

PREDIÇÃO DE RANQUEAMENTO DE LOTES DE SEMENTE DE MILHO POR INTELIGÊNCIA ARTIFICIAL

RUAN BERNARDY¹; NICACIA ANDRADE BORGES MORAES²;
RITA DE CASSIA MOTA MONTEIRO²; ROMARIO DE MESQUITA PINHEIRO²;
GIZELE INGRID GADOTTI³

¹Universidade Federal de Pelotas – ruanbernardy@yahoo.com.br

²Universidade Federal de Pelotas

³Universidade Federal de Pelotas– gizeleingrid@gmail.com

1. INTRODUÇÃO

Para o setor sementeiro, ainda se encontram vários desafios na tomada de decisão rápida e precisa quando se trabalha com significativa quantidade de lotes de sementes, sendo que a manipulação manual se torna quase impossível. O ranqueamento de lotes de sementes para comercialização de uma empresa é fundamental no quesito da rapidez de despacho de lotes para os produtores, porém é necessário informar a qualidade desses lotes. Além disso, uma abordagem entre a produção de grãos que desempenha um papel importante na economia global. Nesse sentido, a demanda por métodos eficientes e seguros de produção de alimentos está aumentando. A tecnologia da informação é uma das ferramentas para tal fim (PATRÍCIO; RIEDER, 2018).

Durante o controle de qualidade de lotes de sementes, os requisitos mínimos serão as exigências legais, proveniente dos resultados de testes que realizam a análise de sementes. Isso gera uma quantidade substancial de informações, que dependendo do porte da empresa somente em uma safra agrícola poderá ter milhares de dados.

Assim, em grandes produtores de sementes se faz necessário um profissional que analise esses dados e classifique esses lotes para que o pessoal de vendas possa trabalhar e a posteriori o pessoal da expedição. Geralmente, devido ao fato de sua responsabilidade ser grande, pois um erro pode fazer a empresa perder seu prestígio e ainda ter um prejuízo financeiro, esse tem um cargo e salário mais alto e altamente especializado.

Nesse contexto, o objetivo deste trabalho foi utilizar o aprendizado de máquinas para a classificação de lotes de sementes de milho.

2. METODOLOGIA

Foram utilizados dados provenientes de um produtor de sementes de milho, sendo os atributos oito safras (14/15, 15/15, 15/16, 16/16, 16/17, 17/17, 17/18, 18/18) e quatro cultivares (Tabela 1), totalizando 5800 lotes de sementes (linhas).

Os atributos utilizados foram relacionados a questões legais (germinação, pureza física, número de outras sementes, porcentagem de sementes infestadas) e outros de identificação do lote (material, peneira, safra).

O pré-processamento utilizado foi verificar dados discrepantes, retirar dados incoerentes e linhas sem dados, trocar vírgulas por pontos e tornar os dados para leitura no software utilizado.

O arquivo de treinamento foi utilizado com 80 *Reject* (34%) e 237 *Accept* (66%) porque os dados possuem muitos aceites e poucos rejeitados e este foi uma forma de balancear os mesmos.

Esses dados foram analisados primeiramente em cultivares e safras em separado. A seguir, a análise foi de cada cultivar com todas as safras e finalmente de todas as cultivares e safras unidas.

Os classificadores utilizados foram J48, Random Forest, Classification Via Regression (CVR), IBk, Multilayer Perceptron (MLP) e Naive Bayes. Foi utilizada a validação cruzada, no qual se dividiu o conjunto de dados, treinamento e teste, em 10 subconjuntos. A média dessas precisões corresponde ao desempenho do algoritmo sobre o conjunto de dados fornecido.

Essa técnica reduz a probabilidade de que coincidências subavaliem ou sobreavaliem o desempenho para uma determinada configuração. Todos os resultados reportados neste trabalho utilizam essa técnica. As etapas descritas foram realizadas no software Weka.

Após a escolha do classificador, por meio da melhor acurácia obtida no teste de treinamento, o modelo foi utilizado com os dados totais. Para a realização de cluster e assim realizar uma avaliação não supervisionada foi realizada com os algoritmos SimpleKMeans e FarthestFirst.

3. RESULTADOS E DISCUSSÃO

Aplicando a técnica de classificação e verificando as acurácias dos algoritmos (Tabela 1), os mais indicados foram o J48 e CVR. Analisando suas matrizes de confusão (Tabela 2), verifica-se que são idênticas e de grande precisão, pois os dados são classificados perfeitamente e sem falsos positivos ou negativos.

Tabela 1. Acurácia dos algoritmos após a classificação

Algoritmo	Acurácia (%)
J48	100,0000
Random Forest	99,6845
CVR	100,0000
lazy.IBk	90,8517
MLP	98,1073
Naive Bayes	96,5300

a	b	< classificado como
237	0	a = Accept
0	80	b = Reject

Tabela 2. Matriz de confusão do algoritmo J48 e do CVR.

Os dados demonstram que, para sementes de milho, o número de aceites dos lotes é alto com baixa rejeição. Com isso se esperava que houvesse um crivo maior sobre o atributo vigor, o que ocorreu conforme a Figura 1.

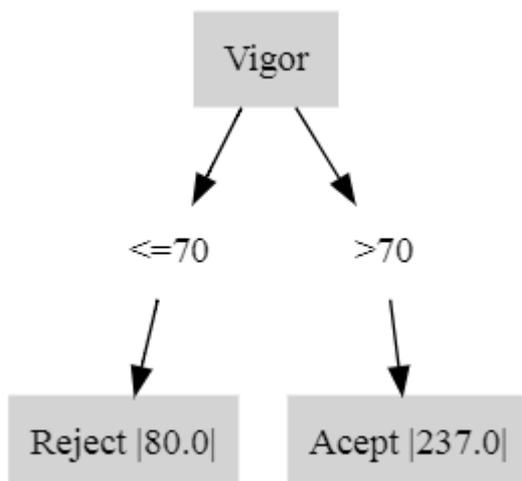


Figura 1. Árvore de decisão para predição de classificação de lotes de sementes de milho

A árvore de decisão gerada demonstra que o atributo que mais possui influência na classificação de um lote de sementes de milho é o atributo vigor. No entanto, para a área de Ciência e Tecnologia de Sementes seria interessante um maior número de atributos ligados ao vigor porque legalmente só há um atributo: germinação. Como o teste de frio não é padronizado e no caso deste produtor só utilizou um único teste, de acordo com a teoria se visualiza a importância de mais testes de vigor para uma classificação mais eficiente de lotes conforme já escrito por Tillmann et al. (2003). Já na empresa, há um histórico sobre o comportamento dessa espécie e de todos os materiais, fazendo que com apenas um teste a empresa tenha resultados satisfatórios e confiáveis.

Quando se realiza a predição sem supervisão, verifica-se a acurácia nos clusters, sendo que o algoritmo FarthestFirst gera maior certeza.

Tabela 3. Acurácia dos algoritmos após o cluster

Algoritmo	Incerteza (%)
SimpleKMeans	32,8076
FarthestFirst	29,6530

É importante visualizar que no treinamento dos agrupamentos, ambos os algoritmos se comportam da maneira semelhante e colocando alguns lotes no agrupamento equivocado, o que é considerado de boa precisão (Tabela 3).

Tabela 3. Agrupamentos gerados da classificação de lotes de milho.

SimpleKMeans	
0	1
0	283 (89%)
1	34 11 (11%)
FathersFirst	
0	281 (89%)
1	36 (11%)

4. CONCLUSÕES

É possível classificar lotes de sementes de milho com grande acurácia e precisão por meio de inteligência artificial e sua técnica de aprendizado de máquina. A inteligência artificial permitirá automatizar processos na empresa de sementes e aumentar a velocidade de conclusão de tarefas.

5. REFERÊNCIAS BIBLIOGRÁFICAS

PATRICIO, D.I.; RIEDER, R. **Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review**. Computers and Electronics in Agriculture, 2018, v.153, p.69-81. DOI: <https://doi.org/10.1016/j.compag.2018.08.001>.

ABRASEM - Associação Brasileira de Sementes e Mudas. **Estatísticas**. Disponível em: <<https://www.conab.gov.br/info-agro/safras/graos>>. Acesso em: 20 nov. 2019.

CARVALHO, N.M.; NAKAGAWA, J. **Sementes: ciência, tecnologia e produção**. 4.ed. Jaboticabal: FUNEP, 2000. 524p.

FINLAY, S. **Artificial intelligence for everyone. Great Britain: Relativistic**. 189p. 2020.

EIBE FRANK, Mark A. Hall, and Ian H. Witten (2016). **The WEKA Workbench**. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.

GRZYBOWSKI, C. R. S.; VIEIRA, R. D.; PANOBIANCO, M. Testes de estresse na avaliação do vigor de sementes de milho. **Revista Ciência Agrônômica**, v.46, n.3, p.590-596, 2015. <https://doi.org/10.5935/1806-6690.20150042>.

KUMAR, S. **Reinventing Agri-produce quality testing using AI**. National Workshop on Artificial Intelligence in Agriculture. Indian Council of Agricultural Research. New Delhi. p.44-45. 2018.

MAGALHÃES NETO, W.; GADOTTI, G. I.; VILLELA, F. A.; MARTINS, A. B. N.; CARVALHO, I. R.; CAVALCANTE, J. A. Critérios de escolha de sementes de milho pelos agricultores da região de Votuporanga-SP. **Revista Brasileira de Engenharia e Sustentabilidade**, v.5, n.2, p.43-50, dez. 2018.

MARCOS FILHO, J. Teste de vigor: importância e utilização. In: KRZYŻANOWSKI, F.C.; VIEIRA, R.D.; FRANÇA NETO, J.B. (Ed.). **Vigor de sementes: conceitos e testes**. Londrina: ABRATES, 1999. 218p. Cap.1, p.1-21

Tillmann M.A.A. Tunes L.M. Almeida A.S. (2019). Análise de sementes. In: Peske ST, Villela FA, Meneghello GE. **Sementes: fundamentos científicos e tecnológicos**. Editora Universitária. Universidade Federal de Pelotas, 2019, p.147-258.