

ANÁLISE DO DESEMPENHO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA CLASSIFICAÇÃO DE LOTES DE SEMENTES DE SOJA

ROMÁRIO DE MESQUITA PINHEIRO¹; GIZELE INGRID GADOTTI²; RUAN BERNARDY³; CARLA APARECIDA ASCOLI¹

¹Programa de Pós-Graduação em Ciência e Tecnologia de Sementes, Universidade Federal de Pelotas – e-mail romario.ufacpz@hotmail.com

²Centro de Engenharias e Programa de Pós-Graduação em Ciência e Tecnologia de Sementes Universidade Federal de Pelotas – e-mail gizele.gadotti@ufpel.edu.br

³Programa de Pós-Graduação em Ciências Ambientais da Universidade Federal de Pelotas – e-mail ruanbernardy@yahoo.com.br

1. INTRODUÇÃO

A estimativa do desempenho da produtividade para produtos agrícolas é um tópico de interesse na agricultura (SARIJALOO et al., 2021). E interpretar os resultados gerados pela técnica de aprendizado de máquinas é um passo essencial de garantir uma informação robusta quando se tem várias métricas de avaliação na saída de dados. Portanto, é muito importante ajudar os produtores e as empresas de sementes com modelos precisos para prever a produtividade das culturas e a qualidade das sementes. Isso os ajudará a avaliar sua situação atual em termos de produtividade e eficiência e identificar possíveis melhorias, como identificar a qualidade das sementes ou escolher lotes de culturas de alto desempenho entre baixo, médio e alto vigor.

Com os avanços na tecnologia computacional e coleta de dados, surgiu uma necessidade essencial de desenvolver novos modelos para previsão de rendimento (SARIJALOO et al., 2021). Para este fim, os pesquisadores apresentaram vários modelos de previsão de rendimento, incluindo algoritmos de aprendizado de máquina tradicionais, métodos de mineração de dados e modelos de rede neural profunda (LIAKOS et al., 2018; CHLINGARYAN et al., 2018; VLONTZOS e PARDALOS, 2017; MUCHERINO et al., 2009; PAPAJORGJI e PARDALOS, 2014). Identificar os melhores lotes entre muitos dados torna-se o processo muito lento, pois a empresa e o produtor necessitam de informações rápidas e a curto prazo. Com isso o objetivo com este trabalho foi analisar as métricas de avaliação em diferentes modelos de aprendizagem de máquinas.

2. METODOLOGIA

Este estudo foi realizado com dados de uma empresa privada, sobre a qualidade fisiológica em lotes de sementes de soja. Devido a questões de confidencialidade, não foi citados nome de empresa e cultivares. Os dados foram enviados em planilha eletrônica organizado com base na análise interna do controle de qualidade da própria empresa. Foram utilizados 45 lotes (peneiras 6,0; 6,5; 7,0 mm) de semente de soja da cultivar A e 36 lotes (peneiras 6,0; 6,5; 7,0 mm) de sementes de soja da cultivar B e 12 lotes (peneiras 6,5; 7,0; 7,5 mm) de sementes da cultivar C. Os testes realizados foram: germinação, envelhecimento acelerado (EA), tetrazólio, emergência de plântulas e peso de mil sementes. Os resultados gerados destas avaliações foram organizados para inserção da plataforma do software Weka.

Foram utilizados quatro modelos para aprendizado de máquina, sendo: Random Forest, MLP (*Multi Layer Perceptron*), CVR (*Classification Via Regression*) e

J48. Procedimento inicial foi validação cruzada, onde se dividiu o conjunto de dados, treinamento e teste com 10 subconjuntos. Para verificar quais algoritmos seriam mais precisos utilizou-se as seguintes métricas de avaliação: Acurácia, Precisão, Recall, F-measure, Área ROC. Com os resultados obtidos determinou-se a melhor técnica de aprendizagem. Abaixo um esquema de fluxograma para os passos adotados no agrupamento dos dados tratados (Figura 1).

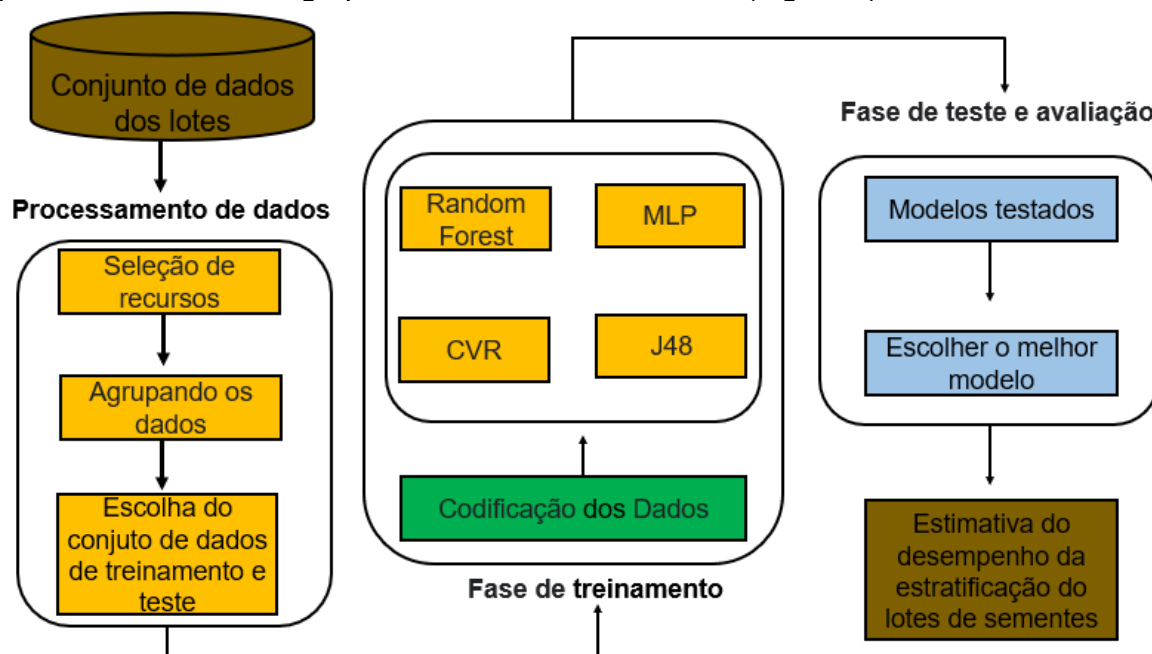


Figura 1- Diagrama de fluxo da análise dos dados de lotes de sementes de soja, mostrando a sequência de etapas do processamento de dados, treinamento do modelo e teste, levando à previsão de qualidade dos lotes das sementes.

Após a mineração foram analisados os dados. Neste caso, consideramos como mais importantes na métrica de avaliação os seguintes componentes: Acurácia: Percentual de aceito na validação do conjunto de dados; Precisão: dentre todas as classificações de classe Positivo que o modelo fez, quantas estão corretas; Recall: dentre todas as situações de classe Positivo como valor esperado, quantas estão corretas; F-Measure é a média harmônica entre Precisão e Recall (RODRIGUES, 2019), este foi transformado em porcentagem; área ROC (*Receiver Operator Characteristic*) mostra o quão bom o modelo criado pode distinguir entre duas coisas (já que é utilizado para classificação). Essas duas coisas podem ser 0 ou 1, ou positivo e negativo (RODRIGUES, 2018). E valores da classe de saída dos dados, como aceitos, rejeitados ou intermediário.

3. RESULTADOS E DISCUSSÃO

Ao utilizar um conjunto de dados sobre a qualidade de sementes em que a combinação e a orientação das sementes sobre os aspectos dos testes de controle de qualidade podem ser eficientes para prever a qualidade fisiológica, é necessário entender as métricas de avaliação proposta para melhorar o desempenho dos modelos testados.

De acordo com Rodrigues (2019), a acurácia é uma boa indicação geral de como o modelo “performou”, porém, pode haver situações em que ela é enganosa. O recall pode ser usado em uma situação em que os Falsos Negativos são

considerados mais prejudiciais que os Falsos Positivos. A precisão, pode ser usada em uma situação em que os Falsos Positivos são considerados mais prejudiciais que os Falsos Negativos. F-Measure é simplesmente uma maneira de observar somente uma métrica ao invés de duas (precisão e recall) em alguma situação, quando tem-se um valor baixo, é um indicativo de que ou a precisão ou o recall está baixo. A área ROC ou curva ROC é capaz de demonstrar o desempenho de um modelo de *machine learning*, que seja um classificador binário, por meio da relação da Taxa de Verdadeiro Positivo (Sensibilidade) e da Taxa de Falso Positivo (Especificidade), variando o ponto de corte na probabilidade estimada (PRATES, 2020).

Os resultados de acurácia na discriminação da avaliação das sementes em suas diferentes classes de decisão foram altos, variando de 0,74 a 0,81. O conjunto de treinamento para cada modelo compreendeu 81,7% dos dados para *Random Forest* e *Classification Via Regression*, seguido de 79,6% para J48 e *Multi Layer Perceptron* com 74,2%, sendo estes últimos os que apresentaram método com desempenho diferente analisados (Tabela 1). Embora, os valores de acurácia para J48 e *Multi Layer Perceptron* foram inferiores, os mesmos apresentaram uma área ROC superior aos dados da classe de rejeitados. Isso implica dizer, que a curva ROC foi mais bem definida na classe dos rejeitados do que para classe dos aceitos como predição de separação dos lotes de sementes de soja.

Tabela 1 - Resultados dos componentes de avaliação nos modelos testados na estratificação de lotes de sementes de soja.

Classificadores	Métricas de avaliação					
	Recall	Precisão	ROC Área	F-Measure	Classe	Acu (%)
Random Forest	0,92	0,90	0,97	91,7%	*	
	0,92	0,73	0,97	82,0%	**	81,7
	0,08	0,25	0,90	12,5%	***	
Multi Layer Perception (MLP)	0,90	0,87	0,93	89,1%	*	
	0,74	0,74	0,94	74,1%	**	74,2
	0,00	0,00	0,59	0,0%	***	
J48	0,92	0,87	0,86	90,1%	*	
	0,81	0,78	0,89	80,0%	**	79,6
	0,16	0,25	0,52	20,0%	***	
Classification Via Regression (CVR)	0,944	0,850	0,96	89,5%	*	
	0,852	0,793	0,96	82,1%	**	81,7
	0,167	0,500	0,74	25,0%	***	

aceito (*), rejeitado (**), intermediário (***); Acu = Acurácia; F-Measure; ROC Área- Receiver Operator Characteristic.

Os resultados das classes de aceito apresentaram maiores valores em todas as métricas de avaliação, exceto na área ROC (Tabela 1). Refletindo nos dados de maior quantidade com determinação para aceitos na tomada de decisão dos modelos testados. Se ROC fosse definido como a única métrica de avaliação os melhores modelos seriam *Random Forest* e CVR, conseqüentemente foram também os que apresentaram resultados de melhores para a acurácia. Entretanto,

cada métrica tem suas peculiaridades que devem ser levadas em consideração na escolha de como o modelo de classificação será avaliado. Não se deve pensar em uma como melhor ou pior que a outra de maneira geral, e sim deve-se analisar o problema e escolher a/as que melhor se adapta(m) (RODRIGUES, 2019).

4. CONCLUSÕES

Os resultados apresentados demonstram que a acurácia e área ROC permitem indicar o melhor modelo de avaliação de aprendizagem de máquina nos algoritmos testados, sendo Random Forest e CVR os que apresentaram melhor desempenho na tomada de decisão sobre os lotes de sementes de soja.

5. REFERÊNCIAS BIBLIOGRÁFICAS

CHLINGARYAN, A.; SUKKARIEH, S.; WHELAN, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review. **Comput. Electron. Agric.**, v. 151, p. 61-69, 2018.

LIAKOS, K.G.; BUSATO, P.; MOSHOU, D.; PEARSON, S.; BOCHTIS, D. Machine learning in agriculture: a review, **Sensors**, v. 18, p. 2674, 2018.

MUCHERINO, A.; PAPAJORGJI, P.; PARDALOS, P. M. A survey of data mining techniques applied to agriculture, **Oper. Res.**, v. 9, p. 121-140, 2009.

PAPAJORGJI, P. J., PARDALOS, P. Software engineering techniques applied to agricultural systems, ed. 2, Springer, v. 93, 301p. 2014.

PRATES, W. R. Curva ROC e AUC em Machine Learning. 2020. Disponível em <<https://cienciaenegocios.com/curva-roc-e-auc-em-machine-learning/>>. Acesso em 29 jul. 2021.

RODRIGUES, V. Entenda o que é AUC e ROC nos modelos de Machine Learning. 2018. Disponível em <<https://medium.com/bio-data-blog/entenda-o-que-%C3%A9-auc-e-roc-nos-modelos-de-machine-learning-8191fb4df772>>. Acesso em 02 ago. 2021.

RODRIGUES, V. Métricas de Avaliação: acurácia, precisão, recall... quais as diferenças? 2019. Disponível em <<https://vitorborbarodrigues.medium.com/m%C3%A9tricas-de-avalia%C3%A7%C3%A3o-acur%C3%A1cia-precis%C3%A3o-recall-quais-as-diferen%C3%A7as-c8f05e0a513c>>. Acesso em 02 ago. 2021.

SARIJALOO, F. B.; PORTA, M.; TASLIMI, B.; PANOS, F.; PARDALOS, M. Yield performance estimation of corn hybrids using machine learning algorithms. **Artificial Intelligence in Agriculture**, v 5, p. 82-89, 2021.

VLONTZOS, G. PARDALOS, P.M. Data mining and optimisation issues in the food industry, **Int. J. Sustain. Agric. Manag. Inform.**, v. 3, p. 44-64, 2017.