

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE CIÊNCIAS AGRÁRIAS  
CURSO DE ZOOTECNIA**

**CAMILA DE ALMEIDA FRANCESCHI**

**MODELOS ESTATÍSTICOS PARA PREDIÇÃO DA  
PRODUÇÃO FUTURA DA FÊMEA SUÍNA COM BASE NO  
PESO AO NASCIMENTO E OUTRAS CARACTERÍSTICAS  
DE INTERESSE**

**FLORIANÓPOLIS - SC**

**2022**

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE CIÊNCIAS AGRÁRIAS  
CURSO DE ZOOTECNIA**

**CAMILA DE ALMEIDA FRANCESCHI**

**MODELOS ESTATÍSTICOS PARA PREDIÇÃO DA  
PRODUÇÃO FUTURA DA FÊMEA SUÍNA COM BASE NO  
PESO AO NASCIMENTO E OUTRAS CARACTERÍSTICAS  
DE INTERESSE**

Trabalho de Conclusão de Curso apresentado  
como exigência para obtenção do Diploma de  
Graduação em Zootecnia da Universidade Federal  
de Santa Catarina.

Orientador (a): Prof.<sup>a</sup> Priscila Arrigucci Bernardes.

**FLORIANÓPOLIS - SC**

**2022**

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Franceschi, Camila de Almeida Franceschi  
Modelos estatísticos para predição da produção futura da  
fêmea suína com base no peso ao nascimento e outras  
características de interesse / Camila de Almeida Franceschi  
Franceschi ; orientadora, Priscila Arrigucci Bernardes  
Bernardes, 2022.  
61 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Centro de Ciências  
Agrárias, Graduação em Zootecnia, Florianópolis, 2022.

Inclui referências.

1. Zootecnia. 2. Análise dos componentes principais. 3.  
Limites de desempenho. 4. Random forest. 5. Regressão  
logística multinomial. I. Bernardes, Priscila Arrigucci  
Bernardes. II. Universidade Federal de Santa Catarina.  
Graduação em Zootecnia. III. Título.

Camila de Almeida Franceschi

**MODELOS ESTATÍSTICOS PARA PREDIÇÃO DA  
PRODUÇÃO FUTURA DA FÊMEA SUÍNA COM BASE NO  
PESO AO NASCIMENTO E OUTRAS CARACTERÍSTICAS  
DE INTERESSE**

Esta Monografia de Trabalho de Conclusão de Curso foi julgada aprovada e adequada para obtenção do grau de Zootecnista.

Florianópolis, 21 de Novembro de 2022.

**Banca Examinadora:**

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Priscila Arrigucci Bernardes  
Orientador  
Universidade Federal de Santa Catarina

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Lucélia Hauptli

---

Prof.<sup>o</sup> Dr.<sup>o</sup> Marcio Cinachi Pereira

## **AGRADECIMENTOS**

Agradeço em primeiro lugar à minha família, por todo apoio dado durante todos esses anos e a toda paciência e tempo investidos em mim principalmente durante esta importante etapa final da minha graduação. Muito obrigada Marinês Joana Soares de Almeida Franceschi, Mauri Marcos Franceschi, Carolina de Almeida Franceschi e Nicolas Evertom Duarte da Silva! Um obrigada também à minha avó e meu avô, por todos os momentos de risadas que temos juntas e por todo compartilhamento e aos meus queridos nonno e nonna que gosto tanto. Vocês possuem um espaço eterno em meu coração e em minha vida para sempre.

A todos os professores do curso que fizeram parte deste caminho da minha vida, em especial à minha professora orientadora Priscila Arrigucci Bernardes, a qual compartilhou esse grande desafio comigo com excelência e muita força de vontade, sempre me incentivando, auxiliando no que fosse necessário e tirando seu tempo para sanar minhas (muitas) dúvidas. Também agradeço à professora Lucélia Hauptli por toda ajuda realizada neste projeto, por sempre incentivar seus alunos e por não medir esforços para garantir o melhor para seus alunos e para o curso. Agradeço também a professora Priscila de Oliveira Moraes, que consegue se desdobrar em um milhão para atender a todos de maneira excepcional e por todos os conselhos e suporte ao longo do curso. Agradeço também aos professores coordenadores do Grupo de Estudos em Nutrição e Produção de Ruminantes (Genpror) por todos os ensinamentos e pelo excelente trabalho realizado para garantir excelentes profissionais no futuro. Muito obrigada! Todos vocês são excelentes pessoas e profissionais! Junto com vocês evolui muito e só tenho a agradecer!

Agradeço às excelentes profissionais Tatiane Chud e Daniela Grossi por todo o tempo disponibilizado para sanar nossas dúvidas e todo apoio com suas excelentes ideias, sem vocês este trabalho não seria nada. Muito obrigada!

Agradeço aos meus amigos e colegas do grupo de estudos Nutrição em Ação Pet (Neapet), por todas as noites, finais de semana e encontros para discussões e compartilhamentos de aprendizados, vocês são demais e fazem acontecer! Também agradeço aos meus amigos e colegas do Genpror, com vocês evolui e aprendi muito

em todas as reuniões e visitas realizadas. Á vocês meu muito obrigada e meus desejos de sucesso, pois vocês merecem e muito!

Agradeço a todos os meus amigos de fora do curso que fazem dos meus dias os mais felizes de todos, obrigada por todo apoio, suporte, compreensão, parceria e amizade muito além do que poderia imaginar, principalmente nesses últimos anos de correria, vocês são incríveis, cada um com seus jeitinhos que admiro muito. Amo todos vocês, muito obrigada!

Por fim, agradeço a você, leitor, por seu interesse neste assunto que amo tanto, espero que este texto lhe ajude de alguma maneira. Obrigada!

## RESUMO

O melhoramento genético tem papel fundamental na criação de linhagens de reprodutores com características que atendam ao objetivo de melhores índices zootécnicos. No entanto, na cadeia de produção de suínos, nem todos os animais apresentam o valor do índice para direcionar quais serão os reprodutores mais produtivos. Nestes casos, a escolha das melhores fêmeas para reposição do plantel possui papel fundamental para o sucesso da produção e lucratividade da cadeia. Dessa forma, o objetivo deste estudo foi avaliar o modelo *random forest* (RF) e de regressão logística multinomial (RLM) para predição da produção futura com base no peso ao nascimento e outras características de interesse, além de definir limites seguros de escolha da fêmea suína com base em tais características. Dados de 4.071 fêmeas suínas, sendo 1.904 fêmeas Landrace (LA) e 2.167 fêmeas Large White (LW) foram utilizados. Para o cálculo da produção futura, o número de leitões nascidos vivos, peso médio dos leitões ao nascimento e número estimado de leitões desmamados de quatro partos foram analisados e os animais foram separados em três classes, sendo elas, alta, moderada e baixa produtividade, as quais foram utilizadas como variáveis dependentes para entrada no modelo. Como preditores, as características utilizadas foram peso ao nascimento (PN), peso médio dos leitões ao nascimento (PMN), desvio padrão do peso médio ao nascimento dos leitões (DPN), número de leitões natimortos (Nat), número de leitões mumificados (Mum), número de leitões nascidos vivos (NLV), número de leitões vivos após 24 horas (V24) e taxa de sexo da leitegada (TS). O modelo de RLM e RF demonstraram ser mais eficientes em predizer os animais que não pertencem a classe analisada do que predizer os animais que realmente pertencem, além disso, obtiveram predição mediana para a classe de animais com maior produtividade. Por fim, os limites sugeridos para a leitegada de origem da matriz para a escolha das marrãs com melhor desempenho futuro para as raças LA e LW foi de PN maior ou igual a 1,40 e 1,20 kg, PMN a partir de 1,30 e 1,20 kg, DPN menor ou igual a 0,30 e 0,40, Nat menor ou igual a 2 e 4, Mum menor ou igual a 2 e 3, NLV a partir de 12 e 13 e V24 maior ou igual a 12 e 13, respectivamente. Utilizando estes limites, para a raça LA 72,8% dos animais escolhidos eram pertencentes a classe de produtividade alta e moderada e para a raça LW o valor foi de 69%. Adicionalmente, caso poucas fêmeas atendam aos critérios mencionados, sugere-se considerar somente as características de PN, PMD e NLV. O modelo de RF apresentou potencial de predição, sendo assim, sugere-se novos estudos utilizando os ajustes para aprimorar a acurácia do mesmo.

**Palavras-chave:** análise dos componentes principais; limites de desempenho; *random forest*; regressão logística multinomial; reprodutoras.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1</b> – Fluxo do aprendizado supervisionado .....	7
<b>Figura 2</b> – Estrutura da árvore de decisões.....	9
<b>Figura 3</b> – Metodologia utilizada para estabelecimento das categorias. ....	14
<b>Figura 4</b> – Exemplo de metodologia utilizada para estabelecimento das classes. ...	15
<b>Figura 5</b> – Separação dos dados dos modelos em grupos de treinamento e teste. .	17
<b>Figura 6</b> – Matriz de confusão considerada para cálculo de sensibilidade, especificidade, valor preditivo positivo e valor preditivo negativo para a classe 1. ...	18
<b>Figura 7</b> – Análise de componentes principais utilizando características da leitegada de origem das fêmeas Landrace (A) e das fêmeas Large White (B) e marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.....	22
<b>Figura 8</b> – Frequência de distribuição das classes com relação ao número de leitões nascidos vivos de fêmeas Landrace (A) e Large White (B) e com relação ao número de leitões vivos após 24 horas para fêmeas Landrace (C) e Large White (D) marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.....	31
<b>Figura 9</b> – Frequência de distribuição das classes com relação ao número de leitões mumificados de fêmeas Landrace (A) e Large White (B) e com relação ao número de leitões natimortos para fêmeas Landrace (C) e Large White (D) marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.....	32
<b>Figura 10</b> – Frequência de distribuição das classes com relação ao peso ao nascimento de fêmeas Landrace (A) e Large White (B) e com relação ao peso médio dos leitões ao nascimento para fêmeas Landrace (C) e Large White (D) marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.....	35
<b>Figura 11</b> – Frequência de distribuição das classes com relação ao desvio padrão do peso médio ao nascimento de fêmeas Landrace (A) e Large White (B) e com relação	



a taxa de sexo da leitegada para fêmeas Landrace (C) e Large White (D) marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.....36

## LISTA DE TABELAS

<b>Tabela 1</b> – Número de leitões para cada taxa de sexo da leitegada conforme a raça. 13	
<b>Tabela 2</b> – Média das características para cada parto. ....	13
<b>Tabela 3</b> – Número de observações com a pontuação total dos quatro partos (PonT). 14	
<b>Tabela 4</b> – Número de animais com a pontuação total dos quatro partos. ....	15
<b>Tabela 5</b> – Análise de Tukey dos animais da raça Landrace.....	20
<b>Tabela 6</b> – Análise de Tukey dos animais da raça Large White. ....	21
<b>Tabela 7</b> – Métricas de sensibilidade, especificidade, verdadeiro preditivo positivo (VPP), verdadeiro preditivo negativo (VPN), acurácia e coeficiente Kappa para o modelo <i>random forest</i> (RF) das fêmeas da raça Landrace. ....	24
<b>Tabela 8</b> – Métricas de sensibilidade, especificidade, verdadeiro preditivo positivo (VPP), verdadeiro preditivo negativo (VPN), acurácia e coeficiente Kappa para o modelo de regressão logística multinomial (RLM) das fêmeas da raça Landrace. ...	25
<b>Tabela 9</b> – Métricas de sensibilidade, especificidade, verdadeiro preditivo positivo (VPP), verdadeiro preditivo negativo (VPN), acurácia e coeficiente Kappa para o modelo <i>random forest</i> (RF) das fêmeas da raça Large White.....	27
<b>Tabela 10</b> – Métricas de sensibilidade, especificidade, verdadeiro preditivo positivo (VPP), verdadeiro preditivo negativo (VPN), acurácia e coeficiente Kappa para o modelo de regressão logística multinomial (RLM) das fêmeas da raça Large White. 28	
<b>Tabela 11</b> – Limites estabelecidos para as características da raça Landrace e Large White. ....	38

## LISTA DE ABREVIATURAS E SIGLAS

ACP - Análise dos componentes principais

ANOVA - Análise de variância

DPN - Desvio padrão do peso médio ao nascimento dos leitões

ED - Número estimado de leitões desmamados

FN - Falso negativo

FP - Falso positivo

LA - Landrace

LH - Hormônio luteinizante

LW - Large White

ML - *Machine learning*

MNL - Modelos não lineares

Mum - Número de leitões mumificados

Nat - Número de leitões natimortos

NLV - Número de leitões nascidos vivos

OOB - *Out-of-bag*

PMN - Peso médio dos leitões ao nascimento

PN - Peso ao nascimento

PonT - Pontuações totais

RF - *Random Forest*

RLM - Regressão logística multinomial

SNPs - Polimorfismo de nucleotídeo único

TS - Taxa de sexo da leitegada

V24 - Número de leitões vivos após 24 horas

VN - Verdadeiro negativo

VP - Verdadeiro positivo

VPN - Valor predito negativo

VPP - Valor predito positivo

# SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	<b>1</b>
<b>2. OBJETIVOS</b> .....	<b>2</b>
<b>3. REVISÃO BIBLIOGRÁFICA</b> .....	<b>2</b>
3.1. Landrace e Large White.....	2
3.2. Relação entre peso ao nascimento e produtividade da matriz .....	3
3.3. Modelos estatísticos .....	6
3.3.1. Regressão logística.....	6
3.3.2. Algoritmos .....	6
3.3.3. Aprendizado de máquina .....	7
3.3.4. Modelos não lineares .....	8
3.3.5. Random forest.....	9
<b>4. METODOLOGIA</b> .....	<b>10</b>
4.1. Descrição dos dados .....	10
4.2. Edição dos dados .....	11
4.2.1. Peso ao nascimento.....	11
4.2.2. Animais avaliados .....	12
4.2.3. Estimação de número de leitões desmamados para produtividade da fêmea suína .....	12
4.2.4. Determinação da taxa de sexo da leitegada de origem da matriz.....	12
4.3. Determinação da produtividade da fêmea (resposta) .....	13
4.4. Preditores .....	16
4.5. Modelos .....	16
4.6. Matriz de confusão .....	17
4.7 Limites seguros para melhor desempenho reprodutivo de matrizes.....	18

4.8 Ganho genético estimado por geração.....	19
<b>5. RESULTADOS E DISCUSSÃO.....</b>	<b>19</b>
5.1. Análise de variância (ANOVA).....	19
5.1.1. Raça Landrace.....	19
5.1.2. Raça Large White.....	21
5.2 Análise dos componentes principais (ACP).....	22
5.3 Random forest (RF) e regressão logística multinomial (RLM).....	22
5.3.1 Raça Landrace.....	23
5.3.2 Raça Large White.....	26
5.4. Limites sugeridos para análise da leitegada.....	29
5.4.1 Características relacionadas ao tamanho da leitegada.....	30
5.4.2 Características relacionadas ao peso da leitegada.....	34
5.4.3 Características relacionadas a proporção de sexo existente na leitegada.....	37
5.4.4. Limites de escolha e ganho genético estimado por geração.....	38
<b>6. CONCLUSÕES.....</b>	<b>40</b>
<b>7. REFERÊNCIAS.....</b>	<b>41</b>

# 1. INTRODUÇÃO

A carne suína ocupa o segundo lugar no mundo como a mais produzida (FAO, 2022) e consumida, dessa forma, a suinocultura possui elevada relevância nos mercados mundiais, chegando a uma estimativa de produção para 2022 de 109,887 milhões de toneladas métricas em peso de carcaça equivalente (USDA, 2022). Sendo assim, a forte liderança da suinocultura acarretou em diversos investimentos no setor para a melhoria da produtividade desses animais.

Neste contexto, além de outros fatores, o melhoramento genético foi fundamental para a criação de linhagens de reprodutores com características desejáveis, levando a melhores índices zootécnicos. No entanto, na cadeia de produção de suínos, nem todos os animais apresentam o valor do índice para direcionar quais serão os reprodutores mais produtivos. Nestes casos, a escolha das melhores fêmeas para reposição do plantel possui papel fundamental para o sucesso da produção e lucratividade da cadeia.

Em geral, o aumento da prolificidade das fêmeas promove aumento na proporção de leitões com menor peso ao nascimento. Esta redução no peso já foi relacionada, por diferentes estudos, com diversos efeitos negativos no desempenho de crescimento e sobrevivência desses animais (QUINIOU et al., 2002; WOLF et al., 2008; FIX et al. 2010a; FIX et al. 2010b; BIDANEL et al., 2011), porém, poucos estudos associaram esta característica com o desempenho reprodutivo futuro das matrizes.

Diferentes características podem ser utilizadas para analisar o desempenho reprodutivo das matrizes suínas, dentre estas, o tamanho e peso da leitegada, número de partos por fêmea por ano e o número de leitões desmamados por fêmea. Além disso, outros fatores que vêm sendo estudados são com relação a escolha da matriz com base na leitegada de origem da mesma, incluindo características como a proporção de sexo e tamanho da leitegada (WARDA et al. 2021; SEYFANG et al., 2017; LAMBERSON et al., 1988).

Uma vez que há o envolvimento de muitas características na leitegada de origem da fêmea e na avaliação da produtividade da mesma, torna-se difícil para o produtor a escolha das melhores fêmeas para a reposição. Assim, uma alternativa para lidar com as relações entre essas características é a utilização de modelos matemáticos.

## 2. OBJETIVOS

- Definir limites para as características de número de leitões nascidos vivos, de leitões vivos após 24 horas, de mumificados, de natimortos, peso ao nascimento, peso médio dos leitões ao nascimento, desvio padrão do peso médio ao nascimento dos leitões e taxa de sexo da leitegada de nascimento das marrãs, com o intuito de selecionar matrizes de reposição dentro de rebanhos suínos;
- Avaliar o modelo de floresta aleatória (*random forest*) e de regressão logística multinomial (RLM) para predição da produção futura das fêmeas suínas das raças Large White (LW) e Landrace (LA) com base no peso ao nascimento das mesmas e outras características de interesse.

## 3. REVISÃO BIBLIOGRÁFICA

### 3.1. Landrace e Large White

Entre as raças puras criadas comercialmente o Landrace (LA) e Large White (LW) possuem destaque, sendo ambos animais de pelagem branca e diferenciando-se principalmente pelo tipo de orelha, a qual é do tipo céltico para os LA e asiática nos animais LW. Os animais da raça LA são caracterizados pela sua precocidade sexual, prolificidade, boa produção de leite, habilidade materna, baixa deposição de gordura, alta taxa de crescimento, eficiência alimentar e alto rendimento de carne de boa qualidade, dessa forma, seus reprodutores são utilizados em cruzamentos tanto na linha materna quanto paterna. Ademais, os animais LW são caracterizados pela sua precocidade sexual, prolificidade, excelente taxa reprodutiva e de crescimento, eficiência alimentar, ausência do alelo Haln e carcaças de alto rendimento com carne de qualidade. Devido a presença das características desejáveis nestas raças elas normalmente são cruzadas para a obtenção de fêmeas F-1 com ganhos em heterose nos programas de melhoramento (IRGANG, 2014).

Para que uma característica seja incluída em um programa de melhoramento genético, a avaliação de alguns parâmetros genéticos é necessária para a correta condução da seleção. A herdabilidade ( $h^2$ ) é um parâmetro de grande importância a

ser considerado e, segundo Pereira (2012), esta é caracterizada como a proporção da variação total de determinada característica que advém da natureza genética, podendo variar de 0,0 a 1,0 ou se utilizando de porcentagem, de 0 a 100%.

Rothschild e Bidanel (1998) verificaram estimativas de herdabilidade para características reprodutivas de machos e fêmeas suínos encontradas na literatura e observaram que estas, de maneira geral, apresentam herdabilidade de baixa a moderada, demonstrando que não há alta correlação entre o genótipo e o fenótipo do animal, o que dificulta o rápido progresso por meio da seleção nessas características.

### **3.2. Relação entre peso ao nascimento e produtividade da matriz**

A produtividade das matrizes suínas dentro do rebanho pode ser avaliada por meio de diferentes características. Dentre estas, destacam-se o número de partos por fêmea por ano e o número de leitões desmamados por fêmea, os quais irão indicar o número de leitões por matriz acasalada por ano (KENNEDY, 2017), que se trata de um indicador de produtividade, com alvo em maior número de animais. Este e outros indicadores podem ser obtidos através do uso de programas de melhoramento genético com o intuito de aprimorar o rebanho, objetivando menor custo de produção e, conseqüentemente, maior retorno econômico (PEREIRA, 2012).

A verificação de parâmetros genéticos irá auxiliar na tomada de decisão para a verificação das características de interesse. No caso da fêmea suína com foco para o desempenho reprodutivo, verifica-se que além do tamanho, o peso da leitegada também possui grande importância e se constitui como uma característica de interesse para o sucesso da produção de suínos, sendo que o tamanho da leitegada contribui amplamente para o peso da mesma (PIRES, 1999). Nesse sentido, segundo Quiniou et al. (2002), em estudo com 965 leitegadas de mães Large White X Landrace inseminadas com sêmen de machos Large White x Pietrain, observaram que leitegadas com 11 ou menos leitões apresentavam um peso ao nascimento (PN) médio de 1,59 kg, que era reduzido para 1,26kg em leitegadas com 16 ou mais leitões. Além disso, os autores também observaram que a variação média do PN individual dentro da leitegada, obtido a partir da média do desvio padrão de cada leitegada do grupo, também diferiu conforme o número de leitões, sendo 0,26 kg para as leitegadas



menores e 0,30 kg para as leitegadas maiores, demonstrando que quanto maior o tamanho da leitegada, menor o peso ao nascimento e maior é a variação desse peso entre os leitões.

Progressos genéticos foram observados para o tamanho da leitegada, conforme demonstrado por Tribout et al. (2002), em um estudo utilizando sêmen congelado de machos nascidos em 1977 ou em 1998. Estes autores verificaram um significativo aumento na característica de tamanho da leitegada ao nascimento, que apresentou uma tendência genética anual de  $+0,15 \pm 0,05$  leitões nascidos vivos. No entanto, a constante seleção com base no aumento do tamanho da leitegada para melhorar a prolificidade das fêmeas suínas levou a um decréscimo no peso médio dos leitões ao nascimento, afetando a sobrevivência e desempenho futuros destes leitões (QUINIOU et al., 2002). Neste contexto, Wolf et al. (2008) verificaram uma correlação fenotípica negativa (-0,50) entre o peso mínimo ao nascimento e o número de leitões natimortos, além de uma correlação de 0,94 entre o PN mínimo e médio da leitegada. Assim, os autores concluíram que a seleção para o tamanho da leitegada deve levar em conta a seleção por características de PN e, para isso, propuseram a inclusão da característica de PN mínimo nas leitegadas como potencial para seleção.

Semelhantemente, Fix et al. (2010a) demonstraram que o PN influenciou significativamente na sobrevivência dos leitões, pois o aumento do PN proporcionou maior probabilidade de leitões nascidos vivos e redução na possibilidade de mortalidade antes do desmame. De fato, estudos demonstraram que mais de 90% dos leitões mais pesados (acima de 1kg de peso vivo) sobreviveram nas primeiras 24 horas de vida, enquanto que 33% dos leitões da classe mais leve (menos de 0,61kg) morreram (QUINIOU et al., 2002). Sendo assim, é possível verificar indiretamente as relações opostas entre as características de tamanho da leitegada e a sobrevivência dos leitões (BIDANEL et al., 2011).

Devido a influência do PN na sobrevivência dos animais, diferentes estudos foram realizados para verificar a relação entre PN e demais características de importância na produtividade de suínos. Almeida et al. (2015) estudaram os efeitos do PN no desenvolvimento do sistema reprodutivo e início da puberdade em fêmeas suínas e verificaram que essas características não foram afetadas pelo PN. No entanto, os autores observaram que as características de peso corporal variaram conforme o PN e sugeriram que demais estudos são necessários para verificar o

desenvolvimento reprodutivo conforme o PN e a viabilidade de inclusão dessa característica nos programas de melhoramento em fêmeas. Neste contexto, ao avaliar características em marrãs que foram afetadas pelo PN, Magnabosco et al. (2016) verificaram que o peso e o ganho de peso médio diário na puberdade, a idade em que foram expostas ao cachaço, e o peso e ganho de peso médio diário no acasalamento variavam de acordo com o PN das marrãs. Além disso, os autores também observaram que as fêmeas mais leves produziram menos leitões ao longo de três partos e também permaneceram por menor tempo no rebanho.

Embora poucos estudos tenham demonstrado a relação entre o PN com o futuro desempenho reprodutivo da matriz (ALMEIDA et al., 2015), a inclusão dessa característica em programas de melhoramento na linhagem materna ocorre principalmente devido à baixa quantidade de marrãs selecionadas para reprodução quando o PN é baixo, pois haverá maiores perdas das mesmas no período pré-seleção, acarretando em perdas econômicas nas unidades de reposição (ALMEIDA et al., 2014).

Outra característica avaliada por Warda et al. (2021) e que pode refletir na vida produtiva futura da matriz foi o tamanho da leitegada em que a mesma nasceu. Estes autores observaram que as matrizes de leitegadas pequenas (com até 9 leitões) tiveram o primeiro parto em uma idade mais avançada e apresentaram menor eficiência, medida pelo número de leitões nascidos vivos até 100 dias da vida produtiva. Já as matrizes advindas de ninhadas maiores (a partir de 14 leitões) apresentaram maior eficiência, maior vida reprodutiva e produziram leitegadas maiores. No mesmo sentido, em outro estudo verificando as características das leitegadas de origem das marrãs, Vallet et al. (2016) encontraram associações do peso ao nascimento e da taxa de crescimento pré-desmame com a característica de idade a puberdade, de forma a refletir no desempenho reprodutivo futuro, sendo que, os resultados indicaram que, ao combinar os efeitos de maior peso ao nascimento com um crescimento mais lento no período pré-desmame, as leitoas têm sua idade a puberdade atrasada.

Além de influências entre tamanho da leitegada e produtividade, outra relação estudada é a proporção do sexo dos animais nascidos na leitegada de origem da matriz com a produtividade futura desta. De acordo com Seyfang et al. (2017), quando comparadas fêmeas nascidas em leitegadas com predominância de machos e as

nascidas em leitegadas com predominância de fêmeas, as primeiras demoram mais tempo, em horas, para iniciarem o aumento do hormônio luteinizante (LH) circulante, assim como o intervalo de tempo do início até atingir o pico de LH foi menor nestas fêmeas nascidas em leitegadas com predominância de machos. Da mesma forma, em um estudo realizado por Lamberson et al. (1988), em que os autores analisaram a idade ao primeiro estro de fêmeas nascidas em leitegadas com diferentes proporções de sexo, foi observado que o maior número de machos na leitegada promovia um declínio na idade ao primeiro parto, entretanto, devido a um baixo coeficiente de determinação, os autores não definiram essa característica como um critério para seleção.

### **3.3. Modelos estatísticos**

Em geral, a análise de dados de uma determinada população tem como objetivo estudar o problema a partir dos dados gerados, para isso, faz-se o uso dos modelos estatísticos, com a finalidade de determinar, aproximadamente, características importantes da população (OLIVE, 2010).

#### **3.3.1. Regressão logística**

Na regressão logística, a resposta será a predição de uma variável categórica, normalmente binária, com base em uma ou mais variáveis independentes, as quais podem ser contínuas e/ou binárias. Quando essa regressão apresentar como resposta uma variável binária e apenas uma variável independente o modelo é chamado de regressão logística univariada e, nos casos de mais de uma variável independente, o modelo se caracteriza por ser uma regressão logística múltipla (GONZALEZ, 2018), também chamada de multinomial (RLM).

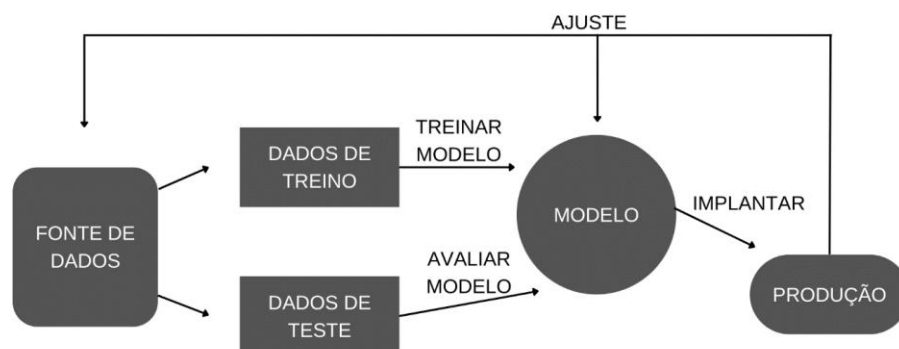
#### **3.3.2. Algoritmos**

Segundo HILL (2016), um algoritmo é uma estrutura de controle composta finita, abstrata, eficaz, dada imperativamente, cumprindo um determinado propósito

sob determinadas disposições. Em outras palavras, os algoritmos podem ser definidos como o conjunto de ações necessárias para executar uma determinada tarefa, sendo que, tais ações necessitam de um nível de precisão para que o algoritmo computacional consiga realizar sua função de trazer a solução utilizando de maneira eficiente os recursos computacionais frente a um problema estipulado (CORMEN, 2017).

### 3.3.3. Aprendizado de máquina

O aprendizado de máquina, chamado de *machine learning* (ML), é um ramo da inteligência artificial que tem como função aprender com base em um banco de dados, de maneira a tomar decisões a partir de experiências anteriores (MONARD e BARANAUSKAS, 2003). Tais métodos são classificados como não supervisionados quando utilizados para compreender a estrutura e os padrões do conjunto de dados, sendo que, todas as variáveis são tratadas como iguais, a exemplo da análise dos componentes principais (DASGUPTA et. al., 2011); e classificados como métodos supervisionados, quando buscam pela relação existente entre um conjunto de variáveis independentes e a variável dependente com base em dados de treinamento. Estes últimos dividem os dados em conjuntos de treino, o qual contém a variável de saída, e teste, no qual o algoritmo irá aplicar os padrões aprendidos no conjunto de treino para realizar a predição ou classificação (MAHESH, 2020) (Figura 1).



**Figura 1** – Fluxo do aprendizado supervisionado. Fonte: Adaptado de MAHESH, 2020.

Juntamente com as tecnologias emergentes da agricultura 4.0, como sensores, *big data* e inteligência artificial, os mecanismos de aprendizado de máquina ajudam a reduzir custos de produção e aumentar a eficiência do sistema (NEETHIRAJAN, 2020). Além disso, tais algoritmos vêm sendo amplamente utilizados na literatura para estimação de parâmetros como, porcentagem de carne na carcaça de suínos (SZYNDLER-NEDZA et al., 2016), relações entre a temperatura corporal de suínos com certas variáveis ambientais (BASAK et al., 2020), predição de valores genéticos genômicos com o uso de marcadores do tipo polimorfismo de nucleotídeo único (JOAQUIM, 2019) e classificação de reprodutores suínos (SOUZA, 2022).

#### **3.3.4. Modelos não lineares**

Os modelos não lineares (MNL) se caracterizam por serem não lineares em seus parâmetros, ou seja, por possuírem pelo menos uma das derivadas parciais como dependente de pelo menos um dos parâmetros do modelo em questão (MATTOS, 2013). Segundo Zeviani et al., (2013), algumas vantagens na utilização dos MNL com relação aos modelos lineares se baseiam no fato de que, diferentemente dos modelos de regressão lineares, os MNL são escolhidos de forma embasada em informações prévias, permitindo realizar predições fora do domínio observado nas informações de entrada. Além disso, são modelos que geralmente são mais parcimoniosos, de forma a apresentarem menos parâmetros, os quais apresentam uma interpretação mais prática, porém, como desvantagens, os MNL possuem métodos de inferência aproximados.

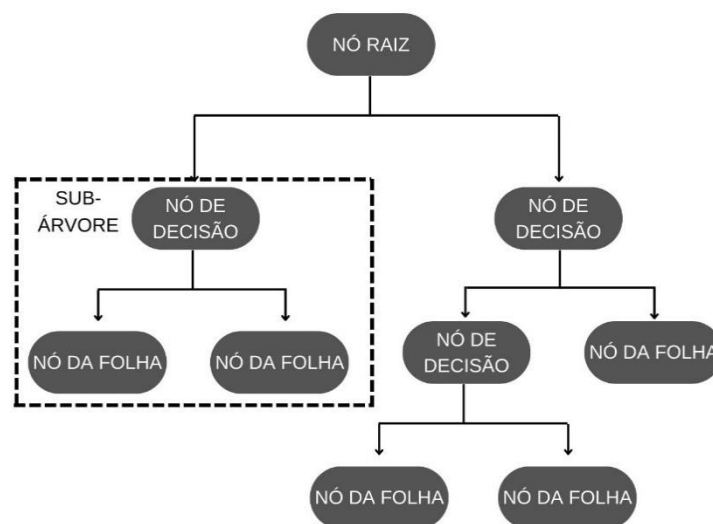
Ademais, os MNL podem ser menos flexíveis do que os modelos lineares, de forma que a boa escolha do modelo é de grande importância. Além disso, tais modelos não possuem solução de forma analítica para a estimação de seus parâmetros, o que torna obrigatório a realização de um método numérico para este fim, utilizando valores iniciais pré-determinados de forma a se sustentar em procedimentos iterativos de estimação e podendo afetar a convergência do algoritmo de estimação (ARCHONTOULIS e MIGUEZ, 2015).

Dessa forma, estudos comparando os MNL com os modelos lineares vêm sendo realizados para verificar qual modelo se aplica melhor para a predição das

características de interesse. Neste sentido, avaliaram-se estes modelos para mensurar o desempenho de vacas em lactação em termos de produção de leite e, foi verificado que o MNL demonstrou resultados melhores do que o modelo linear, por possivelmente superar a suposição de linearidade entre a produção animal e os nutrientes da dieta e também incluir maior complexidade (SAXENA, 2012).

### 3.3.5. *Random forest*

A árvore de decisão é um modelo de aprendizado supervisionado de classificação, na qual o resultado parte da raiz e é descrito por uma sequência de dados até alcançar o nó da folha (CHARBUTY e ABDULAZEEZ, 2021). A Figura 2 ilustra a estrutura básica de uma árvore de decisão.



**Figura 2** – Estrutura da árvore de decisões. Fonte: Adaptado de CHARBUTY e ABDULAZEEZ, 2021.

O *Random Forest* (RF) é o método combinado de aprendizado de máquina mais conhecido entre tais metodologias (LI et al., 2018) e que expandiu seu uso principalmente a partir de 2018 no setor da carne suína (VAN KLOMPENBURG e KASSAHUN, 2022), sendo amplamente utilizado em problemas de regressão e classificação (PONTE et al., 2020). Este tende a obter uma melhor predição do que os modelos de regressão linear, pois se adapta facilmente às não linearidades contidas no conjunto de dados (SCHONLAU e ZOU, 2020). Além disso, é capaz de

produzir interpretações simples a partir da análise de dados de alta dimensionalidade (JOAQUIM, 2019).

Tal modelo se caracteriza por ser um algoritmo de aprendizado baseado em árvore que realiza previsões a partir de múltiplas árvores de decisão (SCHONLAU e ZOU, 2020), as quais são construídas a partir de uma amostra aleatória de inicialização dos dados. Tal metodologia é conhecida como *bagging* ou agregação *bootstrap* e consiste em gerar diversas versões de um preditor, as quais serão utilizadas para gerar uma previsão agregada (BREIMAN, 1996). Além disso, cada árvore recebe um conjunto aleatório dos preditores que atua juntamente com a amostra *bootstrap* buscando as variáveis ideais para cada nó até que a árvore seja construída (DASGUPTA et al., 2011).

Para a análise do modelo é possível realizar a diminuição média do coeficiente de Gini, que demonstra, de forma métrica, a contribuição de cada variável para a homogeneidade dos nós e folhas do modelo RF, sendo que, quanto maior for esse valor, maior será a importância da variável para o modelo (MARTINEZ-TABOADA e REDONDO, 2020). Após gerar as árvores de decisão, as mesmas irão votar na classe mais popular para gerar a resposta final do modelo (BREIMAN, 2001). Uma das vantagens do uso do RF é que não haverá o *over-fit* devido à Lei dos Grandes Números (NICOLA, 2021). Além disso, em comparação a outras metodologias de aprendizado de máquina, o RF apresenta uma estrutura mais simplificada de maneira a facilitar sua interpretação (ALSAHAF et al., 2018b).

Na área das agrárias, o RF se mostrou eficaz para a previsão da idade ao abate de suínos no início da terminação (ALSAHAF et al., 2018a), aumento da uniformidade do peso corporal de suínos em terminação a partir da atribuição de grupos de animais (ALSAHAF et al., 2018b), identificação de um subconjunto de SNPs e seu uso para previsão genômica de valores genéticos (LI et al., 2018) e osteocondrose de suínos (BAKOEV et al., 2020).

## **4. METODOLOGIA**

### **4.1. Descrição dos dados**

O presente estudo foi conduzido de setembro de 2021 até outubro 2022 no Laboratório de Genética Animal localizado no Departamento de Zootecnia e Desenvolvimento Rural, da Universidade Federal de Santa Catarina (UFSC) – Florianópolis.

Foram utilizados para as análises deste trabalho os dados cedidos por uma empresa de genética suína, contendo informações iniciais de 68.391 fêmeas suínas da raça Landrace (LA) e 129.266 fêmeas suínas da raça Large White (LW) provenientes de 24 e 31 rebanhos, respectivamente. Estas informações continham características que foram obtidas do conjunto de dados ou que foram calculadas, sendo referente a produtividade das fêmeas (número de leitões nascidos vivos, peso médio dos leitões ao nascimento e número estimado de leitões desmamados) e referente a leitegada de origem das fêmeas (peso ao nascimento, peso médio dos leitões ao nascimento, desvio padrão do peso médio ao nascimento dos leitões, número de leitões natimortos, número de leitões mumificados, número de leitões nascidos vivos, número de leitões vivos após 24 horas e taxa de sexo da leitegada). As características que foram calculadas e a edição dos dados realizada está descrita nos próximos subtópicos.

## **4.2. Edição dos dados**

### **4.2.1. Peso ao nascimento**

Foram definidos os pesos ao nascimento das fêmeas em avaliação, dentro das suas leitegadas de origem e dos leitões que esta fêmea produziu enquanto matriz avaliada durante sua vida produtiva. Sendo assim, para obtenção da idade que um leitão possuía quando foi pesado, a seguinte subtração foi realizada:

$$\textit{Idade do animal na pesagem} = \textit{Data da pesagem} - \textit{Data de nascimento do animal}$$

O peso dos animais com idade na pesagem maior ou igual a - 2 dias e menor ou igual a 2 dias foi classificado como peso ao nascimento (PN).



Após avaliar todas as informações do banco de dados, para cada característica de desempenho produtivo foram removidos valores mínimos ou máximos que se encontravam fora dos limites fisiológicos de produção, de forma a remover erros tanto da produção da fêmea suína quanto da produção da leitegada de origem da mesma. Assim, para a característica peso médio dos leitões ao nascimento, foram removidos animais que tiveram valores menores que 0,7 kg e para a característica de peso ao nascimento, os animais com valores menores do que 0,7 ou maiores que 2,7 kg foram removidos da análise.

#### **4.2.2. Animais avaliados**

Nas análises, apenas as fêmeas que apresentaram quatro primeiros partos foram avaliadas, ou seja, as informações de fêmeas com menos de quatro partos e informações de partos acima do quarto foram retiradas do banco de dados. Dessa forma, foram analisados os dados de 16.284 partos de 4.071 fêmeas, sendo 7.616 observações de 1.904 fêmeas Landrace (LA) e 8.668 observações de 2.167 fêmeas Large White (LW).

#### **4.2.3. Estimação de número de leitões desmamados para produtividade da fêmea suína**

Foi fornecido pela empresa a taxa de mortalidade esperada até a desmama dos leitões com base no peso ao nascimento dos mesmos, dessa forma, para estimar o número de desmamados, a taxa de mortalidade fornecida para cada faixa de peso ao nascimento foi multiplicada pelo número de leitões nascidos vivos.

#### **4.2.4. Determinação da taxa de sexo da leitegada de origem da matriz**

A taxa de sexo da leitegada (TS) foi realizada com base no número de animais nascidos vivos para verificar se a matriz nasceu em uma leitegada com predominância de machos ou fêmeas, sendo observada a distribuição descrita na Tabela 1.

**Tabela 1** – Número de leitões para cada taxa de sexo da leitegada conforme a raça.

Raça	Taxa de sexo da leitegada (número de leitões)		
	Machos = Fêmeas	Machos > Fêmeas	Machos < Fêmeas
Landrace	179	662	1063
Large White	254	765	1148

### 4.3. Determinação da produtividade da fêmea (resposta)

Para a definição da produtividade da fêmea, resposta do modelo utilizado posteriormente, a média das características de número de leitões nascidos vivos, peso médio dos leitões ao nascimento e número estimado de leitões desmamados de cada parto foi calculada separadamente para cada característica, parto e raça (Tabela 2).

**Tabela 2** – Média das características para cada parto.

Característica	Raça <sup>4</sup>	Parto 1	Parto 2	Parto 3	Parto 4
NLV <sup>1</sup>	LA	12,6544	13,1911	13,8119	13,5572
PMN <sup>2</sup>	LA	1,3966	1,5732	1,5475	1,5100
ED <sup>3</sup>	LA	9,5226	10,6790	11,0811	10,7146
NLV <sup>1</sup>	LW	12,8329	13,7360	14,2182	13,9298
PMN <sup>2</sup>	LW	1,3639	1,4643	1,4244	1,3917
ED <sup>3</sup>	LW	9,4773	10,6605	10,8010	10,3649

<sup>1</sup>NLV: Número de leitões nascidos vivos. <sup>2</sup>PMN: Peso médio dos leitões ao nascimento. <sup>3</sup>ED: Número estimado de leitões desmamados. <sup>4</sup>LA: Landrace e LW: Large White.

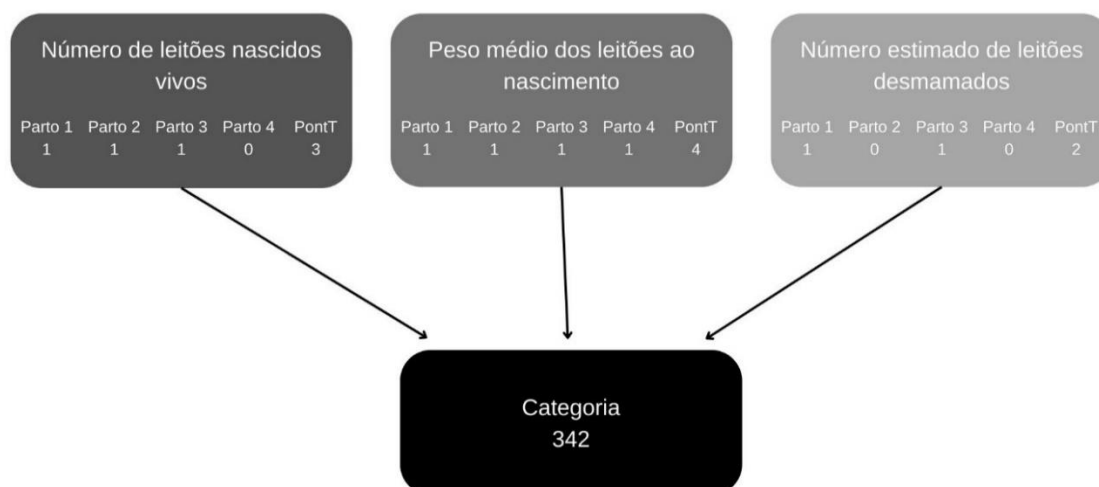
Para as três características selecionadas, os animais que obtiveram desempenho superiores à média dos partos, separadamente, recebiam a pontuação 1 e os animais com desempenho menor ou igual a média recebiam a pontuação zero. Cada característica teve sua pontuação somada nos quatro partos para obtenção da pontuação total da variável (Tabela 3).

**Tabela 3** – Número de observações com a pontuação total dos quatro partos (PonT).

Característica	Raça <sup>4</sup>	Número de observações com a pontuação total dos quatro partos (PonT)				
		0	1	2	3	4
NLV <sup>1</sup>	LA	608	1496	2336	2152	1024
PMN <sup>2</sup>	LA	1656	1820	1516	1400	1224
ED <sup>3</sup>	LA	640	1500	2228	2184	1064
NLV <sup>1</sup>	LW	700	1736	2496	2444	1292
PMN <sup>2</sup>	LW	1904	1944	1840	1656	1324
ED <sup>3</sup>	LW	924	1828	2312	2392	1212

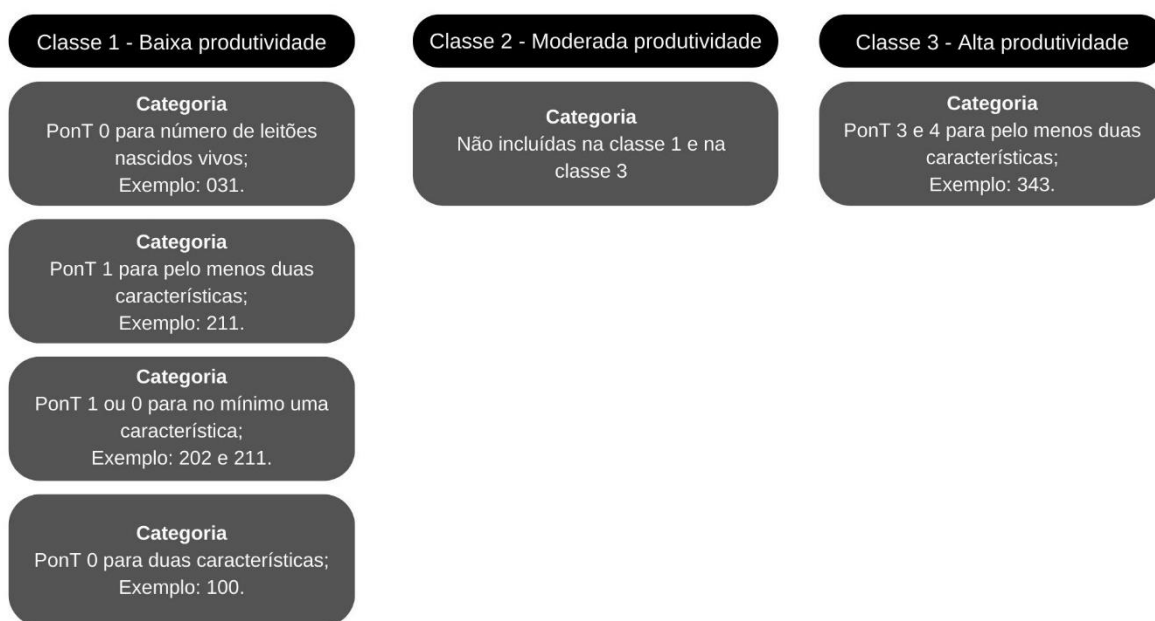
<sup>1</sup>NLV: Número de leitões nascidos vivos. <sup>2</sup>PMN: Peso médio dos leitões ao nascimento. <sup>3</sup>ED: Número estimado de leitões desmamados. <sup>4</sup>LA: Landrace e LW: Large White.

As pontuações totais (PonT) para cada característica foram concatenadas e foram obtidas 80 categorias de animais com três dígitos para LA e 82 para o LW, sendo o primeiro dígito referente a PonT do número de leitões nascidos vivos, o segundo ao peso médio dos leitões ao nascimento e o terceiro ao número estimado de leitões desmamados (Figura 3).



**Figura 3** – Metodologia utilizada para estabelecimento das categorias.

Com a finalidade de reduzir o número de categorias, estas foram reorganizadas para se determinar apenas três classes, sendo que a classe 1 incluiu animais que obtiveram em sua categoria, PonT 0 para o número de leitões nascidos vivos, ou animais que obtiveram PonT 1 para pelo menos duas características, ou animais com PonT 1 ou 0 para no mínimo uma característica, ou animais com PonT 0 para duas características; a classe 2 foi formada por animais que não foram incluídos na classe 1 e na classe 3 e; a classe 3 foi composta por animais que obtiveram em sua categoria PonT 3 e 4 para pelo menos duas características (Figura 4). Dessa maneira, a classe 1 foi composta por fêmeas de baixa produtividade, a classe 2 por fêmeas de produtividade moderada e a classe 3 por fêmeas de alta produtividade (Tabela 4).



**Figura 4** – Exemplo de metodologia utilizada para estabelecimento das classes.

**Tabela 4** – Número de animais com a pontuação total dos quatro partos.

Raça	Número de animais conforme sua classe e produtividade		
	1 (Baixa)	2 (Moderada)	3 (Alta)
Landrace	600	685	619
Large White	693	787	687

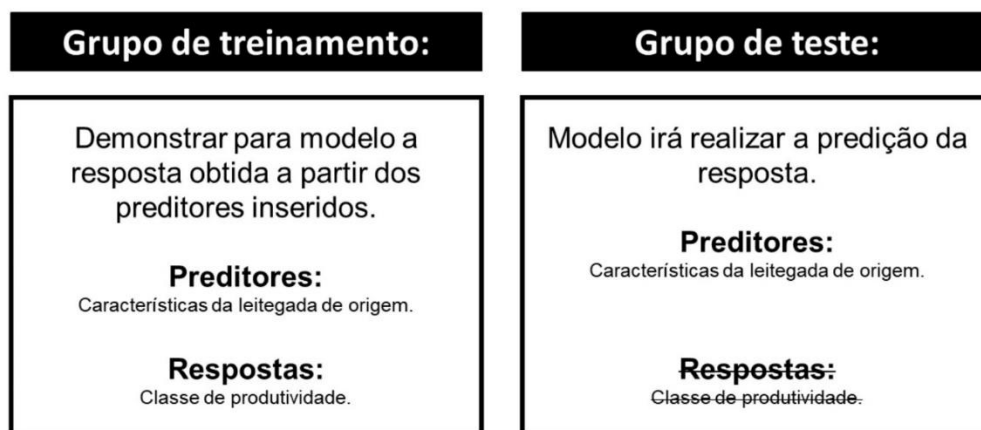
#### 4.4. Preditores

Os preditores utilizados nos modelos foram referentes às características da leitegada de origem destas fêmeas, aquelas que foram classificadas de acordo com a produtividade, sendo considerados o peso ao nascimento (PN), peso médio dos leitões ao nascimento (PMN), desvio padrão do peso médio ao nascimento dos leitões (DPN), número de leitões natimortos (Nat), número de leitões mumificados (Mum), número de leitões nascidos vivos (NLV), número de leitões vivos após 24 horas (V24) e taxa de sexo da leitegada (TS).

#### 4.5. Modelos

Para avaliar de forma simples se cada característica da leitegada de origem apresentava diferença de acordo com as classes de produtividade (baixa, moderada, alta), utilizou-se inicialmente um modelo linear univariado, sendo realizada a análise de variância (ANOVA) e o teste de Tukey. Então, com a finalidade de identificar se animais das diferentes classes (baixa, moderada, alta) de produtividade apresentavam perfis de leitegada de origem diferentes quando consideradas todas as características em simultaneidade, realizou-se uma análise de componentes principais (ACP). Por fim, estudou-se a aplicação dos modelos de *Random Forest* (RF) e regressão logística multinomial (RLM) para verificar se ao utilizar as informações das características da leitegada de origem (preditores) seria possível obter a classificação correta dos animais de acordo com sua classe de produtividade (resposta). Estes modelos foram avaliados pois uma vez que a predição se mostra adequada, é possível utilizar o modelo para, a partir de informações da leitegada de origem (preditores), identificar fêmeas que terão maior produtividade no futuro. Todas as análises foram realizadas utilizando o programa computacional R (R Core Team, 2021). A análise de RF foi feita por meio do pacote *randomForest*, o qual apresentou como padrão para o conjunto de dados do presente trabalho a utilização de 500 árvores de decisão e número de variáveis amostradas aleatoriamente igual a 2 (LIAW e WIENER, 2018).

Para análise dos modelos RF e RLM o conjunto de dados foi dividido em grupo de treinamento, o qual foi composto por preditores (características da leitegada de origem) e as respostas (classe de produtividade), de forma a demonstrar para o modelo qual resposta é obtida a partir dos preditores inseridos e, o grupo de teste, o qual era formado por somente as informações dos preditores (características da leitegada de origem), omitindo a resposta (classe de produtividade), para que o modelo faça a predição da resposta (Figura 5). Para essa separação utilizou-se a metodologia da validação cruzada, separando os dados em três grupos, em que dois eram destinados para o treinamento e um para validação, sendo os grupos alternados, de forma que cada um dos três grupos em algum momento fosse considerado na validação. Assim, o resultado obtido foi a média das três rodadas realizadas. Após o treinamento do modelo e a aplicação do mesmo no grupo de teste, verificou-se o desempenho da predição do modelo pela matriz de confusão.



**Figura 5** – Separação dos dados dos modelos em grupos de treinamento e teste.

#### 4.6. Matriz de confusão

Para avaliar o desempenho de classificação dos diferentes modelos de predição, RF e RLM, no grupo de teste, as classes preditas pelos modelos foram comparadas com as classes reais dos animais (que foram omitidas neste grupo teste). A partir dessa comparação foi possível obter métricas de acurácia, coeficiente Kappa, sensibilidade, especificidade, valor preditivo positivo (VPP) e valor preditivo negativo (VPN), segundo Borchers et al. (2017), conforme demonstrado abaixo:

$$\text{Sensibilidade} = \frac{VP}{(VP + FN)} \times 100;$$

$$\text{Especificidade} = \frac{VN}{(VN + FP)} \times 100;$$

$$VPP = \frac{VP}{(VP + FP)} \times 100;$$

$$VPN = \frac{VN}{(VN + FN)} \times 100.$$

Em que VP é o verdadeiro positivo; VN é o verdadeiro negativo; FP é o falso positivo e FN é o falso negativo (Figura 6).

Matriz de confusão		Valor real		
		1 (Baixa)	2 (Moderada)	3 (Alta)
Valor predito	1 (Baixa)	Verdadeiro positivo	Falso positivo	Falso positivo
	2 (Moderada)	Falso negativo	Verdadeiro negativo	Verdadeiro negativo
	3 (Alta)	Falso negativo	Verdadeiro negativo	Verdadeiro negativo

**Figura 6** – Matriz de confusão considerada para cálculo de sensibilidade, especificidade, valor preditivo positivo e valor preditivo negativo para a classe 1.

Segundo Keceli et al. (2020), a sensibilidade é considerada a proporção de positivos reais que são classificados corretamente; a especificidade é a porção de negativos reais que são detectados corretamente; o VPP é definido como a proporção de resultados positivos na classificação; e o VPV é proporção de resultados negativos na classificação.

#### 4.7 Limites seguros para melhor desempenho reprodutivo de matrizes

Para a definição de limites a serem atendidos por cada característica da leitegada de origem para considerar uma matriz produtiva foram realizados gráficos com a frequência de animais (%) para cada classe utilizando o pacote ggplot2 do programa computacional R (R Core Team, 2021), sendo que, as características que apresentavam valores com mais de uma casa decimal foram arredondadas para a análise. Através destes limites, foi verificado o número de animais que atendiam aos critérios de todas as características e realizado o ganho genético estimado por geração.

## 4.8 Ganho genético estimado por geração

Para o cálculo do ganho genético estimado por geração ( $\Delta G$ ) foi considerada seleção massal, portanto, a acurácia utilizada foi a raiz quadrada da herdabilidade ( $h^2$ ), a qual foi fornecida pela empresa. Sendo que, a fórmula utilizada para o cálculo do  $\Delta G$  foi a seguinte:

$$\Delta G = i \times \sqrt{h^2} \times \sigma_a$$

Onde,

$\Delta G$ : Ganho genético estimado por geração;

$i$ : Intensidade de seleção;

$\sqrt{h^2}$ : Acurácia;

$\sigma_a$ : Desvio padrão da variância genética aditiva.

## 5. RESULTADOS E DISCUSSÃO

### 5.1. Análise de variância (ANOVA)

#### 5.1.1. Raça Landrace

Na análise separadamente de cada característica da leitegada de origem da matriz Landrace a partir da ANOVA verificou-se que as características número de



leitões nascidos vivos ( $P = 0,0007676$ ) e número de leitões vivos após 24 horas ( $P = 0,004543$ ) foram significativos de acordo com a classificação dos animais. Adicionalmente, ao verificar as médias pelo teste de Tukey (Tabela 5) de ambas as características foi observado que para o NLV a classe 3 apresentou maior média e diferiu significativamente da classe 1 ( $P = 0,0005$ ). Com relação a característica de V24 a mesma situação foi observada, sendo que, as classes 1 e 3 diferiram significativamente entre si ( $P = 0,0035$ ).

As características de peso ao nascimento ( $P = 0,3639$ ), peso médio dos leitões ao nascimento ( $P = 0,1066$ ), desvio padrão do peso ao nascimento ( $P = 0,5197$ ), número de leitões natimortos ( $P = 0,7239$ ) e número de leitões mumificados ( $P = 0,5254$ ) e taxa de sexo da leitegada ( $P = 0,5755$ ) não foram significativas na análise da ANOVA.

**Tabela 5** – Análise de Tukey dos animais da raça Landrace.

<b>Característica</b>	<b>Classe 1</b>	<b>Classe 2</b>	<b>Classe 3</b>
Peso ao nascimento (kg)	1,55 <sup>a</sup>	1,55 <sup>a</sup>	1,57 <sup>a</sup>
Peso médio dos leitões ao nascimento(kg)	1,49 <sup>a</sup>	1,47 <sup>a</sup>	1,50 <sup>a</sup>
Desvio padrão do peso ao nascimento	0,285 <sup>a</sup>	0,282 <sup>a</sup>	0,287 <sup>a</sup>
Número de leitões natimortos	0,805 <sup>a</sup>	0,855 <sup>a</sup>	0,835 <sup>a</sup>
Número de leitões mumificados	0,273 <sup>a</sup>	0,270 <sup>a</sup>	0,307 <sup>a</sup>
Número de leitões nascidos vivos	13,5 <sup>b</sup>	13,8 <sup>ab</sup>	14,1 <sup>a</sup>
Número de leitões vivos após 24 horas	12,8 <sup>b</sup>	13,1 <sup>ab</sup>	13,3 <sup>a</sup>
Taxa de sexo predominante na leitegada	-0,240 <sup>a</sup>	-0,209 <sup>a</sup>	-0,184 <sup>a</sup>

Letras minúsculas diferentes na mesma linha indicam diferença significativa no teste de Tukey onde  $p < 0,01$ .

Para a raça LA, as variáveis que possuem uma relação linear univariada com a produtividade dos animais são as que envolvem o tamanho da leitegada, a qual é um dos fatores importantes para o sucesso da criação de suínos (DA ROSA et al., 2014).

De fato, o tamanho da leitegada de origem da matriz reflete em sua vida produtiva futura, sendo que, matrizes de leitegadas pequenas (menos que 9 leitões) alcançam o primeiro parto em uma idade mais avançada e apresentam menor eficiência, medida pelo número de leitões nascidos vivos até 100 dias da vida

produtiva. Já matrizes advindas de ninhadas maiores (mais que 14 leitões) apresentam maior eficiência, maior vida reprodutiva e produzem leitegadas maiores (WARDA et al., 2021). Estudos de Giacomeli et al. (2010) demonstraram efeito da raça com relação ao tamanho da leitegada, no qual avós da raça LA obtiveram maior número médio de leitões por leitegada ( $p < 0,01$ ) em comparação as bisavós LW.

### 5.1.2. Raça Large White

Com relação a raça Large White, na ANOVA, a característica significativa para a classificação dos animais foi apenas o peso médio dos leitões ao nascimento ( $P = 0,01473$ ), sendo que, ao analisar as médias a partir do teste de Tukey (Tabela 6) desta característica, foi verificado que as médias das classes 1, 2 e 3 foram 1,40; 1,42 e 1,43, respectivamente. Além disso, a classe 1 demonstrou diferença significativa da classe 3 para PMN ( $P = 0,0105$ ), dessa maneira, explicitando a superioridade das fêmeas da classe de alta produtividade para as fêmeas de baixa produtividade.

As demais características, como peso ao nascimento ( $P = 0,2268$ ), desvio padrão do peso ao nascimento ( $P = 0,662$ ), número de nascidos vivos ( $P = 0,2077$ ), número de leitões vivos após 24 horas ( $P = 0,171$ ), número de leitões natimortos ( $P = 0,8516$ ), número de leitões mumificados ( $P = 0,3655$ ) e taxa de sexo predominante na leitegada ( $P = 0,4333$ ) não se mostraram significativas na ANOVA.

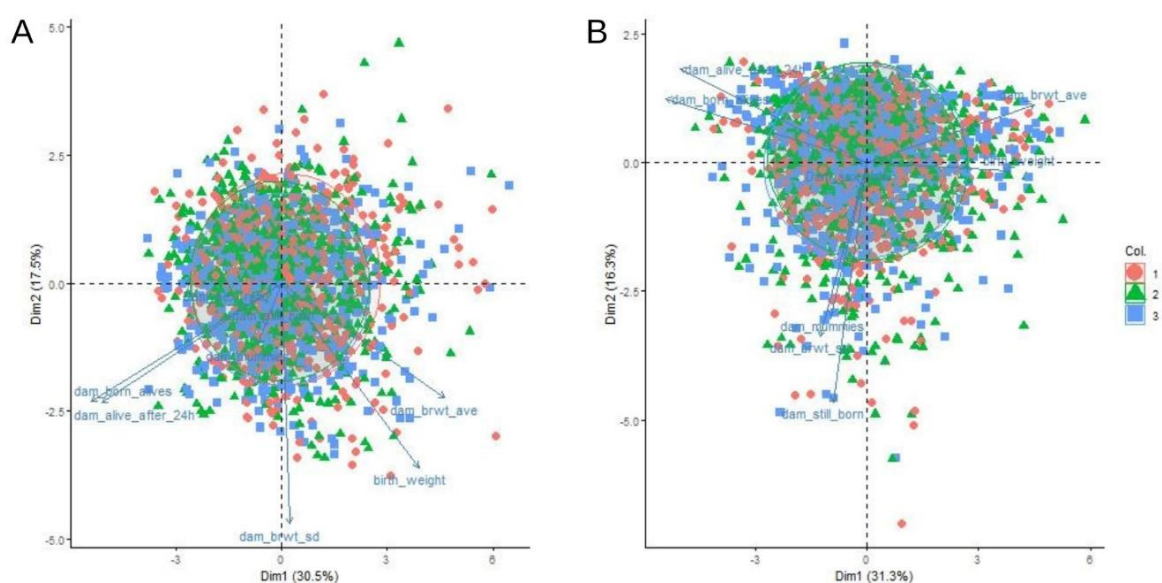
**Tabela 6** – Análise de Tukey dos animais da raça Large White.

<b>Característica</b>	<b>Classe 1</b>	<b>Classe 2</b>	<b>Classe 3</b>
Peso ao nascimento (kg)	1,47 <sup>a</sup>	1,48 <sup>a</sup>	1,50 <sup>a</sup>
Peso médio dos leitões ao nascimento(kg)	1,40 <sup>b</sup>	1,42 <sup>ab</sup>	1,43 <sup>a</sup>
Desvio padrão do peso ao nascimento	0,268 <sup>a</sup>	0,268 <sup>a</sup>	0,727 <sup>a</sup>
Número de leitões natimortos	1,22 <sup>a</sup>	1,18 <sup>a</sup>	1,19 <sup>a</sup>
Número de leitões mumificados	0,341 <sup>a</sup>	0,301 <sup>a</sup>	0,351 <sup>a</sup>
Número de leitões nascidos vivos	13,7 <sup>a</sup>	13,8 <sup>a</sup>	14,0 <sup>a</sup>
Número de leitões vivos após 24 horas	12,9 <sup>a</sup>	13,0 <sup>a</sup>	13,2 <sup>a</sup>
Taxa de sexo predominante na leitegada	-0,150 <sup>a</sup>	-0,210 <sup>a</sup>	-0,166 <sup>a</sup>

Letras minúsculas diferentes na mesma linha indicam diferença significativa no teste de Tukey onde  $p < 0,01$ .

## 5.2 Análise dos componentes principais (ACP)

A análise de todas as características da leitegada de origem das matrizes em conjunto a partir da ACP resultou em uma distribuição homogênea entre as classes das fêmeas LA e LW (Figura 7). Essa distribuição indica que não há uma distinção evidente entre os grupos de fêmeas de produtividade baixa, moderada e alta.



**Figura 7** – Análise de componentes principais utilizando características da leitegada de origem das fêmeas Landrace (A) e das fêmeas Large White (B) e marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.

Conforme verificado por Mills et al. (2020), de forma consistente com o gerenciamento atual das granjas de suínos, efeitos combinados da genética, criação cruzada, cronograma de desmame e seleção de mães, resultam em PN, tamanho de leitegada ao nascimento e durante a amamentação, peso ao desmame e número de dias lactando muito semelhantes entre os animais, dessa forma, resultando em um alto grau de similaridade entre os mesmos. Este perfil também foi verificado no presente estudo, no qual o alto grau de similaridade diminuiu a capacidade de separação de animais.

## 5.3 Random forest (RF) e regressão logística multinomial (RLM)

Após os resultados obtidos analisando as características separadamente por meio da ANOVA e em conjunto através da ACP, modelos como o RF e RLM foram utilizados para verificar a possibilidade de predição de fêmeas mais produtivas a partir de informações da leitegada de origem da mesma. A habilidade de predição dos modelos pode ser verificada por meio da matriz de confusão, a qual forneceu base para cálculo das métricas de sensibilidade, especificidade, verdadeiro predito positivo (VPP) e verdadeiro predito negativo (VPN) para cada uma das rodadas da validação cruzada considerando a raça Landrace com o modelo de RF (Tabela 7) e com o modelo de RLM (Tabela 8), e considerando a raça Large White com o modelo RF (Tabela 9) e com o modelo RLM (Tabela 10).

### **5.3.1 Raça Landrace**

Para os animais da raça LA, os resultados obtidos da classe 1 (animais considerados de baixa produtividade) pelo modelo RF possui sensibilidade mediana, indicando que, dos animais que nos valores reais eram classificados como classe 1, o modelo foi capaz de prever corretamente 31,97%. Para a classe 2 (produtividade moderada) e 3 (alta produtividade) o acerto dos valores positivos foi de 42,79% e 36,33%, respectivamente. Em contrapartida, com relação a classe 1, o modelo RF apresentou alta especificidade, ou seja, dos animais que nos valores reais não pertenciam a classe 1 o modelo conseguiu prever que 71,18% realmente não pertenciam a esta classe. O acerto dos valores negativos para a classe 2 e 3 foram de 61,90% e 72,43%, respectivamente. Dessa forma, ao comparar os valores reais com os valores preditos, o modelo é capaz de predizer com mais assertividade os animais que não pertencem a classe analisada do que predizer os animais que realmente pertencem.

**Tabela 7** – Métricas de sensibilidade, especificidade, verdadeiro preditivo positivo (VPP), verdadeiro preditivo negativo (VPN), acurácia e coeficiente Kappa para o modelo *random forest* (RF) das fêmeas da raça Landrace.

Item	Classes de acordo com produtividade			Acurácia	Kappa <sup>1</sup>
	1 (Baixa)	2 (Moderada)	3 (Alta)		
<i>Random Forest – Rodada 1</i>					
Sensibilidade	0,3480	0,4434	0,3905	0,3953	0,0912
Especificidade	0,7425	0,6522	0,6965		
VPP <sup>2</sup>	0,3901	0,4050	0,3886		
VPN <sup>3</sup>	0,7064	0,6870	0,6981		
<i>Random Forest – Rodada 2</i>					
Sensibilidade	0,2950	0,4612	0,3596	0,3764	0,0579
Especificidade	0,7172	0,5856	0,7546		
VPP <sup>2</sup>	0,3241	0,3905	0,4078		
VPN <sup>3</sup>	0,6887	0,6537	0,7149		
<i>Random Forest – Rodada 3</i>					
Sensibilidade	0,3163	0,3793	0,3398	0,347	0,0174
Especificidade	0,6758	0,6194	0,7220		
VPP <sup>2</sup>	0,3039	0,3651	0,3704		
VPN <sup>3</sup>	0,6883	0,6336	0,6944		
<i>Random Forest – Média das três rodadas</i>					
Sensibilidade	0,3197	0,4279	0,3633	0,3729	0,0555
Especificidade	0,7118	0,6190	0,7243		
VPP <sup>2</sup>	0,3393	0,3868	0,3889		
VPN <sup>3</sup>	0,6944	0,6581	0,7024		

<sup>1</sup>Kappa: Coeficiente Kappa. <sup>2</sup>Verdadeiro preditivo positivo. <sup>3</sup>Verdadeiro preditivo negativo.

**Tabela 8** – Métricas de sensibilidade, especificidade, verdadeiro preditivo positivo (VPP), verdadeiro preditivo negativo (VPN), acurácia e coeficiente Kappa para o modelo de regressão logística multinomial (RLM) das fêmeas da raça Landrace.

Item	Classes de acordo com produtividade			Acurácia	Kappa <sup>1</sup>
	1 (Baixa)	2 (Moderada)	3 (Alta)		
<i>Regressão Logística Multinomial – Rodada 1</i>					
Sensibilidade	0,1225	0,5973	0,2952	0,3449	0,0071
Especificidade	0,8955	0,3865	0,7247		
VPP <sup>2</sup>	0,3571	0,3420	0,3463		
VPN <sup>3</sup>	0,6831	0,6426	0,6754		
<i>Regressão Logística Multinomial – Rodada 2</i>					
Sensibilidade	0,2650	0,5172	0,2758	0,3606	0,0298
Especificidade	0,7701	0,5236	0,7361		
VPP <sup>2</sup>	0,3464	0,3846	0,3294		
VPN <sup>3</sup>	0,6950	0,6533	0,6838		
<i>Regressão Logística Multinomial – Rodada 3</i>					
Sensibilidade	0,1683	0,4784	0,3252	0,3328	-0,0154
Especificidade	0,8356	0,4826	0,6659		
VPP <sup>2</sup>	0,3142	0,3480	0,3190		
VPN <sup>3</sup>	0,6918	0,6159	0,6722		
<i>Regressão Logística Multinomial – Média das três rodadas</i>					
Sensibilidade	0,1852	0,5309	0,2987	0,3461	0,0071
Especificidade	0,8337	0,4642	0,7089		
VPP <sup>2</sup>	0,3392	0,3582	0,3315		
VPN <sup>3</sup>	0,6899	0,6372	0,6771		

<sup>1</sup>Kappa: Coeficiente Kappa. <sup>2</sup>Verdadeiro preditivo positivo. <sup>3</sup>Verdadeiro preditivo negativo.

Os valores de VPP da classe 1 representam uma predição mediana, indicando que dos animais que o modelo classificou como sendo da classe 1, 33,93% realmente eram deste grupo nos valores reais e o restante dos valores preditos como 1 pertenciam ou a classe 2 ou a classe 3. Para a classe 2 e 3 esse resultado foi de 38,68% e 38,89%. Já os valores de VPN da classe 1 indicaram que dos animais que não pertenciam a esta classe, o modelo fez a predição correta de 69,44%, sendo que esse valor foi de 65,81% para a classe 2 e 70,24% para a classe 3. Sendo assim, ao analisar os valores preditos com relação aos valores reais, o modelo é mais eficaz em prever os animais que não pertencem a determinada classe dos que os que pertencem. Além disso, a acurácia (0,3729) e o coeficiente Kappa (0,0555), que demonstra o grau de concordância (confiabilidade e precisão) dos dados na classificação (PERROCA e GAIDZINSKI, 2003) do modelo RF foram baixos.

As métricas de análise para o modelo de RLM evidenciaram que tal modelo foi menos eficaz do que o RF na sensibilidade das classes 1 e 3, na especificidade das classes 2 e 3 e nos VPP e VPN de todas as classes. Além disso, apresentou menor acurácia e coeficiente Kappa quando comparado ao modelo RF

Adicionalmente, a métrica de importância das características analisada através da diminuição média de Gini demonstrou que a característica mais importante para o modelo na raça LA é a do peso ao nascimento, seguida por DPN, PMN, NLV, V24, Nat, TS e Mum.

### **5.3.2 Raça Large White**

Os resultados obtidos para os animais da raça LW foram semelhantes aos dos animais LA para o modelo RF e RLM. Assim, o RF se mostrou mais eficiente que o RLM nas métricas de sensibilidade das classes 2 e 3, especificidade da classe 2, VPP das classes 1 e 3 e VPN da classe 1, além de apresentar melhor coeficiente Kappa. Adicionalmente, a sequência de maior importância das características para o modelo foi de peso ao nascimento, número de leitões nascidos vivos, desvio padrão do peso médio ao nascimento dos leitões, peso médio dos leitões ao nascimento, número de leitões vivos após 24 horas, número de leitões natimortos, taxa de sexo da leitegada e número de leitões mumificados.

**Tabela 9** – Métricas de sensibilidade, especificidade, verdadeiro preditivo positivo (VPP), verdadeiro preditivo negativo (VPN), acurácia e coeficiente Kappa para o modelo *random forest* (RF) das fêmeas da raça Large White.

Item	Classes de acordo com produtividade			Acurácia	Kappa <sup>1</sup>
	1 (Baixa)	2 (Moderada)	3 (Alta)		
<i>Random Forest – Rodada 1</i>					
Sensibilidade	0,2939	0,4609	0,2731	0,3463	0,0129
Especificidade	0,7368	0,5172	0,7582		
VPP <sup>2</sup>	0,3401	0,3440	0,3571		
VPN <sup>3</sup>	0,6933	0,6359	0,6796		
<i>Random Forest – Rodada 2</i>					
Sensibilidade	0,3392	0,4373	0,2918	0,3596	0,0331
Especificidade	0,7218	0,5717	0,7387		
VPP <sup>2</sup>	0,3581	0,3686	0,3469		
VPN <sup>3</sup>	0,7047	0,6399	0,6869		
<i>Random Forest – Rodada 3</i>					
Sensibilidade	0,3361	0,3545	0,3148	0,3366	0,0019
Especificidade	0,6570	0,6410	0,7035		
VPP <sup>2</sup>	0,3252	0,3682	0,3119		
VPN <sup>3</sup>	0,6681	0,6272	0,7063		
<i>Random Forest – Média das três rodadas</i>					
Sensibilidade	0,3230	0,4175	0,2932	0,3475	0,0159
Especificidade	0,7052	0,5766	0,7334		
VPP <sup>2</sup>	0,3411	0,3602	0,3386		
VPN <sup>3</sup>	0,6887	0,6343	0,6909		

<sup>1</sup>Kappa: Coeficiente Kappa. <sup>2</sup>Verdadeiro preditivo positivo. <sup>3</sup>Verdadeiro preditivo negativo.



**Tabela 10** – Métricas de sensibilidade, especificidade, verdadeiro preditivo positivo (VPP), verdadeiro preditivo negativo (VPN), acurácia e coeficiente Kappa para o modelo de regressão logística multinomial (RLM) das fêmeas da raça Large White.

Item	Classes de acordo com produtividade			Acurácia	Kappa <sup>1</sup>
	1 (Baixa)	2 (Moderada)	3 (Alta)		
<i>Regressão Logística Multinomial – Rodada 1</i>				0,367	0,0379
Sensibilidade	0,2632	0,6484	0,1638		
Especificidade	0,7834	0,3820	0,8719		
VPP <sup>2</sup>	0,3593	0,3656	0,3861		
VPN <sup>3</sup>	0,6973	0,6642	0,6795		
<i>Regressão Logística Multinomial – Rodada 2</i>				0,3555	0,0145
Sensibilidade	0,1982	0,6122	0,2188		
Especificidade	0,8709	0,4043	0,7387		
VPP <sup>2</sup>	0,4128	0,3701	0,2849		
VPN <sup>3</sup>	0,7035	0,6458	0,6654		
<i>Regressão Logística Multinomial – Rodada 3</i>				0,3449	-0,0018
Sensibilidade	0,2226	0,5560	0,2176		
Especificidade	0,7479	0,4559	0,7945		
VPP <sup>2</sup>	0,3028	0,3763	0,3113		
VPN <sup>3</sup>	0,6617	0,6350	0,7040		
<i>Regressão Logística Multinomial – Média das três rodadas</i>				0,3558	0,0168
Sensibilidade	0,2280	0,6055	0,2000		
Especificidade	0,8007	0,4474	0,8017		
VPP <sup>2</sup>	0,3576	0,3706	0,3274		
VPN <sup>3</sup>	0,6875	0,6483	0,6829		

<sup>1</sup>Kappa: Coeficiente Kappa. <sup>2</sup>Verdadeiro preditivo positivo. <sup>3</sup>Verdadeiro preditivo negativo.

Wang et al. (2022), ao compararem métodos de *machine learning* (ML) com os métodos convencionais de predição genômica de características reprodutivas em suínos verificaram que, os modelos de ML obtiveram melhor performance e, que mesmo sem diferença significativa, o RF apresentou a menor acurácia entre eles. Além disso, tal estudo demonstrou que o ajuste dos hiper parâmetros é necessário para modelos de ML, sendo que, quando ajustado (*tuning*), o RF aprimorou sua acurácia de predição genômica em até 9,3% para o número de leitões nascidos e até 4% para o número de leitões nascidos vivos.

Tal ajuste dos hiper parâmetros é considerado essencial para modelos de ML (ALVES et al., 2021) e é capaz de impactar na performance de tais modelos (AZODI et al., 2019). Kesbi et al. (2016), ao analisarem diferentes ajustes nos parâmetros do RF verificaram uma menor taxa de erro *out-of-bag* (OOB) e melhor acurácia quando o modelo era ajustado para um número de variáveis amostradas aleatoriamente de 6.000, número de árvores de 1.000 e tamanho do nó igual a 5. Dessa forma, o ajuste dos hiper parâmetros se mostra como uma boa alternativa para aperfeiçoar a acurácia do modelo RF apresentado neste estudo.

Adicionalmente, Alsahaf et al. (2018), ao utilizarem o modelo *random forest* para predição da idade de abate de suínos LW quando os mesmos atingem 120 kg e quantificação da importância das variáveis, utilizaram uma redução de dimensionalidade em 2 etapas (projeção aleatória e análise dos componentes principais) para diminuir a dimensão das características que, segundo os autores, pode acarretar em problemas no momento do modelo selecionar as características mais importantes e também ocasionar em maior demanda computacional. Além disso, tal redução também considerou que as características de pedigree e as semelhanças genéticas do pedigree possuem altas correlações. Dessa forma, tal redução de dimensionalidade afim de reduzir o impacto de características altamente correlacionadas demonstra ser uma segunda alternativa para aprimorar a performance do modelo proposto neste estudo.

#### **5.4. Limites sugeridos para análise da leitegada**

Para a escolha da fêmea com base nas características da leitegada de seu nascimento, os limites sugeridos afim de obter matrizes mais produtivas no futuro foram obtidos pela observação das Figuras 8 a 11.

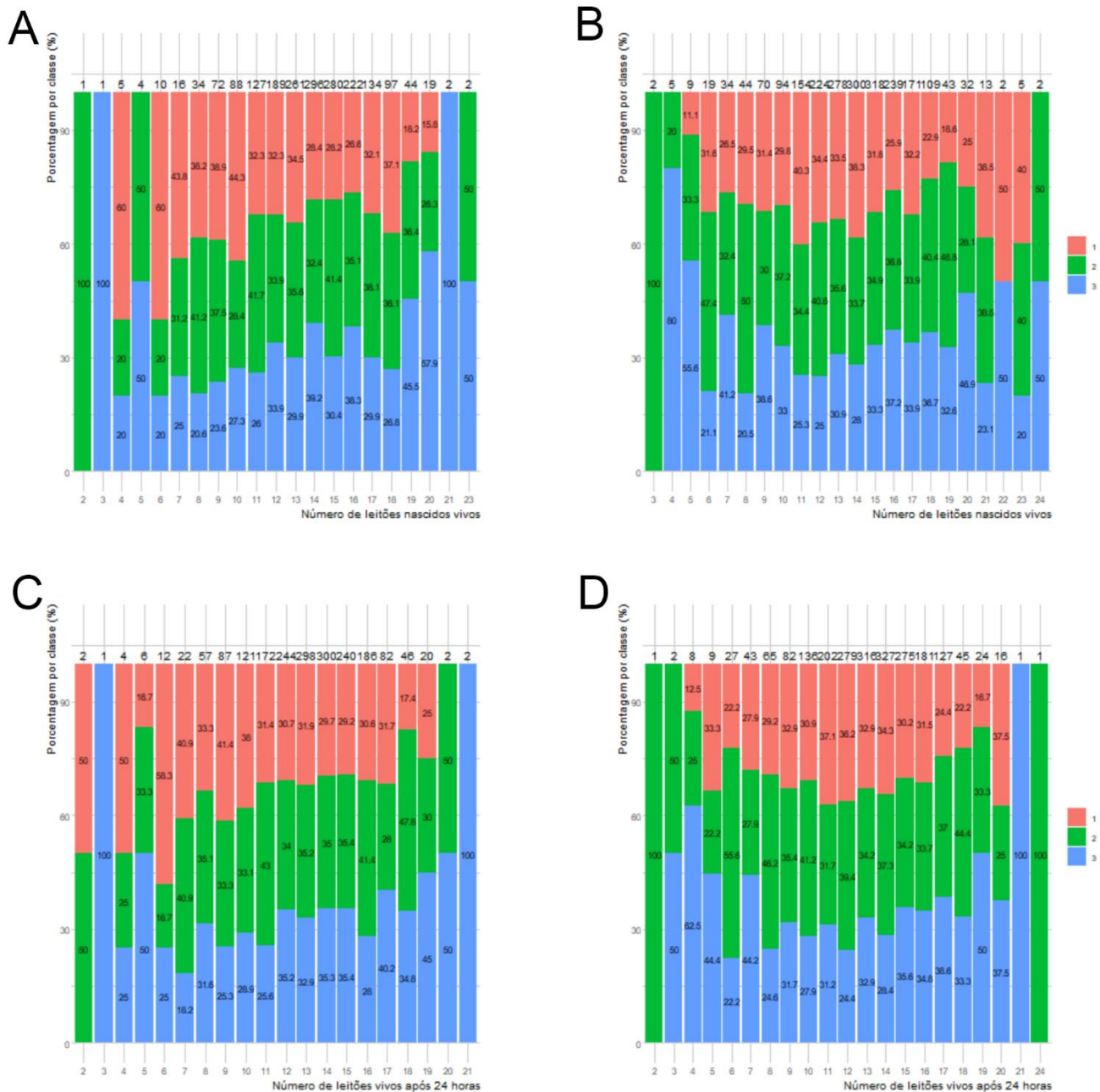
#### **5.4.1 Características relacionadas ao tamanho da leitegada**

A distribuição do número de leitões nascidos vivos para as fêmeas LA (Figura 8A) indicam que um valor mínimo de 12 animais pode ser interessante, uma vez que a frequência de animais das classes de alta (3) e moderada (2) produtividade atingem 33,9% cada um. Além disso, conforme esse número diminui há maior frequência de animais da classe de menor desempenho (1). Para as fêmeas LW (Figura 8B), ao observar a distribuição dos dados, o mínimo sugerido seria de 13 leitões nascidos vivos.

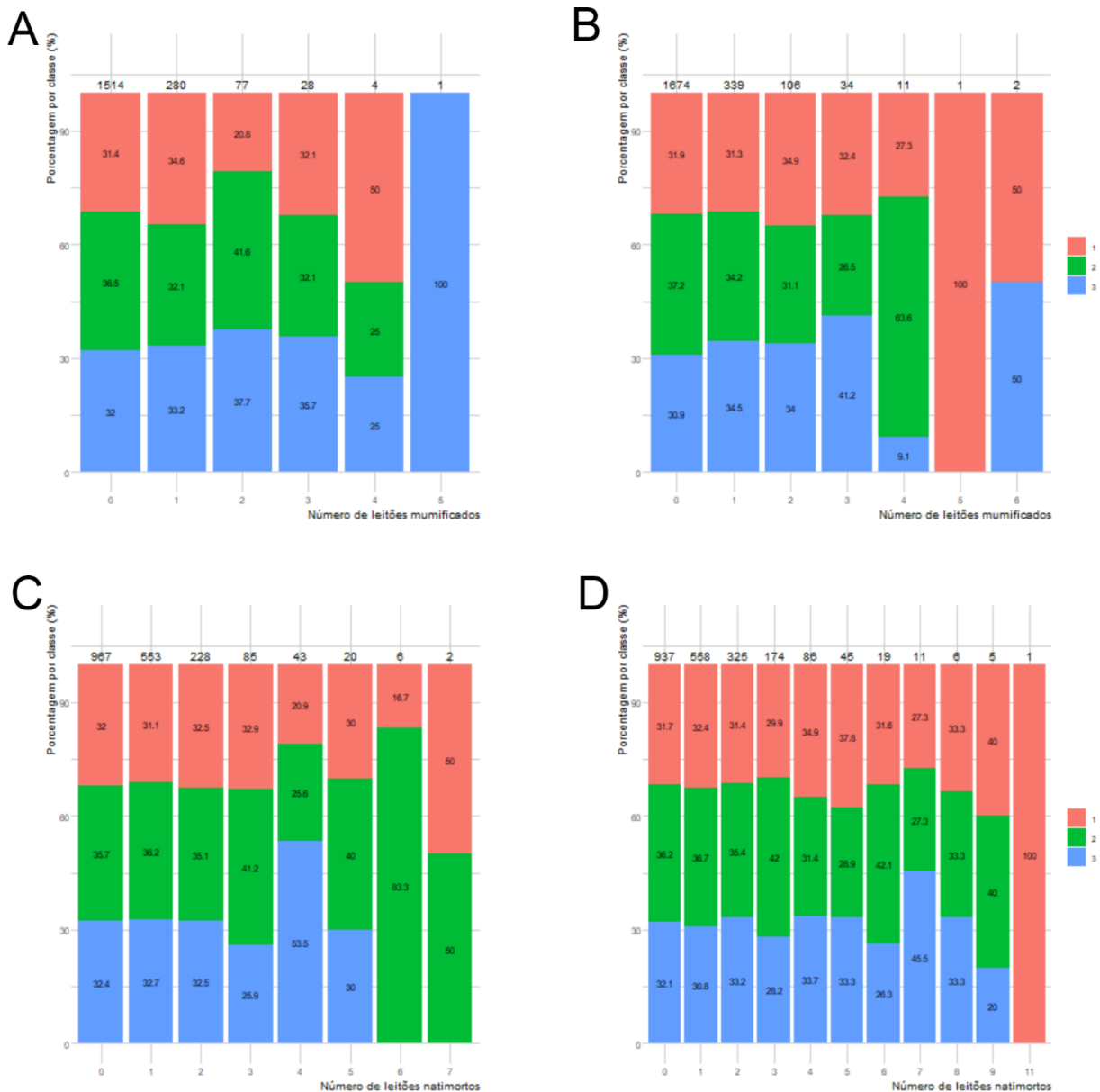
Para a característica número de leitões vivos após 24 horas, o mínimo para as fêmeas LA (Figura 8C) de 12 animais pode ser sugerido, pois acima deste limite há maior predominância de fêmeas de alta produtividade (3) e, abaixo de 12, as fêmeas de produtividade baixa (1) são mais frequentes. Com relação as fêmeas LW (Figura 8D), a partir de 13 animais há um aumento das fêmeas de produtividade alta (3) com relação as de produtividade baixa (1) em um maior número de animais.

Na distribuição do número de leitões mumificados é possível verificar que a grande maioria dos animais possuem baixo número de animais mumificados, sendo que o total de 1.794 fêmeas (94,22%) da raça LA apresentam de 0 a 1 leitões mumificados e, na raça LW, esse número foi de 2.013 (92,89%). Dessa forma, para essa característica nas raças LA e LW (Figura 9A e 9B) valores de no máximo 2 e 3 animais, respectivamente, podem ser interessantes, pois a partir destes, o número de fêmeas da classe de produtividade baixa (1) se sobrepões sobre o das fêmeas de produtividade alta (3).

O número de natimortos também demonstrou maior concentração de animais quanto menor o número de natimortos, sendo que 1.748 (91,80%) das fêmeas LA e 1.820 (83,98%) das fêmeas LW apresentaram valores menor ou igual a 2 leitões natimortos. Dessa maneira, de acordo com as distribuições para LA e para LW (Figura 9C e 9D), os limites que podem ser sugeridos como mais seguros para se ter uma animal produtivo no futuro são de no máximo 2 para LA e 4 para LW.



**Figura 8** – Frequência de distribuição das classes com relação ao número de leitões nascidos vivos de fêmeas Landrace (A) e Large White (B) e com relação ao número de leitões vivos após 24 horas para fêmeas Landrace (C) e Large White (D) marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.



**Figura 9** – Frequência de distribuição das classes com relação ao número de leitões mumificados de fêmeas Landrace (A) e Large White (B) e com relação ao número de leitões natimortos para fêmeas Landrace (C) e Large White (D) marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.

Em um estudo realizado por Rekiel et al. (2013), no qual avaliaram a relação entre diferentes tamanhos de leitegada de origem da matriz com sua produção futura de número de leitões nascidos vivos e número de leitões aos 21 dias de idade, foi observado que para a seleção de fêmeas Landrace polonesas e Large White polonesas, as mesmas devem nascer em leitegadas de no mínimo 13 e 12 leitões, respectivamente. Além disso, os mesmos autores verificaram que as fêmeas advindas

de leitegadas pequenas (com até 9 e/ou 10 leitões), também irão produzir menores leitegadas no futuro.

Adicionalmente, Warda et al. (2021) verificaram que matrizes advindas de leitegadas pequenas (até 9 leitões) apresentavam o primeiro parto em idade mais avançada e uma menor eficiência com relação ao número de leitões nascidos vivos até os 100 dias de vida produtiva da fêmea. Os mesmos autores observaram que matrizes de leitegadas maiores (a partir de 14 leitões) além de mais eficientes, possuíam leitegadas maiores e maior vida produtiva. Resultados semelhantes foram encontrados por Ptak et al. (2016), que relataram que o número de leitões foi maior para fêmeas que nasceram em leitegadas maiores, sendo que, os autores concluíram que tal característica deve ser utilizada em programas de seleção escolhendo fêmeas advindas de leitegadas com pelo menos 12 ou 13 animais.

Com relação a sobrevivência dos leitões, estudos indicam que o tamanho da leitegada indiretamente possui relações opostas com a sobrevivência dos mesmos (BIDANEL et al., 2011), sendo que resultados já demonstraram que mais de 90% dos leitões mais pesados (acima de 1kg de peso vivo) sobreviveram nas primeiras 24 horas de vida, enquanto que 33% dos leitões da classe mais leve (menor que 0,61kg) morreram (QUINIOU et al., 2002), demonstrando a importância da avaliação dessa característica na escolha das futuras matrizes. Além disso, Nam e Sukon (2020) ao analisarem as características associadas ao número de leitões natimortos de leitegadas de fêmeas Landrace x Large White, verificaram aumento da natimortalidade na primeira parição e também quando a fêmea atingia mais que 4 partos, um tamanho de leitegada maior ou igual a 9, menos de 114 dias de gestação, e duração do parto maior ou igual a 5 horas. Udomchanya et al. (2019) também observaram influência significativa do peso ao nascimento dos animais e sua relação com o número de leitões natimortos, sendo o baixo peso ao nascimento um fator de risco. Além disso, estes autores verificaram que fêmeas com igual ou mais de 3 natimortos por leitegada apresentam um maior tempo de duração do parto. Com relação a mortalidade pós-natal, Le Dividich et al. (2017) relataram que as principais causas são o peso ao nascimento e a ingestão de colostro em quantidades insuficientes.

Outro fator importante relacionado ao desempenho reprodutivo e que possui relação com o tamanho da leitegada é o peso da mesma (PIRES, 1999), sendo que, a constante seleção com base no aumento do tamanho da leitegada para melhorar a

prolificidade das fêmeas suínas acarretou em um decréscimo no peso médio dos leitões ao nascimento, afetando a sobrevivência e desempenho futuros dos mesmos (QUINIOU et al., 2002).

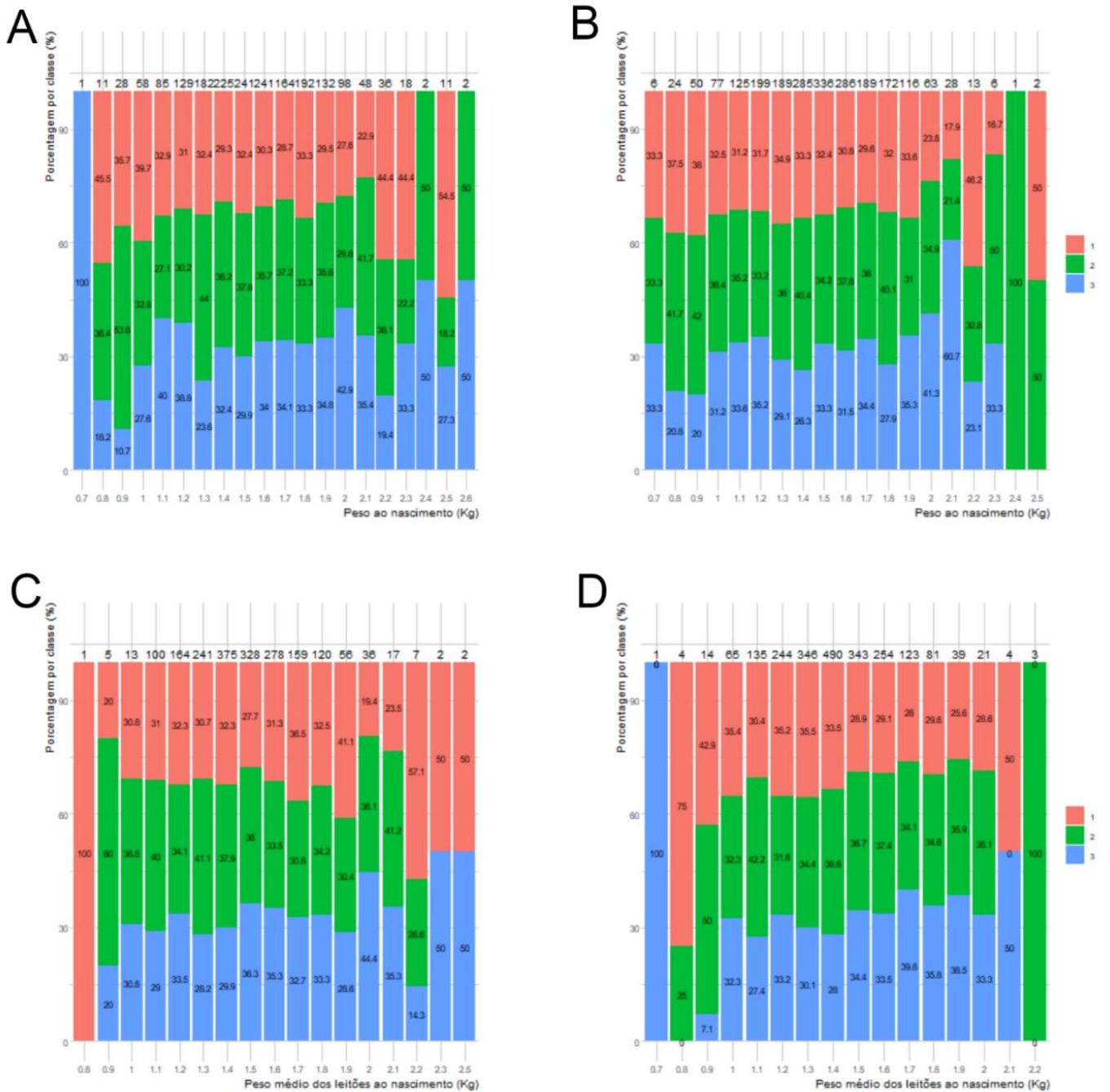
#### **5.4.2 Características relacionadas ao peso da leitegada**

A frequência de animais das classes 1, 2 e 3 com relação ao peso ao nascimento dos mesmos pode ser observada na Figura 10A e 10B. Para as fêmeas LA, o valor a partir de 1,4 kg e, para as fêmeas LW, a partir de 1,2 kg podem ser sugeridos como interessantes para obter animais mais produtivos futuramente. A distribuição de frequências do peso médio dos leitões ao nascimento da leitegada de origem da matriz, sugerem que limites mínimos para a escolha das fêmeas LA e LW (Figura 10C e 10D) podem ser de 1,3 kg e 1,2 kg, respectivamente.

Ao verificar o desvio padrão do peso médio ao nascimento, os valores de no máximo 0,3 para as matrizes LA (Figura 11A), o qual contém os dados de 79,52% dos animais e 0,4 para as matrizes LW (Figura 11B), contendo 97,74% dos animais, são sugeridos como limites que podem ser adotados.

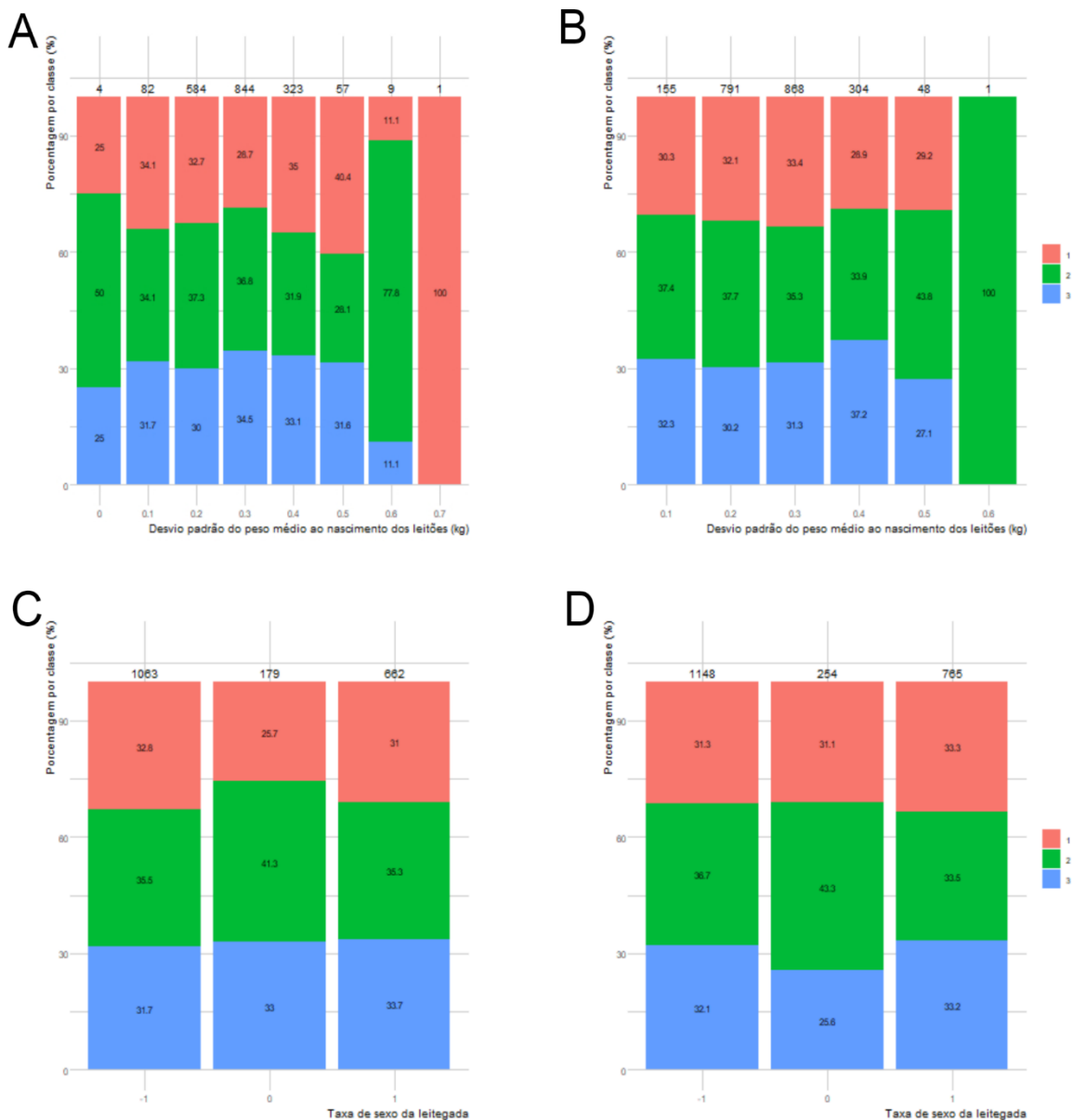
Neste contexto de características relacionadas ao peso da leitegada, Magnabosco et al., (2016), ao avaliar características em marrãs que foram afetadas pelo PN, verificaram que o peso e o ganho de peso médio diário na puberdade, a idade em que foram expostas ao cachaço, e o peso e ganho de peso médio diário no acasalamento variavam de acordo com o PN das mesmas. Além disso, também foi observado que as fêmeas mais leves produziram menos leitões ao longo de três partos e também permaneceram por menor tempo no rebanho.

No mesmo sentido, em outro estudo verificando as características das leitegadas de origem das marrãs, Vallet et al., (2016) encontraram associações do peso ao nascimento e da taxa de crescimento pré-desmame com a característica de idade a puberdade, de forma a refletir no desempenho reprodutivo futuro. Os resultados do estudo citado indicaram que, ao combinar os efeitos de maior peso ao nascimento com um crescimento mais lento no período pré-desmame, as fêmeas têm sua idade a puberdade atrasada. Além disso, estes autores sugerem haver uma relação entre o PN e a capacidade uterina da fêmea suína.



**Figura 10** – Frequência de distribuição das classes com relação ao peso ao nascimento de fêmeas Landrace (A) e Large White (B) e com relação ao peso médio dos leitões ao nascimento para fêmeas Landrace (C) e Large White (D) marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.





**Figura 11** – Frequência de distribuição das classes com relação ao desvio padrão do peso médio ao nascimento de fêmeas Landrace (A) e Large White (B) e com relação a taxa de sexo da leitegada para fêmeas Landrace (C) e Large White (D) marcadas de acordo com a classe de baixa (1), moderada (2) e alta (3) produtividade.

O peso ao nascimento também possui influência no peso corporal futuro, sendo que, animais com maior PN possuem um ganho diário mais rápido e resultam em animais mais pesados no futuro (FIX, 2010b). Com base nisso, através da contagem de corpos lúteos dos ovários de fêmeas suínas Landrace x Large White ( $148 \pm 1.6$  kg de peso corporal) a fim de obter o número de ovulações dos animais, foi verificada uma correlação significativa entre o número de ovulações e o peso corporal das mães, sendo que fêmeas pesando entre 141 a 150 kg obtiveram mais ovulações do

que as que pesavam menos ou igual a 130 kg. (TUMMARUK e KESDANGSAKONWUT, 2015). Dessa forma, sugere-se que quanto maior o peso corporal, o qual é influenciado pelo peso ao nascimento dos animais, maior seria o número de ovulações.

Ao estudar a variação de peso dentro da leitegada, Quesnel et al. (2008) verificaram que o tamanho da leitegada, parição, data de nascimento da matriz e estação durante a concepção explicam apenas 20% desta variação, a qual os autores atrelaram ser devida majoritariamente a outros fatores, como genótipo do embrião e a fatores que afetam o desenvolvimento do embrião e feto.

#### **5.4.3 Características relacionadas a proporção de sexo existente na leitegada**

No estudo da frequência dos animais com relação a taxa de sexo da leitegada, foi verificado que tanto para as fêmeas LA quanto LW (Figura 11C e 11D), tal característica apresentou distribuição semelhante para os animais nas classes de alta, moderada e baixa produtividade. Estudos demonstraram que andrógenos de fetos masculinos podem atravessar as membranas fetais e masculinizar as fêmeas em diversas espécies de mamíferos, afetando sua reprodução, comportamento e fenótipo, porém, esses efeitos não foram muito estudados em suínos (SEYFANG et al., 2018) e, conforme analisado neste estudo, tais fatores não interferiram na distinção das classes dos animais. Em contrapartida, de acordo com Seyfang et al. (2017), quando comparadas fêmeas nascidas em leitegadas com predominância de machos e as nascidas em leitegadas com predominância de fêmeas, as primeiras demoram mais tempo, em horas, para iniciarem o aumento do LH circulante, assim como, o intervalo de tempo do início até atingir o pico de LH foi menor nestas fêmeas nascidas em leitegadas com predominância de machos.

Da mesma forma, em um estudo realizado por Lamberson et al. (1988) em que os autores analisaram a idade ao primeiro estro de fêmeas nascidas em leitegadas com diferentes proporções de sexo, foi observado que o maior número de machos na leitegada promovia um declínio na idade ao primeiro parto. Entretanto, devido a um baixo coeficiente de determinação da análise citada, os autores não definiram essa característica como um critério para seleção, assim como foi determinado no presente estudo. Diferentemente, Rekiel et al. (2012), verificaram que matrizes Large White polonesas nascidas em leitegadas com mais de 70% de fêmeas obtiveram maior

número de leitões nascidos em comparação a matrizes nascidas em leitegadas com até 60% de fêmeas, sendo 8% ( $P \leq 0.05$ ), 9,4% ( $P \leq 0.01$ ) e 6,6% de leitões a mais do que os grupos com menos de 40%, de 40 a 50% e de 50 a 60% de fêmeas, respectivamente.

#### 5.4.4. Limites de escolha e ganho genético estimado por geração

Os limites de todas as características da leitegada de origem da matriz são apresentados na Tabela 11 para a raça LA e a raça LW.

**Tabela 11** – Limites estabelecidos para as características da raça Landrace e Large White.

<b>Característica</b>	<b>Limite estabelecido para fêmeas Landrace</b>	<b>Limite estabelecido para fêmeas Large White</b>
PN <sup>1</sup> (kg)	≥ 1,40	≥ 1,20
PMN <sup>2</sup> (kg)	≥ 1,30	≥ 1,20
DPN <sup>3</sup>	≤ 0,30	≤ 0,40
Nat <sup>4</sup>	≤ 2	≤ 4
Mum <sup>5</sup>	≤ 2	≤ 3
NLV <sup>6</sup>	≥ 12	≥ 13
V24 <sup>7</sup>	≥ 12	≥ 13
TS <sup>8</sup>	-	-

<sup>1</sup>PN: peso ao nascimento. <sup>2</sup>PMN: peso médio dos leitões ao nascimento. <sup>3</sup>DPN: desvio padrão do peso ao nascimento. <sup>4</sup>Nat: número de leitões natimortos. <sup>5</sup>Mum: número de leitões mumificados. <sup>6</sup>NLV: número de leitões nascidos vivos. <sup>7</sup>V24: número de leitões vivos após 24 horas. <sup>8</sup>TS: taxa de sexo predominante da leitegada.

Dessa forma, com base nos limites sugeridos para todas as características da leitegada de origem da matriz da raça LA e considerando uma baixa taxa de reposição das matrizes da granja, o número de animais que atendem esses critérios com relação ao conjunto total foi de 405 (21%) para animais da raça LA. Destes, o ganho genético estimado por geração foi de 0,43 para número de leitões vivos após 24 horas; 0,46 para número de leitões nascidos vivos; 0,04 para peso médio dos leitões ao nascimento; -0,16 para o desvio padrão do peso médio ao nascimento dos leitões; -0,02 para número de leitões mumificados; -0,16 para número de leitões natimortos. Considerando estes animais que passaram pelos critérios, 72,8% são da classe 2 e

3, ou seja, animais de produtividade alta e moderada. Caso o produtor obtenha baixa porcentagem de animais que atendem os critérios, como observado neste conjunto de dados, sugere-se considerar somente as características de PN, PMN e NLV. Utilizando apenas as três características, maior número de animais fica disponível para a escolha (902 animais - 47%), sendo que, dos que atendem o critério, a proporção de fêmeas que pertencerem as classes 2 e 3 nestes dados foi de 71%. Além disso, ao utilizar as três características, seus ganhos genéticos estimados por geração ficam de 0,03 para peso médio dos leitões ao nascimento e 0,3 para número de leitões nascidos vivos.

Da mesma maneira, para a raça LW, o número de animais que atendem os critérios para todas as características com relação ao conjunto total foi de 864 (40%). Destes, o ganho genético estimado por geração foi de 0,30 para número de leitões vivos após 24 horas; 0,35 para número de leitões nascidos vivos; 0,08 para peso médio dos leitões ao nascimento; -0,08 para o desvio padrão do peso médio ao nascimento dos leitões; -0,02 para número de leitões mumificados; -0,14 para número de leitões natimortos. Ao considerar estes animais que passaram por todos os critérios, 69% são da classe 2 e 3, ou seja, animais de produtividade alta e moderada. Diante da necessidade de obter maior quantidade de animais devido a taxa de reposição, é possível considerar somente as características de PN, PMN e NLV, pois haverá um maior número de animais para a escolha (1069 animais - 49%), sendo que, neste conjunto de dados a proporção que pertenceram as classes 2 e 3 foi de 69%. Além disso, ao utilizar as três características, seus ganhos genéticos estimados por geração ficam de 0,07 para peso médio dos leitões ao nascimento e 0,30 para número de leitões nascidos vivos.

Por fim, é importante destacar que o produtor pode considerar os valores sugeridos, porém o mesmo deve realizar uma verificação sistêmica das futuras matrizes envolvendo todas as características apresentadas no estudo, como também análises de outras características importantes e que o mesmo julgar necessárias. Sendo assim, a escolha das melhores futuras matrizes poderá ser realizada de maneira mais segura por meio de características fenotípicas dos animais.

## 6. CONCLUSÕES

Os limites sugeridos para a leitegada de origem da matriz para a escolha de futuras reprodutoras dentro do rebanho das raças Landrace e Large White foi de peso ao nascer maior ou igual a 1,40 e 1,20 kg, peso médio dos leitões ao nascimento a partir de 1,30 e 1,20 kg, desvio padrão do peso médio ao nascimento dos leitões menor ou igual a 0,30 e 0,40, número de leitões natimortos menor ou igual a 2 e 4, número de leitões mumificados menor ou igual a 2 e 3, NLV a partir de 12 e 13 e V24 maior ou igual a 12 e 13, respectivamente. A característica de taxa de sexo da leitegada de nascimento das marrãs não se apresentou como opção de definição de limite. O modelo de *random forest* para predição da produção futura das fêmeas suínas das raças Large White (LW) e Landrace (LA) com base no peso ao nascimento das mesmas e outras características de interesse demonstrou-se mais eficiente em predizer os animais que não pertencem a classe analisada do que predizer os animais que realmente pertencem. Sugere-se novos estudos utilizando os ajustes dos parâmetros e redução da dimensionalidade dos dados com o uso da projeção aleatória e ACP afim aprimorar a acurácia dos modelos apresentados.

## 7. REFERÊNCIAS

ALMEIDA, F. R. C. L., LAURENSSEN, B., PEREIRA, L. X., TEERDS, K. J., & SOEDE, N. M. (2015). **Effects of birthweight on reproductive system development and onset of puberty in gilts.** *Reproduction, Fertility and Development*. doi:10.1071/rd15036

ALSAHAF, Ahmad et al. **Assigning pigs to uniform target weight groups using machine learning.** In: *Proceedings of the World Congress on Genetics Applied to Livestock Production*, vol. Species-Porcine. p. 112. 2018a.

ALSAHAF, Ahmad et al. **Prediction of slaughter age in pigs and assessment of the predictive value of phenotypic and genetic information using random forest.** *Journal of animal science*, v. 96, n. 12, p. 4935-4943, 2018a.

ALVES, A. A. C. et al. **Genome-enabled prediction of reproductive traits in Nellore cattle using parametric models and machine learning methods.** *Animal Genetics*, v. 52, n. 1, p. 32-46, 2021.

ARCHONTOULIS, Sotirios V.; MIGUEZ, Fernando E. **Nonlinear regression models and applications in agricultural research.** *Agronomy Journal*, v. 107, n. 2, p. 786-798, 2015.

AZODI, Christina B. et al. **Benchmarking parametric and machine learning models for genomic prediction of complex traits.** *G3: Genes, Genomes, Genetics*, v. 9, n. 11, p. 3691-3702, 2019.

BAKOEV, Siroj et al. **PigLeg: prediction of swine phenotype using machine learning.** *PeerJ*, v. 8, p. e8764, 2020.

BASAK, Jayanta Kumar et al. **Artificial neural networks and multiple linear regression as potential methods for modelling body surface temperature of pig.** *Journal of Applied Animal Research*, v. 48, n. 1, p. 207-219, 2020.

BIDANEL, J.. ROTHSCHILD, Max F. and Anatoly Ruvinsky, eds. "**Biology and Genetics of Reproduction.**" **Genetics of the Pig, 2nd Edition.** Cambridge, MA: CAB International, 2011. 218-241.

BORCHERS, M. R. et al. **Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle.** Journal of dairy science, v. 100, n. 7, p. 5664-5674, 2017.

BREIMAN, Leo. **Bagging predictors.** Machine learning, v. 24, n. 2, p. 123-140, 1996.

BREIMAN, Leo. **Out-of-bag estimation.** 1996b.

BREIMAN, Leo. **Random forests.** Machine learning, v. 45, n. 1, p. 5-32, 2001.

CHARBUTY, Bahzad; ABDULAZEEZ, Adnan. **Classification based on decision tree algorithm for machine learning.** Journal of Applied Science and Technology Trends, v. 2, n. 01, p. 20-28, 2021.

CORMEN, Thomas. **Desmistificando algoritmos.** Elsevier Brasil, 2017.

DA ROSA, Lorena Silva et al. **Fatores que afetam as características produtivas e reprodutivas de fêmeas suínas.** Boletim de Indústria Animal, v. 71, n. 4, p. 381-395, 2014.

DASGUPTA, Abhijit et al. **Brief review of regression-based and machine learning methods in genetic epidemiology: the Genetic Analysis Workshop 17 experience.** Genetic epidemiology, v. 35, n. S1, p. S5-S11, 2011.

DE ALMEIDA, Mirian et al. **Effect of birth weight and litter size on the performance of landrace gilts until puberty.** Acta Scientiae Veterinariae, v. 42, n. 1, p. 1-8, 2014.

DE MORAES GIACOMELI, Alisson Bruno; KOZICKI, Luiz Ernandes; DE CARVALHO, Naréu Simas. **Emprego de diferentes protocolos de inseminação artificial (IA) em suínos das raças Landrace e Large White.** Revista Acadêmica Ciência Animal, v. 8, n. 3, p. 333-339, 2010.

FAO. **Food Outlook – Biannual Report on Global Food Markets.** Roma, 2022.

FIX, J. S. et al. **Effect of piglet birth weight on body weight, growth, backfat, and longissimus muscle area of commercial market swine.** *Livestock Science*, v. 127, n. 1, p. 51-59, 2010b.

FIX, J. S. et al. **Effect of piglet birth weight on survival and quality of commercial market swine.** *Livestock Science*, v. 132, n. 1-3, p. 98-106, 2010a.

GHAFOURI KESBI, Farhad et al. **Tuning and application of random forest algorithm in genomic evaluation.** *Research On Animal Production (Scientific and Research)*, v. 7, n. 13, p. 185-178, 2016.

GONZALEZ, Leandro de Azevedo. **Regressão logística e suas aplicações.** 2018.

GUIMARÃES, Diego Duque et al. **Suinocultura: estrutura da cadeia produtiva, panorama do setor no Brasil e no mundo e o apoio do BNDES.** *BNDES Setorial*, Rio de Janeiro, n. 45, p. [85] -136, mar. 2017. Disponível em: Acesso em: 07 de dez. 2021.

HILL, R.K. **What an Algorithm Is.** *Philos. Technol.* 29, 35–59 (2016).

IRGANG, Renato. Raças e linhagens na produção de suínos. In. INTEGRALL SOLUÇÕES EM PRODUÇÃO ANIMAL (Org.); ASSOCIAÇÃO BRASILEIRA DE CRIADORES DE SUÍNOS (Org.). **Produção de suínos: teoria e prática.** 1. ed. Brasília, DF, 2014. p. 51-92.

JOAQUIM, Letícia Borges. **Métodos paramétricos e não paramétricos para a predição de valores genéticos genômicos de características de importância econômica em suínos.** 2019.

KECELI, Ali Seydi et al. **Development of a recurrent neural networks-based calving prediction model using activity and behavioral data.** *Computers and electronics in agriculture*, v. 170, p. 105285, 2020.

KENNEDY, Emily Blair et al. **Estimates of Variance Components for Genetic Correlations Among Swine Litter Quality Traits.** 2017.



LAMBERSON, W. R. et al. **Effect of sex ratio of the birth litter on subsequent reproductive performance of gilts.** Journal of animal science, v. 66, n. 3, p. 595-598, 1988.

LE DIVIDICH, Jean; CHARNECA, Rui; THOMAS, Françoise. **Relationship between birth order, birth weight, colostrum intake, acquisition of passive immunity and pre-weaning mortality of piglets.** Spanish Journal of Agricultural Research, v. 15, n. 2, p. e0603, 2017.

LI, Bo et al. **Genomic prediction of breeding values using a subset of SNPs identified by three machine learning methods.** Frontiers in genetics, v. 9, p. 237, 2018.

LIAW, Andy; WIENER, Matthew. **Classification and regression based on a forest of trees using random inputs, based on Breiman (2001).** R Doc. Package Randomforest, v. 4, p. 14, 2018.

LIU, Yanli; WANG, Yourong; ZHANG, Jian. **New machine learning algorithm: Random forest.** In: International Conference on Information Computing and Applications. Springer, Berlin, Heidelberg, 2012. p. 246-252.

MAGNABOSCO, D. et al. **Low birth weight affects lifetime productive performance and longevity of female swine.** Livestock Science, v. 184, p. 119-125, 2016.

MAHESH, Batta. **Machine learning algorithms-a review.** International Journal of Science and Research (IJSR).[Internet], v. 9, p. 381-386, 2020.

MARTINEZ-TABOADA, FERNANDO; REDONDO, JOSE IGNACIO (2020): **Variable importance plot (mean decrease accuracy and mean decrease Gini).** PLOS ONE. Figure. <https://doi.org/10.1371/journal.pone.0230799.g002>

MATTOS, T. do B. **Modelos não lineares e suas aplicações.** Juiz de Fora. 58p. Monografia, 2013.

MILLS, Kayla M. et al. **Evaluation of on-farm indicators of gilt reproductive performance potential at 21 days of age.** Translational Animal Science, v. 4, n. 4, p. txaa210, 2020.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. **Conceitos sobre aprendizado de máquina.** Sistemas inteligentes-Fundamentos e aplicações, v. 1, n. 1, p. 32, 2003.

NAM, Nguyen Hoai; SUKON, Peerapol. **Risk factors associated with stillbirth in swine farms in Vietnam.** World, v. 10, n. 1, p. 74-79, 2020.

NEETHIRAJAN, Suresh. **The role of sensors, big data and machine learning in modern animal farming.** Sensing and Bio-Sensing Research, v. 29, p. 100367, 2020.

NICOLA, Márcio José. **Adoção de random forest e regressão linear para previsão de falhas em equipamentos agrícolas.** 2021. Tese de Doutorado. Universidade de São Paulo.

OLIVE, David J. **Multiple linear and 1D regression.** David J. Olive, 2010.

PEREIRA, Jonas Carlos Campos. **Melhoramento genético aplicado à produção animal.** 6. ed. Belo Horizonte: FEPMVZ, 2012. 758 p. ISBN 9788587144461.

PEREIRA, Jonas Carlos Campos. **Melhoramento genético aplicado à produção animal.** 6. ed. Belo Horizonte: FEPMVZ, 2012. 758 p. ISBN 9788587144461.

PERROCA, Márcia Galan; GAIDZINSKI, Raquel Rapone. **Avaliando a confiabilidade interavaliadores de um instrumento para classificação de pacientes: coeficiente Kappa.** Revista da Escola de Enfermagem da USP, v. 37, p. 72-80, 2003.

PIRES, Aldrin Vieira. **Avaliação genética de características reprodutivas em suínos.** 1999.

PONTE, Caio; CAMINHA, Carlos; FURTADO, Vasco. **Otimização de florestas aleatórias através de ponderação de folhas em árvore de regressão.** In: Anais do XVII Encontro Nacional de Inteligência Artificial e Computacional. SBC, 2020. p. 698-708.

PTAK, J. et al. **Reproductive value of Polish Landrace sows depending on size and number of litter in which the sow was born.** Roczniki Naukowe Zootechniki, v. 43, n. 1, p. 41-49, 2016.

QUESNEL, Helene et al. **Influence of some sow characteristics on within-litter variation of piglet birth weight.** Animal, v. 2, n. 12, p. 1842-1849, 2008.

QUINIOU, Nathalie; DAGORN, J.; GAUDRÉ, D. **Variation of piglets' birth weight and consequences on subsequent performance.** Livestock production science, v. 78, n. 1, p. 63-70, 2002.

R CORE TEAM. **R: A language and environment for statistical computing.** Viena: R Foundation for Statistical Computing. Available: <http://www.R-project.org/>. Acesso em: 16 março 2022.

REKIEL, Anna et al. **Effect of sex ratio in the litter in which Polish Large White and Polish Landrace sows were born on the number of piglets born and reared.** Annals of Animal Science, v. 12, n. 2, p. 179, 2012.

REKIEL, Anna et al. **Effect of size of the litter in which Polish Landrace and Polish Large White sows were born on the number of piglets born and reared.** Roczniki Naukowe Polskiego Towarzystwa Zootechnicznego, v. 9, p. 41-48, 2013.

ROTHSCHILD, M. F.; BIDANEL, Jean Pierre. **Biology and genetics of reproduction.** 1998.

SAXENA, PRATIKSHA. (2012). **Comparison of Linear and Nonlinear Programming Techniques for Animal Diet.** Journal Applied Mathematics. 1. 106-108. 10.5923/j.am.20110102.17.

SCHONLAU, Matthias; ZOU, Rosie Yuyan. **The random forest algorithm for statistical learning.** The Stata Journal, v. 20, n. 1, p. 3-29, 2020.

SEYFANG, J. et al. **The pre-ovulatory luteinising hormone surge is affected by the sex ratio of a gilt's birth litter.** 2017.

SEYFANG, J. et al. **The sex ratio of a gilt's birth litter can affect her fitness as a breeding female.** Animal Production Science, v. 58, n. 9, p. 1567-1574, 2018.

SOUZA, Lucas Fiamoncini de et al. **Predição da fertilidade de reprodutores suínos através de parâmetros de cinética espermática e aprendizado de máquina.** 2022.

SZYNDLER-NEDZA, Magdalena et al. **Prediction of carcass meat percentage in young pigs using linear regression models and artificial neural networks.** Annals of Animal Science, v. 16, n. 1, p. 275, 2016.

TRIBOUT, Thierry et al. **Estimation of realised genetic trends in French Large White pigs from 1977 to 1998 for male and female reproduction traits using stored frozen semen.** In: 7th World Congress on Genetics Applied to Livestock Production. 2002

TUMMARUK, Padet; KESDANGSAKONWUT, Sawang. **Number of ovulations in culled Landrace x Yorkshire gilts in the tropics associated with age, body weight and growth rate.** Journal of Veterinary Medical Science, p. 13-0496, 2015.

UDOMCHANYA, Jatuporn et al. **Association between the incidence of stillbirths and expulsion interval, piglet birth weight, litter size and carbetocin administration in hyper-prolific sows.** Livestock Science, v. 227, p. 128-134, 2019.

UNITED STATES DEPARTMENT OF AGRICULTURE. **Livestock and poultry: world markets and trade.** 2022.

VALLET, J. L. et al. **Litter-of-origin trait effects on gilt development.** Journal of animal science, v. 94, n. 1, p. 96-105, 2016.

VAN KLOMPENBURG, Thomas A.; KASSAHUN, Ayalew. **Data-driven decision making in pig farming: a review of the literature.** Livestock Science, p. 104961, 2022.

WANG, Xue et al. **Using machine learning to improve the accuracy of genomic prediction of reproduction traits in pigs.** Journal of Animal Science and Biotechnology, v. 13, n. 1, p. 1-12, 2022.

WARDA, Agnieszka et al. **The Effect of the Size of the Litter in Which the Sow Was Born on Her Lifetime Productivity.** Animals, v. 11, n. 6, p. 1525, 2021.

WOLF, J.; ŽÁKOVÁ, E.; GROENEVELD, E. **Within-litter variation of birth weight in hyperprolific Czech Large White sows and its relation to litter size traits, stillborn piglets and losses until weaning.** *Livestock Science*, v. 115, n. 2-3, p. 195-205, 2008.

ZEVIANI, W. M.; JÚNIOR, PJR; BONAT, W. H. **Curso-Modelos de regressão não linear.** Laboratório de Estatística e Geoinformação, Departamento de Estatística, Universidade Federal do Paraná, 58º RBRAS e 15º SEAGRO, Campina Grande–PB, 2013.