

**OTIMIZAÇÃO DE PROCESSOS ORIENTADO POR *MACHINE LEARNING*:**

**estudo de caso em uma granja sul mineira**

**MACHINE LEARNING-DRIVEN PROCESS OPTIMISATION:**

**a case study at a farm in southern Minas Gerais**

**Leonardo Romanelli Guimarães<sup>1</sup>, Rodrigo Franklin Frogeri<sup>2</sup>, Ana Amélia Furtado<sup>3</sup>**

<sup>1</sup> Centro Universitário do Sul de Minas - UNISMG, Varginha, MG,

[leonardo.guimaraes@alunos.unis.edu.br](mailto:leonardo.guimaraes@alunos.unis.edu.br); <https://orcid.org/0009-0007-2807-6241>

<sup>2</sup> Centro Universitário do Sul de Minas - UNISMG, Varginha, MG,

[rodrigo.frogeri@professor.unis.edu.br](mailto:rodrigo.frogeri@professor.unis.edu.br); <https://orcid.org/0009-0007-2807-6241>

<sup>3</sup> Centro Universitário do Sul de Minas - UNISMG, Varginha, MG,

[ana.furtado@professor.unis.edu.br](mailto:ana.furtado@professor.unis.edu.br); <https://orcid.org/0009-0006-9782-4736>

**RESUMO**

O presente trabalho apresenta um estudo de caso sobre a aplicação de Machine Learning (ML) na otimização da produção de ovos, vivenciado no contexto de uma granja sul mineira. Através desta pesquisa, busca-se analisar a viabilidade de otimizar os processos produtivos da granja utilizando técnicas preditivas de ML. Para isso, adotou-se uma metodologia quantitativa baseada em três etapas principais: coleta e tratamento de dados históricos da granja, construção de modelos preditivos utilizando o software WEKA, e avaliação do desempenho dos algoritmos aplicados por meio de métricas estatísticas como erro médio absoluto (MAE) e coeficiente de correlação (R). Entre os modelos testados, o Random Forest apresentou o melhor desempenho, com elevado grau de correlação e menor margem de erro, mostrando-se ideal para previsões operacionais. O M5P destacou-se pelo equilíbrio entre precisão e interpretabilidade, enquanto a regressão linear, embora mais simples, ofereceu resultados satisfatórios e de fácil compreensão. Acredita-se que os resultados deste relato possam contribuir para o debate e reflexões sobre a importância da tecnologia na gestão eficiente da avicultura e seu impacto em pequenas e médias granjas da região, otimizando o uso de recursos e aumentando a sustentabilidade.

**Palavras-chave:** Inteligência Artificial, Produção de Ovos, Processo, Avicultura.

# 1 INTRODUÇÃO

Nos últimos anos, o agronegócio brasileiro tem se transformado significativamente, impulsionado pelo avanço das tecnologias digitais e pela crescente demanda por eficiência e sustentabilidade na produção (Ang & Seng, 2021; Kamilaris & Prenafeta-Boldú, 2018). O país, reconhecido por sua liderança em várias cadeias produtivas, como soja, milho e aves, tem visto no uso de tecnologias avançadas, como o *Machine Learning* (ML), uma solução indispensável para enfrentar os desafios de maximização da produtividade e redução de custos operacionais (Ang & Seng, 2021). A aplicação de ML permite automatizar processos e utilizar dados em tempo real para tomadas de decisão estratégicas (Meshram et al., 2021).

Estudos demonstram que o uso de algoritmos de ML pode aumentar a eficiência de operações agropecuárias, reduzindo o desperdício de recursos e melhorando a sustentabilidade dos processos produtivos (Meshram et al., 2021; Sharma et al., 2021). Isso é particularmente relevante no contexto da produção de ovos, onde variáveis como nutrição, saúde das aves, temperatura e umidade precisam ser controladas de forma precisa para garantir a produtividade (Ang & Seng, 2021). Por meio da análise automatizada desses fatores, sistemas baseados em ML oferecem previsões precisas que permitem a adaptação do manejo com antecedência, evitando perdas inesperadas (Sharma et al., 2021).

Com a adoção de soluções tecnológicas, granjas que utilizam ML têm alcançado resultados superiores, tanto em termos de eficiência operacional quanto de sustentabilidade ambiental (Ang & Seng, 2021). Modelos de ML permitem o monitoramento contínuo e automatizado das condições dos galpões, identificando anomalias na produção e sugerindo ajustes imediatos na alimentação ou na ventilação, conforme necessário (Meshram et al., 2021; Sharma et al., 2021). Além disso, algoritmos preditivos ajudam a identificar padrões de doenças nas aves antes que se tornem um problema, garantindo um manejo mais eficiente e preservando a saúde do plantel (LIAKOS et al., 2018).

Outra vantagem significativa da aplicação de ML na produção de ovos é a possibilidade de prever a demanda e planejar a produção com maior precisão (Ang & Seng, 2021). Isso evita tanto a falta quanto o excesso de produção, contribuindo para uma gestão mais sustentável e econômica (Meshram et al., 2021; Sharma et al., 2021). Além disso, sistemas inteligentes ajudam a otimizar o uso de insumos ao longo do ciclo

produtivo, como ração e água, ajustando a oferta conforme a necessidade das aves (Meshram et al., 2021; Sharma et al., 2021).

Diante desse cenário, a presente pesquisa tem como objetivo analisar a viabilidade da previsão da produção de ovos em uma granja sul mineira por meio de técnicas de Machine Learning. Acredita-se que essa análise possa contribuir para os produtores locais e profissionais do setor avícola, já que a otimização da produção pode resultar em aumentos significativos de eficiência e rentabilidade, beneficiando toda a cadeia produtiva na região (Meshram et al., 2021; Sharma et al., 2021).

## **2 REFERENCIAL TEÓRICO**

### **2.1 Algoritmos de *Machine Learning* aplicados à otimização de processos**

*Machine Learning* (ML) tem se mostrado essencial na otimização de processos agrícolas e avícolas, permitindo maior precisão e eficiência (Khan et al., 2025; Meshram et al., 2021; Sharma et al., 2021). Esta seção analisa como algoritmos específicos de ML têm sido aplicados em diversos contextos produtivos, especialmente em granjas de ovos, visando o uso otimizado de recursos e aumento da produtividade.

O uso de *Machine Learning* em sistemas agrícolas permite otimizar o uso de recursos e prever resultados com alta precisão (Treboux & Genoud, 2018). Algoritmos supervisionados, como Regressão Linear e Random Forest, são comumente aplicados para prever o rendimento e identificar padrões nas variáveis que influenciam diretamente a produção, como temperatura e umidade (Ang & Seng, 2021).

Estudos demonstram que, ao empregar esses algoritmos, os produtores conseguem não apenas prever os resultados produtivos, mas também identificar rapidamente eventuais falhas no manejo. Isso é especialmente útil para granjas que operam com margens apertadas, onde a previsão precisa do consumo de ração e água pode evitar desperdícios e reduzir custos operacionais (Khan et al., 2025).

### **2.2 Estudo de Casos e Resultados na Avicultura**

Na produção de ovos, o uso de algoritmos como *Long Short-Term Memory* (LSTM) e Redes Neurais têm mostrado resultados promissores na previsão de picos produtivos e no ajuste dinâmico de condições ambientais (Meshram et al., 2021; Sharma

et al., 2021). Estes algoritmos permitem um controle mais eficaz do ciclo de produção, promovendo a sustentabilidade e garantindo que o consumo de insumos seja otimizado (Ang & Seng, 2021).

Outro estudo realizado por Khan et al. (2025) demonstrou que a integração de algoritmos preditivos com sistemas de IoT melhora significativamente a eficiência operacional. A utilização de sensores para monitorar variáveis em tempo real, como temperatura e iluminação, garante ajustes imediatos, reduzindo perdas produtivas e melhorando o bem-estar das aves.

Com base nas evidências apresentadas, conclui-se que a adoção de algoritmos de Machine Learning em granjas pode trazer benefícios operacionais expressivos, aumentando a eficiência e promovendo uma produção sustentável. No contexto da produção de ovos, tais algoritmos possibilitam a otimização do planejamento produtivo, permitindo que os gestores ajustem recursos e insumos com precisão, reduzindo custos e maximizando a rentabilidade (Khan et al., 2025).

### 3 MATERIAL E MÉTODOS

A abordagem metodológica do estudo é do tipo quantitativa, seguiu uma lógica dedutiva e foi do tipo estudo de caso. Três etapas principais compuseram a organização do estudo: coleta de dados, tratamento de dados e aplicação de técnicas de Machine Learning, seguidas de uma análise dos resultados obtidos.

A pesquisa focou na aplicação de técnicas de *Machine Learning* (ML) em um processo específico da granja de produção de ovos no Sul de Minas Gerais. A análise documental foi fundamental para identificar quais processos poderiam ser otimizados, e o software WEKA 3.8 (<https://ml.cms.waikato.ac.nz/weka>) foi adotado como ferramenta para construção e avaliação dos modelos de ML.

Na primeira etapa, foi realizada uma coleta de dados históricos relacionados aos processos produtivos da granja, com foco em variáveis que influenciam diretamente a produção de ovos. O conjunto de dados incluiu informações como semana do ano, idade das aves, total de aves por lote, número de óbitos semanais, taxa de mortalidade, consumo médio de ração por ave e a produção total de ovos por semana.

Esses dados foram obtidos exclusivamente a partir de registros internos da granja, os quais são atualizados semanalmente e consolidados em planilhas eletrônicas utilizadas para acompanhamento zootécnico. Todo o conteúdo foi estruturado em uma única planilha, que serviu como base para a análise exploratória, tratamento e aplicação dos modelos preditivos.

Na segunda etapa, os dados coletados passaram por um processo de tratamento, que envolveu a eliminação de valores ausentes e a normalização de variáveis para garantir a qualidade do conjunto de dados. A análise exploratória de dados (EDA) foi utilizada para identificar padrões e possíveis correlações entre as variáveis, o que auxiliou na definição dos processos que poderiam ser otimizados por técnicas de *Machine Learning*. Nesta fase, também foi realizada uma etapa de seleção de características (*feature selection*), visando identificar as variáveis mais relevantes para o modelo preditivo. Esse tratamento incluiu:

- **Eliminação de valores ausentes e inconsistências**, garantindo que apenas dados completos fossem utilizados na modelagem.
- **Normalização das variáveis**, ajustando as escalas dos dados para evitar distorções nos modelos.

Na terceira e última etapa, foram aplicadas técnicas de Machine Learning apropriadas para o cenário da granja. Modelos supervisionados e não supervisionados foram testados, como regressão linear, árvores de decisão e *clustering*, com o intuito de prever o consumo de recursos e otimizar o planejamento das atividades operacionais, como a alimentação dos animais e o controle climático. Os modelos foram treinados com uma parte dos dados coletados e testados com o restante, sendo avaliados por métricas como acurácia, precisão e erro médio absoluto (MAE).

A construção e avaliação dos modelos foram realizadas respeitando as seguintes configurações:

- **Regressão Linear**

Caminho: *Classify* → *functions* → *LinearRegression*

Test Mode: *Percentage split* – 80% treino / 20% teste

Atributos utilizados: Idade, TotAves, Mortalidade, MortSem,

Consumo\_Racao\_Ave

- **Árvore de Regressão (M5P)**

Caminho: *Classify* → *trees* → *M5P*

Test Mode: *Percentage split* – 80% treino / 20% teste

Atributos utilizados: SemAno, Idade, TotAves, Mortalidade, MortSem,

Consumo\_Racao\_Ave

Parâmetros adicionais: -M 4.0 (mínimo de instâncias por folha)

- **Random Forest:**

Caminho: *Classify* → *trees* → *RandomForest*

Test Mode: *Percentage split* – 80% treino / 20% teste

Atributos utilizados: SemAno, Idade, TotAves, Mortalidade, MortSem,

Consumo\_Racao\_Ave

Parâmetros adicionais: -I 100 (número de árvores), -P 100 (amostragem 100%), -K 0 ( $\sqrt{n}$  atributos por divisão), -S 1 (semente aleatória)

Ao final, os resultados foram analisados em termos de viabilidade e impacto no processo produtivo da granja. Foi avaliada a possibilidade de redução de custos

operacionais e aumento da eficiência dos processos com o uso de técnicas de *Machine Learning*, além das implicações práticas para a gestão da granja.

## **4 RESULTADOS E DISCUSSÃO**

### **4.1 Estatística Descritiva**

Nesta seção, apresentam-se as estatísticas descritivas das principais variáveis envolvidas no estudo, com o objetivo de oferecer uma visão geral do comportamento dos dados antes da aplicação dos modelos preditivos. São exibidas, para cada variável, medidas de tendência central (média) e de dispersão (desvio-padrão), bem como os valores mínimo e máximo observados. Esses indicadores permitem avaliar a heterogeneidade dos lotes e a variabilidade semanal da produção de ovos, idade das aves, mortalidade, mortes semanais e consumo de ração. A tabela 1 sintetiza essas informações:

Tabela 1. Estatísticas descritiva dos dados analisados

<b>Estatística Descritivas - Interpretação</b>						
<b>Variável</b>	<b>Média</b>	<b>Médiana</b>	<b>Desvio Padrão (DP)</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Comentário</b>
Idade (semanas)	90,42	89,00	11,13	74	120	As aves estão, em média, na fase produtiva madura.
Total de Aves	35.518	12.289	42.362	2.695	124.384	Alta variação entre os galpões
Mortalidade (%)	0,18	0,16	0,18	0,00	0,90	Mortalidade geralmente baixa, mas com picos.
Mortalidade Absoluta	128,44	37,00	215,88	2	1.878	Dados coerente com a variação de lotes.
Produção de ovos (semana)	191.604	65.372	256.457	0	782.439	Grande variação; alguns registros com 0 produção.
Consumo de Ração por Ave	0,11 kg	0,11 kg	0,07 kg	0 kg	0,99 kg	A média está dentro do esperado, mas alguns dados fora da curva.

Fonte: Desenvolvida pelos autores.

Durante a fase de estatística descritiva, observou-se que a produção média semanal de ovos foi de 191 604 unidades (DP = 256 457), com lotes variando de 0 a 782 439 ovos, o que evidenciou a grande dispersão natural do sistema. A idade média das aves foi de 90,4 semanas (DP = 11,1), enquanto o consumo médio de ração por ave situou-se em 0,11 kg/dia (DP = 0,07). A taxa média de mortalidade foi de 0,18 % (DP = 0,18), com óbitos semanais médios de 128 aves (DP = 216). Esses indicadores confirmam que os

dados representam condições produtivas maduras e heterogêneas, essenciais para o treinamento de modelos preditivos (Meshram et al., 2021; Sharma et al., 2021).

## 4.2 Regressão Linear

O modelo de Regressão Linear buscou ajustar uma função linear que relacionasse as variáveis zootécnicas à produção semanal de ovos. O modelo segue a seguinte estrutura:

Variável dependente: TotOvosSem – produção total de ovos na semana.

Variáveis independentes:

- Idade – idade média das aves em semanas.
- TotAves – número total de aves no lote.
- Mortalidade – taxa de mortalidade semanal, em %.
- MortSem – número absoluto de óbitos por semana.
- Consumo\_Racao\_Ave – consumo médio de ração por ave, em kg/dia.

A equação ajustada pelo software WEKA é a seguinte:

$$\text{TotOvosSem} = B_0 + B_1 \times \text{Idade} + B_2 \times \text{TotAves} + B_3 \times \text{Mortalidade} + B_4 \times \text{MortSem} + B_5 \times \text{Consumo\_Racao\_Ave}$$

Onde:

- $b_0$  é o intercepto (valor de TotOvosSem quando todas as variáveis independentes são zero)
- $b_1$  a  $b_5$  são os coeficientes de regressão, que indicam quanto a variável dependente (TotOvosSem) varia, em média, a cada unidade de aumento da respectiva variável independente, mantendo as demais constantes.

A seguir, os dados substituídos na equação:

$$\text{TotOvosSem} = 1460503,86 - 14683,18 \times \text{Idade} - 4,5726 \times \text{TotAves} - 84711,35 \times \text{Mortalidade} - 482,42 \times \text{MortSem} - 234453,05 \times \text{Consumo\_Racao\_Ave}$$

Ao aplicar o modelo de Regressão Linear, verificou-se um coeficiente de correlação de 0,9461, com erro médio absoluto (MAE) de 53.601 ovos e erro relativo de

24,89 %. Em comparação com expectativas iniciais de erro acima de 30 %, esse desempenho foi superior, mostrando que mesmo um modelo simples consegue capturar bem a tendência central dos dados. Coeficientes negativos foram obtidos para idade, mortalidade e consumo de ração, indicando que o aumento dessas variáveis tende a reduzir a produção semanal de ovos, em conformidade com os achados de Khan et al. (2025).

### **4.3 Árvore de Regressão (M5P)**

O M5P (árvore de regressão) elevou o coeficiente de correlação para 0,9544, reduzindo o MAE para 37 688 ovos (-29,7 % em relação à regressão linear) e o erro relativo para 17,50 %. Esta melhoria indica que a partição do espaço de dados e os modelos lineares locais capturaram padrões regionais que o modelo global não distinguia. Uma tendência observada foi que, em lotes maiores (Total aves > 23 393), a idade passou a ser o fator mais determinante, sugerindo transição de influência conforme o porte do plantel, comportamento semelhante ao relatado por Meshram et al. (2021) e Sharma et al. (2021), que destacam a eficácia de árvores de regressão em cenários agrícolas.

### **4.4 Random Forest**

O Random Forest, por sua vez, alcançou o melhor desempenho, com correlação de 0,9869, MAE de 18 319 ovos e erro relativo absoluto de 8,51%. Esse resultado demonstra a eficácia do ensemble em reduzir o erro médio ao combinar diversas árvores de regressão, proporcionando previsões mais robustas frente à variabilidade dos dados, conforme sugerido por Meshram et al. (2021) e Sharma et al. (2021).

Como se observa na Tabela 2, o Random Forest superou amplamente os demais modelos, alcançando  $R = 0,9869$  e  $MAE = 18\,319$  ovos, o que representa um ganho de precisão de 66 % em comparação à regressão linear e de 51 % frente ao M5P. Esse desempenho superior se deve à capacidade do ensemble de árvores de capturar tanto interações complexas entre variáveis quanto efeitos de outliers em lotes atípicos, consolidando previsões robustas mesmo sob elevada variabilidade.

Tabela 2. Comparativo de Desempenho dos Modelos

<b>Comparativo de Desempenho dos Modelos</b>			
<b>Modelo</b>	<b>Correlação</b>	<b>MAE (ovos)</b>	<b>Erro Rel.</b>
Regressão Linear	0,9461	53 601	24,89%
Árvore de Regressão (M5P)	0,9544	37 688	17,50%
Random Forest	0,9869	18 319	8,51%

Fonte: Desenvolvido pelos autores

O M5P, por sua vez, apresentou excelente equilíbrio entre complexidade e interpretabilidade: com  $R = 0,9544$ , reduziu o MAE para 37 688 ovos, evidenciando que o particionamento em subgrupos e ajustes lineares locais são eficazes para modelar diferentes regimes de produção (por exemplo, variações de porte de lote e faixas etárias). Seu uso pode ser recomendado em cenários onde se busca compreender padrões regionais específicos sem perder completamente a simplicidade de um modelo linear.

Já a Regressão Linear, embora tenha menor precisão ( $R = 0,9461$ ; MAE = 53 601 ovos), oferece a cristalina vantagem de fornecer coeficientes diretos que quantificam o impacto individual de variáveis como idade, mortalidade e consumo de ração. Essa transparência é valiosa para diagnosticar rapidamente quais fatores estão mais pressionando a produção, servindo como ponto de partida para intervenções gerenciais imediatas.

Em termos práticos, a adoção do Random Forest pode ser priorizada quando o objetivo for maximizar a acurácia de previsões semanais, apoiando decisões como ajuste fino de alimentação e climatização. O M5P surge como alternativa para análises aprofundadas de segmentação de lotes e identificação de pontos críticos em faixas específicas de produção. E a Regressão Linear deve ser mantida como ferramenta de diagnóstico rápido, facilitando interpretações diretas e comunicação eficiente junto a equipes de campo.

Portanto, a combinação hierárquica desses modelos, diagnóstico inicial por regressão, segmentação exploratória com M5P e previsão operacional robusta com Random Forest, constitui um arcabouço analítico completo, alinhado às melhores práticas de Machine Learning no agronegócio (Meshram et al., 2021; Sharma et al., 2021; Treboux & Genoud, 2018).

## 5 CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo analisar a viabilidade da aplicação de técnicas de Machine Learning (ML) na otimização dos processos produtivos de uma granja de ovos localizada no Sul de Minas Gerais. Através da coleta, organização e modelagem de dados históricos da produção, foi possível explorar o potencial de modelos preditivos como ferramenta de apoio à tomada de decisão no ambiente avícola.

Os principais achados indicam que a aplicação de ML pode contribuir significativamente para o planejamento zootécnico, tornando as estimativas de produção mais confiáveis e apoiando ajustes antecipados no manejo. Além disso, a experiência prática com diferentes algoritmos demonstrou a importância de se considerar tanto a acurácia quanto a interpretabilidade dos modelos, permitindo que produtores e gestores compreendam os fatores que mais impactam a produção.

Durante a realização do trabalho, algumas limitações foram identificadas. A ausência de variáveis ambientais, como temperatura, luminosidade e umidade, que também influenciam a produção de ovos, limitou o alcance preditivo dos modelos. Além disso, a base de dados, embora real e relevante, estava restrita a um único ambiente produtivo, o que pode reduzir a generalização dos resultados para outras granjas com realidades distintas.

Sugere-se que futuras pesquisas explorem a integração de sensores ambientais, dados meteorológicos e informações de mercado, ampliando a complexidade e aplicabilidade dos modelos preditivos. Também se recomenda o estudo da implementação prática desses sistemas em tempo real, avaliando não apenas sua precisão, mas também os impactos econômicos e operacionais na rotina das granjas.

## ABSTRACT

*This study presents a case analysis of the application of Machine Learning (ML) to optimize egg production on a poultry farm in southern Minas Gerais, Brazil. The research investigates the feasibility of improving production processes through predictive ML techniques. A quantitative methodology was employed, consisting of three main stages: collecting and preprocessing the farm's historical data, developing predictive models using the WEKA software, and evaluating algorithm performance using statistical metrics such as Mean Absolute Error (MAE) and the correlation coefficient (R). Among the models tested, the Random Forest algorithm achieved the best performance, demonstrating a high correlation and a low error margin—making it particularly suitable for operational forecasting. The M5P model was notable for its balance between accuracy and interpretability, while linear regression, though simpler, delivered satisfactory and easily understandable results. The findings of this study may contribute to broader discussions about the role of technology in the efficient management of poultry farming. Additionally, the results have the potential to benefit small and medium-sized farms in the region by optimizing resource use and promoting greater sustainability.*

**Keywords:** *Machine Learning, egg production, predictive modeling, poultry farm; process optimization.*

## REFERÊNCIAS

- Ang, K. L.-M., & Seng, J. K. P. (2021). Big Data and Machine Learning With Hyperspectral Information in Agriculture. *IEEE Access*, *9*, 36699–36718. <https://doi.org/10.1109/ACCESS.2021.3051196>
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*, 70–90. <https://doi.org/10.1016/j.compag.2018.02.016>
- Khan, S. A., Shukla, A. K., Yadav, S. K., & Vishwakarma, G. K. (2025). Machine learning models for analysis and prediction of optimal egg production. *Quality & Quantity*. <https://doi.org/10.1007/s11135-025-02309-1>
- Meshram, V., Patil, K., Meshram, V., Hanchate, D., & Ramkteke, S. D. (2021). Machine learning in agriculture domain: A state-of-art survey. *Artificial Intelligence in the Life Sciences*, *1*, 100010. <https://doi.org/10.1016/j.aills.2021.100010>
- Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2021). Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access*, *9*, 4843–4873. <https://doi.org/10.1109/ACCESS.2020.3048415>
- Treboux, J., & Genoud, D. (2018). Improved Machine Learning Methodology for High Precision Agriculture. *2018 Global Internet of Things Summit (GIoTS)*, 1–6. <https://doi.org/10.1109/GIOTS.2018.8534558>