

Métodos Preditivos Espaço-Temporais e sua Aplicação na Agronomia

Marcos Vinícius Pretti Dias¹, Daniel dos Santos Kaster¹

¹ Departamento de Computação – Universidade Estadual de Londrina (UEL)
Caixa Postal 10.011 – CEP 86057-970 – Londrina – PR – Brasil

marcosvinicius.pretti@uel.br, dskaster@uel.br

Abstract. *The study of spatiotemporal predictive methods holds great significance for agronomy, particularly for predicting diseases in agricultural crops such as Asian soybean rust, one of the main diseases affecting soybean production and causing considerable economic losses. This work aims to study and compare existing spatiotemporal predictive models, applying them to agronomic problems with a focus on predicting Asian soybean rust. Advanced data analysis and machine learning techniques, including statistical modeling, artificial neural networks, and time series analysis, will be used to assess the effectiveness of these models on historical data such as climate, soil, and disease incidence. The expected results aim to provide greater accuracy in disease prediction, contributing to more efficient management strategies and reduction of losses in the production of various crops.*

Resumo. *O estudo de métodos preditivos espaço-temporais é de grande relevância para a agronomia, especialmente para a previsão de doenças em culturas agrícolas, como a ferrugem asiática da soja que é uma das principais doenças que afetam a produção dessa leguminosa, causando consideráveis prejuízos econômicos. Este trabalho tem como objetivo estudar e comparar modelos preditivos espaço-temporais já existentes, visando aplicar esses modelos a problemas agrônômicos, com foco na previsão da ferrugem asiática da soja. Técnicas avançadas de análise de dados e aprendizado de máquina, incluindo modelagem estatística, redes neurais artificiais e análise de séries temporais, serão utilizadas para avaliar a eficácia desses modelos em dados históricos como de clima, solo e incidência da doença. Espera-se que os resultados proporcionem uma maior precisão na previsão de doenças, contribuindo para estratégias de manejo mais eficientes e redução de perdas na produção de diferentes culturas.*

1. Introdução

Com a evolução da coleta de dados e formas de armazenamento, dados espaço-temporais têm crescido de forma exponencial, e com isso vários métodos de previsão tem surgido (Sun et al., 2024)[17]. A previsão de séries espaço-temporais (ST) tornou-se uma área de estudo vital para enfrentar desafios como a ferrugem asiática da soja, uma das principais doenças que afetam essa cultura e resultam em significativas perdas econômicas. A ferrugem asiática da soja é causada pelo fungo *Phakopsora pachyrhizi*, que pode causar desfolha precoce, resultando em redução da fotossíntese e, conseqüentemente, em perdas significativas na produtividade da soja (Godoy et al., 2016)[7].

Com o avanço das técnicas de aprendizado de máquina e análise de dados, novos modelos preditivos têm sido desenvolvidos para capturar padrões complexos em dados espaço-temporais. A literatura recente tem mostrado que redes neurais artificiais, especialmente redes neurais convolucionais (CNNs) e redes neurais recorrentes (RNNs), são particularmente eficazes para a predição de séries temporais devido à sua capacidade de capturar dependências temporais e espaciais (Zhang et al., 2020)[22]. Por exemplo, a utilização de CNNs para extrair características espaciais combinada com RNNs para capturar dependências temporais tem mostrado resultados promissores em várias aplicações, incluindo previsão de tráfego e predição de demanda energética (Wang et al., 2019; Li et al., 2018)[21].

No contexto agrônomico, a predição precisa de doenças em culturas agrícolas é essencial para o manejo eficaz e a mitigação de perdas. Estudos têm demonstrado que a integração de dados meteorológicos, condições do solo e incidência de doenças em modelos de aprendizado de máquina pode melhorar significativamente a acurácia das previsões (Chlingaryan et al., 2018)[5]. Por exemplo, o uso de modelos baseados em aprendizado profundo para a previsão de ferrugem do trigo utilizando dados climáticos históricos tem se mostrado eficaz na identificação de padrões que precedem surtos da doença (Liu et al., 2020)[14].

Este trabalho se propõe a comparar a eficácia de diferentes abordagens preditivas, incluindo modelos de *machine learning* tradicionais e métodos de aprendizado profundo, como CNNs e RNNs, na predição da ferrugem asiática da soja. Modelos baseados em redes grafo neurais (GNNs) também serão explorados devido à sua capacidade de capturar relações complexas em dados espaciais (Wu et al., 2020)[20] adicionalmente, a utilização de técnicas de ensemble learning, que combinam múltiplos modelos preditivos para melhorar a robustez e acurácia das previsões, será investigada. Estudos recentes sugerem que abordagens de ensemble podem superar modelos individuais ao lidar com a variabilidade e incerteza inerentes aos dados espaço-temporais (Dietterich, 2000)[6].

A metodologia deste estudo inclui a coleta e pré-processamento de dados históricos de clima, solo e incidência de ferrugem asiática da soja. Serão utilizadas técnicas de análise exploratória de dados para identificar padrões e correlações nos dados, seguidas pela aplicação e avaliação de diferentes modelos preditivos estes modelos podem incluir algoritmos de aprendizado de máquina, redes neurais e outras técnicas estatísticas. A eficácia dos modelos será avaliada utilizando métricas de desempenho como erro médio absoluto (MAE), raiz do erro quadrático médio (RMSE) e coeficiente de determinação (R^2).

Espera-se que os resultados deste estudo não apenas melhorem a precisão na predição da ferrugem asiática da soja, mas também forneçam *insights* valiosos para a aplicação de técnicas de aprendizado de máquina em outros problemas agrônomicos. A implementação de modelos preditivos eficazes pode contribuir significativamente para estratégias de manejo mais eficientes, ajudando agricultores a tomar decisões informadas e reduzir perdas na produção.

2. Fundamentação Teórico-Metodológica e Estado da Arte

2.1. Predição espaço-temporal

A predição espaço-temporal é um campo de estudo que se concentra na previsão de valores futuros de uma variável ou conjunto de variáveis que variam tanto no espaço quanto no tempo. Portanto, envolve prever como um fenômeno ou processo se desenvolverá ao longo do tempo e em diferentes locais geográficos. Os principais componentes desse tipo de projeção são uma estimativa no espaço que se refere às características geográficas ou posições físicas onde os dados são coletados ou onde os eventos ocorrem. Pode ser representado por coordenadas geográficas (latitude e longitude), regiões administrativas (cidades, estados, países), ou outras formas de segmentação espacial dependendo do problema. Já o componente do tempo considera a mudança dos dados num período específico de tempo podendo ser segundos, minutos, horas, dias ou qualquer outra unidade de tempo. Para realizar previsões no espaço e no tempo, foram desenvolvidos diversos métodos que utilizam conceitos de diferentes áreas do conhecimento. Esses métodos se beneficiam de abordagens interdisciplinares, combinando técnicas estatísticas, algoritmos de machine learning, modelagem baseada em física e análise de séries temporais. A integração dessas diversas áreas permite a criação de modelos robustos e precisos, capazes de capturar a complexidade inerente aos dados espaço-temporais.

2.2. Modelos baseados em aprendizado de máquina tradicional

Algoritmos como *K-Nearest Neighbor* (KNN), *support vector machines* (SVM) e baseados em árvores de decisão são comumente utilizados para problemas espaço-temporais [23] pela sua alta capacidade de manejo com muitos dados, entretanto é importante ressaltar que muito de seu desempenho é diretamente afetado pela forma que são extraídas as características a partir dos dados [17].

Descrição de cada um destes modelos

- KNN é um algoritmo de aprendizado supervisionado que classifica uma amostra com base nas categorias dos seus K vizinhos mais próximos, porém ele não faz suposições sobre a distribuição dos dados. No contexto espaço-temporal o KNN pode usar tanto características espaciais quanto temporais, considerando a proximidade geográfica e a similaridade temporal para fazer previsões.
- SVM encontra um hiperplano que separa os dados em diferentes classes com a máxima margem, se torna eficiente para altas dimensionalidades e pode ser usado para regressão e classificação, para previsões espaço-temporais ele pode ser estendido para lidar com dados espaço-temporais incorporando características temporais e espaciais nas amostras de entrada, como demonstrado em estudos de previsão de tráfego e padrões climáticos.
- Árvores de decisão dividem o espaço de características em regiões homogêneas, utilizando esta estrutura de árvore onde cada nó representa uma decisão baseada em um atributo para predição espaço-temporal, árvores de decisão podem ser utilizadas para identificar padrões de decisão baseados em atributos espaciais e temporais, sendo eficazes principalmente em dados com relações complexas.

Modelos de aprendizado de máquina tradicional, como *K-Nearest Neighbor* (KNN), *Support Vector Machines* (SVM), e árvores de decisão, Dentro da agronomia

são especialmente utilizados para gestão de recursos, detecção e classificação de pragas, e previsões de colheita. KNN, por exemplo, pode ser usado para classificação de doenças na soja [13], enquanto SVM pode ser usada para um melhor gerenciamento de culturas [11]. Árvores de decisão são empregadas para diversas áreas como eficiência energética [1] e agricultura de precisão [18].

2.3. Modelos baseados em *deep learning*

O campo de predição espaço-temporal com o uso de deep learning tem atraído atenção significativa, com aplicações em diversas áreas como a previsão de tráfego [10]. Revisões recentes destacam o uso de arquiteturas específicas, como redes neurais recorrentes (RNNs) [9] e modelos baseados em grafos [15]. Lim e Zohren (2021) [12] discutem de forma abrangente os principais componentes das redes neurais, abordando, porém, de forma limitada a extração de características espaciais. Benidis et al. (2022) [2] forneceram um tutorial detalhado sobre previsão de séries temporais, mas sem um foco específico em redes neurais de grafos (GNNs). Já Wang, Cao, e Philip (2020) [19] exploraram métodos de predição de séries temporais multivariadas baseados em *deep learning*, sugerindo estratégias para predições em múltiplos passos (*multi-step*). O estado da arte atual demonstra o potencial das técnicas de aprendizado profundo, mas também aponta para a necessidade de mais pesquisas que integrem de maneira eficaz características espaciais e temporais, além de explorarem novas arquiteturas e modelos para aprimorar a precisão e aplicabilidade das previsões. A seguir, serão descritos brevemente os algoritmos abordados neste trabalho, com ênfase nas suas características e aplicações em contextos espaciais e/ou temporais. Esta análise visa fornecer uma compreensão clara de como cada algoritmo contribui para a predição espaço-temporal.

Breve Descrição dos Algoritmos

- **Redes Neurais Recorrentes (RNNs)** As RNNs são uma classe de redes neurais especialmente eficazes para dados sequenciais, devido à sua capacidade de manter informações de estados anteriores. Elas são amplamente utilizadas em previsão de séries temporais e processamento de linguagem natural. Autores como [8] dão detalhes sobre o funcionamento das RNNs e suas variantes
- **Redes Neurais de Grafos Espaciais-Temporais (ST-GNNs)** Os modelos de grafos espaciais-temporais utilizam a estrutura de grafos para capturar dependências espaciais e temporais em dados complexos, como tráfego de rede ou atividades cerebrais. Essas técnicas são particularmente úteis em cenários onde a topologia espacial é uma variável crítica [3].
- **Redes Neurais Convolucionais (CNNs) e Modelos Híbridos** As CNNs, comumente usadas para análise de imagens, também são aplicáveis a dados espaço-temporais quando combinadas com RNNs ou LSTMs, formando modelos híbridos. Essas combinações permitem capturar tanto características espaciais quanto temporais, funcionamento destes algoritmos também são abordados por [8].

Estudos recentes têm demonstrado a eficácia de redes neurais profundas em aplicações agrícolas. Por exemplo, o uso de redes neurais informadas por princípios físicos (PINNs) tem se mostrado promissor na modelagem morfológica da ferrugem asiática da soja, proporcionando *insights* sobre a progressão da doença e a eficácia dos tratamentos [4].

Além disso, técnicas de classificação de doenças por aprendizado profundo, como o uso de arquiteturas U-Net para identificar ferrugem em culturas, têm melhorado significativamente a precisão da detecção e a resposta a doenças [16].

3. Objetivos

Objetivo Geral

No trabalho atual, o propósito é analisar e estudar cada método de predição de forma individual para identificar suas vantagens e os problemas que cada modelo pode resolver, especificamente no contexto da agronomia.

Objetivos Específicos

- Pretende-se comparar essas diferentes arquiteturas para determinar qual apresenta melhor desempenho em problemas agrícolas, como a previsão de doenças em culturas.
- Além disso, busca-se explorar como essas técnicas podem ser combinadas para abordar problemas agrícolas de forma mais eficaz, utilizando o potencial de cada abordagem para melhorar a produtividade e a sustentabilidade na agricultura.

4. Procedimentos metodológicos/Métodos e técnicas

O plano de trabalho para a elaboração do TCC começará com a coleta e o pré-processamento de dados, seguido de uma revisão abrangente da literatura no campo das predições espaço-temporais. Posteriormente, serão implementados e aplicados modelos preditivos, com uma análise detalhada das possibilidades de integração desses modelos para abordar diferentes etapas de um mesmo problema. O objetivo é identificar a abordagem de modelagem mais adequada para cada tipo de problema estudado. Com os resultados obtidos, será realizado um estudo comparativo entre os diferentes métodos, destacando suas aplicações práticas, como no caso da movimentação dos esporos de ferrugem asiática. Esse estudo comparativo permitirá uma melhor compreensão da eficácia e das limitações de cada modelo, fornecendo *insights* valiosos para a aplicação prática dos resultados em cenários reais.

5. Cronograma de Execução

Atividades:

1. Levantamento bibliográfico;
2. Coleta de dados;
3. Pre-processamento dos dados;
4. Estudo e implementação dos métodos propostos;
5. Verificar a possível integração entre os métodos;
6. Comparação de resultados;
7. Análise dos resultados em uma aplicação prática;
8. Escrita da versão final do TCC;

Tabela 1. Cronograma de Execução

	jun	jul	ago	set	out	nov	dez	jan	fev
Atividade 1	•								
Atividade 2		•							
Atividade 3			•						
Atividade 4			•	•	•				
Atividade 5					•				
Atividade 6						•			
Atividade 7							•		
Atividade 8								•	•

6. Contribuições e/ou Resultados esperados

Este estudo tem como objetivo principal elucidar o funcionamento e a aplicação de diversos métodos preditivos em problemas que envolvem variáveis espaço-temporais. Espera-se demonstrar aos leitores quais arquiteturas de modelagem são mais eficazes para abordar problemas específicos nesse contexto. Além disso, o estudo busca contribuir significativamente para a área agrônômica, evidenciando a eficácia desses algoritmos na previsão de desafios agrícolas, como proliferação de doenças, padrões de cultivo, e gestão de recursos. A pesquisa pretende não apenas identificar os modelos mais adequados, mas também explorar como esses métodos podem ser integrados e aplicados na prática para melhorar a tomada de decisões e a eficiência na agricultura.

7. Espaço para assinaturas

Londrina, 29 de julho de 2024.

Aluno

Orientador

Referências

- [1] Omid Abrishambaf, Pedro Faria, and Zita Vale. Energy resource scheduling in an agriculture system using a decision tree approach. In *2019 20th International Conference on Intelligent System Application to Power Systems (ISAP)*, pages 1–5, 2019.
- [2] K. Benidis et al. Deep learning for time series forecasting: tutorial and literature survey. *ACM Computing Surveys*, 55(6):1–36, 2022.
- [3] M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst. *Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges*. arXiv preprint arXiv:1611.08097, 2017.
- [4] Henry Cavanagh, Andreas Mosbach, Gabriel Scalliet, Rob Lind, and Robert Endres. Physics-informed deep learning characterizes morphodynamics of asian soybean rust disease. *Nature Communications*, 12, 11 2021.

- [5] Anna Chlingaryan, Salah Sukkarieh, and Brett Whelan. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*, 151:61–69, 08 2018.
- [6] Thomas G Dietterich. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, pages 1–15. Springer, Berlin, Heidelberg, 2000.
- [7] Claudio V Godoy et al. Asian soybean rust in brazil: past, present, and future. *Pesquisa Agropecuária Brasileira*, 51(5):407–421, 2016.
- [8] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- [9] H. Hewamalage, C. Bergmeir, and K. Bandara. Recurrent neural networks for time series forecasting: current status and future directions. *International Journal of Forecasting*, 37(1):388–427, 2021.
- [10] Y. Hou et al. Deep learning methods in short-term traffic prediction: a survey. *Information Technology and Control*, 51(1):139–157, 2022.
- [11] Zhi Hong Kok, Abdul Rashid Mohamed Shariff, Meftah Salem M. Alfatni, and Siti Khairunniza-Bejo. Support vector machine in precision agriculture: A review. *Computers and Electronics in Agriculture*, 191:106546, 2021.
- [12] B. Lim and S. Zohren. Time-series forecasting with deep learning: a survey. *Philosophical Transactions of the Royal Society A*, 379(2194):209, 2021.
- [13] Chandrakant Mahobiya, Sailesh Iyer, and Savita Kolhe. A novel approach for detection of soybean leaf disease using bayesian optimized-knn classifier with nca-based feature selection. In Arti Noor, Kriti Saroha, Emil Pricop, Abhijit Sen, and Gaurav Trivedi, editors, *Proceedings of Third Emerging Trends and Technologies on Intelligent Systems*, pages 261–279, Singapore, 2023. Springer Nature Singapore.
- [14] Qian Pan, Maofang Gao, Pingbo Wu, Jingwen Yan, and Shilei Li. A deep-learning-based approach for wheat yellow rust disease recognition from unmanned aerial vehicle images. *Sensors*, 21(19), 2021.
- [15] Z. A. Sahili and M. Awad. Spatio-temporal graph neural networks: a survey. *arXiv preprint arXiv:2301.10569*, 2023.
- [16] Shivani Sood, Harjeet Singh, and Suruchi Jindal. Rust disease classification using deep learning based algorithm: The case of wheat. In Ana I. Ribeiro-Barros, Daniel S. Tevera, Luís F. Goulao, and Lucas D. Tivana, editors, *Food Systems Resilience*, chapter 10. IntechOpen, Rijeka, 2022.
- [17] Feiyan Sun, Wenning Hao, Ao Zou, and Qianyan Shen. A survey on spatio-temporal series prediction with deep learning: taxonomy, applications, and future directions. *Neural Computing and Applications*, 36:9919–9943, 2024.
- [18] Jérôme Treboux and Dominique Genoud. High precision agriculture: An application of improved machine-learning algorithms. In *2019 6th Swiss Conference on Data Science (SDS)*, pages 103–108, 2019.
- [19] S. Wang, J. Cao, and S. Y. Philip. Deep learning for spatiotemporal data mining: a survey. *IEEE Transactions on Knowledge and Data Engineering*, 34(8):3681–3700, 2020.

- [20] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1):4–24, 2021.
- [21] Zhen Yan, Hongyu Yang, Fan Li, and Yi Lin. A deep learning approach for short-term airport traffic flow prediction. *Aerospace*, 9(1), 2022.
- [22] Bing Yu, Haoteng Yin, and Zhanxing Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-2018*. International Joint Conferences on Artificial Intelligence Organization, July 2018.
- [23] Wei Zhang, Yong Zhang, Zhihui Song, and Yuhong Sun. Spatiotemporal prediction and mapping of heavy metals at regional scale using regression methods and landsat 7. *Remote Sensing*, 13(24):4615, 2021.