

Garantindo a Segurança e Bem-Estar dos Pandas: Detecção de Animais no Zoológico com Aprendizado de Máquina

Alisson Rodrigo Carneiro da Silva, João dos Santos Neto,
Marcos Vinicius Batista Sampaio, Jorge Luis Ferreira Luz

¹PAAD - Núcleo Pesq. Avanç. Anál. Dados - Universidade Federal do Piauí (UFPI)
Picos – PI – Brazil

²PAVIC - Lab. Pesq. Aplic. Visão e IA – Universidade Federal do Piauí (UFPI)
Picos – PI – Brazil

jorge.luz@ufpi.edu.br, alisson.silva@ufpi.edu.br, joaonetoprivado2001@ufpi.edu.br,
socialcraft.sc@gmail.com

Abstract. *Detecting pandas in a zoo is crucial for the safety of animals, visitors, and monitoring the well-being of resident pandas. This process involves visual identification and distinguishing between pandas, dogs, and cats. To address this challenge, we employ computer vision and deep learning techniques to train a model capable of recognizing distinct patterns among these three species. This summary discusses the construction of a machine learning model for panda detection in the zoo, with a focus on image attribute concepts.*

Resumo. *A detecção de pandas em um zoológico é crucial para a segurança dos animais, visitantes e o monitoramento do bem-estar dos pandas residentes. Este processo envolve a identificação visual e a distinção entre pandas, cachorros e gatos. Para solucionar esse desafio, utilizamos visão computacional e aprendizado profundo para treinar um modelo capaz de reconhecer padrões distintos entre essas três espécies. Este resumo aborda a construção de um modelo de aprendizado de máquina para a detecção de pandas no zoológico, com foco em conceitos de atributos de imagem.*

1. Introdução

A detecção de animais em rodovias é uma tarefa crucial para garantir a segurança da fauna selvagem e dos motoristas, bem como para contribuir para a preservação das espécies. Neste contexto, abordaremos a detecção de pandas, um exemplo emblemático, mas também de outros animais, como cachorros e gatos, em determinadas áreas de interesse ao longo das rodovias, fazendo uso de um modelo de aprendizado de máquina.[Sato 2022] A detecção de pandas vai além da simples identificação visual, envolvendo a capacidade de distinguir esses animais de outras espécies, como cachorros e gatos, que podem cruzar as estradas. Para superar esse desafio, faremos uso de técnicas avançadas de visão computacional e aprendizado profundo para treinar um modelo capaz de identificar padrões distintivos que diferenciam essas espécies. Neste contexto introdutório, abordaremos a detecção de animais em rodovias, focando especificamente na detecção de pandas, como um estudo de caso, por meio da construção de um modelo de aprendizado de máquina. Para isso, exploraremos conceitos relacionados aos atributos das imagens que nos permitirão realizar a detecção dos respectivos animais na área de interesse. Este estudo tem implicações

importantes na promoção da coexistência segura entre a vida selvagem e as rodovias, contribuindo assim para a conservação e segurança de ambas as partes envolvidas.

2. Metodologia

2.1. Base de dados

Em nossa pesquisa, utilizamos uma rica base de dados composta por 3.000 imagens, com igual distribuição de 1.000 imagens de pandas, 1.000 de gatos e 1.000 de cachorros. Essa base de dados desempenhou um papel fundamental em nossos estudos, permitindo o treinamento e teste de algoritmos de aprendizado de máquina para classificação de imagens. Essa diversidade de classes tornou possível abordar desafios de identificação e distinção entre essas três espécies, com implicações práticas em áreas como reconhecimento de animais, automação de processos e segurança.

2.2. Atributo de Média

A média é uma medida de tendência central que representa o valor médio dos tons de cinza na imagem. O cálculo desse processo ocorre da seguinte maneira: somam-se todos os valores dos pixels na imagem e divide-se pela quantidade total de pixels. A fórmula é basicamente a média, dada por:

$$\text{Média} = \frac{1}{N} \sum_{i=1}^N X_i$$

onde N é o número total de pixels, e X_i representa cada valor de pixel na imagem.

Este atributo fornece uma representação numérica da intensidade média dos tons de cinza na imagem. É uma medida fundamental em análises de imagem e processamento de dados visuais, contribuindo para a compreensão da luminosidade geral da cena capturada.

2.3. Atributo de Variância

A variância é uma medida que avalia a dispersão dos valores de pixel em uma imagem, indicando o quão diferentes esses valores são da média. O cálculo envolve a média dos quadrados das diferenças entre os valores de pixel e a média da imagem. A fórmula é dada por:

$$\text{Variância} = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2$$

onde N é o número total de valores de pixel, X_i representa cada valor de pixel, e \bar{X} é a média dos valores de pixel na imagem.

A variância fornece insights sobre a distribuição dos valores de pixel, sendo uma medida importante em análises de qualidade de imagem e processamento de dados visuais. Valores mais altos indicam uma maior dispersão em torno da média, enquanto valores mais baixos sugerem uma distribuição mais concentrada.

2.4. Atributo de *Skewness*

O skewness mede o grau de assimetria da distribuição dos valores de pixel. Um valor positivo indica uma cauda longa à direita, e um valor negativo indica uma cauda longa à esquerda. Calcula-se a média dos cubos das diferenças entre os valores de pixel e a média, normalizada pela variância, representada pela fórmula:

$$\text{Skewness} = \frac{\sum_{i=1}^N (X_i - \bar{X})^3}{N \cdot \sigma^3}$$

onde N é o número de observações, X_i é cada valor de pixel, \bar{X} é a média dos valores de pixel e σ^2 é a variância.

2.5. Atributo de *Kurtosis*

O kurtosis mede o grau de concentração dos valores de pixel em torno da média. Valores altos indicam uma distribuição mais concentrada. O cálculo envolve a média dos quartos poderes das diferenças entre os valores de pixel e a média, normalizada pela variância, conforme a fórmula:

$$\text{Kurtosis} = \frac{\sum_{i=1}^N (X_i - \bar{X})^4}{N \cdot \sigma^4}$$

onde N é o número de observações, X_i é cada valor de pixel, \bar{X} é a média dos valores de pixel e σ^2 é a variância.

2.6. Atributo de Energia

Definição: A energia é uma medida da quantidade de informação ou desordem na imagem. Imagens com alta entropia são mais complexas. Para calcular a energia, somamos os quadrados dos valores de pixel na imagem. A fórmula é dada por:

$$\text{Energia} = \sum_{i=1}^N (X_i)^2$$

onde N é o número de valores de pixel na imagem, e X_i representa cada valor de pixel. A energia é uma métrica que ajuda a avaliar o nível de complexidade ou informação em uma imagem, sendo útil em várias aplicações de processamento de imagens e análise de texturas.

2.7. Atributo de Entropia

Definição: A entropia é uma medida da incerteza ou desordem na distribuição dos valores de pixel em uma imagem. Ela indica o grau de informação contida na imagem. Calcula-se a entropia a partir da distribuição de probabilidade dos valores de pixel na imagem. Quanto mais uniforme a distribuição, maior a entropia, indicando mais informação na imagem. A fórmula é dada por:

$$\text{Entropia} = - \sum_{i=1}^N (\text{Probabilidade de intensidade do pixel} \cdot \log_2(\text{Probabilidade de intensidade do pixel}))$$

onde a soma é realizada para todas as intensidades possíveis de pixels na imagem.

Após a implementação do modelo de aprendizado de máquina para a detecção de pandas no zoológico, obtivemos resultados que são essenciais para avaliar o desempenho da solução. A acurácia do modelo foi medida em 34%, o que significa que o modelo classificou corretamente 34% das instâncias.

2.8. Alargamento de Contraste

O alargamento de contraste é uma técnica utilizada no processamento de imagens para realçar a diferença entre os tons de cinza, resultando em uma imagem com maior gama de intensidades. A fórmula que descreve o alargamento de contraste é dada por:

$$s = T(r) = \frac{1}{1 + \left(\frac{k}{r}\right)^E}$$

onde s é o valor do pixel após o alargamento, r é o valor original do pixel, k é um parâmetro de ajuste e E é a potência aplicada.

O alargamento de contraste é especialmente útil quando a distribuição dos tons de cinza na imagem está concentrada em uma faixa estreita, resultando em uma imagem com detalhes menos distintos. A aplicação desta fórmula amplia a gama de intensidades, melhorando a visibilidade de detalhes na imagem.

Essa técnica é valiosa em diversas aplicações, incluindo melhorias visuais em fotografias e análises de imagem em processos de visão computacional.

2.9. Negativo

O negativo é uma transformação comum aplicada a imagens para inverter os tons de cinza. Essa técnica produz uma imagem em que os pixels mais escuros tornam-se mais claros e vice-versa. A fórmula que descreve a transformação para o negativo é dada por:

$$s = L - 1 - r$$

onde s é o valor do pixel após a transformação, r é o valor original do pixel e L é o número total de níveis de intensidade (geralmente 256 em imagens de 8 bits).

A aplicação da transformação de negativo resulta em uma imagem com inversão de cores, proporcionando destaque aos detalhes que podem não ser prontamente visíveis na imagem original.

Essa técnica é amplamente utilizada em processamento de imagem para realçar características específicas, como bordas e contornos, e pode ser aplicada em diversas aplicações, desde melhorias visuais até análises mais avançadas de dados visuais.

2.10. Logarítmico

A transformação logarítmica é uma técnica empregada em processamento de imagens para realçar detalhes em regiões de baixo contraste, especialmente quando os valores dos pixels são predominantemente escuros. A fórmula que descreve a transformação logarítmica é dada por:

$$s = c \cdot \log(1 + r)$$

onde s é o valor do pixel após a transformação, r é o valor original do pixel e c é uma constante de ajuste.

A aplicação da transformação logarítmica tem o efeito de espalhar os tons de cinza, realçando detalhes em áreas de baixo contraste e melhorando a visualização de características sutis na imagem.

Esta técnica é valiosa em situações em que há predominância de pixels escuros e é especialmente útil em aplicações de melhoria visual, análise de texturas e detecção de padrões em imagens.

2.11. Potência

A transformação de potência é uma técnica comumente utilizada em processamento de imagens para ajustar o contraste, destacando detalhes específicos na imagem. A fórmula que descreve a transformação de potência é dada por:

$$s = c \cdot r^\gamma$$

onde s é o valor do pixel após a transformação, r é o valor original do pixel, c é uma constante de ajuste e γ é o parâmetro de potência.

A aplicação da transformação de potência tem o efeito de realçar detalhes em áreas específicas da imagem, tornando-a mais adaptável a diferentes condições de contraste.

Essa técnica é amplamente utilizada em diversas aplicações de processamento de imagens, proporcionando flexibilidade no ajuste de contraste para melhorar a visualização e análise de detalhes em diferentes tipos de imagens.

3. Resultados

3.1. Resultado dos treinamentos utilizando atributos de primeira ordem

Após a implementação do modelo de aprendizado de máquina para a detecção de pandas no zoológico, a acurácia foi de 34%, indicando que o modelo classificou corretamente 34% das instâncias, o que é essencial para avaliar o desempenho da solução.

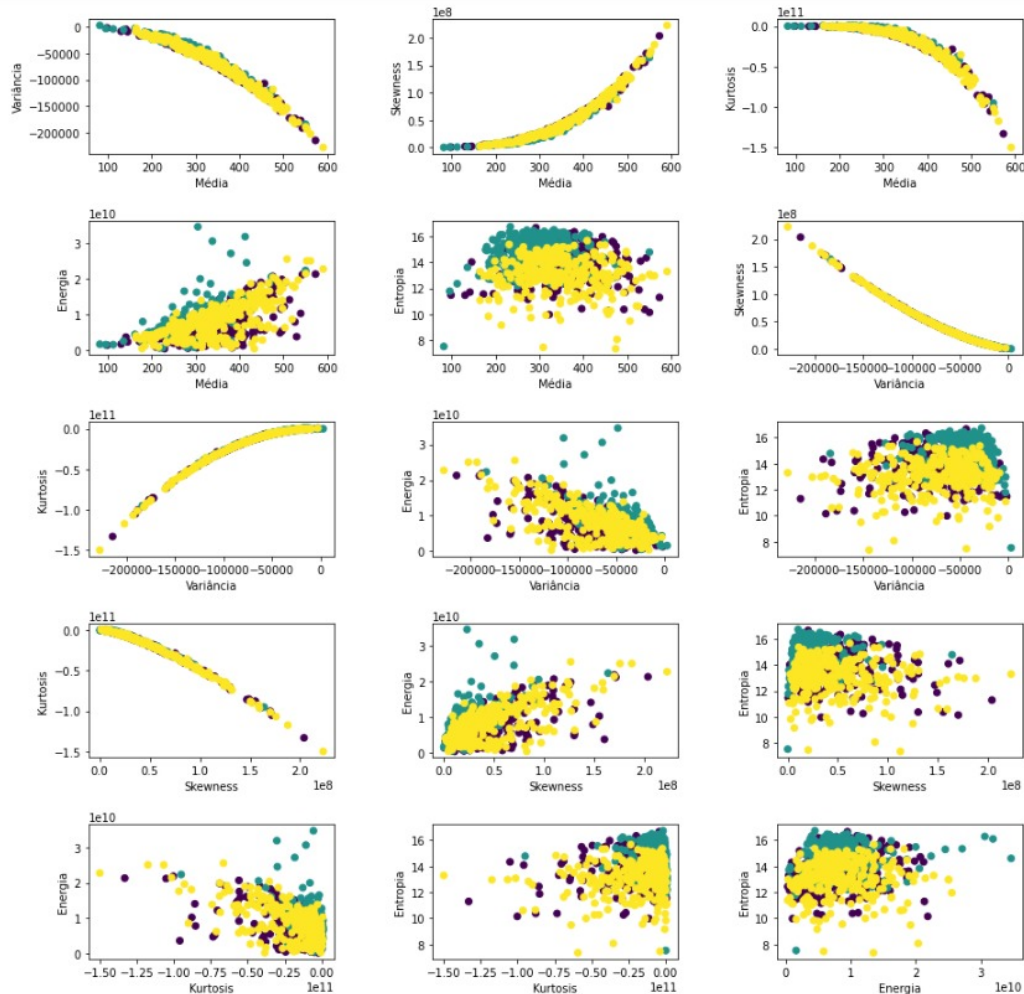


Figura 1. Resultado dos treinamentos

3.2. Resultados após atributos de transformação

Explorar as nuances das transformações de dados, como alargamento, logaritmo e negativa, revela facetas distintas em análises estatísticas. O alargamento expande a faixa de valores, impactando variância e média, enquanto o logaritmo lida com distribuições assimétricas, alterando variância, média e entropia. A transformação negativa, ao inverter valores, destaca padrões nos extremos inferiores, afetando variância, média e entropia. Essas técnicas, embora enriqueçam a compreensão dos dados, desafiam a previsibilidade da acurácia, exigindo uma abordagem criteriosa para interpretar seu impacto nas análises estatísticas.

3.3. Resultados após atributos de Transformação: Alargamento de contraste

Após a aplicação da transformação de alargamento, conduzimos testes de acurácia para avaliar a eficácia da técnica. Surpreendentemente, os resultados revelaram uma acurácia de apenas 0.48. Isso sugere que, embora a transformação de alargamento possa ter influenciado certos aspectos estatísticos dos dados, ela pode não ter sido eficaz para melhorar a capacidade preditiva ou discriminativa do modelo.

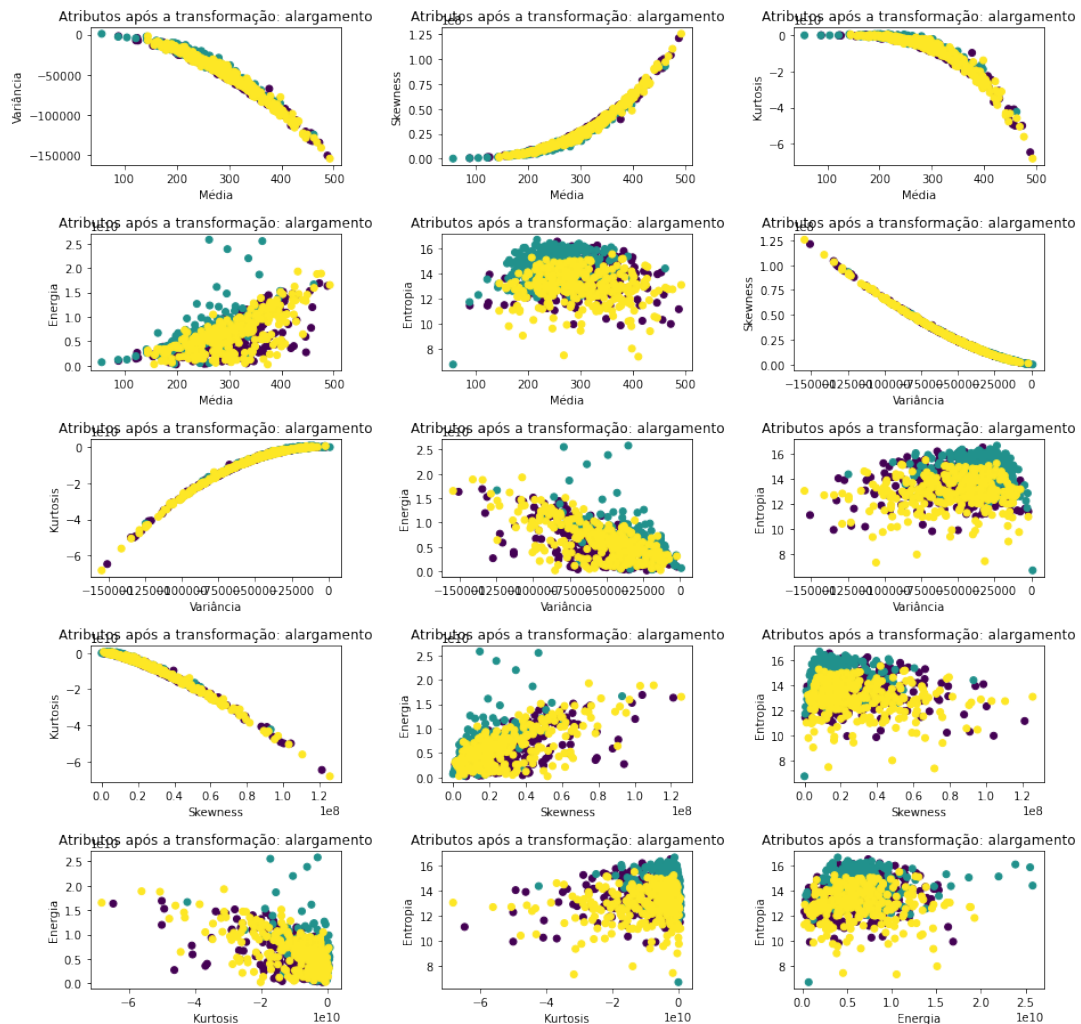


Figura 2. Resultado após transformação: Alargamento de contraste

3.4. Resultados após atributos de Transformação: Logarítmico

Surpreendentemente, os resultados dos testes de acurácia após a aplicação da transformação logarítmica revelaram uma taxa de 0.49. Este resultado sugere que, embora a transformação logarítmica tenha influenciado certos aspectos estatísticos, ela ainda não conseguiu proporcionar uma melhoria significativa na capacidade preditiva ou discriminativa do modelo em comparação com a acurácia inicial.

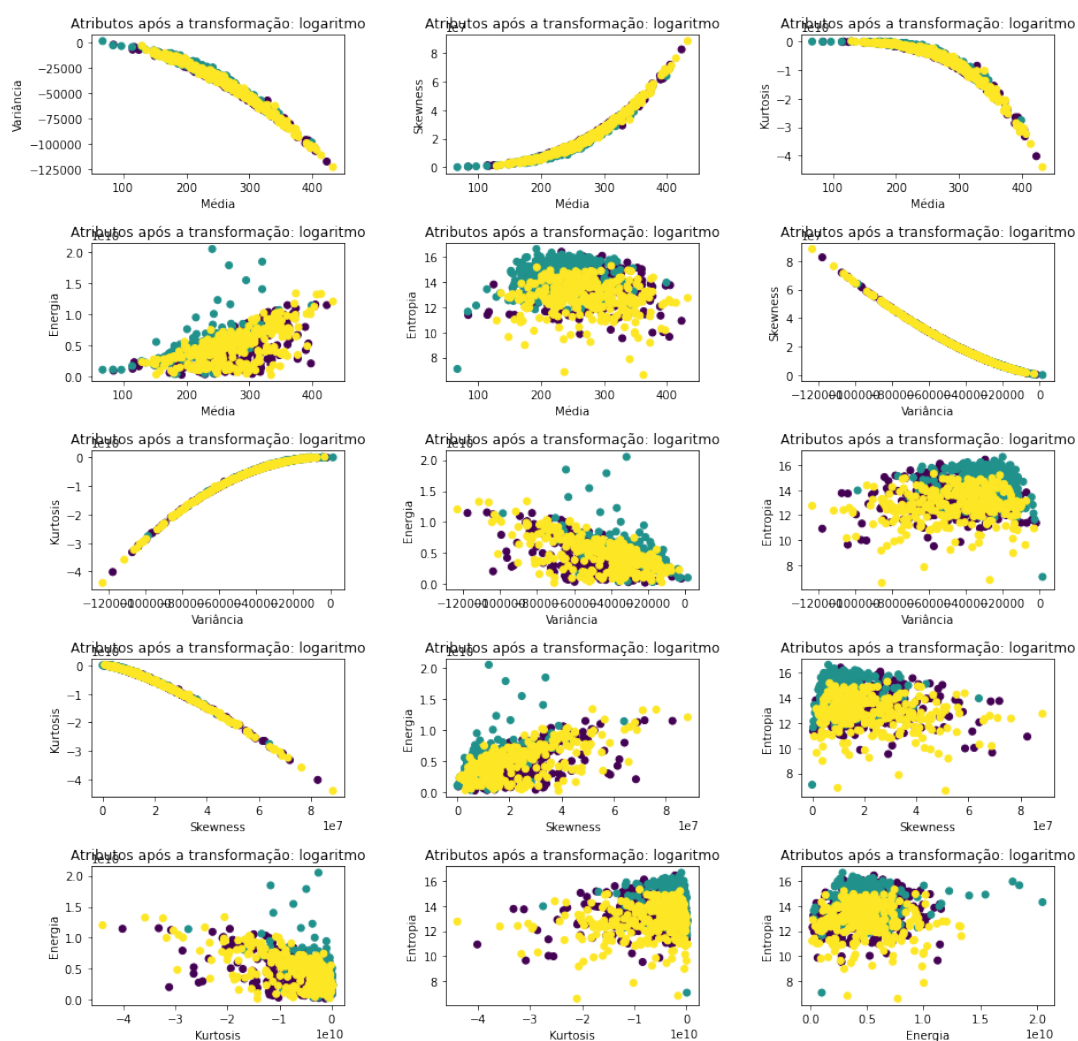


Figura 3. Resultado após transformação: Logaritmo

3.5. Resultados após atributos de Transformação: Negativo

Os testes de acurácia revelaram um resultado de 0.29 após a implementação da transformação negativa. Este valor sugere que, apesar das alterações nos atributos estatísticos, a transformação negativa não resultou em uma melhoria significativa na capacidade preditiva ou discriminativa do modelo.

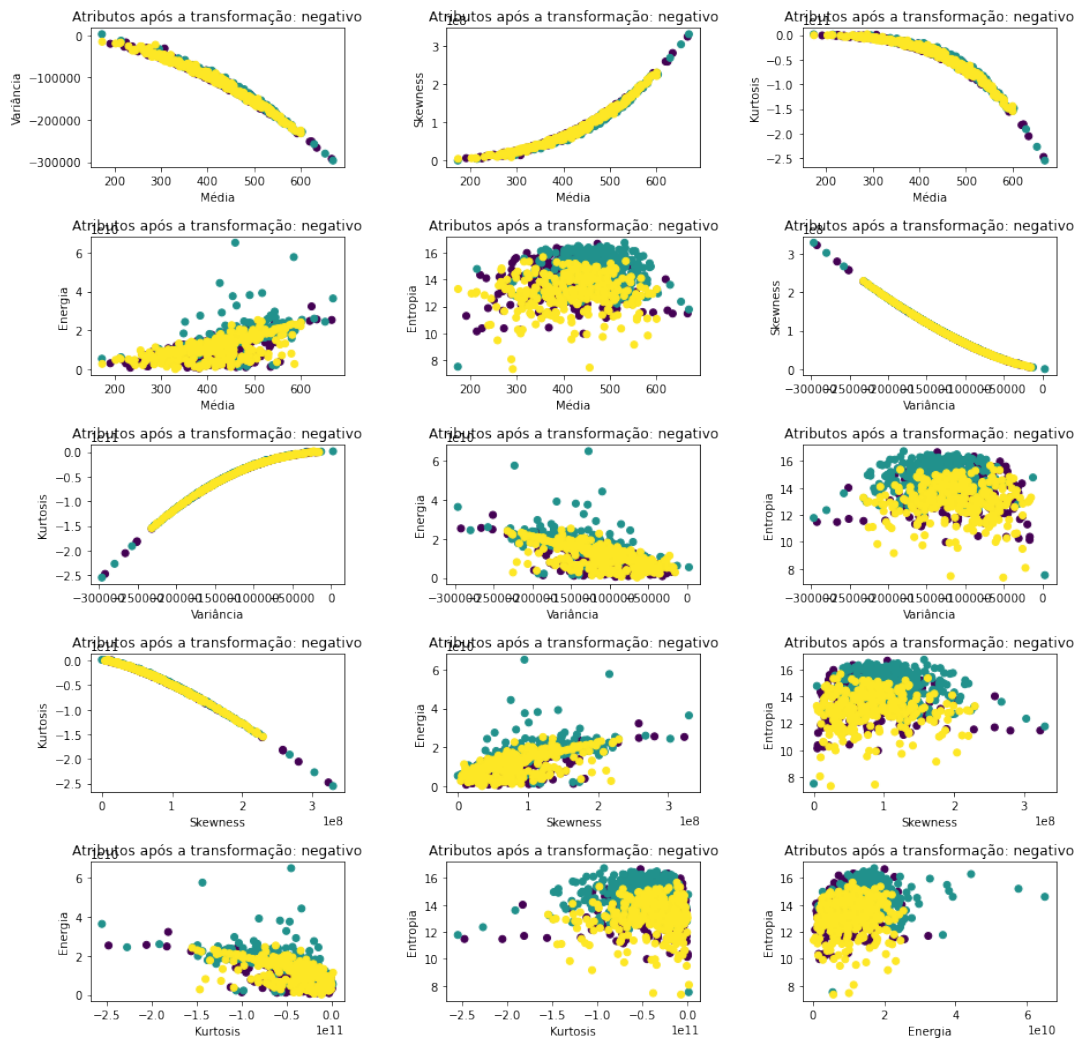


Figura 4. Resultado após transformação: Negativo

3.6. Resultados após atributos de Transformação: Porência

Os testes de acurácia revelaram um resultado de 0.45 após a implementação das transformações negativas. Este valor sugere que, embora tenha havido alterações nos atributos estatísticos, as transformações negativas não resultaram em uma melhoria substancial na capacidade preditiva ou discriminativa do modelo.

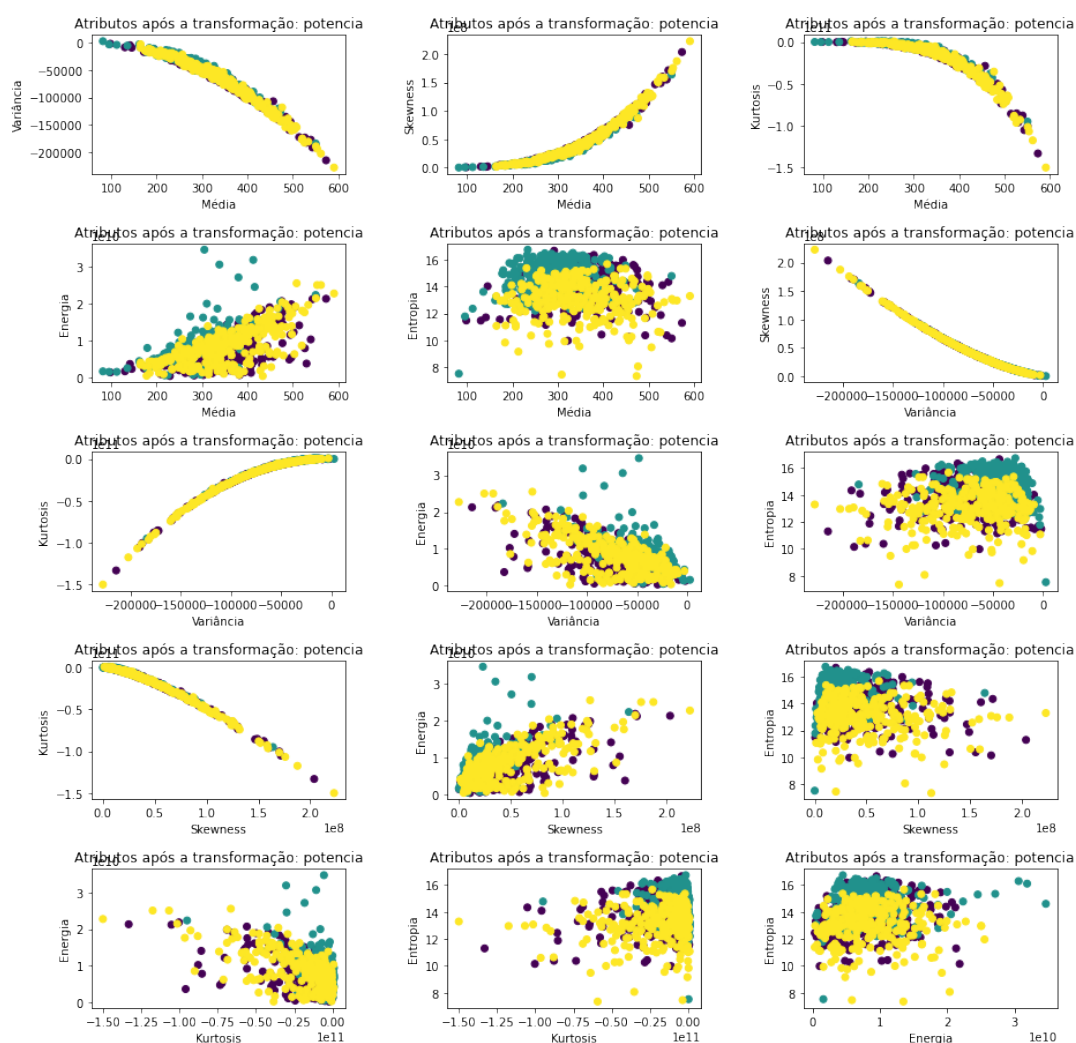


Figura 5. Resultado após transformação: Potência

Referências

Sato, D. (2022). Sistema de detecção de animais nas rodovias com aprendizado de máquina e visão computacional.