I Workshop de Computação Aplicada no Agronegócio

INTEGRAÇÃO DE DADOS AGRÍCOLAS

Emili Everz Golombiéski – Mestrado em Computação Aplicada, UEPG – emilieverz043@gmail.com
Maria Salete Marcon Gomes Vaz – Mestrado em Computação Aplicada, UEPG – salete@uepg.br

Resumo: A agricultura enfrenta desafios devido à crescente frequência e intensidade dos eventos climáticos extremos, como secas prolongadas, tempestades intensas e ondas de calor. Essas condições adversas afetam a produtividade agrícola e a segurança alimentar, complicando a gestão de recursos hídricos e a previsão climática. A integração de dados, utilizando processos como ETL - Extração, Transporte e Carregamento, enfrenta esses desafios. Ao transformar dados brutos de diversas fontes em informações acionáveis, essa abordagem facilita a tomada de decisões estratégicas e operacionais. Além disso, o modelo constelação no contexto de *data warehouses* organiza os dados, permitindo análises rápidas e detalhadas. Este artigo explora a integração de dados climáticos e imagens de satélite para a melhoria da resiliência e da sustentabilidade agrícola, destacando as melhores práticas e desafios envolvidos.

Palavras-chave: ETL. Agricultura. Impactos Climáticos. Decisões Estratégicas.

AGRICULTURAL DATA INTEGRATION

Abstract: *Agriculture faces challenges due to the increasing frequency and intensity of extreme weather events, such as prolonged droughts, severe storms, and heatwaves. These adverse conditions affect agricultural productivity and food security, complicating water resource management and climate forecasting. The integration of data using processes like ETL - Extraction, Transformation, and Loading - addresses these challenges. By transforming raw data from various sources into actionable information, this approach facilitates strategic and operational decision-making. Additionally, the constellation model in the context of data warehouses organizes the data, enabling rapid and detailed analyses. This article explores the integration of climate data and satellite images to enhance agricultural resilience and sustainability, highlighting best practices and the challenges involved.*

Key-words: *Agriculture. Climate Impacts. Strategic Decisions.*

1. Introdução

Cogato et al. (2019) destacam que a agricultura enfrenta desafios devido aos eventos climáticos extremos, frequentes e severos. Esses eventos, como secas prolongadas, tempestades intensas e ondas de calor, afetam diretamente a produtividade agrícola e a segurança alimentar.

As mudanças climáticas resultam em padrões imprevisíveis de precipitação e temperatura, tornando dificil a previsão do tempo e complexa gestão de recursos hídricos (Elahi, et al., 2021).

As secas podem levar à escassez de água, comprometendo a saúde das culturas e reduzindo as colheitas. As fortes tempestades podem causar danos às plantações e ao solo.

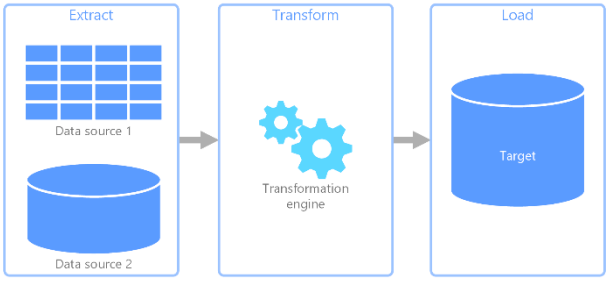
Esses impactos aumentam os riscos para os agricultores e desafiam a resiliência das práticas agrícolas tradicionais, exigindo soluções adaptativas e inovadoras para diminuir os efeitos adversos e garantir a sustentabilidade da produção agrícola.

Para enfrentar esses desafios, é necessário adotar a diminuição dos efeitos adversos e promover uma integração eficaz de dados agrícolas.

Cima et al. (2021) explicam a integração de dados, exigindo harmonização de informações provenientes de diversas fontes, como de estações de dados, sensores de campo, imagens de satélite, entre outros, para criar um panorama completo e preciso do área agrícola

Para integração de dados, demanda utilizar o Processo ETL - *Extract, Transform, Load* (Extrair, Transformar e Carregar), metodologia para integração e gerenciamento de dados, especialmente em ambientes de *data warehouse* (Figura 1). Cada uma das três etapas permite a preparação e organização dos dados para análise.

Figura 1. Fluxo ETL.



Fonte: Microsoft Lern (2024).

No contexto atual da agricultura, a integração de dados foi inserida para a tomada de decisão estratégica e operacional. Com a evolução das tecnologias digitais e a crescente disponibilidade de dados provenientes de diversas fontes, como sensores de campo, satélites e plataformas de monitoramento em tempo real, os produtores agrícolas e gestores de agronegócios enfrentam o desafio de converter dados em informações acionáveis.

A integração desses dados melhora a precisão das decisões e promove sustentabilidade e eficiência no uso de recursos naturais (Goldstein, Fink e Ravid, 2022).

A abordagem tradicional, baseada em experiências e observações empíricas, está sendo complementada por métodos avançados que incorporam *big data*, inteligência artificial e Internet das Coisas.

Os avanços permitem análise detalhada e preditiva das condições ambientais, dos processos produtivos e das dinâmicas de mercado (Montalvo-Romero, et al., 2023).

A integração de dados possibilita a criação de modelos preditivos e sistemas de suporte à decisão, podendo antecipar problemas, otimizar práticas de cultivo e minimizar impactos negativos.

No entanto, a integração de dados agrícolas enfrenta desafios, como a heterogeneidade das fontes de dados, a complexidade dos sistemas de informação e a necessidade de sistemas robustos para gestão e análise de grandes volumes de dados.

Para superar esses desafios e melhorar a integração de dados agrícolas, demanda a adoção de abordagens estruturadas e eficientes. No ambiente de um *data warehouse*, o uso de modelo constelação oferece uma solução para tratar com a heterogeneidade das fontes de dados e a complexidade dos sistemas.

O modelo é uma estrutura organizada, com dados centralizados em tabelas de fatos, rodeadas por tabelas de dimensões, facilitando a integração e análise de grandes volumes de dados (Chen, et al., 2022).

A adoção do modelo constelação permite uma análise rápida e detalhada das informações, oferecendo uma visão clara e consolidada dos dados agrícolas, para enfrentar os desafios inerentes à gestão e análise de dados complexos e variados.

A estrutura incorpora os modelos estrela e floco de neve, em sua forma normalizada. A principal vantagem é permitir análise rápida e eficiente dos dados, proporcionando visão clara e detalhada das informações.

2. Material e Métodos

Para o progresso da pesquisa, foi necessário realizar revisão do estado da arte de integração de dados, incluindo uma análise dos diferentes tipos de dados gerados no setor agrícola.

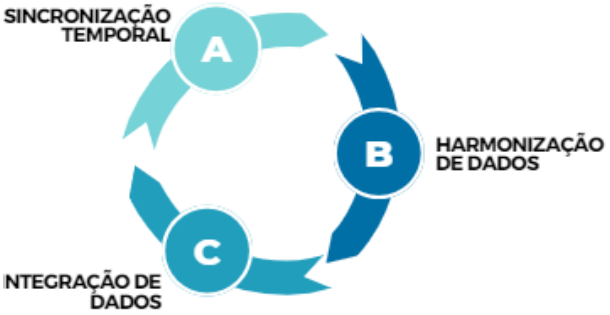
Foram analisados os métodos de acesso e manipulação dos dados, com detalhamento das etapas necessárias para especificação de um modelo de integração de dados agrícolas.

A revisão resultou na identificação das melhores práticas e desafios associados à integração de dados, provenientes de diversas fontes no contexto agrícola, bem como a apresentação das técnicas utilizadas e das ferramentas a serem utilizadas.

3. Resultados e Discussões

Para a criação do modelo de integração de dados climáticos e imagens de satélite, demanda melhorar a tomada de decisão agrícola, seguindo um processo estruturado de dados (Figura 2).

Figura 2. Etapas do Processo para Integração



Fonte: As autoras.

Ao realizar a integração de dados, passa por várias etapas. Inicialmente, deve ser realizada a sincronização, onde os dados climáticos e as imagens são alinhados temporalmente, garantindo a comparação e relevância das informações no período analisado.

Na sequência, a harmonização dos dados foi feita, com técnicas de pré-processamento, ajustando as escalas e formatos dos dados para garantir que sejam compatíveis. As imagens são georreferenciadas para corresponder às coordenadas das estações meteorológicas.

Finalmente, foi realizada a integração de dados utilizando ferramentas ETL para modelar os dados climáticos e as imagens em um único banco de dados integrado. Essa etapa envolveu a criação de um modelo de dados, relacionando as variáveis climáticas com as características de defeitos relativos aos grãos, observados nas imagens.

Para compreender as técnicas de fusão de dados e algoritmos de integração, foi necessário combinar variáveis climáticas com características das imagens (Tabela 1).

Tabela 1. Principais Técnicas e Algoritmos de Integração.

Modelos	Benefícios
Método de Fusão por Média	Simplifica a combinação de dados, proporcionando uma visão geral das variáveis.
Análise de Componentes Principais (PCA)	Reduz a dimensionalidade dos dados, destacando as principais variáveis.
Classificação por Redes Neurais	Permite a identificação de padrões complexos e a previsão baseada em dados integrados.
Algoritmos de Fusão Baseada em Regra	Integra dados com base em regras, aumentando a precisão na combinação dos dados.
Modelos de Fusão Baseada em Estatísticas	Melhora a análise estatística das variáveis integradas, oferecendo informações precisas.
Métodos de Interpolação Espacial	Facilita a combinação de dados em diferentes escalas espaciais, melhorando a precisão espacial.
Mapeamento de Coeficientes de Correlação	Avalia a relação entre variáveis climáticas e características das imagens, melhorando a interpretação.

Fonte: As autoras.

Cada técnica de integração de dados climáticos e imagens de satélite oferece abordagens distintas para tratar dados complexos e variados. O método de fusão por média é útil na busca de uma visão geral e simplificada, combinando dados para obter uma média geral.

A análise de componentes principais (PCA) é apropriada quando demanda reduzir a complexidade dos dados, destacando as variáveis significativas para uma análise focada.

As redes neurais são eficazes para identificar padrões complexos e realizar previsões baseadas em grandes volumes de dados integrados.

Os algoritmos de fusão baseada em regra são ideais na aplicação de regras específicas para combinar dados, aumentando a precisão em contextos bem definidos.

Os modelos de fusão baseada em estatísticas oferecem uma análise mais detalhada das relações estatísticas entre variáveis integradas, proporcionando informações profundas.

Para integrar dados em diferentes escalas espaciais, métodos de interpolação espacial são necessários, enquanto o mapeamento de coeficientes de correlação é útil para entender como diferentes variáveis climáticas e características das imagens se relacionam entre si, ajudando na interpretação dos dados.

A inovação na integração de dados climáticos e imagens de satélite é necessária para enfrentar os desafios complexos e dinâmicos da agricultura.

Para contribuir com a adoção tecnológica no campo, novas metodologias estão sendo desenvolvidas para melhorar a precisão, eficiência e aplicabilidade dos dados integrados. A Tabela 2 apresenta metodologias para integração de dados agrícolas, destacando suas características e benefícios.

Tabela 2.Metodologias de Integração de Dados Agrícolas.

Nome da Metodologia	Descrição	Benefícios
Integração de Dados com *Blockchain*	Uso de *blockchain* para garantir a segurança e a transparência dos dados integrados.	Aumenta a confiança na integridade e na origem dos dados.
Análise de Dados Espaciais com *Machine Learning*	Aplicação de algoritmos de *machine learning* para analisar padrões espaciais	Melhora a identificação de tendências e anomalias espaciais.

Nome da Metodologia	Descrição	Benefícios
	em dados integrados.	
Modelos de Previsão Baseados em Internet das Coisas	Integração de dados de sensores em internet das coisas, para criar modelos preditivos em tempo real.	Permite respostas rápidas a mudanças ambientais e operacionais.
Plataformas de Visualização Interativa	Desenvolvimento de ferramentas interativas para visualizar e explorar dados integrados.	Facilita a tomada de decisões através de informações visuais e intuitivas.
Algoritmos de Fusão de Dados Multi-escalares	Integração de dados em diferentes escalas espaciais e temporais utilizando técnicas avançadas de fusão.	Melhora a precisão das análises ao considerar múltiplas escalas.

Fonte: As autoras.

Nessa tabela, é apresentada uma série de metodologias em desenvolvimento para aprimorar a integração de dados agrícolas. A integração de dados com *blockchain oferece* uma camada adicional de segurança e transparência, necessária para a confiabilidade dos dados.

A análise de dados espaciais com *machine learning* permite identificar padrões e tendências, podendo não ser visíveis através de métodos tradicionais.

Modelos de previsão baseados em Internet das Coisas utilizam dados de sensores em tempo real para prever condições ambientais, permitindo ações proativas.

As plataformas de visualização interativa transformam dados complexos em informações visuais intuitivas, facilitando a tomada de decisões.

Os algoritmos de fusão de dados multi-escalares garantem que os dados sejam integrados de maneira coerente e precisa, considerando variações em diferentes escalas espaciais e temporais. Essas metodologias refletem avanços na forma como os dados agrícolas são gerenciados e utilizados, promovendo agricultura eficiente e resiliente.

4. Conclusões e Perspectivas de Pesquisas Futuras

A integração de dados agrícolas, especialmente no contexto de eventos climáticos extremos, representa um avanço na forma como os dados são geridos e analisados para enfrentar desafios complexos.

A crescente frequência e severidade dos eventos climáticos exigem uma abordagem para a coleta e análise de informações, possibilitando resposta ágil e informada.

A integração eficaz de dados climáticos e imagens de satélite, facilitada por modelo multidimensional, como modelo constelação, oferece uma estrutura robusta para tratar a complexidade e a heterogeneidade dos dados.

O resultado corresponde a uma visão clara e detalhada dos impactos climáticos e permite uma tomada de decisão estratégica e eficiente.

A aplicação de metodologias avançadas para a integração de dados melhora a gestão das informações e contribui para a sustentabilidade e resiliência da produção agrícola.

A utilização de técnicas como a fusão de dados e algoritmos de integração possibilita a criação de modelos preditivos e sistemas de suporte à decisão, essenciais para enfrentar as incertezas e otimizar as práticas agrícolas.

O desenvolvimento contínuo de novas ferramentas e técnicas para a integração de dados permite aos agricultores adaptar-se às mudanças climáticas e garantir a segurança alimentar.

Como perspectivas de pesquisas futuras, sugere-se utilizar o Sistema de Informação Geográfica para criar plataformas interativas que integram dados climáticos e imagens de satélite com outros dados, como informações de uso do solo e práticas agrícolas. Isso permite uma visualização holística e interativa dos dados, facilitando a análise espacial e a tomada de decisões.

Outra possibilidade é usar a teoria das redes para analisar as interações complexas entre diferentes variáveis climáticas, características das imagens e práticas agrícolas. Isso pode ajudar a identificar relações não lineares e padrões emergentes que influenciam a saúde das culturas e a produtividade agrícola.

Ao aplicar técnicas de análise preditiva com base em dados históricos e em tempo real, pode incluir a previsão de eventos climáticos extremos, pestes e doenças, ajudando a antecipar e diminuir riscos às culturas.

Agradecimentos
As autoras expressam seus agradecimentos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de mestrado, apoiando financeiramente às pesquisas interdisciplinares. E, ao Programa de Pós-Graduação em Computação Aplicada pela infraestrutura e apoio incondicionais.

Referências

CHEN, N. et al. **Diagrama de constelação tridimensional com um design de nível hierárquico para transmissão multinúcleo.** Optics express. 2022.

CIMA, G. et al. **Abstração na integração de dados.** 36º Simpósio Anual ACM/IEEE sobre Lógica em Ciência da Computação (LICS). 2021.

COGATO, A.; et al. **Eventos climáticos extremos na agricultura: uma revisão sistemática.** Sustentabilidade. 2019.

DU, Y. E SHIUE, M. **Análise e computação de algoritmos de assimilação contínua de dados para o sistema Lorenz 63 com base em técnicas de nudging não lineares.** Journal of Computer Applied in Mathematics. 2021.

ELAHI, E. et al. **Eventos climáticos extremos representam risco para a produção agrícola e a adaptação de estratégias de gestão inovadoras para mitigar o risco: uma pesquisa retrospectiva da área rural de Punjab, Paquistão.** Technovation. 2021.

GOLDSTEIN, A.; FINK, L. E RAVID, G. **Uma estrutura baseada em nuvem para integração de dados agrícolas: uma abordagem de cima para baixo e de baixo para cima.** IEEE Access. 2022.

MICROSOFT LEARN. **ETL (extração, transformação e carregamento)**. EUA: Microsoft, 2024.

MONTALVO-ROMERO, N. et al. **Sistemas Agrotecnológicos na Assistência à Agricultura Tradicional: Uma Revisão Sistemática**. IEEE Acess. 2023.