



Ministério da Educação
Secretaria de Educação Profissional e Tecnológica
Instituto Federal Catarinense
Câmpus Videira

CLAUDIONEI LOVATO SERAFINI

**APRIMORANDO A DETECÇÃO DE ANOMALIAS EM
SISTEMAS AGRÍCOLAS INTELIGENTES.**

Videira
2024

CLAUDIONEI LOVATO SERAFINI

**APRIMORANDO A DETECÇÃO DE ANOMALIAS EM
SISTEMAS AGRÍCOLAS INTELIGENTES.**

Trabalho de Conclusão de Curso apresentado ao
Curso de graduação em Ciência da Computação
do Instituto Federal Catarinense – Câmpus Vi-
deira para obtenção do título de bacharel em Ci-
ência da Computação.

Orientador: Angelita Rettore De Araujo Zanella
Coorientador: Rafael Antonio Zanin

Videira
2024

RESUMO

A agricultura desempenha um papel crucial no desenvolvimento sustentável e econômico global, impulsionada significativamente pelos avanços tecnológicos recentes como a Internet das Coisas (IoT) e a agricultura de precisão. Estas inovações têm transformado práticas agrícolas, permitindo uma gestão mais eficiente dos recursos e uma produção mais sustentável. Dentro deste contexto, o CEIFA, um sistema híbrido de detecção de anomalias de baixo custo, foi desenvolvido para melhorar a detecção de anomalias em sistemas agrícolas inteligentes. Apesar de alcançar mais de 98% de precisão, o CEIFA prioriza a detecção de dados anômalos, o que faz com que toda massa de dados normais passe por todos os detectores, resultando em alta latência e um consumo excessivo de recursos computacionais. O objetivo principal deste trabalho é inverter a lógica matemática tradicionalmente utilizada na detecção de anomalias do CEIFA, priorizando inicialmente a identificação de dados normais para otimizar o uso dos recursos computacionais, reduzindo o custo e aumentando a eficiência do sistema. Espera-se que esta nova abordagem não apenas melhore a eficiência do CEIFA, mas também contribua significativamente para uma gestão agrícola mais inteligente e sustentável, alinhando-se com os esforços globais para promover a inovação tecnológica e a sustentabilidade no setor agrícola.

Palavras-chave: detecção de anomalias, sistemas agrícolas inteligentes, otimização de recursos, CEIFA.

ABSTRACT

Agriculture plays a crucial role in sustainable and global economic development, significantly driven by recent technological advancements such as the Internet of Things (IoT) and precision agriculture. These innovations have transformed agricultural practices, enabling more efficient resource management and more sustainable production. Within this context, CEIFA, a low-cost hybrid anomaly detection system, has been developed to enhance anomaly detection in smart agricultural systems. Although it achieves over 98% accuracy, CEIFA prioritizes the detection of anomalous data, which causes all normal data to pass through all detectors, resulting in high latency and excessive consumption of computational resources. The main objective of this work is to reverse the traditional mathematical logic used in CEIFA's anomaly detection, initially prioritizing the identification of normal data to optimize the use of computational resources, reducing costs and increasing system efficiency. This new approach is expected not only to improve the efficiency of CEIFA but also to contribute significantly to smarter and more sustainable agricultural management, aligning with global efforts to promote technological innovation and sustainability in the agricultural sector.

Key-words: anomaly detection, smart agricultural systems, resource optimization, CEIFA.

LISTA DE ILUSTRAÇÕES

Figura 1 – 17 Objetivos de Desenvolvimento Sustentável	7
Figura 2 – Detector de anomalias	11
Figura 3 – Arquitetura do Edge Core	13
Figura 4 – Fluxo de Processamento do Edge Core	14
Figura 5 – Memória alocada na borda	17

LISTA DE ACRÔNIMOS

CDA Cloud Data Analytics.

CEIFA *Cloud-Edge Identifier of Farming Anomalies.*

CML Cloud Machine Learning.

EC Edge Core.

EDA Edge Data Analytics.

EML Edge Machine Learning.

FAO Organização das Nações Unidas para a Alimentação e a Agricultura.

IoT *Internet of Things.*

ODS Objetivos de Desenvolvimento Sustentável.

ONU Organização das Nações Unidas.

SUMÁRIO

1	INTRODUÇÃO	7
1.1	Objetivos	10
1.1.1	Objetivo Geral	10
1.1.2	Objetivos Específicos	10
2	VISÃO GERAL DO DETECTOR DE ANOMALIAS	11
2.1	Edge Core	13
3	METODOLOGIA	18
4	RESULTADOS ESPERADOS	19
4.1	CONSIDERAÇÕES FINAIS	19
4.2	CRONOGRAMA (MODELO PROFESSORA)	20
	REFERÊNCIAS	22

1 INTRODUÇÃO

A agricultura desempenha um papel fundamental na alimentação mundial, sendo o centro do desenvolvimento econômico global (RAJALAKSHMI; MAHALAKSHMI, 2016). À medida que a população mundial cresce¹ e a demanda por alimentos de qualidade também aumenta exponencialmente, a necessidade de inovações na agricultura torna-se cada vez mais necessária. (ALEXANDRATOS; BRUINSMA, 2012). Nesse contexto, a ONU, através de suas agências especializadas como a Organização das Nações Unidas para a Alimentação e Agricultura (FAO, do inglês *Food and Agriculture Organization*), tem enfatizado a importância de transformar e revitalizar o setor agrícola para atender às crescentes necessidades alimentares globais de forma sustentável (Food and Agriculture Organization of the United Nations, 2018).

Um dos principais marcos para a sustentabilidade global é a Agenda 2030, que inclui 17 Objetivos de Desenvolvimento Sustentável (Figura 1) (NACÕES UNIDAS, 2015). Tais objetivos visam erradicar a fome, alcançar a segurança alimentar, melhorar a nutrição e promover a agricultura sustentável. A Agenda 2030 destaca a agricultura como uma área crítica de investimento e inovação, necessária não apenas para alimentar a crescente população mundial, mas também para combater a mudança climática, preservar os recursos naturais e sustentar as comunidades rurais (Food and Agriculture Organization of the United Nations (FAO), 2021).

Figura 1 – 17 Objetivos de Desenvolvimento Sustentável



Fonte: Nações Unidas no Brasil

¹ Estima-se que a população alcance quase 10 bilhões até 2050.

Neste cenário desafiador, a agricultura inteligente surge como uma solução promissora, incorporando tecnologias da Indústria 4.0 e da Internet das Coisas (IoT) para otimizar a produção agrícola (FERREIRA; NASCIMENTO; SANTOS, 2023; ZANELLA, 2022). A agricultura inteligente utiliza uma variedade de tecnologias, como sensores, sistemas de irrigação automatizados, veículos aéreos não tripulados (drones), *big data* e aprendizado de máquina, para aumentar a eficiência das práticas agrícolas (QUY et al., 2022). Estas tecnologias permitem aos agricultores monitorar e analisar as condições do solo, a água, o clima e a saúde das plantas em tempo real, possibilitando tomar decisões que maximizam a produção, reduzem o desperdício e minimizam o impacto ambiental.

A agricultura inteligente está crescendo rapidamente como uma resposta inovadora aos desafios enfrentados pelo setor agrícola tradicional (DAS et al., 2023). Por exemplo, o gerenciamento preciso de recursos hídricos através de sistemas de irrigação inteligente, pode reduzir significativamente o consumo de água, enquanto as análises preditivas podem ajudar na prevenção de pragas e doenças, reduzindo a necessidade de pesticidas químicos (ALI et al., 2023). Além disso, a agricultura de precisão, que se concentra na coleta e análise de dados em larga escala, pode potencializar a aplicação de fertilizantes e sementes, aumentando os rendimentos e preservando a biodiversidade do solo (FRIMPONG et al., 2023).

No entanto, apesar de a agricultura inteligente ter potencial para transformar o setor agrícola, ainda existem barreiras significativas para sua adoção em larga escala (SILVEIRA, 2023). Estas incluem o alto custo inicial das tecnologias, a falta de conhecimento técnico entre os agricultores, e a necessidade de políticas e regulamentações adequadas que promovam a inovação, ao mesmo tempo em que protegem os interesses dos pequenos agricultores e das comunidades rurais (CAMPUZANO et al., 2023). Portanto, é essencial que governos, organizações internacionais, setor privado e a comunidade científica trabalhem juntos para superar esses desafios, promovendo o acesso às tecnologias inovadoras e às práticas agrícolas sustentáveis.

Com o avanço da Internet das Coisas (IoT, do inglês *Internet of Things*) na agricultura, a segurança emerge como uma necessidade, especialmente quando sistemas agrícolas começam a incorporar tecnologias conectadas à Internet, comunicação sem fio e tomada de decisão baseada na nuvem (ELIJAH et al., 2018). A integração entre dispositivos e tecnologias de diferentes fornecedores, a garantia da privacidade e confidencialidade, o gerenciamento eficaz e o armazenamento seguro de informações, assim como a manutenção da confiança e confiabilidade, são aspectos cruciais para o desenvolvimento de soluções agrícolas eficientes e seguras

(ZANELLA; SILVA; ALBINI, 2020). Esses fatores são vitais, não só para prevenir danos nas lavouras e evitar perdas financeiras, mas também para potencializar a produtividade e melhorar a renda dos agricultores, especialmente no contexto da agricultura familiar, onde pequenas variações na produção podem ter impactos significativos (ZANELLA, 2022).

Em meio aos avanços e desafios da agricultura inteligente, a segurança emerge como uma preocupação primordial, impulsionada pela necessidade de proteger os sistemas agrícolas de atividades mal-intencionadas (GUPTA et al., 2020). A proteção contra ataques cibernéticos, que podem comprometer ou explorar a agricultura digital como vetor para ataques mais amplos, destaca a segurança não apenas como um requisito crítico, mas também como um elemento chave para a eficiência e a confiabilidade dos sistemas agrícolas digitais. Neste contexto, o CEIFA (do inglês *Cloud-Edge Identifier of Farming Anomalies*) surge como um sistema projetado para fortalecer a segurança da agricultura inteligente, por meio da detecção de anomalias, assegurando a integridade e a precisão dos dados, o que é fundamental para a tomada de decisões e o controle automatizado dos processos agrícolas (ZANELLA, 2022).

O CEIFA, um detector de anomalias híbrido de baixo custo, pode atuar em uma variedade de cenários agrícolas, promovendo a transição para práticas agrícolas mais inteligentes, seguras e sustentáveis (ZANELLA; SILVA; ALBINI, 2022). Esta característica é particularmente relevante para a agricultura familiar, onde os recursos podem ser limitados e a necessidade de soluções eficazes e acessíveis é crítica. O CEIFA possui a capacidade de identificar uma vasta gama de falhas, erros e potenciais ataques cibernéticos, bem como outras alterações maliciosas nos dados, atendendo tanto a pequenas quanto a grandes propriedades agrícolas. Sua grande inovação reside em uma arquitetura híbrida, que alia a operação tanto na borda quanto na nuvem, permitindo sua adaptação a diferentes contextos do sistema agrícola. Essa flexibilidade é essencial para a implementação do CEIFA em dispositivos com capacidades computacionais restritas, comuns na agricultura familiar. Isso permite que mesmo os produtores com recursos limitados possam beneficiar-se da proteção avançada e das funcionalidades de otimização oferecidas pelo sistema.

Além de proteger contra intrusões² e de identificar anomalias, o CEIFA é um catalisador para a adoção de práticas tecnológicas na agricultura. Ao maximizar a segurança e a precisão dos dados, o CEIFA habilita os agricultores familiares a aumentarem a eficiência de seus cultivos, minimizarem os riscos e reduzirem o desperdício, contribuindo significati-

² Intrusão, refere-se a qualquer acesso ou manipulação de dados não autorizados no sistema.

vamente para a sustentabilidade e a rentabilidade de suas operações (ZANELLA, 2022). Esta solução tecnológica não apenas reforça a resiliência dos sistemas agrícolas frente às ameaças cibernéticas, mas também se alinha com os esforços globais para promover a inovação e a sustentabilidade na agricultura, assegurando que as comunidades rurais possam prosperar em um ambiente cada vez mais digital e interconectado.

O CEIFA, embora seja uma ferramenta inovadora e eficaz, ainda apresenta espaço, para melhorias que possam aumentar ainda mais sua eficiência. Uma das principais áreas de desenvolvimento contínuo é a redução do custo e a melhora do desempenho computacional associado ao funcionamento do sistema. Ao minimizar o uso de processamento e memória, o CEIFA poderá operar de maneira mais leve e melhorar sua eficiência, tornando o processo mais rápido.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Com o propósito de reduzir o custo computacional e minimizar o uso de processamento e memória, este projeto tem como objetivo otimizar o CEIFA, reduzindo sua latência, custo computacional e maximizando sua eficiência, através da inversão da lógica de detecção de anomalias.

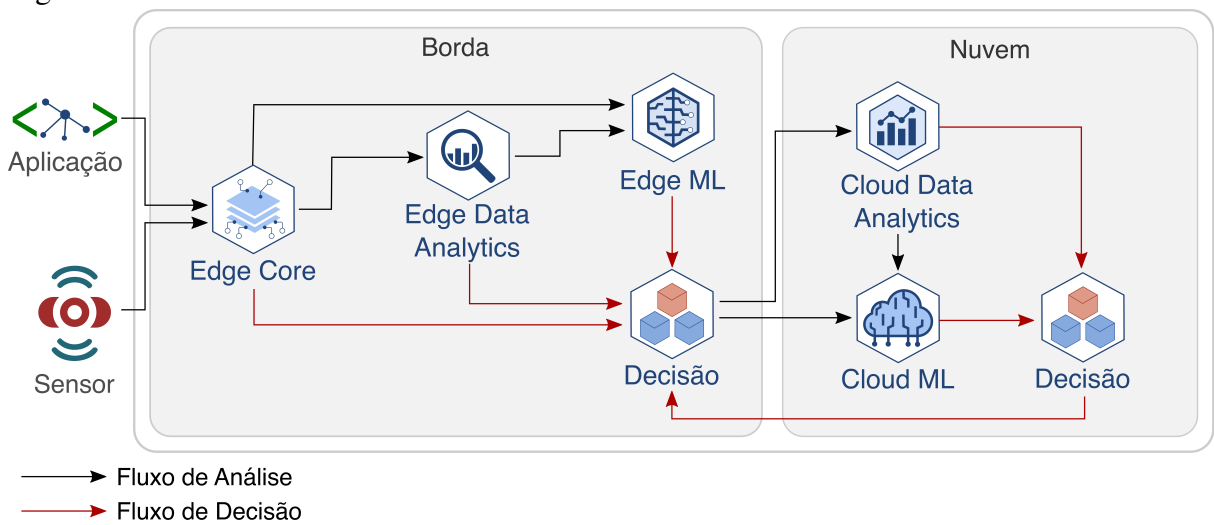
1.1.2 Objetivos Específicos

- Desenvolver um modelo matemático que permita a inversão da lógica de detecção de anomalias, priorizando a identificação preliminar de dados potencialmente normais.
- Avaliar o impacto da inversão da lógica de detecção na eficiência e no consumo computacional do CEIFA.
- Identificar e quantificar os ganhos de desempenho resultantes da aplicação do novo modelo de detecção, em termos de redução do tempo de resposta e economia de recursos.
- Adaptar a arquitetura híbrida do CEIFA (borda e nuvem) para suportar a nova lógica de detecção, garantindo a integração eficaz e a flexibilidade do sistema.

2 VISÃO GERAL DO DETECTOR DE ANOMALIAS

O CEIFA é um sistema estruturado para operar em dois níveis principais: borda e nuvem. Cada nível está equipado com módulos de classificação e módulos de decisão, adaptáveis às capacidades computacionais e às características do ambiente agrícola em que serão implantados. Esses módulos trabalham em conjunto para analisar, classificar e tomar decisões com base nos dados recebidos dos sensores, conforme ilustrado na Figura 2. Essa estrutura modular permite ao CEIFA ser altamente eficiente e adaptável em diferentes ambientes agrícolas.

Figura 2 – Detector de anomalias



Fonte: (ZANELLA, 2022)

O primeiro módulo dentro da arquitetura do CEIFA é o Edge Core (EC), que destaca-se como um componente essencial, atuando como o primeiro nível de filtragem e classificação de dados. O EC é responsável por uma análise inicial dos dados recebidos diretamente dos sensores. Este módulo verifica se os dados estão dentro dos parâmetros normais ou se apresentam anomalias. Dependendo do resultado desta primeira análise, os dados podem ser imediatamente classificados como anômalos ou encaminhados para análises adicionais através de outros módulos na borda ou na nuvem. O EC utiliza técnicas de análise estatística para identificar padrões e possíveis desvios nos dados coletados, aplicando uma combinação de verificação de limites estabelecidos e análise de tendências. Esta abordagem permite uma classificação rápida e eficiente, essencial para responder dinamicamente às condições em tempo real no campo agrícola.

O módulo Edge Data Analytics (EDA) compara os dados em tempo real com informações históricas e previsões, permitindo uma análise contextualizada que pode revelar ten-

dências sazonais ou padrões de longo prazo. Este módulo é crucial para entender as nuances dos dados agrícolas, que frequentemente apresentam variações significativas devido a fatores externos, como clima e sazonalidade.

O Edge Machine Learning (EML) utiliza algoritmos de aprendizado de máquina para aprimorar a precisão das classificações. Este módulo é especialmente útil em cenários onde os recursos computacionais são adequados e há uma necessidade de processamento analítico mais sofisticado. O aprendizado de máquina no EML aprimora a precisão na detecção de anomalias, reduzindo a latência na comunicação e o tráfego de dados para a nuvem, otimizando o uso dos recursos na nuvem (ZANELLA, 2022).

O módulo de decisões recebe os dados e as classificações dos módulos classificadores (EC, EDA e EML) e, com base nas classificações, toma uma decisão, que pode ser: devolver o dado para a aplicação junto com a informação sobre a classificação gerada, reter o dado e informar sobre a classificação (em casos de anomalias severas) ou encaminhar para a nuvem para nova análise.

Na nuvem, os módulos Cloud Data Analytics (CDA) e Cloud Machine Learning (CML) oferecem funcionalidades semelhantes aos seus correspondentes de borda, mas com acesso a uma capacidade computacional significativamente maior e conjuntos de dados mais completos. Eles permitem análises mais profundas e complexas, capazes de integrar múltiplas fontes de dados e executar algoritmos de aprendizado de máquina mais sofisticados. O módulo de decisões na nuvem finaliza o processo, tomando ações baseadas nas análises realizadas, que podem incluir ajustes automatizados ou alertas para gestão centralizada.

O módulo CDA é executado na nuvem e se utiliza de dados históricos e previsões meteorológicas para identificar comportamentos anômalos. Este módulo aproveita a maior capacidade computacional e a conectividade constante disponíveis na nuvem para processar dados mais complexos e realizar análises mais profundas. O CDA emprega técnicas de análise de tendências e probabilidade, baseando-se em informações de longo prazo e previsões para detectar padrões e comportamentos anômalos. A análise é feita tanto de forma combinada quanto individual e pode incluir previsões meteorológicas mais precisas e atualizadas, obtidas de múltiplas fontes.

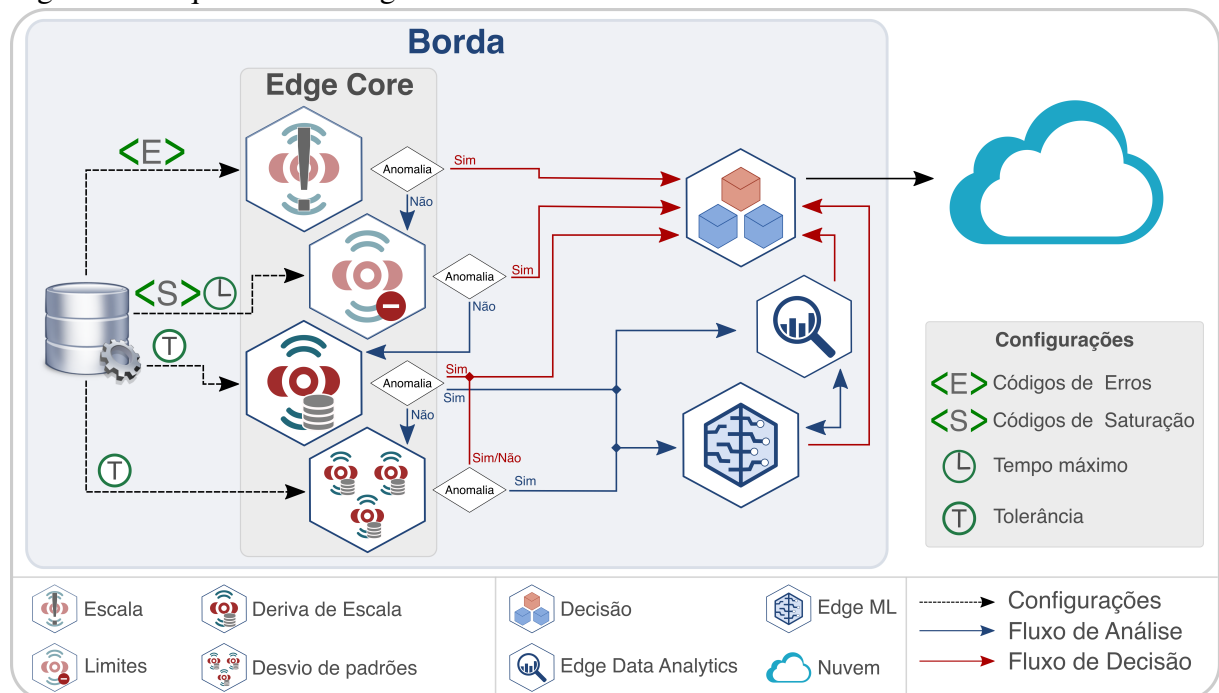
O módulo CML emprega algoritmos de aprendizado de máquina mais robustos para a detecção de anomalias e ataques, beneficiando-se dos recursos computacionais avançados disponíveis na nuvem. Diferentemente dos algoritmos usados na borda, que são restritos pelas ca-

pacidades dos dispositivos, o CML pode implementar técnicas sofisticadas como aprendizagem profunda e algoritmos específicos para grandes volumes de dados. Este módulo pode operar tanto em complemento aos módulos de borda quanto independentemente, proporcionando uma camada adicional de filtragem ou agindo como um reforço na detecção de anomalias não identificadas nos níveis inferiores.

2.1 EDGE CORE

O EC é o componente principal no sistema de detecção de anomalias do CEIFA, representando um módulo essencial, enquanto os outros módulos são opcionais. O EC é responsável pela análise estatística inicial dos dados recebidos diretamente dos sensores. Este módulo classificador é essencial para o primeiro nível de filtragem e classificação dos dados, permitindo uma rápida verificação da normalidade dos dados ou da presença de anomalias. Conforme mostrado na Figura 3, o EC desempenha um papel fundamental na estrutura do sistema.

Figura 3 – Arquitetura do Edge Core



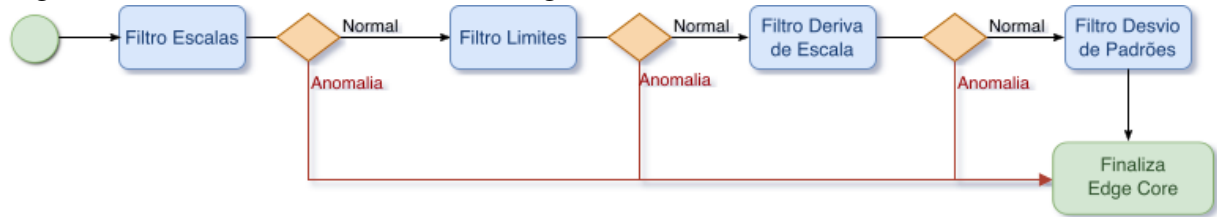
Fonte: (ZANELLA, 2022)

A arquitetura do EC é projetada para operar eficientemente, mesmo em dispositivos com capacidades computacionais limitadas, típicas em ambientes de agricultura familiar. Este módulo serve como uma linha de defesa primária, processando dados em tempo real e decidindo

rapidamente sobre a normalidade dos valores antes de passá-los para análises mais complexas na nuvem ou outros módulos na borda.

O fluxo de processamento no EC, apresentado na Figura 4, é estruturado para minimizar o uso de recursos ao mesmo tempo em que mantém uma resposta rápida e eficiente, essencial para a tomada de decisões imediatas em ambientes agrícolas dinâmicos. Inicialmente, o processo começa com o filtro de *Escala*, responsável por assegurar que todos os valores recebidos dos sensores estejam dentro dos limites especificados.

Figura 4 – Fluxo de Processamento do Edge Core



Fonte: (ZANELLA, 2022)

Este filtro utiliza a Equação 2.1 para analisar os dados recebidos, verificando se cada valor medido está dentro de um intervalo mínimo e máximo aceitável. Caso um valor esteja fora desses limites, ele é marcado como potencialmente anômalo e pode ser encaminhado para análises adicionais ou classificado diretamente como uma anomalia.

$$(v \geq v_{min}) \wedge (v < v_{max}) = normal$$

onde: v = valor,

(2.1)

v_{min} = menor valor,

v_{max} = maior valor

Em seguida, o dado é encaminhado para o Filtro *Limites* (Equação 2.2), que foca na detecção de valores que atingem os extremos de suas capacidades de medição, o que pode indicar condições de saturação ou degradação do sensor. Este filtro é essencial para identificar falhas técnicas ou alterações ambientais extremas que causem leituras consistentemente no máximo ou mínimo da escala do sensor, sugerindo a necessidade de manutenção ou ajuste no equipamento.

$$(v = v_{min} \wedge \overline{H_s} = v_{min}) \vee (v = v_{max} \wedge \overline{H_s} = v_{max}) = anomalia$$

onde: v = valor,

$$v_{min} = \text{menor valor}, \quad (2.2)$$

$$v_{max} = \text{maior valor},$$

$$\overline{H_s} = \text{valor hist3rico}$$

Valores que falham em ambos os filtros anteriores s3o processados pelo filtro de *Deriva de Escala*, que utiliza a Equa33o 2.3 para examinar mudan3as nos dados que desviam dos padr3es hist3ricos. Este filtro, ao aplicar m3todos estat3sticos, compara a m3dia dos valores hist3ricos do sensor com a m3dia e o desvio padr3o de sensores similares, analisando os dados atuais em rela33o a estat3sticas de longo prazo, para identificar qualquer comportamento an3malo que possa sinalizar problemas emergentes ou mudan3as significativas no ambiente monitorado.

$$\overline{H_s} + \sigma_t + \varepsilon < \overline{H_t}$$

onde: $\overline{H_s}$ = valor hist3rico

$$\sigma_t = \text{margem de erro} \quad (2.3)$$

$$\varepsilon = \text{toler3ncia}$$

$$\overline{H_t} = \text{valor m3dio hist3rico}$$

O filtro de *Deriva de Escala*, que identifica diferentes problemas, emprega uma an3lise minuciosa dos dados hist3ricos do sensor para detectar mudan3as graduais ou abruptas nas condi33es monitoradas. Utilizando estat3sticas de longo prazo, como m3dia e desvio padr3o, ele compara as tend3ncias e dispers3o dos dados atuais com os registros hist3ricos. Esta compara33o ajuda a identificar quaisquer desvios dos padr3es normais que possam indicar problemas emergentes ou mudan3as significativas no ambiente monitorado. Se os dados n3o correspondem aos crit3rios de normalidade estabelecidos, eles s3o encaminhados para o pr3ximo n3vel de an3lise.

O Filtro de *Desvio de Padr3es* (Equa33o 2.4) foi inspirado na Desigualdade de Hoeffding (RAO, 2018). Esse filtro utiliza m3todos estat3sticos para examinar se os valores medidos se desviam significativamente dos padr3es observados em sensores similares ou sob condi33es semelhantes. O filtro Desvio de Padr3es 3 essencial para capturar anomalias sutis

que podem não ser imediatamente óbvias através das verificações de escala ou limite, permitindo uma identificação mais fina de irregularidades que poderiam passar despercebidas, mas que podem ter implicações significativas para a operação agrícola.

$$|v - \overline{H_s}| - \varepsilon < \sigma_s$$

onde: v = valor,

$\overline{H_s}$ = valor histórico, (2.4)

ε = tolerância,

σ_s = margem de erro histórico

Juntos, os filtros apresentados formam o principal módulo de detecção de anomalias do CEIFA, o Edge Core. Este módulo demonstrou ser altamente eficaz na detecção de anomalias, alcançando uma precisão de 65,17% conforme resultados disponíveis na Tabela 1. Esse resultado destaca a capacidade do EC em comparação aos outros componentes do CEIFA, proporcionando uma detecção confiável e eficiente em ambientes de agricultura inteligente.

Tabela 1 – Classificação dos dados anômalos pelos módulos do CEIFA

Anomalias	Borda				Nuvem		
	Edge Core	EDA	EML	Total	CDA	CML	Total
Falhas Aleatórias	100,00%	-	-	100,00%	-	-	100,00% ¹
Saturação	51,00%	43,00%	-	94,00%	54,10%	44,80%	98,90%
Degradação	50,60%	29,10%	0,20%	79,90%	71,50%	27,80%	99,33%
Danificação	77,10%	10,80%	0,50%	88,40%	72,40%	8,63%	97,30% ²
Ruídos	99,80%	0,20%	-	100,00%	0,20%	-	100,00% ³
Injeção de Dados Falsos	11,31%	7,60%	1,01%	19,37%	28,90%	68,90%	97,80%
Geral	65,17%	13,94%	0,69%	79,60%	54,74%	20,91%	98,31%⁴

¹ A nuvem recebeu 100% das falhas com a marca “A1”, por isso elas foram marcadas como anomalia sem passar pelos classificadores da nuvem.

² A borda enviou 16,27% das anomalias com a marca “A1”.

³ A nuvem recebeu 99,8% dos registros de ruídos com a marca “A1”.

⁴ A nuvem recebeu 22,70% das anomalias com a marca “A1”.

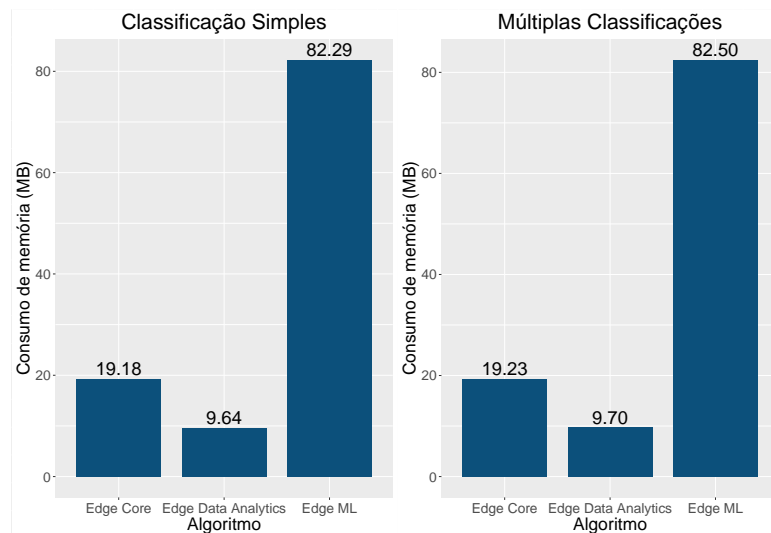
Fonte: (ZANELLA, 2022)

No entanto, o processo, apesar de eficiente em detectar anomalias, demanda que os dados normais, que são a grande maioria, passem por todos os filtros antes de serem classificados. Isso implica em um consumo maior de recursos computacionais para identificar dados, os normais, do que para dados anômalos. Acredita-se que é possível otimizar o consumo de recursos computacionais invertendo a forma de detecção, de modo que a maior parte dos dados normais seja identificada já nas fases iniciais do processo. Essa mudança poderia reduzir significativamente o tempo necessário para a detecção e o consumo de recursos computacionais,

aumentando a eficiência do sistema ao diminuir a carga nos filtros mais complexos e consumidores de recursos, que atualmente são aplicados a todos os dados.

De acordo com os dados apresentados na Figura 5, o EC consome 19,18MB de memória para classificar um único registro e 19,23MB para classificar múltiplos registros, representando um incremento de pouco mais de 49KB de memória. Medições feitas com o *smaps*¹ mostram que o Edge Data Analytics é mais econômico, consumindo apenas 9,64MB de memória para classificar um registro, com um acréscimo de 61KB para classificação de múltiplos registros, totalizando 9,70MB. O maior consumo de memória é atribuído ao Edge ML, que utiliza mais de 82MB. O processamento de múltiplos registros por esse classificador adiciona mais de 217KB de memória.

Figura 5 – Memória alocada na borda



Fonte: (ZANELLA, 2022)

Essa análise detalhada das métricas de classificação e memória nos diversos módulos do sistema CEIFA evidência a complexidade e as exigências variadas dos diferentes componentes no contexto da borda. Esses dados são fundamentais para entender melhor como otimizar o desempenho dos sistemas de detecção de anomalias na agricultura inteligente, permitindo que a operacionalidade e eficiência sejam mantidas sem sobrecarregar os recursos computacionais disponíveis.

¹ Smaps é uma ferramenta baseada no sistema de arquivos do Linux que fornece informações detalhadas sobre o uso de memória de processos, incluindo a memória usada por diferentes segmentos de um programa.

3 METODOLOGIA

Para o desenvolvimento de uma nova abordagem na detecção de anomalias no CEIFA, o foco inicial é aprimorar a priorização dos dados normais, alterando a lógica matemática atual. Essa inversão visa otimizar o processamento e reduzir a carga computacional, concentrando-se primeiro em identificar e confirmar a normalidade dos dados antes de avançar para a análise de dados com potenciais anomalias. A pesquisa será de natureza experimental, qualitativa e aplicada. A primeira etapa consistirá em uma análise estatística dos dados existentes para explorar a viabilidade e as implicações práticas dessa inversão lógica.

Em seguida, utilizando conjuntos de dados já estabelecidos e novos, provenientes de sensores virtuais, serão utilizados para simular condições reais de operação. Isso permitirá a comparação entre os resultados novos e os já existentes, avaliando assim a eficácia em relação à abordagem tradicional. Essa comparação direta visa não apenas validar a precisão e destacar potenciais melhorias, mas também identificar potenciais equações matemáticas capazes de discernir dados normais de forma mais eficiente.

Dessa forma, será realizada uma análise detalhada das fórmulas matemáticas e estatísticas com o objetivo de aprimorar a detecção de dados normais, reduzindo a carga sobre os filtros subsequentes dedicados a identificar anomalias. Este processo envolverá a revisão e possivelmente a reformulação das equações existentes para maximizar a eficiência na classificação inicial dos dados. Ao concentrar-se em reconhecer a maior parte dos dados como normais com alta precisão, pode-se delegar a análise de dados potencialmente anômalos para etapas mais avançadas do sistema.

Ao final deste processo, espera-se estabelecer uma fórmula matemática mais eficiente que melhore significativamente a eficiência dos sistemas para detecção de dados normais, reduzindo a quantidade de dados que passarão pelos filtros, permitindo uma resposta mais rápida e precisa nas operações.

4 RESULTADOS ESPERADOS

O CEIFA como apresentado, desempenha um papel crucial no contexto da agricultura inteligente, sendo um sistema avançado para a detecção de anomalias que analisa dados de forma intensiva, o que consequentemente demanda uma quantidade significativa de recursos computacionais. Ao longo deste trabalho, foi apresentado a possibilidade de inverter a lógica matemática utilizada pelo CEIFA, com o objetivo de priorizar a identificação de dados normais antes de submeter os dados potencialmente anômalos aos filtros de análise.

Esta abordagem propõe uma nova maneira de processar os dados, visando reduzir o consumo de recursos computacionais. Visto que a maior parte dos dados coletados em ambientes agrícolas tende a ser normal, focar inicialmente na confirmação dessa normalidade permite que apenas uma parcela menor de dados, aqueles que apresentam potenciais anomalias, seja submetida aos processos mais intensivos. Espera-se que essa inversão na lógica matemática não só melhore a eficiência do sistema como também otimize o uso da capacidade computacional disponível.

Implementar essa nova metodologia envolverá uma série de análises estatísticas para determinar a viabilidade dessa inversão, assim como experimentações com dados de sensores virtuais que simulam condições reais, cruzando com dados já registrados anteriormente, permitindo avaliar a eficácia da abordagem proposta. Ao final desse processo, a inversão, contribuirá para uma operação mais sustentável e eficiente do sistema CEIFA, para detecção de anomalias em ambientes de agricultura inteligente, alinhando-se com os esforços globais para promover a inovação tecnológica e a sustentabilidade no setor agrícola.

4.1 CONSIDERAÇÕES FINAIS

A abordagem proposta requer uma revisão cuidadosa e potencial reestruturação das fórmulas matemáticas existentes. Isso envolverá uma análise estatística para determinar se algumas das equações atuais podem ser reutilizadas ou se é necessário desenvolver novas formulações que se alinhem melhor com os objetivos de detecção de dados normais. Este processo será apoiado por testes extensivos usando dados provenientes de sensores virtuais, simulando condições reais para comparar os novos resultados com registros anteriores.

Esses testes não apenas validam a nova metodologia, mas também ajudam a ajustar as equações para maximizar a precisão na identificação e classificação inicial de dados normais.

Ao conseguir identificar a maioria dos dados como normais com alta confiabilidade, o sistema poderá então focar os recursos computacionais na análise de dados que realmente apresentam potenciais anomalias.

4.2 CRONOGRAMA (MODELO PROFESSORA)

O desenvolvimento do trabalho e as atividades planejadas até a defesa final, prevista para dezembro de 2024, estão organizadas conforme a tabela 2.

Tabela 2 – Cronograma de atividades

Atividades	2024					
	jul	ago	set	out	nov	dez
Análise e reformulação de equações	•	•				
Testes com dados de sensores	•	•	•			
Coleta de dados		•	•			
Avaliação dos resultados dos testes			•	•		
Análise estatística de desempenho			•	•	•	
Finalização escrita TCC final					•	•
Defesa TCC						•

Análise e reformulação de equações - Esta etapa envolve o exame detalhado das equações matemáticas atuais utilizadas pelo CEIFA para identificar potenciais melhorias ou necessidades de adaptação. O objetivo é desenvolver fórmulas que priorizem a eficiência e a eficácia na detecção de dados normais;

Testes com dados de sensores virtuais - Realização de experimentos utilizando conjuntos de dados gerados por sensores virtuais que simulam condições ambientais reais;

Coleta de dados - Os dados coletados serão analisados e processados, para entender os padrões normais e anormais dos dados;

Avaliação dos resultados dos testes - Análise dos dados coletados nos testes para avaliar o desempenho das novas equações e a eficiência geral do sistema CEIFA com as modificações implementadas;

Finalização escrita final TCC - Consolidação de todos os resultados, análises e processos. Essa fase incluirá a revisão do texto;

Defesa TCC - Preparação e realização da apresentação final do TCC perante uma banca examinadora.;

REFERÊNCIAS

- ALEXANDRATOS, Nikos; BRUINSMA, Jelle. **World agriculture towards 2030/2050: the 2012 revision**. 2012. <<https://www.fao.org/global-perspectives-studies/resources/detail/en/c/411108/>>. Food and Agriculture Organization of the United Nations.
- ALI, Awais et al. Application of smart techniques, internet of things and data mining for resource use efficient and sustainable crop production. **Agriculture**, v. 13, n. 2, 2023.
- CAMPUZANO, Laura Restrepo et al. Barriers to the adoption of innovations for sustainable development in the agricultural sector—systematic literature review (slr). **Sustainability**, v. 15, n. 5, 2023.
- DAS, Kunal et al. Smartfields: Revolutionizing agriculture with iot-based monitoring and auto-irrigation. **International Journal of Research and Analytical Reviews (IJRAR)**, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune-43, Maharashtra, India, v. 10, n. 2, p. 866, 2023. Disponível em: <<http://www.ijrar.org/IJAR23B2783>>.
- ELIJAH, Olakunle et al. An overview of internet of things (iot) and data analytics in agriculture: Benefits and challenges. **IEEE Internet of things Journal**, IEEE, v. 5, n. 5, p. 3758–3773, 2018.
- FERREIRA, Amanda Machado; NASCIMENTO, Gabrielly Cardoso; SANTOS, Leandro Colavati dos. Tecnologia da internet das coisas na agricultura 4.0: Uma revisão sistemática. **Advances in Global Innovation & Technology**, v. 1, n. 2, p. 50–57, 2023.
- Food and Agriculture Organization of the United Nations. **The future of food and agriculture – Alternative pathways to 2050**. 2018. <<https://www.fao.org/global-perspectives-studies/food-agriculture-projections-to-2050/en/>>.
- Food and Agriculture Organization of the United Nations (FAO). **Introducing Climate-Smart Agriculture**. 2021. <<https://www.fao.org/climate-smart-agriculture-sourcebook/concept/module-a1-introducing-csa/chapter-a1-1/en/>>.
- FRIMPONG, Felix et al. Water-smart farming: review of strategies, technologies, and practices for sustainable agricultural water management in a changing climate in west africa. **Frontiers in Sustainable Food Systems**, CSIR-Crops Research Institute of Ghana, Kumasi, Ghana; Department of Biological Production Science, Tokyo University of Agriculture and Technology, Tokyo, Japan, 2023.
- GUPTA, Maanak et al. Security and privacy in smart farming: Challenges and opportunities. **IEEE Access**, v. 8, p. 34564–34584, 2020.
- NACÕES UNIDAS. **Objetivos de Desenvolvimento Sustentável**. 2015. <<https://nacoesunidas.org/pos2015/>>. Acesso em: 23/03/2024.
- QUY, Vu Khanh et al. Iot-enabled smart agriculture: Architecture, applications, and challenges. **Applied Sciences**, v. 12, n. 7, 2022.
- RAJALAKSHMI, P.; MAHALAKSHMI, S. Devi. Iot based crop-field monitoring and irrigation automation. In: **10th International Conference on Intelligent Systems and Control (ISCO 2016)**. Coimbatore, India: IEEE, 2016. p. 1–6.

RAO, Shravas. A hoeffding inequality for markov chains. **Electronic Communications in Probability**, The Institute of Mathematical Statistics and the Bernoulli Society, v. 24, p. 1–11, 2018.

SILVEIRA, Franco da. Gerenciamento das barreiras no desenvolvimento da agricultura 4.0 na cadeia de produção agrícola da região sul do brasil. Porto Alegre, 2023.

ZANELLA, Angelita Rettore de Araujo. **Detector Híbrido de Anomalias para Agricultura Inteligente**. Tese (Doutorado) — Universidade Federal do Paraná, Curitiba, PR, 2022.

ZANELLA, Angelita Rettore de Araujo; SILVA, Eduardo da Silva; ALBINI, Luiz Carlos Pessoa Albini. Security challenges to smart agriculture: Current state, key issues, and future directions. **Array**, v. 8, p. 100048, 2020.

ZANELLA, Angelita Rettore de Araujo; SILVA, Eduardo da; ALBINI, Luiz Carlos Pessoa. Ceifa: A multi-level anomaly detector for smart farming. **Computers and Electronics in Agriculture**, Elsevier, v. 202, p. 107279, 2022.