



Comportamento dos fretes rodoviários dos grãos: uma aplicação de uma técnica de mineração de dados

Palavras-Chave: FRETE, GRÃOS, MINERAÇÃO DE DADOS

Autores/as:

Samuel Teixeira UNICAMP

Prof^a Dr^a Andréa Leda Ramos (orientador/a) UNICAMP

Ma. Karina Braga Marsola UNICAMP

1. INTRODUÇÃO

A economia brasileira se torna cada vez mais dependente da agronegócio, principalmente dos grãos como a soja, milho, arroz e feijão representando 23,5% do Produto Interno Bruto (PIB) (ABNER *et al.*, 2020)(COSTA; SANTOS; BARTHOLOMEU, 2020). Segundo a Companhia Nacional de Abastecimento (CONAB, 2021), a produção de grãos da safra 2020/2021, colheu 268,9 milhões de toneladas, sendo 11,9 milhões a mais que a safra anterior, tal desempenho está ligado ao aumento da produtividade ligado as inovações tecnológicas e a área plantada, que por sua vez cresceu cerca de 1,8% em relação à safra de 2019/2020 chegando à marca de 67,1 milhões de hectares. Ainda segundo a CONAB (2021), a cultura de soja deverá produzir 135 milhões de toneladas em uma área de 38,2 milhões de hectares, tornado o Brasil o maior produtor mundial de soja na safra de 2020/2021. Em relação a exportação, a expectativa para a soja é de 82,7 milhões de toneladas, visto que no período de janeiro a outubro de 2021 já foram exportados 81,4 milhões de toneladas.

Para comportar o escoamento dessa grande produção é necessário infraestrutura de transporte. Apesar do grande potencial para a expansão rede ferroviária e hidrológica, a via modal mais utilizada é a rodoviária. Atualmente cerca de 61% do total de cargas transportada são através das rodovias, 21% por ferrovias, 14% pelas hidrovias e apenas 0,4% por vias aéreas (CNT, 2020). Conhecer o comportamento e prever o valor do frete agrícola se torna uma ferramenta importante para a tomada de decisão dos *stakeholders*. A fim de conhecer quais são os parâmetros que estão relacionados como valor do frete agrícola praticado, este estudo utilizará a técnica de mineração conhecida como árvore de decisão. A ferramenta usada nesta pesquisa para a mineração de dados foi Waikato Environment for Knowledge Analysis (WEKA).

Este método já foi utilizado no meio agrícola para a classificação espectral de uma determinada área com plantação de cana-de-açúcar em Minas gerais, o resultado se mostrou

fundamental na identificação e delimitação da área plantada da cana-de-açúcar mostrando que a técnica Árvore de Decisão é eficiente para o mapeamento de áreas (DELGADO *et al.*, 2013). A Árvore de Decisão também foi utilizada por MEIRA *et. al* (2008) para compreender como as condições do ambiente, espaçamento entre as plantas na lavoura e cargas pendentes dos frutos influenciam na manifestação da ferrugem do cafeeiro, foi detectado que a temperatura é o fator mais relevante na taxa de contaminação do café segundo a árvore (MEIRA; RODRIGUES; MORAES, 2008)

2. METODOLOGIA

A árvore de decisão divide um conjunto de dados em classes através de regras hierárquicas e divisões em grupos onde cada segmento é um nó, organizando e permitindo a análise de uma grande quantidade de dados de maneira mais clara e efetiva (PITOMBO; COSTA, 2015). Para tal, o processo foi dividido em 3 partes principais: Seleção de dados, Limpeza de dados e Mineração de dados.

Os dados utilizados para montar o banco de dados estão disponíveis na Tabela 1.

Tabela 1. Descrição dos dados utilizados

| Dados | Recorte Temporal | Fonte |
|------------------------------------|-----------------------|---------------|
| Frete | Mensal de 2012 a 2019 | SIFRECA, 2020 |
| Produção estadual | Anual de 2012 a 2019 | IBGE, 2020 |
| Preço da soja negociada em Chicago | Mensal de 2012 a 2019 | CONAB, 2020a |
| Preço médio do diesel | Mensal de 2012 a 2019 | ANP, 2020 |
| Capacidade de armazenagem estadual | Mensal de 2012 a 2019 | CONAB , 2020b |
| Capacidade de moagem estadual | Anual de 2012 a 2019 | ABIOVE, 2020 |

A partir dos dados, o banco de dados foi criado com as seguintes variáveis: UF de origem, UF de destino, distância (km), produção por estado (toneladas), preço médio do diesel (R\$/L), capacidade de moagem do estado de origem (toneladas/dia), mercado internacional de Chicago (US\$/T), capacidade de armazenagem estadual do estado de origem (toneladas).

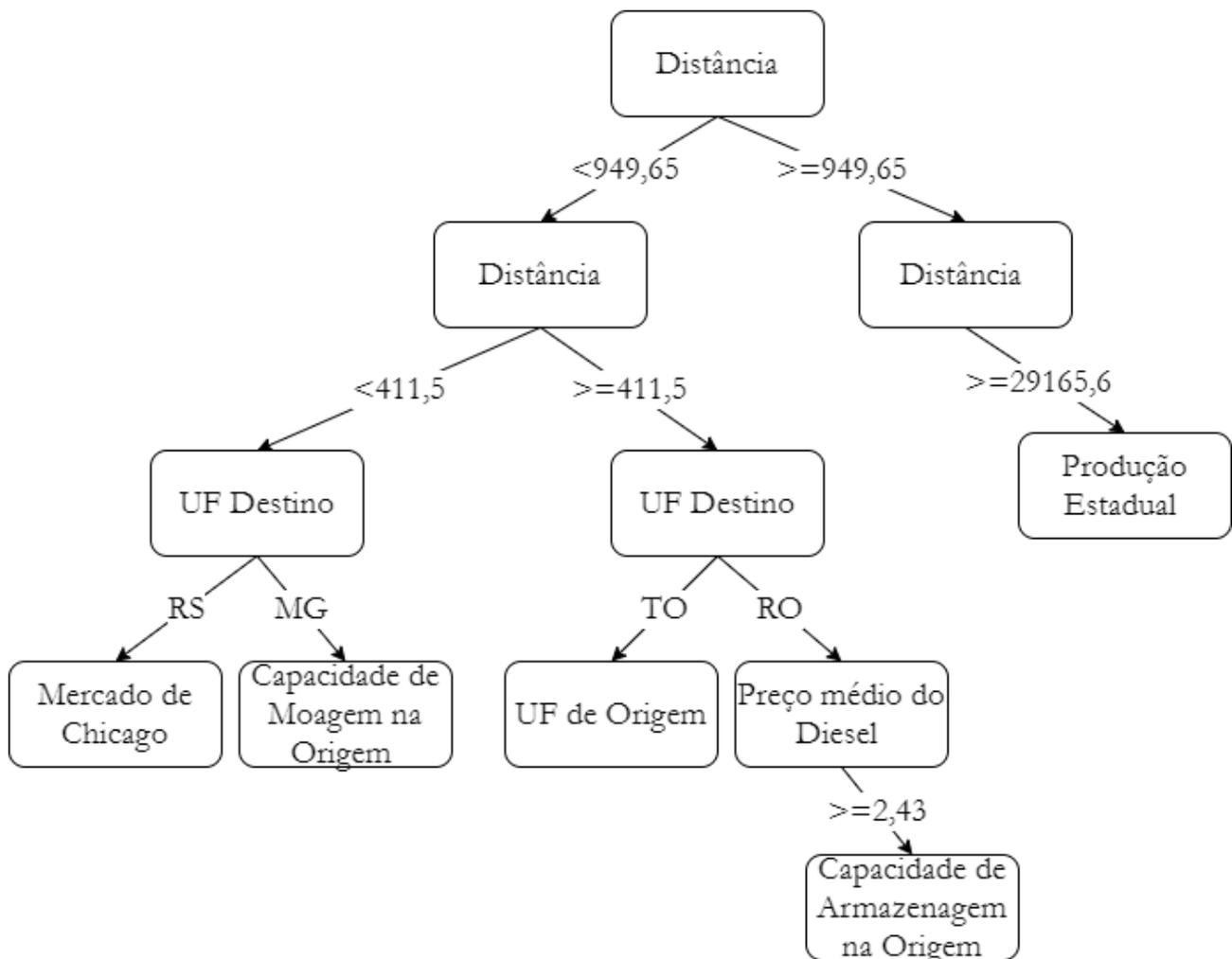
O banco de dados passou pelo processo de limpeza dos dados, para a remoção de erros de digitação, caracteres que atrapalham a leitura do *Weka*. Também foram analisadas as redundâncias e inconsistências das variáveis.

A terceira etapa do consistiu na aplicação das técnicas de mineração, para encontrar padrões no conjunto de dados avaliado. O método escolhido foi o *RandomTree*, as seguintes modificações foram feitas: *batchsize* de 100 para 50. A opção de teste no *cross-validation -10 folds*.

3. RESULTADOS E DISCUSSÃO

Como resultado da metodologia aplicada, o erro absoluto foi de 21,78% e coeficiente de correlação de 0,96. Dubath et al., (2011) evidencia que para um bom resultado o erro dever ser de no máximo 25% e uma correlação acima de 0,80. Os resultados aqui obtidos são superiores aos mínimos recomendados. A Figura 1 ilustra a sequência na qual as variáveis do estudo apareceram pela primeira vez na árvore de decisão, no Anexo I é disponibiliza a árvore completa.

Figura 1. Recorte das variáveis na árvore de decisão



A partir da árvore de decisão é possível observar que a variável distância é que a que tem maior impacto na predição do valor do frete da soja, vai de encontro as discussões de OLIVEIRA; CICOLIN; FILASSI, (2019), que ressaltam que além dos avanços tecnológicos das lavouras

brasileiras, os custos logísticos de transporte e armazenagem ainda são um problema para o produtor. Nesse mesmo sentido, SOUZA; CREMASCO; GABRIEL FILHO, (2016) reafirmam que as longas distâncias de transporte rodoviário para transportar a soja, não é o modelo mais econômico.

O estado para o qual a soja será destinada, bem como a produção estadual, aparecem no segundo nível da árvore de decisão. Já no terceiro nível aparecem as variáveis preço do diesel, mercado de Chicago e por fim, no quarto nível a armazenagem no estado de origem. A relação entre produção e armazenagem foi identificada por MOREIRA *et al.*, (2017), que utilizou o algoritmo Apriori, para gerar regras de associação que descrevessem padrões para os dados relacionados a logísticas de grãos no Brasil, seus achados apontam que a produção de soja e milho no estado de Mato Grosso é escoada pelo porto de Santos, os silos de armazenamento de grãos localizados no estado de Mato Grosso são abastecidos pela produção do próprio estado.

Por fim, a formação do preço do transporte rodoviário, além dos custos fixos e variáveis da atividade, incorpora também fatores locais e conjunturais. Se destacam as variáveis que influenciam o custo de transporte rodoviários de cargas agrícola: distância percorrida; especificidade da carga transportada; sazonalidade da produção; peculiaridades regionais; possibilidade de carga de retorno e custos operacionais (ORTÚZAR; WILLUMSEN, 1995; PÉRA *et al.*, 2019). Os resultados encontrados corroboram com as discussões das variáveis destacadas e dão direção para a predição do valor de frete brasileiro.

4. CONCLUSÕES

O Brasil é o segundo maior produtor de soja do mundo, com a maior parte da produção voltada para a exportação e que um dos principais entraves é o escoamento em função dos altos preços de transporte. Além das condições das rodovias, existem outros fatores que influenciam diretamente no preço final dos fretes como: as longas distâncias percorridas, a sazonalidade da produção, inflação e capacidade de armazenagem. Este estudo avaliou oito variáveis relacionadas ao frete rodoviário de soja.

Dentre as variáveis analisadas, a distância percorrida pra a realização do frete, foi o nó principal. No segundo nível da árvore, a variável distância novamente apareceu. As variáveis analisadas aparecerem pela primeira vez no terceiro nível da árvore, produção estadual e estado de destino. Já as variáveis mercado de Chicago, Capacidade de Moagem no estado de origem, estado de origem e preço médio do diesel apareceram na quarta camada, e por fim, na quinta camada, temos a capacidade de armazenamento na origem.

Referências Bibliográficas

- ANP. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. **Preços**. 2020. Disponível em: <https://preco.anp.gov.br/>. Acesso em: 01 out. 2020
- ABIOVE. Associação Brasileira das Indústrias de Óleos Vegetais. **Estatísticas**. 2020. <https://abiove.org.br/estatisticas/>. Acesso em: 20 out. 2020
- ABNER, J. *et al.* Estudo sobre a influência do frete no comportamento da margem do produtor rural. 2020, **Sociedade Brasileira de Economia, Administração e Sociologia Rural - SOBER** Foz do Iguaçu, 2020.
- CONAB. Companhia Nacional de Abastecimento. **Safras**. 2021. Disponível em: <https://www.conab.gov.br/info-agro/safras>. Acesso em: 01 jul. 2021.
- CONAB. Companhia Nacional de Abastecimento. **Série histórica das safras**. 2020a. Disponível em: <https://www.conab.gov.br/info-agro/safras/serie-historica-das-safras>. Acesso em: 10 out. 2020.
- CONAB. Companhia Nacional de Abastecimento. **Indicadores da Agropecuária**. 2020b. Disponível em: <https://www.conab.gov.br/indicadores-da-agropecuaria>. Acesso em: 10 out. 2020.
- CNT. Confederação Nacional do Transporte. **Boletim Unificado**. 2020 Disponível em: <https://www.cnt.org.br/boletins>. Acesso em: 10 nov. 2020.
- DUBATH, P. *et al.* Random forest classification of Hipparcos periodic variable stars. **Monthly Notices of the Royal Astronomical Society**, v. 414, n. 3, p. 2602-2617, 2011.
- DELGADO, R.C. *et al.* Classificação espectral de áreas plantadas com as culturas do milho e feijão por meio da árvore de decisão. **Engenharia Agrícola**, Jaboticabal, ano 2012, v. 32, n. 2, p. 369-380, mar./abr. 2012.
- Instituto Brasileiro de Geografia e Estatística. Disponível em: <https://sidra.ibge.gov.br/pesquisa/lspa/tabelas>. Acesso em: 15 out. 2020.
- MEIRA, C. A. A.; RODRIGUES, L. H. A.; MORAES, S. A. Análise da epidemia da ferrugem do cafeeiro com árvore de decisão. **Tropical Plant Pathology**, v. 33, n. 2, p. 114-124, 2008.
- MOREIRA, C. E. S. *et al.* Identification of freight patterns via association rules: The case of agricultural grains. **Bulgarian Journal of Agricultural Science**, v. 23, n. 6, p. 887-893, 2017.
- OLIVEIRA, A. L. R. DE; CICOLIN, L. DE O. M.; FILASSI, M. The logistic performance of Brazilian grains transportation. **Agrarian**, v. 12, n. 46, p. 521-527, 2019.
- ORTÚZAR, J. D.; WILLUMSEN, L. G. **Modelling transport**. 2nd. ed. [S.l.]: Chichester: John Wiley & Sons, 1995.
- PÉRA, T. G. *et al.* Evaluation of green transport corridors of Brazilian soybean exports to China. **Brazilian Journal of Operations & Production Management**, v. 16, n. 3, p. 398-412, 2019.
- PITOMBO, C. S.; COSTA, A. S. G. DA. Aplicação conjunta de modelos não paramétricos e paramétricos para previsão de escolha modal. **Journal of Transport Literature**, v. 9, n. 1, p. 30-34, 2015.
- SIFRECA. Sistema de Informações de Fretes. **Mercado de Fretes**. Disponível em: <https://sifreca.esalq.usp.br/mercado-de-fretes>. Acesso em: 10 out. 2020.
- SOUZA, R. O.; CREMASCO, C. P.; GABRIEL FILHO, L. R. A. Análise dos valores de frete da soja a granel nos sistemas unimodal e multimodal de transporte. **Revista em Agronegócio e Meio Ambiente**, v. 9, n. 4, p. 819-837, 2016.