

Análise da variabilidade espacial, temporal e previsão de resultado de produção agrícola com redes Bayesianas dinâmicas

Guilherme Afonso Halal, discente de graduação, Universidade Federal do Pampa, Campus Bagé

Ana Paula Lüdtke Ferreira, docente, Universidade Federal do Pampa

guilhermehalal.aluno@unipampa.edu.br

A agricultura é considerada por muitos historiadores como a principal tecnologia que moldou o desenvolvimento humano, permitindo a concentração populacional e o surgimento das sociedades modernas. A segurança alimentar depende de que a agricultura supere os desafios que o crescimento populacional e as mudanças climáticas impõem. A tecnologia da informação é uma aliada importante para a superação destas dificuldades de maneira sustentável, incrementando a produtividade com melhor aproveitamento dos recursos já disponíveis. Neste contexto, modelos de inferência probabilística são técnicas que podem prover os melhores resultados, visto que são modelos que podem ser gerados tanto por dados da realidade quanto por especialistas, são explicáveis e podem incorporar novos dados à medida que eles são produzidos. O objetivo deste trabalho é desenvolver um pacote R, com base em redes Bayesianas dinâmicas, que permita a previsão de resultados de colheita, levando em consideração a variabilidade espacial e temporal das culturas, a partir de divisões da área de produção que apresentem características de produção similares e das fases fenológicas da cultura. Uma rede Bayesiana é um modelo probabilístico que computa inferências por meio de grafos acíclicos dirigidos, onde os nós correspondem às variáveis do sistema e os arcos entre os nós representam relações causais existentes entre as variáveis. O valor de cada arco é computado como a probabilidade condicional da variável destino dados os valores da variável de origem. Uma rede Bayesiana dinâmica é construída a partir de instâncias repetidas de uma rede Bayesiana estática em que as dependências temporais são arcos entre os nodos de diferentes instâncias. O método usado para o desenvolvimento do trabalho incluiu uma revisão de escopo da literatura sobre previsão de resultados de colheita e aplicação de modelos probabilísticos na agricultura. As questões de pesquisa da revisão foram: i) quais são as técnicas utilizadas na previsão de resultado de colheita?; ii) existem implementações que utilizam modelos probabilísticos como abordagem a temas ligados à agricultura? O levantamento foi realizado a partir dos termos chave: previsão, agricultura, colheita, variabilidade e modelos probabilísticos, com suas respectivas traduções para Inglês e termos sinônimos. As fontes de referências bibliográficas utilizadas foram *ACM Digital Library*, *IEEE Xplorer*, *ScienceDirect* e *Wiley Online Library*, que são as mais representativas das áreas de conhecimento relacionadas. Os critérios de inclusão consideraram trabalhos publicados a partir de janeiro de 2016 até a data da realização das buscas, em fevereiro de 2021; trabalhos relativos à previsão de resultado de colheita, independente da técnica utilizada; trabalhos que abordem análises sobre variabilidade espacial ou temporal de colheitas; trabalhos que usem modelos probabilísticos na agricultura, mesmo que fora do domínio de previsão de resultados de colheita. As buscas nos repositórios listados obtiveram 993 resultados.

Após a realização de duas triagens, observando os títulos dos artigos e a leitura dos resumos dos trabalhos, eliminaram-se as publicações que não tratam das questões de pesquisa, restando 59 publicações. A previsão de produtividade agrícola foi abordada em 13 trabalhos que aplicam técnicas de regressão, 15 trabalhos que empregam técnicas de aprendizado de máquina e 4 que utilizam outras técnicas que não se enquadram em nenhuma das duas categorias anteriores. Os trabalhos que utilizam modelos probabilísticos aplicados à agricultura usaram modelos de Markov, com 13 trabalhos levantados e redes Bayesianas, com 11. Identificação de padrões é a aplicação com maior recorrência encontrada nos trabalhos de modelos de Markov e modelos de apoio à tomada de decisão e previsão de resultado de produção são mais recorrentes nas aplicações de redes Bayesianas. O levantamento bibliográfico permitiu verificar que nossa abordagem é inovadora. O modelo proposto neste trabalho faz uso de redes dinâmicas, em que cada instância estática representa os estados de variáveis relevantes à produtividade agrícola em diferentes estágios do desenvolvimento da cultivar. A ideia é capturar como as diferentes variáveis do sistema solo-planta-atmosfera influem no desenvolvimento de cada cultura em cada fase fenológica. Para garantir a validade estatística dos resultados, estamos gerando os dados de treinamento de forma que possamos avaliar a correção dos resultados. No pacote de software desenvolvido, o usuário poderá entrar com um conjunto de áreas de produção e suas respectivas variáveis. O pacote construirá as redes, exibindo as previsões de produção, gráficos e relatórios explicativos, a partir dos dados inseridos ao longo do crescimento da cultura. O pacote AgroGraph será registrado e publicado no repositório padrão da linguagem R (<https://cran.r-project.org/>).

Palavras-chave: Agricultura digital; Agricultura de precisão; AgroGraph; Inferência probabilística; Ciência de dados.