

Departamento de Computação

Estratégias de Sistemas de Informação

Trabalho sobre tomada de decisão

Milena de Toledo Araujo (221024743)

Pedro Henrique Roratto (221026932)

Renan Arques (221026517)

Trabalho sobre tomada de decisão	1
1. Avaliação do cenário	3
1.1. Descrição	3
1.2. Problema	3
1.3. Oportunidades	3
1.4. Objetivos	4
2. Dados e informações	4
3. Descrição do modelo	5
3.1. Uso de Informações	5
3.2. Indicadores e valores utilizados	5
3.3. Funcionamento do algoritmo Random Forest Regressor	5
3.4. Critério e regras	6
4. Problemas e limitações	7
5. Considerações finais	7
6. Referências	7

1. Avaliação do cenário

“Uma empresa agrícola quer comprar uma fazenda para plantio no Estado de São Paulo. O objetivo da empresa é escolher uma cidade adequada para a plantação de um tipo de cultura em uma determinada época do ano. O sistema deve exibir em um mapa quais são as cidades mais adequadas para o tipo de cultura e época do ano que o usuário escolher. O sistema deve permitir escolher entre ao menos 20 opções de culturas e deve conter ao menos 30 cidades.”

1.1. Descrição

A empresa agrícola está planejando adquirir uma fazenda para ampliar suas operações de plantio no Estado de São Paulo. A região é conhecida pela sua diversidade climática e condições variadas de solo, oferecendo uma ampla gama de possibilidades para diferentes culturas agrícolas. A empresa está interessada em escolher a cidade mais adequada para o planejamento de uma cultura específica em uma época específica do ano. O objetivo final é otimizar a produção, levando em consideração fatores como clima, solo e sazonalidade.

1.2. Problema

O principal desafio enfrentado pela empresa é a falta de informações centralizadas e acessíveis sobre as características específicas de cada cidade em relação às diferentes culturas e épocas do ano. Uma escolha errada da cidade ou da cultura pode resultar em baixos rendimentos, perdas financeiras e ineficiência operacional.

1.3. Oportunidades

A implementação de um sistema eficiente pode oferecer diversas oportunidades para a empresa que se encontra no cenário proposto. Algumas delas podem ser pontuadas como: otimização da produção, aumento significativo na produtividade e, conseqüentemente, aumento nos ganhos..

Ao mapear informações sobre cada cidade, o sistema capacita os gestores para tomar decisões informadas e estratégicas. Essa capacidade de tomada de decisão informada é essencial, já que, dessa maneira é possível proporcionar uma redução de riscos, a fim de evitar o planejamento inadequado em determinadas cidades ou épocas do ano. Essa abordagem cuida para minimizar o risco de perdas, maximizando, assim, o retorno sobre o investimento.

A inclusão de opções para 20 culturas oferece à empresa a chance de diversificar suas operações. Essa flexibilidade permite que a empresa se adapte às mudanças nas condições do mercado e nas demandas do consumidor, garantindo uma abordagem resiliente e dinâmica.

Por fim, ao ter acesso a informações centralizadas, a empresa pode atualizar o planejamento logístico e operacional de forma eficiente. Essa eficiência operacional

resulta em um fluxo de trabalho otimizado, proporcionando ganhos adicionais em termos de produtividade e gestão eficaz dos recursos.

1.4. Objetivos

Primeiramente, a meta é implementar um Sistema de Informações Geográficas (GIS) capaz de mapear não apenas as temperaturas climáticas, mas também as características do solo em pelo menos 30 cidades. Este sistema será uma base para a tomada de decisões estratégicas em relação ao planejamento.

A personalização é uma peça-chave, e o sistema é projetado para permitir que o usuário selecione uma cultura específica e a época do ano desejada. Isso se traduzirá em um mapa interativo que destaca as cidades mais adequadas para o planejamento da cultura escolhida, garantindo uma abordagem sob medida para as necessidades da empresa.

A interface do usuário também está no centro do planejamento, o desenvolvimento de uma interface intuitiva que facilite a interação com o sistema é uma prioridade. Isso inclui opções claras e acessíveis para a escolha de culturas e épocas do ano, garantindo que a utilização do sistema seja eficiente e amigável.

2. Dados e informações

O sistema desenvolvido foi alimentado por dados e informações provenientes de fontes confiáveis e atualizadas, garantindo a precisão e confiabilidade das recomendações oferecidas. Três fontes principais foram fundamentais na elaboração dos conjuntos de dados:

- Banco de Dados e Informações Ambientais: essa fonte fornece informações sobre o tipo de solo predominante nas cidades incluídas no conjunto de dados. O objetivo é coletar dados abrangentes sobre as características do solo, essenciais para determinar as previsões agrícolas em cada localidade.
- *Weather Spark* : o *site Weather Spark* é uma fonte poderosa para consultar as temperaturas médias das cidades ao longo das quatro estações do ano. Esses dados meteorológicos são fundamentais para avaliar as condições climáticas e, consequentemente, influenciar nas decisões sobre o plantio de culturas específicas em épocas determinadas.
- Empresa Brasileira de Pesquisa Agropecuária (Embrapa): o site da Embrapa foi utilizado para coletar informações sobre as condições ideais das culturas escolhidas para formar o segundo conjunto de dados. Essas informações são essenciais para entender os requisitos específicos de cada cultura, garantindo que as escolhas de plantio estejam alinhadas com as condições ideais.

Ressalta-se que todas as fontes utilizadas para a obtenção de dados e informações são confiáveis e passíveis de atualização. A constante manutenção e verificação da relevância e precisão dos dados garantem que o sistema forneça

recomendações fundamentadas e as mais próximas das condições reais, promovendo assim a eficácia operacional da empresa.

3. Descrição do modelo

3.1. Uso de Informações

O modelo proposto utiliza as seguintes informações para recomendar as cinco cidades mais semelhantes para cultivar uma determinada cultura na estação do ano escolhida pelo usuário: a temperatura máxima e mínima e o tipo de solo predominante. As informações são utilizadas para calcular uma pontuação de similaridade para cada cidade, que é dada pela seguinte fórmula:

$$\text{pontuação_similaridade} = 1 / (1 + \text{similaridade_temp_max} + \text{similaridade_temp_min} + (1 - \text{similaridade_tipo_solo}))$$

Onde:

- $\text{similaridade_temp_max}$ é a diferença entre a temperatura máxima da cidade e a temperatura máxima ideal para a cultura, dividida por 2
- $\text{similaridade_temp_min}$ é a diferença entre a temperatura mínima da cidade e a temperatura mínima ideal para a cultura, dividida por 2
- $\text{similaridade_tipo_solo}$ é 1 se o tipo de solo da cidade for o mesmo que o tipo de solo ideal para a cultura, ou 0 caso contrário

As cidades são então ordenadas em ordem decrescente de pontuação de similaridade e as três primeiras são recomendadas.

3.2. Indicadores e valores utilizados

O único indicador utilizado para escolher entre as possíveis alternativas é a pontuação de similaridade. As cidades com maior pontuação são consideradas mais semelhantes às condições ideais para a cultura e, portanto, são as mais recomendadas.

3.3. Funcionamento do algoritmo Random Forest Regressor

O funcionamento do algoritmo de aprendizado de máquina *Random Forest Regressor*, em termos de fluxo de dados, funciona da seguinte maneira:

1. Os dados de treinamento são divididos em um conjunto de dados de treino e um conjunto de dados de teste.
2. Cada árvore de decisão é treinada no conjunto de dados de treino.

3. As árvores de decisão são então combinadas para produzir uma única previsão.

No modelo proposto, os dados de treinamento são as cidades da base B, que são filtradas de acordo com a estação do ano escolhida pelo usuário. As variáveis utilizadas para treinar as árvores de decisão são as temperaturas máxima e mínima e o tipo de solo predominante.

A primeira etapa do fluxo de dados é a divisão dos dados em um conjunto de dados de treino e um conjunto de dados de teste. O conjunto de dados de treino é utilizado para treinar as árvores de decisão, enquanto o conjunto de dados de teste é utilizado para avaliar o desempenho do modelo.

Na segunda etapa, cada árvore de decisão é treinada no conjunto de dados de treino. O algoritmo Random Forest Regressor utiliza um processo de *bootstrap* para criar subamostras aleatórias dos dados de treinamento. Cada árvore de decisão é treinada em uma subamostra diferente.

Na terceira etapa, as árvores de decisão são então combinadas para produzir uma única previsão. O algoritmo Random Forest Regressor utiliza um processo de votação para combinar as previsões das árvores de decisão.

No modelo proposto, o algoritmo Random Forest Regressor está sendo usado para calcular as pontuações de similaridade das cidades. Cada árvore de decisão calcula uma pontuação de similaridade para cada cidade. As pontuações de similaridade das cinco árvores de decisão são então combinadas para produzir uma única pontuação de similaridade para cada cidade.

As cidades com maior pontuação de similaridade são consideradas mais semelhantes às condições ideais para a cultura e, portanto, são as mais recomendadas.

3.4. Critério e regras

O critério adotado pelo modelo é recomendar, através de uma previsão realizada pelo algoritmo de *machine learning Random Forest Regressor*, as cinco cidades com maior pontuação de similaridade, que é calculada com base nas condições ideais (temperatura máxima e mínima e tipo de solo) das culturas escolhidas para compor a base de dados.

As regras adotadas no modelo são:

- A pontuação de similaridade é calculada usando as temperaturas máxima e mínima e o tipo de solo predominante da cidade;
- As cidades são ordenadas em ordem decrescente de pontuação de similaridade;
- As cinco primeiras cidades são recomendadas.

4. Problemas e limitações

O modelo apresenta uma limitação significativa relacionada aos tipos de solo, uma vez que considera apenas o solo mais abundante de cada região durante o desenvolvimento. Essa abordagem restringe a amplitude da pesquisa, prejudicando a representatividade do modelo em cenários nos quais solos menos predominantes também desempenham um papel crucial nas condições ideais de cultivo. Além disso, a restrição às cidades do estado de São Paulo pode ser considerada outras restrições, especialmente para algoritmos de aprendizado de máquina, pois as variações de solo e temperatura nessa região são relativamente pequenas.

Uma questão adicional a ser considerada é que determinados cultivos podem receber uma pontuação bastante baixa, mesmo quando representam casos ideais. Isso ocorre porque o modelo busca assimilar as informações anteriores e encontrar um consenso entre elas, o que pode resultar na subestimação das condições ideais para determinadas culturas. Essa limitação pode impactar a precisão das recomendações, especialmente quando a variação nas condições ideais é significativa. Portanto, é importante estar ciente dessa ponderação durante a interpretação dos resultados do modelo.

5. Considerações finais

O modelo proposto é uma ferramenta simples e eficiente para recomendar cidades para cultivar uma determinada cultura na estação do ano escolhida pelo usuário. O modelo utiliza apenas três variáveis para calcular uma pontuação de similaridade e um algoritmo de *machine learning* para realizar uma previsão de quais seriam, na realidade, as cinco cidades ideais do estado de São Paulo para realizar o cultivo nas condições escolhidas, as quais são ordenadas em ordem decrescente de similaridade.

6. Referências

1. <https://pt.weatherspark.com/>
2. <https://bdiaweb.ibge.gov.br/#/home>
3. [https://www.embrapa.br/hortalicas/batata-doce/ecofisiologia-e-exigencias-climaticas#:~:text=A%20temperatura%20ideal%20para%20o,%C2%BAC%20\(diurna%2Fnoturna\)](https://www.embrapa.br/hortalicas/batata-doce/ecofisiologia-e-exigencias-climaticas#:~:text=A%20temperatura%20ideal%20para%20o,%C2%BAC%20(diurna%2Fnoturna))
4. http://www.sapc.embrapa.br/arquivos/consorcio/publicacoes_tecnicas/livro_implantacao_cafe_zais.pdf
5. <https://www.esalq.usp.br/visaoagricola/sites/default/files/va06-manejo01.pdf>
6. <https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/arroz/pre-producao/caracteristicas/clima#:~:text=A%20temperatura%20%C3%B3tima%20para%20o,abaixamento%20brusco%2C%20ap%C3%B3s%20a%20flora%C3%A7%C3%A3o>

7. <https://www.embrapa.br/hortalicas/batata/clima#:~:text=A%20temperatura%20ideal%20para%20o,pouco%20acima%20de%2015%20%C2%BAC>
8. <https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/feijao/pre-producao/clima#:~:text=Para%20que%20o%20feijoeiro%20possa,preju%C3%ADzos%20ao%20rendimento%20do%20feijoeiro>
9. <https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/milho/pre-producao/caracteristicas-da-especie-e-relacoes-com-o-ambiente/relacoes-com-o-clima#:~:text=A%20temperatura%20ideal%20para%20o,gr%C3%A3os%20na%20temperatura%20de%2021%C2%BAC>
10. <https://www.cati.sp.gov.br/porta1/produtos-e-servicos/publicacoes/acervo-tecnico/aspectos-climatologicos-na-cultura-da-banana#:~:text=Temperaturas%20entre%2020%20a%2029,ideal%20para%20o%20bom%20desenvolvimento>
11. <http://www.iea.sp.gov.br/out/TerTexto.php?codTexto=13614#:~:text=Em%20rela%C3%A7%C3%A3o%20%C3%A0%20temperatura%2C%20a,%C2%B0C%20%C3%A9%20nulo11>
12. <https://www.embrapa.br/hortalicas/cebola/clima#:~:text=Considerando%2Dse%20a%20velocidade%20e,de%2020%2D25%C2%B0C>
13. <https://www.embrapa.br/hortalicas/tomate-de-mesa/clima#:~:text=A%20temperatura%20%C3%B3tima%20para%20germina%C3%A7%C3%A3o,muito%20influenciadas%20por%20fatores%20ambientais>
14. <https://globo.rural.globo.com/vida-na-fazenda/como-plantar/noticia/2013/12/como-plantar-laranja.html>
15. <https://www.cnabrazil.org.br/assets/arquivos/215-CACAU.pdf>
16. <https://globo.rural.globo.com/vida-na-fazenda/como-plantar/noticia/2018/04/como-plantar-limao.html>
17. <https://globo.rural.globo.com/vida-na-fazenda/como-plantar/noticia/2014/10/como-plantar-uva.html>
18. <https://www.embrapa.br/hortalicas/cenoura/clima2#:~:text=A%20cenoura%20%C3%A9%20uma%20hortali%C3%A7a,10%20dias%20ap%C3%B3s%20o%20plantio>
19. [https://www.embrapa.br/hortalicas/batata-doce/ecofisiologia-e-exigencias-climaticas#:~:text=A%20temperatura%20ideal%20para%20o,%C2%BAC%20\(diurna%2Fnoturna\)](https://www.embrapa.br/hortalicas/batata-doce/ecofisiologia-e-exigencias-climaticas#:~:text=A%20temperatura%20ideal%20para%20o,%C2%BAC%20(diurna%2Fnoturna))
20. <https://agronline.com.br/porta1/artigo/clima-e-solo-ideais-para-o-cultivo-de-soja#:~:text=N%C3%A3o%20cultivar%20a%20soja%20em,acima%20de%201%2C0%20m>
21. https://agroadvance.com.br/blog-plantio-de-cafe-7-pontos-relevantes/#15_Solo

22. <https://agriculture.basf.com/br/pt/conteudos/cultivos-e-sementes/algodao/preparacao-sementes.html#:~:text=Como%20o%20algod%C3%A3o%20%C3%A9%20bastante,e%20f%C3%B3sforo%20em%20concentra%C3%A7%C3%B5es%20moderadas>
23. <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/574117/os-solos-e-a-cultura-do-arroz#:~:text=Os%20solos%20mais%20usados%20para,aproximadamente%2059%25%20do%20territ%C3%B3rio%20brasileiro>
24. <https://eos.com/pt/blog/tipos-de-solo/>
25. https://www.sbcs-nrs.org.br/docs/Manual_de_Calagem_e_Adubacao_para_os_Estados_do_RS_e_de_SC-2016.pdf