



UNIVERSIDADE FEDERAL DO PARÁ
CAMPUS UNIVERSITARIO DE CASTANHAL
FACULDADE DE MATEMÁTICA

ELIZÂNGELA MARIA GONÇALVES SILVA

MODELAGEM MATEMÁTICA E INTELIGÊNCIA ARTIFICIAL: :
Implicações na formação docente em licenciatura em matemática

CASTANHAL - PA
2025

ELIZÂNGELA MARIA GONÇALVES SILVA

MODELAGEM MATEMÁTICA E INTELIGÊNCIA ARTIFICIAL: :

Implicações na formação docente em licenciatura em matemática

Portfólio acadêmico apresentado à Faculdade de Matemática - Campus Castanhal. Esta pesquisa foi elaborada por fim de Trabalho de Conclusão de Curso, sob a orientação do Prof. Dr. Renato Germano Reis Nunes, da Faculdade de Matemática - UFPA, como requisito parcial para a obtenção do título de Licenciado em Matemática.

Orientador: Prof. Dr. Renato Germano Reis Nunes

CASTANHAL - PA
2025

ELIZÂNGELA MARIA GONÇALVES SILVA

MODELAGEM MATEMÁTICA E INTELIGÊNCIA ARTIFICIAL: :

Implicações na formação docente em licenciatura em matemática

Portfólio acadêmico apresentado à Faculdade de Matemática - Campus Castanhal. Esta pesquisa foi elaborada por fim de Trabalho de Conclusão de Curso, sob a orientação do Prof. Dr. Renato Germano Reis Nunes, da Faculdade de Matemática - UFPA, como requisito parcial para a obtenção do título de Licenciado em Matemática.

DATA DE APROVAÇÃO: 15/09/2025

CONCEITO: Excelente

Prof. Dr. Renato Germano Reis Nunes
Orientador - Universidade Federal do Pará

Prof. M. Eng. José Geraldo Gonçalves da
Silva
Membro - Universidade Federal do Pará

Prof. Dr. Nildsen Fernando Lisbôa da Silva
Membro - Universidade Federal do Pará

CASTANHAL - PA
2025

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

G635m Gonçalves Silva, Elizângela Maria.
 Modelagem Matemática e Inteligência Artificial : Implicações
 na formação docente em licenciatura em matemática / Elizângela
 Maria Gonçalves Silva. — 2025.
 66 f. : il. color.

 Orientador(a): Prof. Dr. Renato Germano Reis Nunes
 Trabalho de Conclusão (Graduação) - Universidade Federal do
 Pará, Campus Universitário de Castanhal, Faculdade de
 Matemática, Castanhal, 2025.

 1. Formação docente. 2. Modelagem matemática. 3.
 Iniciação científica. 4. Inteligência artificial. I. Título.

CDD 510

In Memória: Sra Antônia Gonçalves , minha amada avó , que me incentivou profundamente nessa trajetória acadêmica.

AGRADECIMENTOS

Agradeço à ciência que nos inspira a busca pelo saber e nos transforma positivamente para a vida.

A filosofia, a educação e a matemática, que me permitiram ver o mundo com outros olhos, sem dúvida, mudaram a minha realidade.

À minha querida avó materna, Sra. Antônia Gonçalves, meu profundo agradecimento pelo seu incentivo fundamental à minha vida acadêmica. Suas palavras foram o alicerce que me guiou até aqui.

Aos meus pais, Maria e Joaquin, por todo o apoio incondicional e, sobretudo, por sempre acreditarem no meu potencial.

Ao meu irmão, Gabriel, pelo amor e pelo afeto que transcendem as palavras. Seus abraços foram meu porto seguro nos momentos mais desafiadores desta trajetória e foram essenciais para que eu não desistisse.

Aos meus amigos, Elizabeth Lima, José Bruno e José Luiz, pela amizade e pelo incentivo durante o curso.

Ao meu orientador, Renato Germano, que foi essencial no desenvolvimento deste portfólio.

A cada pessoa que contribuiu direta e indiretamente em minha trajetória acadêmica e de vida.

“Modelagem matemática é a matemática por excelência” (Ubiratan D’Ambrosio)

RESUMO

Este portfólio acadêmico reflete a trajetória de formação docente na Licenciatura em Matemática da UFFA, Campus Castanhal, centrada na articulação entre iniciação científica, modelagem matemática e inteligência artificial e visão computacional. A pesquisa apresenta uma análise reflexiva de duas produções principais: um capítulo de livro sobre o desenvolvimento de um protótipo de baixo custo para separação de grãos utilizando visão computacional e redes neurais convolucionais (CNNs) com Arduino, e um artigo que aplica a Lei de Zipf na análise comparativa de textos em português e inglês. A metodologia qualitativa e memorialista, na qual permite explorar como a modelagem matemática e o uso crítico de tecnologias, contribuíram para a constituição de uma identidade de professora-pesquisadora. Os resultados demonstram não apenas a viabilidade técnica e educacional dos projetos desenvolvidos, mas também a ampliação da perspectiva sobre o ensino da matemática, enfatizando a interdisciplinaridade, contextualização e resolução de problemas reais. Tais contribuições sobre integração entre teoria e prática, mediada pela pesquisa, é fundamental para uma formação docente crítica, criativa e alinhada às demandas contemporâneas, capacitando futuros educadores para implementar abordagens inovadoras em sala de aula e para atuar como agentes de transformação social.

Palavras-chave: Formação docente. Modelagem matemática. Iniciação científica. Inteligência Artificial.

ABSTRACT

This academic portfolio reflects the teacher training journey in the Mathematics Degree course at UFPA, Castanhal Campus, centered on the connection between scientific initiation, mathematical modeling, artificial intelligence, and computer vision. The research presents a reflective analysis of two main works: a book chapter on the development of a low-cost prototype for grain separation using computer vision and Convolutional Neural Networks (CNNs) with Arduino, and an article that applies Zipf's Law in the comparative analysis of texts in Portuguese and English. The qualitative and memorialistic methodology allows for an exploration of how mathematical modeling and the critical use of technologies contributed to the constitution of a teacher-researcher identity. The results demonstrate not only the technical and educational feasibility of the developed projects but also the broadening of the perspective on mathematics teaching, emphasizing interdisciplinarity, contextualization, and the resolution of real-world problems. These contributions highlight that the integration between theory and practice, mediated by research, is fundamental for a critical and creative teacher education aligned with contemporary demands, empowering future educators to implement innovative approaches in the classroom and to act as agents of social transformation.

Keywords: Teacher education. Mathematical modeling. Scientific initiation. Artificial Intelligence.

LISTA DE QUADROS

| | |
|--|----|
| Quadro 1 – Principais Produções acadêmicas | 17 |
|--|----|

LISTA DE SIGLAS

| | |
|--------|--|
| CCAST | Campus Universitário de Castanhal |
| CNNs | Redes Neurais Convolucionais |
| EPAEM | Encontro Paraense de Educação Matemática |
| EPAMM | Encontro Paraense de Modelagem Matemática |
| IA | Inteligência Artificial |
| LEMM | Laboratório Experimental de Modelagem Matemática |
| SAMATC | Semana Acadêmica de Matemática de Castanhal |
| UFPA | Universidade Federal do Pará |

SUMÁRIO

| | | |
|-----|--|-----------|
| 1 | INTRODUÇÃO | 12 |
| 2 | MODELAGEM MATEMÁTICA NA FORMAÇÃO DOCENTE . . . | 13 |
| 3 | USO DE TECNOLOGIAS NO ENSINO DE MATEMÁTICA | 15 |
| 4 | CONTRIBUIÇÕES ACADÊMICAS | 17 |
| 5 | METODOLOGIA | 19 |
| 5.1 | TIPO DE PESQUISA | 19 |
| 5.2 | INSTRUMENTOS METODOLÓGICOS | 19 |
| 5.3 | ANÁLISE REFLEXIVA DAS PRODUÇÕES | 20 |
| 6 | CONSIDERAÇÕES SOBRE CAPÍTULO DE LIVRO: UTILIZAÇÃO DE VISÃO COMPUTACIONAL E DEEP LEARNING PARA SE- PARAÇÃO DE GRÃOS: UM PROTÓTIPO DE BAIXO CUSTO BASEADO NO ARDUINO. | 21 |
| 7 | CONSIDERAÇÕES SOBRE O ARTIGO: MODELAGEM MATEMÁ- TICA DA LEI DE ZIPF: UMA ANÁLISE COMPARATIVA ENTRE TEXTOS EM PORTUGUÊS E INGLÊS | 24 |
| 8 | CONSIDERAÇÕES FINAIS | 26 |
| | REFERÊNCIAS | 27 |
| | APÊNDICES | 29 |
| | APÊNDICE A – UTILIZAÇÃO DE VISÃO COMPUTACIONAL E DEEP LEARNING PARA SEPARAÇÃO DE GRÃOS: UM PROTÓTIPO DE BAIXO CUSTO BASEADO NO ARDUINO | 30 |
| | APÊNDICE B – MODELAGEM MATEMÁTICA DA LEI DE ZIPF: UMA ANÁLISE COMPARATIVA ENTRE TEX- TOS EM PORTUGUÊS E INGLÊS | 53 |
| | ANEXOS | 61 |
| | ANEXO A – CERTIFICADO DE ACEITE PARA PUBLICAÇÃO NA EDITORA CIENTÍFICA DIGITAL | 62 |
| | ANEXO B – CERTIFICADO DE ACEITE NO IX ENCONTRO PA- RAENSE DE MODELAGEM MATEMÁTICA | 64 |

1 INTRODUÇÃO

A minha trajetória na licenciatura em matemática na Universidade Federal do Pará, Campus Castanhal, está atrelada com a iniciação científica e as produções deste portfólio. Desde os primeiros períodos, sob o incentivo do professor Renato Germano, me levaram a um engajamento com modelagem matemática e tecnologia para resolver problemas do mundo real. Esta vivência prática permitiu unir a teoria aprendida em sala de aula com atividades de pesquisa, despertando em mim a identidade de uma professora-pesquisadora.

Paralelamente, a escrita acadêmica e a participação nas disciplinas laboratório experimental de modelagem matemática (LEMM) e prática de ensino por modelagem matemática foram processos indissociáveis dessa jornada. Cada produção científica, submetido a eventos nacionais, representou muito mais do que uma produção técnica foi um espaço de incentivo a formação docente alinhada a pesquisa.

A imersão em projetos com inteligência artificial e visão computacional, ampliou minha percepção sobre o potencial das tecnologias educacionais. Compreendi que ferramentas como o Arduino, o Teachable Machine e o Pictoblox são recursos pedagógicos que, quando mediados de forma crítica e criativa, podem tornar a aprendizagem da matemática mais dinâmica, investigativa e conectada com as demandas contemporâneas.

Considerando as prerrogativas Universidade Federal do Pará (2024, p. 4-5) estabelece as normas regimentais para a elaboração e apresentação do trabalho de conclusão de curso (TCC) do curso de Licenciatura em Matemática, opto-se por trazer um capítulo de livro e um pôster.

Os artigos **Utilização de visão computacional e deep learning para separação de grãos**: um protótipo de baixo custo baseado no Arduino e **Modelagem matemática da lei de Zipf**: uma análise comparativa entre textos em português e inglês tem como objetivos trazer o reflexo desse processo integrado na minha formação docente. Mais do que listar produções acadêmicas, pretendo apresentar como a modelagem matemática e o uso crítico de tecnologias emergentes atuaram como eixos centrais na construção do saber/fazer pedagógico.

2 MODELAGEM MATEMÁTICA NA FORMAÇÃO DOCENTE

Essa seção representa uma dialogo conceitual entre os principais autores da modelagem matemática brasileira, sem um aprofundamento teórico. Visto que, o proposito é trazer suas principais ideias que possam influenciar na formação docente, trazidas aqui, em uma perspectiva de complementaridade de diferentes ênfases.

Inicialmente, parte-se da conceituação informal de modelagem matemática:

A modelagem matemática consiste na arte de transformar problemas da realidade em problemas matemáticos e resolvemos interpretando suas soluções na linguagem do mundo real (BASSANEZI, 2002, p. 16).

Pode aparecer a primeira vista, uma visão focada método e no processo de modelagem. Entretanto, Bassanezi (2002, p. 291) ressalta que esse método de trabalho na iniciação científica tem sido desenvolvido com sucesso entre estudantes universitários de graduação e pós-graduação de Matemática. Ele é mais fácil e produtivo quando os alunos fazem parte de um grupo maior de pessoas que pesquisam um determinado tema. A iniciação científica, como um projeto de estudo dirigido, facilita a combinação de teoria e prática, fomentando a aprendizagem de disciplinas básicas ao valorizá-las e recriar suas ideias quando aplicadas à realidade. Ela serve como uma ponte natural entre o ensino e a pesquisa.

Em uma outra linha, modelagem matemática é vista como:

A Modelagem Matemática constitui-se em um conjunto de procedimentos cujo objetivo é construir um paralelo para tentar explicar, matematicamente, os fenômenos presentes no cotidiano do ser humano, ajudando-o a fazer predições e a tomar decisões. (BURAK, 1992, p. 62).

Partindo desse pressuposto, modelagem é, antes de tudo, um ambiente de aprendizagem. Um espaço que busca contribuir para o preparo e a mudança de postura do professor em relação ao ensino de Matemática. A modelagem visa transformar o professor em um efetivo agente de transformação social (BURAK, 1992). Tal colocação afasta a visão puramente técnica da modelagem matemática, aproximando-se de uma preparação didático-pedagógica centrada na dimensão social e mediada pela interação.

No processo formativo em matemática, observa que, modelagem não é só um método, é uma postura pedagógica que requer professores preparados para mediar conflitos,

estimular diálogos e guiar os alunos em um processo de construção social do conhecimento matemático (BURAK; ZONTINI, 2020).

Trazendo esse universo conceitual, com foco na realidade local do nordeste paraense, o Laboratório Experimental de Modelagem Matemática (LEMM), situado na Universidade Federal do Pará (UFPA), Campus Universitário de Castanhal (CCAST), destaca-se como um espaço que contribui significativamente para o processo de formação docente e para a prática da modelagem matemática (SOARES, 2023).

Nesse ambiente, os estudantes são ativamente engajados em seus processos de aprendizagem, o que resulta em um aprendizado reflexivo, crítico e significativo. O laboratório permite aos participantes interagirem motivados por um tema de investigação, com o objetivo de solucionar coletivamente situações da vida real. As práticas desenvolvidas incluem pesquisas programadas, experimentos simples, revisão bibliográfica e a produção de textos acadêmicos para submissão em eventos científicos, além do estímulo à criação de atividades didático-pedagógicas com o uso da modelagem matemática (BRAGA, 2019; SOARES, 2023).

As vivências/experiências mencionadas por Soares (2023), Braga (2021) nesse espaço pedagógico, proporcionam um ambiente de aprendizagem que estimula a iniciação científica e a produção de trabalhos acadêmico-científicos. Essa experiência amplia a perspectiva dos graduandos sobre o ensino da matemática, motivando-os a serem mais comprometidos e proativos em sua jornada acadêmica. O LEEM trás ao futuro professor de matemática uma preparação para implementar a modelagem matemática em sala de aula de maneira efetiva na educação básica.

3 USO DE TECNOLOGIAS NO ENSINO DE MATEMÁTICA

A seção apresenta uma discussão introdutória sobre o uso de inteligência artificial (IA) com implicações no ensino de matemática. Visto que, uma das produções científicas deste portfólio utiliza tal ferramenta. Assim, o leitor situa-se conceitualmente, das motivações que levaram a pesquisadora a investigar sobre uso de IA e visão computacional.

A Inteligência Artificial (IA) emergiu como uma das inovações tecnológicas mais significativas do século XXI, com profundas implicações para o ensino da Matemática. Sua integração na educação é vista como um caminho para aprimorar o processo de ensino-aprendizagem, embora também apresente desafios consideráveis (COSTA *et al.*, 2025).

Um dos maiores benefícios da IA no ensino da Matemática é sua capacidade de personalização. Ela permite adaptar o conteúdo às necessidades individuais dos alunos, identificando dificuldades específicas e ajustando o ritmo de aprendizado, o que é crucial em turmas com disparidade de compreensão. A IA pode fornecer feedback imediato e detalhado sobre exercícios e problemas, explicando erros passo a passo e ajudando na correção de lacunas conceituais (OLIVEIRA; SILVA, 2023).

Essa tecnologia auxilia na criação de materiais didáticos contextualizados e culturalmente relevantes, como problemas matemáticos que dialogam com a realidade amazônica, simulando dados e adaptando trilhas de aprendizagem. Para os professores pode aliviar a carga de trabalho, auxiliando na elaboração de planos de aula, listas de exercícios e até na correção automática de avaliações, permitindo que se concentrem na orientação pedagógica. A integração de tecnologias como o ChatGPT e o Bing pode tornar o ensino de Matemática mais dinâmico, interativo e acessível, promovendo o engajamento dos alunos e o desenvolvimento do pensamento computacional e do raciocínio lógico (CAETANO; OLIVEIRA, 2024)

A sua implementação na educação matemática enfrenta desafios e limitações significativas. Uma preocupação primordial é a possibilidade de os alunos desenvolverem uma dependência excessiva da tecnologia, utilizando-a para copiar respostas sem um esforço genuíno de compreensão, o que pode levar a um aprendizado superficial e enfraquecer o pensamento crítico. E também pode gerar respostas imprecisas, incorretas ou descontextualizadas, conhecidas como "alucinações", exigindo supervisão humana e verificação crítica

das informações. Questões éticas, como plágio, privacidade de dados e vieses algorítmicos, são cruciais e necessitam de uma abordagem reflexiva (MEDEIROS *et al.*, 2024).

Um aspecto a ser considerado, é a formação inadequada dos educadores para utilizar essas tecnologias de forma pedagógica e crítica é uma barreira considerável, com muitos professores resistindo à integração devido à falta de conhecimento ou sobrecarga de trabalho. A falta de infraestrutura tecnológica e o abismo digital, especialmente em regiões remotas como a Amazônia, podem acirrar as desigualdades educacionais se a IA for implementada sem mediações adequadas. Há também a preocupação de que, possa distanciar a relação professor-aluno, substituindo o papel essencial do educador (BORBA; JUNIOR, 2023).

Nesse contexto a integração da IA no ensino de matemática pode ser comparada a um novo e potente recurso pedagógico, mas necessita de constante intermediações e uma compreensão holística dessas ferramentas.

4 CONTRIBUIÇÕES ACADÊMICAS

Durante minha formação acadêmica no curso de Licenciatura em Matemática, dediquei-me ao desenvolvimento de diversas pesquisas nas áreas de modelagem matemática, tecnologia e educação. Este tópico tem como objetivo apresentar uma visão das minhas principais contribuições nesse período. Entre 2022 e 2024, explorei uma série de temas relevantes, e, para os propósitos deste documento, destacarei e discutirei detalhadamente dois trabalhos específicos. Essa escolha, esta relacionada ao tema central do portfólio e a trajetória acadêmica da pesquisadora.

Quadro 1 – Principais Produções acadêmicas

| Título | Autores | Evento/Editora | Ano | Prática |
|--|--|---|------------|----------------|
| Tangram e o GeoGebra no Ensino de Conceitos de Frações e Geometria | Elizangela Silva David Soares Renato Germano | XIII EPAEM | 2022 | Publicação |
| EXPLORANDO A POTENCIAÇÃO ATRAVÉS DO TRIÂNGULO DE SIERPINSKI: uma abordagem para o 6 ^o ano do ensino fundamental | Elizangela Silva Renato Germano | III SAMATC | 2023 | Publicação |
| Um estudo da difusão tecnológica com uso de equações diferenciais ordinária aplicada á venda de smartphones | Elizangela Silva Renato Germano | 34 ^o Colóquio Brasileiro de Matemática | 2023 | Publicação |
| UTILIZAÇÃO DE VISÃO COMPUTACIONAL E DEEP LEARNING PARA SEPARAÇÃO DE GRÃOS: um protótipo de baixo custo baseado no Arduino. | Elizangela Silva Renato Germano | Editora Científica Digital | 2024 | Publicação |
| MODELAGEM MATEMÁTICA DA LEI DE ZIPF: uma análise comparativa entre textos em português e inglês | Elizangela Silva Renato Germano | IX EPAMM | 2024 | Publicação |

Fonte: Autora, 2025.

Os projetos **Utilização de visão computacional e deep learning para separação de grãos**: um protótipo de baixo custo baseado no Arduino. e **Modelagem matemática da lei de Zipf**: uma análise comparativa entre textos em português e inglês foram selecionados para oferecer uma visão holística do conhecimento matemático e demonstrar a interdisciplinaridade intrínseca à área. Ambos se alinham, direta e indiretamente, com os princípios da modelagem e dos modelos matemáticos. Na formação docente, a interligação de conteúdos de diferentes áreas do conhecimento é um aspecto de extrema relevância, pois expande a visão de mundo do futuro professor e enriquece a sua perspectiva sobre a transmissão do saber.

5 METODOLOGIA

5.1 TIPO DE PESQUISA

Este portfólio, de natureza básica, visa apresentar detalhadamente duas contribuições acadêmicas citadas da seção anterior, que foram desenvolvidas pela autora durante o curso de licenciatura em matemática no período de 2020 a 2025. Enquanto abordagem, se norteia por Soares (2023) e Chaves (2022) adotam o método qualitativo e memoria-lista, buscando uma compreensão profunda e interpretativa das produções e vivências acadêmicas.

Esses apontamentos, retratam o conceito de lugar epistemológico do pesquisador, que é a sua posição na produção de conhecimento, moldada por sua história de vida, formação e experiências, e que o leva a uma visão sistêmica e íntima com sua área de interesse.

5.2 INSTRUMENTOS METODOLÓGICOS

O primeiro instrumento de pesquisa foi levantamento da literatura, tendo como base teórica Bassanezi (2002), em um diálogo da ótica de Burak (1992) e Braga (2019) voltado para modelagem matemática na formação docente. Essa leitura se baseia em Gil (2002) é suporte para pesquisa bibliográfica e documental.

No eixo bibliográfico, realizou-se uma análise das principais obras dos autores mencionados, com o objetivo de extrair conceitos-chave, categorias de análise e perspectivas que permitissem interpretar e interligar os textos que serão selecionados na segunda fase da pesquisa.

No corpo documental, compilou as principais publicações da autora em eventos nacionais e editoriais. Os documentos foram organizados cronologicamente e por temática. Com base no inventário, procedeu-se trazer uma amostra representativa da trajetória e da diversidade de pesquisas desenvolvidas. O critério de seleção privilegiou: (i) modelagem matemática; (ii) modelos matemáticos e (iii) aplicações com inteligência artificial . Os trabalhos selecionados compõem o corpus principal de análise do portfólio e são apresentados em forma de apêndice.

5.3 ANÁLISE REFLEXIVA DAS PRODUÇÕES

Nessa fase, cada texto selecionado foi interpretado a luz das categorias definidas e das referências teóricas. Assim, na seção 6 e 7 apresentam um resumo sobre cada obra, seguido de uma narrativa que não apenas descreve um percurso, mas o explica e valoriza, a importância desses conhecimentos para a futura professora de matemática, que não apenas acumula experiências/vivências, mas é capaz de aprender com elas, articulando, de forma consciente e fundamentada, a teoria e a prática que constituem o seu próprio processo, enquanto professora-pesquisadora em formação.

6 CONSIDERAÇÕES SOBRE CAPÍTULO DE LIVRO: UTILIZAÇÃO DE VISÃO COMPUTACIONAL E DEEP LEARNING PARA SEPARAÇÃO DE GRÃOS: UM PROTÓTIPO DE BAIXO CUSTO BASEADO NO ARDUINO.

Esse capítulo de livro descreve o desenvolvimento de um sistema automatizado e de baixo custo para a separação mecânica de grãos de feijão preto e carioca. O objetivo central é empregar visão computacional e Deep Learning, especificamente redes neurais convolucionais (CNNs), para classificar os grãos, integrando essa inteligência com a programação em linguagem de blocos e uma placa Arduino para a execução física do processo.

Ao mesmo tempo, destaca a crescente relevância da automação agrícola para aumentar a eficiência e reduzir custos, especialmente em países em desenvolvimento, e propõe uma solução econômica para o tratamento pós-colheita de grãos. O protótipo é uma solução integrada que combina componentes mecânicos e eletrônicos. A metodologia adotada é aplicada, exploratória e experimental. A estrutura principal inclui uma esteira de distribuição ajustável (feita de papel cartão e vibrada por um motor de corrente contínua) para movimentar os grãos, uma área de filmagem (caixa transparente iluminada por LEDs) onde uma webcam captura as imagens dos grãos, e um sistema de controle baseado em um Arduino Uno e um computador.

Os mecanismos de separação consistem em um servo motor que move um anteparo para direcionar os grãos classificados para recipientes de coleta (copos plásticos). O computador executa os algoritmos de Deep Learning através de duas plataformas principais: o Teachable Machine, da Google, para o treinamento interativo do modelo de aprendizado de máquina, e o Pictoblox, uma plataforma de programação visual baseada em blocos, para a importação do modelo treinado e a criação do código de controle.

O modelo de classificação foi treinado no Teachable Machine com 2000 imagens de feijões pretos, 2000 de feijões cariocas e 2000 do dispositivo vazio, utilizando os hiperparâmetros padrão e validação cruzada para maximizar a precisão. O código em blocos no Pictoblox interpreta as classificações da webcam e ajusta o servo motor para ângulos específicos, garantindo a separação precisa de cada tipo de grão.

As redes neurais convolucionais (CNNs) são fundamentais para este processo, utilizando camadas de convolução para extrair características (bordas, texturas), camadas

de pooling para reduzir a dimensionalidade e camadas totalmente conectadas para a classificação final, aprendendo hierarquias espaciais de características de forma automática.

Os resultados demonstram a alta eficácia do sistema, com uma precisão global de classificação de 94,5%. Para o feijão carioca, a especificidade alcançou 100%, enquanto para o feijão preto foi de 89%. Essa diferença na precisão para o feijão preto é atribuída a variações maiores em sua forma e cor, além de possíveis erros no processo de separação mecânica, indicando que há espaço para melhorias.

Apesar dessas limitações, a alta taxa de precisão é considerada satisfatória, especialmente dado o baixo custo e a simplicidade dos componentes utilizados. O estudo sugere otimizações futuras, como o ajuste da iluminação e o refinamento dos hiperparâmetros da rede neural, para aprimorar ainda mais o desempenho e expandir a aplicabilidade para outros tipos de grãos.

No que tange às reflexões para a modelagem matemática, o trabalho exemplifica como modelos matemáticos complexos são implementados em aplicações práticas através do Deep Learning. A base da inteligência artificial, o Perceptron, é descrita como um modelo matemático que mimetiza neurônios biológicos, ajustando pesos (parâmetros matemáticos) para minimizar erros com base em uma função de ativação. Embora limitado em problemas não linearmente separáveis, sua evolução para redes neurais multicamadas e o algoritmo de retropropagação, e subsequentemente para CNNs, demonstra a progressão da modelagem matemática para resolver problemas mais complexos.

As CNNs, com suas camadas convolucionais e de pooling, representam uma sofisticada modelagem matemática de extração de características e redução de dimensionalidade, onde filtros são treinados para detectar padrões específicos e otimizar a precisão de classificação, ajustando os pesos dos filtros para minimizar uma função de perda.

A capacidade do Deep Learning de aprender representações hierárquicas de dados por meio de múltiplas camadas de neurônios artificiais é, em essência, uma construção de funções matemáticas complexas a partir de funções mais simples, permitindo ao computador construir conceitos complexos de forma robusta e adaptativa.

O estudo oferece para o ensino de modelagem, aplicações com inteligência artificial e visão computacional. A utilização de plataformas como Teachable Machine e Pictoblox é um diferencial pedagógico, pois tornam a criação e manipulação de modelos de aprendizado de máquina simples e intuitiva, mesmo para iniciantes. Essas ferramentas de livre acesso permitem que estudantes e educadores explorem conceitos fundamentais de

inteligência artificial e programação de forma interativa, sem a necessidade de aprofundado conhecimento em linguagens de programação tradicionais.

O projeto, sendo de baixo custo e com resultados concretos, serve como um recurso educacional valioso, incentivando a adoção de tecnologias avançadas em ambientes de aprendizagem e preparando os estudantes para entender e projetar soluções tecnológicas em diversos contextos, incluindo a automação agrícola. Ao transformar um problema real (separação de grãos) em um projeto prático que integra hardware (Arduino), software visual (Pictoblox) e IA (Teachable Machine/CNNs), o trabalho facilita a compreensão de como a modelagem matemática subjacente aos algoritmos de Deep Learning se traduz em funcionalidades do mundo real, estimulando o desenvolvimento de habilidades de resolução de problemas e promovendo um ensino inovador e interativo.

7 CONSIDERAÇÕES SOBRE O ARTIGO: MODELAGEM MATEMÁTICA DA LEI DE ZIPF: UMA ANÁLISE COMPARATIVA ENTRE TEXTOS EM PORTUGUÊS E INGLÊS

O presente artigo propõe-se a explorar a aplicação da Lei de Zipf em dois textos de gêneros distintos e suas respectivas traduções: a Bíblia Sagrada e o romance Dom Casmurro, ambos nas versões em português e inglês. A Lei de Zipf descreve uma relação matemática inversa entre a frequência de ocorrência das palavras em um texto e sua posição no ranking de frequência, sugerindo que os padrões de uso das palavras seguem princípios universais.

O estudo utiliza a Bíblia Sagrada, conhecida por seu vocabulário religioso específico e repetitivo, e Dom Casmurro, que apresenta maior variedade lexical e nuances estilísticas, para investigar como a Lei de Zipf se manifesta em diferentes contextos linguísticos e culturais. O objetivo principal é identificar as variações nas frequências das palavras entre os idiomas e os gêneros textuais, utilizando gráficos em escala log-log para calcular os coeficientes de Zipf e, assim, proporcionar uma visão quantitativa da distribuição de palavras e sua aderência à lei.

A metodologia empregada envolveu a seleção de versões digitais completas dos textos em português e inglês. Os textos foram submetidos a um pré-processamento digital, que incluiu a remoção de pontuações, números e caracteres especiais, além da conversão de todas as palavras para minúsculas para garantir consistência na contagem de frequência. Em seguida, foi realizada a contagem da frequência de ocorrência de cada palavra distinta, organizando-as em ordem decrescente de frequência para formar um ranking.

Para visualizar a distribuição e verificar a aderência à Lei de Zipf, foram gerados gráficos em escala log-log, plotando o logaritmo da posição da palavra no ranking contra o logaritmo de sua frequência. Finalmente, a regressão linear nesses gráficos permitiu determinar o coeficiente de Zipf para cada conjunto de dados, sendo a inclinação da reta ajustada uma medida da distribuição de frequências das palavras, permitindo a comparação entre os textos e idiomas.

Os resultados revelaram que a Bíblia Sagrada adere mais rigidamente à Lei de Zipf em ambas as línguas, com coeficientes próximos do valor clássico de -1. Na versão em português, o coeficiente foi de -1,13, e na versão em inglês, -1,33. Essa conformidade mais

rigorosa é atribuída à natureza formal e repetitiva do texto bíblico, com alta frequência de termos religiosos e expressões fixas.

Em contraste, Dom Casmurro apresentou coeficientes ligeiramente menores, refletindo sua maior variação lexical e o estilo literário mais diversificado de Machado de Assis. O coeficiente para a versão em português de Dom Casmurro foi de aproximadamente -0,83, enquanto a versão em inglês ficou mais próxima do valor clássico, com cerca de -0,93. As traduções para o inglês, embora tenham mantido os padrões gerais de distribuição de palavras, introduziram pequenas variações nos coeficientes, especialmente na Bíblia Sagrada, evidenciando o impacto das nuances linguísticas e culturais no processo de tradução e na conformidade com a lei.

No âmbito da modelagem matemática, o estudo exemplifica a Lei de Zipf como um modelo preditivo para a distribuição de palavras em línguas naturais. A aplicação de gráficos log-log e regressão linear para quantificar o coeficiente de Zipf demonstra a potência da modelagem para analisar e comparar fenômenos complexos como a estrutura linguística. O coeficiente de Zipf, como uma medida da inclinação da curva, atua como um parâmetro crucial do modelo, cujas variações permitem inferir sobre as características intrínsecas do texto; estilo, gênero e o impacto da tradução.

A observação de que o modelo, embora universalmente aplicável, mostra desvios quantificáveis do valor teórico (-1) em diferentes contextos (textos religiosos versus literários, e suas traduções) sublinha a importância de considerar o contexto na interpretação dos modelos matemáticos e a capacidade destes em quantificar e contextualizar essas variações.

Um contexto prático de como conceitos matemáticos se aplicam a linguística. Assim, o estudo fomenta a compreensão da flexibilidade dos modelos matemáticos para descrever uma variedade de fenômenos, estimulando a curiosidade e o pensamento crítico sobre as relações entre matemática e outras áreas do saber.

8 CONSIDERAÇÕES FINAIS

O valor deste portfólio destaca a influência da modelagem matemática e tecnologias na minha formação acadêmica. Para além do processo formativo em licenciatura em matemática, consolida-se como elemento fundamental na constituição de professora-pesquisadora, uma vez que possibilita a articulação entre teoria e prática.

Esse posicionamento molda o pensamento crítico e criativo, e talvez o mais importante, impacta na formação docente e cidadã. Uma vez que, a educação matemática pautada na modelagem e na pesquisa prepara o indivíduo para compreender e intervir na sociedade de maneira informada e responsável.

Para além disso, a vivência na iniciação científica instrumentaliza o futuro professor de matemática para implementação de abordagens interdisciplinares em sala de aula. Ao articular conhecimentos de matemática com áreas como linguística, computação e inteligência artificial, o educador pode deixar o aprendizado mais significativo e cultivar uma visão sistêmica e crítica.

REFERÊNCIAS

- BASSANEZI, R. C. **Ensino-aprendizagem com modelagem matemática**. 4^a. ed. São Paulo: Contexto, 2002. ISBN 978-8572442077. Citado 2 vezes nas páginas 13 e 19.
- BORBA, M. d. C.; JUNIOR, V. R. B. O chatgpt e educação matemática. **Educação Matemática Pesquisa Revista do Programa de Estudos Pós-Graduados em Educação Matemática**, v. 25, n. 3, p. 142–156, out. 2023. Disponível em: <<https://revistas.pucsp.br/index.php/emp/article/view/63304>>. Citado na página 16.
- BRAGA, R. M. O laboratório experimental de modelagem matemática na formação inicial. In: **Anais do Seminário de Projetos Educacionais da UFPA**. [s.n.], 2019. v. 3, p. 380–383. ISSN 2764-1880. Disponível em: <<https://drive.google.com/file/d/1c5K-1XhOBRxJO0-GxxCEWpYsNKT0NdiD/view>>. Citado 2 vezes nas páginas 14 e 19.
- BRAGA, R. M. Práticas no laboratório experimental de modelagem matemática na formação docente. In: **Anais do Seminário de Projetos Educacionais da UFPA**. [s.n.], 2021. v. 4, p. 858–870. ISSN 2764-1880. Disponível em: <<https://drive.google.com/file/d/1sAxjhuMlnnCd7lxmfxEJN2xBOXPFxg4K/view>>. Citado na página 14.
- BURAK, D. **Modelagem Matemática: ações e interações no processo de ensino-aprendizagem**. 460 f. Tese (Doutorado em Educação) — Universidade Estadual de Campinas - Faculdade de Educação, Campinas, SP, Brasil, 1992. Disponível em: <https://www.psiem.fe.unicamp.br/pf-psiem/burak_dionisio_d.pdf>. Citado 2 vezes nas páginas 13 e 19.
- BURAK, D.; ZONTINI, L. d. R. S. Práticas com modelagem na formação do professor da educação básica: a busca por uma nova racionalidade. **Práxis Educativa**, v. 15, p. 1–20, fev. 2020. Disponível em: <<https://revistas.uepg.br/index.php/praxiseducativa/article/view/14239>>. Citado na página 14.
- CAETANO, J. J.; OLIVEIRA, P. Gonçalves de. O uso do chatgpt no ensino da matemática. **Com a Palavra, o Professor**, v. 9, n. 25, p. 150–169, dez. 2024. Disponível em: <<http://revista.geem.mat.br/index.php/PPP/article/view/1076>>. Citado na página 15.
- CHAVES, A. da C. **Do quilombo à universidade: As Escrevivências de formação de uma educadora matemática nos fazeres decoloniais da extensão Universitária**. 63 f. Monografia (Trabalho de Conclusão de Curso - Licenciatura em Matemática) — Universidade Federal do Pará, Castanhal, PA, Brasil, 2022. Disponível em: <<https://bdm.ufpa.br/server/api/core/bitstreams/fc131a35-e8c2-409e-bbcf-4d0442ba48de/content>>. Acesso em: 25 ago. 2025. Citado na página 19.
- COSTA, S. P. d. *et al.* O uso de ferramentas de inteligência artificial generativa no desenvolvimento de materiais didáticos interativos para o ensino de matemática na amazônia: um ensaio teórico. **Caderno Pedagógico**, v. 22, n. 9, p. e17992, jul. 2025. Disponível em: <<https://ojs.studiespublicacoes.com.br/ojs/index.php/cadped/article/view/17992>>. Citado na página 15.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4^a ed.. ed. São Paulo: Atlas, 2002. ISBN 85-224-3169-8. Citado na página 19.

MEDEIROS, T. K. S. d. *et al.* A utilização da inteligência artificial no ensino de matemática. **Cognitionis Scientific Journal**, v. 7, n. 2, p. e490, set. 2024. Disponível em: <<https://revista.cognitioniss.org/index.php/cogn/article/view/490>>. Citado na página 16.

OLIVEIRA, R. M.; SILVA, R. d. S. O uso da inteligência artificial no ensino da matemática. **Caderno Intersaberes**, v. 12, n. 44, 2023. Disponível em: <<https://www.cadernosuninter.com/index.php/intersaberes/article/view/2964>>. Citado na página 15.

SOARES, D. G. **Laboratório Experimental de Modelagem Matemática: Práticas e Formação Docente**. 51 f. Monografia (Trabalho de Conclusão de Curso - Licenciatura em Matemática) — Universidade Federal do Pará, Castanhal, PA, Brasil, 2023. Disponível em: <<https://bdm.ufpa.br/server/api/core/bitstreams/7539795b-4f20-4d6a-861a-6c615c9c1d27/content>>. Acesso em: 25 ago. 2025. Citado 2 vezes nas páginas 14 e 19.

Universidade Federal do Pará. **Regulamento de Trabalho de Conclusão de Curso Resolução nº 04/2024**. Campus Universitário de Castanhal: Faculdade de Matemática, 2024. Disponível em: <<https://facmatcastanhal.ufpa.br>>. Citado na página 12.

Apêndices

**APÊNDICE A – UTILIZAÇÃO DE VISÃO COMPUTACIONAL E DEEP
LEARNING PARA SEPARAÇÃO DE GRÃOS: UM PROTÓTIPO DE BAIXO
CUSTO BASEADO NO ARDUINO**

06

UTILIZAÇÃO DE VISÃO COMPUTACIONAL E *DEEP LEARNING* PARA SEPARAÇÃO DE GRÃOS: UM PROTÓTIPO DE BAIXO CUSTO BASEADO NO ARDUINO

Elizângela Maria Gonçalves Silva
Universidade Federal do Pará (UFPA)

Renato Germano
Universidade Federal do Pará (UFPA)



RESUMO

Objetivo: Apresentamos o desenvolvimento de um sistema automatizado de baixo custo para a separação mecânica de dois tipos de grãos de feijão, utilizando visão computacional (*Deep Learning*) e a placa Arduino. Utilizando redes neurais convolucionais para a classificação dos feijões, este sistema integra-se com programação em linguagem de blocos, demonstrando sua eficácia com uma taxa de acurácia de 94,5% na classificação dos feijões. Esta pesquisa contribui para a automação na agricultura ao oferecer uma solução econômica e destacar o valor educacional de tais tecnologias. Ao empregar plataformas de livre acesso, sem custos, que permitem o treinamento interativo de modelos de aprendizado de máquina, este projeto avança não apenas a automação agrícola, mas também serve como um recurso educacional valioso, fomentando a adoção de tecnologias avançadas em ambientes de aprendizagem.

Palavras-chave: visão computacional; redes neurais convolucionais; arduino; *teachable machine*; pictoblox; *deep learning*; *machine learning*.

INTRODUÇÃO

A automação agrícola, cada vez mais relevante, reflete a necessidade contínua de aumentar a eficiência e reduzir custos no setor. Diante deste panorama, a separação automatizada de grãos, como os feijões preto e carioca, destaca-se como uma aplicação promissora. Este trabalho desenvolve um sistema que integra visão computacional e redes neurais convolucionais (CNNs), utilizando um Arduino para a execução física do processo, visando a classificação e separação eficiente destes grãos.

A inserção da automação em contextos agrícolas nos países em desenvolvimento, onde a agricultura desempenha um papel crucial para a economia, tem experimentado um crescimento substancial. Este estudo concentra-se no desenvolvimento de um protótipo de baixo custo que emprega técnicas avançadas de visão computacional e aprendizado profundo (*Deep Learning*), demonstrando a viabilidade de soluções tecnológicas acessíveis para a automação do tratamento pós-colheita de grãos. Além de garantir a precisão na identificação e classificação dos grãos, estas tecnologias têm um impacto significativo na qualidade do produto final e na eficiência do processamento.

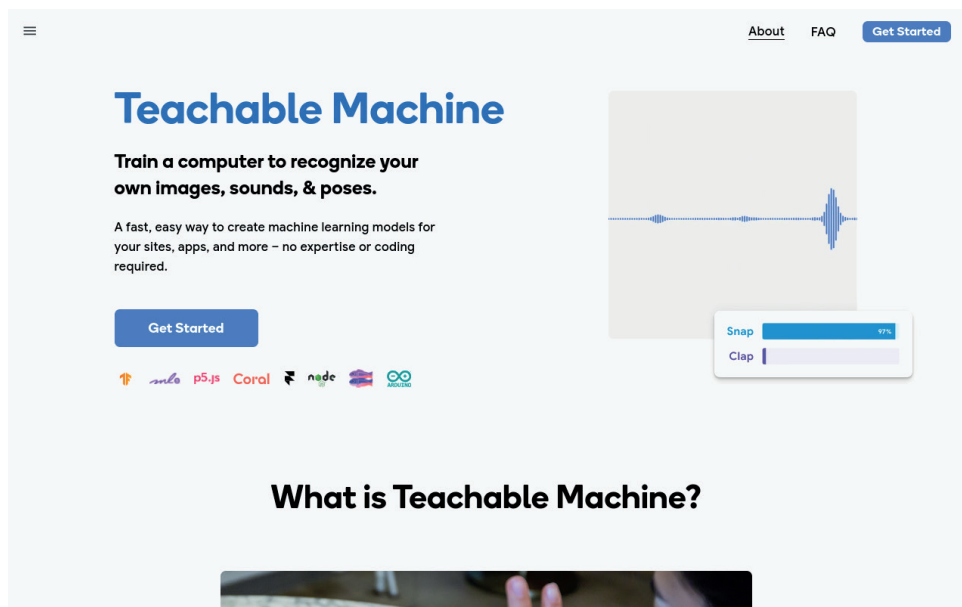
A importância deste estudo estende-se também ao âmbito educacional, onde as tecnologias envolvidas podem ser exploradas como ferramentas didáticas. O uso de plataformas como o Teachable Machine e o Pictoblox, que permitem a criação e manipulação de modelos de aprendizado de máquina de forma interativa, oferece um potencial educativo significativo. Essas plataformas podem ser utilizadas para ensinar conceitos fundamentais de inteligência artificial e programação a estudantes, preparando-os para entender e projetar soluções tecnológicas em uma variedade de contextos, incluindo a automação agrícola.

Por fim, a capacidade de integrar tecnologias de baixo custo em soluções automatizadas não apenas promove a inovação acessível, mas também está alinhada com as diretrizes de desenvolvimento rural sustentável e segurança alimentar. Este trabalho, portanto, não apenas contribui para a automação na agricultura, mas também serve como um recurso educacional valioso, incentivando a adoção de tecnologias avançadas em ambientes de aprendizagem.

Teachable Machine

O Teachable Machine é uma plataforma interativa desenvolvida pela Google, projetada para permitir que usuários de todos os níveis de conhecimento criem modelos de aprendizado de máquina de forma simples e intuitiva (Google, 2024). Esta plataforma oferece uma flexibilidade significativa na criação de modelos de aprendizado de máquina, permitindo o uso de diferentes tipos de dados para o treinamento, tais como: imagens de arquivos pré-existentes ou capturando novas imagens através de uma *webcam*; sons, como diferentes tipos de áudio e poses, como posições corporais, utilizando arquivos ou capturando poses ao vivo via *webcam*.

Figura 1 - imagem da página inicial da plataforma Teachable Machine.



Fonte: <https://teachablemachine.withgoogle.com/>.

Uma característica importante do Teachable Machine é a ênfase na privacidade do usuário. A plataforma pode ser utilizada inteiramente no dispositivo local, garantindo que os dados da *webcam* ou do microfone não sejam transmitidos para fora do computador do usuário. Isso é especialmente relevante para projetos que lidam com dados sensíveis ou que necessitam de um alto nível de privacidade.

Diversas aplicações práticas podem ser desenvolvidas utilizando o Teachable Machine, demonstrando sua versatilidade, tipo projetos conectando o Teachable Machine com Arduino. Reconhecer e classificar objetos rapidamente utilizando

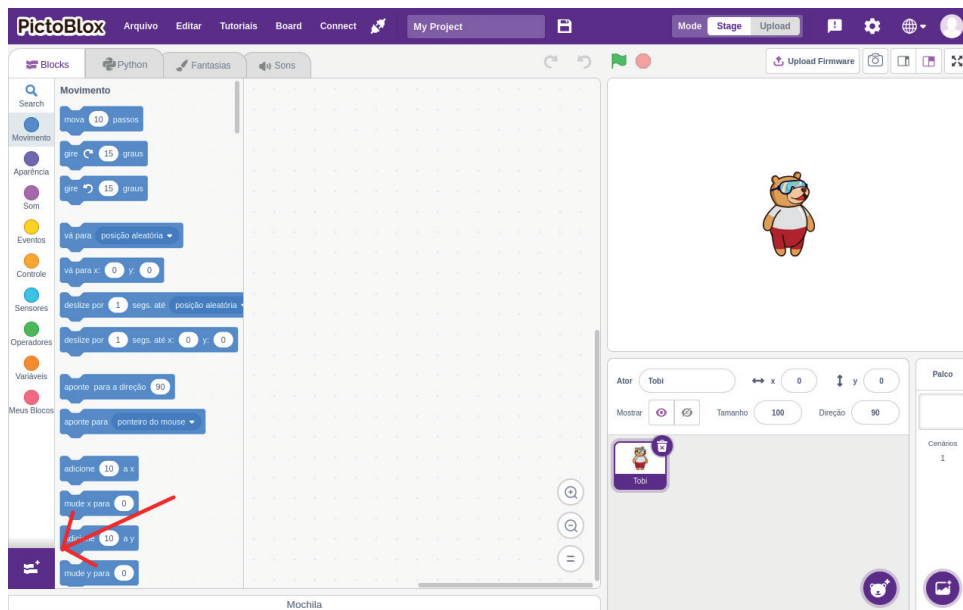
modelos de aprendizado de máquina personalizados ou facilitar a comunicação por meio de gestos faciais, acionando sons específicos, bem como o uso de uma *webcam* e um pedaço de papel para controlar jogos digitais, destacando a flexibilidade da plataforma em contextos lúdicos (Google, 2024).

Além das aplicações práticas, o Teachable Machine possui um grande potencial educativo, sendo utilizado em várias lições e atividades para ensinar conceitos fundamentais de aprendizado de máquina, possibilitando o desenvolvimento de soluções eficientes e de baixo custo.

Pictoblox

Pictoblox é uma plataforma de programação visual baseada em blocos, desenvolvida para tornar a codificação acessível e divertida para estudantes, educadores e entusiastas da tecnologia (Pictoblox, 2024). Ela oferece uma interface intuitiva que permite aos usuários programar sem a necessidade de conhecimentos profundos em linguagens de programação tradicionais. Suas principais características incluem: programação em blocos, isto é, uma abordagem de arrastar e soltar para criar programas, tornando a codificação acessível para iniciantes (Marji, 2014; Sweigart, 2016); suporte multimídia, o que permite criar projetos interativos e educativos; extensões personalizáveis, como controle de hardware, inteligência artificial e visão computacional e multiplataforma, o que facilita o acesso e uso em diferentes dispositivos.

Figura 2 - imagem da páginas inicial da plataforma Pictoblox.



Fonte: <https://pictoblox.ai/>.

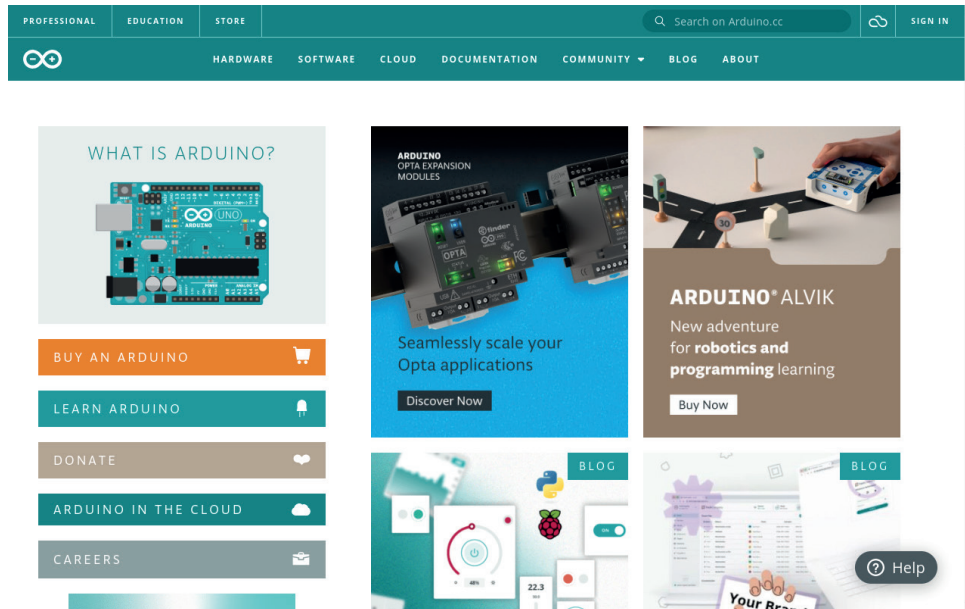
O Pictoblox é utilizado em ambientes educacionais para ensinar conceitos de ciência, tecnologia, engenharia e matemática de maneira prática e envolvente. Suporta a criação de projetos em que os usuários podem programar e controlar dispositivos eletrônicos como Arduino e placas de desenvolvimento similares. Possui integração com bibliotecas de visão computacional, inclusive com a plataforma Teachable Machine, que permite a criação de projetos que envolvem reconhecimento de objetos, classificação de imagens e outras aplicações de inteligência artificial.

Arduino

Arduino é uma plataforma de código aberto utilizada para a criação de projetos eletrônicos interativos (Arduino, 2024). O Arduino oferece uma variedade de características que o tornam uma escolha popular para desenvolvedores de projetos eletrônicos de baixo custo, a qual inclui uma série de placas de desenvolvimento, como Arduino Uno, Mega, Nano, entre outras, cada uma adequada para diferentes tipos de projetos. Possui um Ambiente de Desenvolvimento Integrado (IDE), que é uma ferramenta de software gratuita que facilita a programação das

placas Arduino. Ele suporta a linguagem de programação C/C++ e inclui diversas bibliotecas para simplificar o desenvolvimento.

Figura 3 - imagem da páginas inicial da plataforma do Arduino.



Fonte: <https://www.arduino.cc/>.

Arduino é utilizado em uma ampla gama de projetos, desde simples experimentos educacionais até complexos sistemas de automação. É utilizado também no ensino de conceitos de ciência, tecnologia, engenharia e matemática de forma prática. Em projetos de automação residencial, como controle de iluminação, sistemas de segurança e automação de eletrodomésticos. Em desenvolvimento de robôs para diversas finalidades, incluindo robótica educacional, competições e projetos pessoais e a chamada Internet das Coisas (IoT).

REDES NEURAIS ARTIFICIAIS

A primeira rede neural artificial surgiu em 1958 com a criação do Perceptron pelo psicólogo Frank Rosenblatt (Rosenblatt, 1962) e sua primeira implementação prática foi inspirada nos trabalhos teóricos de McCulloch e Pitts (1943), os quais desenvolveram um modelo matemático de neurônios artificiais. O Perceptron foi projetado para reconhecer padrões e aprender a partir de exemplos, mimetizando o funcionamento dos neurônios biológicos. Ele consistia em uma única camada de

neurônios com pesos ajustáveis e uma função de ativação que determinava se um neurônio deveria “disparar” ou não, baseado na soma ponderada das entradas.

A Figura 4 representa a estrutura básica de um Perceptron, as quais são:

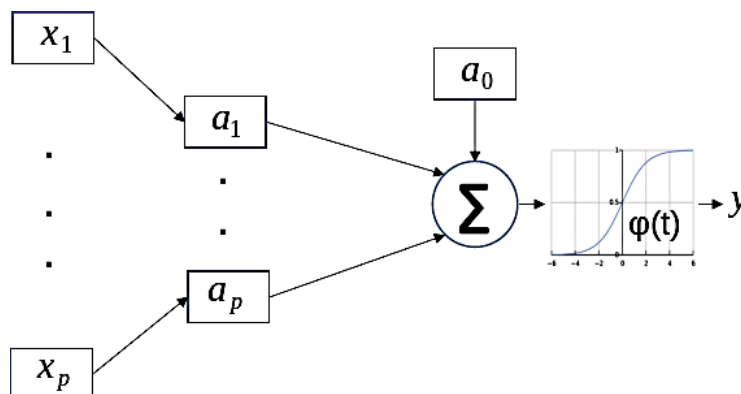
1. Valores de entrada (x_1, x_2, \dots, x_p): cada x_i representa uma característica ou um atributo do conjunto de dados de entrada.
2. Pesos (a_1, a_2, \dots, a_p): cada entrada x_i está associada a um peso a_i , esses pesos são ajustados durante o processo de treinamento.
3. *Bias* (a_0): termo adicional que é somado à combinação linear das entradas ponderadas. Esse termo é utilizado para ajustar a função de ativação, permitindo modelar dados que não passam pela origem.
4. Somatório (Σ): representa a combinação linear das entradas com seus respectivos pesos: $\Sigma = a_0 + a_1x_1 + a_2x_2 + \dots + a_px_p$.
5. Função de Ativação $\varphi(t)$: função que aplica uma transformação ao valor do somatório. Comumente, a função de ativação é uma função não-linear como a função degrau, sigmoide ou ReLU (*Rectified Linear Unit*).
6. Saída (y): esta saída é usada para tomar decisões ou fazer previsões.

Assim, a estrutura básica do Perceptron pode ser resumida na fórmula

$$y = \varphi(a_0 + a_1x_1 + a_2x_2 + \dots + a_px_p),$$

em que φ é a função de ativação e y é a saída.

Figura 4 - estrutura básica de um Perceptron.



Fonte: autores, 2024.

O Perceptron funciona ajustando os pesos das conexões de entrada de forma que a saída do neurônio corresponda ao valor desejado. O algoritmo de aprendizado ajusta os pesos incrementando-os ou decrementando-os com base no erro entre a saída prevista e a saída desejada. Este processo é repetido iterativamente até que os pesos convergissem para valores que minimizem o erro.

Embora o Perceptron tenha sido um avanço significativo, ele tinha algumas limitações. Foi demonstrado que Perceptrons de camada única não podiam resolver problemas não linearmente separáveis, como a função XOR (Minsky e Papert, 1969). Esta limitação foi superada com o desenvolvimento de redes neurais multicamadas (*multilayer perceptrons*) e a introdução do algoritmo de retropropagação (*backpropagation*). Estas redes neurais multicamadas permitiram o aprendizado de funções não lineares complexas e contribuíram para o avanço das redes neurais profundas e do *Deep Learning* (Zollanvari, 2023).

Arquiteturas de Redes Neurais

As arquiteturas de redes neurais incluem diversas variantes, cada uma projetada para tipos específicos de problemas, especialmente em problemas onde os dados são abundantes e complexos:

- *Multilayer Perceptrons*: redes neurais com várias camadas de neurônios, utilizadas para aproximar funções complexas.
- Redes Neurais Convolucionais (CNNs): inspiradas pela estrutura visual dos animais, são usadas para reconhecimento de imagem e vídeo, sistemas de recomendação e processamento de linguagem natural.
- Redes Neurais Recorrentes (RNNs): projetadas para lidar com dados sequenciais, como séries temporais e texto, considerando dependências temporais entre os dados.
- Máquinas de Boltzmann Restritas (RBMs): redes neurais não supervisionadas que utilizam minimização de energia e são úteis para criar modelos generativos.

- Redes de Funções de Base Radial (RBF): redes que utilizam funções de base radial como neurônios e são eficientes em classificação de padrões não lineares.

Deep Learning

Deep Learning é um subcampo do aprendizado de máquina (*Machine Learning*) que utiliza redes neurais artificiais para modelar padrões complexos em grandes quantidades de dados. Conforme destacado por Petrelli (2023), *Deep Learning* resolve o problema central no aprendizado de representações ao introduzir representações expressas em termos de outras, mais simples, permitindo ao computador construir conceitos complexos a partir de conceitos mais simples.

Deep Learning se distingue dos métodos tradicionais de *Machine Learning* pela sua capacidade de aprender representações hierárquicas de dados. Isso é realizado através de redes neurais profundas, que consistem em múltiplas camadas de neurônios artificiais. Cada camada em uma rede profunda captura diferentes níveis de abstração dos dados de entrada, desde características básicas até conceitos de alto nível. Goodfellow *et al.* (2016) descrevem que redes neurais profundas, como os *multilayer perceptrons*, são formadas por funções matemáticas que combinam muitas funções mais simples.

Classificação e Detecção de Objetos

As Redes Neurais Convolucionais (CNNs) são um tipo específico de rede neural projetada para processar dados que possuem uma grade topológica, como imagens. Inspiradas pela organização do córtex visual de animais, as CNNs são utilizadas em tarefas de visão computacional, como classificação de imagens, segmentação de imagens, e detecção de objetos. De acordo com Fukushima (1980), CNNs são construídas para “aprender hierarquias espaciais de características de maneira automática e adaptativa através da retropropagação”.

A capacidade das CNNs de aprender hierarquias de características e adaptar-se automaticamente a diferentes tipos de dados as torna uma ferramenta indispensável em muitas aplicações, desde o reconhecimento de imagem até a análise de dados geoespaciais. Um exemplo clássico é a identificação de objetos em imagens, como cães, gatos, ou carros. Durante o treinamento a CNN ajusta

os pesos dos filtros convolucionais para maximizar a precisão de classificação, minimizando a função de perda.

Uma CNN típica é composta por várias camadas, cada uma desempenhando um papel crucial no processamento e interpretação dos dados de entrada. As principais camadas de uma CNN incluem.

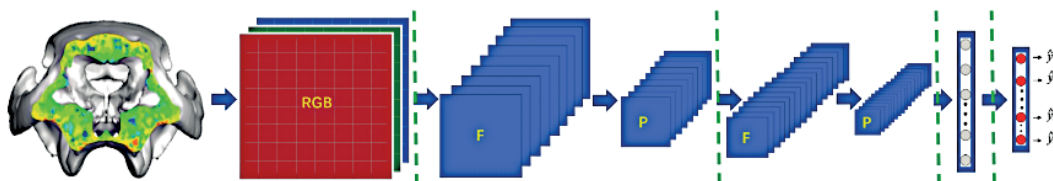
- Camadas de Convolução: essas camadas aplicam filtros convolucionais aos dados de entrada para extrair características locais. Cada filtro é treinado para detectar um padrão específico, como bordas, texturas, ou formas após cada convolução aplicada, conforme ilustrado por Zeiler e Fergus (2014).
- Camadas de *Pooling*: também conhecidas como camadas de amostragem, elas reduzem a dimensionalidade dos dados, agregando valores dentro de uma janela de operação, como o máximo ou a média. Isso ajuda a diminuir o número de parâmetros e a complexidade computacional e o risco de *overfitting*.
- Camadas Totalmente Conectadas: Estas camadas conectam todos os neurônios de uma camada à próxima. Elas são responsáveis por combinar as características extraídas em camadas anteriores para produzir a saída final, seja uma classificação ou uma localização de objetos.

A Figura 5 ilustra os passos de uma CNN utilizada para processamento de imagens (LIU, 2023):

1. Imagem Inicial (RGB): a imagem que é convertida em uma representação RGB.
2. Primeiro Conjunto de Filtros (F): a imagem RGB é passada através do primeiro conjunto de filtros, que gera várias imagens filtradas (também conhecidas como mapas de características).
3. Primeira Operação de *Pooling* (P): as imagens filtradas são então submetidas a uma operação de *pooling*, reduzindo sua resolução enquanto mantém as informações mais importantes.

4. Segundo Conjunto de Filtros (F): as imagens reduzidas pelo *pooling* são novamente passadas através de outro conjunto de filtros, gerando novos mapas de características.
5. Segunda Operação de *Pooling* (P): mais uma vez, as imagens resultantes são submetidas a *pooling* para reduzir ainda mais a dimensionalidade.
6. Camadas Densas: finalmente, as representações de características extraídas são alimentadas em uma sequência de camadas densas para a combinação e interpretação final.
7. Previsão (*Output*): previsão final gerada pela rede, que pode representar a classificação ou identificação desejada com base na imagem de entrada.

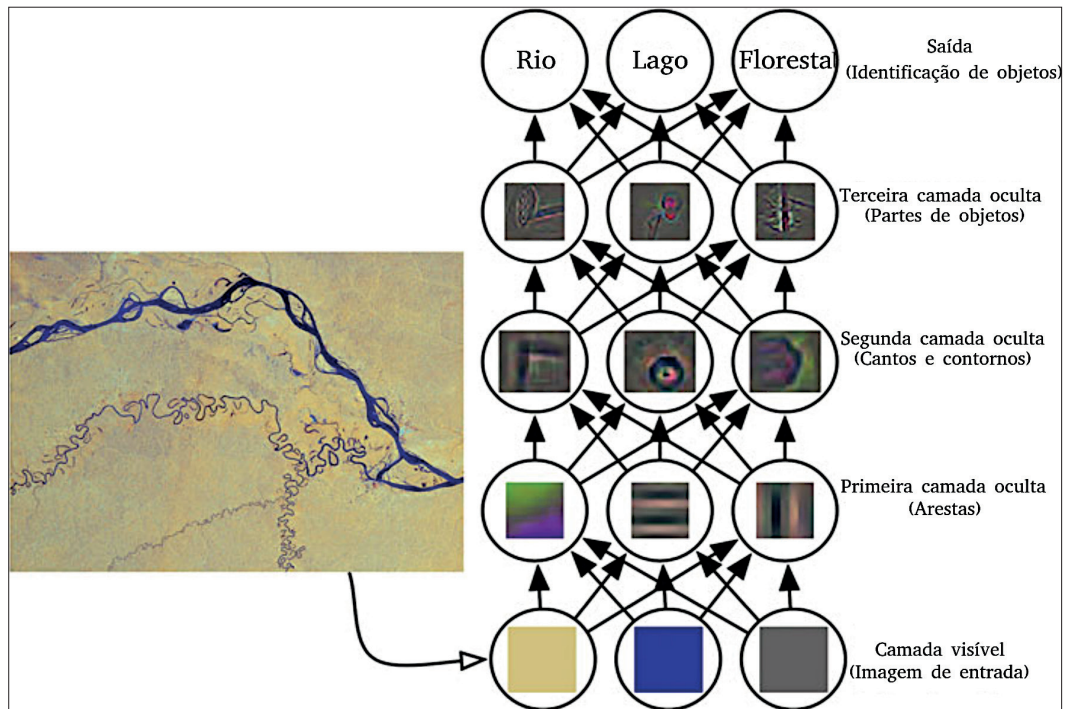
Figura 5 - passos de uma CNN utilizada para processamento de imagens.



Fonte: figura adaptada de Liu (2023).

Numa descrição mais específica de uma CNN com três camadas ocultas, de acordo com Goodfellow *et al.* (2016), temos que a entrada do modelo consiste em *pixels* de uma imagem. A primeira camada oculta detecta bordas na imagem. Esta camada é responsável por identificar as transições de intensidade de pixel que correspondem a contornos e formas básicas. A segunda camada oculta detecta cantos e contornos. Esta camada combina as bordas detectadas pela primeira camada para identificar formas e padrões mais complexos. A terceira camada oculta detecta partes de objetos. Esta camada identifica componentes específicos de objetos maiores. A camada de saída classifica a imagem em diferentes categorias de objetos. Esta camada usa as características extraídas pelas camadas ocultas para fazer a classificação final.

Figura 6 - passos de uma CNN utilizada para processamento de imagens.



Fonte: figura adaptada de Goodfellow *et al.* (2016).

A Figura 6 ilustra estes passos, na prática. Os quais serão adotados nesse trabalho:

1. A imagem de entrada é uma captura do rio Amazonas obtida via satélite.
2. As bordas da imagem são detectadas pela primeira camada oculta, que destaca transições abruptas nos valores de *pixel*.
3. A segunda camada oculta analisa as bordas detectadas e identifica cantos e contornos, formando uma representação mais detalhada das características da imagem.
4. Na terceira camada, partes específicas dos objetos são identificadas, permitindo uma compreensão mais granular dos componentes presentes na imagem.
5. Finalmente, a camada de saída usa as informações das camadas anteriores para classificar a imagem em categorias específicas, como rio, lago e floresta.

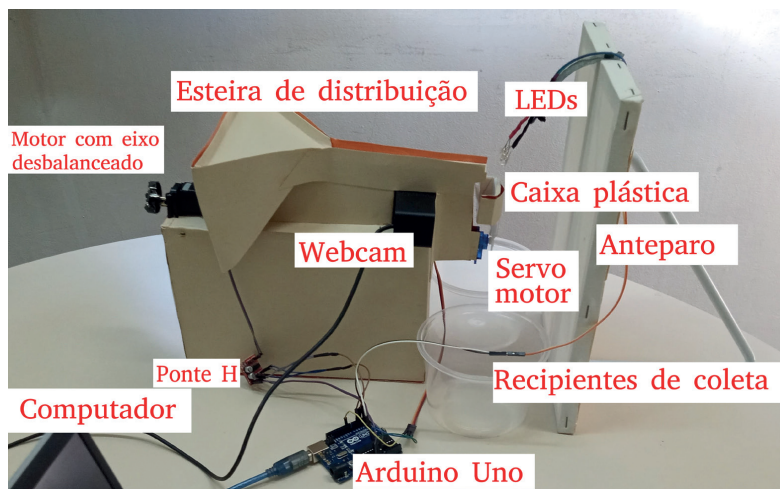
METODOLOGIA

Esta pesquisa é de natureza aplicada, exploratória, cujo procedimento é experimental, de acordo com Prodanov e Freitas (2013). O protótipo desenvolvido consiste em um sistema integrado que utiliza um microcontrolador Arduino, uma *webcam* e um computador para implementar algoritmos de *Deep Learning*. A escolha do Arduino foi motivada por sua acessibilidade, facilidade de uso e a vasta comunidade de suporte, que oferece diversas bibliotecas e tutoriais que facilitam a implementação de projetos complexos.

Aparato Experimental

O protótipo desenvolvido para a separação de grãos utiliza uma combinação de componentes mecânicos e eletrônicos para automatizar o processo de classificação de dois tipos de grãos (ver Figura 7).

Figura 7 - aparato experimental.



Fonte: autores, 2024.

Os componentes da estrutura principal é descrita no Quadro 1.

Quadro 1 - descrição dos principais componentes do aparato experimental.

| Esteira de Distribuição Ajustável |
|--|
| Material: Papel cartão, utilizado devido à sua flexibilidade e facilidade de manipulação, montado com cola. |
| Funcionalidade: A esteira serve para distribuir os grãos uniformemente. Um motor de corrente contínua, equipado com um eixo desbalanceado, provoca vibrações na esteira, fazendo com que os grãos deslizem até a área de filmagem. |

| Área de Filmagem |
|--|
| Estrutura: Caixa transparente feita de uma embalagem de balas tic tac, escolhida por sua clareza e tamanho adequado para conter os grãos durante a captura de imagens. |
| Iluminação: LEDs são usados para iluminar a caixa transparente, garantindo que as imagens capturadas pela webcam sejam claras e bem iluminadas. |
| Webcam: Captura imagens dos grãos para análise. Está posicionada de modo a ter uma visão direta e clara dos grãos que passam pela caixa transparente. |
| Sistema de Controle |
| Arduino Uno: Utilizado para controlar o acionamento dos mecanismos de separação, incluindo o servo motor que direciona os grãos para os recipientes de coleta após a classificação. |
| Computador: Executa os algoritmos de deep learning através do Pictoblox e do site Teachable Machine para identificar e classificar os grãos com base nas imagens capturadas pela webcam. |
| Mecanismos de Separação |
| Servo Motor: Responsável por mover um anteparo que direciona os grãos classificados para os respectivos recipientes de coleta. |
| Anteparo: Facilita a identificação visual clara das imagens ao minimizar a poluição visual ao redor dos grãos. |
| Receptáculos de Coleta |
| Dois Copos Plásticos: Coletam os grãos após serem classificados e separados, dividindo-os conforme o tipo identificado pelo sistema. |

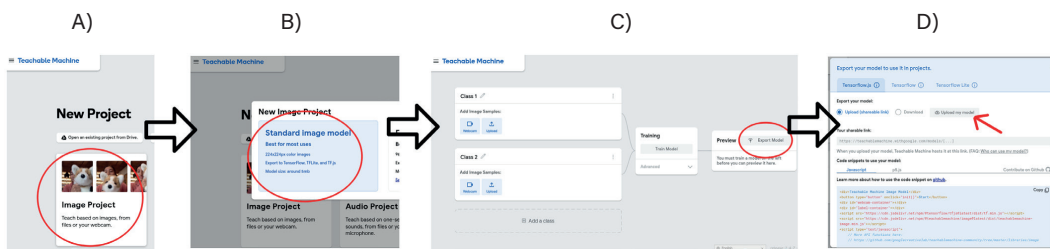
Fonte: autores, 2024.

Procedimento experimental

A seguir descreveremos o processo de treinamento do algoritmo no portal Teachable Machine.

A Figura 8 mostra os passos iniciais para criar um novo projeto de treinamento de modelo de redes neurais no portal Teachable Machine: A) Início do Projeto, a interface inicial mostra uma opção para iniciar um novo projeto. O usuário pode escolher entre diferentes tipos de projetos, com destaque para o projeto de imagens, indicado pelo círculo vermelho. B) Seleção do Tipo de Projeto, após clicar em New Project, aparecem várias opções de projetos. Para este exemplo, a escolha é um projeto de imagem, com opções adicionais disponíveis para seleção, tais como Standard Image Model. C) Configuração Inicial, neste passo, o usuário define as classes que compõem o modelo e o usuário tem a opção de adicionar mais classes se necessário. O usuário pode carregar imagens para cada classe criada, clicando no botão Upload, conforme indicado pelo círculo vermelho. As imagens carregadas serão usadas para treinar o modelo de rede neural. D) Indica a exportação do link do modelo treinado.

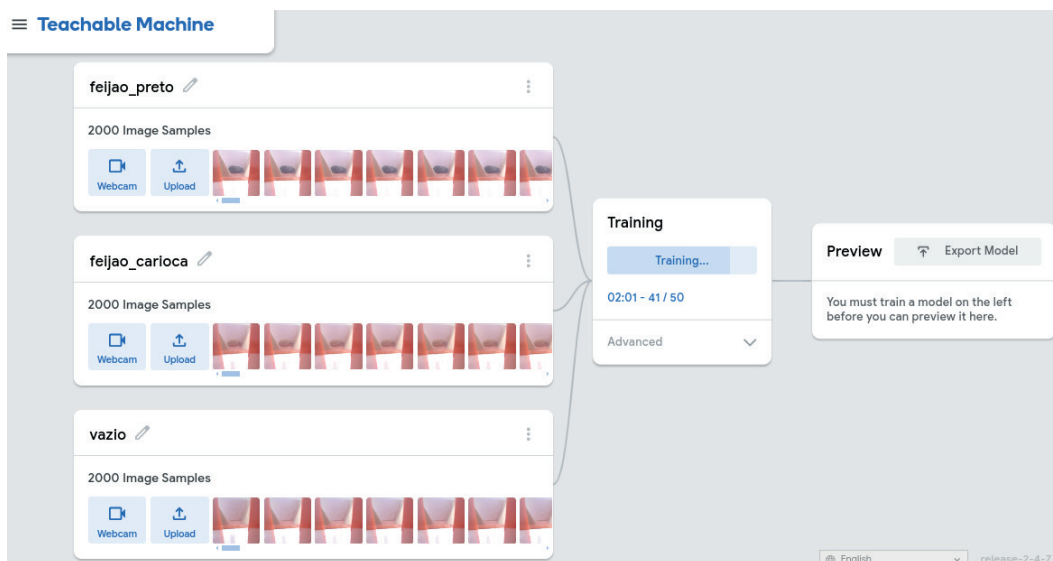
Figura 8 - passos no portal Teachable Machine para iniciar um novo projeto de treinamento de um modelo de redes neurais.



Fonte: autores, 2024.

Para o caso específico deste trabalho, o modelo foi treinado utilizando um conjunto de 2000 imagens de feijões pretos, 2000 imagens de feijões cariocas e 2000 imagens do dispositivo vazio, formando assim três classes as quais foram denominadas no código de: feijao_preto, feijao_carioca e vazio (ver Figura 9).

Figura 9 - definição das classes no treinamento do modelo para a separação de feijões no portal Teachable Machine.



Fonte: autores, 2024.

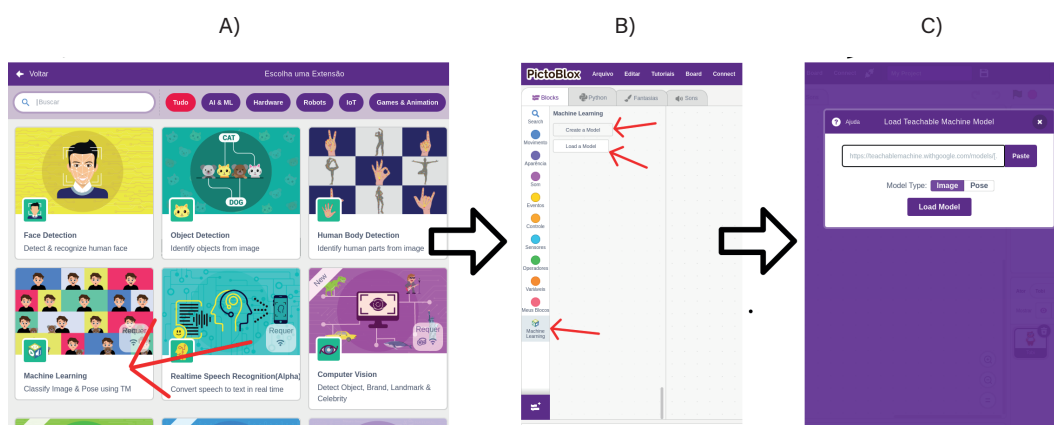
Os hiperparâmetros do algoritmo, como taxa de aprendizado, número de épocas de treinamento e número de camadas ocultas, foram utilizadas os valores padrões do portal, o qual utiliza uma abordagem de validação cruzada para maximizar a precisão da classificação.

Importação do modelo no pictoblox

No portal do Pictoblox, a ativação do complemento de *Machine Learning* é feita clicando no botão indicado pela seta vermelha na Figura 2.

A Figura 10 indica os passos de como carregar um modelo já treinado no Pictoblox, que é composto por três etapas: A) Seleção de Projeto, a tela mostra diversas opções de projetos que podem ser criados no Pictoblox, destacando funcionalidades como detecção facial e detecção de objetos. A seta vermelha aponta para a escolha específica de *Machine Learning*. B) Escolha de Blocos de Função, as setas indicam as opções para carregar o modelo treinado ou iniciar o processo de treinamento. C) Esta mostra a interface onde o usuário indica o link do modelo já treinado, advindo do Teachable Machine.

Figura 10 - treinamento do modelo no portal Teachable Machine.



Fonte: autores, 2024.

Desenvolvimento do código em blocos

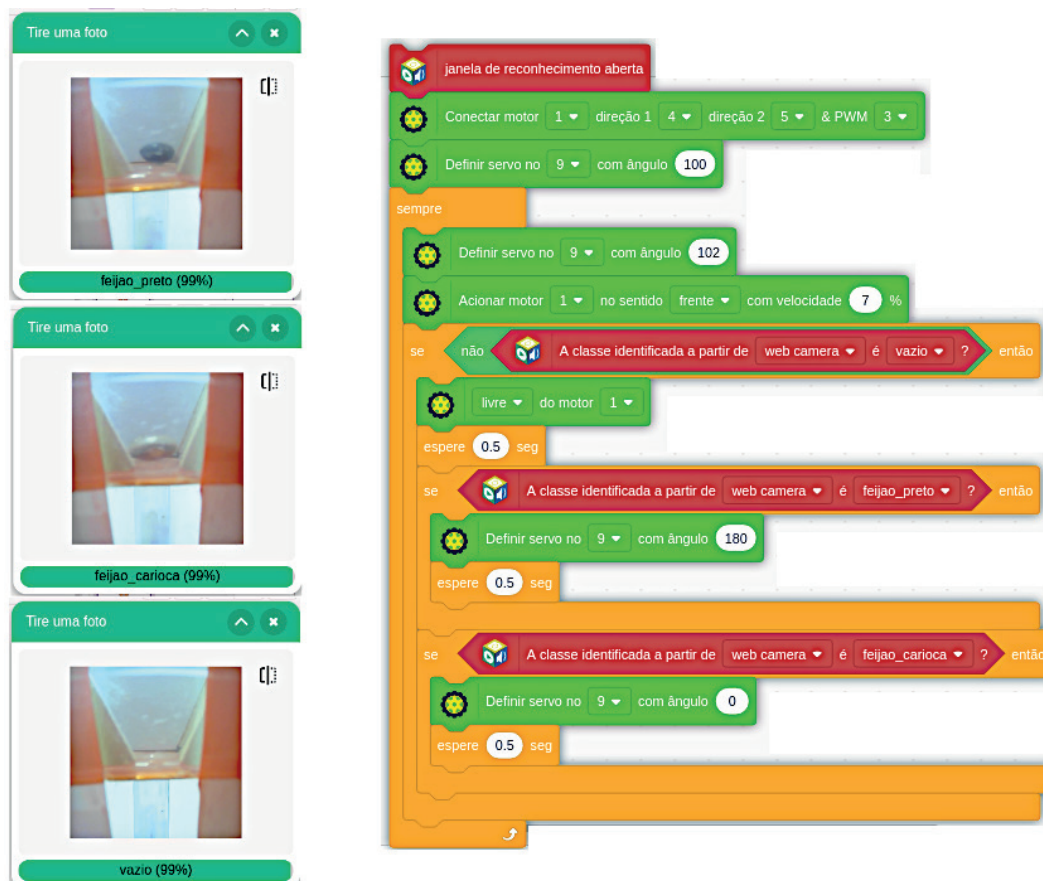
A Figura 11 mostra uma combinação de imagens capturadas por uma *webcam* e o código de programação em blocos para o reconhecimento e classificação dos feijões.

Na esquerda podemos visualizar três imagens do recipiente com grãos. Cada imagem está associada a uma classificação - "feijão preto" com 99% de confiança, "feijão carioca" com 99% de confiança, e um recipiente "vazio" com 99% de confiança. Estes indicadores de confiança, quantificam a certeza do sistema sobre a classificação realizada.

No lado direito estão os blocos de código. O bloco verde dentro do “sempre” configura a posição inicial do servo motor e a direção do motor que regula o sistema de movimentação. Blocos “se” amarelos utilizam as condições baseadas nas classificações feitas pela *webcam*. Dependendo da classificação (“vazio”, “feijão preto”, “feijão carioca”), diferentes blocos são executados para ajustar o servo motor para ângulos específicos, que direcionam os grãos classificados para os recipientes correspondentes.

A parte do código responsável pela mecânica de separação, detecta quando a classe é identificada (e.g., feijão preto), o servo motor é ajustado para um ângulo específico (e.g., 180 graus), o que direciona os grãos para um recipiente específico. Essa programação garante que cada tipo de grão seja separado com precisão conforme identificado pela câmera.

Figura 11 - Imagens da *webcam* dos feijões e do código em blocos final.



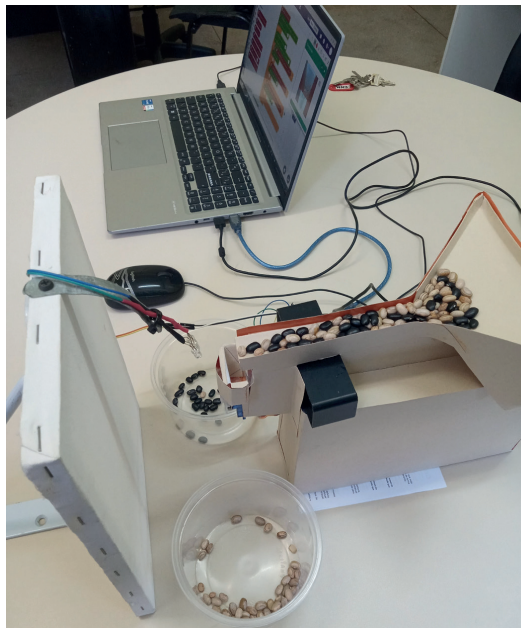
Fonte: autores, 2024.

Desta forma, em resumo, o procedimento experimental foi composto por: ajuste da câmera e da iluminação para otimizar a captura de imagens dos grãos. Início da esteira transportadora para mover os grãos sob a câmera, onde as imagens são capturadas e processadas em tempo real pelo computador. As imagens são analisadas pela CNN e classificadas de acordo com as classes definidas. Decisões de separação são tomadas com base na classificação. O sistema de atuador mecânico, controlado pelo Arduino, é utilizado para direcionar os grãos para recipientes separados conforme sua classificação. A eficiência do sistema é avaliada pela comparação entre a quantidade de grãos corretamente classificados e o total de grãos processados, calculando-se a porcentagem de acertos.

RESULTADOS E DISCUSSÃO

O protótipo foi submetido a uma série de testes para avaliar sua eficácia na separação de grãos de feijão preto e feijão carioca. Os resultados demonstraram que o sistema alcançou uma precisão global de 94,5% nesta classificação. A especificidade para o feijão preto foi de 89%, enquanto para o feijão-carioca foi de 100%. Estes resultados indicam uma alta eficiência do protótipo em condições controladas.

Figura 12 - aparato experimental em funcionamento.



Fonte: autores, 2024.

A alta taxa de precisão alcançada pelo protótipo é satisfatória, especialmente considerando a simplicidade e o baixo custo dos componentes utilizados. A maior taxa de erros para feijão pode ser atribuída a variações maiores em forma e cor nesta categoria de grãos, que apresentam desafios adicionais para o algoritmo de visão computacional, bem como erros no processo de separação mecânica, o qual cabe melhorias.

Além disso, experimentos adicionais para ajustar os hiperparâmetros da rede neural, como a taxa de aprendizado e o número de camadas, poderiam otimizar ainda mais o desempenho.

Este protótipo também serve como uma prova de conceito para a aplicação de técnicas avançadas de inteligência artificial em aplicações de automação agrícola de baixo custo, abrindo caminho para futuras inovações no setor.

Figura 13 - imagem de um dos resultados experimental de separação.



Fonte: autores, 2024.

CONSIDERAÇÕES FINAIS

Este estudo demonstrou a viabilidade de desenvolver um protótipo de baixo custo para a separação de grãos de feijão preto e feijão carioca, utilizando tecnologias avançadas de visão computacional e *Deep Learning*. Através da captura e classificação imediata de imagens, o sistema ativa mecanismos de hardware, como servos e motores, para realizar a separação física dos grãos, exemplificando a integração eficaz entre visão computacional e automação física. O seu funcionamento pode ser visto em Germano (2024).

Os resultados alcançados revelam uma alta eficácia do sistema, com uma precisão global de classificação de 94,5%. Este desempenho é notável, considerando

a simplicidade e o custo acessível dos componentes utilizados. Ainda assim, o protótipo enfrenta desafios que impactam a taxa de erro na classificação dos grãos, destacando áreas para melhorias futuras, como a otimização da iluminação e ajustes finos nos hiperparâmetros da rede neural. Estas modificações têm o potencial de aprimorar ainda mais a precisão do sistema e expandir a aplicabilidade para outros tipos de grãos, além de integrar outras tecnologias emergentes como IoT (Internet das Coisas).

Além de suas aplicações práticas, o protótipo apresenta um valor educacional considerável. A integração de plataformas como o Teachable Machine e o Pictoblox, utilizadas para treinar e implementar modelos de aprendizado de máquina, proporciona uma oportunidade rica para a introdução de conceitos de inteligência artificial e automação em ambientes educacionais. Esta abordagem prática não apenas facilita a compreensão dos estudantes sobre tecnologias complexas, mas também estimula o desenvolvimento de habilidades de resolução de problemas em contextos reais.

Portanto, este trabalho não apenas confirma a eficácia de um sistema acessível para a separação de grãos, mas também estabelece um precedente para futuras pesquisas e desenvolvimento no campo da automação com Arduino e inteligência artificial. Adicionalmente, ele abre caminho para a expansão da aplicabilidade dessas tecnologias em educação, promovendo um ensino inovador e interativo que pode transformar o aprendizado técnico e científico.

REFERÊNCIAS

ARDUINO. Arduino - Home. Disponível em: <https://www.arduino.cc/>. Acesso em: 6 jun. 2024.

FUKUSHIMA, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. **Biological Cybernetics**, v. 36, p. 193–202, 1980.

GERMANO, R. Visão Computacional + Deep Learning + Scratch + Arduino. Disponível em: <https://www.youtube.com/watch?v=3UAW3IGeF2o>. Acesso em: 31 de ago. 2024.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. Disponível em: <https://www.deeplearningbook.org/>. Acesso em: 6 jun. 2024.

GOOGLE. Teachable Machine. Disponível em: <https://teachablemachine.withgoogle.com/>. Acesso em: 6 jun. 2024.

LIU, G. R. **Machine Learning with Python: Theory and Applications**. New Jersey: World Scientific Publishing Co. Pte. Ltd., 2023.

MARJI, M. **Learn to Program with Scratch: A Visual Introduction to Programming with Games, Art, Science and Math**. San Francisco: No Starch Press, 2014.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, v. 5, p. 115-133, 1943.

MINSKY, Marvin; PAPERT, Seymour. **Perceptrons: An Introduction to Computational Geometry**. Cambridge, Mass.: M.I.T. Press, 1969.

PETRELLI, Maurizio. **Machine Learning for Earth Sciences: Using Python to Solve Geological Problems**. Cham: Springer, 2023.

PICTOBLOX. PictoBlox - AI Programming for Kids. Disponível em: <https://pictoblox.ai/>. Acesso em: 6 jun. 2024.

PRODANOV, C. C.; FREITAS, E. C. **Metodologia do Trabalho Científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico**. 2. ed. Novo Hamburgo: Universidade Feevale, 2013.

ROSENBLATT, F. **Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms**. Washington: Spartan Books, 1962.

SWEIGART, A. **Scratch Programming Playground: Learn to Program by Making Cool Games**. San Francisco: No Starch Press, 2016.

ZEILER, M. D.; FERGUS, R. Visualizing and Understanding Convolutional Networks. *In*: FLEET, D. *et al.* (Ed.). **ECCV 2014, Part I, LNCS 8689**. Springer International Publishing, p. 818-833, 2014.

ZOLLANVARI, Amin. **Machine Learning with Python: Theory and Implementation**. Cham: Springer, 2023.

**APÊNDICE B – MODELAGEM MATEMÁTICA DA LEI DE ZIPF: UMA
ANÁLISE COMPARATIVA ENTRE TEXTOS EM PORTUGUÊS E INGLÊS**

IX EPAMM
18 a 19
OUTUBRO 2024
UFPA - BELÉM



IX EPAMM

ENCONTRO PARAENSE DE MODELAGEM MATEMÁTICA

MODELAGEM MATEMÁTICA: INTERLOCUÇÕES DE MÚLTIPLOS
SABERES NO MUNDO CONTEMPORÂNEO

MODELAGEM MATEMÁTICA DA LEI DE ZIPF: UMA ANÁLISE COMPARATIVA ENTRE TEXTOS EM PORTUGUÊS E INGLÊS

Elizangela Silva

UFPA- Campus Castanhal

elizansl02@gmail.com

Renato Germano

UFPA – Campus Castanhal

rgermano@ufpa.br

Resumo:

Este artigo explora a aplicação da Lei de Zipf em dois textos amplamente conhecidos: a Bíblia Sagrada e o romance Dom Casmurro, de Machado de Assis, ambos em suas versões em português e inglês. A Lei de Zipf descreve a relação inversa entre a frequência das palavras e sua posição no ranking de frequência. Utilizando gráficos em escala log-log, os coeficientes de Zipf foram calculados para cada texto. Os resultados mostram que a Bíblia Sagrada adere mais rigidamente à Lei de Zipf, enquanto Dom Casmurro apresenta maior variação lexical. As traduções para o inglês mantiveram os padrões gerais, mas com pequenas variações, destacando o impacto do estilo e da tradução na conformidade com a lei.

Palavras-chave: Lei de Zipf. Dom Casmurro. Bíblia Sagrada. distribuição de palavras.

Introdução

A Lei de Zipf descreve uma relação matemática entre a frequência de ocorrência das palavras em um texto e sua classificação de frequência e é amplamente aplicada em estudos de linguística quantitativa. Segundo Zipf, a frequência de uma palavra em uma língua natural é inversamente proporcional à sua posição no ranking de frequências, o que resulta em uma curva descendente quando os dados são plotados em uma escala log-log (PIANTADOSI, 2014). Essa regularidade foi observada em diversos idiomas e tipos de texto, sugerindo que os padrões de uso das palavras seguem princípios universais, independentemente do estilo ou gênero textual.

Os primeiros estudos de Zipf apontaram para a aplicação dessa lei em corpora variados, abrangendo textos literários, religiosos e científicos (POWERS, 1998). O padrão de decaimento exponencial é observado tanto em textos clássicos quanto em obras contemporâneas, o que torna a Lei de Zipf uma ferramenta poderosa para investigar as

IX EPAMM
18 a 19
OUTUBRO 2024
UFPA - BELÉM



IX EPAMM

ENCONTRO PARAENSE DE MODELAGEM MATEMÁTICA

MODELAGEM MATEMÁTICA: INTERLOCUÇÕES DE MÚLTIPLOS
SABERES NO MUNDO CONTEMPORÂNEO

estruturas de linguagem em diferentes contextos. Essa abordagem tem sido utilizada em análises de grandes corpora, como a Bíblia Sagrada, e em obras literárias significativas, como os romances de Machado de Assis, o que possibilita um olhar comparativo sobre a distribuição de frequências em diferentes gêneros e idiomas.

A Bíblia Sagrada, por exemplo, é conhecida por seu uso recorrente de vocabulário religioso específico e altamente repetitivo, especialmente em termos de nomes, verbos e expressões fixas. Já textos literários, como *Dom Casmurro*, de Machado de Assis, apresentam uma maior variedade lexical, com nuances estilísticas que influenciam diretamente a distribuição das palavras (MOREIRA, 2021). Essas características tornam a análise desses dois tipos de texto uma oportunidade valiosa para verificar como a Lei de Zipf se manifesta em diferentes contextos linguísticos e culturais.

Neste artigo, exploramos a aplicação da Lei de Zipf em dois textos amplamente conhecidos: a Bíblia Sagrada e o romance *Dom Casmurro*. Ambos os textos foram analisados em suas versões em português e inglês, com o objetivo de identificar como as frequências das palavras variam entre os idiomas e os gêneros textuais. Além disso, a utilização de gráficos em escala log-log permitiu calcular os coeficientes de Zipf para cada um dos textos, fornecendo uma visão quantitativa da distribuição de palavras e da aderência à Lei de Zipf (PIANTADOSI, 2014; POWERS, 1998).

Os resultados indicam variações interessantes entre os dois textos e suas respectivas traduções, sugerindo que diferentes estilos de escrita, gêneros literários e contextos culturais podem influenciar a distribuição de frequências das palavras. A análise comparativa entre as versões em português e inglês também contribuiu para uma melhor compreensão de como as estruturas linguísticas e estilísticas são mantidas ou alteradas em traduções literárias e religiosas.

Metodologia

Este estudo teve como objetivo analisar a aplicabilidade da Lei de Zipf em dois textos amplamente conhecidos e de grande relevância: a Bíblia Sagrada e o romance *Dom Casmurro*, de Machado de Assis. Para tal, foram analisadas as versões originais em português de ambos os textos, bem como suas traduções para o inglês. A metodologia

IX EPAMM
18 a 19
OUTUBRO 2024
UFPA - BELÉM



IX EPAMM

ENCONTRO PARAENSE DE MODELAGEM MATEMÁTICA

MODELAGEM MATEMÁTICA: INTERLOCUÇÕES DE MÚLTIPLOS
SABERES NO MUNDO CONTEMPORÂNEO

adotada compreendeu as etapas de seleção e preparação dos textos, contagem de frequência de palavras, geração de gráficos log-log e cálculo dos coeficientes da Lei de Zipf, com o intuito de comparar as distribuições de palavras entre os textos e entre os idiomas.

Inicialmente, foram selecionadas as versões digitais completas da Bíblia Sagrada, tanto em português quanto em inglês e do romance Dom Casmurro, em suas respectivas versões nas duas línguas. Os textos foram processados digitalmente para remover elementos não linguísticos, como pontuações, números e outros caracteres especiais, e todas as palavras foram convertidas para minúsculas, a fim de garantir consistência na contagem de frequência.

Em seguida, foi realizada a contagem da frequência de ocorrência de cada palavra distinta nos textos, utilizando-se técnicas de processamento de linguagem natural para identificar e registrar a frequência de cada termo. A contagem foi organizada em ordem decrescente de frequência, possibilitando a geração de uma lista classificada (rank) de todas as palavras presentes nos textos.

Com os dados de frequência em mãos, foram gerados gráficos em escala log-log, nos quais o logaritmo da posição de cada palavra no ranking foi plotado contra o logaritmo de sua frequência de ocorrência permitindo a visualização da distribuição das palavras e a verificação da aderência dos dados à Lei de Zipf.

Para determinar os coeficientes da Lei de Zipf em cada conjunto de dados, foi realizada uma regressão linear nos gráficos log-log. A inclinação da reta ajustada aos dados representou o coeficiente de Zipf, que foi interpretado como uma medida da distribuição de frequências das palavras. Por fim, os resultados foram comparados entre si, com o intuito de identificar possíveis variações na distribuição de palavras causadas pelas diferenças linguísticas e de estilo entre um texto religioso e um texto literário.

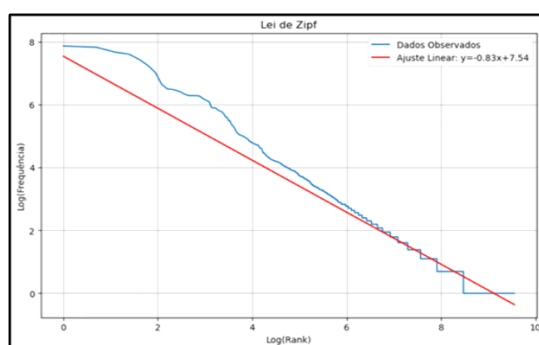
Resultados e Discussão

Na versão em português de Dom Casmurro, o coeficiente de Zipf encontrado foi de aproximadamente -0,83, valor ligeiramente inferior ao coeficiente clássico de -1, sugerido pela Lei de Zipf. Esse resultado indica que, embora o texto siga a distribuição



esperada, a inclinação da curva é menos acentuada, o que pode ser explicado pela rica variação lexical e pelo uso repetido de determinados termos estilísticos característicos da obra de Machado de Assis. A estrutura narrativa do romance, com variações linguísticas e expressões peculiares, parece influenciar diretamente essa suavização na distribuição de frequências.

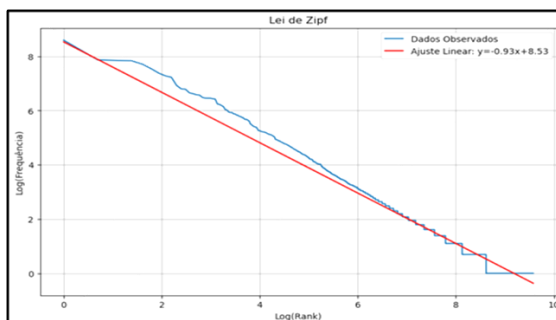
Figura 1 – Coeficiente de Zipf na versão em português de Dom Casmurro



Fonte: Autores, 2024.

Por outro lado, a versão em inglês de Dom Casmurro apresentou um coeficiente de Zipf mais próximo do valor clássico, aproximadamente -0,93. Esse resultado sugere que a tradução preservou razoavelmente o padrão de frequência de palavras do texto original, mantendo uma aderência maior à Lei de Zipf. No entanto, pequenas diferenças são esperadas, uma vez que a tradução envolve adaptações estilísticas e culturais, o que pode impactar a frequência de certos termos. Esse comportamento reforça a ideia de que, embora as traduções se esforcem para manter a estrutura semântica, nuances linguísticas específicas de cada idioma podem influenciar a distribuição das palavras.

Figura 2 – Coeficiente de Zipf na versão em inglês de Dom Casmurro



Fonte: Autores, 2024.

IX EPAMM
18 a 19
OUTUBRO 2024
UFPA - BELÉM



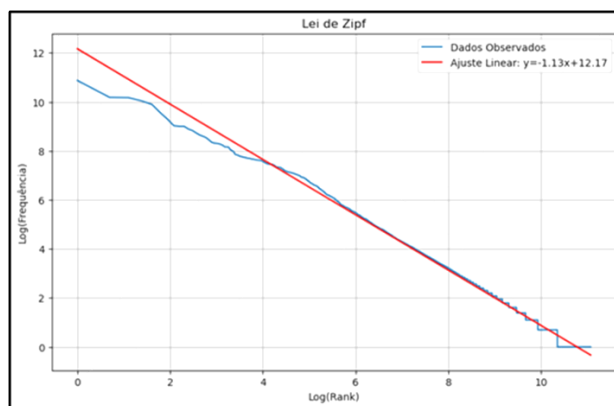
IX EPAMM

ENCONTRO PARAENSE DE MODELAGEM MATEMÁTICA

MODELAGEM MATEMÁTICA: INTERLOCUÇÕES DE MÚLTIPLOS
SABERES NO MUNDO CONTEMPORÂNEO

Na análise da Bíblia Sagrada, os resultados mostraram uma maior aderência à Lei de Zipf. A versão em português apresentou um coeficiente de Zipf de $-1,13$, indicando uma conformidade mais rigorosa com a distribuição clássica. Esse comportamento pode ser atribuído à natureza formal e repetitiva do texto bíblico, em que termos religiosos e expressões fixas aparecem com grande frequência ao longo do texto. A estrutura particular da Bíblia, com sua linguagem repetitiva e solene, parece seguir de maneira mais estrita o padrão proposto por Zipf, em comparação com um texto literário.

Figura 3 – Coeficiente de Zipf na versão em português da Bíblia Sagrada



Fonte: Autores, 2024.

A versão em inglês da Bíblia Sagrada resultou em um coeficiente de Zipf ainda mais íngreme, aproximadamente $-1,33$. Esse valor sugere que a tradução intensificou a repetição de certos termos comuns, ampliando a diferença de frequência entre as palavras mais usadas e as menos frequentes. A tradução da Bíblia, em particular, pode ter acentuado essas diferenças devido à busca por manter uma fidelidade semântica rigorosa, que tende a reforçar a utilização de um vocabulário religioso específico, muitas vezes repetido ao longo das passagens.

IX EPAMM
18 a 19
OUTUBRO 2024
UFPA - BELÉM

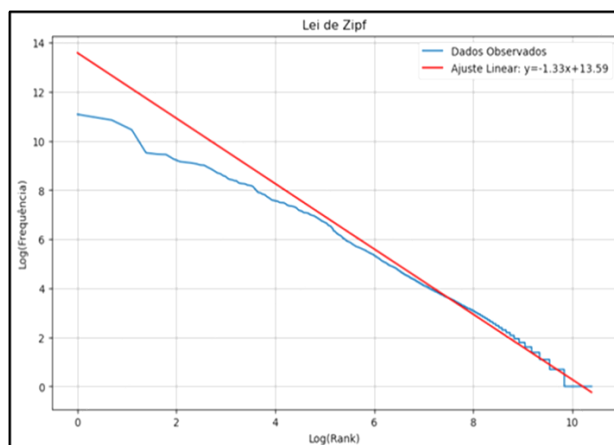


IX EPAMM

ENCONTRO PARAENSE DE MODELAGEM MATEMÁTICA

MODELAGEM MATEMÁTICA: INTERLOCUÇÕES DE MÚLTIPLOS
SABERES NO MUNDO CONTEMPORÂNEO

Figura 4 – Coeficiente de Zipf na versão em inglês da Bíblia Sagrada



Fonte: Autores, 2024.

Ao comparar os dois textos, percebe-se que a Bíblia Sagrada se conforma mais estritamente à Lei de Zipf em ambas as línguas, com coeficientes próximos de -1, especialmente na versão em português. Essa característica está relacionada ao estilo mais formal e repetitivo do texto bíblico, em contraste com o estilo mais livre e lexicalmente diversificado de Dom Casmurro. Além disso, a análise das traduções revelou que, apesar da tentativa de manter o conteúdo e o estilo do texto original, as traduções introduzem variações na distribuição das palavras, como observado nos coeficientes de Zipf calculados.

Desta forma, os resultados indicam que a Lei de Zipf se manifesta de forma consistente nos dois gêneros textuais, com variações influenciadas tanto pelo estilo dos textos quanto pelas diferenças entre as línguas. Enquanto a Bíblia Sagrada apresenta uma distribuição mais estável e repetitiva de palavras, Dom Casmurro mostra uma maior flexibilidade linguística, refletida em coeficientes de Zipf menores. Além disso, as traduções para o inglês mostraram-se capazes de manter os padrões gerais de distribuição de palavras, embora pequenas discrepâncias tenham sido observadas, especialmente na Bíblia. Esses achados ressaltam o impacto das traduções e dos estilos textuais na conformidade com a Lei de Zipf.

IX EPAMM
18 a 19
OUTUBRO 2024
UFPA - BELÉM



IX EPAMM

ENCONTRO PARAENSE DE MODELAGEM MATEMÁTICA

MODELAGEM MATEMÁTICA: INTERLOCUÇÕES DE MÚLTIPLOS
SABERES NO MUNDO CONTEMPORÂNEO

Considerações Finais

Este estudo investigou a aplicação da Lei de Zipf na Bíblia Sagrada e no romance Dom Casmurro, comparando suas versões em português e inglês. Os resultados mostraram que a Bíblia Sagrada adere mais rigorosamente à Lei de Zipf, com coeficientes próximos de -1, tanto em português quanto em inglês, devido à sua natureza repetitiva e formal. Em contraste, Dom Casmurro apresentou coeficientes ligeiramente menores, refletindo a maior variação lexical e o estilo literário mais diversificado de Machado de Assis. As traduções para o inglês mantiveram os padrões gerais de distribuição de palavras, mas introduziram pequenas variações, especialmente na Bíblia Sagrada, evidenciando o impacto das nuances linguísticas e culturais no processo de tradução.

Conclui-se que a Lei de Zipf se aplica de forma consistente aos textos analisados, mas o estilo textual e as adaptações de tradução influenciam sua conformidade. Este estudo abre possibilidades para futuras pesquisas sobre a aplicação dessa lei em outros gêneros e sobre o impacto das traduções na distribuição de palavras.

Referências

MOREIRA, Olga (Ed.). **Fractal Analysis**. Burlington, ON: Arcler Press, 2021.

Disponível em: <https://arclerpress.com>. Acesso em: 10 set. 2024.

PIANTADOSI, Steven T. **Zipf's word frequency law in natural language: A critical review and future directions**. *Psychonomic Bulletin & Review*, v. 21, n. 5, p. 1112-1130, 2014. Disponível em: <https://doi.org/10.3758/s13423-014-0585-6>. Acesso em: 10 set. 2024.

POWERS, David M. W. **Applications and Explanations of Zipf's Law**. In: **NeMLaP3/CoNLL98: New Methods in Language Processing and Computational Natural Language Learning**, ACL, 1998. p. 151-160

Anexos

**ANEXO A – CERTIFICADO DE ACEITE PARA PUBLICAÇÃO NA EDITORA
CIENTÍFICA DIGITAL**



CERTIFICADO DE ACEITE PARA PUBLICAÇÃO

Submetido em: **11/06/2024**

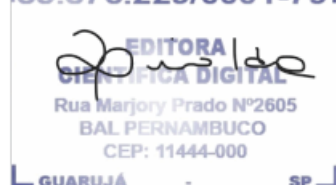
Aceito em: **15/06/2024**

Certificado N°: **240616822**

Certificamos que o trabalho intitulado "UTILIZAÇÃO DE VISÃO COMPUTACIONAL E DEEP LEARNING PARA SEPARAÇÃO DE GRÃOS UM PROTÓTIPO DE BAIXO CUSTO BASEADO NO ARDUÍNO" de autoria de Elizângela Maria Gonçalves Silva, Renato Germano, foi submetido para revisão de pares externos (Peer Review) e atendeu a todos os critérios de qualidade estipulados na política editorial obtendo parecer de ACEITO para publicação. O trabalho encontra-se em processo editorial e será publicado no formato de Capítulo de Livro em 30/09/2024. A obra será publicada com ISBN, DOI, Conselho Editorial e Certificado de Publicação.

Reinaldo Cardoso

Editor-Geral | Fone: (11) 229/0001-791



**ANEXO B – CERTIFICADO DE ACEITE NO IX ENCONTRO PARAENSE DE
MODELAGEM MATEMÁTICA**

18 a 19
OUTUBRO 2024
UFPA - BELÉM
ISSN 1982-8691

IX EPAMM

ENCONTRO PARAENSE DE MODELAGEM MATEMÁTICA

Modelagem Matemática: interlocuções de Múltiplos saberes no
mundo contemporâneo



CERTIFICADO

Certificamos que o trabalho intitulado **MODELAGEM MATEMÁTICA DA LEI DE ZIPF: UMA ANÁLISE COMPARATIVA ENTRE TEXTOS EM PORTUGUÊS E INGLÊS** de autoria de Elizangela Maria Gonçalves Silva e RENATO GERMANO REIS NUNES, foi submetido, aceito e apresentado na modalidade Pôster, no evento **IX ENCONTRO PARAENSE DE MODELAGEM MATEMÁTICA**, realizado no período de 18 a 19 de novembro de 2024, na cidade cidade de Belém-PA.

Belém-PA, 19 de novembro de 2024.

Líder do GEMM

Grupo de Estudos em Modelagem Matemática

Coordenadora IX EPAMM

Vice Líder do GEMM

