**ANÁLISE DA DIFUSÃO DE NOVOS HÍBRIDOS DE MILHO NO POSICIONAMENTO E INCREMENTO DE VENDAS ATRAVÉS DE MINERAÇÃO DE DADOS**

**BRITO, MARCELO G.1, GADOTTI, GIZELE I. 2, WENDLING, GRAZIELE D.3 , MONTEIRO, RITA DE CÁSSIA MOTA4 , BERNARDY, RUAN5**

1 MsC Eng. Agr., Faculdade de Agronomia Eliseu Maciel UFPel, marceloriberkws@gmail.com

2 Dra. Eng. Agric., Centro de Engenharias UFPel, gizele.gadotti@ufpel.edu.br

3 Eng. Agrônoma, Faculdade de Agronomia Eliseu Maciel, UFPel, grazi.wendling@gmail.com

4 MsC. Eng. Agric., MAPA, ritamonteiroo@gmail.com

5 MsC. Eng. Agric., Faculdade de Agronomia Eliseu Maciel UFPel, ruanbernardy@yahoo.com.br

Apresentado no

Congresso Técnico Científico da Engenharia e da Agronomia – CONTECC

08 a 11 de agosto de 2023

**RESUMO**: Para implantação da cultura do milho, o nível de tecnologia da semente aliada ao manejo adequado da lavoura impacta diretamente na produtividade dos campos brasileiros, sendo o uso de híbridos adaptados à diferentes regiões e épocas de semeadura um diferencial da produtividade. Este estudo tem por objetivo analisar a difusão do posicionamento de diferentes híbridos de milho através de ferramentas de mineração de dados. Para isso avaliou-se campos de difusão de milho no Maranhão na safra 2020/2021 e safrinha 2021/2022 com três diferentes níveis de produtividade, sendo cada nível submetido à três níveis de investimentos (baixo, médio e alto), em diferentes épocas de semeadura (abertura, meio e fechamento). Através da mineração de dados, utilizou-se como classificador o J48, sendo que o modelo apresentou 94,71% de correlação para os diferentes anos de plantio, 91,62% para as diferentes épocas de semeadura e 84,67% para produtividade, apresentando alta sensibilidade e precisão quanto aos diferentes anos de plantio. A mineração de dados é uma ferramenta que pode útil ao avaliar o posicionamento e recomendação de novos híbridos de milho, mas para isso é indispensável que haja um banco de dados robusto e com informações do local, como: época de plantio, altitude, condições edafoclimáticas e manejo a ser aplicado na área.

**PALAVRAS-CHAVE:** Inteligência artificial, estratificação, rentabilidade

**ANALYSIS OF THE DIFFUSION OF NEW CORN HYBRIDS IN POSITIONING AND INCREASE IN SALES THROUGH DATA MINING**

**ABSTRACT**: For the implantation of the corn crop, the level of seed technology combined with the proper management of the crop has a direct impact on the productivity of Brazilian fields, with the use of hybrids adapted to different regions and sowing times a productivity differential. This study aims to analyze the diffusion of positioning of different maize hybrids through data mining tools. For this, corn diffusion fields in Maranhão were evaluated in the 2020/2021 and 2021/2022 harvest with three different levels of productivity, each level being subjected to three levels of investment (low, medium, and high), at different times of sowing (opening, middle and closing). Through data mining, the J48 was used as a classifier, and the model presented 94.71% of correlation for the different years of planting, 91.62% for the different sowing times and 84.67% for productivity, presenting high sensitivity and precision regarding the different years of planting. Data mining is a tool that can be useful when evaluating the positioning and recommendation of new corn hybrids, but for that it is essential to have a robust database with local information, such as: planting time, altitude, edaphoclimatic and management to be applied in the area

**KEYWORDS:** Artificial intelligence, stratification, profitability

**INTRODUÇÃO**

Existem hoje no mercado híbridos de milho com diferentes características e que se adaptam melhor às condições de cada ambiente de produção visando atender as finalidades do agricultor (NASCIMENTO et al., 2011). Dependendo do nível de tecnologia de uma semente, o agricultor terá uma lavoura de baixo, médio ou alto nível tecnológico. Portanto, a adoção de sementes de alta tecnologia, aliada a um manejo de plantas adequado e à dedicação do agricultor, permitiu colocar o Brasil entre os três primeiros produtores e exportadores de milho globais (FILHO; BORGHI, 2022). A procura por híbridos adaptados e estáveis para cultivo em diferentes regiões e épocas de semeadura é uma das principais dificuldades dos programas de melhoramento genético de milho, principalmente quando considerada a alta manipulação gênica necessária para incorporação dos diferentes eventos tecnológicos existentes (STORCK et al., 2014). Esta dificuldade pode ser reduzida, estratificando-se a região considerada em sub-regiões com características ambientais semelhantes, buscando genótipos específicos que se apresentem adaptados em tais ambientes (CARVALHO et al., 2002). Esse trabalho objetiva avaliar, por meio de ferramentas de mineração de dados, a difusão de diferentes híbridos de milho no estado do Maranhão, em área com diferentes produtividades e em diferentes épocas de semeadura.

**MATERIAL E MÉTODOS**

 Neste estudo foram utilizados dados de campo provenientes de ensaios de difusão para posicionamento de milhos híbridos no Estado do Maranhão nas safras 2020/21 e na safrinha de 2021/2022, onde avaliou-se 3 níveis de produtividade de sementes de milho (ton/ha): baixa produtividade (<6ton/ha), média produtividade (>que 6 e ≤8 ton/ha) e alta produtividade (>8ton/ha), em duas altitudes diferentes: acima de 500m e abaixo de 500m.

Os diferentes níveis de produtividade avaliados forma submetidos a três diferentes níveis de investimentos: baixo, médio e alto. E a diferentes épocas de semeadura: abertura (de 01 de janeiro a 10 de fevereiro), meio (de 11 a 25 de fevereiro) e fechamento (de 26 de fevereiro a 10 de março). Os dados foram pré-processados e realizada a validação cruzada, subdividindo os dados em 10 subconjuntos (10 folds), retirando-se um dado para validação do modelo, buscando reduzir a probabilidade de subavaliar ou sobreavaliar o desempenho do modelo. Para realizar a mineração de dados supervisionados foi utilizado o classificador J48 e para a não supervisionada foi utilizado Cluster, utilizando o software Weka, versão 3.8.5. Para determinar a assertividade dos algoritmos foram utilizadas como métricas de avaliação: Acurácia, Precisão, Recall, F-measure e Área ROC. Através da matriz de confusão, foram extraídos os valores positivos (TP), falsos positivos (FP), verdadeiros negativos (TN) e falsos negativos (FN), com a finalidade de calcular as métricas Recall e Precisão, como proposto por Medeiros et al. (2020), determinando-se a melhor técnica de aprendizagem. Os mesmos conjuntos de dados foram utilizados para a análise não supervisionada por meio de clusters e por meio de agrupamento. O modo de classificação foi realizado através do clusters evaluation que tem como função comparar os agrupamentos criados com a classe do conjunto de dados.

**Tabela 1.** Descrição dos atributos analisados pela mineração de dados baseado em ano plantio.

|  |  |  |
| --- | --- | --- |
| **Atributos** | **Descrição** | **Valor** |
| Ano Protocolo | Ano | {2019; 2022} |
| Altitude | Altitude | {240; 585} |
| Cultivar ID | CID | {1; 1692} |
| Produtividade | Prod | {0-17.500} |
| Sacas/ha | Sa/ha | {5,6; 289} |
| Época de semeadura | Decisão tomada | {abertura, meio, fechamento} |

Fonte: Elaborado pelo autor, 2022.

**RESULTADOS E DISCUSSÃO**

Em seu estudo, Gadotti et al (2022) utiliza a aprendizagem de máquinas para classificar lotes de semente de milho quanto à sua qualidade fisiológica, onde o classificador J48 100% de precisão. Neste estudo, este mesmo classificador apresentou 94,71% de correlação com o ano de plantio (2019 e 2022), 91,62% com a época de semeadura (abertura, meio e fechamento) e de 84, 67% com as diferentes produtividades testadas. Ao avaliar as diferentes épocas de plantio, o algoritmo J48 apresentou 93,7% de recall e 90,4% de precisão quando se avalia a semeadura na época de abertura. Para a época de semeadura denominada Meio, apresentou 94,7% de recall e 94,2% de precisão. Já para época de semeadura denominada fechamento apresentou 93,7% de recall e 96,4% de precisão. Ao classificar os dados por agrupamento através de clustering foram obtidas as porcentagens de instâncias classificadas corretamente. No caso do ano plantio foi de 74,5%, para a classe plantio foi de 90,41% e para produtividade 87,22%. No ano de plantio em diferentes épocas de semeadura, ao avaliar as matrizes de confusão, o algoritmo J48 (supervisionado) apresentou melhor correlação do que por cluster (não supervisionado). Ao analisar os resultados de verdadeiro positivo na matriz de confusão para os clusters do ano plantio nota-se uma tendência característica de cada classificação de acordo com a época de plantio.

**CONCLUSÃO**

 A mineração de dados se mostrou uma ferramenta útil para auxiliar no posicionamento e recomendação de novos híbridos de milho. Contudo, se faz necessária a construção de um banco de dados robusto e com informações do local, como: época de plantio, altitude, condições edafoclimáticas e manejo a ser aplicado na área a fim de melhorar o indicativo de posicionamento desses novos híbridos de milho nas diferentes regiões do Maranhão.

**AGRADECIMENTOS**

Os autores agradecem a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão das bolsas.

**REFERÊNCIAS**

CARVALHO CGP et al. 2002. Interação genótipo x ambiente no desempenho produtivo da soja no Paraná. Pesquisa Agropecuária Brasileira 37: 989-1000.

FILHO, I.A.P.; BORGHI, A.P. Disponibilidade de cultivares de milho para o mercado de sementes do Brasil: safra 2021/2022 Documentos / Embrapa Milho e Sorgo, 2022, 16p.

GADOTTI, G.I.; MORAES, N.A.B.; SILVA, J.G.; PINHEIRO, R.M.; MONTEIRO, R.C. Prediction of ranking of lots of corn seeds by artificial intelligence. Scientific Paper, v. 42 n.4, e.20210005, 2022.

NASCIMENTO, E. S.; GILO, E. G.; TORRES, F. E.; SILVA JUNIOR, C. A.; OLIVEIRA, L. V. A.; LOURENÇÃO, A. S. Resposta de híbridos de milho a diferentes espaçamentos entre linhas. Nucleus, Ituveraba, v. 9, n. 2, p. 131-140, 2011.

STORCK L et al. 2014. Análise conjunta de ensaios de cultivares de milho por classes de interação genótipo x ambiente. Pesquisa Agropecuária Brasileira 49: 163-172.