

UNIVERSIDADE FEEVALE

RENAN BREMM

PREVISÃO DE COLHEITA DE UVAS COM TÉCNICAS DE
MACHINE LEARNING

(Título Provisório)

Anteprojeto de Trabalho de Conclusão

Novo Hamburgo
2022

RENAN BREMM

PREVISÃO DE COLHEITA DE UVAS COM TÉCNICAS DE
MACHINE LEARNING

(Título Provisório)

Anteprojeto de Trabalho de Conclusão de Curso, apresentado como requisito parcial
à obtenção do grau de Bacharel em
Sistemas de Informação pela
Universidade Feevale

Orientador: Juliano Varella de Carvalho

Novo Hamburgo
2022

RESUMO

O impacto dos fatores climáticos - como temperatura, umidade e precipitação - sobre as culturas perenes são notáveis no seu desempenho produtivo anual sucessivo. Na viticultura, mais especificamente, as condições meteorológicas exercem um papel preponderante sobre os resultados qualitativos e quantitativos das safras, sendo até mesmo mais importantes do que muitas intervenções executadas pelos viticultores. Por este motivo, conhecer antecipadamente os resultados potenciais do vinhedo constitui-se um fator crítico e estratégico para o setor vitivinícola. Com isso, é possível alcançar um melhor equilíbrio no manejo da cultura, melhorar as intervenções durante o desenvolvimento vegetativo e reprodutivo das videiras, além de planejar a logística desde a colheita até a vinificação, venda e distribuição dos produtos produzidos a partir das uvas, e provisionar os impactos financeiros. A predição dos resultados da colheita baseada em observações das condições meteorológicas, fenológicas e fenotípicas não é novidade entre os produtores, já que estes, de forma empírica, conseguem estimar os potenciais resultados de uma dada safra. Por outro lado, existe um avanço significativo no uso de tecnologias computacionais aplicadas à agricultura de precisão que estão tornando tarefas manuais e muitas vezes impraticáveis em aplicações totalmente automatizadas. Dentre as muitas tecnologias computacionais, estão as técnicas de *Machine Learning*, as quais têm sido amplamente exploradas para auxiliar na produção agrícola como técnicas que permitem, por exemplo, através de dados históricos prever a produtividade das culturas. Com base nos aspectos citados acima, esta pesquisa tem como objetivo geral criar um modelo preditivo a partir do uso de técnicas de *Machine Learning* para previsão de resultado de colheitas futuras de uvas, baseado em dados históricos de clima, considerando os efeitos do El Niño e La Niña; a produtividade e acúmulo de açúcares em uvas cultivadas em Santana do Livramento na Campanha Gaúcha. A metodologia de pesquisa escolhida foi *Design Science Research* (DSR) que é propícia para integrar aspectos teóricos com o desenvolvimento de software. Serão consideradas as seis etapas definidas pela DSR: identificação do problema e motivação; definição dos objetivos para a solução; projeto e desenvolvimento; demonstração; avaliação; comunicação.

Palavras-chave: Uva, *Machine Learning*, Previsão, Clima

SUMÁRIO

MOTIVAÇÃO	5
OBJETIVOS	8
METODOLOGIA	9
CRONOGRAMA	11
BIBLIOGRAFIA	13

MOTIVAÇÃO

A previsão sobre os resultados quantitativos e qualitativos dos vinhedos é de grande interesse para os produtores de uva e vinho. Dentre os aspectos que mais influenciam na produção estão a irrigação, fertilização, seleção de variedades e, sobretudo, as condições climáticas. A videira é fortemente sensível ao clima e, em muitas regiões vinícolas do mundo, as últimas décadas foram caracterizadas por grandes variações interanuais na produção de uvas com efeitos adversos para os produtores de uva e vinho (Cunha e Richter, 2012).

As condições climáticas, como temperatura, velocidade do vento, umidade e precipitação, influenciam significativamente a fisiologia e os desenvolvimentos vegetativo e reprodutivo das videiras. Isso pode resultar em mudanças substanciais na fenologia¹ eventos de crescimento da cultura, como brotação, floração, *veraison* e colheita. Segundo Keller (2010), no caso da videira, essas mudanças geralmente levam a alterações substanciais no teor de açúcares da baga e na produtividade.

A previsão de colheita baseada em fatores climáticos não é algo novo. Obviamente, sabe-se que os resultados das culturas dependem do clima. Dessa forma, os produtores, mesmo sem qualquer aparato tecnológico, já conseguem visualizar e estimar possíveis resultados. Porém, esta previsão puramente empírica possui limitações, principalmente relacionadas à quantidade de dados em que essa previsão se apoia e, também, ao grau de precisão dos resultados. Buscando resolver estas limitações, há um grande avanço dos sistemas de agricultura inteligente e de precisão, que começam a desempenhar um importante papel na melhoria das atividades agrícolas. De acordo com Issad et al.(2019), atualmente é possível coletar dados do campo provenientes das mais diversas fontes: temperatura do solo, umidade do solo, umidade, radiação solar, direção do vento, nível de chuva, que tornam qualquer análise de fatores relacionados à agricultura muito mais completa, organizada e precisa.

No entanto, o processamento manual dessa enorme quantidade de dados é difícil ou impossível para o ser humano, por isso é essencial automatizar essa tarefa usando métodos e ferramentas de análise para transformar esses dados em conhecimento que servirá no processo de tomada de decisão. As técnicas de mineração de dados têm um papel fundamental na análise de dados, já que é o processo que permite a descoberta de padrões de configuração em

¹ Estudo de como a planta se desenvolve ao longo de suas diferentes fases: germinação, emergência, crescimento e desenvolvimento vegetativo, florescimento, frutificação, formação das sementes e maturação.

grandes conjuntos de dados. O uso de técnicas de *Data Mining* acabou assumindo diversas tarefas no campo agrícola como identificação, detecção e classificação de pragas e previsão de doenças de culturas, previsão de produtividade (ISSAD et al., 2019), sugestão de fertilizantes (PADALU et al., 2017), e determinação da umidade do solo em tempo real (MATEI et al., 2017).

Estudos sobre o tema têm sido relatados nas mais diversas áreas da agricultura, muitos deles empregando aplicações de aprendizado de máquina para análise preditiva. E nos últimos anos, o uso de Redes Neurais Artificiais (RNA) para análise de previsão tem chamado a atenção da comunidade científica. Os algoritmos de aprendizagem desenvolvidos abordaram a estimativa de produção de diversas maneiras: Ballesteros et al. (2020) combinaram técnicas de sensoriamento remoto, visão computacional e redes neurais artificiais para estimar colheitas. Da mesma forma, Monga et al. (2018) usaram redes neurais de convolução para desenvolver modelos que podem estimar o rendimento usando imagens RGB.

Visto a demanda existente por aplicações automatizadas de predição de colheitas e os avanços relacionados à coleta e armazenamento de dados históricos na área da agricultura, este trabalho tem como objetivo geral estudar e criar modelos preditivos de resultados de colheitas a partir de técnicas de *Machine Learning*. Tendo em vista as diversas potenciais aplicações do estudo nas culturas agrícolas do estado do Rio Grande do Sul, encontramos junto a colegas pesquisadores do Programa de Pós-Graduação em Agronomia (Área de Concentração – Fruticultura de Clima Temperado) Universidade Federal de Pelotas (UFPel), uma excepcional base de dados climatológicos e de qualidade e produtividade de uvas produzidas pela Vinícola Almadén. A base de dados abarca mais de 20 variedades de uva e um período de quase 3 décadas.

As variações interanuais do clima na produção de vinho são uma grande ameaça para a indústria do vinho (Cunha et al., 2016), o planejamento incorreto da safra acarreta em grandes perdas financeiras e pode afetar diretamente a fisiologia das videiras, não apenas em safras isoladas, mas ao longo do seu ciclo de vida de várias décadas (Kaltbach et al. 2022). Portanto, há uma forte demanda por modelos que possam estimar a produção dos próximos anos, e com isto guiar os produtores nas suas estratégias de cultivo e comerciais.

A Vinícola Almadén, que situa-se na Campanha Gaúcha, e que assim como toda região sul brasileira é afetada diretamente pelo fenômeno El Niño-Oscilação Sul (ENSO) e suas fases quente (El Niño) e fria (La Niña). O El Niño é caracterizado por temperaturas mais quentes das águas superficiais no Oceano Pacífico e enfraquecimento de ventos alísios equatoriais. La

Niña apresenta características opostas. As condições neutras são compreendidas entre os limites mínimos que caracterizam El Niño e La Niña (Timmermann et al., 2018). Tanto a fase fria quanto a quente ocorrem com frequência irregular e diferentes níveis de intensidade (por exemplo, fraco, moderado ou forte), e geralmente persistem por 12–18 meses (Cunha, 2001).

Segundo o estudo de Kaltbach et al.(2022), já existe uma comprovada associação entre o fenômeno ENSO e o desempenho produtivo das videiras na região de Santana do Livramento (RS). Portanto, assume-se que é possível fazer estimativas da quantidade e qualidade da produção vitícola anual, baseando-se nas variáveis meteorológicas da região e nas oscilações de temperatura das águas superficiais do pacífico. Levando estes aspectos em consideração, o presente trabalho estudará essa correlação na óptica da tecnologia, através de técnicas de *Machine Learning*.

Acredita-se que este trabalho possa comprovar que as técnicas de *Machine Learning* são capazes de prever a produtividade e qualidade dos frutos das videiras em níveis satisfatórios de assertividade a partir de dados climatológicos. E, com isso, em uma situação ideal, o produtor poderá melhorar a produtividade da sua área de cultivo e aumentar os rendimentos da safra por meio dos modelos preditivos obtidos através da confecção deste trabalho de conclusão.

OBJETIVOS

Objetivo geral:

Criar um modelo preditivo a partir do uso de técnicas de *Machine Learning* para previsão de resultado de colheitas futuras de uvas, baseado em dados históricos de clima, considerando os efeitos do El Niño e La Niña; a produtividade e acúmulo de açúcares em uvas de videiras cultivadas em Santana do Livramento, na Campanha Gaúcha (RS).

Objetivos específicos:

- Interagir de forma interdisciplinar e holística com o grupo de pesquisa do Programa de Pós-Graduação em Agronomia, na área da Agronomia da Universidade Federal de Pelotas(UFPel);
- Criar um *dataset* de treinamento, integrando dados históricos de clima e de produtividade/qualidade de uvas;
- Escolher, treinar e avaliar os modelos gerados;
- Analisar a viabilidade de utilização de técnicas de *Machine Learning* para predição de colheitas futuras;

METODOLOGIA

Neste trabalho será realizada uma pesquisa de natureza aplicada, uma vez que objetiva gerar conhecimentos para aplicação prática buscando a solução de um problema específico (PRODANOV; FREITAS, 2013), neste trabalho o problema a ser resolvido será a previsão de resultado de colheitas de uvas com base em dados climáticos, utilizando técnicas de *Machine Learning*.

A metodologia usada para o desenvolvimento desta pesquisa será a *Design Science Research* (DSR). Esta metodologia é composta por seis etapas principais que permitem avaliar a eficácia dos resultados e publicá-los na comunidade.

Segundo JUNIOR et al. (2017), baseado em PEFFERS (2007), as etapas são descritas abaixo juntamente com a aplicação no presente estudo de criação de um modelo de previsão de colheitas de uva:

1. Identificação do problema e motivação: fazer um levantamento bibliográfico de tecnologias e ferramentas utilizadas para construção de modelos preditivos baseados em *Machine Learning* aplicados a agricultura (mais especificamente a produção de uvas), com objetivo de modelar a arquitetura que será utilizada e compor o referencial teórico do trabalho.
2. Definição dos objetivos para a solução: definir os softwares utilizados no processamento de dados e criação do modelo preditivo, assim como as técnicas de *Machine Learning* que serão utilizadas. Também será analisada a base de dados das informações da Vinícola Almadén e serão definidas as informações que serão consideradas no estudo.
3. Projeto e desenvolvimento: fazer a preparação do *dataset*, e então aplicar as técnicas de *Machine Learning* selecionadas, treinando os modelos de algoritmos, avaliando resultados e ajustando os parâmetros de forma a alcançar resultados de previsão de colheitas mais confiáveis e assertivas.
4. Demonstração: executar os modelos de algoritmos criados com o *dataset* de teste.
5. Avaliação: comparar os resultados obtidos a partir dos modelos de *Machine Learning* da presente pesquisa de forma a definir o modelo mais assertivo e com melhor performance para prever o resultado de colheitas de uvas na região da Campanha Gaúcha.
6. Comunicação: os resultados obtidos nesse trabalho serão possivelmente publicados em congressos científicos da computação e comunicados ao Programa de Pós-graduação em Agronomia da UFPel e a Vinícola Almadén que proveu os dados utilizados no estudo.

CRONOGRAMA

Trabalho de Conclusão I

Etapa	Meses			
	Ago	Set	Out	Nov
Identificação do problema e motivação				
Busca por Base de Dados				
Interagir com o grupo de pesquisa da UFPel				
Escrita do Anteprojeto				
Estudo sobre a Viticultura				
Levantamento de técnicas de Machine Learning				
Escrita do TCC I				
Definição dos objetivos para a solução				
Definir <i>Softwares</i> utilizados na criação do modelo				
Definir técnicas de <i>Machine Learning</i>				
Projeto e desenvolvimento				
Preparação do <i>dataset</i>				
Treinamento, avaliação e ajuste dos modelos de algoritmos				
Entrega TCC I				

Trabalho de Conclusão II

Etapa	Meses			
	Mar	Abr	Mai	Jun
Projeto e desenvolvimento				
Treinamento, avaliação e ajuste dos modelos de algoritmos				

Interagir com o grupo de pesquisa da UFPel				
Escrita do TCC II				
Demonstração:				
Execução dos modelos com o <i>dataset</i> de teste				
Avaliação:				
Avaliar acurácia e performance das técnicas de <i>Machine Learning</i>				
Comunicação:				
Apresentação dos resultados às partes interessadas				
Entrega do TCC II				

BIBLIOGRAFIA

- AIT ISSAD, H; AOUDJIT, R; RODRIGUES, J. J. P. C. **A comprehensive review of Data Mining techniques in smart agriculture**. Engineering in Agriculture, Environment and Food, [s. l.], v. 12, n. 4, p. 511–525, 2019. DOI 10.1016/j.eaef.2019.11.003. Disponível em: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=edselp&AN=S1881836619301533&lang=pt-br&site=eds-live&scope=site>. Acesso em: 6 setembro. 2022.
- BALLESTEROS, R et al. **Vineyard yield estimation by combining remote sensing, computer vision and artificial neural network techniques**. Precision Agriculture, 21(6), 1242–1262, 2020. <https://doi.org/10.1007/s11119-020-09717-3>
- CUNHA, M., RICHTER, C. **Measuring the impact of temperature changes on the wine production in the douro region using the short time fourier transform**. International Journal of Biometeorology 56, 357–370, 2012.
- CUNHA, G; DALMAGO G; ESTEFANEL, V. **El Niño Southern Oscillation Influences on Wheat Crop in Brazil**. I In: Bedö Z & Láng L (Eds.) Wheat in a Global Environment. Developments in Plant Breeding, Volume 9. Dordrecht, Springer. p. 445–450. 2001.
- JUNIOR, Vanderlei et al. **Design Science Research Methodology As Methodological Strategy for Technological Research**. Revista Espacios, 2017. vol. 38, p. 25.
- KALTBACH, Pedro et al. **Relationship between the El Niño-Southern Oscillation and yield and sugar content of wine grapes grown in Santana do Livramento, RS, Brazil**. Semina: Ciênc. Agrár. Londrina, v. 43, n. 5, p. 2031-2044, set./out. 2022
- KELLER, Markus. **The Science of Grapevines**. 3. ed. Prosser: Academic Press, 2020. 521 p.
- MATEI, Oliviu et al. **A Data Mining System for Real Time Soil Moisture Prediction**. Procedia Engineering, [s. l.], Volume 181, p. 837-844. 2017
- MONGA, T. **Estimating vineyard grape yield from images**. In E. Bagheri & J. C. K. Cheung (Eds.), Advances in artificial intelligence p. 339–343. 2018.
- PADALU, P et al. **Smart water dripping system for agriculture/farming**. 2nd International Conference for Convergence in Technology (I2CT), [s. l.], p. 659-662. 2017.
- PEFFERS, Ken et al. **A Design Science Research Methodology for Information Systems Research**. Journal of Management Information Systems, v. 24, 3ed., p. 45–78. 2007.

PRODANOV, C. C.; FREITAS, E. C. (2013). **Metodologia do trabalho científico: Métodos e Técnicas da Pesquisa e do Trabalho Acadêmico**. Universidade Feevale. Disponível em: <<http://www.feevale.br/Comum/midias/8807f05a-14d0-4d5b-b1ad-1538f3aef538/E-book%20Metodologia%20do%20Trabalho%20Cientifico.pdf>>. Acesso em 26 ago. 2022.