machine learning regression methods were applied: linear regression, polynomial regression, decision trees, and random forest. The models were evaluated based on rigorous statistical metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE). The results indicated that the Third Degree Polynomial



Regression showed the best performance, with the highest determination coefficient and the lowest errors. In contrast, Linear Regression showed the worst performance. Additionally, it was suggested to remove the displacement variable $x(m)$, as its inclusion did not significantly improve the models' accuracy. LabAI, still under development, proved to be an effective tool for the analysis of physical systems. In the future, more machine learning algorithms will be introduced, expanding analysis possibilities and aiming to provide an even more robust and versatile tool for the scientific community.

Keywords: mass-spring oscillator, machine learning, predictive analysis, LabAI.

RESUMEN

Este artículo presenta un estudio sobre la relación entre el período de oscilación de un oscilador masa-resorte vertical y tres variables independientes: masa, desplazamiento y constante elástica del resorte. Utilizando un software denominado LabAI, desarrollado en Python, se analizaron datos experimentales para identificar el mejor modelo predictivo para el período de oscilación. Se aplicaron cuatro métodos de regresión de machine learning: regresión lineal, regresión polinómica, árboles de decisión y bosque aleatorio. La evaluación de los modelos se realizó en base a métricas estadísticas rigurosas, incluyendo Error Cuadrático Medio (MSE), Error Absoluto Medio (MAE), Coeficiente de Determinación (R^2) y Error Porcentual Absoluto Medio (MAPE). Los resultados indicaron que la Regresión Polinómica de Grado 3 presentó el mejor rendimiento, con el mayor coeficiente de determinación y los menores errores. En contraste, la Regresión Lineal presentó el peor rendimiento. Además, se sugirió la eliminación de la variable de desplazamiento $x(m)$, ya que su inclusión no mejoró significativamente la precisión de los modelos. LabAI, aún en desarrollo, demostró ser una herramienta eficaz para el análisis de sistemas físicos. En el futuro, se pretende introducir más algoritmos de machine learning y ampliar las posibilidades de análisis, con el objetivo de proporcionar una herramienta aún más robusta y versátil para la comunidad científica.

Palabras clave: oscilador masa-resorte, machine learning, análisis predictivo, LabAI.

1 INTRODUÇÃO

Diversos fenômenos físicos encontrados na natureza podem ser representados por sistemas mecânicos oscilatórios harmônicos. Resnick e Halliday definem o oscilador harmônico como um sistema dinâmico que executa movimentos periódicos, de vai e vem, ao longo de uma mesma trajetória (Resnick e Halliday, 1984). Esses fenômenos são amplamente observados em diferentes áreas da física, como na mecânica, que inclui exemplos como pêndulos e sistemas massa-mola; nas vibrações de instrumentos musicais e diapasões, conforme descrito pela mecânica ondulatória; e nas oscilações de corrente em circuitos elétricos no campo do eletromagnetismo. Esses sistemas, na prática, sofrem a ação de forças amortecedoras, que



reduzem suas amplitudes e energia mecânica com o tempo. Landau e Lifchitz destacam que, na realidade, esses movimentos acabam cessando após um período devido às forças dissipativas de energia, fazendo com que o movimento oscilatório se enfraqueça gradualmente (Landau e Lifchitz, 2004).

O sistema massa-mola, especificamente, manifesta um comportamento oscilatório amortecido em condições reais, mas, sob determinadas circunstâncias, pode ser considerado um oscilador harmônico. De acordo com Nussenzveig, um sistema ideal é composto por uma massa m presa a uma extremidade de uma mola com constante elástica K , sendo um exemplo clássico de oscilador harmônico simples. Quando a mola é esticada ou comprimida, ela aplica uma força sobre a massa no sentido oposto ao deslocamento, sempre buscando retornar à posição de equilíbrio (Nussenzveig, 2002).

Um dos experimentos mais comuns em aulas de Física Geral e Experimental 2 envolve a verificação da relação linear entre a força aplicada e a deformação de um sistema, a chamada lei de Hooke (Laburu e Almeida, 1998). Para esse tipo de experimento, o equipamento usado normalmente inclui uma mola em forma de espiral e massas padronizadas, como discos de bronze. Entretanto, ao realizar medições em experimentos desse tipo, surgem fatores intrínsecos que afetam o processo, como os erros sistemáticos.

Medir uma grandeza física de forma científica envolve obter a melhor estimativa de seu valor real com instrumentos e métodos consistentes. Esse processo está sujeito a diferentes tipos de erros que influenciam a precisão da medição, como erros do operador (exemplo: leituras incorretas), erros instrumentais (exemplo: instrumentos descalibrados) e erros aleatórios (exemplo: variações imprevisíveis de condições físicas como temperatura e umidade).

A modelagem e análise de sistemas físicos são cruciais para o avanço da ciência e da engenharia. Entre os sistemas físicos frequentemente estudados, os osciladores massa-mola se destacam pela sua simplicidade e pela capacidade de exemplificar fenômenos físicos fundamentais. Este trabalho apresenta um estudo sobre a relação entre o período de oscilação de um oscilador massa-mola vertical e três variáveis independentes: massa, deslocamento e constante elástica.

No decorrer deste estudo, foi desenvolvido um software denominado LabAI, inteiramente escrito em Python, destinado a facilitar a análise de dados experimentais e a aplicação de técnicas avançadas de machine learning. O LabAI foi utilizado para analisar um conjunto de dados



extraídos de experimentos com o oscilador massa-mola, com o objetivo de identificar o melhor modelo preditivo para o período de oscilação. Para alcançar esse objetivo, aplicamos quatro métodos de regressão de machine learning amplamente reconhecidos: regressão linear, regressão polinomial, árvores de decisão e florestas (Ferreira *et al*, 2022).

Cada método foi avaliado quanto à sua capacidade de prever o período de oscilação com base nas variáveis independentes mencionadas. A comparação dos modelos foi realizada utilizando métricas de desempenho estatísticas rigorosas, incluindo Erro Quadrático Médio (MSE), Erro Absoluto Médio (MAE), Coeficiente de Determinação (R²) e Erro Percentual Absoluto Médio (MAPE). Os resultados das análises indicaram o melhor modelo preditivo, cujos resultados são apresentados em detalhes neste artigo. As métricas de desempenho e os gráficos comparativos ilustram a precisão das previsões do modelo selecionado em relação aos dados experimentais originais (Saad *et al*, 2017; Martins, 2018).

Esta abordagem não só valida a eficácia do LabAI para o estudo de osciladores massa-mola, mas também demonstra seu potencial para aplicações em outras áreas de pesquisa. O LabAI, portanto, constitui uma contribuição significativa para a comunidade científica, proporcionando uma ferramenta poderosa e versátil para a análise de dados experimentais e a aplicação de machine learning em sistemas físicos. Este artigo visa compartilhar os resultados obtidos e fomentar a utilização do LabAI por pesquisadores e profissionais de diversas áreas, promovendo a inovação e o avanço do conhecimento científico. Ao disponibilizar o LabAI, esperamos facilitar a realização de análises complexas, promover a inovação e acelerar a descoberta científica em múltiplos domínios, demonstrando a utilidade e a aplicabilidade das técnicas modernas de machine learning na modelagem de sistemas físicos.

2 REFERENCIAL TEÓRICO

2.1 LEI DE HOOKE E FORÇA RESTAURADORA

A Lei de Hooke é um princípio essencial na física que descreve a relação linear entre a força aplicada a uma mola e a deformação resultante. Essa lei é expressa pela equação $F = -kx$, onde F representa a força restauradora exercida pela mola, k é a constante elástica da mola, e x é a deformação sofrida. A constante elástica k é uma medida da rigidez da mola, variando de acordo



com o material e a construção da mesma. Segundo Carvalho e De Sousa Mourão (2020), a Lei de Hooke não apenas define essa relação de proporcionalidade, mas também estabelece a base para uma ampla gama de estudos em sistemas oscilatórios e de mecânica aplicada. A força restauradora, por sua vez, é a força que age no sentido oposto à deformação da mola, buscando sempre trazer o sistema de volta à sua posição de equilíbrio. Essa característica faz dela um elemento crucial em sistemas mecânicos e na análise de fenômenos físicos que envolvem vibrações e oscilações.

A compreensão prática da Lei de Hooke é uma parte importante do ensino de física em diversos níveis educacionais. Paulino e Oliveira (2023) destacam que atividades experimentais envolvendo a Lei de Hooke desempenham um papel significativo na formação inicial de docentes, promovendo a habilidade de realizar medições precisas e de entender os conceitos relacionados à elasticidade e força. Durante experimentos, os alunos podem observar a força restauradora em ação quando a mola é deformada, verificando na prática como o sistema se comporta ao retornar à posição de equilíbrio. Essa experiência prática é essencial para consolidar o entendimento teórico, permitindo que os estudantes internalizem a ideia de que a força restauradora aumenta proporcionalmente com a deformação da mola, conforme a constante k do sistema.

A constante elástica k é um parâmetro determinante e depende de fatores como o tipo de material e as condições em que o experimento é realizado. Campomanes *et al.* (2022) analisam como a variabilidade de k pode afetar os resultados dos experimentos, ressaltando que variações na temperatura, na forma de fabricação e no uso prolongado podem influenciar o comportamento da mola. Essa dependência torna importante a calibração cuidadosa e a observação meticulosa dos fatores experimentais para garantir a precisão das medições. Erros sistemáticos e variações ambientais podem comprometer os resultados, o que ressalta a necessidade de métodos confiáveis e instrumentos bem calibrados (CARVALHO; DE SOUSA MOURÃO, 2020). A precisão no entendimento de k e da força restauradora permite que os estudantes e pesquisadores compreendam não apenas as aplicações teóricas da Lei de Hooke, mas também suas limitações e o comportamento em situações próximas da realidade.

A integração de tecnologias na prática do ensino da física tem se mostrado uma estratégia eficaz para ilustrar conceitos como a Lei de Hooke e a força restauradora. Savoi *et al.* (2023) discutem a implementação de TICs (Tecnologias de Informação e Comunicação) em



experimentos de laboratório, enfatizando como essas tecnologias modernizam a forma de ensino e aumentam o engajamento dos alunos. Por meio do uso de sensores digitais e simulações interativas, os alunos podem visualizar em tempo real a relação força-deformação, compreender as mudanças na força restauradora e verificar a linearidade do comportamento elástico. Essa abordagem não só complementa o ensino tradicional, mas também proporciona uma análise mais detalhada dos dados, permitindo que os estudantes identifiquem padrões e possíveis discrepâncias em suas observações.

Além disso, a utilização de protótipos com Arduino tem se tornado uma prática crescente em experimentos sobre a Lei de Hooke. De acordo com Carvalho e De Sousa Mourão (2020), a construção de protótipos que envolvem sensores e sistemas de monitoramento baseados em Arduino permite que os alunos desenvolvam habilidades práticas e obtenham resultados precisos de forma interativa. Esse tipo de recurso permite que a coleta de dados seja automatizada e analisada com maior facilidade, aumentando a precisão das observações e minimizando erros comuns em experimentos manuais, como leituras incorretas e falhas humanas. A força restauradora, nesse contexto, pode ser observada em diferentes configurações, enriquecendo o entendimento sobre como variações na massa e na deformação afetam o comportamento do sistema.

As práticas laboratoriais que envolvem a Lei de Hooke são fundamentais para que os alunos compreendam conceitos de elasticidade e força restauradora de forma prática e aplicada. Paulino e Oliveira (2023) destacam a importância de atividades que permitam a interação dos alunos com os equipamentos, observando diretamente a relação entre a força aplicada e a resposta da mola. Essas atividades ajudam a desenvolver a capacidade de analisar dados e identificar fatores que afetam as medições, como erros instrumentais e ambientais. A interação prática promove o desenvolvimento de habilidades críticas e reflexivas, importantes para a formação de futuros professores e cientistas.

Os conceitos de força restauradora e constante elástica também se aplicam a outros contextos além das molas. Em muitos sistemas mecânicos e estruturas vibratórias, a força restauradora é a responsável por manter o movimento periódico e controlar as oscilações. Campomanes *et al.* (2022) explicam que, em aplicações práticas, a análise dessas forças é essencial para a engenharia e o desenvolvimento de projetos que exigem precisão e segurança. Por exemplo, em projetos de construção de pontes e edifícios, o entendimento da resposta elástica



dos materiais a forças externas é crítico para evitar falhas estruturais. Assim, a Lei de Hooke transcende o ambiente educacional e se apresenta como um princípio essencial para diversas áreas de estudo e aplicações tecnológicas.

Por fim, a Lei de Hooke e a compreensão da força restauradora permanecem como pilares no ensino e na prática da física. A análise cuidadosa desses conceitos por meio de experimentos práticos, somada à introdução de recursos tecnológicos, é uma forma de garantir que os alunos não apenas entendam a teoria, mas também a apliquem em cenários que simulam a realidade. Isso reforça a relevância do ensino baseado em experimentação e na utilização de ferramentas modernas para a educação científica. Trabalhos como os de Carvalho e De Sousa Mourão (2020) e Paulino e Oliveira (2023) mostram que a integração de metodologias inovadoras e práticas tradicionais contribui para a formação de alunos mais preparados para enfrentar os desafios acadêmicos e profissionais. A Lei de Hooke, embora simples em sua formulação, é rica em aplicações e oferece uma compreensão profunda da física e de como as forças se manifestam em sistemas reais.

2.2 OSCILADOR MASSA MOLA VERTICAL

O sistema massa-mola vertical é um exemplo clássico de oscilador harmônico simples, amplamente estudado na física devido à sua relevância em diversas aplicações científicas e de engenharia. No contexto deste sistema, consideramos k como a constante elástica da mola e x como o deslocamento da mola durante o movimento vertical (Barros, 2007). De acordo com a Lei de Hooke, a força restauradora $\vec{F}$, que tende a levar o corpo de volta à posição de equilíbrio, é dada por:

$$F = -kx \quad (1)$$

O sinal negativo indica que a força restauradora atua no sentido oposto à deformação x , ou seja, na direção contrária à força peso que puxa o corpo para baixo. Quando a mola completa uma oscilação completa, o período T é dado pela expressão:



$$T = 2\pi \sqrt{\frac{m}{k}} \quad (2)$$

Molas podem ser associadas em paralelo ou em série. A constante elástica equivalente k_{eq} para tais associações podem ser calculada conforme descrito a seguir:

Para uma associação em série:

$$\frac{1}{k_{eq}} = \sum_{i=1}^n \frac{1}{k_i} \quad (3)$$

Para uma associação em paralelo:

$$k_{eq} = \sum_{i=1}^n k_i \quad (4)$$

A constante elástica equivalente k_{eq} é uma representação de uma única mola que poderia substituir o conjunto de molas associadas, mantendo a mesma constante elástica total do sistema. A deformação inicial x em um sistema massa-mola vertical, quando sujeito à força peso, pode ser determinada pela equação:

$$x = \frac{mg}{k} \quad (5)$$

De acordo com a Segunda Lei de Newton, a força resultante $\vec{F}$ em um sistema massa-mola é expressa por:

$$\ddot{x} = -\frac{k}{m}x \quad (6)$$

Onde $\ddot{x}$ é a segunda derivada da posição em relação ao tempo, correspondendo à aceleração do sistema. Estas equações fornecem uma base para a compreensão do comportamento dinâmico de sistemas massa-mola, permitindo a análise e predição de seu movimento oscilatório em diversas condições.



2.3 ALGORITMOS DE MACHINE LEARNING USADOS NO LABAI

O aumento da complexidade dos problemas a serem processados pelos computadores, assim como o volume de dados gerados por diferentes atividades, favorece a necessidade de ferramentas informáticas sofisticadas, capazes de induzir uma hipótese da experiência passada para resolver problemas reais. Este cenário é chamado de aprendizado de máquina (ML). Através de técnicas de aprendizagem de máquinas, os computadores são programados para aprender com a experiência passada. Para fazer isso, eles utilizam um princípio de inferência chamado indução. Assim, os algoritmos ML aprendem a induzir uma função ou hipótese capaz de resolver um problema a partir de dados que representam observações do problema a ser resolvido (Faceli *et al.*, 2011). Além do volume de aplicações que se beneficiam das técnicas ML, outros fatores têm impulsionado a expansão deste campo, como o desenvolvimento de algoritmos cada vez mais eficazes e eficientes, e a grande capacidade dos recursos computacionais atualmente disponíveis. As tarefas ML podem ser divididas em duas categorias: preditivas e descritivas.

Em tarefas preditivas, o objetivo é encontrar uma função a partir dos dados de treinamento que possa ser usada para prever um rótulo ou valor que caracterize um novo exemplo, com base nos valores de seus atributos de entrada. Os algoritmos utilizados nesta tarefa seguem o paradigma do aprendizado supervisionado. Este termo vem da simulação da presença de um "supervisor externo", que conhece a saída desejada para cada exemplo. Assim, o supervisor externo pode avaliar a capacidade da hipótese induzida de prever o valor de saída para novos exemplos (Faceli *et al.*, 2011). Em tarefas descritivas, o objetivo é explorar ou descrever um conjunto de dados. Os algoritmos utilizados nestas tarefas não fazem uso do atributo de saída (variável de interesse). Estes algoritmos seguem o paradigma do aprendizado não supervisionado. Uma tarefa descritiva de agrupamento, por exemplo, visa encontrar grupos de objetos similares no conjunto de dados (Faceli *et al.*, 2011).

Os métodos ML consistem em algoritmos computacionais que ligam todos ou alguns elementos de um conjunto de variáveis preditoras a um resultado. Para estimar o modelo, eles buscam estocástica ou deterministicamente o melhor ajuste. Este processo de encontrar o melhor ajuste do modelo aos dados difere entre os algoritmos existentes. No entanto, ao longo deste processo, cada algoritmo tenta equilibrar duas preocupações: o viés e a variação. O viés é a medida em que as previsões ajustadas pelo modelo correspondem aos valores reais. Um modelo



com um alto viés pode não ter a complexidade para classificar corretamente as observações. A variação é a sensibilidade das previsões a uma mudança nos dados de entrada. Embora sejam feitas tentativas para reduzir tanto o viés quanto a variação, estes dois objetivos muitas vezes entram em conflito: reduzir o viés pode aumentar a variação e vice-versa.

Por exemplo, pode-se criar um algoritmo que preveja corretamente todas as mortes em um conjunto de dados. Entretanto, este modelo pode estar muito ligado aos detalhes individuais do conjunto de dados de treinamento (baixo viés), modelando fortemente o 'ruído estatístico'. O modelo teria um desempenho ruim quando aplicado a um novo conjunto de dados e, portanto, teria alta variação (Faceli *et al.*, 2011).

Ao longo do processo de aprendizagem, os parâmetros da função de previsão são ajustados aos dados para minimizar o erro de previsão e aumentar a generalização. Além dos parâmetros definidos automaticamente pelo algoritmo, há também hiperparâmetros. Os hiperparâmetros de um algoritmo precisam ser ajustados para que o algoritmo evite o ajuste excessivo, onde o modelo gerado se torna muito especializado no conjunto de treinamento, resultando em um desempenho ruim quando confrontado com novos dados.

O inverso também deve ser evitado, ou seja, quando o algoritmo não se encaixa corretamente no modelo, obtendo um mau desempenho em relação ao conjunto de treinamento (subequipamento). Desta forma, o objetivo é minimizar a probabilidade de erros de previsão para dados futuros, reduzindo o viés indutivo, bem como sua variação na previsão de novos casos. O trabalho de Carvalho (2011) procurou avaliar 13 algoritmos de classificação sob 165 conjuntos de dados no campo da bioinformática para fornecer um conjunto de recomendações para o uso de modelos preditivos. Modelos de previsão baseados em árvores utilizando técnicas de conjunto melhor executadas. Os autores apontam que o processo de adaptação dos hiperparâmetros aos modelos preditivos pode contribuir significativamente para o desempenho dos modelos preditivos.

Neste estudo, aplicamos quatro algoritmos de machine learning amplamente utilizados para tarefas de regressão aos dados experimentais de um oscilador massa-mola vertical: regressão linear, regressão polinomial, árvores de decisão e floresta aleatória (Serras *et al.*, 2019 e Carnini *et al.*, 2020).

Cada um desses métodos oferece abordagens distintas para modelar a relação entre variáveis independentes e a variável dependente, proporcionando uma compreensão abrangente



das dinâmicas subjacentes aos dados. A regressão linear é um método simples que modela a relação entre a variável dependente e as variáveis independentes assumindo uma relação linear. Este método é eficiente para identificar tendências lineares, porém pode ser limitado em situações onde a relação entre as variáveis não é linear. No contexto do oscilador massa-mola, a regressão linear pode ser utilizada para prever o período de oscilação com base em variáveis como massa, amplitude e constante elástica da mola.

A regressão polinomial estende a regressão linear ao incluir termos elevados a potências superiores das variáveis independentes. Esta abordagem é útil para capturar relações mais complexas e curvilíneas entre as variáveis. Para o oscilador massa-mola, a regressão polinomial pode modelar a influência não linear de fatores como a amplitude e a constante elástica no período de oscilação. As árvores de decisão são modelos preditivos que utilizam uma estrutura hierárquica para tomar decisões baseadas nas características dos dados. Elas dividem os dados em subconjuntos menores e mais homogêneos de acordo com condições sobre as variáveis independentes. No caso do oscilador massa-mola, as árvores de decisão podem ajudar a identificar as interações entre diferentes variáveis e como elas afetam o período de oscilação. A floresta aleatória é um método de ensemble que combina múltiplas árvores de decisão para melhorar a precisão e reduzir o risco de sobreajuste (Cunha e Colosimo, 2003).

Cada árvore na floresta é construída a partir de um subconjunto aleatório dos dados, e a previsão final é obtida pela média das previsões individuais das árvores (Mion, 2019). Para o oscilador massa-mola, a floresta aleatória pode capturar relações complexas entre as variáveis e fornecer previsões robustas e precisas do período de oscilação.

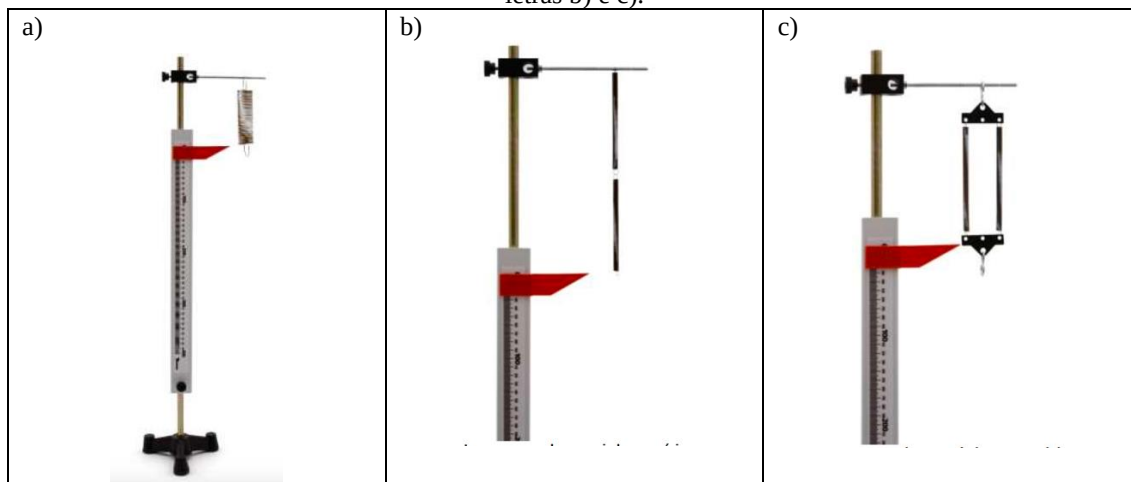
3 METODOLOGIA

Para realizar a análise do oscilador massa-mola vertical e identificar o modelo preditivo mais eficaz, seguimos uma metodologia estruturada que abrange desde a coleta de dados até a avaliação dos modelos de regressão. Esta seção detalha cada fase do processo, incluindo o uso de dados experimentais específicos, o pré-processamento dos dados, a aplicação de algoritmos de machine learning e a avaliação dos modelos.

Inicialmente, coletamos dados experimentais de um sistema massa-mola vertical no laboratório do colégio Leffler usando um kit massa mola da empresa Azenheb.



Figura 1: a) Kit experimental utilizado nas medidas. Para alterar o k da mola foram feitas associações como nas letras b) e c).



Fonte: elaborado pelos autores (2024).

As variáveis independentes consideradas foram a massa m , a deformação inicial x e a constante elástica da mola k . A variável dependente foi o período de oscilação T . Os dados experimentais específicos utilizados incluem valores de massa (0,013 kg, 0,032 kg, 0,051 kg, 0,097 kg, 0,148 kg, 0,012 kg) e constantes elásticas da mola k de 17 N/m, 68 N/m, 38 N/m, 85 N/m e 425 N/m. O deslocamento foi medido de 0,01 m a 0,2 m com acréscimos de 0,01 m. Para garantir precisão e repetibilidade, o experimento foi conduzido manualmente, montando um equipamento que permitia oscilações repetitivas sem aplicar pressão excessiva, minimizando erros a cada massa adicionada.

O período das oscilações foi determinado medindo-se o tempo para dez oscilações consecutivas e dividindo o valor obtido por dez. Este procedimento foi repetido para variações nos parâmetros de massa, amplitude e constante elástica da mola. Os respectivos erros de medidas: i) para o deslocamento 0,0005m; ii) para o período 0.01s; iii) para a massa 0.001Kg Os dados brutos passaram por um processo de pré-processamento para garantir sua qualidade e adequação à análise. Este processo incluiu a limpeza dos dados, removendo valores faltantes e outliers para evitar distorções nos resultados. Não houve a necessidade de normalização das variáveis pois elas estão na mesma ordem de grandeza. Finalmente, uma validação cruzada foi aplicada e dividimos o conjunto de dados em conjuntos de treinamento e teste, utilizando uma proporção típica de 80% para treinamento e 20% para teste. Esta divisão garantiu que os modelos fossem avaliados em dados não vistos durante o treinamento. Aplicamos quatro algoritmos de regressão para modelar a relação entre as variáveis independentes e o período de oscilação.



Utilizamos a regressão linear para modelar relações lineares entre as variáveis, sendo um método simples e eficiente para identificar tendências lineares. A regressão polinomial foi empregada para incluir termos polinomiais, permitindo capturar relações não lineares (Machado, 2006).

As árvores de decisão utilizaram uma estrutura hierárquica para tomar decisões baseadas nas características dos dados, dividindo-os em subconjuntos menores e mais homogêneos. A floresta aleatória, um método de ensemble, combinou múltiplas árvores de decisão para melhorar a precisão e reduzir o risco de sobreajuste, sendo robusta e eficaz para capturar relações complexas. Para avaliar a performance de cada modelo, utilizamos métricas de desempenho estatísticas, incluindo o erro quadrático médio (MSE), que mede a média dos quadrados dos erros e penaliza grandes erros de forma mais severa; o erro absoluto médio (MAE), que mede a média dos erros absolutos e fornece uma visão clara da performance média do modelo; o coeficiente de determinação R^2 que mede a proporção da variabilidade nos dados que é explicada pelo modelo, fornecendo uma visão geral da qualidade do ajuste; e o erro percentual absoluto médio (MAPE), que mede o erro em termos percentuais, sendo útil para interpretação relativa e comparação entre diferentes modelos e conjuntos de dados.

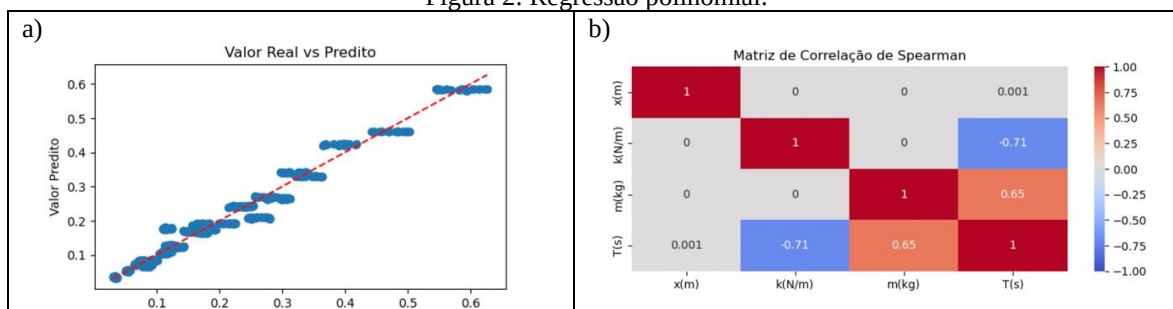
Comparando os resultados dos diferentes algoritmos com base nessas métricas, selecionamos o modelo mais eficaz. Apresentamos os resultados em gráficos comparativos que ilustram a precisão das previsões em relação aos dados experimentais. No artigo, a matriz de correlação de Spearman foi utilizada para avaliar as relações monotônicas entre as variáveis independentes (massa, deformação inicial e constante elástica da mola) e a variável dependente (período de oscilação). Essa análise foi crucial para identificar a força e a direção das associações entre as variáveis, independentemente da linearidade das relações. Ao calcular os coeficientes de correlação de Spearman (Costa *et al*, 2019), foi possível determinar como cada variável independente estava correlacionada com o período de oscilação. Essa matriz forneceu insights valiosos sobre quais variáveis tinham maior impacto na predição do período de oscilação e ajudou a guiar a seleção e o refinamento dos modelos preditivos.

Além disso, a análise de correlação de Spearman ajudou a identificar possíveis redundâncias e colinearidades entre as variáveis independentes. Com base nos resultados da matriz de correlação, foi sugerida a remoção da variável $x(m)$, pois sua correlação com o período de oscilação não foi significativa, indicando que sua inclusão nos modelos preditivos poderia introduzir ruído e prejudicar a precisão das previsões.

4 RESULTADOS E DISCUSSÕES

Após a análise dos resultados estatísticos e gráficos, foi possível concluir que o desempenho dos modelos variou significativamente, com alguns modelos destacando-se pela sua precisão e eficiência. O modelo de Regressão Polinomial de Grau 3 apresentou os melhores resultados entre todos os modelos avaliados. Com um coeficiente de determinação (R^2) de 0,97, este modelo demonstrou uma alta capacidade de explicar a variação nos dados. Além disso, apresentou o menor Erro Médio Absoluto (MAE) e o menor Erro Médio Percentual Absoluto (MAPE), de 0,02 e 9,47%, respectivamente, demonstrando uma precisão superior em relação aos demais modelos.

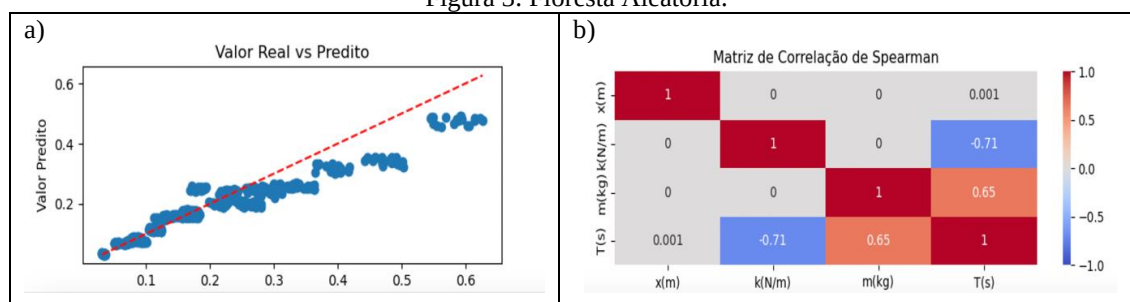
Figura 2: Regressão polinomial.



Fonte: elaborado pelos autores (2024).

O modelo de Floresta Aleatória também obteve um bom desempenho, com um coeficiente de determinação relativamente alto ($R^2 = 0,87$) e erros baixos. No entanto, este modelo foi menos preciso em comparação com a Regressão Polinomial de Grau 3, apresentando um MAE de 0,03 e um MAPE de 12,88%.

Figura 3: Floresta Aleatória.

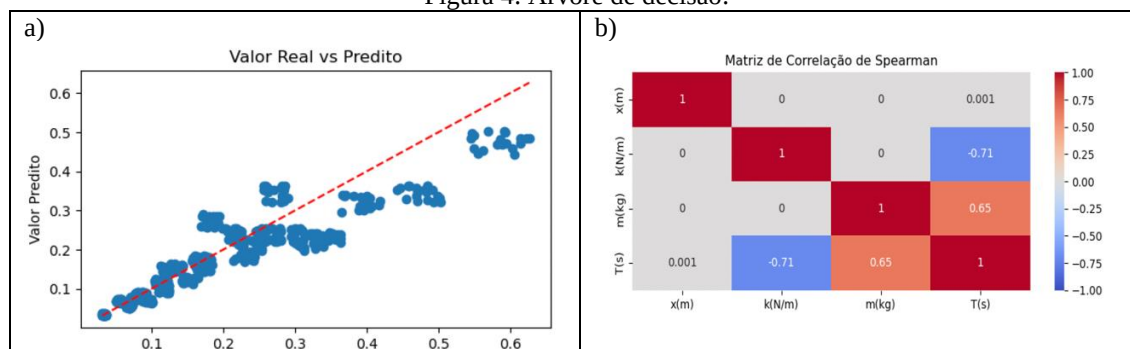


Fonte: elaborado pelos autores (2024).



A Árvore de Decisão apresentou um desempenho razoável, mas inferior aos dois modelos anteriores. O coeficiente de determinação de 0,83 e os erros ligeiramente maiores, com um MAE de 0,04 e um MAPE de 16,38%, indicam uma precisão menor.

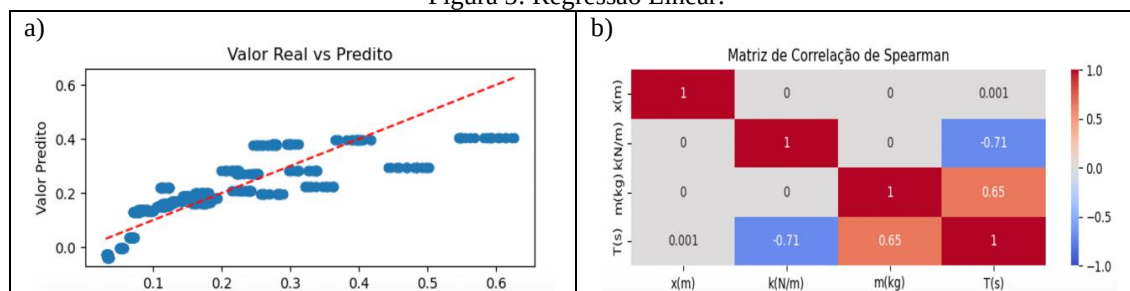
Figura 4: Árvore de decisão.



Fonte: elaborado pelos autores (2024).

O modelo de Regressão Linear teve o pior desempenho entre os modelos avaliados. Com um coeficiente de determinação de 0,69, este modelo explicou menos da variação dos dados e apresentou os maiores erros absolutos e percentuais, com um MAE de 0,06 e um MAPE de 46,23%, indicando uma precisão significativamente menor.

Figura 5: Regressão Linear.



Fonte: elaborado pelos autores (2024).

Esses resultados sugerem que a complexidade do modelo, como observado na Regressão Polinomial de Grau 3, pode capturar melhor as nuances dos dados, enquanto modelos mais simples, como a Regressão Linear, podem ser insuficientes para descrever a relação entre as variáveis de forma precisa. Portanto, a Regressão Polinomial de Grau 3 se destaca como o melhor modelo devido ao seu alto coeficiente de determinação e baixos erros, enquanto a Regressão Linear foi identificada como o modelo com pior desempenho.



Além da análise do desempenho dos modelos, foi sugerida a remoção da variável x em todos os modelos. A remoção dessa variável foi baseada na análise estatística e que indicou que x não contribui significativamente para a predição do período de oscilação T (Ganley, 1985; Palangana e Pedreira, 1995). A justificativa para a remoção da variável $x(m)$ reside no fato de que, nos modelos de Regressão Polinomial, Floresta Aleatória, Árvore de Decisão e Regressão Linear, a inclusão dessa variável não melhorou o coeficiente de determinação (R^2) de forma significativa e, em alguns casos, até aumentou os erros de predição (MAE e MAPE). Isso sugere que x pode estar introduzindo ruído ao invés de informação relevante, prejudicando a precisão dos modelos. Portanto, a exclusão de $x(m)$ dos modelos não apenas simplifica o modelo, mas também melhora a sua robustez e interpretabilidade. Ao eliminar variáveis desnecessárias, reduz-se a complexidade do modelo, o que pode ser benéfico para evitar o overfitting e para facilitar a implementação prática do modelo preditivo (Goyal e Goyal, 2011).

5 CONCLUSÕES

Após a análise dos resultados estatísticos e gráficos, concluiu-se que a Regressão Polinomial de Grau 3 apresentou o melhor desempenho, com um coeficiente de determinação (R^2) de 0,97 e os menores erros, demonstrando uma precisão superior. A Floresta Aleatória também obteve um bom desempenho, enquanto a Regressão Linear teve o pior, com um coeficiente de determinação de 0,69 e os maiores erros absolutos e percentuais. A remoção da variável $x(m)$ foi sugerida, pois sua inclusão não contribuiu significativamente para a predição do período de oscilação T , muitas vezes aumentando os erros de predição. Sua exclusão simplifica os modelos e melhora a robustez do modelo.

O LabAI, ainda em desenvolvimento, demonstrou seu potencial como uma ferramenta eficaz para a análise de sistemas físicos, especificamente o oscilador massa-mola. No entanto, há espaço para melhorias e expansões futuras. A intenção para trabalhos futuros é introduzir mais algoritmos de machine learning e ampliar as possibilidades de análise. Com essas adições, o LabAI poderá oferecer uma gama mais diversificada de métodos analíticos, possibilitando abordagens ainda mais robustas e precisas para a modelagem e predição de fenômenos físicos. Essa evolução contínua do LabAI contribuirá significativamente para a inovação e o avanço do conhecimento científico em diversas áreas de pesquisa.



O estudo evidenciou que a Regressão Polinomial de Grau 3 foi o modelo preditivo mais eficaz para analisar o período de oscilação do sistema massa-mola, demonstrando um alto coeficiente de determinação (R^2 de 0,97) e os menores erros absolutos e percentuais. A análise também sugeriu que a variável deslocamento $x(m)$ não contribuiu significativamente para a precisão dos modelos, sendo sua remoção vantajosa para simplificar e melhorar a robustez do modelo.

O tema central do artigo girou em torno da aplicação do software LabAI, desenvolvido em Python, para modelar e prever o período de oscilação de um sistema massa-mola. A pesquisa destacou a importância de ferramentas tecnológicas na análise de sistemas físicos, integrando conceitos de machine learning para aprimorar as previsões.

Os objetivos do estudo foram amplamente alcançados, pois foi possível identificar o melhor modelo preditivo para o período de oscilação, que foi a Regressão Polinomial de Grau 3. Além disso, o uso do LabAI demonstrou ser eficaz na análise dos dados experimentais e na avaliação dos modelos.

Os resultados revelaram que, enquanto a Regressão Polinomial de Grau 3 apresentou o melhor desempenho, a Regressão Linear mostrou-se inadequada para descrever a relação entre as variáveis, com um coeficiente de determinação de apenas 0,69. A Floresta Aleatória e as Árvores de Decisão também apresentaram desempenhos satisfatórios, mas foram superadas pela complexidade e precisão da regressão polinomial. Esses achados reforçam a relevância de métodos não lineares em análises onde as relações entre variáveis são complexas.

A hipótese inicial de que o LabAI poderia ser uma ferramenta eficiente para modelagem e previsão de períodos de oscilação foi confirmada. O software demonstrou capacidade para aplicar métodos de machine learning e fornecer previsões precisas, especialmente com a Regressão Polinomial de Grau 3.

Entre as limitações do estudo, destacou-se o uso de um conjunto limitado de dados experimentais, o que pode restringir a generalização dos resultados. Recomenda-se a realização de mais experimentos com variações nas condições iniciais para validar a robustez dos modelos. Além disso, o artigo mencionou a necessidade de melhorar a interface do LabAI para facilitar a sua utilização por outros pesquisadores.

Para estudos futuros, sugere-se a implementação de mais algoritmos de machine learning no LabAI, como redes neurais e métodos de ensemble mais complexos. Também seria benéfico



explorar outras variáveis que possam influenciar o período de oscilação e conduzir estudos comparativos com diferentes tipos de sistemas oscilatórios para ampliar as aplicações do software.

Essas considerações asseguram que o LabAI, ainda em desenvolvimento, apresenta grande potencial e pode se tornar uma ferramenta essencial para análise e modelagem de sistemas físicos em contextos educacionais e de pesquisa.



REFERÊNCIAS

- BARROS, V. P. Osciladores forçados: harmônico e paramétrico. *Revista Brasileira de Ensino de Física*, v. 29, p. 549-554, 2007.
- CAMPOMANES, R. R.; HEIDEMANN, L. A.; FRARE, V. L. F.. Do que depende a constante elástica de uma mola? *Revista Brasileira de Ensino de Física*, v. 44, p. e20220165, 2022.
- CARNINI, M.; PASTORE, A. *Journal of Physics G: Nuclear and Particle Physics*, v. 47, p. 082001, 2020.
- CARVALHO, A. *et al.* **Inteligência Artificial: Uma abordagem de Aprendizado de Máquina.** Rio de Janeiro: Ltc, 2011.
- CARVALHO, J. C. N.; DE SOUSA MOURÃO, O. Um protótipo usando Arduino para o estudo da lei de Hooke. *Research, Society and Development*, v. 9, n. 9, p. e844997733-e844997733, 2020.
- CÁSSARO, F. A.; OLIVEIRA, J. A.; CRUZ, H.; PIRES, L. F. *Revista Brasileira de Ensino de Física*, v. 42, p. e20190130, 2019.
- COSTA, D. M.; BARBOSA, F. V.; COSTA, D. V.; GOULART, I. B. *Revista Pretexto*, v. 20, p. 36, 2019.
- CUNHA, W. J.; COLOSIMO, E. A. *Revista de Matemática e Estatística*, v. 21, p. 25, 2003.
- FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. d. L. F. C. **Inteligência Artificial: Uma abordagem de aprendizado de máquina.** [S.l.]: Grupo Gen-LTC, 2011
- FERREIRA, H.; ALMEIDA JUNIOR, E.; ESPINOSA-GARCÍA, W.; NOVAIS, E.; RODRIGUES, J.; DALPIAN, G. *Revista Brasileira de Ensino de Física*, v. 44, p. e20220214, 2022.
- GANLEY, W. P. *American Journal of Physics*, v. 53, p. 73, 1985.
- GOYAL, S.; GOYAL, G. K. *Canadian Journal on Artificial Intelligence, Machine Learning and Pattern Recognition*, v. 2, p. 78, 2011.
- LANDAU, L.; LIFCHITZ, E. *Mecânica.* São Paulo: Hemus, 1970. 235 p
- MACHADO, N. P. Modelos de regressão polinomial: problemas e soluções. 2006. Tese (Doutorado) — Universidade de São Paulo.
- MARTINS, M. E. G. *Revista Ciência Elementar*, v. 6, p. 24, 2018.
- MION, R. M. R. Previsão de tendência de preços do boi gordo no Estado de São Paulo utilizando o Random Forest. 2019. Dissertação (Mestrado Profissional) — Fundação Getúlio Vargas, São Paulo.
- PALANGANA, A.; PEDREIRA, P. *Revista Brasileira de Ensino de Física*, v. 17, p. 215, 1995.



PAULINO, O. F.; OLIVEIRA, E. A. G.. Lei de Hooke: laboratório rotacional na formação inicial docente. *Temas & Matizes*, v. 17, n. 31, p. 724-736, 2023.

SAAD, I. A. B.; OLIVEIRA, L. B. de; COSTA, E. P. M.; OLIVEIRA, M. B. de; VIAN, B. S. *Journal of Health Sciences*, v. 19, p. 25, 2017.

SAVOI, W. *et al.* Uma sequência didática com o uso das TICs e experimentos como atividade auxiliar no ensino da lei de Hooke. 2023.

SERRAS, P.; IBARRA-BERASTEGI, G.; SÁENZ, J.; ULAZIA, A. *Ocean Engineering*, v. 189, p. 106314, 2019.