

APLICAÇÃO DE APRENDIZADO DE MÁQUINA PARA AUMENTO DE PRECISÃO DE UM SISTEMA AUTOMATIZADO DE NUTRIÇÃO DE SUÍNOS

APPLICATION OF MACHINE LEARNING TO INCREASE THE PRECISION OF AN AUTOMATED SWINE NUTRITION SYSTEM

Giuliano Dal Bo* E-mail: g_dalbo@hotmail.com

Leandro Luis Corso* E-mail: lcorso@ucs.br

*Universidade de Caxias do Sul (UCS), Caxias do Sul, Rio Grande do Sul, Brasil.

Resumo: Com a constante evolução computacional, a viabilidade de implementação de modelos de aprendizado de máquina vem se desenvolvendo em diversas áreas de atuação. Uma das áreas de grande relevância para o desenvolvimento econômico nacional é o agronegócio, e dentre suas subáreas, a suinocultura apresenta parcela representativa para o mercado. Dentre as rubricas de custo da suino cultura, a que tem maior representatividade e necessidade de controle preciso é o consumo de ração. Visando o aumento da precisão de um sistema automatizado de nutrição de suínos, este artigo tem como objetivo, demonstrar uma composição de modelo supervisionado de inteligência artificial para previsão das variáveis mais relevantes para atingir o menor erro possível entre a dosagem programa e a dosagem realizada pelo equipamento. O modelo de aprendizado de máquina considerado é de regressão linear. Para verificação da acuracidade e performance do modelo são utilizadas as principais métricas de erro de análises quantitativas. Os resultados apresentados no artigo indicam que o modelo de aprendizado de máquina do tipo regressão linear consegue uma melhor generalização da variável a ser prevista e pode reduzir o erro absoluto da dosagem de cada ciclo de trato dos animais em média 7,4 vezes.

Palavras-chave: Aprendizado de Máquina. Inteligência Artificial. Rede Neural Artificial. Regressão Linear. Nutrição Suína.

Abstract: With the constant computational evolution, the feasibility of implementing machine learning models has been developing in several areas of activity. One of the areas of great relevance for national economic development is agribusiness, and among its sub-areas, pig farming has a representative share of the market. Among the swine production cost items, the one with the greatest representation and the need for precise control is feed consumption. Aiming at increasing the accuracy of an automated swine nutrition system, this article aims to demonstrate a composition of supervised artificial intelligence model to predict the most relevant variables to achieve the smallest possible error between the programmed dosage and the dosage performed by the equipment. The considered machine learning model is linear regression. To verify the accuracy and performance of the models, the main error metrics of quantitative analysis are used. The results presented in the paper indicate the linear regression machine learning model achieves a better generalization of predicted variable and can reduce the absolute error of the dosage of each animal treatment, in average, by 7,4 times.

Keywords: Machine Learning. Artificial Intelligence. Artificial Neural Network. Linear Regression. SwineNutrition

1 INTRODUÇÃO

Segundo Martins et al. (2013, p. 106), a produção mundial de suínos tem aumento relevante para atendimento da demanda por fontes de proteína animal da população nacional, sendo que novas técnicas de manejo e procedimentos que tenham como objetivo a otimização da suinocultura são cada vez mais pesquisados e implementados nas granjas de cooperativas agrícolas.

De acordo com o último Censo Agro, realizado pelo IBGE (Instituto Brasileiro de Geografia e Estatística) (2017), o Brasil possui um rebanho de suínos com 39.346.192 cabeças, distribuídas em 1.471.270 estabelecimentos. Conforme divulgado pela Embrapa (Empresa Brasileira de Pesquisa Agropecuária) (2003), a média de representatividade da ração no custo total de produção de suínos no Brasil é de 65%, podendo chegar a valores até 75% em períodos de crise do setor ou outros fatores econômicos.

Para aumento da taxa de produtividade das granjas, o modelo de nutrição com alimentação restrita é usualmente utilizado e recomendado pelas cooperativas agrícolas. No sistema de alimentação com restrição, um ou mais nutrientes são fornecidos na quantidade ou proporção não suficiente para permitir máximo ganho de peso (EMBRAPA, 2003). Em complemento a afirmação da EMBRAPA em relação a alimentação restrita, ela não considera o ganho de peso máximo como um ganho com qualidade, ou seja, a restrição alimentar gera um ganho de peso com um menor desenvolvimento de gordura na carne do animal e que proporciona maior qualidade da produção suína. A alimentação restrita necessita seguir rigorosamente a dosagem prescrita de acordo com a idade do animal e relacionada ao seu ciclo de crescimento. De acordo com Marcato e Lima (2005, p. 855), a restrição alimentar, com seu correto e preciso controle, pode reduzir o consumo diário de ração em 8,5% e ainda auxiliar na redução de poluentes presentes nos excrementos, representado redução dos custos de produção, bem como maior apelo em relação a sustentabilidade.

Conforme os dados oriundos do IBGE, Embrapa, bem como as informações referentes a metodologia de alimentação restrita, corroboram com necessidade de desenvolvimento de novas tecnologias dedicadas à suino cultura. A taxa constante

da evolução computacional e dos adventos provenientes do fomento da Indústria 4.0, inclusive no meio do agronegócio, vem de encontro com essa demanda disruptiva e possibilitam o emprego da inteligência artificial e do aprendizado de máquina para otimização do processo de nutrição de suínos, de uma maneira automatizada e com maior controle de perdas, reduzindo a intervenção humana no ciclo produtivo. Segundo Muhuri, Shukla e Abraham (2019, p. 218), o termo Indústria 4.0 não se refere somente a indústria de manufatura em si, mas uma transformação abrangente da forma de produzir, utilizando a integração digital e um conceito de engenharia inteligente.

Este artigo tem como objetivo apresentar o desenvolvimento de um modelo de aprendizado de máquina com capacidade de realizar a previsão da dosagem realizada por um equipamento automatizado e aplicar as devidas correções, estas necessárias para minimizar o erro entre o valor programado e o valor efetivamente realizado pela máquina. Sendo o objetivo de maior precisão atingido, isto apresenta grande valia para o desenvolvimento ótimo e com qualidade de cada lote de animais, bem como a redução do respectivo custo com consumo de ração e maior apelo sustentável para o processo produtivo da suinocultura.

2 REFERENCIAL TEÓRICO

2.1 Abrangência da Indústria 4.0 no desenvolvimento tecnológico

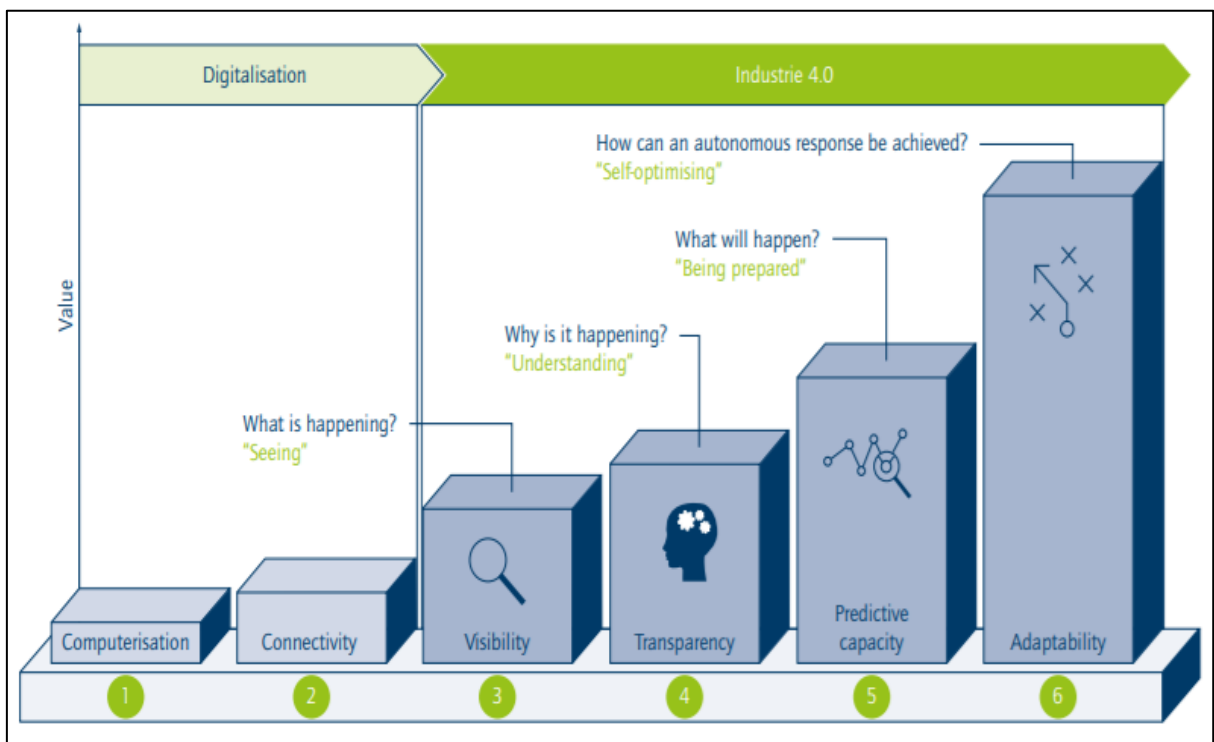
A quarta revolução industrial está chegando e está avançando rápido, dia após dia (MUHURI, SHUKLA e ABRAHAM, 2019, p. 218). Conforme Kagermann et al. (2011, p.2), a indústria 4.0 realiza uma junção entre os mundos real e virtual, com grande ênfase no desenvolvimento de diversas aplicações de engenharia, onde a conectividade e a assistência a tomadas de decisão de forma descentralizada, são necessidades indispensáveis na atualidade.

O tema Indústria 4.0 e conseqüentemente o desenvolvimento tecnológico que o assunto carrega consigo estão em considerável evidência atualmente. De acordo com Nakagawa et al. (2021, p. 106562), o aumento da popularidade do termo “Indústria 4.0” é evidenciado em uma pesquisa no Google Acadêmico, que retorna mais de 95.000 publicações sobre o assunto, a partir do ano de 2015.

Com esse crescimento exponencial do alto apelo do termo e a ideologia de Indústria 4.0, ou também chamada quarta revolução industrial, o desenvolvimento da inteligência artificial e de técnicas de aprendizado de máquina têm seu crescimento correlacionado positivamente. Segundo Shukla (2018, p. 29), inteligência artificial e técnicas de aprendizado de máquina estão se desenvolvendo de uma forma sofisticada, com uma grande taxa de crescimento, alavancando investimentos nos diversos setores onde são aplicados e atribuindo diversas melhorias nos processos industriais ou processos produtivos.

Conforme Schuh et al. (2017), o processo de transformação de qualquer instituição que deseje se beneficiar da evolução trazida pela Indústria 4.0, necessita considerar uma jornada de desenvolvimento contínua e ir de encontro ao equilíbrio, entre os custos aplicáveis a cada processo e seus respectivos benefícios. Para implementação de modelos de inteligência artificial e aprendizado de máquina, todos os estágios evolutivos ilustrados pela Figura 1 necessitam da verificação e da confirmação de um grau de maturidade elevado nos estágios prévios, para que o estágio de autonomia do conjunto apresente resultados efetivos.

Figura 1 – Estágios de desenvolvimento aplicados à indústria 4.0



Fonte: SCHUH et al. (2017).

A aplicação dos conceitos da Indústria 4.0 não estão sendo aplicados somente nos setores industriais de manufatura, mas também no setor do agronegócio. Conforme Newell, Newman e Mendly-Zambo (2021, p. 1), a quarta revolução agrícola, diretamente interligada aos preceitos da Indústria 4.0, resulta em tecnologias que podem, de forma significativa, apoiar esforços direcionados a segurança alimentar e sustentabilidade ambiental, incluindo o âmbito de produção de proteína animal. Segundo Spanaki (2020, p. 1), a evolução da forma de coleta dos dados dos processos e a criação de tecnologias disruptivas, mudaram o setor do agronegócio, do convencional e altamente dependente de ações humanas, para abordagens mais inteligentes, seguindo as novas tecnologias de análise baseadas em dados e computação de alto desempenho.

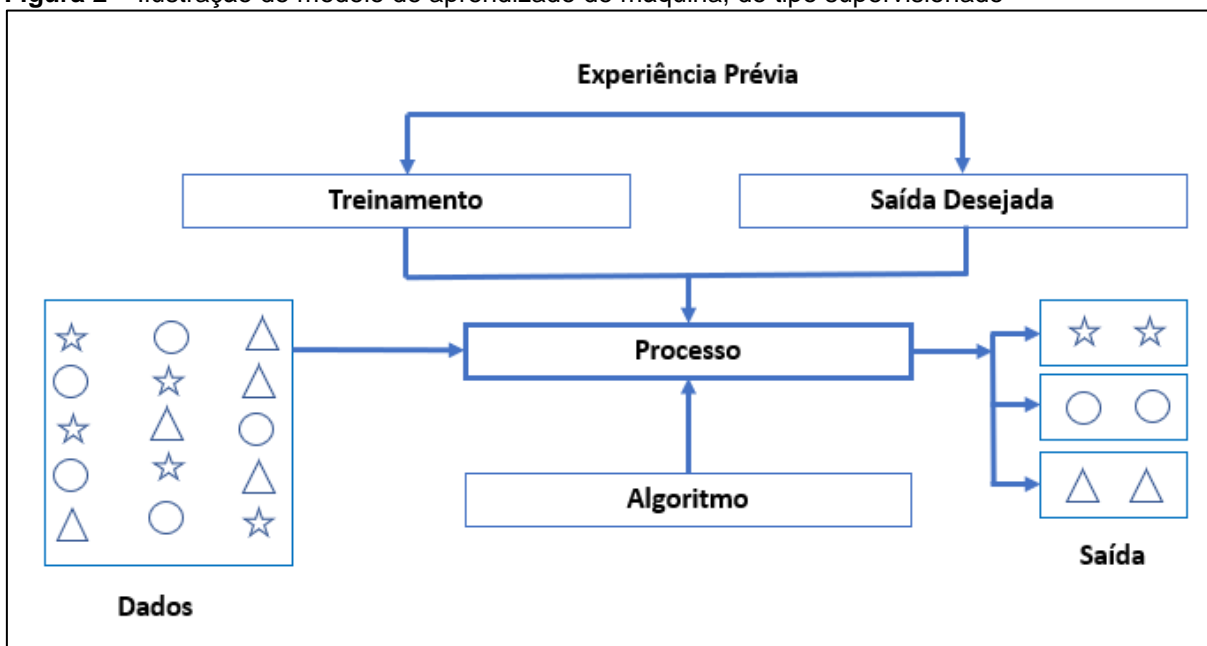
2.2 Aprendizado de Máquina

Conforme Jordan e Mitchell (2019, p. 255), o termo denominado como aprendizado de máquina, refere-se a construir e implementar um código computacional que inclui um modelo matemático com a capacidade de melhorar seu comportamento e resultados de forma automática, baseado em um banco de dados coletados, ou seja, nas experiências adquiridas pela máquina durante o seu funcionamento.

As técnicas de aprendizado de máquina geraram impactos sociais em uma gama de aplicações, como, por exemplo, em visão computacional, processamento de fala, compreensão de linguagem natural, neurociência, saúde e Internet das coisas (ZHOU et al., 2017, p. 350).

Conforme Basha, Bharathi e Venusamy (2021, p. 12067), modelos de aprendizado de máquina com a denominação e classificação de supervisionados, objetos de estudo deste artigo, utilizam os pacotes de dados reais coletados e organizados pela máquina, os empregam para os processos de treinamento e teste do modelo matemático e apresentam a capacidade de realizar uma ou mais previsões de um determinado dado ou conjunto de dados de saída, de acordo com os respectivos dados de entrada pré-determinados, conforme a ilustração apresentada na Figura 2.

Figura 2 – Ilustração de modelo de aprendizado de máquina, do tipo supervisionado



Fonte: Adaptado de BASHA, BHARATHI E VENUSAMY (2021).

Segundo Bonsignorio et al. (2020, p. 20), é necessário considerar um compartilhamento das formas de supervisão e controle dos processos de aprendizagem de máquina, com intuito de superar as limitações de um sistema cognitivo criado artificialmente. Esse compartilhamento e integração da supervisão aos processos de aprendizagem necessitam ser suficientemente conhecidos e detalhados para que o processo decisório do modelo seja satisfatório.

A importância de estudos e do desenvolvimento de modelos de aprendizado de máquina, pode ser visto no artigo publicado por Beroza, Segou e Mousavi (2021, p. 1). O trabalho dos pesquisadores mostra o desenvolvimento de um modelo de aprendizado de máquina com a funcionalidade de realizar a previsão de eventos sísmicos com base em um banco de dados sismológicos. Estas previsões em tempo e espaço são importantes para realizar, caso necessário, procedimentos de evacuação de determinadas localidades próximas aos epicentro ou outras ações mitigadoras de riscos atribuídos a essa natureza de fenômeno de movimentação tectônica terrestre.

Em contrapartida às vantagens e as possibilidades de desenvolvimento comentadas a respeito de modelos de aprendizado de máquina, Bury e Delaere (2021, p. 1), expõem que a principal dificuldade dos métodos desenvolvidos é

garantir que o que foi considerado como processo de aprendizado do modelo apresente a capacidade de generalizar e aplicar a tomada de decisão da maneira correta e esperada em relação a saída. A capacidade de generalização do modelo de aprendizado de máquina tem como objetivo principal manter versatilidade do modelo em relação a tomada de decisão, mesmo com diferentes tipos de comportamento associados aos dados utilizados para sua construção.

Em relação aos avanços obtidos com o desenvolvimento das técnicas de aprendizado de máquina, especificamente aplicadas a suinocultura, pode-se citar os trabalhos realizados por He et al. (2021, p. 106085) e Lee et al. (2019, p. 161716). Ambos os artigos são trabalhos de desenvolvimento de modelos de aprendizado de máquina com a capacidade de realizar a predição de ganho de peso de suínos durante sua fase de crescimento e desenvolvimento no ciclo produtivo da granja. Ambos os trabalhos publicados obtiveram ganhos de precisão da predição do ganho de peso em função das variáveis de entrada definidas e em relação aos métodos estatísticos convencionalmente empregados para inferência desta natureza de variável.

2.3 Modelo de Regressão Linear

De acordo com Ighalo, Adeniyi e Marques (2020, p. 1286), um modelo de regressão linear pode ser classificado como um modelo de aprendizado de máquina, com necessidade de ser supervisionado, e com capacidade de realizar uma predição com base nas variáveis independentes que apresentam características para gerar influência na variável dependente, ou a variável a ser prevista. Conforme Paialunga (2022), um modelo de regressão linear pode ser classificado como o modelo mais acessível e imediato de aprendizado de máquina, que pode ser um algoritmo simples de ser programado, com baixo custo computacional e utilizado com percentuais baixos de erro em modelos com comportamentos lineares ou até mesmo polinomiais.

O método de regressão linear é efetivamente a construção de uma reta com a finalidade de aproximar, com o melhor ajuste possível, o comportamento dos dados existentes para realizar uma previsão. Conforme Brownlee (2020), o processo de

aprendizado de um modelo de regressão linear tem como objetivo estimar os valores dos coeficientes empregados na representação da reta, com os dados que foram disponibilizados para realizar esta aproximação. As condições iniciais, algumas obrigatórias e outras recomendadas, são importantes para que um modelo de regressão linear apresente resultados satisfatórios:

- a) A variável dependente necessita ser contínua em seu domínio (condição obrigatória);
- b) As variáveis, dependente e independentes necessitam ser do tipo quantitativo (condição obrigatória);
- c) A relação entre a variável dependente e independentes necessitam apresentar linearidade em seu comportamento (condição recomendada);
- d) O índice de correlação (Pearson) entre as variáveis independentes necessita, preferencialmente, ser baixo (condição recomendada).

Conforme Amouri, Alapathy e Morgera (2020, p. 461), um modelo de regressão linear pode esclarecer a dependência entre a variável dependente Y_i e a variável independente X_i e pode ser descrita pela Equação(1).

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1)$$

Onde, β_0 representa o intercepto da reta no eixo da variável dependente, β_1 os coeficientes que representam a inclinação da reta e ε_i os seus respectivos resíduos.

Segundo Basha, Bharathi e Venusamy (2021, p. 12067), um algoritmo de regressão linear pode ser extremamente versátil, com base na simplicidade da suposição que todas as restrições do modelo apresentem um comportamento com tendência de linearidade. Conforme Swaminathan (2018), a predição de um modelo de regressão linear é proveniente do somatório de uma parte determinística e outra parte estocástica. A parte determinística de um modelo de regressão linear é coberta pela variável preditora ou independente e a parte estocástica demonstra que o valor esperado ou variável dependente e observado é também classificada como “imprevisível”, sempre faltando algumas informações para completar corretamente o

modelo e atingir o acerto da previsão com erro zero. Esta informação faltante pode ser obtida do termo residual do modelo.

2.4 Métricas de Erro

De acordo com Srivastava (2019), a construção de modelos de aprendizado de máquina tem seu funcionamento com base em um princípio de *feedback* construtivo, ou seja, a partir da construção do algoritmo são obtidas as métricas de erro, que tem a função de demonstrar e auxiliar na avaliação do desempenho do modelo e a partir disso realizar as ações de melhoria contínua do processo. Oriunda da explicação que devesse julgar a adaptação de qualquer medida de erro utilizada pelo resultado gerado da precisão das previsões fora da amostra do modelo de previsão, assim uma medida de erro para ser selecionada precisa aliar-se a finalidade da predição e os objetivos dos decisores (MAKRIDAKIS, 1993).

O *Mean Absolute Deviation* (MAD), do inglês ou desvio médio absoluto em português, mensura a acuracidade global e gera um indicativo de disseminação global em que todos os erros são dados pesos iguais (LAW, 2000).

O *Mean Absolute Percentual Error* (MAPE), do inglês ou erro percentual médio em português, fornece os resultados em porcentagem, o que facilita a interpretação pelos decisores, bem como gera um indicativo qualitativo da precisão do modelo de previsão testado (FLORES, 1986; MAKRIDAKIS, 1993).

O MAD e o MAPE podem ser representados pelas Equação 2 e 3, a seguir. Onde R representa o valor real, D representa o valor da previsão e n representa a quantidade de observações.

$$MAD = \frac{\sum |R-D|}{n} \quad (2)$$

$$MAPE = \sum \frac{|R-D/R|}{n} \cdot 100 \quad (3)$$

3 METODOLOGIA

Este artigo adota como referência para desenvolvimento do seu método, uma abordagem quantitativa dos dados disponíveis com o intuito de fundamentar a decisão para sua aplicação em campo. Tem como objetivo explorar os dados

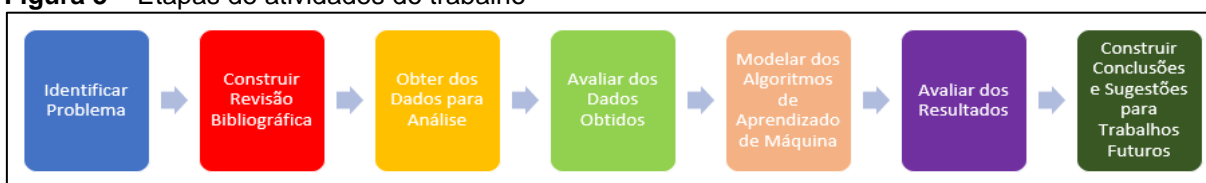
disponíveis para desenvolver um estudo de caso para avaliação da melhora de desempenho em relação a precisão de dosagem de ração de um sistema de alimentação de suínos automatizado, realizando uma verificação de desempenho do modelo de aprendizado de máquina denominado como regressão linear.

Segundo Gil (2002), o estudo de caso tem sua fundamentação no profundo estudo de uma ou de poucas variáveis, para que seja possível detalhamento aprofundado e amplo do objeto de trabalho.

3.1 Sequenciamento de Atividades do Estudo

O desenvolvimento do trabalho foi dividido em sete etapas, conforme demonstrado na Figura 4. Iniciando pela identificação e mapeamento do problema de precisão, construção da base teórica por meio da revisão bibliográfica, obtenção dos dados das máquinas para realização da análise e preparação dos dados, modelagem dos algoritmos de aprendizado de máquina, avaliação dos resultados obtidos, até as respectivas conclusões do estudo de caso, incluindo um comparativo para verificação dos respectivos ganhos na precisão do processo de tratamento de suínos.

Figura 3 – Etapas de atividades do trabalho



Fonte:Elaborado pelo autor (2021).

3.2 Obtenção dos Dados

A obtenção dos dados da pesquisa se dá diretamente do fabricante do equipamento para alimentação automatizada de suínos. Durante os ciclos de tratamento, denominados como tratos, o equipamento coleta os dados por meio de sensores instalados (células de carga e encoder incremental) e definições programadas na interface homem-máquina (definições de dosagem, quantidade de

ciclos por dia, horário de cada ciclo de tratamento, especificidades de cada recinto onde são alocados um conjunto de suínos a ser alimentado etc.). Todos os dados do processo são armazenados em seu controlador lógico programável, os quais podem ser acessados diretamente no equipamento via sua interface homem-máquina ou são armazenados em um servidor dedicado a esta função. Os dados inseridos no servidor ficam disponibilizados para uso, tanto do fabricante, como para monitoramento do processo pelo produtor e podem ser acessados e filtrados via *software* especificamente desenvolvido para esta função pelo fabricante. Por meio deste monitoramento o produtor pode verificar a acuracidade do seu equipamento e utilizar os dados para construção de seus indicadores de gerenciamento da produção.

Em função da Lei Geral de Proteção de Dados (LGPD), os dados das curvas de tratamento aplicadas, que são foco de estudo de veterinários e fisiologistas especializados em suínos e resultado de uma série de testes, sendo estes de propriedade de cada cooperativa agrícola produtora de suínos, não serão demonstrados neste trabalho. Serão demonstradas apenas as métricas de erro que explicitem os ganhos de acuracidade das dosagens programadas do modelo de aprendizado de máquina em relação ao modelo atualmente empregado no equipamento.

3.3 Dados Obtidos para Análise

Os dados para análise e implementação dos modelos de aprendizado de máquina foram obtidos a partir de três equipamentos, já em operação, e com acesso de forma remota do servidor ao dedicado ao armazenamento dos dados de cada ciclo de tratamento.

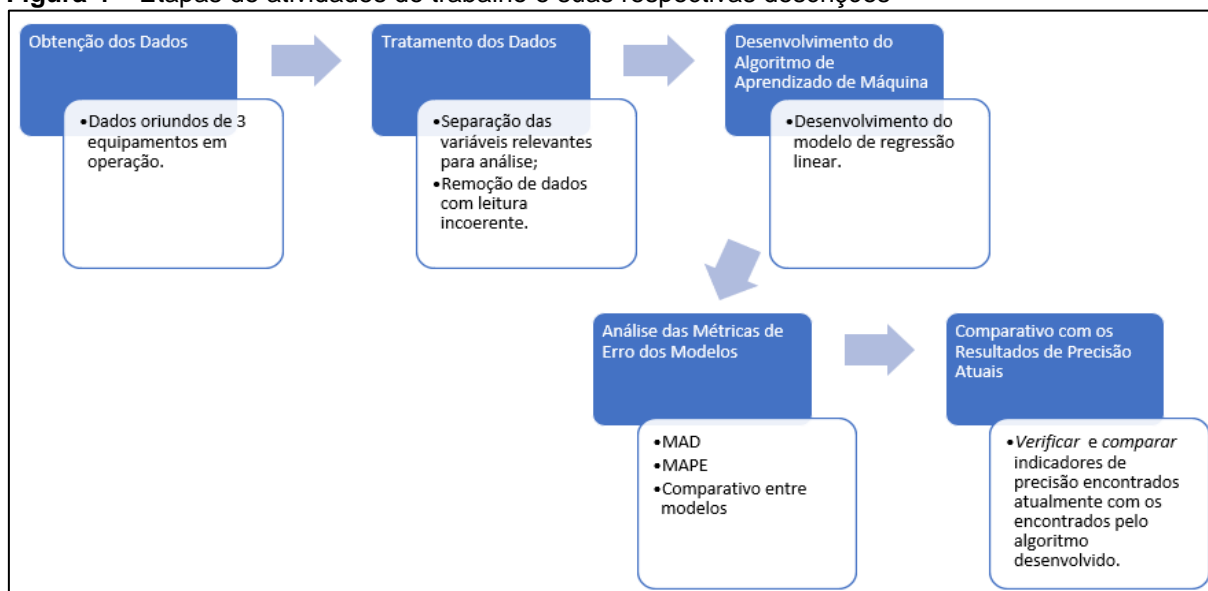
Dentre uma série de dados que são armazenados pelo equipamento, apenas os que foram definidos na etapa de análise e tratamento dos dados e descritos como diretamente influentes para o aspecto de precisão foram considerados para o estudo, sendo definidos como as variáveis independentes e o alvo (variável dependente) do trabalho. As variáveis consideradas para elaboração dos modelos matemáticos podem ser listadas como:

- a) Gramas por rotação (GPR): variável que representa a relação entre massa e volume da ração dosado pelo equipamento. Variável que apresenta grande volatilidade em função do comportamento discreto do material sólido e sua variação de densidade de acordo com algumas condições de contorno não consideradas na modelagem matemática como: temperatura, umidade, tempo de armazenagem no silo, tipo de ração etc. Esta variável é o alvo da previsão a ser realizada pelo modelo matemático e descrita como dependente. Por meio desta variável é possível determinar o número de rotações aplicadas à rosca de transferência;
- b) Valor total em gramas por animal (GPA): variável que representa o que foi disponibilizado em massa de ração para cada animal, para cada ciclo de trato. Valor obtido pela verificação via células de carga instaladas na base do silo do equipamento, no início e no final do ciclo de tratamento e pela quantidade de animais alojados na granja. Calculada com base no número de rotações da rosca de transferência e número de animais da granja. Variável dependente a ser comparada com o valor programado em gramas para cada animal e utilizada para comparativo da precisão entre o modelo atual utilizado pelo equipamento e o modelo de aprendizado de máquina;
- c) Gramas programado por animal (GP): variável programada previamente a cada início de ciclo de trato com base na curva de nutrição definida por cada cooperativa agrícola. Usualmente apresenta um crescimento linear de acordo com o período de engorda dos animais, desde o início do processo até a sua comercialização. Descrita como variável independente;
- d) Número de rotações (NR): variável do modelo descrita como dependente e diretamente atuante na rosca de transferência do equipamento. Determinada em função do valor de gramas por rotação (GPR) previsto pelo modelo de aprendizado de máquina;
- e) Número de animais (NA): quantidade total de animais a serem tratados em cada ciclo de tratamento. Variável independente que é previamente programada por meio da interface homem-máquina do equipamento.

3.4 Etapas de Desenvolvimento dos Modelos de Aprendizado de Máquina

Para desenvolvimento dos modelos de aprendizado de máquina, com foco na melhoria da precisão de dosagem do alimentador de suínos automatizado, as etapas de obtenção dos dados de três equipamentos, tratamento dos dados com a separação das variáveis utilizadas na análise e remoção de leituras inconsistentes, modelagem do modelo de regressão linear, implementação das métricas de erro para avaliação do desempenho de modelo matemático de regressão e por fim, o comparativo dos resultados obtidos com o modelo de aprendizado de máquina em relação ao modelo atualmente empregado nos equipamentos, conforme mostrado na Figura 4.

Figura 4 – Etapas de atividades do trabalho e suas respectivas descrições



Fonte: Elaborado pelo autor (2021).

Da etapa de tratamento dos dados até a etapa de comparativo dos resultados, o desenvolvimento será realizado utilizando a programação e ferramentas dos pacotes presentes no *software* R.

4 RESULTADOS E DISCUSSÕES

Para construção dos modelos de aprendizado de máquina, foram considerados 50 ciclos completos de tratamento, sendo estes separados em dois

conjuntos de dados: dados de treino e dados de teste e com uma distribuição entre dados respectivamente definida em 70% e 30% em relação a quantidade total de ciclos de tratamento e mantendo a sua sequência real de ocorrência, conforme a coleta realizada.

O pacote de dados de treino é utilizado para que o modelo matemático aprenda e realize o processo de regressão linear e os dados de teste são utilizados para verificar a capacidade preditiva calculada em relação aos dados reais coletados dos equipamentos.

O modelo de aprendizado de máquina, do tipo regressão linear, foi aplicado a três equipamentos, todos eles já em operação, utilizando os dados obtidos e detalhados na seção 3.3 deste artigo.

As variáveis independentes aplicadas ao modelo de aprendizado de máquina foram as apresentadas na seção 3.3 deste artigo e suas respectivas derivações, verificadas na fase de análise dos dados, são elas:

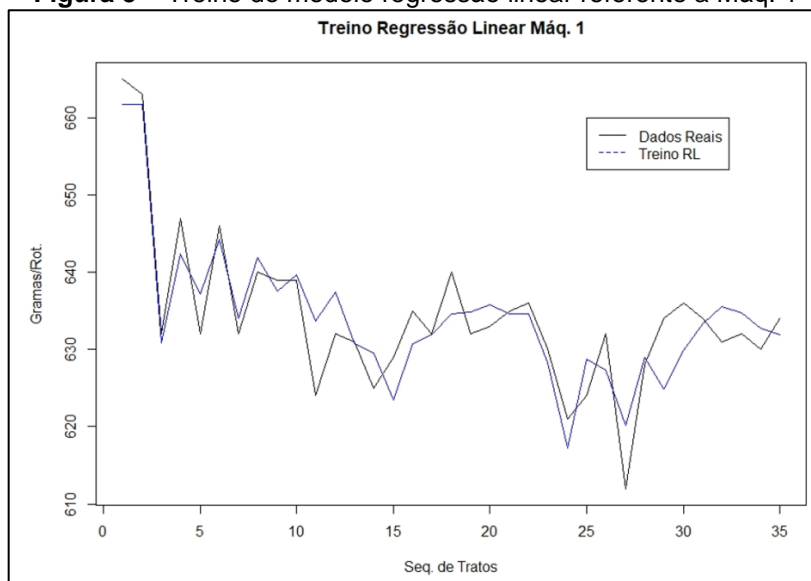
- a) Gramas programado por animal (GP);
- b) A média móvel dos últimos 5 ciclos de tratamento da variável gramas por rotação (GPR);
- c) *Lags* do inglês, ou retardos em português considerando os 5 ciclos de tratamento anteriores da variável gramas por rotação (GPR).

A variável prevista pelo modelo, ou variável denominada como dependente, será a variável de gramas por rotação (GPR). A partir da predição desta variável dependente, que conforme mencionado na seção 3.3 deste artigo, é uma relação entre as grandezas volume e massa, habilita a possibilidade de ajustar a quantidade de rotações aplicadas a rosca de transferência do equipamento. Com a habilitação do ajuste no número de rotações da rosca de transferência de uma maneira mais assertiva, o controle e precisão da dosagem de ração apresenta uma melhoria da característica de precisão de acordo com o comparativo entre o valor efetivamente programado no equipamento e o valor efetivamente realizado, o qual é verificado a cada ciclo por meio dos sensores mencionados no item 3.2 deste artigo.

4.1 Resultados do Modelo de Regressão Linear

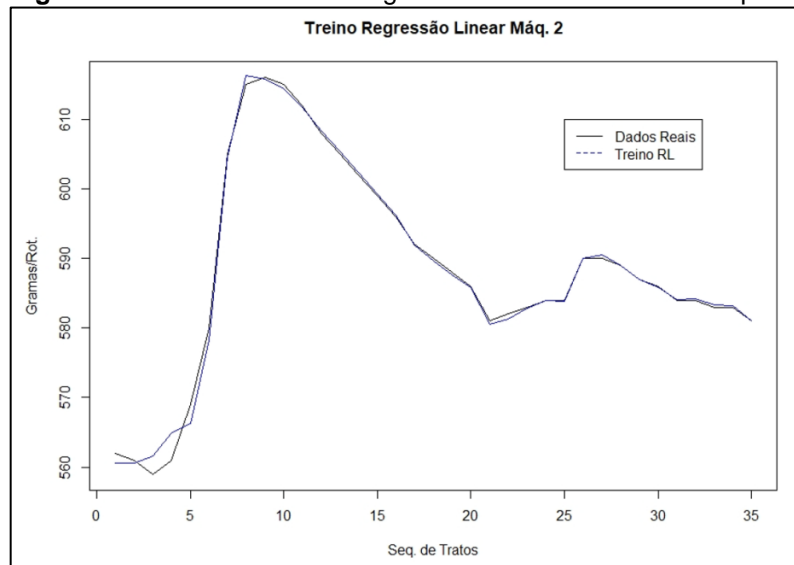
O modelo de aprendizado de máquina baseado na regressão linear foi construído considerando as variáveis independentes mencionadas na seção 4 deste documento. O modelo apresentou uma boa capacidade de aprendizado em relação aos pacotes de dados destinados ao processo de treinamento. O modelo de regressão linear tende a apresentar bons resultados, mesmo em casos com valores de gramas por rotação que apresentam uma alta volatilidade entre cada ciclo de tratamento, conforme demonstrado nas Figuras 5 a 7. Os resultados de aprendizado da regressão linear, para os três equipamentos demonstrados denota que a volatilidade do valor gramas por rotação é um com grande fator de relevância no desempenho do equipamento e que o modelo de aprendizado de máquina é efetivo para melhoria quesito precisão de dosagem de ração em função de sua capacidade de acompanhar os dados reais obtidos.

Figura 5 – Treino do modelo regressão linear referente a Máq. 1



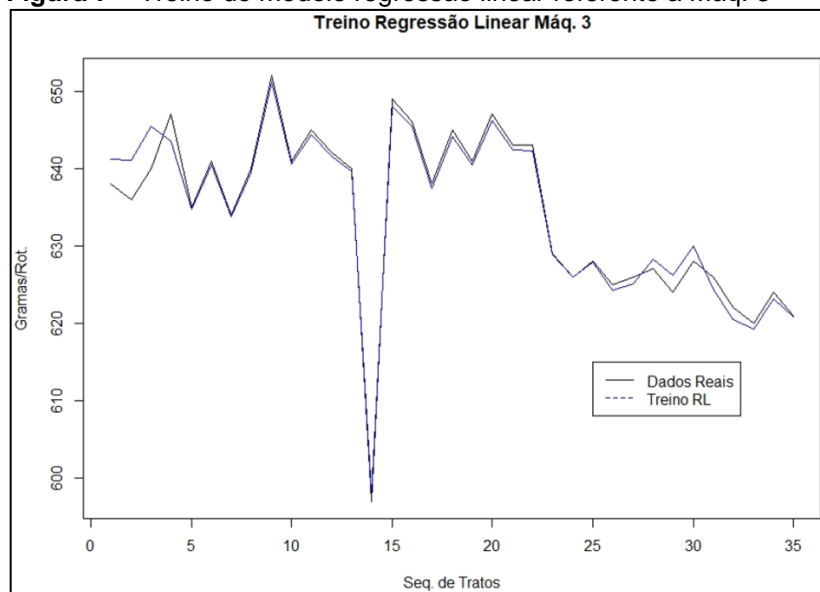
Fonte: Elaborado pelo autor (2021).

Figura 6 – Treino do modelo regressão linear referente a Máq. 2



Fonte: Elaborado pelo autor (2021).

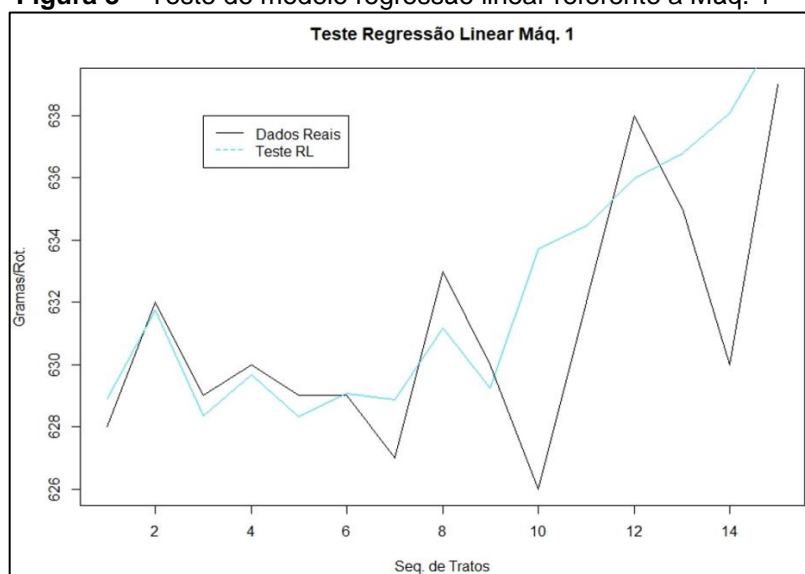
Figura 7 – Treino do modelo regressão linear referente a Máq. 3



Fonte: Elaborado pelo autor (2021).

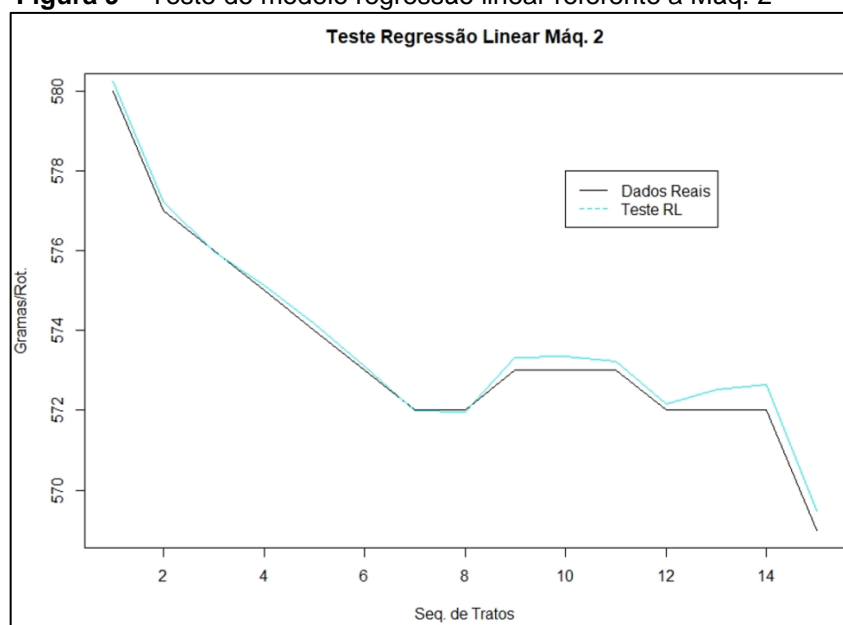
Em relação ao processo de teste do modelo de regressão linear para as três máquinas, mesmo com a grande volatilidade dos pacotes de dados de treino considerados, a capacidade do modelo de aprendizado de máquina de generalizar as condições em que é aplicado e de realizar as respectivas previsões dos valores de gramas por rotação com melhor precisão pode ser atingida, conforme as Figuras 8 a 10.

Figura 8 – Teste do modelo regressão linear referente a Máq. 1



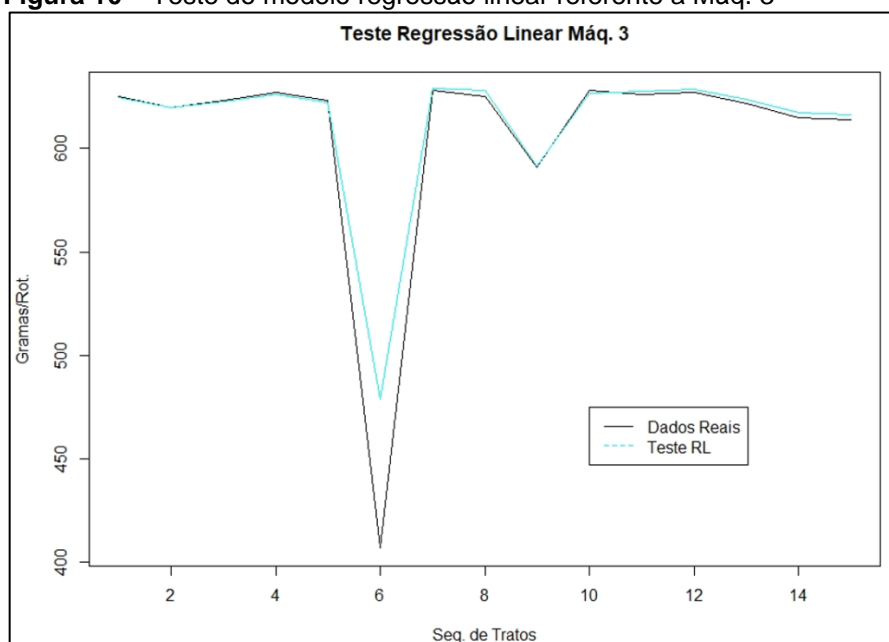
Fonte: Elaborado pelo autor (2021).

Figura 9 – Teste do modelo regressão linear referente a Máq. 2



Fonte: Elaborado pelo autor (2021).

Figura 10 – Teste do modelo regressão linear referente a Máq. 3



Fonte: Elaborado pelo autor (2021).

As considerações a respeito das verificações das métricas de erro em relação aos testes de validação do modelo de regressão linear para as máquinas 1, 2 e 3 mostra os valores de MAD e MAPE com níveis baixos e indicam que a aplicabilidade e capacidade de generalizar a volatilidade do comportamento das máquinas, principalmente em relação a variável de gramas por rotação, pode ser atendida para realizar as previsões da relação entre o volume e a massa de ração de forma mais precisa, conforme demonstrado na Tabela 1. Apesar dos valores da MAD e MAPE da máquina 3 apresentarem valores relativamente inferiores em relação aos apresentados pela máquina 1 na fase de aprendizado, na fase de teste a máquina 1 apresentou métricas de erro inferiores em relação a máquina 3. Isso ocorreu em função da máquina 3 ter um dos ciclos de tratamento com valor muito abaixo do que vinha ocorrendo historicamente de acordo com os dados avaliados, representando, do inglês, um *outlier*, ou em português um ponto fora da curva conjunto de dados. Mesmo com essa variação abruptada variável a ser prevista verificada na fase de teste, o modelo matemático demonstrou capacidade de atingir o valor com diferença drástica de comportamento, conforme demonstrado na Figura 10.

Tabela 1 – Métricas de erro do modelo de regressão linear

	MAD	MAPE
Treino Maq.1	3,37	0,53
Treino Maq.2	0,60	0,10
Treino Maq.3	1,15	0,18
Teste Máq. 1	2,07	0,33
Teste Máq. 2	0,24	0,04
Teste Máq. 3	5,97	1,36

Fonte: Elaborado pelo autor (2021).

4.2 Comparativo entre o Modelo de Regressão Linear e o Modelo Atual

Para uma visualização direta dos resultados do valor de gramas realizado por animal (GPA), que são os valores de comparação direta com dados reais programados e que foram atingidos pelo modelo de regressão linear e pelo modelo atual empregado nos equipamentos considerados para este artigo é demonstrado o valor da métrica de erro MAPE, onde o ganho percentual entre as duas metodologias consideradas pode ser visualizado de uma maneira mais intuitiva.

Por meio do cálculo das rotações aplicadas (NR) é possível realizar o cálculo do valor de GPA, conforme demonstrado pelas Equações 4 e 5.

$$NR = \frac{(GP.NA)}{GPR} \quad (4)$$

$$GPA = \frac{(NR.GPR)}{NA} \quad (5)$$

Com a finalidade de verificação do MAPE do modelo de regressão linear, o procedimento de cálculo para obtenção do GPA é calcular as rotações com base na variável de GPR prevista pelo modelo e em seguida o GPA é calculado utilizando o valor real de GPR demonstrado diretamente por cada máquina. A partir do GPA calculado para o modelo de regressão linear, o comparativo é realizado com o valor real de GPA armazenado em cada máquina, conforme a Tabela 2.

Tabela 2 – Comparativo MAPE Regressão Linear x Real

	MAPE Regressão	MAPE Real
Máquina 1	0,33	3,24
Máquina 2	0,04	0,80
Máquina 3	1,19	7,48
Média	0,52	3,84

Fonte: Elaborado pelo autor (2021).

De acordo com as métricas de erro verificadas na Tabela 2, o modelo de regressão linear adotado para as três máquinas atingiu um MAPE de 0,52% enquanto o MAPE do modelo real (atual utilizado pela máquina) calculado foi de 3,84%. Estes resultados de MAPE representaram um ganho médio de precisão da dosagem por ciclo de tratamento, considerando os resultados supracitados, de 7,4 vezes em relação ao modelo atual adotado (resultados reais verificados nos pacotes de dados).

5 CONSIDERAÇÕES FINAIS

O artigo teve como objetivo aumentar a precisão de dosagem de um sistema automatizado de alimentação de suínos, por meio do emprego de um modelo supervisionado de aprendizado de máquina. O modelo de aprendizado empregado, em função da avaliação do comportamento dos dados, foi o modelo de regressão linear, o qual pôde atender de forma satisfatória a necessidade de generalização das condições verificadas nos dados e realizar as previsões da variável dependente (relação entre as grandezas de massa e volume de ração e é dado de entrada para determinar a aplicação das rotações necessárias a rosca de transferência do equipamento). O emprego do modelo de regressão linear, em comparação com os dados reais armazenados pela máquina, apresentou ganhos significativos para a precisão de dosagem de ração, sendo que para os três equipamentos avaliados, a precisão obtida foi em média 7,4 vezes superior ao atual modelo utilizado pela máquina. Os resultados obtidos demonstraram que o objetivo proposto pelo trabalho, de incrementar a precisão de dosagem de ração e por consequência

otimizar os lotes de produção de suínos, bem como reduzir o consumo de ração foi atendido pelo modelo matemático proposto neste artigo.

Em relação a sugestão de trabalhos futuros sobre o tema, sugere-se a coleta de outras variáveis que seriam relevantes para a precisão de dosagem e utilizá-las no modelo de aprendizado de máquina, como temperatura, umidade, características físicas da ração etc. Outra sugestão seria a avaliação de outros modelos de aprendizado de máquina como redes neurais artificiais, árvore de decisão ou *random forest*.

REFERÊNCIAS

AMOURI, Amar; ALAPARTHY, Vishwa T.; MORGERA, Salvatore D.. A Machine Learning Based Intrusion Detection System for Mobile Internet of Things. **Sensors**, [S.L.], v. 20, n. 2, p. 461, 14 jan. 2020. DOI <http://dx.doi.org/10.3390/s20020461>.

BASHA, Rani Fathima Kamal; BHARATHI, M.L; VENUSAMY, Kanagaraj. Dynamic prediction of energy and power usage cost using linear regression-machine learning analysis. **Journal Of Physics: Conference Series**, [S.L.], v. 1921, p. 012067, 2021. IOP Publishing. DOI <http://dx.doi.org/10.1088/1742-6596/1921/1/012067>.

BEROZA, G. C.; SEGOU, M.; MOUSAVI, S. Mostafa. Machine learning and earthquake forecasting - next steps. **Nature Communications**, [S.L.], v. 12, n. 1, p. 1-3, 6 ago. 2021. DOI <http://dx.doi.org/10.1038/s41467-021-24952-6>.

BONSIGNORIO, Fabio; HSU, David; JOHNSON-ROBERSON, Matthew; KOBER, Jens. Deep Learning and Machine Learning in Robotics [From the Guest Editors]. **IEEE Robotics & Automation Magazine**, [S.L.], v. 27, n. 2, p. 20-21, jun. 2020. DOI <http://dx.doi.org/10.1109/mra.2020.2984470>.

BURY, F.; DELAERE, C. Matrix element regression with deep neural networks - Breaking the CPU barrier. **Journal Of High Energy Physics**, [S.L.], v. 2021, n. 4, p. 1-26, abr. 2021. Springer Science and Business Media LLC. DOI [http://dx.doi.org/10.1007/jhep04\(2021\)020](http://dx.doi.org/10.1007/jhep04(2021)020).

BROWNLEE, Jason. **Linear Regression for Machine Learning**. 2020. Elaborada por Machine Learning Algorithms. Disponível em: <https://machinelearningmastery.com/linear-regression-for-machine-learning/>. Acesso em: 24 fev. 2022.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Produção de Suínos, Embrapa Suínos e Aves**. 2003. Disponível em: <http://www.cnpsa.embrapa.br/SP/suinos/nutricao.html>. Acesso em: 06 out. 2021.

FLORES, B. E. A pragmatic view of accuracy measurement in forecasting. **Omega**, v. 14, n. 2, p. 93–98, 1986.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4 ed. São Paulo: Atlas, 2002.

HE, Yuqing; TIEZZI, Francesco; HOWARD, Jeremy; MALTECCA, Christian. Predicting body weight in growing pigs from feeding behavior data using machine learning algorithms. **Computers And Electronics in Agriculture**, [S.L.], v. 184, p. 106085-106100, 2021. DOI <http://dx.doi.org/10.1016/j.compag.2021.106085>.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. O’Texts, 2018. *E-book*.

IGHALO, Joshua O.; ADENIYI, Adewale George; MARQUES, Gonçalo. Application of linear regression algorithm and stochastic gradient descent in a machine-learning environment for predicting biomass higher heating value. **Biofuels, Bioproducts and Biorefining**, [S.L.], v. 14, n. 6, p. 1286-1295, 2020. DOI <http://dx.doi.org/10.1002/bbb.2140>.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Resultados Definitivos Censo Agro 2017, Portal IBGE, 2017. Disponível em: https://censoagro2017.ibge.gov.br/templates/censo_agro/resultadosagro/pecuaria.html?localidade=0&tema=1. Acesso em: 06 out. 2021.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: trends, perspectives, and prospects. **Science**, [S.L.], v. 349, n. 6245, p. 255-260, jul. 2015. <http://dx.doi.org/10.1126/science.aaa8415>.

KAGERMANN, H.; LUKAS, W.; WAHLSTER, W. Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. Industriellen Revolution. **VDI Nachrichten**, v. 13, n. 1, p. 2-3, 2011.

LAW, Rob. Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. **Tourism Management**, [S.L.], v. 21, n. 4, p. 331-340, 2000. DOI [http://dx.doi.org/10.1016/s0261-5177\(99\)00067-9](http://dx.doi.org/10.1016/s0261-5177(99)00067-9).

LEE, Woongsup; HAM, Younghwa; BAN, Tae-Won; JO, Ohyun. Analysis of Growth Performance in Swine Based on Machine Learning. **IEEE Access**, [S.L.], v. 7, p. 161716-161724, 2019. DOI <http://dx.doi.org/10.1109/access.2019.2951522>.

MAKRIDAKIS, S. Accuracy concerns measures: theoretical and practical concerns. **International journal of forecasting**, v. 9, p. 527–529, 1993.

MARCATO, Simara Márcia; LIMA, Gustavo Júlio Mello Monteiro de. Efeito da restrição alimentar como redutor do poder poluente dos dejetos de suínos. **Revista Brasileira de Zootecnia**, [S.L.], v. 34, n. 3, p. 855-863, jun. 2005. DOI <http://dx.doi.org/10.1590/s1516-35982005000300017>.

MARTINS, P.C., ALBUQUERQUE, M.P.; MACHADO, Mesquita, A. A. Implicações da Imuno castração na Nutrição de Suínos e nas características de Carcaça. **Archivos de Zootecnia**. 2013, 62, 105-118. Disponível em: <https://www.redalyc.org/articulo.oa?id=49558826008>

MUHURI, P. K.; SHUKLA, A. K.; ABRAHAM, Ajith. Industry 4.0: a bibliometric analysis and detailed overview. **Engineering Applications Of Artificial Intelligence**, [S.L.], v. 78, p. 218-235, 2019. <http://dx.doi.org/10.1016/j.engappai.2018.11.007>.

NAKAGAWA, Elisa Yumi; ANTONINO, Pablo Oliveira; SCHNICKE, Frank; KUHN, Thomas; LIGGESMEYER, Peter. Continuous Systems and Software Engineering for Industry 4.0: a disruptive view. **Information And Software Technology**, [S.L.], v. 135, p. 106562-106566, 2021. DOI <http://dx.doi.org/10.1016/j.infsof.2021.106562>.

NEWELL, Robert; NEWMAN, Lenore; MENDLY-ZAMBO, Zsofia. The Role of Incubators and Accelerators in the Fourth Agricultural Revolution: A Case Study of Canada. **Agriculture**, v. 11, p. 1-15, 2021. DOI <http://doi.org/10.3390/agriculture11111066>.

PAIALUNGA, Piero. **Deep Learning vs Linear Regression**: theoretical differences and hands on examples about MLP and linear regressor or in a machine learning problem. Theoretical differences and hands on examples about MLP and Linear Regressor in a Machine Learning problem.. 2022. Elaborada por Towards Data Science. Disponível em: <https://towardsdatascience.com/deep-learning-vs-linear-regression-ea74aca115ea>. Acesso em: 24 fev. 2022.

SCHUH, G. *et al.*, Industrie 4.0 maturity index: managing the digital transformation of companies. **Herbert Utz Verlag GmbH**, 2017.

SHUKLA, S. K. (Ed.). Industry 4.0—A Confluence of Embedded Artificial Intelligence, **Machine Learning, Robotics and Security**. 2018.

SRIVASTAVA, Tavish. **11 Important Model Evaluation Metrics for Machine Learning Everyone Should Know**. 2019. Elaborada por Analytics Vidhya. Disponível em: <https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/>. Acesso em: 24 fev. 2022.

SPANAKI, Konstantina; SIVARAJAH, Uthayasankar; FAKHIMI, Masoud; DESPOUDI, Stella; IRANI, Zahir. Disruptive technologies in agricultural operations: a systematic review of ai-driven agritech research. **Annals of Operations Research**, [S.L.], p. 1-34, 2021. DOI <http://dx.doi.org/10.1007/s10479-020-03922-z>.

SWAMINATHAN, Saishruthi. **Linear Regression — Detailed View**. 2018. Elaborada por Towards Data Science. Disponível em: <https://towardsdatascience.com/linear-regression-detailed-view-ea73175f6e86>. Acesso em: 24 fev. 2022.

ZHOU, Lina; PAN, Shimei; WANG, Jianwu; VASILAKOS, Athanasios V.. Machine learning on big data: opportunities and challenges. **Neurocomputing**, [S.L.], v. 237, p. 350-361, 2017. <http://dx.doi.org/10.1016/j.neucom.2017.01.026>.



Artigo recebido em: 24/02/2022 e aceito para publicação em: 30/08/2022

DOI: <https://doi.org/10.14488/1676-1901.v22i1.4586>