

## Modelo Matemático para classificação de culturas em Mato Grosso utilizando NDVI/MODIS

**Kelly Marques de Oliveira Lopes**      **Laércio Luis Vendite**

Depto de Matemática Aplicada, IMECC, UNICAMP,  
13083-970, Campinas, SP  
E-mail: kelly@ime.unicamp.br, vendite@ime.unicamp.br,

**Stanley Robson M.Oliveira**

Embrapa Informática Agropecuária,  
13083-970, Campinas, SP  
E-mail: stanley.oliveira@embrapa.br.

### RESUMO

O desenvolvimento de estudos na área de geotecnologia e o aumento na capacidade de armazenar dados têm melhorado a exploração e os estudos de imagens de satélites obtidas através de sensores orbitais. O mapeamento da cobertura da terra, estimativas de produtividade de culturas e a previsão de safras são informações importantes para o agricultor e para o governo, pois essas informações são essenciais para subsidiar decisões relacionadas à produção, estimativas de compra e venda, e cálculos de importação e exportação. Uma das alternativas para analisar dados de uso e cobertura da terra, obtidos por meio de sensores, é o uso de técnicas de mineração de dados, uma vez que essas técnicas podem ser utilizadas para transformar dados e informações em conhecimentos que irão subsidiar decisões relativas ao planejamento agrícola. Neste trabalho, foram utilizados dados multitemporais sobre o índice de vegetação NDVI, derivados de imagens do sensor MODIS, para o monitoramento das culturas de algodão, soja e milho no estado do Mato Grosso, no período do ano-safra de 2008/2009. O conjunto de dados, fornecido pela Embrapa Informática Agropecuária, foi composto por 24 colunas e 728 linhas, onde as 23 primeiras colunas referem-se aos valores do NDVI, e a última, à cobertura do solo. A metodologia utilizada teve como base o modelo CRISP-DM (Cross Industry Standard Process for Data Mining). Modelos preditivos para classificar dados sobre essas culturas foram elaborados e avaliados por algoritmos de aprendizado de máquina, tais como árvores de decisão (J48 e PART), florestas aleatórias (Random Forest). A seleção de atributos melhorou os valores do índice *Kappa* e a acurácia dos modelos. Foram geradas regras de classificação para mapear as culturas estudadas (soja, milho e algodão). Os resultados revelaram que os algoritmos de aprendizado de máquina são promissores para o problema de classificação de cobertura do solo. Em particular o algoritmo J48, utilizado em conjunto com a seleção de atributos feito por meio de Análise de Componentes Principais, destacou-se em relação ao demais pela simplicidade e pelos valores apresentados. Os resultados também evidenciaram a presença de regiões de cultivo do algodão em outras áreas do estado, fora daquelas estudadas.

**Palavras-chave:** *Mineração de dados, Imagens de satélites, Modelos preditivos, Algoritmos de aprendizado de máquina, Mapeamento da cobertura da terra, NDVI.*

### Referências

- [1] E. L. D. M. Novo, Sensoriamento remoto: princípios e aplicações. Edgard Blucher, 2008.
- [2] J. K. HAN, AND P. JIAN, Data Mining: concepts and techniques. Morgan Kaufmann Publishers, San Francisco, CA, USA, 2005.

- [3] R. M. HOFFER, Biological and physical considerations in applying computer aided analysis techniques to remoto sensor data. Mc Graww-Hill, 1978.
- [4] I. H. WITTEN, AND E. FRANK, Data Mining: Pratical machine learning tools and techniques. Morgan Kaufmann, 2005.