

AJC

Academic Journal on Computing, Engineering and Applied Mathematics

EAM

2025

Volume 6 Issue 1

Academic Journal on Computing, Engineering and Applied Mathematics

Universidade Federal do Tocantins

Reitor

Prof. Dr. Luís Eduardo Bovolato

Vice-Reitor

Prof. Dr. Marcelo Leineker Costa

Pró-Reitoria de Graduação

Prof. Dr. Eduardo José Cezari

Pró-Reitoria de Pesquisa e Pós-Graduação

Prof. Dra. Karylleila dos Santos Andrade Klinger

Pró-Reitoria de Extensão e Cultura

Profa. Dra. Maria Santana Ferreira dos Santos

Pró-Reitoria de Administração e Finanças

Me. Carlos Alberto Moreira de Araújo Júnior

Pró-Reitoria de Assuntos Estudantis e Comunitários

Prof. Dr. Kherlley Caxias Batista Barbosa

Pró-Reitoria de Avaliação e Planejamento

Prof. Dr. Eduardo Andrea Lemus Erasmo

Pró-reitoria de Gestão e Desenvolvimento de Pessoas

Profa. Dra. Vânia Maria de Araújo Passos

Pró-Reitoria de Tecnologia da Informação e Comunicação

Me. Werley Teixeira Reinaldo

Direção do Campus de Palmas

Prof. Dr. Moisés de Souza Arantes Neto

Coordenação do Curso de Ciência da Computação

Prof. Dr. Andreas Kneip

Dados Internacionais de Catalogação na Publicação (CIP)

Academic Journal on Computing, Engineering and Applied Mathematics (AJCEAM) [recurso eletrônico] / Universidade Federal do Tocantins, Curso de Ciência da Computação. – vol. 06, n. 01 ([september/march], 2025) – Palmas - TO, UFT, 2025. ISSN nº 2675-3588.

Quadrimestral no primeiro ano de publicação 2020

Semestral.

Disponível em:

<https://sistemas.uft.edu.br/periodicos/index.php/AJCEAM/index>

1. Ciência da Computação - periódico. 2. Matemática Aplicada. 3. Computação Aplicada. 4. Engenharias. 5. Ciências Exatas. I. Universidade Federal do Tocantins.

CDD 22.ed. 004

Ficha Catalográfica elaborada por Edson de Sousa Oliveira – CRB/2 – 1069.

Expediente

Editor-Chefe

Dr. Warley Gramacho da Silva (UFT), Brasil

Editores

Dr. Edeilson Milhomem Silva (UFT), Brasil

Dr. Marcos Antônio Estremeto (ETEC-SP), Brasil

Dr. Rafael Lima de Carvalho (UFT), Brasil

Dr. Tanilson Dias dos Santos (UFT), Brasil

Me. Tiago da Silva Almeida (UFT), Brasil

Dr. Warley Gramacho da Silva (UFT), Brasil

Realização

Fundação Universidade Federal do Tocantins (UFT)

Quadra 109 Norte, Avenida NS-15, ALCNO-14 | Bloco III | sala 214 | Plano Diretor Norte | 77001-090 | Palmas / TO | Brasil

Periodicidade

Este periódico possui periodicidade semestral e utiliza a Licença Creative Commons 4.0 - CC BY-NC 4.0. Contudo, a publicação dos artigos em modalidade avançada ou ahead of print, ou seja, tão logo os manuscritos aprovados sejam editados para publicação, é possível. O AJCEAM não possui taxas de publicação, tanto pouco de submissão de manuscritos, sendo totalmente gratuita para autores e leitores.

Indexadores

Google Acadêmico, desde 9 de maio de 2020

International Standard Serial Number – ISSN, desde 28 de maio de 2020

Crossref, desde 7 de junho de 2020

Revistas de Livre Acesso – LivRe, desde 24 de junho de 2020

Diretório das revistas científicas eletrônicas brasileiras – Miguilim, desde novembro de 2022

Sumário

- | | |
|---|----------|
| 1 Challenges and Opportunities in the Adoption of Automation Technologies in Small-Scale Agriculture: A Path Toward Sustainability | 1 |
| BARBOSA ET. AL. | |
| 2 Formal description of activation functions of machine learning models | 9 |
| LEAL, APOLINÁRIO AND ROCHA | |

Challenges and Opportunities in the Adoption of Automation Technologies in Small-Scale Agriculture: A Path Toward Sustainability

Vinícius Cristóvão da Silva Barbosa¹, Kathy Camila Cardozo Osinski Senhorini¹, Stefani Carolline Leal de Freitas¹ and Jadiel Caparrós da Silva¹

¹ Universidade Federal do Tocantins (UFT), Tocantins, Brasil

Reception date of the manuscript: 17/10/2024

Acceptance date of the manuscript: 27/11/2024

Publication date: 02/12/2024

Abstract— This article explores the critical challenges and opportunities faced by small-scale farmers in adopting automation technologies in their agricultural operations. Automation has the potential to optimize resource usage, reduce operational costs, and streamline tasks that traditionally rely on manual labor. Despite the clear benefits, small farmers encounter a range of barriers, including financial constraints, limited access to advanced technology, and the need for technical expertise. However, these challenges can be addressed through means economically viable and scalable, tailored to the unique needs of small-scale agriculture. The role of engineers is pivotal in this process, as they develop innovative solutions such as integrating sensors for real-time monitoring or employing precision agriculture techniques with drones. By identifying and implementing appropriate technologies, small-scale farmers can enhance productivity, sustainability, and their participation in competitive markets.

Keywords—Agricultural Automation, Small-Scale Agriculture, Precision Technology, Sustainability, Technological Innovation, Technical Training

I. INTRODUCTION

Automation refers to the use of new technologies to perform tasks and operations autonomously or with minimal human intervention [1]. In the context of agriculture, it involves the application of sensors, actuators, software, and intelligent machinery that optimize agricultural production by automating activities that were previously dependent solely on manual labor [2]. Technologies have revolutionized agriculture by enabling real-time data collection, precise analysis, and automated actions, which collectively enhance the efficiency of agricultural management.

In addition to automation, precision agriculture has emerged as a central strategy for improving resource use and increasing productivity in a sustainable way [3]. This concept involves detailed monitoring of environmental variables such as soil moisture, temperature, nutrient levels, and plant disease detection [4]. With these insights, more accurate decisions can be made, ensuring optimized use of inputs while minimizing waste. Technologies such as drones, GPS systems, and mapping software are essential tools for

managing both small and large agricultural properties with greater precision [5].

Automation in small-scale agriculture has the potential to transform the sector, offering significant benefits such as increased productivity, reduced costs, and environmental sustainability. However, the adoption of these technologies faces considerable barriers [6]. Small farmers, especially those engaged in family or subsistence farming, often encounter challenges such as high initial implementation costs, lack of appropriate infrastructure, and insufficient technical expertise to operate automated systems [7].

Despite these obstacles, the development of affordable and tailored technological solutions aimed at the specific needs of small-scale agriculture can help overcome these barriers [8]. Initiatives involving low-cost sensors, accessible drones, and the implementation of renewable energy sources offer new opportunities for modernizing agricultural operations in small properties [9]. Furthermore, engineers and technicians have played a fundamental role in creating automation tools and methods that are financially viable and adapted to the needs of small producers [10].

Address certain topics, such as the main challenges in implementing automation in agriculture, can help clarify for the reader the key risks and precautions involved in automation. However, by exploring these issues comprehensively, it is also possible to highlight the opportunities that

arise with the advancement of automation. Initiatives aimed at reducing costs, improving infrastructure in rural areas, and offering training programs can significantly contribute to overcoming these barriers. This approach not only promotes the inclusion of small-scale farmers in the use of cutting-edge technologies but also drives agricultural sustainability by enabling more efficient management of natural resources and more profitable, resilient production. Therefore, understanding the challenges and opportunities related to the adoption of automation technologies is a key step in guiding public policies and development strategies that can support the digital transformation of small-scale agriculture.

This paper explores the main challenges faced by small-scale agriculture in adopting automation, as well as the opportunities for technological innovation that can contribute to the sustainable development of the sector. We analyze the various forms of automation applicable to this context, focusing on technologies that enable the reduction of operational costs, efficient use of natural resources, and improvement of agricultural productivity.

II. MAIN CHALLENGES

The adoption of automation technologies in small-scale agriculture faces a series of challenges that go beyond technical and financial issues. It is crucial to also consider social, political, environmental, and climatic factors, which directly affect the ability of small farmers to implement automation solutions in their operations. To address these difficulties in a structured manner, the following subsections will list the main challenges faced by this group, each accompanied by a viable solution proposal. This approach not only details the barriers encountered in the context of precision agriculture and automation but also suggests practical alternatives that can mitigate these obstacles and facilitate the adoption of these innovative technologies in the field.

a. High Costs

The cost of acquiring automated equipment is one of the most common challenges when discussing automation in small-scale agriculture [7]. In specific contexts, Depending on the type of farming or scale, the initial investment required for machinery, such as automated tractors, smart irrigation systems, and sensors, can be quite high [8, 9]. In addition to the upfront investment, ongoing costs related to maintenance, operation, repairs, and part replacement further increase the financial burden for small farmers [11].

According to [12], many of small-scale farmers are unable to secure adequate financing for automation projects, significantly limiting their capacity to modernize their operations. As a result, they often remain reliant on subsistence or low-scale agriculture, which depends almost exclusively on manual labor.

Proposed Solution: To overcome these challenges, the use of cost-effective technologies can be explored. For example, small-scale farmers can start with affordable options like temperature and humidity sensors, which can monitor soil and environmental conditions. Such incremental adoption of

automation, starting with the most economically accessible solutions, can help lower initial barriers and gradually introduce more sophisticated technologies as financial resources allow. Another proposed solution is the creation of public policies that encourage rural credit with lower interest rates and longer payment terms, which can be a way to allow farmers to have access to new technologies. Government programs or partnerships with private financial institutions can also provide specific lines of credit for the acquisition of automated systems.

The authors of [13] show the improvements achieved through a study of former revolutionary areas in China and what improvements in consumer welfare of rural residents were achieved through public support policies. This work relates to cost reductions by demonstrating how public support policies, such as fiscal transfers and tax incentives, help alleviate economic burdens in underdeveloped regions. The policies narrow urban–rural income gaps, enhance local financial capacities, and improve resource allocation efficiency, which collectively reduce the relative costs of living and consumption for rural residents.

In addition to the financial burden of automation, small farmers often face difficulties in accessing rural credit, limiting their ability to invest in innovative technologies [8]. According to recent studies [14, 15], a significant percentage of small farmers are unable to secure financing for automation projects due to a lack of collateral and high-interest rates. Furthermore, the absence of government incentives and policies that promote technological adoption in small-scale agriculture exacerbates the problem [15].

Proposed Solution: Policymakers need to address this challenge by providing easier access to credit and offering government subsidies or low-interest loans aimed specifically at small farmers who wish to adopt automation technologies. Establishing cooperative models and partnerships between farmers, engineers, and technology companies can also foster shared ownership of resources, reducing individual investment costs and making automation more accessible.

According to [14], a study shows that a poorly functioning rural credit market is an example of a delay in the application of new technologies, and obviously the cost-benefit, for rural producers. In [15] a study in India is shown on how costs create a financial burden and hinder investment in new rural automation technologies.

b. Limited Infrastructure

Infrastructure is another major challenge in small-scale agriculture, particularly in rural or remote areas where access to stable electricity and communication networks is often limited or unreliable [16, 17]. Many automated systems require a constant and reliable energy source to function, and in areas without consistent access to electricity, this can be a significant impediment. Similarly, the lack of internet connectivity in rural areas hinders the implementation of technologies that rely on real-time data transmission and monitoring [17].

Proposed Solution: Renewable energy solutions, such as solar power, can provide a reliable and sustainable energy source for automation systems in remote areas [18]. Solar-powered irrigation systems, for example, can be installed in

farms without access to the electrical grid. Additionally, communication systems that do not rely on conventional internet access, such as those operating via SMS or radio frequencies, can be implemented to overcome the lack of connectivity and still allow effective automation. Figure 1 brings a representation of a functional irrigation system powered by solar energy [19].

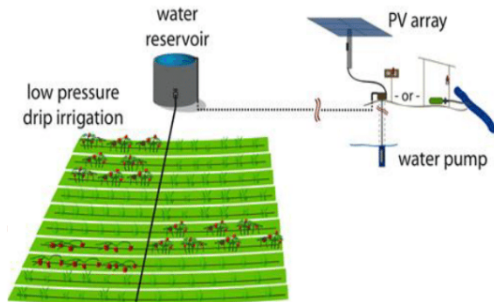


Fig. 1: Representation of an irrigation system that uses solar energy [19].

The study in [18] explores the potential of renewable energy in rural areas to support sustainable development and improve quality of life. It highlights the significant benefits to the local economy, education, and environmental awareness, with active community involvement being crucial for project success. The research emphasizes the importance of inclusive planning and community empowerment for sustainable energy initiatives. In [17] shows that expanding connectivity in Brazil through mobile and fixed broadband can significantly benefit underserved populations by improving access to services, reducing poverty, and supporting education. Small Internet providers, present in 70% of municipalities, play a key role in bridging the digital divide, particularly in less commercially attractive regions. Implementing automation in agriculture often faces significant challenges due to infrastructure limitations, particularly in rural or remote areas where advanced systems are most needed. Poor internet connectivity and unreliable power supplies hinder the adoption of Internet of Things (IoT) devices and automated machinery that require stable networks and consistent energy sources. For instance, farmers in regions with limited broadband coverage struggle to utilize precision agriculture tools that rely on real-time data exchange. By addressing those systemic issues, such as expanding rural broadband or investing in renewable microgrids, the potential of automation in agriculture can be more fully realized, bridging the gap between technology availability and practical implementation.

c. Need for Skilled Labor and Technical Knowledge

Operating modern and precise automated equipment often requires specialized training, which small-scale farmers typically do not possess [20]. This creates a reliance on hiring skilled labor, which can be expensive and difficult to find, especially in rural regions where qualified technicians may not be readily available [21].

Proposed Solution: Providing training and capacity-building

programs through rural extension services can equip farmers with the knowledge and skills needed to operate and maintain automated systems. This would reduce the need for external technical assistance and empower farmers to take full advantage of the available technologies, as shown in [22], farmers involved in Bangladesh's National Agricultural Technology Program (NATP) have significantly enhanced their capabilities. Data from 120 rice and vegetable farmers reveal that over 75% of respondents believe that NATP's group-based extension services have improved their skills in cultivation techniques, resource use, farm inputs, communication, market access, and participation in farmer and community organizations. The study suggests that integrating the Common Interest Group model into the extension system could be beneficial after donor support ends.

In [20] it was examined how robotics, AI, and advanced data analytics are reshaping agriculture to meet the demands of a growing global population, which is expected to increase by nearly two billion in the next three decades. Automation in farming, including robotic systems and smart technologies, offers significant benefits like optimized resource use and reduced environmental impact. However, the study also identifies key challenges such as high initial costs, connectivity problems, the lack of technical knowledge of small farmers, and data security issues that need addressing. This analysis provides insights into the potential and hurdles of adopting these technologies in modern agriculture and how the dissemination of technological knowledge applied to the needs of each reality can be done.

d. Environmental and Climatic Variability

Small-scale agriculture is particularly vulnerable to environmental and climatic fluctuations, which can affect the effectiveness of automation technologies [23]. For example, extreme weather events, such as droughts or floods, can damage automated systems or render them ineffective if they are not properly designed to withstand such conditions [24].

Furthermore, many farmers have limited awareness of the long-term benefits of automation as a tool for climate change adaptation. They may prioritize traditional farming practices that, in the short term, appear safer, despite being less efficient and sustainable [25].

Proposed Solution: To mitigate these risks, automation systems must be designed with environmental resilience in mind. This includes developing weather-resistant equipment and integrating climate monitoring technologies that can adapt to environmental changes in real time. Additionally, government and institutional support in the form of climate adaptation programs could help small-scale farmers invest in more robust automation systems that can endure environmental stressors.

The study in [25] reveals significant methodological gaps in examining the impact of weather variation on agricultural output and migration. By reviewing 22 out of 18,929 articles published between January 2010 and June 2022, it identifies issues such as inadequate sample size reporting, inconsistent adherence to model assumptions, and disorganized methodology sections. Despite the frequent use of time series data and various statistical techniques,

many studies lacked methodological rigor, affecting their reliability. This review underscores the need for improved methodological documentation in researching weather-related migration through agricultural channels, providing critical insights for both researchers and policymakers.

e. Cultural Issues and Resistance to Change

Lastly, one cannot overlook the cultural aspects and resistance to change among small-scale farmers [26]. Many family-owned farms have relied on established agricultural practices for decades and often view automation with suspicion or as a threat to their autonomy [25].

Moreover, more traditional farming communities may resist the adoption of technologies that seem disconnected from their everyday realities [27]. The introduction of automation requires a gradual adaptation process that respects traditional knowledge while clearly demonstrating the tangible benefits of modernization.

Proposed Solution: The development of educational campaigns that include practical demonstrations of the benefits of automation, along with success stories from similar communities, can help reduce resistance. It is crucial that automation is presented not as a replacement for traditional practices, but as a tool to improve agricultural sustainability and efficiency.

As highlighted in [27], the current geopolitical and socioeconomic landscape presents significant challenges for farming and agri-food businesses. Technological innovation, while not a cure-all, can help reduce environmental impacts and improve efficiency. However, barriers such as economic conditions, bureaucratic hurdles, market uncertainties, and personal factors impede its adoption. This paper explores these social, technical, and systemic obstacles and identifies research gaps for future study. Overcoming these barriers with adaptive strategies at various levels is essential for advancing agri-technology in a complex and uncertain world. And shown in [26], China's agricultural technology extension system faces a critical issue called "technological disembeddedness", where formal extension methods, fail to meet farmers' practical needs. Fieldwork in Shandong Province (2019-2020) reveals that varied stakeholder interests have turned extension into a profit-driven tool, marginalizing small-scale farmers and diverging from its original goals. The study advocates for a shift towards a community-centered approach, suggesting that aligning technology extension with local realities and community dynamics will better serve farmers' needs.

f. Considerations on these issues

In conclusion of this subject, the challenges faced by small-scale farmers in adopting automation technologies are multifaceted, spanning economic, technical, social, and environmental domains. While high costs, limited access to credit, insufficient infrastructure, and a lack of technical knowledge remain significant barriers, targeted solutions can pave the way for broader implementation. By addressing these issues through a combination of affordable technologies, government support, and capacity-building programs, small-scale farmers can not only overcome these

obstacles but also harness the potential of automation to improve productivity, sustainability, and resilience. It is clear that a concerted effort from both public and private sectors will be necessary to fully unlock the benefits of automation in small-scale agriculture.

III. OPPORTUNITIES IN THE ADOPTION OF AUTOMATION TECHNOLOGIES

Automation presents numerous opportunities for farmers, especially small-scale, offering not only increased operational efficiency but also the potential to transform the sustainability and competitiveness of family farming [1, 2, 3, 7]. To explore these opportunities more clearly, they can be categorized into three subcategories: Technological Opportunities, Economic Benefits, and Environmental Sustainability.

a. Technological Opportunities

The fast advancement of technology provides small farmers with tools that were once only accessible to large-scale producers [20]. Access to technologies such as drones, soil sensors, smart irrigation systems, and farm management software allows farmers to optimize production with precision and reduced manual labor [21]. Automation can be introduced gradually, starting with more affordable and easy-to-implement technologies [22].

Additionally, the use of drones for aerial monitoring can provide real-time data on crop conditions, detecting pests and diseases before they spread through aerial images as shown in Fig. 2. This facilitates a quicker response and saves resources by avoiding excessive application of pesticides [28].

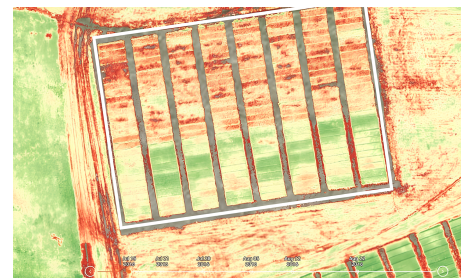


Fig. 2: Measuring the effectiveness of fungicide treatments in a barley research field [28].

According to [5], the integration of precision agriculture with photogrammetry software allows for a more detailed analysis of spatial and temporal variability in crops, optimizing farm management and increasing sustainability. By utilizing drones for data capture, farmers can analyze spatial and temporal variability, leading to improved yields and sustainability. The paper emphasizes the benefits of advanced technologies like reflectance measurement and orthophoto mapping, which increase accuracy in assessing crop phenotyping and optimizing farm production.

Also, drone data offers valuable insights for assessing and recording crop damage resulting from natural disasters such as floods, fires, and severe weather conditions, as can be seen in Fig. 3, [28].

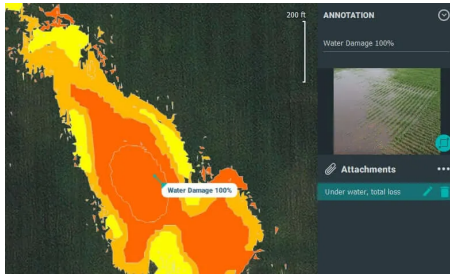


Fig. 3: Water damage map generated using drone data [28].

In [29] a study is presented that shows that soil sensor systems, which monitor moisture and nutrients, enable precise irrigation, saving up to 30% of water and improving crop health. Small-scale farmers who have adopted these technologies reported significant reductions in irrigation costs while increasing productivity. In an applied case study presented by [9], it explores the development of a tomato crop prediction system by leveraging sensor networks, IoT, and AI technologies in precision agriculture. Distributed sensors collected environmental data, which were processed by AI models hosted on a cloud platform. These models, based on recurrent neural networks, were trained and validated using real harvest data. The system demonstrated high accuracy, predicting weekly harvest volumes with an average error of just 3.2%. The integration of sensors and AI proved highly effective, showcasing the potential for broader applications in precision farming.

Figure 4 shows the components of a system developed using FIWARE technology, which utilizes artificial intelligence, complex event processing, and the Internet of Things (IoT) to significantly enhance precision agriculture, reduce resource waste, and other costs such as labor, agricultural inputs, and fertilizers [7].



Fig. 4: Sensors and components of the QUHOMA platform [7].

The authors of [30] emphasizes the role of Internet of Things (IoT) in enhancing irrigation accuracy by integrating sensors and wireless communication devices to collect critical data such as soil moisture, temperature, and humidity. He highlights how IoT-based smart irrigation can address challenges like labor shortages and water wastage, which are particularly relevant in regions affected by extreme weather, such as South Asia. In contrast, in [31] focuses on smart irrigation systems within the broader context of maximizing agricultural productivity in India. He proposes a system that integrates IoT with predictive analytics to optimize water usage throughout various stages of plant growth. Ingole's approach includes the use of sensors to monitor

soil moisture, temperature, and humidity, facilitating precise irrigation planning and decision-making to ensure optimal crop growth and water conservation. Both authors underline the importance of IoT in improving irrigation practices, yet Ingole expands the concept by incorporating an agriculture cultivation recommender, aiming for even greater efficiency and yield optimization in precision agriculture.

The book [32] examines the adoption of precision agriculture through advanced farm management software and information systems (FMSIS). It highlights how current FMSISs allows farmers to optimize production with precision and reduced manual labor. Using statistical models, the authors identify factors influencing farmers' preferences for precision technologies, such as farm size and soil texture. They propose a web-based FMSIS to meet the specific needs of precision agriculture, emphasizing the importance of integrating advanced technologies for improved farm management, automation and technical knowledges..

b. Economic Benefits

Automation has the potential to profoundly reshape the economic landscape of small-scale farming by enhancing productivity and lowering production costs, as evidenced in the research [8]. These advancements make small farmers more competitive in both local and global markets, reducing their reliance on manual labor and mitigating risks associated with climate variability and market volatility [23]. As smallholders adopt these technologies, they can improve the standardization and quality of their products, making them more attractive to higher-value markets that demand precision and consistency in agricultural output [33].

The economic benefits of automation in small-scale agriculture are evident in various studies. Gazieva's research [8] presents a thorough cost-benefit analysis, revealing significant efficiency gains, including better disease control and improved economic performance. Compared to traditional methods, automation optimizes resource use and reduces labor requirements, enabling more efficient production management. This creates a pathway for small farmers to improve sustainability and profitability in micropropagation. In a broader context, [33] synthesized data from various sources to highlight the economic and environmental advantages of Digital Agricultural Technologies (DATs). Technologies such as Farm Management Information Systems (FMIS), Variable Rate Technologies (VRT), and Robotic Systems or Smart Machines (RSSM) demonstrate significant reductions in fertilizer and pesticide use (up to 80%) and substantial yield increases (up to 62%). The use of RSSM reduced labor by 97% and diesel consumption by 50%, while FMIS contributed to yield improvements of 10%-15%, showing how integrated digital solutions enhance both efficiency and sustainability.

Automation also opens new revenue streams and cost-saving opportunities. Technologies like automated harvesters and smart irrigation systems help reduce dependence on seasonal labor, a major expense for small farms [20, 30, 31]. These technologies ensure optimal harvest timing through data-driven decisions, minimizing post-harvest losses [32]. For instance, Kakarla [34] demonstrates how automation is revolutionizing the harvesting process for leafy

greens. Traditionally labor-intensive, the manual cutting and bundling of produce has been replaced by automated systems that use robotics and sensors, improving efficiency, reducing labor dependency, and ensuring consistent quality. This approach not only increases productivity but also lowers operational costs, offering a sustainable solution for modern agriculture. In addition, automation can streamline farm operations, enabling farmers to diversify their production and optimize resources across different types of crops [20].

By reducing input costs such as water, fertilizers, and pesticides while increasing yields, automation enhances profitability for small-scale farmers [8]. Precision farming tools allow farmers to minimize costs while boosting yields, improving economic resilience [7, 33]. This makes farmers less vulnerable to market fluctuations and climate variability, giving them greater control over their production outcomes [24]. However, as Zolkin [35] raises important concerns, highlighting that not all sectors experience uniform benefits from automation. His analysis shows that highly mechanized agro-enterprises still face substantial losses, especially in the production of fruits, vegetables, and root crops, with losses reaching up to 79%. Similarly, cereal production suffers losses during processing and consumption. These findings illustrate that while automation offers significant advantages, certain sectors still face challenges, underscoring the need for ongoing research, adaptation, and investment in technology tailored to specific regional and sectoral needs.

Access to advanced technologies allows small farmers to scale their operations more efficiently, making family farming a more sustainable and financially viable livelihood. As demonstrated by these studies, automation fosters rural development, reduces poverty, and enhances food security across various regions.

c. Environmental Sustainability

Automation also plays a crucial role in enhancing environmental sustainability in agriculture [33]. The adoption of precision farming technologies allows for more efficient use of natural resources, reducing waste and minimizing the ecological footprint of farming practices [5]. Technologies such as automated irrigation systems, soil sensors, and drones for monitoring and input application enable farmers to make real-time adjustments based on precise environmental conditions [9, 29]. Raihan's review [36] further supports this by highlighting recent advancements in digital agriculture that not only improve resource efficiency but also aim to mitigate climate risks and alleviate food insecurity. By adopting digital tools and management techniques, farmers can reduce greenhouse gas emissions, enhance productivity, and contribute to more sustainable food systems, all while addressing the growing challenges posed by climate change.

One of the key environmental benefits of automation is the reduction in water consumption through smart irrigation systems [31]. These systems monitor soil moisture and weather conditions, ensuring that crops receive the right amount of water at the right time [30]. The adoption of micro-irrigation significantly reduces soil water evaporation and enhances Water-Use Efficiency (WUE) with lower water consumption. In cotton production, micro-irrigation decreased water usage by 37%, though it led to a 21%

reduction in yield. Conversely, in wheat cultivation, it reduced water use by 23% while increasing output by 37%. Micro-irrigation systems effectively minimize evaporation, particularly from the plant canopy, and limit soil water loss between plant rows during early growth stages. These findings demonstrate that WUE can be significantly improved through strategic water management, providing a positive impact on irrigated crop performance [37]. Automated irrigation systems are essential for optimizing water use in agriculture, as Goyal's research highlights [38]. These systems utilize crop, soil, and weather data to ensure precise water management, reducing waste and improving productivity in a process as shown in Fig. 5 [39]. The integration of machine learning further enhances water distribution efficiency, adapting to real-time conditions. Overall, these technologies are key to modern precision farming, conserving water while increasing crop yields [38].

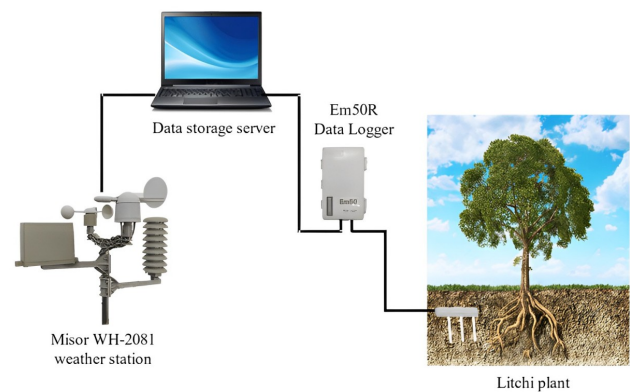


Fig. 5: A weather station and soil moisture sensor to collect data [39].

A weather station and soil moisture sensor to collect data

As previously stated in the section Technological Opportunities, automation can also contribute to mitigating the effects of climate change by increasing farm resilience. Automated systems enable farmers to respond more quickly to extreme weather events such as droughts, floods, or heatwaves [24]. For instance, sensors that monitor soil conditions and plant health can provide early warnings, allowing for timely interventions to protect crops from environmental stressors [23]. Automated systems in agriculture play a critical role in mitigating the impacts of climate change, as highlighted by the studies of Potts [24] and Sarker [23]. Potts, [24] emphasizes how automation in energy systems can enhance resilience to extreme weather events by improving infrastructure management and decision-making processes, thereby minimizing damage during crises like Winter Storm Uri. In [23], on the other hand, illustrates the relationship between climate variation, agricultural production, and migration, showing that automation, such as precision irrigation, helps farmers better adapt to changing weather conditions, optimizing resources and reducing risks. Thus, automation not only boosts agricultural productivity but also contributes to climate resilience. In the long term, these technologies promote sustainable farming practices that help conserve

soil health, biodiversity, and ecosystems, all of which are essential for maintaining agricultural productivity in a changing climate.

Furthermore, by optimizing resource use and reducing waste, automation aligns with global sustainability goals, such as reducing greenhouse gas emissions and conserving natural habitats [40]. The integration of these technologies into small-scale farming contributes to the development of an agriculture sector that is not only more efficient but also more environmentally responsible. The book [41] highlights how IoT and AI are transforming agriculture by enabling real-time, data-driven decision-making. These technologies help optimize resources and improve efficiency in both crop and livestock management. For small-scale farming, their integration not only boosts productivity but also enhances environmental sustainability by reducing water and chemical use. This shift supports the development of a more efficient and ecologically responsible agricultural sector, essential for ensuring a sustainable future in food production.

IV. CONCLUSION

The adoption of automation technologies in small-scale agriculture presents substantial opportunities for enhancing productivity, sustainability, and resource efficiency. However, challenges related to the high costs of implementation, lack of adequate infrastructure, and the need for technical training persist. For automation technologies to be successfully adopted by smallholders, they must be adapted to the realities of small-scale farming, offering affordable and context-appropriate solutions tailored to the financial and operational capacities of these farmers.

Additionally, the role of public policies and support programs cannot be understated. Government initiatives aimed at providing access to credit, subsidies for technology acquisition, and technical training programs are critical to lowering the entry barriers for smallholders. By addressing these challenges, small-scale farmers will be better positioned to leverage the benefits of automation, ultimately contributing to more resilient, competitive, and sustainable agricultural practices.

Further research and innovation should focus on developing low-cost automation solutions and scalable models that align with the diverse needs of smallholders in different regions. Collaborative efforts between governments, private sectors, and research institutions will be essential to ensure that small-scale farmers can fully participate in and benefit from the ongoing technological advancements in agriculture.

a. Future Work

The adoption of automation technologies in small-scale farming presents numerous opportunities for further research and development. One promising avenue is the exploration of cost-effective automation systems specifically designed for smallholders. This includes the development of low-cost IoT devices, drones, and smart irrigation systems tailored to the financial constraints and operational scale of small farms.

Additionally, future studies could investigate the socio-economic impacts of automation adoption, particularly in terms of labor dynamics, rural development, and gender

roles in agriculture. Understanding how automation affects employment, skills requirements, and income distribution among smallholder communities is essential for ensuring that these technologies contribute to inclusive growth.

Another key area for future research involves the environmental implications of widespread automation adoption. While existing studies suggest that automation can improve resource efficiency and reduce environmental impact, long-term assessments are needed to determine how these technologies affect biodiversity, soil health, and water resources over time.

Finally, policy-oriented research is crucial to explore how governments and institutions can create supportive frameworks that encourage small-scale farmers to adopt automation. Investigating financial incentives, infrastructure investments, and education programs would help address current barriers and maximize the potential benefits of automation for small-scale agriculture.

V. ACKNOWLEDGMENT

The authors would like to thank the Federal University of Tocantins, their Electronic Laboratory and the Tutorial Education Program-PET of the Electrical Engineering for their collaboration in carrying out this research.

REFERENCES

- [1] M. Escribà-Gelonch, S. Liang, P. van Schalkwyk, I. Fisk, N. V. D. Long, and V. Hessel, "Digital twins in agriculture: Orchestration and applications," *Journal of Agricultural and Food Chemistry*, vol. 72, no. 19, pp. 10 737–10 752, 2024.
- [2] R. Karothia and M. K. Chattopadhyay, "Review of various technologies involved in precision farming automation," in *Precision Agriculture for Sustainability*. Apple Academic Press, 2024, pp. 3–20.
- [3] A. O. Adewusi, O. F. Asuzu, T. Olorunsogo, C. Iwuanyanwu, E. Adaga, and D. O. Daraojimba, "Ai in precision agriculture: A review of technologies for sustainable farming practices," *World Journal of Advanced Research and Reviews*, vol. 21, no. 1, pp. 2276–2285, 2024.
- [4] F. Fuentes-Peñailillo, K. Gutter, R. Vega, and G. C. Silva, "Transformative technologies in digital agriculture: Leveraging internet of things, remote sensing, and artificial intelligence for smart crop management," *Journal of Sensor and Actuator Networks*, vol. 13, no. 4, p. 39, 2024.
- [5] I. Ioja, V. Nedeff, M. Agop, F. M. Nedeff, and C. Tomozei, "Software uses in precision agriculture based on drone image processing: A review," in *2024 9th International Conference on Energy Efficiency and Agricultural Engineering (EE&AE)*. IEEE, 2024, pp. 1–6.
- [6] V. Nandeshwar, I. Borade, A. Borade, A. Bonde, T. Bora, O. Bobade, and V. Bokare, "Farm automation using nodemcu," in *Advances in AI for Biomedical Instrumentation, Electronics and Computing*. CRC Press, 2024, pp. 560–566.
- [7] K. Bazargani and T. Deemyad, "Automation's impact on agriculture: opportunities, challenges, and economic effects," *Robotics*, vol. 13, no. 2, p. 33, 2024.
- [8] L. Gazieva, "Cost-benefit analysis of implementing automated systems in agricultural micropropagation," in *BIO Web of Conferences*, vol. 84. EDP Sciences, 2024, p. 01025.
- [9] M. Á. G. Pérez, A. G. González, F. J. C. Rodríguez, I. M. M. Leon, F. A. L. Abrisqueta *et al.*, "Precision agriculture 4.0: Implementation of iot, ai, and sensor networks for tomato crop prediction," *Buletin Ilmiah Sarjana Teknik Elektro*, vol. 6, no. 2, pp. 172–181, 2024.
- [10] S. Ghazal, A. Munir, and W. S. Qureshi, "Computer vision in smart agriculture and precision farming: Techniques and applications," *Artificial Intelligence in Agriculture*, 2024.

- [11] Q. Lu, C. Liao, M. Chen, V. Shi, X. Hu, and W. Hu, "Platform financing or bank financing in agricultural supply chains: The impact of platform digital empowerment," *European Journal of Operational Research*, vol. 315, no. 3, pp. 952–964, 2024.
- [12] I. Ochilov, "Issues of improving financing of agricultural clusters," in *BIO Web of Conferences*, vol. 82. EDP Sciences, 2024, p. 02035.
- [13] C. He, C. Zhou, and H. Wen, "Improving the consumer welfare of rural residents through public support policies: A study on old revolutionary areas in china," *Socio-Economic Planning Sciences*, vol. 91, p. 101767, 2024.
- [14] J. Nasir, "Constraints on rural growth," in *Development Challenges of Pakistan: Constraints and Choices*. Springer, 2024, pp. 235–256.
- [15] V. Sahni and M. Singh, "Smart indian agriculture farm using an iot dashboard," *Educational Administration: Theory and Practice*, vol. 30, no. 5, pp. 2872–2888, 2024.
- [16] T. Raji, A. Allee, and N. Williams, "Productive use of electricity in agriculture in sub-saharan africa: A systematic review," *Progress in Energy*, 2024.
- [17] S. V. Moreira, N. R. Del Bianco, and C. F. Martins, "Small internet providers as agents: Internalizing digital infrastructure in brazil," in *Geo Spaces of Communication Research*. Emerald Publishing Limited, 2024, pp. 9–29.
- [18] M. R. Pugu, L. Judijanto, and J. Bungai, "Use of renewable energy in rural areas: Literature review and application in community service," *JOURNAL OF COMMUNITY DEDICATION*, vol. 4, no. 2, pp. 441–454, 2024.
- [19] V. Shinde and S. Wandre, "Solar photovoltaic water pumping system for irrigation: A review," *African journal of agricultural research*, vol. 10, no. 22, pp. 2267–2273, 2015.
- [20] R. Raj, S. Kumar, S. P. Lal, H. Singh, J. Pradhan, and Y. Bhardwaj, "A brief overview of technologies in automated agriculture: Shaping the farms of tomorrow," *International Journal of Environment and Climate Change*, vol. 14, no. 7, pp. 181–209, 2024.
- [21] Y. B. KRAJA, E. Bejleri, and P. SARAÇI, "Examining the relationship between technical competencies, personal skills, and employment opportunities. case of albania," *International Journal of Religion*, vol. 5, no. 8, pp. 161–171, 2024.
- [22] M. G. Farouque, K. H. Kabir, D. Roy, M. M. Rana, and J. L. Donaldson, "An assessment of capacity building activities for common interest farmer groups in bangladesh," *Development in Practice*, vol. 34, no. 4, pp. 491–511, 2024.
- [23] B. Sarker, M. Alam, and M. J. Uddin, "Relationship among weather variation, agricultural production, and migration: A systematic methodological review," *Health Science Reports*, vol. 7, no. 4, p. e2002, 2024.
- [24] J. Potts, H. R. Tiedmann, K. K. Stephens, K. M. Faust, and S. Castellanos, "Enhancing power system resilience to extreme weather events: A qualitative assessment of winter storm uri," *International Journal of Disaster Risk Reduction*, vol. 103, p. 104309, 2024.
- [25] D. Borah, J. Rout, and T. Nooruddin, "Traditional knowledge-based sustainable agriculture in the eastern himalayas in india," in *Addressing the Climate Crisis in the Indian Himalayas: Can Traditional Ecological Knowledge Help?* Springer, 2024, pp. 95–125.
- [26] X. Hu, B. Xiao, and Z. Tong, "Technological integration and obstacles in china's agricultural extension systems: A study on disembeddedness and adaptation," *Sustainability*, vol. 16, no. 2, p. 859, 2024.
- [27] L. Manning, "Innovating in an uncertain world: Understanding the social, technical and systemic barriers to farmers adopting new technologies," *Challenges*, vol. 15, no. 2, p. 32, 2024.
- [28] Wingtra. (2024) Drone mapping applications: Use of drones in agriculture. Applications Agriculture <https://wingtra.com/drone-mapping-applications/use-of-drones-in-agriculture/> (10/10/2024).
- [29] B. P. d. R. LEÃO, "Implementação de sensores de baixo custo para otimização da gestão de recursos naturais na agricultura de pequena escala: um estudo de viabilidade e impacto," 2024.
- [30] M. S. Hossain, U. K. Ghosh, M. N. Islam, and M. A. R. Khan, "Precision agriculture practices for smart irrigation," in *Remote Sensing in Precision Agriculture*. Elsevier, 2024, pp. 175–188.
- [31] K. Ingoles and D. Padole, "An internet of things (iot)-based smart irrigation and crop suggestion platform for enhanced precision agriculture," *Journal of Information and Optimization Sciences*, vol. 45, no. 4, pp. 873–883, 2024.
- [32] S. Pramanik, S. Roy, and R. Bose, *Data Driven Mathematical Modeling in Agriculture: Tools and Technologies*. CRC Press, 2024.
- [33] G. Papadopoulos, S. Arduini, H. Uyar, V. Psiroukis, A. Kasimati, and S. Fountas, "Economic and environmental benefits of digital agricultural technologies in crop production: A review," *Smart Agricultural Technology*, p. 100441, 2024.
- [34] U. S. Kakarla, A. Venkataramana, S. Gopinath, K. Shivaram, K. Yugendhar, K. Harika, and S. Dixit, "Revolutionizing agricultural harvesting with iot application," in *MATEC Web of Conferences*, vol. 392. EDP Sciences, 2024, p. 01014.
- [35] A. Zolkin, E. Matvienko, O. Shevchenko, A. Bityutskiy, and V. Ozhiganov, "Digital technologies and automation in breeding and agriculture," in *E3S Web of Conferences*, vol. 474. EDP Sciences, 2024, p. 01017.
- [36] A. Raihan, "A review of digital agriculture toward food security and environmental sustainability," in *Proceedings of The International Conference on Digital Agriculture, Food Security, and Environmental Sustainability*, 2024.
- [37] S. Alharbi, A. Felemban, A. Abdelrahim, and M. Al-Dakhil, "Agricultural and technology-based strategies to improve water-use efficiency in arid and semiarid areas," *Water*, vol. 16, no. 13, p. 1842, 2024.
- [38] R. Goyal, A. Nath, U. Niranjana, and R. Niyogi, "Analyzing monitoring and controlling techniques for water optimization used in precision irrigation," in *International Conference on Advanced Information Networking and Applications*. Springer, 2024, pp. 157–168.
- [39] P. Gao, H. Qiu, Y. Lan, W. Wang, W. Chen, X. Han, and J. Lu, "Modeling for the prediction of soil moisture in litchi orchard with deep long short-term memory," *Agriculture*, vol. 12, no. 1, p. 25, 2021.
- [40] R. C. Bunkar, K. Ananda, K. Kumar, B. Sadangi, B. S. Gupta, D. Saikant, and A. Shukla, "A review on leveraging technology for sustainable development in agricultural extension program," *Archives of Current Research International*, vol. 24, no. 5, pp. 543–556, 2024.
- [41] S. Gupta, W. Hasan, S. Singh, D. Kumar, M. J. Ansari, and S. Nisar, *Agriculture 4.0: Smart Farming with IoT and Artificial Intelligence*. CRC Press, 2024.

Descrição formal das funções de ativação de modelos de aprendizado de máquina

Formal description of activation functions of machine learning models

Julia Assunção Leal¹, Hellena Christina Fernandes Apolinário¹ e Rogério Azevedo Rocha¹

¹ Curso Bacharelado em Ciência da Computação, Universidade Federal do Tocantins (UFT), Palmas/TO

Data de recebimento do manuscrito: 29/11/2024

Data de aceitação do manuscrito: 25/02/2025

Data de publicação: 13/03/2025

Resumo— Modelos de inteligência artificial são cada vez mais comuns em vários aspectos do dia a dia. Não só nos casos mais emblemáticos, mas também para casos corriqueiros como em sistemas de recomendação em sites de compras. Nesse sentido, é muito importante o entendimento de como esses modelos funcionam por parte dos desenvolvedores. Contudo, o uso massivo de bibliotecas para utilização desses modelos pode desfavorecer esse entendimento. Assim, esse trabalho traz a definição e demonstração formal das funções de ativação em modelos de aprendizado de máquina. Esse é um ponto fundamental para a introdução do assunto à novos desenvolvedores e cientistas que trabalharão na área. A descrição formal se aplica às clássicas funções ReLU, Sigmóide, tangente hiperbólica, softmax e gradiente descendente. Além disso, também são discutidos o impacto dessas funções no modelo LeNet-5 aplicado à base de dados MNIST.

Palavras-chave—Inteligência Artificial, Matemática Aplicada, Funções de Ativação, ReLU, Sigmóide, TanH, Softmax, Método do gradiente

Abstract— Artificial intelligence models are increasingly common in many aspects of everyday life, not only in the most emblematic cases, but also in everyday cases such as recommendation systems on shopping websites. In this sense, developers need to understand how these models work. However, the massive use of libraries to use these models can hinder this understanding. Therefore, this work defines and formally demonstrates activation functions in machine learning models. This is a fundamental point for introducing the subject to new developers, and scientists, who will work in the area. The formal description applies to the classic ReLU, Sigmóide, hyperbolic tangent, softmax and gradient descent functions. In addition, the impact of these functions on the LeNet-5 model applied to the MNIST database is also discussed.

Keywords—Artificial Intelligence, Applied Mathematics, Activation Functions, ReLU, Sigmoid, TanH, Softmax, Gradient Descent

I. INTRODUÇÃO

Com o avanço da tecnologia nos campos da computação e da matemática, o desenvolvimento da Inteligência Artificial (IA) tornou-se mais eficiente e próximo da realidade. A matemática, nesse contexto, configura-se como a base que fundamenta a IA, fornecendo teorias e ferramentas essenciais para o aprendizado das máquinas (ML ou *Machine Learning*). De algoritmos complexos a modelos preditivos, a matemática transforma grandes volumes de dados brutos em informações compreensíveis, contribuindo para o progresso em diversas áreas [1, 2, 3, 4].

Esse artigo foca na base matemática para compreender

o funcionamento de uma IA, com interesse em apoiar pesquisadores iniciantes na área. Entre os conceitos matemáticos indispensáveis para a IA, destacam-se as funções de ativação, fundamentais para a arquitetura das redes neurais. Essas funções transformam a entrada linear de cada neurônio em uma saída não-linear, permitindo que as redes aprendam padrões complexos nos dados. Sem elas, uma rede neural seria apenas um modelo linear, incapaz de realizar tarefas mais elaboradas, como o reconhecimento de fala, o processamento de imagens ou a interpretação de linguagem natural [5]. Assim, o estudo das funções é crucial para tornar os sistemas mais eficientes e funcionais.

Outro conceito relevante é a otimização por gradientes, que ajusta os parâmetros dos modelos de IA. Por meio do cálculo do gradiente (a derivada do erro de uma função em relação aos pesos e vieses) é possível minimizar o erro e aprimorar o aprendizado do modelo. O método do gradiente descendente, por exemplo, permite encontrar o mínimo de uma função de custo de forma iterativa, sendo

essencial no treinamento de redes neurais [6]. Além disso, a probabilidade desempenha um papel vital, possibilitando que modelos lidem com incertezas e façam inferências sobre dados variáveis. Redes bayesianas e processos gaussianos, por exemplo, são amplamente utilizados em áreas como diagnósticos médicos e sistemas de recomendação [7].

Portanto, a matemática não apenas sustenta a Inteligência Artificial, mas também potencializa seu impacto em diversas áreas, como a medicina e a indústria. Para garantir o contínuo avanço nessa área, é fundamental que instituições de ensino ampliem os investimentos em disciplinas que integrem matemática aplicada e computação.

II. EXPLICAÇÃO GERAL SOBRE O FUNCIONAMENTO DE UMA IA

Segundo [1], o desenvolvimento de uma Inteligência Artificial (IA) envolve diversas etapas cruciais. Inicialmente, ocorre a coleta e preparação dos dados, etapa na qual os dados são limpos, normalizados e organizados para serem utilizados no treinamento do modelo. Durante o treinamento, o modelo aplica funções de ativação, que introduzem não linearidade e permitem que ele se adapte a funções mais complexas, mas essas funções não processam os dados diretamente, mas transformam os resultados intermediários gerados por uma combinação linear dos dados e dos pesos do modelo. Em cada neurônio, após a multiplicação dos dados de entrada pelos seus respectivos pesos e a adição de um viés, a função de ativação entra em ação.

Em seguida, o modelo faz previsões com base nos dados de entrada, e essas previsões são comparadas aos resultados esperados. A diferença entre eles é mensurada por meio de uma função de perda, que calcula o erro. Esse erro é então utilizado no processo de retropropagação, onde ele é propagado para trás pelas camadas do modelo, ajustando os pesos dos neurônios para melhorar a precisão. Esse ciclo se repete até que o erro seja minimizado ou atinja um nível aceitável, otimizando o desempenho do sistema.

III. DEFINIÇÃO E PROPRIEDADES DA FUNÇÃO RELU

A função de ativação da Unidade Linear Rectificada (ReLU), amplamente conhecida e adotada em redes neurais profundas e no contexto de aprendizado de máquina [1] [5], foi introduzida por [8]. Ela se destaca pela simplicidade e eficácia na modelagem da ativação de neurônios artificiais.

Para compreender o comportamento dessa função, que é constante em $x = 0$, até o ponto $y = 0$ onde os valores aumentam, tem-se a Figura 1, feita a partir do modelo proposto pelo estudo.

Matematicamente, a função ReLU é definida como:

- **Fórmula:** $f(x) = \max(0, x)$
- **Domínio:** $\mathbb{R}$ (todos os números reais)
- **Imagem:** $[0, \infty)$

ReLU pode ser expresso de diferentes maneiras equivalentes. Formalmente, temos:

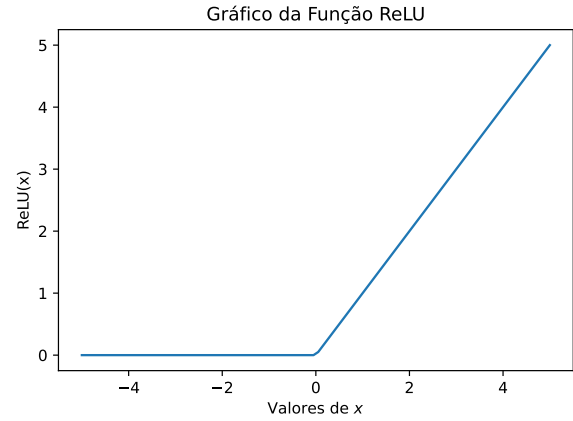


Figura 1: Função ReLU testada com valores do MNIST.

$$f(x) = x^+ = \max(0, x) = \frac{x + |x|}{2}$$

A função atua como uma unidade linear para valores de entrada positivos, enquanto zera os valores de entrada negativos:

$$f(x) = \begin{cases} x & \text{se } x > 0, \\ 0 & \text{se } x \leq 0. \end{cases}$$

a. Derivada da função ReLU:

$$f'(x) = \begin{cases} 1 & \text{se } x > 0, \\ 0 & \text{se } x < 0. \end{cases}$$

Observa-se que em $x = 0$, a derivada não existe, pois:

$$\lim_{h \rightarrow 0^+} \frac{f(a+h) - f(a)}{h} \neq \lim_{h \rightarrow 0^-} \frac{f(a+h) - f(a)}{h}$$

Se a variável é positiva ($x > 0$):

$$f(x) = \frac{x + |x|}{2} = \frac{x + x}{2} = \frac{2x}{2} = x$$

E se a variável é negativa ($x < 0$):

$$f(x) = \frac{x + |x|}{2} = \frac{x - x}{2} = \frac{0}{2} = 0$$

A função de ativação ReLU, devido à sua simplicidade e capacidade de diminuir o problema do desaparecimento do gradiente [1], tornou-se uma escolha normal em redes neurais simples. Este estudo destacou suas propriedades, representação gráfica e aplicação em contextos de aprendizado de máquina.

b. Aplicação de ReLU em redes neurais

A classificação de imagens é uma tarefa fundamental no campo da visão computacional, e o dataset MNIST, composto por dígitos manuscritos, consiste em imagens de 28x28 pixels em tons de cinza, divididas em conjuntos de treinamento e teste. Neste estudo, foi implementada a arquitetura LeNet-5, uma rede neural convolucional clássica, para realizar a classificação dos dígitos do MNIST [9]. O objetivo principal foi comparar o desempenho do modelo

ao utilizar diferentes funções de ativação, analisando sua influência na precisão e na capacidade de generalização do modelo.

A arquitetura LeNet-5 foi escolhida por sua simplicidade em tarefas de classificação de imagens. Ela é composta por camadas convolucionais, seguidas por camadas de pooling para redução de dimensionalidade, e finalizada com camadas totalmente conectadas. A sua última camada utiliza a função Softmax, que é ideal para problemas de classificação multiclasse, pois produz uma distribuição de probabilidade sobre as classes. Para investigar o impacto das funções de ativação, foram testadas três opções: ReLU, Sigmóide e TanH. Cada uma dessas funções de ativação tem características distintas, influenciando a capacidade do modelo de aprender padrões complexos nos dados.

O treinamento foi realizado por 10 epochs (uma passagem completa pelo conjunto de dados de treinamento), com um *batch size* (número de amostras que são processadas pelo modelo antes que os pesos da rede sejam atualizados) 128. Durante o treinamento, o modelo foi avaliado no conjunto de validação, permitindo monitorar seu desempenho e evitar **overfitting** (ajuste excessivo a um conjunto de dados ao ponto de não generalizar para novos dados). Após o treinamento, o modelo foi avaliado no conjunto de teste de dados, onde foram calculadas a perda e a precisão.

A função de ativação ReLU (*Rectified Linear Unit*) destacou-se como a mais eficiente na classificação dos dígitos do dataset MNIST utilizando a arquitetura LeNet-5. Durante o treinamento, a ReLU proporcionou uma convergência rápida e estável, alcançando uma precisão final de 97,76% por cento no conjunto de teste, com uma perda de 0,0706. A característica principal da ReLU é sua capacidade de ativar apenas valores positivos, o que evita o problema de gradientes desaparecendo, comum em outras funções de ativação. Isso permitiu ao modelo capturar padrões complexos nos dados de forma eficaz, resultando em uma generalização robusta.

Além disso, para uma análise mais detalhada, foram geradas matrizes de confusão 2, 4 e 6.

Na Figura 2 tem-se a matriz referente à função ReLU:

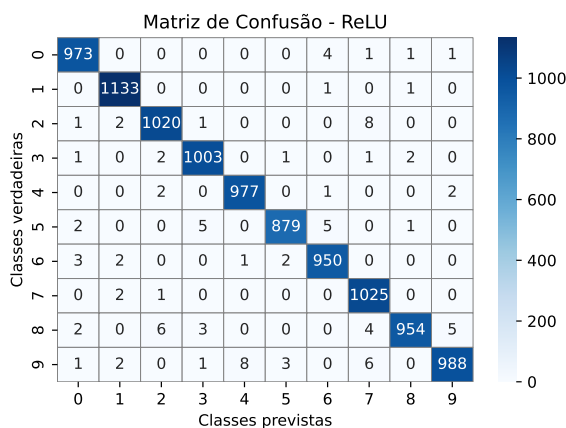


Figura 2: Matriz de confusão testada com valores do MNIST.

IV. DEFINIÇÃO E PROPRIEDADES DA FUNÇÃO SIGMÓIDE

A função de ativação Sigmóide é comumente usada em redes neurais devido às suas propriedades de suavização e transformação não linear [1, 5]. Essa função mapeia a entrada em um intervalo entre 0 e 1, o que é útil para a modelagem de probabilidades [7, 10]. Além disso, a Sigmóide facilita a interpretação dos resultados em tarefas de classificação binária, onde as saídas podem ser interpretadas como probabilidades de pertencer a uma determinada classe [2, 6].

Da mesma forma que a função ReLU, para compreender o comportamento dessa função, a Figura 3 foi criada utilizando-se da inteligência artificial proposta pelo artigo.

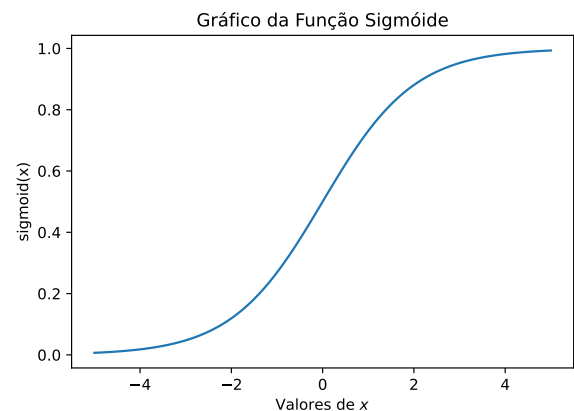


Figura 3: Como a função Sigmóide testada com valores do MNIST.

A expressão matemática para a função Sigmóide é dada por:

- **Fórmula:** $f(x) = \frac{1}{1+e^{-x}}$
- **Domínio:** $\mathbb{R}$
- **Imagem:** $(0, 1)$

a. Derivada da função Sigmóide

Considerando a função Sigmóide $f(x) = \frac{1}{1+e^{-x}}$, a sua derivada em relação a x , é da seguinte forma:

$$f'(x) = \frac{0 \cdot (1 + e^{-x}) - 1 \cdot (1 + e^{-x})'}{(1 + e^{-x})^2} = \frac{-(e^{-x} \cdot (-1))}{(1 + e^{-x})^2} \quad (1)$$

Segue daí que:

$$f'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} \quad (2)$$

Finalmente, expressamos a derivada em termos da própria função Sigmóide:

$$f'(x) = f(x) \cdot (1 - f(x)) \quad (3)$$

b. Aplicação de Sigmóide em redes neurais

A função de ativação Sigmóide destacou-se como a menos eficiente na classificação dos dígitos do dataset MNIST

utilizando a arquitetura LeNet-5. Durante o treinamento, ela alcançou uma precisão final de 96,60% no conjunto de teste, com uma perda de 0.1131. Foi gerada a matriz de confusão para ela na Figura 4, que permite visualizar a distribuição dos erros de classificação entre as diferentes classes.

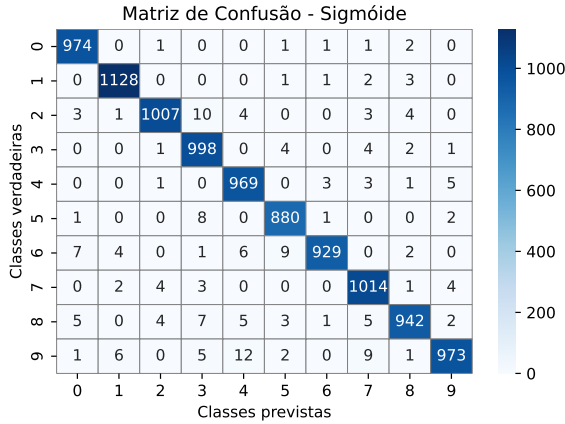


Figura 4: Matriz de confusão testada com valores do MNIST.

V. DEFINIÇÃO E PROPRIEDADES DA FUNÇÃO TANH

A função tangente hiperbólica ($\tanh$) é uma função de ativação comum em redes neurais, definida matematicamente por:

- **Fórmula:** $f(x) = \tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$
- **Domínio:** $\mathbb{R}$
- **Imagem:** $(-1, 1)$

Esta função mapeia os valores de entrada para o intervalo entre -1 e 1, o que é bom para a modelagem de características não lineares em modelos de aprendizado de máquina [1], [5].

Da mesma forma das outras funções, para compreender o comportamento dessa função, a Figura 5 foi criada utilizando-se da inteligência artificial proposta pelo artigo.

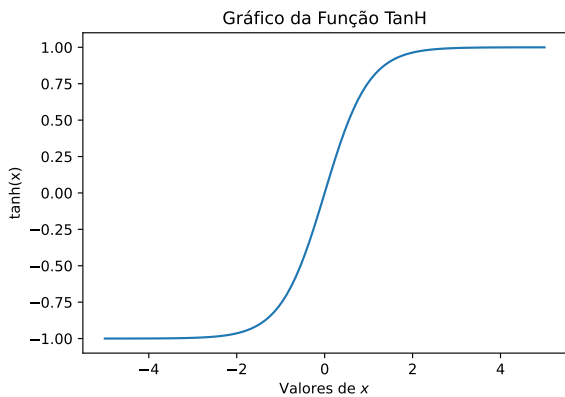


Figura 5: Como a função Tahn testada com valores do MNIST.

a. Derivada da função TanH

A derivada, usando a regra da divisão do quociente, pode ser expandida como:

$$f'(x) = \frac{(e^x + e^{-x})(e^x - e^{-x})' - (e^x - e^{-x})(e^x + e^{-x})'}{(e^x + e^{-x})^2} \quad (1)$$

$$f'(x) = \frac{(e^x + e^{-x})(e^x + e^{-x}) - (e^x - e^{-x})(e^x - e^{-x})}{(e^x + e^{-x})^2} \quad (2)$$

$$f'(x) = \frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{(e^x + e^{-x})^2} \quad (3)$$

E portanto, obtemos que:

$$f'(x) = \frac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{(e^x + e^{-x})^2} = 1 - \tanh^2(x) \quad (4)$$

E assim a derivada de $\tanh(x)$ em relação a x é crucial para o processo de retropropagação em redes neurais e é dada por:

$$f'(x) = 1 - \tanh^2(x) \quad (5)$$

b. Aplicação da função TanH em redes neurais

A função de ativação TanH ficou entre as duas anteriores em termos de eficiência na classificação dos dígitos do dataset MNIST utilizando a arquitetura LeNet-5. Durante o treinamento, ela alcançou uma precisão final de 98,54% no conjunto de teste, com uma perda de 0.0501. Foi gerada a matriz de confusão para ela na Figura 6, que permite visualizar a distribuição dos erros de classificação entre as diferentes classes:

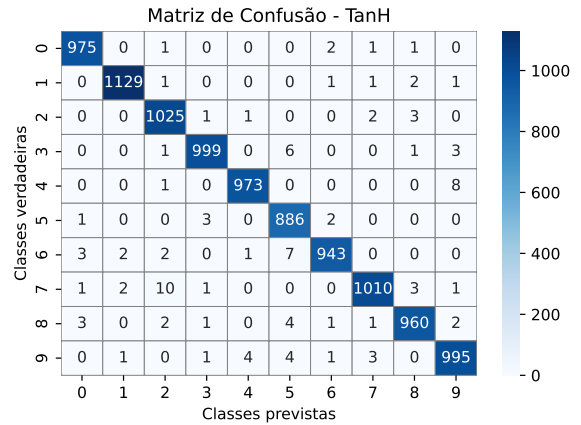


Figura 6: Função Tahn testada com valores do MNIST.

VI. DEFINIÇÃO E PROPRIEDADES DA FUNÇÃO SOFTMAX

A função Softmax também é utilizada para interpretar os resultados como probabilidades, mas não é convencionalmente uma função de ativação, pois é mais utilizada para normalizar as saídas dos neurônios. Isto é especialmente útil em problemas de classificação onde precisamos determinar a probabilidade de uma instância pertencer a uma das várias classes possíveis, ao contrário da função Sigmóide [1, 7]. A função Softmax transforma as pontuações (também conhecidas como logits) produzidas pelos modelos

de aprendizado de máquina em probabilidades que somam 1, facilitando a interpretação dos resultados [1], [5]. Essa propriedade é fundamental para tarefas como classificação multiclasse, onde cada classe é mutuamente exclusiva, permitindo que o modelo selecione a classe mais provável para uma dada entrada [7].

Da mesma forma que a função anterior, para melhor compreender o comportamento dessa função, utilizando os dados do modelo testado, foi feita a Figura 7:

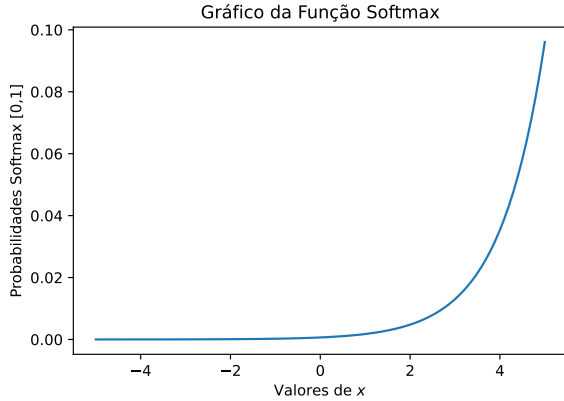


Figura 7: Função Softmax testada com valores do MNIST.

a. Conceito matemático da função SoftMax

A função Softmax é uma função matemática que transforma um vetor de números reais em um vetor de probabilidades, onde a soma total é 1. Ela é utilizada em modelos de classificação multiclasse em aprendizado de máquina, especialmente em redes neurais.

Definição Formal:

Dado um vetor de entrada $\mathbf{z} = (z_1, z_2, \dots, z_K) \in \mathbb{R}^K$, a função Softmax é definida para cada componente z_i como:

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

onde:

- e é denominado número de Euler.
- K é o número de classes ou categorias.

Domínio e Contradomínio:

- **Domínio:** $\mathbb{R}^K$, o espaço de todos os vetores reais de dimensão K .
- **Contradomínio:** Vetores $\mathbf{p} = (p_1, p_2, \dots, p_K)$ tais que $p_i \in (0, 1)$ e $\sum_{i=1}^K p_i = 1$.

Propriedades Importantes:

1. Normalização:

$$\sum_{i=1}^K \text{Softmax}(z_i) = 1$$

Isso significa que a saída da função Softmax forma uma distribuição de probabilidade sobre as classes.

2. Positividade:

$$\text{Softmax}(z_i) > 0 \quad \text{para todo } z_i \in \mathbb{R}$$

3. Monotonicidade Relativa:

A ordem relativa das entradas é preservada nas saídas.

Se $z_i > z_j$, então $\text{Softmax}(z_i) > \text{Softmax}(z_j)$.

b. Derivados da função SoftMax

Vamos considerar duas derivadas principais da função softmax: a derivada em relação ao mesmo elemento ($i = j$) e a derivada em relação a outro elemento ($i \neq j$).

c. Derivada em relação ao mesmo elemento ($i = j$)

Queremos calcular a derivada da função softmax em relação a z_i :

$$\frac{\partial}{\partial z_i} (\text{softmax}(z_i)) = \frac{\partial}{\partial z_i} \left(\frac{e^{z_i}}{\sum_k e^{z_k}} \right) \quad (1)$$

Aplicamos a regra do quociente, onde:

- $f(z_i) = e^{z_i}$
- $g(z_i) = \sum_k e^{z_k}$

As derivadas são:

- $f'(z_i) = e^{z_i}$
- $g'(z_i) = \frac{\partial}{\partial z_i} (\sum_k e^{z_k}) = e^{z_i}$

Aplicando a regra do quociente:

$$\frac{\partial}{\partial z_i} \left(\frac{f(z_i)}{g(z_i)} \right) = \frac{f'(z_i) \cdot g(z_i) - f(z_i) \cdot g'(z_i)}{[g(z_i)]^2} \quad (2)$$

$$= \frac{e^{z_i} \cdot (\sum_k e^{z_k}) - e^{z_i} \cdot e^{z_i}}{(\sum_k e^{z_k})^2} \quad (3)$$

$$= \frac{e^{z_i} \cdot (\sum_k e^{z_k} - e^{z_i})}{(\sum_k e^{z_k})^2} \quad (4)$$

$$= \frac{e^{z_i} (\sum_k e^{z_k} - e^{z_i})}{(\sum_k e^{z_k})^2} \quad (5)$$

$$= \frac{e^{z_i} (\sum_k e^{z_k} - e^{z_i})}{(\sum_k e^{z_k})^2} \quad (6)$$

Reconhecendo que:

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_k e^{z_k}} \quad (7)$$

Portanto:

$$\frac{\partial}{\partial z_i} (\text{softmax}(z_i)) = \text{softmax}(z_i) \cdot \left(\frac{\sum_k e^{z_k} - e^{z_i}}{\sum_k e^{z_k}} \right) \quad (8)$$

$$= \text{softmax}(z_i) \cdot (1 - \text{softmax}(z_i)) \quad (9)$$

d. Derivada em relação a um outro elemento ($i \neq j$)

Agora, calculamos a derivada de $\text{softmax}(z_i)$ em relação a z_j , onde $i \neq j$:

$$\frac{\partial}{\partial z_j} (\text{softmax}(z_i)) = \frac{\partial}{\partial z_j} \left(\frac{e^{z_i}}{\sum_k e^{z_k}} \right) \quad (1)$$

Aplicando novamente a regra do quociente. Neste caso:

- $f(z_j) = e^{z_i}$ (não depende de z_j , logo $f'(z_j) = 0$)
- $g(z_j) = \sum_k e^{z_k}$
- $g'(z_j) = \frac{\partial}{\partial z_j} (\sum_k e^{z_k}) = e^{z_j}$

Aplicando a regra:

$$\frac{\partial}{\partial z_j} \left(\frac{f(z_j)}{g(z_j)} \right) = \frac{f'(z_j) \cdot g(z_j) - f(z_j) \cdot g'(z_j)}{[g(z_j)]^2} \quad (2)$$

$$= \frac{0 \cdot (\sum_k e^{z_k}) - e^{z_i} \cdot e^{z_j}}{(\sum_k e^{z_k})^2} \quad (3)$$

$$= -\frac{e^{z_i} \cdot e^{z_j}}{(\sum_k e^{z_k})^2} \quad (4)$$

Novamente, reconhecendo as expressões de softmax:

$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_k e^{z_k}}$$

$$\text{softmax}(z_j) = \frac{e^{z_j}}{\sum_k e^{z_k}}$$

Portanto:

$$\frac{\partial}{\partial z_j} (\text{softmax}(z_i)) = -\text{softmax}(z_i) \cdot \text{softmax}(z_j) \quad (5)$$

e. Exemplo prático

Seja $\mathbf{z} = (2, 1, 0)$ um vetor do $\mathbb{R}^3$

1. Subtraímos o máximo (2) de cada componente para estabilidade:

$$z_{\max} = 2 \Rightarrow \mathbf{z}' = (0, -1, -2) \quad (1)$$

2. Calculamos as exponenciais:

$$e^0 = 1 \quad e^{-1} \approx 0,3679 \quad e^{-2} \approx 0,1353 \quad (2)$$

3. Calculamos a soma:

$$S = 1 + 0,3679 + 0,1353 \approx 1,5032 \quad (3)$$

4. Calculamos as probabilidades:

$$\text{Softmax}(z_1) = \frac{1}{1,5032} \approx 0,6652 \quad (4)$$

$$\text{Softmax}(z_2) = \frac{0,3679}{1,5032} \approx 0,2447 \quad (5)$$

$$\text{Softmax}(z_3) = \frac{0,1353}{1,5032} \approx 0,0900 \quad (6)$$

As probabilidades resultantes indicam a propensão de cada classe.

VII. FUNÇÃO DE CUSTO QUADRÁTICA (ERRO QUADRÁTICO MÉDIO)

Em Inteligência Artificial, especialmente em aprendizado de máquina e aprendizado profundo, a função de custo (também conhecida como função de perda ou função objetivo) é uma medida que quantifica a diferença entre as previsões do modelo e os valores reais esperados (Goodfellow et al., 2016; Bishop, 2006). Ela serve como um guia para o processo de treinamento do modelo, indicando o quão bem o modelo está performando e como ele pode ser melhorado. O objetivo para melhorar a predição do modelo é minimizar essa função de custo, uma vez que ao fazer isso, a diferença entre o valor previsto pela IA e o valor real se torna mínima.

Existem várias funções de custo, esse artigo abordará a quadrática:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (\text{predição}_i - \text{valor real}_i)^2$$

onde:

- m : é o número de exemplos.
- predição_i : é a previsão do modelo para o i -ésimo exemplo.
- valor real_i : é o valor verdadeiro do i -ésimo exemplo.

Com variáveis:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

onde $h_{\theta}(x)$ é a função do modelo, $x^{(i)}$ são as entradas de treinamento e $y^{(i)}$ são as saídas.

Ao calcular a diferença entre o valor real e o previsto ($h_{\theta}(x^{(i)}) - y^{(i)}$), obtemos o erro para cada exemplo. Elevando esse erro ao quadrado, garantimos que todos os valores sejam positivos, o que evita que erros negativos e positivos se cancelem ao somar. Além disso, o erro ao quadrado penaliza erros maiores de forma mais significativa, incentivando o modelo a minimizar grandes discrepâncias.[1].

O coeficiente $\frac{1}{2}$ é incluído por conveniência matemática. Ao derivar a função de custo durante o processo de otimização (por exemplo, usando gradiente descendente), o expoente 2 do termo quadrático cancela com o $\frac{1}{2}$, simplificando os cálculos das derivadas.

Dividindo a soma total pelo número de exemplos n , obtemos a média dos erros quadráticos. Isso normaliza o custo, permitindo comparações consistentes, independentemente do tamanho do conjunto de dados.

VIII. MÉTODO DOS MÍNIMOS QUADRADOS

O método dos mínimos quadrados é uma técnica estatística utilizada para encontrar a melhor aproximação linear entre um conjunto de dados observados (Montgomery et al., 2012). Por exemplo, pode ser utilizado para minimizar a diferença entre os valores previstos pelo modelo e os valores reais observados, ou seja, o erro do modelo (Draper & Smith, 1998). No caso da IA, o objetivo da utilização do método

dos mínimos quadrados seria minimizar a função de custo discutida anteriormente.

a. Definição

Dado um conjunto de n pontos (x_i, y_i) , $i = 1, 2, 3, \dots, n$, queremos encontrar os parâmetros a e b para a reta $r(x) = a + bx$ de tal forma que minimize a soma das distâncias verticais (d) ao quadrado entre cada ponto y_i e a reta ajustada. Essa soma das distâncias ao quadrado é representada por:

$$M(a, b) = \sum_{i=1}^n [d]^2 \quad (1)$$

Que pode ser substituído por:

$$M(a, b) = \sum_{i=1}^n [y_i - (a + bx_i)]^2 \quad (2)$$

onde $[y_i - (a + bx_i)]^2$ representa a distância ao quadrado entre o valor observado y_i e o valor estimado $r(x_i)$ na reta.

b. Expansão da função objetivo

Substituímos d_i por $y_i - (a + bx_i)$. Assim, a função $M(a, b)$ é dada por:

$$M(a, b) = \sum_{i=1}^n (y_i - a - bx_i)^2 \quad (3)$$

c. Derivação para encontrar o ponto crítico

Para minimizar $M(a, b)$, precisamos encontrar os valores de a e b que tornam $M(a, b)$ mínimo. Segundo o método, isso é feito tomando as derivadas parciais de $M(a, b)$ em relação a a e b e igualando a zero.

Derivada em Relação a a

$$\frac{\partial M}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i) \quad (4)$$

Igualando a zero para minimizar:

$$\sum_{i=1}^n (y_i - a - bx_i) = 0 \quad (5)$$

Derivada em Relação a b

$$\frac{\partial M}{\partial b} = -2 \sum_{i=1}^n x_i (y_i - a - bx_i) \quad (6)$$

Igualando a zero para minimizar:

$$\sum_{i=1}^n x_i (y_i - a - bx_i) = 0 \quad (7)$$

Essas duas equações formam um sistema de equações lineares nos parâmetros a e b . Podemos resolver esse sistema para encontrar os valores de a e b que minimizam $M(a, b)$.

A minimização se dá ao derivar $S(a, b)$ em relação a a e b utilizando a regra da cadeia e, então, igualar a zero:

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial x} \cdot \frac{\partial x}{\partial a} \quad (8)$$

$$\frac{\partial S}{\partial x} = 2 \sum_{i=1}^n (y_i - a - bx_i) \cdot (-1) \quad (9)$$

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i) = 0 \quad (10)$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^n x_i (y_i - a - bx_i) = 0 \quad (11)$$

Distribuindo e dividindo a primeira expressão por $2n$ temos:

$$-2 \sum_{i=1}^n y_i + 2 \sum_{i=1}^n a + 2 \sum_{i=1}^n bx_i = 0 \quad (12)$$

$$\frac{-2 \sum_{i=1}^n y_i}{2n} + \frac{2 \sum_{i=1}^n a}{2n} + \frac{2 \sum_{i=1}^n bx_i}{2n} = \frac{0}{2n} \quad (13)$$

$$-\frac{\sum_{i=1}^n y_i}{n} + a + b \frac{\sum_{i=1}^n x_i}{n} = 0 \quad (14)$$

$$-\bar{y} + a + b\bar{x} = 0 \quad (15)$$

$$\bar{y} - b\bar{x} = a \quad (16)$$

onde $\bar{y}$ é a média amostral de y e $\bar{x}$ é a média amostral de x .

Substituindo esse resultado na segunda expressão, temos:

$$-2 \sum_{i=1}^n x_i (y_i - \bar{y} + b\bar{x} - bx_i) = 0 \quad (17)$$

$$\sum_{i=1}^n [x_i (y_i - \bar{y}) + x_i b (\bar{x} - x_i)] = 0 \quad (18)$$

$$\sum_{i=1}^n x_i (y_i - \bar{y}) + b \sum_{i=1}^n x_i (\bar{x} - x_i) = 0 \quad (19)$$

$$\frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{\sum_{i=1}^n x_i (\bar{x} - x_i)} = b \quad (20)$$

Existe uma fórmula diferente que gera o mesmo resultado:

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (21)$$

IX. GRADIENTE DESCENDENTE

Sobre o método do gradiente descendente, é uma técnica de otimização utilizada em aprendizado de máquina e estatística para minimizar funções (Goodfellow et al., 2016; Bishop, 2006). A ideia é ir na direção inversa da apontada pelo gradiente, uma vez que essa aponta para o máximo da função (Haykin, 2009). Vamos detalhar esse método em termos matemáticos, que foi o tema inicial do estudo. Será ilustrado o método, e colocado um exemplo.

a. Inicialização

Comece com valores iniciais para os parâmetros $\theta = [\theta_1, \theta_2, \dots, \theta_n]$.

b. Função de custo

Utilizando a função de custo:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m \left(h_{\theta}(x^{(i)}) - y^{(i)} \right)^2$$

onde $h_{\theta}(x)$ é a função do modelo, $x^{(i)}$ são as entradas de treinamento e $y^{(i)}$ são as saídas.

c. Cálculo do gradiente

O gradiente da função de custo em relação aos parâmetros θ é um vetor de derivadas parciais:

$$\nabla J(\theta) = \left[\frac{\partial J(\theta)}{\partial \theta_1}, \frac{\partial J(\theta)}{\partial \theta_2}, \dots, \frac{\partial J(\theta)}{\partial \theta_n} \right] \quad (1)$$

Cada componente do gradiente é calculado como:

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^m \left(h_{\theta}(x^{(i)}) - y^{(i)} \right) x_j^{(i)} \quad (2)$$

d. Atualização dos parâmetros

Atualize os parâmetros na direção oposta ao gradiente:

$$\theta_j := \theta_j - \alpha \frac{\partial J(\theta)}{\partial \theta_j} \quad (3)$$

onde α é a taxa de aprendizado.

e. Repetição

Repita os passos 2 e 3 até que a função de custo convirja, ou seja, até que as mudanças nos valores de θ sejam menores que um determinado limiar.

X. EXEMPLO DE GRADIENTE DESCENDENTE

Para ilustrar o método do gradiente descendente, vamos aplicar os conceitos apresentados a um exemplo prático de regressão linear simples. Suponha que desejamos ajustar uma reta a um conjunto de dados para prever valores futuros.

a. Dados de exemplo

Considere o seguinte conjunto de dados:

$$\begin{array}{c|c} x & y \\ \hline 1 & 2 \\ 2 & 4 \\ 3 & 6 \\ 4 & 8 \end{array} \quad (1)$$

Nosso objetivo é encontrar os parâmetros θ_0 e θ_1 da reta $h_{\theta}(x) = \theta_0 + \theta_1 x$ que melhor se ajustam aos dados.

b. Inicialização

Inicializamos os parâmetros com valores arbitrários:

$$\theta_0 = 0, \quad \theta_1 = 0 \quad (2)$$

c. Uso da função de custo

Utilizamos a função de custo de mínimos quadrados:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m \left(h_{\theta}(x^{(i)}) - y^{(i)} \right)^2 \quad (3)$$

d. Cálculo do gradiente

Calculamos as derivadas parciais da função de custo em relação a θ_0 e θ_1 :

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{m} \sum_{i=1}^m \left(h_{\theta}(x^{(i)}) - y^{(i)} \right) \quad (4)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{m} \sum_{i=1}^m \left(\left(h_{\theta}(x^{(i)}) - y^{(i)} \right) x^{(i)} \right) \quad (5)$$

e. Atualização dos parâmetros

Atualizamos os parâmetros na direção oposta ao gradiente:

$$\theta_0 := \theta_0 - \alpha \frac{\partial J}{\partial \theta_0} \quad (6)$$

$$\theta_1 := \theta_1 - \alpha \frac{\partial J}{\partial \theta_1} \quad (7)$$

Escolhemos uma taxa de aprendizado $\alpha = 0,01$.

f. Iterações do gradiente descendente

Realizaremos três iterações para observar o ajuste dos parâmetros.

Iteração 1

Passo 1: Calculamos as previsões iniciais:

$$h_{\theta}(x^{(i)}) = \theta_0 + \theta_1 x^{(i)} = 0 \quad (8)$$

Passo 2: Calculamos os erros:

i	$h_{\theta}(x^{(i)})$	Erro $= h_{\theta}(x^{(i)}) - y^{(i)}$
1	0	-2
2	0	-4
3	0	-6
4	0	-8

(9)

Passo 3: Calculamos as derivadas:

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{4} (-2 - 4 - 6 - 8) = -5 \quad (10)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{4} ((-2)(1) + (-4)(2) + (-6)(3) + (-8)(4)) = -15 \quad (11)$$

Passo 4: Atualizamos os parâmetros:

$$\theta_0 = 0 - 0,01 \times (-5) = 0,05 \quad (12)$$

$$\theta_1 = 0 - 0,01 \times (-15) = 0,15 \quad (13)$$

1. Iteração 2

Passo 1: Novas previsões:

$$h_{\theta}(x^{(i)}) = 0,05 + 0,15x^{(i)} \quad (14)$$

Passo 2: Calculamos os erros:

i	$h_{\theta}(x^{(i)})$	Erro
1	0,20	-1,80
2	0,35	-3,65
3	0,50	-5,50
4	0,65	-7,35

(15)

Passo 3: Calculamos as derivadas:

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{4}(-1,80 - 3,65 - 5,50 - 7,35) \quad (16)$$

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{4}(-18,30) \quad (17)$$

$$\frac{\partial J}{\partial \theta_0} = -4,575 \quad (18)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{4}(-1,80 + (-3,65)(2) + (-5,50)(3) + (-7,35)(4)) \quad (19)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{4}(-1,80 - 7,30 - 16,50 - 29,40) \quad (20)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{4}(-55,00) \quad (21)$$

$$\frac{\partial J}{\partial \theta_1} = -13,75 \quad (22)$$

Passo 4: Atualizamos os parâmetros:

$$\theta_0 = 0,05 - 0,01 \times (-4,575) = 0,09575 \quad (23)$$

$$\theta_1 = 0,15 - 0,01 \times (-13,75) = 0,2875 \quad (24)$$

2. Iteração 3

Passo 1: Novas previsões:

$$h_{\theta}(x^{(i)}) = 0,09575 + 0,2875x^{(i)} \quad (25)$$

Passo 2: Calculamos os erros:

i	$h_{\theta}(x^{(i)})$	Erro
1	0,38325	-1,61675
2	0,67075	-3,32925
3	0,95825	-5,04175
4	1,24575	-6,75425

(26)

Passo 3: Calculamos as derivadas:

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{4}(-1,61675 - 3,32925 - 5,04175 - 6,75425) \quad (27)$$

$$= -4,1855 \quad (28)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{4}((-1,61675)(1) + (-3,32925)(2) \quad (29)$$

$$+ (-5,04175)(3) + (-6,75425)(4)) \quad (30)$$

$$= -12,604375 \quad (31)$$

Passo 4: Atualizamos os parâmetros:

$$\theta_0 = 0,09575 - 0,01 \times (-4,1855) = 0,137605 \quad (32)$$

$$\theta_1 = 0,2875 - 0,01 \times (-12,604375) = 0,41354375 \quad (33)$$

g. Convergência

Continuando esse processo por várias iterações, os parâmetros θ_0 e θ_1 convergirão para os valores que minimizam a função de custo. No nosso exemplo, como os dados seguem uma relação linear perfeita $y = 2x$, os valores ideais são $\theta_0 = 0$ e $\theta_1 = 2$.

h. Visualização

Utilizando duas classes de dígitos selecionadas do conjunto de dados MNIST (neste caso, os números 2 e 7, a título de exemplo), foi gerado um gráfico que apresenta uma superfície tridimensional, e ele representa a função de custo em relação aos dois primeiros parâmetros do modelo de regressão logística.

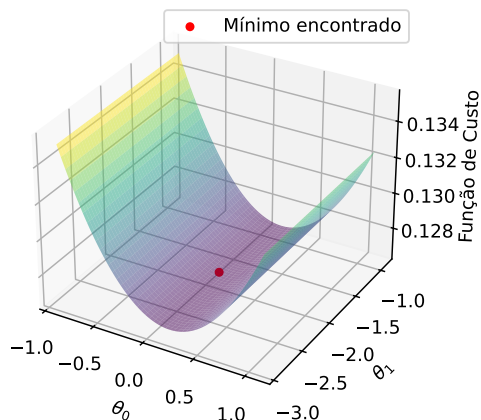
A superfície exibe a variação da função de custo conforme os valores dos parâmetros (Theta 0 e Theta 1) mudam, permitindo uma análise visual do comportamento da otimização. A coloração no gráfico, obtida a partir do mapa de cores viridis, destaca as regiões de maior e menor custo, onde áreas mais elevadas indicam valores de erro maiores, enquanto áreas mais baixas representam regiões de menor erro.

Além disso, um ponto vermelho destacado no gráfico representa o mínimo encontrado pela descida do gradiente, evidenciando a convergência do algoritmo para a melhor solução dentro do espaço de parâmetros. Esse ponto é acompanhado por um contorno preto que enfatiza sua posição na superfície. Dessa forma, o gráfico auxilia na compreensão do processo de ajuste dos pesos do modelo, demonstrando como a otimização ocorre para minimizar a função de custo e melhorar a capacidade preditiva da regressão logística.

XI. CONSIDERAÇÕES FINAIS

Este estudo evidenciou, de maneira detalhada, o papel central da matemática no desenvolvimento da Inteligência Artificial (IA), ressaltando conceitos fundamentais como funções de ativação, o método do gradiente descendente e os mínimos quadrados. A análise destacou que funções de ativação, como ReLU, Sigmóide, TanH e Softmax, são determinantes para a introdução de não-linearidades nas redes neurais. Essa característica permite que essas estruturas matemáticas reconheçam padrões complexos e ofereçam soluções sofisticadas para uma ampla variedade de problemas. Este

Função de Custo para os números 2 e 7 do MNIST

**Figura 8:** Plano do ajuste dos dados utilizando gradiente descendente

estudo evidenciou, de maneira detalhada, o papel central da matemática no desenvolvimento da Inteligência Artificial (IA), ressaltando conceitos fundamentais como funções de ativação, o método do gradiente descendente e os mínimos quadrados. A análise destacou que funções de ativação, como ReLU, Sigmóide, TanH e Softmax, são determinantes para a introdução de não-linearidades nas redes neurais. Essa característica permite que essas estruturas matemáticas reconheçam padrões complexos e ofereçam soluções sofisticadas para uma ampla variedade de problemas. A otimização baseada em gradientes também foi enfatizada como uma das principais técnicas de treinamento de redes neurais, demonstrando como o cálculo do gradiente da função de custo viabiliza ajustes iterativos precisos nos pesos e vieses dos modelos. Esse processo minimiza o erro, acelera a convergência e melhora o desempenho geral dos algoritmos. Além disso, o estudo destacou a importância do método dos mínimos quadrados, ferramenta estatística em modelos preditivos que requerem alta precisão. Embora esses conceitos matemáticos sejam cruciais para o avanço da IA, são raros os artigos que exploram profundamente essa base teórica, frequentemente focando apenas nos resultados práticos ou técnicos. No entanto, estudar mais a fundo essa base matemática oferece vantagens significativas, pois permite a criação de modelos mais robustos, eficientes e inovadores. Assim, ficou evidente que a matemática não apenas melhora o desempenho dos algoritmos, mas também expande o potencial de criação de soluções tecnológicas. Por fim, o estudo reforça que o progresso contínuo da IA depende de uma sólida compreensão matemática e de esforços interdisciplinares. Ao aprofundar o conhecimento nessa área, é possível não apenas elevar o desempenho técnico das redes neurais, mas também maximizar o impacto transformador dessas tecnologias na sociedade contemporânea.

REFERÊNCIAS

- [1] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.
- [2] F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. New York: Spartan Books, 1986.
- [3] N. R. Draper and H. Smith, *Applied Regression Analysis*, 3rd ed. New York: Wiley, 1998.
- [4] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 5th ed. New York: Wiley, 2012.
- [5] S. Haykin, *Neural Networks and Learning Machines*. Pearson Education: Pearson, 2009.
- [6] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [7] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- [8] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010. [Online]. Available: <https://www.cs.toronto.edu/~hinton/absps/reluICML.pdf>
- [9] Y. Le Cun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," in *Proceedings of the 3rd International Conference on Neural Information Processing Systems*, ser. NIPS'89. Cambridge, MA, USA: MIT Press, 1989, p. 396–404.
- [10] M. A. Nielsen, *Neural Networks and Deep Learning*. Available online: Determination Press, 2015. [Online]. Available: <http://neuralnetworksanddeeplearning.com/>



2025

Volume 6 Issue 1

Support