



**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I - CAMPINA GRANDE
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA
CURSO DE GRADUAÇÃO EM BACHARELADO EM ESTATÍSTICA**

HILLNARA DE PAIVA ALMEIDA FERREIRA CARVALHO

**APLICAÇÃO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA NA VALIDAÇÃO DE
MODELOS DE CRESCIMENTO**

CAMPINA GRANDE - PB

2024

HILLNARA DE PAIVA ALMIDA FERREIRA CARVALHO

**APLICAÇÃO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA NA VALIDAÇÃO DE
MODELOS DE CRESCIMENTO**

Trabalho de Conclusão de Curso apresentado ao Departamento de Estatística da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Estatística.

Área de concentração: Estatística

Orientadora: Profa. Dra. Ana Patrícia Bastos Peixoto de Oliveira

Coorientadora: Profa. Ma. Débora de Sousa Cordeiro

CAMPINA GRANDE

2024

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

C331a Carvalho, Hillnara de Paiva Almeida Ferreira.

Aplicação de técnicas de aprendizado de máquina na validação de modelos de crescimento [manuscrito] / Hillnara de Paiva Almeida Ferreira Carvalho. - 2024.

43 p. : il. colorido.

Digitado. Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2024. "Orientação : Profa. Dra. Ana Patrícia Bastos Peixoto de Oliveira, Coordenação do Curso de Estatística - CCT. " "Coorientação: Profa. Dra. Débora de Sousa Cordeiro , Coordenação do Curso de Estatística - CCT.

1. Modelo de crescimento. 2. Modelo não linear. 3. Aprendizado de máquina. I. Título

21. ed. CDD 519.5

HILLNARA DE PAIVA ALMEIDA FERREIRA CARVALHO

**APLICAÇÃO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA NA VALIDAÇÃO
DE MODELOS DE CRESCIMENTO**

Trabalho de Conclusão de Curso apresentado ao Departamento de Estatística da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de bacharel em Estatística.

Área de concentração: Estatística

Aprovado em: 09 / 09 / 2024.

BANCA EXAMINADORA



Profa. Dra. Ana Patrícia B. P. de Oliveira (Orientadora)
Universidade Estadual da Paraíba (UEPB)



Prof. Hiago José A. de Albuquerque Martins
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. Silvio Fernando Alves Xavier Junior
Universidade Estadual da Paraíba (UEPB)

“A Deus, à minha família e aos meus amigos por
todo o carinho e incentivo.”

AGRADECIMENTOS

A Deus, pois Ele me deu a vida, a capacidade para alcançar mais uma meta, é meu refúgio e minha fortaleza, estando sempre comigo, dando-me força, coragem e demonstrando seu eterno amor e felicidade.

Aos meus pais, Heize e Francisco Almeida Ferreira, e ao meu irmão, Fernando, que continuam e continuarão sendo o meu porto seguro, meu norte e o meu forte. Obrigada pelo amor, carinho, dedicação, incentivo e por mais uma vez acreditarem em mim.

Ao meu esposo, Khevenny Karllos, que acreditou em meu potencial ajudando-me com sua proteção nas suas orações diárias e seu apoio emocional e financeiro.

A minha filha, Nayara de Paiva, que me dá muita alegria, amor e a certeza que dias melhores estão por vir.

A professora Dra. Ana Patricia Bastos Peixoto de Oliveira, orientadora, mentora e amiga, que sempre esteve presente para ajudar nas horas de aflição. Ela acompanhou minha evolução, desde então tem me apoiado. A ela todo meu carinho e amizade.

A todos os meus familiares, por me acolherem com tanto amor e carinho e me apoiar nesta jornada desde o começo.

Aos meus amigos João Vitor e Rivaldo, pela companhia, amizade, apoio, convivência e momentos de descontração.

Aos professores do Departamento de Estatística da Universidade Estadual da Paraíba, responsáveis pela base da minha formação profissional, muito obrigada pela amizade, apoio e confiança.

Aos participantes da banca examinadora pelas valiosas contribuições para a finalização deste trabalho.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

“Uma resposta aproximada para o problema certo vale muito mais do que uma resposta exata para um problema aproximado.”
(John Wilder Tukey, 1962)

RESUMO

A aplicação de métodos de aprendizado de máquina na validação de modelos de crescimento que incorporam medidas de não linearidade apresenta-se como uma abordagem promissora para aprimorar a compreensão e a precisão das previsões de crescimento em diversas áreas de estudo. Ao explorar modelos não lineares, capazes de capturar as complexas relações entre variáveis preditoras e a variável de resposta, torna-se possível obter parâmetros com interpretação biológica direta, facilitando tanto a tomada de decisões quanto a compreensão do fenômeno de crescimento. Este estudo propõe o desenvolvimento de modelos de aprendizado de máquina, utilizando algoritmos de aprendizado de máquina (*Machine Learning*), para validar o ganho de peso (kg) do touro ao longo dos anos. De maneira geral, o modelo de crescimento Logístico apresentou o maior coeficiente de determinação e um dos menores Critério de Informação de Akaike (AIC) fornecendo um ajuste mais adequado aos dados experimentais em comparação com os outros modelos.

Palavras-chaves: modelo de crescimento; modelo não linear; aprendizado de máquina.

ABSTRACT

The application of machine learning methods to validate growth models that incorporate nonlinearity measures is a promising approach to improve the understanding and accuracy of growth predictions in several areas of study. By exploring nonlinear models capable of capturing the complex relationships between predictor variables and the response variable, it becomes possible to obtain parameters with direct biological interpretation, facilitating both decision-making and understanding of the growth phenomenon. This study proposes the development of machine learning models, using machine learning algorithms (*Machine Learning*), to validate the weight gain (kg) of the bull over the years. In general, the Logistic growth model presented the highest coefficient of determination and one of the lowest *Akaike Information Criterion* (AIC), providing a better fit to the experimental data compared to the other models.

Keywords: growth model; nonlinear model; machine learning.

LISTA DE ILUSTRAÇÕES

Figura 1 – Processo de treinamento, validação e teste de modelos.	24
Figura 2 – Processo de treinamento, validação cruzada <i>k-fold</i> ($k = 5$).	25
Figura 3 – Modelos subajustado, apropriado e sobreajustado.	26
Figura 4 – Matriz de confusão.	26
Figura 5 – Curva ROC.	28
Figura 6 – Relação entre o peso do touro <i>Holstein-Friesiande</i> ao longo dos anos.	30
Figura 7 – Ajuste dos modelos Gompertz, Logístico e Von Bertalanffy para o peso (kg) do touro.	33
Figura 8 – Normalidad dos resíduos.	34
Figura 9 – Relação entre valores ajustado e resíduos.	35
Figura 10 – Matriz de confusão.	38
Figura 11 – Curva de ROC.	40

LISTA DE TABELAS

Tabela 1 – Relação do crescimento do peso (kg) do touro Holandês ao longo dos anos.	29
Tabela 2 – Estimativas (E) dos parâmetros (P) para o modelo <i>Logístico</i> , <i>Gompertz</i> e <i>Von Bertalanffy</i> , erro padrão da estimativa (E.P.E.), teste <i>t</i> e seus respectivos intervalos de confiança (IC) a 95%	32
Tabela 3 – Estatísticas de qualidade de informação de ajuste.	32
Tabela 4 – Teste de <i>Shapiro-Wilk</i>	34
Tabela 5 – Estimativas <i>bootstrap</i> (E.B.), vício relativo a estimativa <i>bootstrap</i> (V.R.B.), vício de Box, intervalo de confiança <i>bootstrap</i> a 95% para os parâmetros (p).	36
Tabela 6 – Testes de normalidade <i>bootstrap</i>	36
Tabela 7 – Valores das métricas para os modelos.	39

SUMÁRIO

1	INTRODUÇÃO	11
2	FUNDAMENTAÇÃO TEÓRICA	13
2.1	Modelo de regressão não linear	13
2.2	Propriedades dos estimadores	13
2.3	Métodos de Estimação	14
2.3.1	<i>Método dos Mínimos Quadrados</i>	14
2.3.2	<i>Método dos Mínimos Quadrados Ordinários</i>	15
2.3.3	<i>Método de Máxima Verossimilhança</i>	16
2.3.3.1	<i>Função de Verossimilhança</i>	16
2.3.3.2	<i>Estimador de Máxima Verossimilhança</i>	16
2.4	Métodos Iterativos	17
2.4.1	<i>Método de Gauss-Newton</i>	17
2.4.2	<i>Método Newton-Raphson</i>	17
2.5	Análise de Tendências no Vício	18
2.5.1	<i>Vício de box</i>	19
2.6	Modelos de Crescimento	19
2.6.1	<i>Modelo de regressão Gompertz</i>	19
2.6.2	<i>Modelo de regressão Logístico</i>	20
2.6.3	<i>Modelo de Von Bertalanffy</i>	20
2.7	Método de Bootstrap	21
2.7.1	<i>Bootstrap paramétrico</i>	21
2.7.2	<i>Intervalo de confiança bootstrap</i>	22
2.7.3	<i>Teste de normalidade bootstrap</i>	22
2.8	Machine learning	23
2.9	Seleção e validação	24
2.10	Problemas do aprendizado de máquina	25
2.10.1	<i>Sobreajuste e Subajuste</i>	25
2.10.2	<i>Dados desbalanceados</i>	26
2.11	Métricas de avaliação em aprendizado de máquina	26
3	MATERIAL E MÉTODOS	29
4	RESULTADOS E DISCUSSÃO	32
5	CONCLUSÃO	41
	REFERÊNCIAS	42

1 INTRODUÇÃO

Conforme apontado por França e Thornely (1984), o uso de modelos matemáticos tem sido uma prática tradicional para explicar a relação entre idade e crescimento em animais. Esses modelos são ferramentas cruciais para estimar o ganho de peso e analisar parâmetros biológicos importantes, como as necessidades alimentares diárias e a influência das condições ambientais no crescimento dos animais. Além disso, esses modelos são aplicados para determinar a idade ideal para o abate, a faixa de peso esperada e a capacidade reprodutiva, contribuindo para melhorar o manejo e a produtividade na pecuária. No caso da raça Holandesa (Holstein-Frísia), essa aplicação se torna ainda mais significativa.

Segundo Mayk Alves (2021), o touro Holandês é amplamente reconhecido na pecuária leiteira por sua elevada produção de leite e também desempenha um papel central em programas de melhoramento genético. A utilização de modelos matemáticos para prever o desenvolvimento e o rendimento dessa raça não apenas facilita o controle sobre o ganho de peso, como também potencializa a produtividade reprodutiva e leiteira. Dessa forma, o uso desses modelos matemáticos contribui diretamente para a otimização do manejo dos touros Holandeses, proporcionando estimativas mais precisas em relação ao desenvolvimento ideal para reprodução e produção de leite, gerando impactos positivos na pecuária.

Conforme Ratkowsky (1990), as medidas de peso dos animais ao longo do tempo geralmente seguem um padrão de curva sigmoide, o que pode ser bem descrito por modelos que capturam a relação não linear entre peso e idade. Por isso, para estudar o crescimento biológico de diferentes espécies, os modelos não lineares são mais apropriados, pois conseguem explicar de forma mais precisa o processo de crescimento. Para Souza et al. (2010) e Silva et al. (1998), esses modelos frequentemente proporcionam um bom ajuste com um número menor de parâmetros comparado aos modelos lineares. Além disso, os parâmetros ou suas funções possuem interpretações biológicas claras, o que facilita o estudo do peso dos animais na fase adulta, da velocidade de crescimento e dos pontos críticos de mudança na taxa de crescimento.

De acordo com Obermeyer (2016), o aprendizado de máquina é uma área da ciência da computação que combina métodos matemáticos e estatísticos com algoritmos computacionais. Para Waljee (2010), o aprendizado de máquina faz uso de algoritmos baseados em inteligência artificial e é empregado em situações onde se deseja identificar padrões em um conjunto de variáveis para prever um resultado específico de interesse.

Neste estudo, prever o peso dos touros ao longo do tempo é essencial para um manejo eficaz, principalmente em atividades de reprodução e seleção de reprodutores. Portanto, o objetivo é modelar o crescimento do touro Holandês ao longo de sua vida. É conhecido que, à medida que o touro envelhece, seu peso aumenta, mas esse crescimento não é linear; ele costuma seguir uma curva sigmoide, onde o crescimento é mais rápido nos primeiros anos e se estabiliza quando o animal atinge a maturidade. Partindo deste pressuposto, pretende-se ajustar modelos de crescimento não lineares — como os modelos *Gompertz*, *Logístico* e *Von Bertalanffy* — que

são ideais para capturar a evolução do peso ao longo do tempo, permitindo identificar qual deles melhor representa a dinâmica do crescimento do animal entre 6 meses e 8 anos de idade. No entanto, para garantir a confiabilidade e robustez dessas previsões, técnicas de Machine Learning, como *bootstrap*, matriz de confusão e curva ROC, são utilizadas para avaliar o desempenho preditivo dos modelos, especialmente quando é definido um limiar de peso crítico.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, serão explorados os Modelos Não Lineares com especial atenção aos modelos de crescimento, e os princípios teóricos por trás dos algoritmos de Aprendizado de Máquina.

2.1 Modelo de regressão não linear

Os métodos de regressão linear são amplamente empregados em uma variedade de contextos para analisar a relação entre uma variável dependente e uma ou mais variáveis independentes. No entanto, em certas circunstâncias, um modelo que não segue uma forma linear pode ser mais adequado, dado que muitos fenômenos não se comportam linearmente. O emprego de modelos de regressão não linear é comum entre pesquisadores em diversos campos do conhecimento, como agricultura, biologia, economia, engenharia, química, entre outros. Frequentemente, as formulações de tais modelos são derivadas a partir de suposições teóricas que representam adequadamente o fenômeno em questão e os parâmetros resultantes são passíveis de interpretação. O modelo não linear com erros que seguem uma distribuição normal é uma das abordagens mais adotadas quando a variável resposta é contínua, pois é de fácil implementação e possui um vasto embasamento teórico associado.

Conforme Mazucheli e Achcar (2002), os modelos não lineares desempenham um papel importante, possibilitando aos pesquisadores modelar fenômenos do mundo real que escapam à descrição dos modelos lineares. O conceito de regressão não linear pode ser definido da seguinte maneira:

$$y_i = f(x_i, \theta) + \varepsilon_i, \quad i = 1, \dots, n, \quad (2.1)$$

em que, y_i é a variável resposta, sendo $f(x_i, \theta)$ a função das variáveis preditoras, x_i o vetor de variável explicativas e θ os parâmetros a serem estimados. Os erros, ε_i , são independentes e identicamente distribuídos, seguindo uma distribuição $N \sim (0, \sigma^2 \mathbf{I})$. Segundo Bates e Watts (1988), para os modelos de regressão não linear, pelo menos uma das derivadas parciais da função $f(X_i, \theta)$, em relação aos parâmetros θ_i depende de pelo menos um dos parâmetros θ_i .

2.2 Propriedades dos estimadores

As propriedades dos estimadores são características importantes para Estatística, pois elas ajudam a avaliar a qualidade do estimador. Dentre as varias propriedades temos:

- i) Normalidade: Um estimador é considerado normal, quando sua distribuição é próxima da distribuição Normal;
- ii) Independência: Quando seus valores não são correlacionados entre si;
- iii) Não Viesado: Ocorrera quando um estimador é não tendencioso, isto é, quando o seu valor esperado é igual ao valor real do parâmetro que está sendo estimado;

- iv) Eficiência: Será observada quando um estimador tem a menor variância possível entre todos os estimadores não tendenciosos do mesmo parâmetro;
- v) Consistência: Será consistente quando o valor da estimação converge para o valor real do parâmetro a medida que o tamanho da amostra aumenta;
- vi) Linearidade: O estimador é linear quando expresso como uma combinação linear das observações da amostra.

Conforme explicado por Mazucheli (2002), para modelos que não são lineares, essas propriedades são válidas apenas em termos assintóticos, ou seja, quando o tamanho da amostra é consideravelmente grande. Em geral, para amostras pequenas, essas características não são conhecidas. No entanto, é afirmado que à medida que o tamanho da amostra aumenta, os resultados assintóticos se tornam mais aplicáveis.

De acordo com Bates e Watts (1988), a escolha do melhor estimador para um dado problema é influenciada por diversos aspectos, como a natureza do parâmetro a ser estimado, a quantidade de dados disponíveis e as propriedades desejadas do estimador. Podemos mencionar os seguintes estimadores, de acordo com suas propriedades específicas.

- i) Média Amostral: Que representa um estimador não tendencioso da média populacional;
- ii) Variância Amostral: Estimador não tendencioso da variância populacional;
- iii) Intervalo de Confiança de uma Estimador: É um estimador consistente do parâmetro que está sendo estimado.

Segundo Mattos (2013), para modelar os dados observados de um modelo não linear, é preciso estimar os parâmetros desconhecidos que definem a forma da função não linear. Essa estimativa é geralmente feita utilizando métodos iterativos.

2.3 Métodos de Estimação

Existem várias abordagens para obter estimativas dos parâmetros, incluindo o método dos mínimos quadrados e o método da máxima verossimilhança. No caso de modelos de regressão não lineares, resolver o sistema de equações normais não é simples, exigindo métodos iterativos para obter essas estimativas. Os métodos utilizados para tanto será o de Gauss-Newton e Newton-Raphson. A seguir, serão discutidos alguns métodos de estimação.

2.3.1 Método dos Mínimos Quadrados

Considerando o modelo apresentado em (2.1), a estimativa de mínimos quadrados para θ , indicada como $\hat{\theta}$, é aquela que reduz ao mínimo a soma dos quadrados dos resíduos.

$$SQRes(\theta) = \sum_{i=1}^n [y_i - f(x_i, \theta)]^2, \quad \theta \in \Theta \subset \mathbb{R}^p. \quad (2.2)$$

Para obter o estimador de mínimos quadrados, é essencial realizar a diferenciação da expressão (2.2) em relação a cada um dos parâmetros e então igualar as expressões a zero. Isso implica em,

$$\left. \frac{\partial SQRes(\theta)}{\partial \theta_r} \right|_{\theta=\hat{\theta}} = 0 \quad r = 1, 2, \dots, p. \quad (2.3)$$

$$\sum_{i=1}^n [y_i - f(x_i, \theta)] \left. \frac{\partial f(x_i, \theta)}{\partial \theta_r} \right|_{\theta=\hat{\theta}} = 0 \quad r = 1, 2, \dots, p. \quad (2.4)$$

Seguindo esse método, obtemos o sistema de p equações normais (2.4) para o modelo que não é linear. Segundo Mattos (2013), em muitos casos de modelos não lineares, esses sistema de equações normais não pode ser resolvido de forma analítica, o que requer o uso de métodos iterativos.

2.3.2 Método dos Mínimos Quadrados Ordinários

De acordo com Thomas (2016), a teoria da estimação por mínimos quadrados ordinários, as estimativas obtidas para por esse método para o parâmetro θ em um modelo não linear forma (2.1), são aquelas que minimizam a soma dos quadrados dos erros, dados por:

$$S(\theta) = \sum_{i=1}^n [y_i - f(x_i, \theta)]^2. \quad (2.5)$$

Para encontrar os estimadores que geram tais estimativas, é necessário derivar (2.5) com relação a cada componente do vetor θ e igualar a zero, obtendo-se o seguinte vetor de equações normais:

$$\sum_{i=1}^n [y_i - f(x_i, \theta)] \frac{\partial f(x_i, \theta)}{\partial \theta_j} = 0. \quad (2.6)$$

Para os modelos não lineares, os estimadores $\hat{\theta}$ resultantes de (2.6) normalmente não apresentam forma analítica e as soluções precisam ser encontradas por meio de métodos iterativos de aproximação. Dentre as principais metodologias de otimização utilizadas na estimação de parâmetros, podemos citar o método de Gauss-Newton, método de Davidon-Fletcher-Powell, método de descida mais íngreme (*steepest descent*), método da procura pelo gradiente conjugada, método da procura de Nelder-Mead, algoritmo de Golub-Pereyra, entre outros. A seguir, será descrito o Método de Máxima Verossimilhança.

2.3.3 Método de Máxima Verossimilhança

O Método de Máxima Verossimilhança é uma abordagem fundamental para encontrar a estimativa mais provável do parâmetro desconhecido de uma população, com base em uma amostra. Basicamente, ele procura determinar o valor do estimador que torna a ocorrência da amostra em questão mais provável. Para isso, a função de probabilidade das variáveis aleatórias que representam a amostra é tratada como uma função do parâmetro populacional correspondente, permitindo assim a estimativa desse parâmetro desconhecido. Na próxima seção, será destacada a importância dessa função, conhecida como função de verossimilhança, no processo de estimativa.

2.3.3.1 Função de Verossimilhança

Quando obtemos uma amostra aleatória simples de uma variável x com uma densidade de probabilidade é obtida $f(x, \theta)$, em que θ é um parâmetro desconhecido, introduziu-se a função $L(\theta, x)$. Essa função avalia a probabilidade da amostra com base em diferentes valores para θ , auxiliando na busca pelo valor de θ que melhor descreve os dados. Em essência, procurava-se o θ que maximiza a verossimilhança dos nossos dados amostrais. A expressão da função é dada por:

$$L(\theta, x) = \prod_{n=1}^n f(x_i, \theta), \quad \theta \in \Theta. \quad (2.7)$$

Chamada de Função de Verossimilhança, essa função é influenciada pelo parâmetro populacional, enquanto os valores da amostra permanecem fixos. O logaritmo natural da função de verossimilhança, denotado por $l(\theta, x)$ é conhecido como Função de Log-Verossimilhança, frequentemente preferido devido à sua praticidade em cálculos. Ele é definido como:

$$l(\theta, x) = \log L(\theta, x) = \log \prod_{n=1}^n f(x_i, \theta) = \sum_{i=1}^n \log f(x_i, \theta). \quad (2.8)$$

O logaritmo é uma operação que conserva a ordem dos valores, tornando mais fácil lidar com otimização. Assim, neste cenário, escolhemos utilizar a função de log-verossimilhança em vez da verossimilhança diretamente.

2.3.3.2 Estimador de Máxima Verossimilhança

O estimador de máxima verossimilhança para um parâmetro θ é o valor que torna θ mais provável com base nos dados da amostra observada. Em outras palavras, o estimador de máxima verossimilhança de θ como qualquer valor $\hat{\theta}$, tal que $L(\theta, x)$ é maximizada. Isso implica que,

$$\hat{L}(\theta, x) \geq L(\theta, x) \quad \forall x \in \mathbb{R}, \forall \theta \in \Theta. \quad (2.9)$$

O valor do parâmetro que maximiza a função de verossimilhança é o mesmo que leva a função de log-verossimilhança ao seu ponto mais alto. Isso acontece devido ao fato de que o logaritmo é uma função crescente, então os extremos da função de log-verossimilhança

correspondem aos extremos da função de verossimilhança. Para determinar esse valor, busca-se o ponto onde a primeira derivada se anula e verifica-se se a matriz Hessiana é negativa nesse ponto. Geralmente, a função de log-verossimilhança é côncava, o que implica que o ponto onde a primeira derivada se iguala a zero, é o ponto onde o máximo da função ocorre.

De acordo com Amaral, Souza e Catalan (2015), o método de Newton-Raphson é uma técnica iterativa usada para determinar as raízes de equações que não são lineares. Esse método começa com uma suposição inicial e, com base nela, calcula uma nova estimativa utilizando uma fórmula que incorpora a derivada da função. Esse procedimento é repetido até que a estimativa seja precisa o bastante.

2.4 Métodos Iterativos

Os métodos iterativos consistem em iterar sequencialmente através de cálculos repetidos para se aproximar da solução desejada. Em cada etapa, uma estimativa aprimorada da solução é obtida com base na iteração anterior, e esse processo continua até que uma condição de convergência seja alcançada.

Para reduzir a $SQRes(\theta)$ em (2.2) ou maximizar $L(\theta, \sigma^2)$, é essencial recorrer a algum método numérico iterativo. Diversos métodos desse tipo são sugeridos na literatura, como o Newton-Raphson, Gauss-Newton, Steepest-Descent, Marquardt, entre outros.

2.4.1 Método de Gauss-Newton

Suponha que $\theta^{(a)}$, seja uma estimativa aproximada do mínimo dos quadrados $\hat{\theta}$ de um modelo não linear. Quando θ está próximo de $\theta^{(a)}$, considera-se uma expansão de série de Taylor de primeira ordem,

$$f(x, \theta) \approx f(x, \theta^{(a)}) + F(\theta^{(a)})(\theta - \theta^{(a)}),$$

sabendo que $F(\theta) = \left[\frac{\partial f(x_i, \theta)}{\partial \theta_j} \right]$. Sendo $r(\theta)$ um vetor de resíduos, pode-se escrever

$$r(\theta) \approx r(\theta^{(a)}) - F^{(a)}(\theta - \theta^{(a)}).$$

Substituindo $r(\theta)'r(\theta)$ em $SQRes(\theta)$ pelo método de mínimos quadrados, tem-se que esta aproximação será

$$\theta^{(a+1)} = \theta^{(a)} + [F'^{(a)}F^{(a)}]^{-1}F'^{(a)}r(\theta^{(a)}). \quad (2.10)$$

Segundo Mattos (2013), este processo de Gauss-Newton é repetido até obter a convergência.

2.4.2 Método Newton-Raphson

De acordo com Amaral, Souza e Catalan (2015), o método de Newton-Raphson é uma técnica iterativa usada para determinar as raízes de equações que não são lineares. Esse método

começa com uma suposição inicial e, com base nela, calcula uma nova estimativa utilizando uma fórmula que incorpora a derivada da função. Esse procedimento é repetido até que a estimativa seja precisa o bastante.

$$SQRes(\theta) \approx SQRes(\theta^{(a)}) + U_0(\theta^{(a)})(\theta - \theta^{(a)}) + (\theta - \theta^{(a)})' \frac{H(\theta^{(a)})}{2} (\theta - \theta^{(a)}). \quad (2.11)$$

A aproximação da função de soma de quadrados terá um ponto estacionário quando seu gradiente for igual a zero. Logo,

$$U(\theta) = \frac{\partial SQRes(\theta)}{\partial \theta} = 0.$$

A matriz Hessiana da $SQRes(\theta)$, denominada por $H(\theta)$, deve satisfazer:

$$H(\theta) = \frac{\partial^2 SQRes(\theta)}{\partial \theta \partial \theta'} = 0, \text{ positiva (semi)definida.}$$

Essas circunstâncias sugerem que, para localizar um ponto estacionário da função de quadrados, é preciso encontrar um valor de θ em que o gradiente seja zero ($U(\theta) = 0$) e a matriz Hessiana seja positiva semi-definida. Se $H(\theta^{(a)})$ é positiva definida, o método de Newton-Raphson pode ser:

$$\theta^{(a+1)} = \theta^{(a)} - [H(\theta^{(a)})]^{-1} U(\theta^{(a)}).$$

O procedimento é continuado até que a convergência seja alcançada. Entre as diversas abordagens disponíveis, um método para verificar a convergência poder ser:

$$|SQRes(\theta^{(a+1)}) - SQRes(\theta^{(a)})| < 0.000001. \quad (2.12)$$

De acordo com Thomas (2016), o método de Gauss-Newton é uma versão simplificada do método de Newton-Raphson.

2.5 Análise de Tendências no Vício

Segundo Mazucheli e Achcar (2002), a investigação de tendências no vício é um procedimento estatístico que busca compreender como as variáveis independentes afetam os desvios em relação aos padrões desejados, especialmente no contexto de vício em modelos estatísticos. Esse estudo identifica padrões não lineares nos dados, analisa as relações complexas entre as variáveis e detecta desvios sistemáticos, sendo essencial na validação dos modelos. Essa abordagem oferece uma compreensão mais aprofundada das sutilezas dos dados, o que contribui para uma modelagem mais precisa e embasada.

2.5.1 Vício de box

O conceito de "vício de Box" se refere a uma técnica estatística conhecida como "diagnóstico de Box", a qual é empregada para determinar quais parâmetros têm uma influência mais significativa no desvio da linearidade em um modelo estatístico, ou seja, a análise de Box identifica quais variáveis ou fatores têm um efeito mais notável na não conformidade do modelo com uma relação linear. Segundo Mazucheli e Achcar (2002), a estatística empregada para avaliar os desvios nos estimadores de mínimos quadrados dos parâmetros pode ser expressa como:

$$\begin{aligned} \text{Vício}(\hat{\theta}) &= -\frac{\sigma^2}{2} \left[\sum_{i=1}^n F(\theta) F_t(\theta)^{-1} \right] - \sum_{i=1}^n F(\theta). \\ \text{traço} &= \left(\sum_{i=1}^n F(\theta) F_t(\theta)^{-1} H(\theta) \right). \end{aligned} \quad (2.13)$$

$F(\theta)$ representa o vetor ($p \times 1$) de primeiras derivadas da $f(x_i, \theta)$, denominando de vetor velocidade e $H(\theta)$ é uma matriz ($p \times p$) das segundas derivadas em relação a cada elemento de θ . Pode-se resolver a equação (2.13), assumindo inicialmente que $\hat{\theta}$ e $\hat{\sigma}^2$ são valores verdadeiros de θ e σ^2 . Logo, o vetor de dimensões $p \times 1$ representa a discrepância entre as estimativas e os valores reais dos parâmetros.

2.6 Modelos de Crescimento

Conforme Seber e Wild (2003), a análise de dados sobre o crescimento é de suma importância em diversas esferas de pesquisa. Na biologia, por exemplo, há um interesse em descrever o crescimento de animais e plantas para entender os mecanismos fundamentais. Na química, o foco reside na evolução dos produtos de reações ao longo do tempo. Na agricultura, compreender o crescimento da produção, sua taxa e como esses fatores respondem a diferentes condições ou tratamentos ambientais é crucial por motivos econômicos e administrativos. Na medicina, o crescimento infantil e o desenvolvimento de tumores, bem como os efeitos dos tratamentos sobre eles, são áreas de interesse. Finalmente, cientistas sociais se dedicam ao estudo do crescimento populacional, à garantia de segurança alimentar e às demandas energéticas da sociedade.

Segundo Santos (2011), as curvas de crescimentos geralmente imagem de curvas com formato sigmoidais que representam o tempo de vida de medidas de dimensão, como altura e peso. Ao modelarmos dados de crescimento, buscamos entender o significado físico dos parâmetros, visando criar um modelo de referência para as observações em estudo.

2.6.1 Modelo de regressão Gompertz

Conforme mencionado por Máximo (2023), o modelo de *Gompertz* é uma equação matemática que descreve o crescimento de uma população ou sistema ao longo do tempo.

Inicialmente, o crescimento é rápido, porém diminui à medida que a população ou sistema se aproxima de um limite máximo. O modelo de *Gompertz* pode ser aplicado para analisar diversos fenômenos, como o crescimento populacional, a progressão de doenças e a assimilação de novos conceitos. Para Domingues (2011), a aplicação do modelo de Gompertz na área da biologia, tem sido empregado para analisar o crescimento populacional de várias espécies e para antecipar o aumento em diferentes cenários. Essa formulação é dada por:

$$y_t = \alpha e^{-\beta e^{-\gamma t}}, \quad (2.14)$$

em que, y_t é o valor da variável dependente, α é a assíntota do modelo, que indica o valor de estabilização da variável dependente em relação ao tempo t , β corresponde ao parâmetro de locação, sem interpretação biológica, γ determina a taxa de expansão do crescimento e t refere-se ao tempo de análise da variável dependente.

2.6.2 Modelo de regressão Logístico

Conforme mencionado por Souza (2010), o modelo logístico é aplicado utilizando uma função que permite a modelagem do crescimento de uma população, sendo frequentemente utilizado para descrever crescimentos sigmoidais, que se caracterizam pelo ponto de inflexão próximo da metade do último valor observado. Segundo Zeviani e Bonat (2013), a resposta média do modelo linear pode ser ajustada para uma escala mais apropriada ao problema em questão através de uma função de ligação, a qual é definida como:

$$y_t = \alpha(1 + \beta e^{-\gamma t})^{-\tau} + \varepsilon. \quad (2.15)$$

Os pontos em que a curva muda de concavidade são chamados de pontos de inflexão y . No contexto da curva exponencial, estes pontos são $y = \ln(\alpha k)$ e $E(y) = 2\alpha$. Esses valores não dependem do valor de y em que a curva é avaliada.

Os parâmetros da curva exponencial, como α , β e γ , podem ser estimados utilizando o método da máxima verossimilhança, que procura determinar os valores que tornam a curva mais provável, considerando a distribuição dos dados observados. Sabe-se que: α é o valor assintótico da variável resposta, β é uma constante relacionada ao valor inicial observado da variável resposta e γ representa a taxa de crescimento da variável resposta, ou seja, a rapidez com que a variável resposta se aproxima do seu valor assintótico. Além disso, τ é um parâmetro que influencia a curvatura dos modelos, determinando o momento em que ocorre a transição de um crescimento inibitório para um crescimento menos restrito. Isso indica o ponto em que o crescimento se torna menos previsível, com t representando o tempo em meses.

2.6.3 Modelo de Von Bertalanffy

Conforme mencionado por Oliveira (2007), o modelo não linear de *Von Bertalanffy* é amplamente empregado na análise do crescimento em peso de peixes, pois tem a capacidade de

explicar a variação de comprimento de uma determinada variável ao longo do tempo. Dentro do modelo de *Von Bertalanffy*, a variável Y cresce para uma assíntota horizontal superior, enquanto a distância restante a ser percorrida é representada por uma Exponencial decrescente. Assim, a formulação do modelo dentro da família *Von Bertalanffy*, consiste em três parâmetros, que pode ser definida como:

$$y_t = \alpha(1 - \beta e^{-\gamma t})^3 + \varepsilon. \quad (2.16)$$

Conforme descrito por Santos (2011), o modelo mencionado é adequado para descrever processos de crescimento sigmoidais, nos quais o ponto de inflexão ocorre em torno de 30% do último valor registrado. Os parâmetros α representam o valor assintótico da variável de resposta, enquanto β refere-se a uma constante associada ao valor observado inicial e γ indica a taxa de crescimento da variável resposta, representando a velocidade com que o valor observado se aproxima do valor máximo, determinando a eficiência do crescimento.

2.7 Método de Bootstrap

Segundo Araujo et al. (2013), o método *bootstrap* pode ser aplicado para estimar uma variedade de parâmetros, tais como média, variância, correlação e coeficientes de regressão. Além disso, sua utilização se estende à construção de intervalos de confiança e à realização de testes de hipóteses. Essa estratégia apresenta vantagens significativas em relação a outros métodos estatísticos. Por ser não paramétrica, isto é, não requer que a distribuição da população seja conhecida. Além de sua robustez, que implica em uma relativa insensibilidade às violações das suposições inerentes aos métodos paramétricos.

Conforme Efron e Tibshirani (1986), a técnica de *bootstrap* representa uma valiosa ferramenta que pode ser aplicada em diversas situações. Ela oferece uma alternativa viável aos métodos estatísticos convencionais como nos casos de, estimar a média de uma população quando a distribuição da mesma é desconhecida, criar um intervalo de confiança para a correlação entre duas variáveis, testar a hipótese de que a média de uma população é igual a um valor específico ou ainda avaliar o desempenho de um modelo de regressão.

De acordo com Araujo et al. (2013), o método de *bootstrap* é uma abordagem de reamostragem que permite realizar inferências sobre um determinado parâmetro. Esse método envolve a repetição da amostra original várias vezes, criando novas amostras com substituição. A média dessas novas amostras é calculada, e o desvio padrão é usado para estimar o erro padrão da média.

2.7.1 Bootstrap paramétrico

O método *bootstrap* paramétrico implica na geração de várias reamostras da amostra original, na qual a distribuição de probabilidade dos dados originais é considerada. Em cada reamostra, os parâmetros do modelo de regressão não linear são estimados. Esses valores estimados dos parâmetros são então empregados na construção de uma distribuição de probabilidade para

os estimadores de mínimos quadrados. Essa distribuição de probabilidade pode ser empregada na estimativa de intervalos de confiança para os estimadores de mínimos quadrados, além de possibilitar a realização de testes de hipóteses sobre os parâmetros do modelo.

2.7.2 Intervalo de confiança bootstrap

Conforme descrito por Souza et al. (2010), a abordagem *bootstrap* pode ser empregada para obter intervalos de confiança confiáveis para os parâmetros de um modelo de regressão não linear. Essa aplicação é fundamental, pois permite aos pesquisadores estimar com confiança a faixa na qual os parâmetros do modelo provavelmente se encontram. Efron e Tibshirani (1986), detalham dois métodos para a obtenção de intervalos de confiança *bootstrap*: o método percentual e o método percentual t . Os intervalos derivados pelo método percentual t seguem o formato:

$$\gamma^* - t_{1-\frac{\alpha}{2}} \cdot \frac{S(\gamma^*)}{\sqrt{n}} \leq \gamma \leq \gamma^* + t_{1-\frac{\alpha}{2}} \cdot \frac{S(\gamma^*)}{\sqrt{n}}. \quad (2.17)$$

$S(\gamma^*)$ representa o desvio padrão estimado de γ^* . Os valores t_i são calculados utilizando a distribuição *bootstrap* de γ^* .

2.7.3 Teste de normalidade bootstrap

Conforme Razali, Wah et al. (2011), o teste de Shapiro-Wilk é considerado um teste não paramétrico, indicando que não pressupõe nenhuma distribuição específica dos dados. Essa característica o torna robusto, ou seja, menos suscetível a desvios da normalidade em comparação com outros testes de normalidade. A hipótese do teste é formulada como:

$$\begin{cases} H_0 : \text{Os parâmetros provêm de uma distribuição Normal,} \\ H_1 : \text{Os parâmetros não provêm de uma distribuição Normal.} \end{cases}$$

De acordo com Araujo et al. (2013), a assimetria é uma medida que indica a falta de simetria em uma distribuição, mensurando a diferença entre a mediana e a média da distribuição. Na presença de uma distribuição normal, em que a média e mediana coincidem, a distribuição é considerada simétrica. Isso implica que metade das observações está ao lado da média e a outra metade está do outro lado. No entanto, quando a distribuição não é simétrica, a média e mediana não coincidem. Nesses casos, a assimetria pode ser positiva ou negativa.

A assimetria positiva sugere que a distribuição possui mais observações com valores maiores do que a média do que observações com valores menores do que a média. Por outro lado, uma assimetria negativa indica que há distribuição tem mais observações com valores menores do que a média do que maiores. O teste de assimetria é um procedimento estatístico que avalia se uma distribuição é simétrica. Ele se baseia na estatística p , que quantifica a assimetria da distribuição. O teste de assimetria é fundamentado na seguinte hipótese nula:

$$\begin{cases} H_0 : AS = 0, \\ H_1 : AS \neq 0. \end{cases} \quad (2.18)$$

Da mesma forma, o coeficiente de curtose (C) precisa ser igual a 3 para que a distribuição dos dados seja a Normal. A formulação da hipótese do teste é:

$$\begin{cases} H_0 : C = 3, \\ H_1 : C \neq 3. \end{cases} \quad (2.19)$$

2.8 Machine learning

De acordo com Deisenroth (2020), o campo do aprendizado de máquina, também denominado *Machine Learning*, consiste em algoritmos capazes de extrair informações significativas dos dados. Essa habilidade conduz à obtenção de modelos eficazes que podem ser generalizados para dados futuros e desconhecidos pelo algoritmo. O processo de aprendizado em si representa o mecanismo pelo qual o algoritmo reconhece padrões e estruturas nos dados, iniciando o processo de aprendizado por meio de memorização ou experiência.

De acordo com Mueller e Guido (2016), existem vários tipos de aprendizado de máquina, todos os quais envolvem o algoritmo aprendendo com os dados. A principal diferença entre o aprendizado supervisionado e o não supervisionado é a presença de rótulos nos dados. No aprendizado supervisionado, o algoritmo recebe dados com rótulos e sua tarefa é identificar padrões e, em seguida, aplicá-los a novos dados. Por outro lado, no aprendizado não supervisionado, o algoritmo aprende a identificar padrões e a agrupar observações com base em características comuns. No aprendizado não supervisionado, apenas os dados de entrada são conhecidos e nenhum dado de saída conhecido é fornecido ao algoritmo.

Segundo Ignacio (2021) destaca, outro método de aprendizado a ser considerado é o aprendizado semi-supervisionado. Nesse cenário, os dados consistem em observações com e sem rótulos, e o procedimento subsequente assemelha-se ao do aprendizado supervisionado. Por outro lado, o aprendizado por reforço é baseado na experiência, em que o algoritmo aprende a executar tarefas corretamente por meio de tentativa e erro, recebendo recompensas por acertos e punições por erros. Em resumo, a escolha entre esses tipos de aprendizado depende das características do problema em questão e da estrutura dos dados disponíveis.

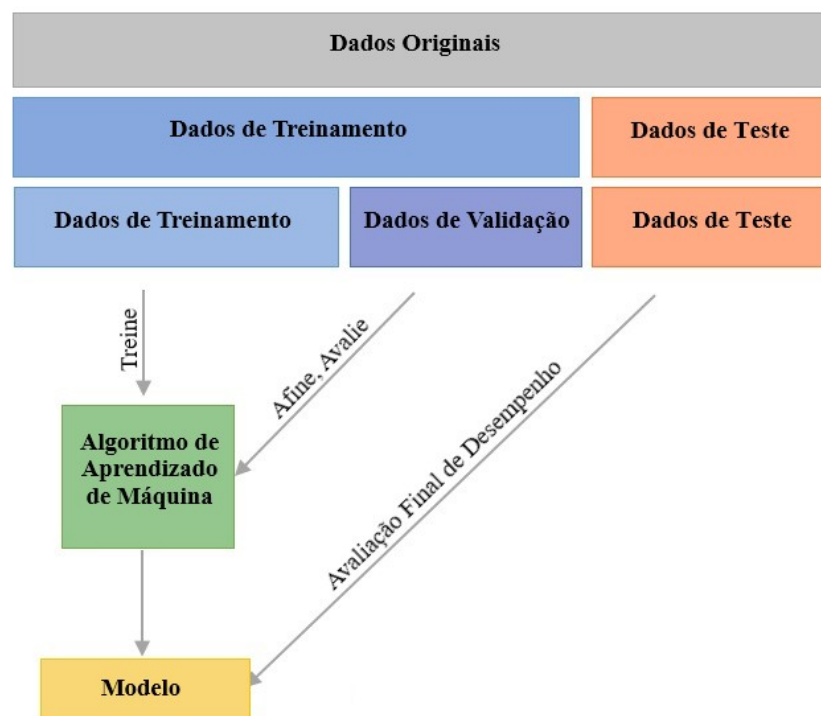
Algoritmos de Aprendizado de Máquina (*Machine Learning*) são métodos que permitem que computadores aprendam a partir de dados, identificando padrões e realizando previsões ou decisões sem a necessidade de uma programação específica para cada tarefa. Esses algoritmos utilizam conceitos de matemática, estatística e ciência da computação para reconhecer relações e ajustar seus modelos, melhorando o desempenho ao longo do tempo. Entre os exemplos mais comuns estão *Árvore de Decisão*, *Random Forest*, *EXtreme Gradient Boosting (XGBoost)*, *Light Gradient Boosting Machine (LightGBM)* e *CatBoost*. No contexto deste trabalho, o aprendizado de máquina é aplicado para entender as métricas de validação.

2.9 Seleção e validação

De acordo com Shalev-Shwartz e Ben-David (2014), é crucial tomar certas precauções antes de escolher o modelo final, e a validação surge como uma maneira de diminuir as incertezas sobre qual modelo oferecerá os melhores resultados. A abordagem básica envolve dividir o conjunto de dados de treinamento em dois conjuntos distintos. Conforme destacado por Santos (2011), quando a base de dados é extensa, é mais apropriado dividir os dados aleatoriamente em três conjuntos: treinamento, validação e teste.

Ainda conforme Shalev-Shwartz e Ben-David (2014), a validação é utilizada para a seleção de modelos, no qual diferentes algoritmos são inicialmente treinados com o conjunto de treinamento. Em seguida, o modelo que será utilizado na base de teste, amostra-se o conjunto de validação e seleciona-se o modelo que minimiza o erro sobre conjunto de validação. Como mostrado na Figura 1 nesse contexto, o conjunto de validação é independente do conjunto de treinamento. Essa técnica de validação é aplicada quando os dados são abundantes. Em situações em que não é possível separar um novo conjunto de dados para validação, recorre-se a outra técnica chamada validação cruzada *k-fold* para validar os modelos de treinamento.

Figura 1 – Processo de treinamento, validação e teste de modelos.

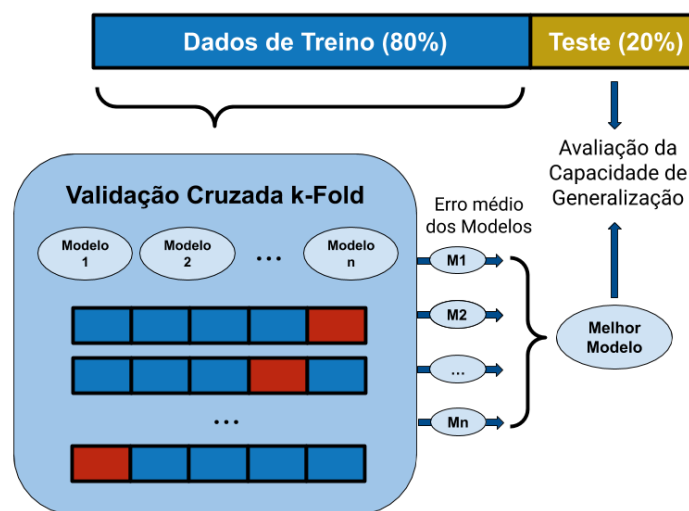


Fonte: Freecodecamp, 2018.

Segundo Santos (2011), na técnica da validação cruzada *k-fold*, o conjunto de dados é dividido aleatoriamente em k subconjuntos de tamanhos iguais, geralmente 5 ou 10, na prática. Desses subconjuntos, $k - 1$ são destinados ao treinamento do modelo e o restante é reservado para a fase de validação. O funcionamento desse tipo de validação é ilustrado na Figura 2 em

que o processo de validação cruzada $k - fold$ é repetido até que todas as partições tenham sido usadas tanto para treinamento quanto para validação. Isso resulta em k estimativas que são posteriormente resumidas calculando-se a média e o erro padrão, conforme mencionado por Santos (2018).

Figura 2 – Processo de treinamento, validação cruzada $k - fold$ ($k = 5$).



Fonte: Scaccia, 2020.

2.10 Problemas do aprendizado de máquina

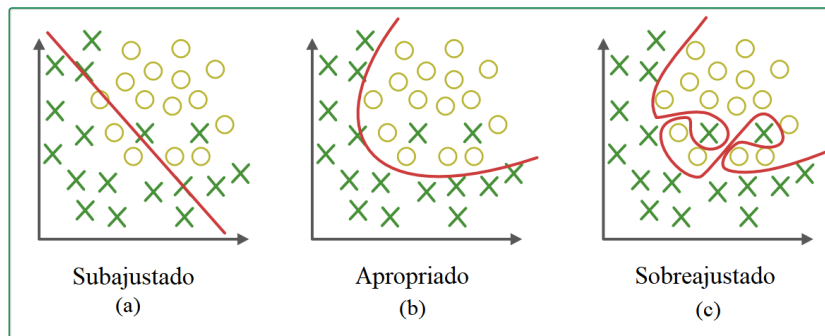
Nesta seção, discutiremos os principais desafios enfrentados ao lidar com o aprendizado de máquinas. Utilizando algoritmos simples e robustos a depender dos dados utilizados, pode-se levar a modelos subajustados e sobreajustados.

2.10.1 Sobreajuste e Subajuste

Conforme apontado por Shalev-Shwartz e Ben-David (2014), é essencial ter cautela ao lidar com modelos de aprendizado de máquina. Um pequeno ajuste em algum parâmetro do modelo pode resultar em um grande erro de aproximação, enquanto um ajuste excessivo pode levar a um problema de sobreajuste (*overfitting*).

Na Figura 3 são ilustrados três cenários. A situação (a) ocorre quando há subajuste (*underfitting*) nos modelos, indicando que o modelo resultante é muito simples para explicar a variância. O subajuste resulta em um alto erro tanto nos dados de treinamento quanto nos dados de teste. Existem duas maneiras de identificar o subajuste nos modelos: visualmente, por meio de gráficos, e através da análise do erro. Na situação (b), é apresentado um exemplo de um modelo bem ajustado. Por outro lado, na situação (c), surge um problema de sobreajuste, no qual o modelo memoriza a estrutura dos dados de treinamento. O sobreajuste ocorre quando o modelo se adapta excessivamente aos dados de treinamento, resultando em um desempenho deficiente ao generalizar para novos dados.

Figura 3 – Modelos subajustado, apropriado e sobreajustado.



Fonte: Geeksforgeeks, 2022.

2.10.2 Dados desbalanceados

Segundo Santos (2011), a abordagem mais comum envolve a amostragem estratificada para criar subamostras da classe menos frequente. Isso visa otimizar o modelo, aumentando a sensibilidade da classe menos frequente e equilibrando as classes do atributo de interesse. De acordo com Lopes (2018), esse processo de reamostragem permite que não haja uma predominância de uma classe específica no conjunto de treinamento, gerando cópias artificiais com reposição da classe menos frequente. Conforme Frizzarini e Lauretto (2013), isto reduz os erros de previsão significativos na classe minoritária, especialmente em contextos médicos, onde essa classe pode representar a presença de uma condição grave.

2.11 Métricas de avaliação em aprendizado de máquina

Conforme Mueller e Guido (2016), na área de classificação de modelos, diversas métricas são empregadas para avaliar seu desempenho, incluindo acurácia, precisão, sensibilidade, especificidade e *F-score*. Estas métricas derivam de uma matriz de confusão, exemplificada na Figura 4 em que a matriz apresenta uma tabulação cruzada entre as classes observadas e preditivas. Analisar diretamente a matriz de confusão não é suficiente para tirar conclusões precisas, no entanto, ao calcular métricas específicas, é possível alcançar conclusões mais adequadas.

Figura 4 – Matriz de confusão.

		Detectada	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Fonte: Rodrigues, 2019.

Uma forma de interpretar as informações da matriz de confusão é calcular a acurácia, que representa a proporção de predições corretas feitas pelo modelo em ambas as classes. No

entanto, a acurácia não é sempre a métrica mais indicada para avaliar modelos, pois não leva em consideração a distribuição das classes. A precisão mede a proporção de verdadeiros positivos em relação a todas as predições positivas do modelo, enquanto a sensibilidade (*recall*) avalia a capacidade do modelo em detectar corretamente os casos positivos. Por sua vez, a especificidade avalia a capacidade de detectar corretamente os casos negativos. O *F-score* é uma métrica que combina precisão e sensibilidade, proporcionando uma medida ponderada da performance do modelo. A seguir, serão apresentadas as equações utilizadas para calcular essas métricas fundamentais de avaliação de modelos em *Machine Learning*.

$$\text{i) Acurácia} = \frac{VP+VN}{VP+FN+VN+FP};$$

$$\text{ii) Precisão} = \frac{VP}{VP+FP};$$

$$\text{iii) Sensibilidade} = \frac{VP}{VP+FN};$$

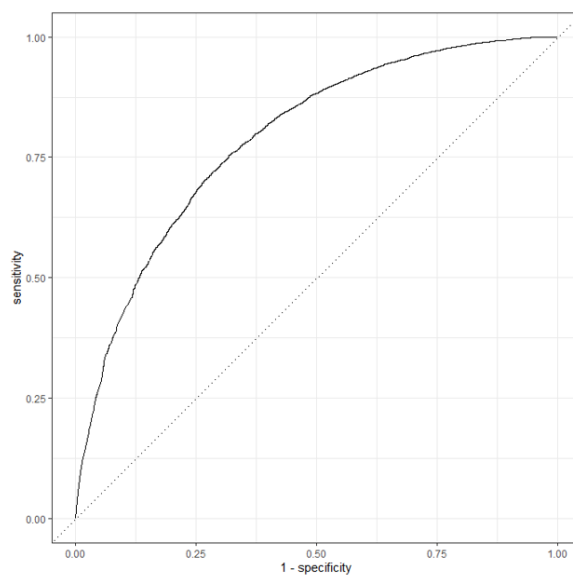
$$\text{iv) Especificidade} = \frac{VN}{VN+FP};$$

$$\text{v) } F\text{-score} = 2 \times \frac{\text{Precisão} \times \text{Sensibilidade}}{\text{Precisão} + \text{Sensibilidade}}.$$

Conforme Mueller e Guido (2016), o *F-score* é uma métrica que combina precisão e sensibilidade de forma mais equilibrada, sendo uma opção mais robusta do que a simples acurácia. Uma ferramenta frequentemente utilizada para avaliar o desempenho de modelos de classificação é a Curva ROC (*Receiver Operating Characteristic Curve*), a qual é um gráfico capaz de distinguir entre duas classes. Esta curva considera todos os possíveis limites para um classificador específico, exibindo a taxa de falsos positivos em relação à taxa de verdadeiros positivos.

A Figura 5 apresenta uma curva ROC acompanhada da sua AUC (*area under the ROC curve*). Neste exemplo específico, são exibidas as curvas para dois modelos: Regressão Logística e Rede Neural. De acordo com Mueller e Guido (2016), o ideal é que a curva ROC se aproxime do canto superior esquerdo, indicando um classificador que alcance alta sensibilidade com baixa taxa de falsos positivos. É especialmente recomendado utilizar a área sob a curva ROC para avaliar modelos em conjuntos de dados desequilibrados. A área sob a curva ROC não possui um limite predefinido, e em algumas circunstâncias, pode ser necessário ajustar um limite de decisão para alcançar resultados com alta área sob a curva ROC.

Figura 5 – Curva ROC.



Fonte: Rodrigues, 2018.

3 MATERIAL E MÉTODOS

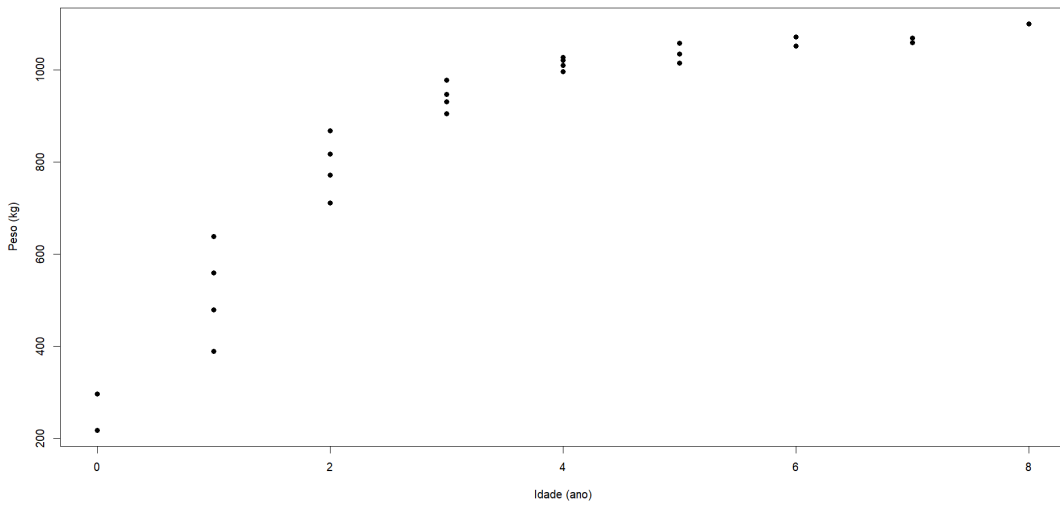
Neste trabalho foram utilizados dados obtidos a partir do peso dos touros *Holstein-Friesiande* de 6 meses a 8 anos de idade, com um total de 26 medições para peso do touro, conforme dados obtidos no artigo de Teleken (2017). Neste estudo, a variável de interesse, representada por y , corresponde ao peso (kg), enquanto a variável preditiva x refere-se à idade em anos. A Tabela 1 tem-se as observações do peso junto com suas respectivas idades em anos.

Tabela 1 – Relação do crescimento do peso (kg) do touro Holandês ao longo dos anos.

Anos	Peso
0	218
0	297
1	389
1	479
1	559
1	638
2	711
2	771
2	817
2	868
3	905
3	930
3	946
3	977
4	996
4	1010
4	1021
4	1027
5	1014
5	1034
5	1058
6	1051
6	1071
7	1059
7	1069
8	1099

Fonte: *Teleken, 2017*.

Na Figura 6 é exibido o gráfico das medias do peso (kg) com relação a idade em anos para o touro *Holstein-Friesiande*. A análise visual revela um crescimento não linear do peso (kg). Portanto, sugere-se a aplicação de técnicas de modelagem não linear para uma descrição mais precisa dos dados analisados.

Figura 6 – Relação entre o peso do touro *Holstein-Friesiande* ao longo dos anos.

Fonte: Elaborado pela autora, 2024.

Neste contexto, foram ajustados os modelos *Logístico*, *Gompertz* e *Von Bertalanffy* para estimar o peso do touro ao longo dos anos. Os parâmetros dos modelos foram estimados, juntamente com seus erros padrão, estatísticas do teste *t* para verificar a significância dos parâmetros e intervalos de confiança correspondentes. O ajuste do modelo foi realizado utilizando o método iterativo de *Gauss-Newton*, que inicia-se com valores iniciais dos parâmetros e itera até convergir, utilizando o método dos Mínimos quadrados para obter as estimativas dos parâmetros do problema de regressão não linear. Para avaliar os modelos, foram utilizados os critérios de coeficiente de determinação (R^2) e o critério de informação de Akaike (AIC).

O R^2 indica a proporção da variabilidade total dos dados que é explicada pelo modelo, variando de 0 (o modelo não explica nenhuma variabilidade nos dados) e 1 (as previsões do modelo correspondem exatamente aos dados observados).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}. \quad (3.1)$$

em que, y_i são os valores observados; \hat{y}_i são os valores preditos pelo modelo; \bar{y} é a média dos valores observados; n é o número de observações.

Enquanto o *Akaike Information Criterion* (AIC) considera a complexidade do modelo em relação à sua capacidade de ajuste aos dados.

$$AIC = 2k - 2\ln(\hat{L}). \quad (3.2)$$

- k é o número de parâmetros do modelo;
- \hat{L} é o valor da função de verossimilhança maximizada.

O AIC é uma ferramenta fundamental para comparar modelos concorrentes, auxiliando na escolha do modelo que melhor se ajusta aos dados de crescimento de forma eficaz, priorizando aquele com o menor valor de AIC.

Em seguida foram empregadas técnicas de aprendizado de máquina, como o *bootstrap*, para estimar o erro padrão e o viés dos estimadores, comparar a qualidade do ajuste dos modelos e validar os resultados com maior confiabilidade. O viés de Box foi aplicado para identificar possíveis desvios da linearidade nos parâmetros. A matriz de confusão possibilitou uma análise detalhada dos acertos e erros do modelo, facilitando ajustes estratégicos para melhorar o desempenho. Para avaliar o desempenho com base na matriz de confusão foram consideradas métricas como acurácia, precisão, sensibilidade, especificidade e *F-score*. Além disso, a curva ROC foi utilizada para avaliar não apenas o ajuste dos modelos aos dados de peso ao longo do tempo, mas também a capacidade de classificar corretamente os animais em categorias de interesse. A métrica da área sob a curva (AUC) próxima de 1 indica que o modelo apresenta excelente desempenho na tarefa de classificação (acima ou abaixo do limiar de 1000 kg). Essas análises foram realizadas usando a função "nls" do *software R* versão 4.3.1, no ambiente *RStudio e COLAB*.

4 RESULTADOS E DISCUSSÃO

O método iterativo de Gauss-Newton é uma técnica de otimização utilizada para ajustar modelos não lineares aos dados, minimizando a soma dos quadrados dos resíduos entre os valores observados e os valores preditos pelo modelo para aplicar esse método os valores iniciais dos parâmetros foram obtidos através da linearização do modelo, permitindo o ajuste dos modelos aos dados. A seguir são apresentadas as estimativas dos parâmetros para os modelos não lineares *Logístico*, *Gompertz* e *Von Bertalanffy*, conforme a metodologia estabelecida, juntamente com os erros padrão das estimativas, os resultados do teste *t* e os intervalos de confiança a 95%. Dessa forma, apresentam-se adiante os resultados pertinentes aos modelos ajustados.

Tabela 2 – Estimativas (E) dos parâmetros (P) para o modelo *Logístico*, *Gompertz* e *Von Bertalanffy*, erro padrão da estimativa (E.P.E.), teste *t* e seus respectivos intervalos de confiança (IC) a 95% .

Modelo	P	E	E.P.E.	teste <i>t</i>	p-valor	IC
<i>Gompertz</i>	α	1081.000	20.76	52.06	<0.0001	[1040.6643; 1126.4691]
	β	1.524	0.1187	12.85	<0.0001	[1.3032; 1.8106]
	γ	0.7760	0.07134	10.88	<0.0001	[0.63921; 0.9342]
<i>Logístico</i>	α	1062.000	17.64	60.164	<0.0001	[1026.0333; 1100.5970]
	β	1.119	0.1250	8.952	<0.0001	[0.8715; 1.3980]
	γ	1.074	0.09577	11.213	<0.0001	[0.8862; 1.2989]
<i>Von Bertalanffy</i>	α	1091.000	23.06	47.32	<0.0001	[1047.4637; 1141.9286]
	β	0.4069	0.02610	15.59	<0.0001	[0.3571; 0.4678]
	γ	0.6770	0.06487	10.44	<0.0001	[0.5537; 0.8167]

Fonte: Elaborado pelo autora, 2024.

Através da Tabela 2 observa-se que, com base nos *p*-valores dos testes *t* os intervalos de confiança dos parâmetros (α , β , γ) não abrangem o valor zero, indicando que os parâmetros estimados são estatisticamente significativos exercem um impacto relevante no ajuste dos modelos aos dados. Contudo, o modelo Von Bertalanffy se destaca por ter o menor erro padrão geral e uma amplitude de intervalo menor para o parâmetro β .

A adequação dos modelos aos dados foi avaliada usando dois critérios: o Critério de Informação de *Akaike* (AIC) e o Coeficiente de Determinação R^2 . Na Tabela 3, estão listados os valores dessas estatísticas. Portanto, o modelo mais adequado é aquele que possui o menor valor de AIC e o maior valor de R^2 .

Tabela 3 – Estatísticas de qualidade de informação de ajuste.

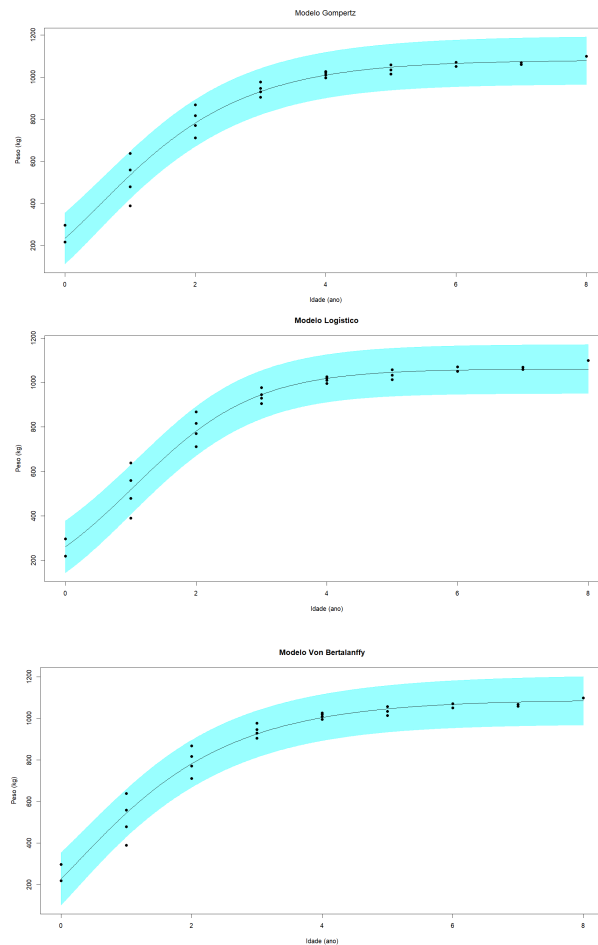
	<i>Gompertz</i>	<i>Logístico</i>	<i>Von Bertalanffy</i>
R^2	0.960211	0.9613296	0.957985
AIC	207.1503	206.4156	208.5653

Fonte: Elaborado pelo autora, 2024.

Os resultados da Tabela 3 indicam que os coeficientes de determinação são bons, aproximadamente 96% . Apesar disso, embora os coeficientes de determinação sejam muito próximos

nos três modelos, o modelo *Logístico* se destaca ao apresentar o maior coeficiente de determinação (R^2) o que explica um melhor ajuste aos dados, e um dos menores valores para o Critério de Informação de Akaike (AIC), indicando que é ligeiramente mais adequado aos dados analisados. A Figura 7 ilustra as curvas de crescimento do peso (kg) ajustadas pelos modelos *Von Bertalanffy*, *Logístico* e *Gompertz*.

Figura 7 – Ajuste dos modelos Gompertz, Logístico e Von Bertalanffy para o peso (kg) do touro.

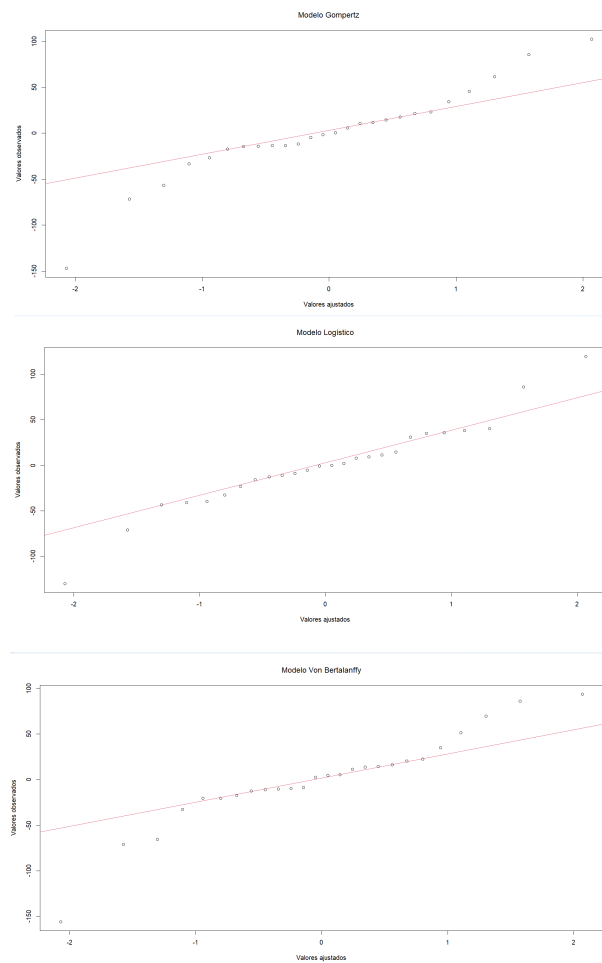


Fonte: Elaborado pelo autora, 2024.

Dessa forma, o modelo *Gompertz* é uma alternativa viável, pois, possui um R^2 muito próximo ao do modelo *Logístico* e um AIC apenas ligeiramente superior.

A análise prosseguiu com a verificação gráfica da normalidade dos resíduos dos modelos Figura 8. Esses gráficos são valiosos para avaliar visualmente o desempenho de cada modelo e destacar pontos em que o modelo pode não estar representando adequadamente a relação entre os valores observados e ajustados. Ao comparar os três gráficos, o modelo *Logístico* aparenta ter uma leve vantagem em termos de ajuste mais próximo à linha reta que indica uma possível normalidade, embora essa diferença seja pequena.

Figura 8 – Normalidad dos resíduos.



Fonte: Elaborado pelo autora, 2024.

Na Tabela 4 são apresentados os resultados do Teste de *Shapiro-Wilk*, utilizado para avaliar se os resíduos dos dados seguem uma distribuição normal. Caso o p-valor seja superior a 0,05, os dados podem ser considerados normalmente distribuídos. Caso contrário, com um p-valor inferior a 0,05, sugere-se que os dados não seguem uma distribuição normal. Os resíduos dos modelos *Gompertz* e *Logístico* indicam uma distribuição normal, enquanto os resíduos do modelo *Von Bertalanffy* não seguem essa distribuição, sendo menos favorável.

Tabela 4 – Teste de *Shapiro-Wilk*.

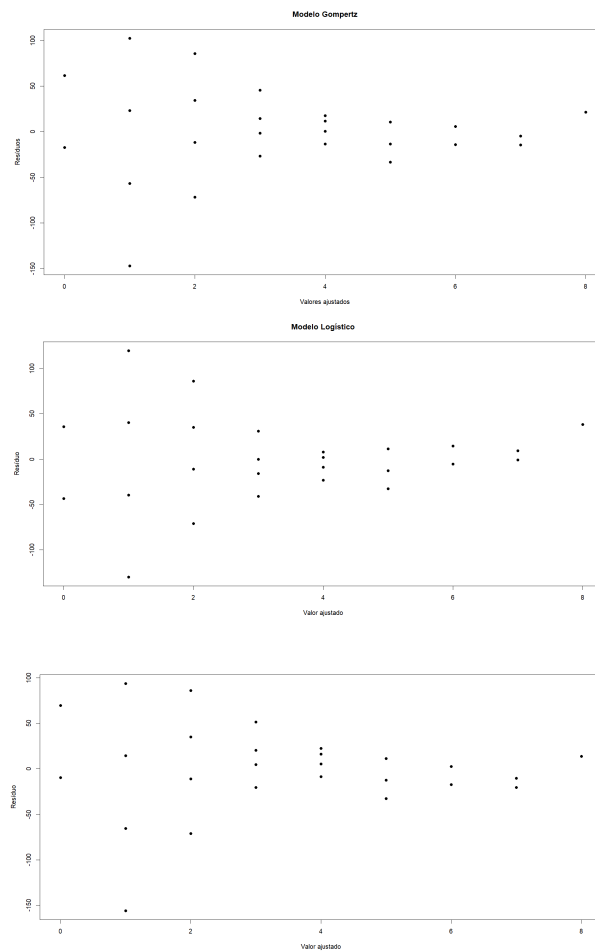
Modelo	p-valor
<i>Gompertz</i>	0.08438
<i>Logístico</i>	0.26810
<i>Von Bertalanffy</i>	0.03458

Fonte: Elaborado pelo autora, 2024.

A homocedasticidade é um requisito crucial na análise de regressão, e a sua violação pode comprometer significativamente a validade das conclusões. Dessa forma, a Figura 9 apresenta a verificação visual da homocedasticidade, indicando que nenhum dos modelos proporciona um

ajuste ideal. Entretanto, o modelo *Logístico* apresenta uma menor variabilidade nos resíduos em relação aos modelos *Gompertz* e *Von Bertalanffy*, o que pode sugerir um melhor ajuste geral aos dados.

Figura 9 – Relação entre valores ajustado e resíduos.



Fonte: Elaborado pelo autora, 2024.

Para validar os modelos não lineares de crescimento (*Gompertz*, *Logístico* e *Von Bertalanffy*), foram utilizados métodos de avaliação de *Machine Learning*, como o *bootstrap*, a curva ROC e a matriz de confusão, para analisar a precisão, robustez e capacidade discriminativa dos modelos. A validação pelo método *bootstrap* foi realizada por meio de simulações, gerando amostras com base nas estimativas dos parâmetros do modelo *Logístico*, que apresentou o melhor ajuste para o peso (kg) dos touros Holandeses ao longo dos anos. Foram criadas 1000 amostras simuladas para cada modelo, cada uma contendo 26 observações.

A Tabela 5 apresenta as estimativas dos parâmetros obtidas através da técnica de *bootstrap* para os três modelos, acompanhadas dos valores do viés relativo da estimativa *bootstrap* (V.R.B.), do viés de Box, e do respectivo intervalo de confiança.

Tabela 5 – Estimativas *bootstrap* (E.B.), vício relativo a estimativa *bootstrap* (V.R.B.), vício de Box, intervalo de confiança *bootstrap* a 95% para os parâmetros (p).

Modelo	P	E.B.	V.R.B	Vício de Box	Intervalo de Confiança
<i>Gompertz</i>	α	1079.2314	-0.0014	-1.5446	[1054.6142; 1103.0766]
	β	1.5594	0.0229	0.0346	[1.2765; 2.0964]
	γ	0.7933	0.0223	0.0173	[0.67741; 0.98748]
<i>Logístico</i>	α	1061.3776	-0.0001	-0.1876	[1041.2108; 1079.9506]
	β	3.1210	0.0189	0.0579	[2.0863; 4.3657]
	γ	1.0815	0.0071	0.0076	[0.8783; 1.2921]
<i>Von Bert.</i>	α	1089.7147	-0.0012	-1.3863	[1063.4100; 1117.9619]
	β	0.41789	0.0270	0.0110	[0.34670; 0.56712]
	γ	0.69556	0.0274	0.0160	[0.58614; 0.89061]

Fonte: Elaborado pelo autora, 2024.

Na Tabela 5 é possível observar que as estimativas dos parâmetros e seus intervalos de confiança obtidos pelo *bootstrap* são bastante semelhantes aos valores assintóticos. Os resultados da simulação *bootstrap* sugerem que o modelo *Logístico* é o mais adequado para descrever o crescimento do peso do touro holandês ao longo dos anos. Esse modelo não só se ajusta bem aos dados, como também apresentou intervalo mais estreitos para α , sugerindo maior precisão.

O modelo *Von Bertalanffy* também se ajusta bem aos dados, seus intervalos de confiança são mais estreitos para β e γ indicando maior precisão nessas estimativas. Já o modelo de *Gompertz* apresenta um ajuste inferior em relação aos outros dois modelos.

Tabela 6 – Testes de normalidade *bootstrap*.

Modelo	P	<i>Shapiro-Wilk</i>	p-valor	Assimetria	p-valor	Curtose	p-valor
<i>Gompertz</i>	α	0.9968	0.0433	-0.6282	0.5298	2.8333	0.0046
	β	0.7854	<0.0001	20.3654	<0.0001	14.3445	<0.0001
	γ	0.9015	<0.0001	15.2394	<0.0001	11.0171	<0.0001
<i>Logístico</i>	α	0.9978	0.2361	-1.9171	0.0552	1.2801	0.2004
	β	0.9202	<0.0001	14.3049	<0.0001	10.8930	<0.0001
	γ	0.9674	<0.0001	8.8060	<0.0001	7.5994	<0.0001
<i>Von Bert.</i>	α	0.9946	0.0013	2.2806	0.0225	3.1199	0.0018
	β	0.8372	<0.0001	18.3725	<0.0001	13.0138	<0.0001
	γ	0.9226	<0.0001	13.6318	<0.0001	9.4519	<0.0001

Fonte: Elaborado pelo autora, 2024.

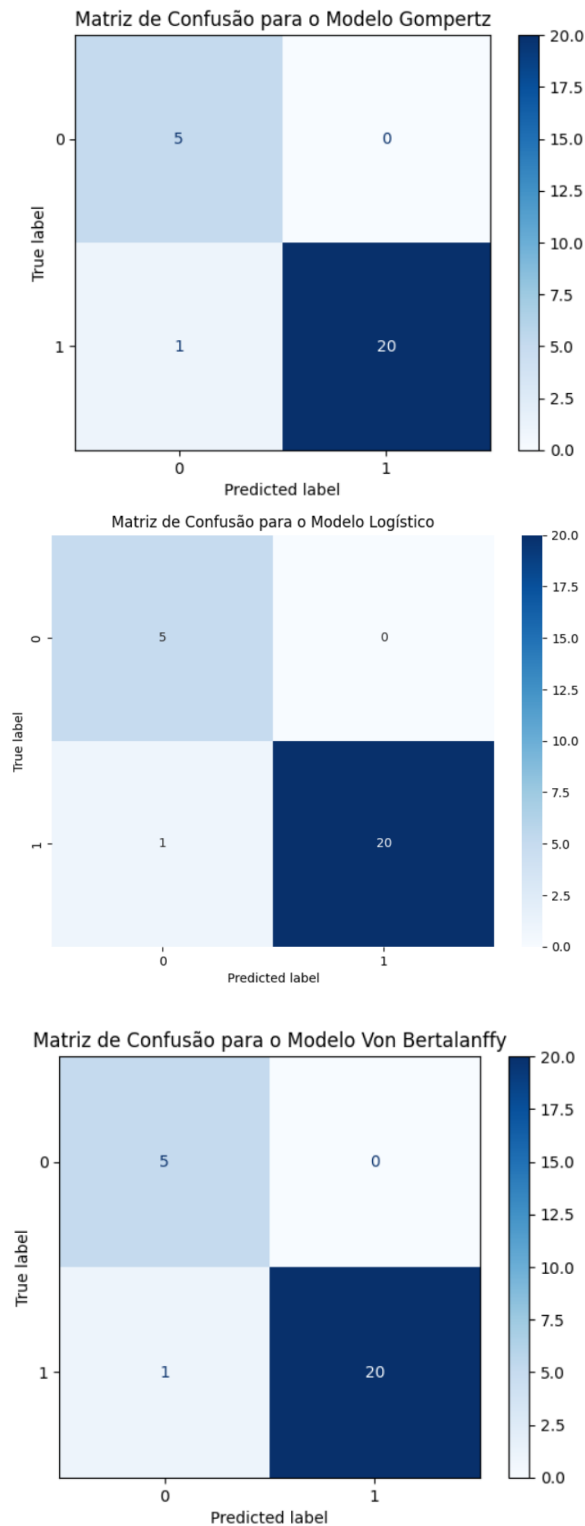
O parâmetro α nos três modelos tende a se aproximar mais de uma distribuição normal, exceto no modelo *Von Bertalanffy*, em que tanto a assimetria quanto a curtose apresentam valores significativos. Por outro lado, os parâmetros β e γ , em todos os modelos, evidenciam desvios claros da normalidade, com valores elevados de assimetria e curtose, além de p-valores muito baixos nos testes de *Shapiro-Wilk*, assimetria e curtose. Portanto, o parâmetro α geralmente apresenta uma distribuição mais estável e possivelmente mais próxima da normalidade, enquanto os parâmetros β e γ exibem distribuições significativamente distorcidas, especialmente nos modelos *Gompertz* e *Von Bertalanffy*.

Ao considerar a normalidade dos resíduos, assimetria e curtose dos parâmetros, juntamente com o desempenho geral do modelo (R^2 e AIC), o modelo *Logístico* se destaca como o "melhor" entre os três avaliados. Ele oferece um ajuste mais equilibrado aos dados, com menos violações dos pressupostos de normalidade, e apresenta os melhores indicadores de ajuste estatístico (R^2 e AIC).

Segundo Augusto (2023) para a análise da Figuras 10 matriz de confusão e da Tabela 7 curvas ROC para cada modelo, foi utilizado um limiar de 600 kg para categorizar os touros que estão em condições ideais para reprodução, uma prática adotada na pecuária. Os touros abaixo desse peso podem não ter a maturidade física necessária para uma reprodução eficiente (classe 0, não aptos), enquanto aqueles com peso igual ou superior a 600 kg geralmente são considerados maduros e aptos para reprodução (classe 1, aptos).

A Figura 10 refere-se às matrizes de confusão, que detalham o desempenho dos modelos *Gompertz*, *Von Bertalanffy* e *Logístico* na classificação correta das observações em classes positivas e negativas.

Figura 10 – Matriz de confusão.



Fonte: Elaborado pelo autora, 2024.

Os três modelos apresentam a mesma matriz de confusão, com 5 Verdadeiros Negativos (VN): 5 touros foram corretamente identificados como não aptos para reprodução; 0 Falsos Positivos (FP): nenhum touro não apto foi incorretamente identificado como apto; 1 Falso

Negativo (FN): 1 touro apto foi incorretamente classificado como não apto; e 20 Verdadeiros Positivos (VP): 20 touros aptos foram corretamente classificados. Isso indica que, em termos de classificação, todos os modelos tiveram desempenho igual nos dados analisados. O modelo *Logístico* continua sendo a melhor escolha, em consonância com os resultados anteriores de R^2 e AIC.

A Tabela 7 apresenta as métricas de desempenho de três modelos preditivos *Gompertz*, *Logístico* e *Von Bertalanffy*, considerando idade e peso. Os três modelos mostraram resultados muito semelhantes, com 96% de acurácia (classificação correta), 100% de precisão ao identificar os touros prontos para reprodução, 95% de sensibilidade ao reconhecer corretamente os touros aptos, e 100% de especificidade ao identificar os touros não prontos. Portanto, qualquer um desses modelos é eficaz para prever a aptidão reprodutiva dos touros com base no peso de 600 kg.

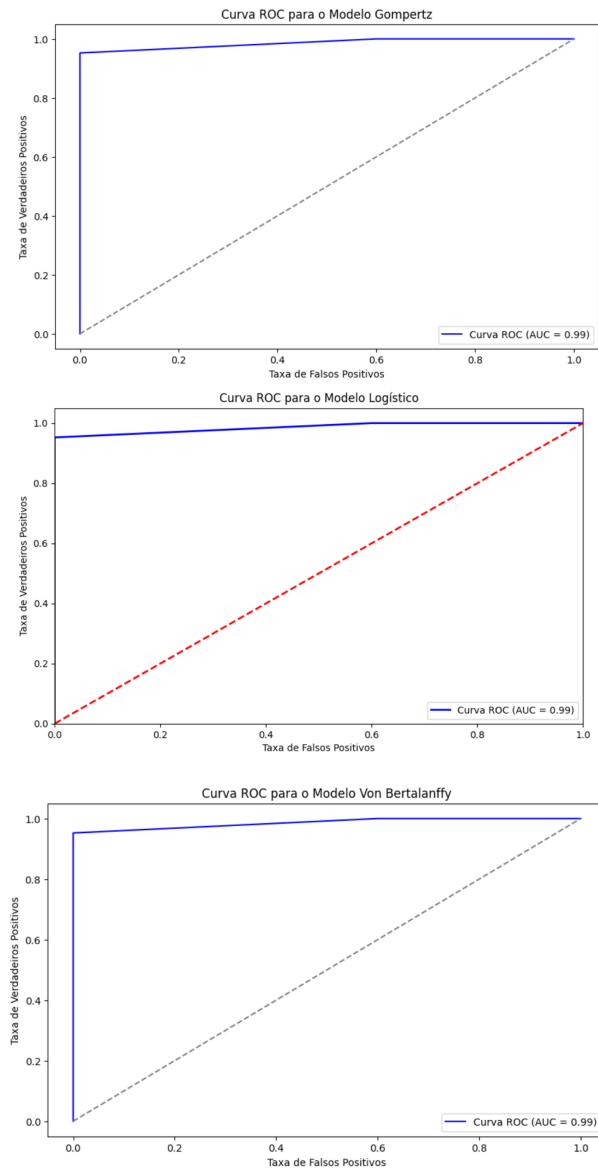
Tabela 7 – Valores das métricas para os modelos.

Modelo	Acurácia	Precisão	Sensibilidade	Especificidade	F-Score
<i>Gompertz</i>	0.96	1.00	0.95	1.00	0.98
<i>Logístico</i>	0.96	1.00	0.95	1.00	0.98
<i>Von Bert.</i>	0.96	1.00	0.95	1.00	0.98

Fonte: Elaborado pelo autora, 2024.

Para a análise da Figura 11 as curvas ROC, foi utilizado um limiar de 600 kg para categorizar os touros que estão em condições ideais para reprodução, uma prática amplamente adotada na pecuária. Touros abaixo desse peso podem não ter a maturidade física necessária para uma reprodução eficiente (classe 0, negativa), enquanto aqueles com peso igual ou superior a 600 kg geralmente são considerados maduros e aptos para reprodução (classe 1, positiva). Os modelos preditivos aplicados, como *Gompertz*, *Von Bertalanffy* e *Logístico*, apresentaram uma Área Sob a Curva (AUC) de 0,99, demonstrando que todos são eficazes em prever a precisão reprodutiva dos touros com base na idade e no peso. No manejo, o uso desses modelos pode aprimorar as decisões de reprodução, assegurando que apenas os touros que atingiram o peso ideal sejam utilizados para reprodução.

Figura 11 – Curva de ROC.



Fonte: Elaborado pelo autora, 2024.

5 CONCLUSÃO

Neste estudo, buscou-se realizar uma revisão detalhada dos modelos de crescimento não lineares. Inicialmente, foram apresentadas as definições desses modelos, destacando suas diferenças em relação aos modelos lineares, bem como os métodos de estimação de parâmetros, com ênfase no método iterativo de *Gauss-Newton* para ajuste. Foram discutidas também as principais técnicas de avaliação e classificação desses modelos. Posteriormente, foram introduzidas as medidas de não linearidade e as técnicas de diagnóstico aplicáveis. Na aplicação prática, foram analisados dados reais compostos por 26 observações que relacionam o peso (em kg) de touros da raça Holandesa, desde os 6 meses até os 8 anos de idade, no contexto de programas de melhoramento genético voltados ao aumento da produtividade do gado leiteiro.

Três modelos de crescimento — *Gompertz*, *Logístico* e *Von Bertalanffy* — foram avaliados utilizando técnicas de aprendizado de máquina, como simulação de *bootstrap*, matriz de confusão e curva ROC. Os resultados das simulações de *bootstrap*, considerando a normalidade dos resíduos, assimetria e curtose dos parâmetros, indicaram que o modelo Logístico oferece um ajuste aos dados de forma satisfatória e confiável. A análise da matriz de confusão mostrou que os três modelos apresentaram o mesmo desempenho, com 5 touros não apto para reprodução e 20 touros aptos.

Por fim, a curva ROC revelou desempenhos semelhantes para os três modelos, com AUC = 0,99, indicando que qualquer um deles pode ser utilizado sem comprometer a precisão das previsões. O modelo Logístico demonstrou que o crescimento dos touros segue a curva esperada, com um aumento rápido de peso nos primeiros anos, seguido por uma estabilização. A alta qualidade do ajuste, evidenciada pelos valores de AIC e R^2 , juntamente com a significância dos parâmetros, reforça que o modelo escolhido descreve adequadamente a evolução do peso dos animais ao longo dos anos, proporcionando boas previsões sobre o crescimento dos touros. Isso assegura que as decisões de manejo sejam fundamentadas em dados sólidos e alinhados com o comportamento biológico real.

REFERÊNCIAS

- AK, H. P. W. Machine learning in medicine: A primer for physicians. *Am J Gastroenterol.*, 2010. Citado na página 11.
- AMARAL C.; SOUZA, M. C. T. Um estudo do método de newton-rophson. *Revista eletrônica matemática e estatística em foco.*, p. 42, 2015. Citado na página 17.
- ARAUJO, F. M. Uso da técnica bootstrap em modelos de regressão não linear. *Programa e resumos*, 2013. Citado 2 vezes nas páginas 21 e 22.
- BATES D. M.; WATTS, D. G. Nonlinear regression: iterative estimation and linear approximations. *Nonlinear regression analysis and its applications*, Wiley Hoboke, NJ, p. 32–66, 1988. Citado 2 vezes nas páginas 13 e 14.
- DEISENROTH M. P.; FAISAL, A. A. *ONG, C. S. Mathematics for Machine Learning*. [S.l.]: Cambridge University Press, 2020. Citado na página 23.
- DOMINGUES, J. S. Análise do modelo de gompertz no crescimento de tumores sólidos e inserção de um fator de tratamento. *Biomatemática*, v. 21, p. 103–12, 2011. Citado na página 20.
- EFRON B.; TIBSHIRANI, R. Bootstrap methods for standard erros. *Confidence Intervals, and Other Measures of Statistical Accuracy*, v. 1, 1986. Citado 2 vezes nas páginas 21 e 22.
- FRIZZARINI C.; LAURETTO, M. S. Proposta de um algoritmo para indução de árvores de classificação para dados desbalanceados.in: Sbc. *Anais do IX Simpósio Brasileiro de Sistemas de Informação*. [S.l.], p. 722–733, 2013. Citado na página 26.
- IGNACIO, L. F. F. *Aprendizado de máquinas: da teoria à aplicação.*, Volta Redonda, 2021. Citado na página 23.
- LOPES, L. P. Poder preditivo de métodos clássicos e de statistical machine learning na classificação de dados desbalanceados em seguros. *Revista de Finanças e Contabilidade da Unimep*, v. 5, n. 2, p. 88–109, 2018. Citado na página 26.
- MATTOS, T. d. B. Modelos não lineares e suas aplicações. Juíz de Fora, 2013. Citado 3 vezes nas páginas 14, 15 e 17.
- MAZUCHELI J.; ACHCAR, J. A. Algumas considerações em regreção não linear. *Acta Scientiarum.*, v. 24, n. 6, p. 1761–1770, 2002. Citado 4 vezes nas páginas 13, 14, 18 e 19.
- MUELLER A. C.; GUIDO, S. *Introduction to Machine Learning with Python: A guide for data scientist*. [S.l.]: O’Reilly Media, 2016. Citado 3 vezes nas páginas 23, 26 e 27.
- MÁXIMO, K. A. M. Compreendendo o modelo de gompertz e suas aplicações. *Revista Ibero-Americano de Humanidades, Ciências e Educação*, v. 9, n. 3, p. 2134–2141, 2023. Citado na página 19.
- OLIVEIRA, L. Modelo de van bertalanffy generalizado aplicado ao crescimento de sumos de corte. *IMECC-UNICAMP*, p. 101, 2007. Citado na página 20.
- RATKOWSKY, D. A. Handbook of nonliear regression models. *New Yourk: Marcel Dekker*, p. 241, 1990. Citado na página 11.

RAZALI N. M.; WAH, Y. B. e. a. Power comparisons of sapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, v. 2, n. 1, p. 21–23, 2011. Citado na página 22.

S., A. B. *IMPORTÂNCIA DA RAÇA HOLANDESA NA PECUÁRIA*. 2023. Disponível em: <https://repositorio.unesp.br/server/api/core/bitstreams/3017bbbb-7c37-49d4-9ee0-672f5e31aa00/content>. Citado na página 37.

SANTOS, A. L. P. Estudo de modelo de crescimento via dados simulados. *Monografia (Graduação) - Universidade Estadual da Paraíba*, Campina Grande, 2011. Citado 4 vezes nas páginas 19, 21, 24 e 26.

SEBER G.; WILD, C. Measures of curvatures and nonlinearity. *Nonlinear Regression*, Hoboken, NJ: John Wiley & Sons, p. 127–190, 2003. Citado na página 19.

SHALEV-SHWARTZ S.; BEN-DAVID, S. *Understanding Machine Learning: From Theory to Algorithms.*, [S.1]: Cambridge University Press, 2014. Citado 2 vezes nas páginas 24 e 25.

SOUZA, E. M. Modelagem não linear da extração de zinco de um solo tratado com lodo de esgoto. *Acta Scientiarum. Technology*, Universidade Estadual de Maringa, v. 32, n. 2, p. 193–199, 2010. Citado 3 vezes nas páginas 11, 20 e 22.

SOUZA, G. d. S. e. Introdução aos modelos de regressão linear e não-linear. *Brasília: Embrapa SPI: Embrapa-seA*, p. 489, 1998. Citado na página 11.

TELEKEN J. T.; GALVÃO, A. C. R. W. S. Comparing non-linear mathematical models to describe growth of different animals. *Acta Scientiarum. Animal Sciences*, v. 39, n. 1, p. 73–81, 2017. Citado na página 29.

THOMAS, G. Regressão não linear. *Trabalho de conclusão do curso de pós-graduação em estatística e experimentação agro pecuária. Escola superior de Agricultura Luiz de Queiroz-USP*, 2016. Citado 2 vezes nas páginas 15 e 18.

TUKEY, J. W. *O futuro da análise de dados. Anais de Estatística Matemática*. 1962. Disponível em: <https://citacoes.in/autores/john-tukey/>. Citado na página 5.

Z., E. E. O. Predictiing the future-big data, machine learning, and clinical medicibe. *N Engl J Med.*, 2016. Citado na página 11.

ZEVIANI P. J. W.; BONAT, W. Modelos de regressão não linear. *Curitiba: W.M.Z., P.J.R.Jr; W.H.B.*, p. 101, 2013. Citado na página 20.