

INFERÊNCIA DE PRECIFICAÇÃO DE PLANTAS ORNAMENTAIS UTILIZANDO INTELIGÊNCIA COMPUTACIONAL COM BASE NO CICLO DE VIDA DE PLANTAS

Professor orientador: Pedro Henrique Rodrigues
Mendes

Aluno: Leão Pereira Neto

PROGRAMA DE
INICIAÇÃO CIENTÍFICA
PIC/CEUB

RELATÓRIOS DE PESQUISA
VOLUME 9 Nº 1- JAN/DEZ
•2023•





**CENTRO UNIVERSITÁRIO DE BRASÍLIA - CEUB
PROGRAMA DE INICIAÇÃO CIENTÍFICA**

LEÃO PEREIRA NETO

**INFERÊNCIA DE PRECIFICAÇÃO DE PLANTAS ORNAMENTAIS UTILIZANDO
INTELIGÊNCIA COMPUTACIONAL COM BASE NO CICLO DE VIDA DE
PLANTAS**

Relatório final de pesquisa de Iniciação Científica apresentado à Assessoria de Pós-Graduação e Pesquisa.

Orientação: Pedro Henrique Rodrigues Mendes

BRASÍLIA

2024

RESUMO

O crescente volume de dados gerado pela expansão tecnológica em diversos setores intensificou a importância da ciência de dados e da inteligência artificial, a pesquisa visa conectar essas duas áreas em ascensão: aprendizado de máquina e botânica com finalidade para melhorar a qualidade e competitividade no varejo de plantas. Focando na espécie *Sunpatiens*, conhecida por sua resistência e capacidade de floração consistente em diversas condições climáticas, a pesquisa explora o desenvolvimento de algoritmos preditivos para aprimorar a detecção de padrões na comercialização e manejo das *Sunpatiens*. O estudo considera os custos empresariais relacionados ao cultivo e comercialização da *Sunpatiens*, integrando tecnologia para melhorar a qualidade e competitividade no mercado e aplicando nas inferências fatores de custos e preços de venda. Foram realizadas previsões sobre os custos e preços de venda, com a possibilidade de expandir essa análise para outras espécies e incluir tecnologias de análise de solo. Neste projeto, ficou evidente que os custos diários tendem a superar os lucros, resultando em perdas significativas após a primeira semana. O estudo identificou que a venda da planta deve ocorrer na primeira semana após a floração, antes que os custos superem o preço de venda. O uso de algoritmos, como o Random Forest, concluiu na necessidade de uma rápida comercialização para circular o produto e minimizar as perdas empresariais, se aproveitando do ciclo de floração. Os resultados da pesquisa sugerem duas principais ações, a primeira que é a alteração programada do preço de venda das amostras com base no o ciclo de vida da espécie, a segunda ação indica um controle de estoque mais eficiente, determinando os melhores dias para venda e separando os principais fatores que mais influenciam as vendas. Portanto, ao aplicar todas as indicações compiladas neste estudo de otimização de precificação e manejo da planta, proporciona um lucro registrado acima de 40% no comércio de *Sunpatiens*.

Palavras-chave: Comércio de plantas; Deep Learning; *Sunpatiens*.

FIGURAS

Figura 1 – crescimento do aprendizado de máquina

Figura 2 – Sunpatiens florida

Figura 3 – Diferença de performance entre métodos modernos e antigos de aprendizado de máquina

Figura 4 – Funcionamento de RNN

Figura 5 – Funcionamento de uma rede LSTM

Figura 6 – Gráfico de evolução dos preços de *Sunpatiens* sem aplicação de adubo.

Figura 7 – Gráfico de evolução dos preços de *Sunpatiens* com aplicação de adubo.

SUMÁRIO

RESUMO	2
FIGURAS	3
INTRODUÇÃO	5
FUNDAMENTAÇÃO TEÓRICA	9
Custos empresariais	10
Legislação federal de impostos e taxas	12
Sunpatiens	14
Aprendizado de máquina por redes neurais	17
Redes Neurais Recorrentes (RNN)	19
Redes Neurais de Memória de Longo e Curto prazo (LSTM)	21
Random Forest	24
MÉTODO	25
Formação de preços	25
Custos da empresa	25
Formação de preços – Sunpatiens	28
Seletiva das plantas e manutenção	29
Observação e divisão das espécies	33
Codificação	34
RESULTADOS E DISCUSSÃO	38
Dos cálculos empresariais	38
Preço de venda	46
Machine Learning	48
CONSIDERAÇÕES FINAIS	52
REFERÊNCIAS	54

INTRODUÇÃO

O volume crescente de dados com a expansão do uso da tecnologia em diversos setores trouxe consigo um crescimento da importância da área de ciência de dados e também do uso da inteligência artificial¹. Outra área que cresceu os estudos e aplicação da tecnologia foi o voltado às plantas (estudos de espécies, melhoramento genético, vendas, etc.). Com o intuito de desenvolver uma ligação entre essas duas áreas crescentes, foi desenvolvido essa pesquisa.

A pandemia COVID-19, com suas consequências no comportamento humano, trouxe consigo um aumento de 40% na demanda por plantas² como forma de alívio do estresse e melhoria do bem-estar durante períodos de isolamento social. A transição para o home office incentivou muitos a reimaginar seus espaços de vida, principalmente no que tange ao serviço público³.

Esse fenômeno não se limitou apenas à aquisição pessoal para embelezar residências; o distanciamento social redefiniu as flores como um meio de expressar afeto e lembrança em ocasiões especiais, como o Dia das Mães e o Dia dos Namorados, através do serviço de entrega de flores⁴. Com a retomada das visitas sociais, incluindo idas a cemitérios, o setor de flores e arranjos florais viu um aumento adicional na demanda.

Diante desse cenário de mercado em ascensão, diversos fatores influenciam o preço das plantas, incluindo datas comemorativas, condições climáticas, dificuldades

¹ IBM. O que é ciência de dados? **Tópicos Ciência de dados**, 2024. Disponível dia <https://www.ibm.com/br-pt/topics/data-science>. Acesso dia 22 julho 2024.

² JORNAL NACIONAL. Pandemia faz a venda de plantas e flores aumentar em todo país. **Globo - G1**, 2021. Disponível dia <https://g1.globo.com/jornal-nacional/noticia/2021/01/29/pandemia-faz-a-venda-de-plantas-e-flores-aumentar-em-todo-pais.ghtml>. Acesso dia 22 julho 2024.

³ LIMA, T. et al. Mudança organizacional: o impacto do home office no comportamento das servidoras e estagiárias da diretoria do Ministério Público de Contas. **Revista de Gestão e Secretariado**, São Paulo, 14 junho 2023. 9033-9051.

⁴ FREDENBURGH, J. How the Covid-19 pandemic hit the cut-flower chain. **BBC**, 2020. Disponível dia <https://www.bbc.com/future/bspoke/made-on-earth/how-the-covid-19-pandemic-hit-the-cut-flower-chain.html>. Acessado dia 18/02/2024, as 08:57.

de cultivo, raridade das espécies, tendências de mercado e até mesmo a idade da planta, que afeta diretamente sua qualidade e tamanho. Assim, torna-se imperativo desenvolver um método eficiente para prever o preço de uma planta com base em seu ciclo de vida, visando otimizar o marketing desses produtos, fornecer informações precisas aos consumidores e, sobretudo, oferecer aos lojistas uma noção mais acurada para a gestão de seus estoques.

Portanto, a priori a integração da tecnologia, especialmente a inteligência artificial, na agricultura e no cultivo de plantas tem sido gradual, mas constante. No Brasil, a implementação dessas tecnologias é relativamente recente, mas já se observa seu uso para detecção de doenças em plantas. A coleta de dados e o uso de computação em nuvem estão começando a ser aplicados para prever o tempo ideal de plantio, avaliar a qualidade das plantas e outros aspectos vitais, de maneira semelhante ao cuidado com a saúde humana, com o objetivo de prolongar a vida e a qualidade dos vegetais.

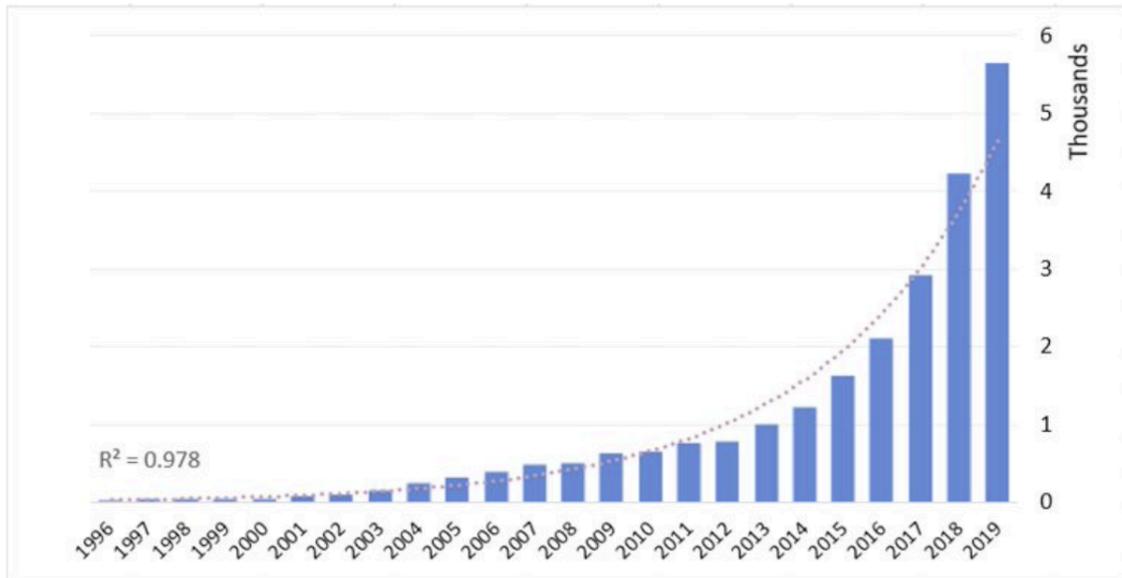
Dentro da agricultura e mesmo o cultivo de plantas, vem aos poucos sendo complementado com recursos tecnológicos de inteligência artificial. Muito embora sua implementação aqui no Brasil seja algo mais recente, é possível ver o uso de inteligência artificial no para detecção de doenças em 2022⁵.

Com a chegada da tecnologia, não foi possível manter a eficiência de análise dos dados gerados sobre a biologia, necessitando de uma implementação de aprendizado de máquina⁶. Isso devido a capacidade de elaborar aprendizados e testes em uma capacidade sobre humana o que ajuda em diversas aplicações. Considere a ajuda do aprendizado de máquina conforme a Figura 1.

⁵ COSTA, V. C. Inteligência artificial identifica plantas doentes simulando processo cerebral. **Embrapa**, 2023. Disponível dia <https://www.embrapa.br/busca-de-noticias/-/noticia/78204383/inteligencia-artificial-identifica-plantas-doentes-simulando-processo-cerebral>. Acessado dia 13/04/2023 as 08:56

⁶ WALSH, I.; FISHMAN, D.; AL, E. Recommendations for machine learning validation in biology. **Aristotle University of Thessaloniki**, 2020. Disponível dia https://www.researchgate.net/publication/342547853_Recommendations_for_machine_learning_validation_in_biology. Acessado dia 10 de julho de 2024, as 07:52.

Figura 1 - aumento exponencial dos estudos sobre aprendizado de máquina em biologia



Fonte: Recommendations for machine learning validation in biology

Além disso, a coleta de dados e utilização de nuvens vem sendo aos poucos aplicada para antecipação do tempo adequado de plantio⁷, verificação da qualidade das plantas e outros fatores que tratem da vida das plantas, similar ao aplicado aos humanos, para preservar por mais tempo a vida e a qualidade dos vegetais.

Embora programas de computador auxiliem na coleta de dados e na análise de tendências de vendas, há aspectos do comércio de plantas que permanecem desafiadores para prever, mesmo com anos de experiência. O presente trabalho visa explorar a aplicação da ciência de dados e da inteligência artificial no comércio varejista e atacadista de plantas ornamentais tendo como ideal facilitar o empreendedorismo em uma área que apresenta muito espaço para aplicação da tecnologia e representa boa parte da força trabalhadora local⁸.

Diversos programas de informática ajudam a coletar dados e entregar ao final do mês ou do ano o fluxo de plantas, em que meses houve alta ou baixa

⁷ BANCO DO BRASIL. Inteligência artificial no agronegócio. **Blog BB**, 2021. Disponível dia <https://blog.bb.com.br/inteligencia-artificial-agronegocio/>. Acessado dia 13/04/2023 as 09:00

⁸ CARAMORI, I. DF tem o mesmo número de empreendedores e servidores públicos. **Metrópolis**, 2019. Disponível dia <https://www.metropoles.com/distrito-federal/economia-df/df-tem-o-mesmo-numero-de-empreendedores-e-servidores-publicos>. Acesso dia 27 julho 2024, as 22:34.

comercialização de produtos. São informações que a depender da experiência do comerciante, já se tornam previsíveis mesmo sem ajuda da coleta de dados. Contudo, há áreas dentro do comércio de plantas que não são de grande facilidade de prever mesmo com anos de experiência.

O controle de estoque, geralmente baseado na análise do fluxo de produtos, é afetado por variáveis sazonais e pelo ciclo de vida das plantas, que determinam o valor comercial do produto. Portanto, é crucial para os comerciantes terem acesso a previsões confiáveis sobre o preço de venda ideal para evitar perdas e negociar valores de forma mais informada.

Para o lojista, o ciclo de vida de cada planta pode determinar o valor de comércio do produto. Mesmo em condições ideais de tratamento, há uma perda da qualidade da planta que impacta diretamente no seu valor de revenda.

Nesse espectro o comerciante necessita de uma previsão sobre qual deve ser o seu preço de venda para não perder totalmente o produto e ainda poder negociar na hora da compra o valor, caso já não esteja em condições ideais de venda.

Este estudo se concentra na necessidade de uma análise preditiva que considere o ciclo de vida das plantas, permitindo antecipar flutuações no valor de venda e, assim, auxiliar os lojistas no controle de estoque. Através de uma abordagem inovadora que combina ciência de dados e inteligência artificial, buscamos desenvolver um modelo que possa prever a variação de preço das plantas, considerando fatores como idade, sazonalidade e demanda do mercado.

Neste contexto, o presente trabalho foca na espécie Sunpatiens, conhecida por sua vivacidade e resistência, e como ela se enquadra nesse mercado em transformação. Primeiramente, avaliamos os custos empresariais associados ao cultivo e comercialização da espécie selecionada, considerando os investimentos em tecnologia e inovação para manter a qualidade e competitividade no mercado. Em seguida, desenvolvemos um modelo de codificação que utiliza ciência de dados e inteligência artificial para prever o preço das Sunpatiens, levando em conta variáveis como ciclo de vida, sazonalidade e tendências de consumo. Por fim, apresentaremos os resultados obtidos com o modelo proposto, discutindo sua eficácia e possíveis

implicações para o mercado de plantas ornamentais, especialmente para os lojistas que buscam uma gestão de estoque mais eficiente e lucrativa.

FUNDAMENTAÇÃO TEÓRICA

A ideia desenvolvida neste trabalho perpassa algumas etapas do comércio de produtos e também sobre a análise de dados e predição de valores futuros através de algoritmos de aprendizado de máquina. Em virtude dessas etapas foram estudados autores diversos sobre o tema de Gestão de Custos e Formação de preços em empresas, além de colaboração com uma empresa de pequeno porte que forneceu uma visão prática sobre o que os autores tratavam em suas obras acadêmicas.

Há muitas décadas tem havido crescente interesse no desenvolvimento de inteligências artificiais que promovam o auxílio em diversas questões do dia a dia⁹. Isso implicou em uma diversidade de aplicações comerciais¹⁰ e não comerciais como os trazidos pela Embrapa¹¹. Contudo a aplicação especificamente para os objetivos tratados neste artigo não existe de forma abrangente, tendo que se recorrer a assuntos conexos e afins.

Esclarecido a questão de relacionamento com obras existentes e os assuntos aqui tratados, esta fundamentação será desenvolvida trazendo em primeiro plano as questões comerciais e logo após as relacionadas ao aprendizado de máquina.

⁹ COSTA, Anna Helena; COZMAN, Fabio G. Revista USP, São Paulo, n 141, p. 133-46, abril/maio/junho 2024, pg. 136

¹⁰ Posso citar como exemplo a questão dos produtos relacionados e propagandas anexadas nos buscadores, que utilizam programação baseada em dados capturados para realizar com eficiência as indicações

¹¹ COSTA, V. C., op. Cit., acessado dia 14/04/2023 as 10:13.

Custos empresariais

Antes de mais nada é preciso determinar o que seriam os custos para após determinar como seria formados os preços, ao menos é um caminho seguido tanto por autores como Dubois e Souza¹², assim como por Antonik que relembra que o objetivo de uma empresa ainda que não divulgado é a maximização da riqueza e para isso é preciso determinar os custos e como avaliar o retorno diante dos riscos do negócio¹³.

Custos podem ser considerados diretos ou indiretos. E como o próprio nome indica os gastos são assim classificados em relação a ligação com a empresa, por exemplo, é considerado custo direto o valor pago aos trabalhadores, calculando conforme o valor da hora trabalhada pela quantidade de horas trabalhadas ao final do mês¹⁴. Já os custos indiretos são aqueles que envolvem cálculos de risco e depreciação ou de produto ou de equipamentos¹⁵.

Há ainda outros dois tipos de custos que são relacionados ao volume de produção: fixos e variáveis. Aqueles são os que sempre estarão presentes durante a existência da empresa: água, luz, telefone, IPTU, seguro. Estes valores independem inclusive da quantidade de comercialização de produtos ou mesmo de sua produção¹⁶. Já os variáveis seriam aqueles que no decorrer do tempo variam como a flutuação de preços no mercado: preço da matéria prima, combustível, horas extras são bons exemplos desse tipo de custo.

A soma de todos os custos formam o chamado custo total, que pode gerar uma ideia para o proprietário do negócio, o valor do custo unitário¹⁷. Calculados esses custos pode-se começar a calcular o preço final do produto. Note que, além dos custos e despesas da empresa, uma regra clara que deve-se considerar para formação do preço de venda é a lei da oferta e demanda¹⁸.

¹² DUBOIS, A.; KUPLA, L.; SOUZA, L. E. D. **Gestão de Custos e Formação de Preços - Conceitos, Modelos e Ferramentas**. 4ª. ed. São Paulo: Atlas, 2024, p. 27

¹³ ANTONIK, L. R. **Empreendedorismo**. Kindle. ed. São Paulo: Alta Books, 2023, p. 6.

¹⁴ DUBOIS, A.; KUPLA, L.; SOUZA, L. E. D. Op. Cit.

¹⁵ DUBOIS, A.; KUPLA, L.; SOUZA, L. E. D. Ibidem, p. 28.

¹⁶ DUBOIS, A.; KUPLA, L.; SOUZA, L. E. D. Ibidem, p. 29.

¹⁷ DUBOIS, A.; KUPLA, L.; SOUZA, L. E. D. Ibidem, p. 33

¹⁸ ANDRADE, R.; PALLUETO, A. W. Aspectos institucionais da formação de preços. **Revista Nova Economica**, Minas Gerais, 11 outubro 2019. 565-90, p. 567.

O produto em questão, planta ornamental, apesar de variações e um aumento durante a pandemia, apresenta-se com uma demanda mais constante, o que ajuda na formação do valor de venda. Tirando essa questão de demanda do produto, o que se considera como necessário para a determinação do valor de venda é o custo de produção (C), tributos(T) e despesas(D), margem de lucro desejada(L). Portanto o cálculo faria:

$$P = C + D + T + L^{19}$$

Assim temos a formação dos custos totais e o método para a formação do preço de venda. Contudo, essa fórmula pode se tornar mais complexa. Segundo Antonik²⁰, essa média de valores que são influenciados pela oferta e procura, podem ser obtidos através de uma administração por objetivos. Esse tipo de administração que se apresenta por uma revisão constante dos objetivos a serem alcançados podem apresentar uma complexidade para os algoritmos de aprendizagem de máquina quando aplicados sem prazo determinado, já que envolveriam uma atualização constante.

Outro tipo de análise para composição do valor da mercadoria seria a análise SWOT que envolve além de estudos sobre fatores internos da empresa, também uma análise externa de oportunidades e ameaças como forma de validação do cenário atual e formação do valor. Tanto a metodologia SWOT quanto a APO, envolvem os cálculos citados anteriormente por Alexy e Luiz Eurico²¹, adicionados de uma análise do cenário que sugere a observação da concorrência e outros fatores, sem necessariamente apresentar uma fórmula exata de como será constituído o valor final de venda.

Outros tipos de estudos sobre outros índices como os de inflação e análise vertical e horizontal que também são citados por Antonik²². São estilos de formação de valor que envolvem atualizações baseadas no histórico de preços ao longo dos anos e projeções de atualização conforme o índice da inflação. Métodos que podem ser utilizados para uma possível atualização do projeto em outro momento.

¹⁹ DUBOIS, A.; KUPLA, L.; SOUZA, L. E. D. Op. Cit, p. 228

²⁰ ANTONIK, L. R., Op. Cit, p. 126.

²¹ DUBOIS, A.; KUPLA, L.; SOUZA, L. E. D Op. Cit, p. 28

²² ANTONIK, L. R., Op. Cit, p. 135

Legislação federal de impostos e taxas

O pequeno comerciante, um dos objetos de estudo desta pesquisa, não segue o mesmo regime tributário de grandes empresas. Por isso é de grande valia entender um pouco sobre a legislação nacional e as determinações legais das empresas que atuam no Distrito Federal como forma de fundamentação dos métodos aplicados a este projeto.

A Lei Complementar nº 123 de dezembro de 2006²³, institui o Estatuto Nacional da Microempresa e da Empresa de Pequeno Porte. Essa lei estabelece normas gerais relativas ao tratamento diferenciado e favorecido a ser dispensado às microempresas e empresas de pequeno porte no âmbito dos Poderes da União, dos Estados, do Distrito Federal e dos Municípios.

Especificamente para empresas que atuam na área de venda de plantas e produtos de decoração de pequeno porte, ela estabelece o chamado simples nacional²⁴, uma modalidade destinada ao cálculo da chamada taxa real: um percentual aplicado sobre a receita bruta mensal da empresa para determinar o montante do desconto sobre o lucro.

Abaixo está a Tabela 1 com as taxas e a Partilha do Simples Nacional – Comércio em vigor desde 01/01/2018.

Tabela 1 - Partilha do Simples Nacional

	Receita Bruta em 12 Meses	Alíquota	Valor a Deduzir
1ª Faixa	Até R\$ 180.000,00	4,00%	–

²³ BRASIL. Lei Complementar Nº 123, de 14 de dezembro de 2006. **Estatuto Nacional da Microempresa e da Empresa de Pequeno Porte**, Brasília, 2006. Disponível dia https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp123.htm. Acessado dia 17/01 as 10:05

²⁴ Um regime tributário que oferece simplificação no pagamento de tributos e benefícios para essas empresas.

2ª Faixa	De R\$180.000,01 a R\$360.000,00	7,30%	R\$5.940,00
3ª Faixa	De R\$360.000,01 a R\$720.000,00	9,50%	R\$13.860,00
4ª Faixa	De R\$720.000,01 a 1.800.000,00	10,70%	R\$22.500,00
5ª Faixa	De R\$1.800.000,01 a R\$3.600.000,00	14,30%	R\$87.300,00
6ª Faixa	De R\$3.600.000,01 a R\$4.800.000,00	19,00%	R\$378.000,00

Através dessa tabela, conseguiremos encontrar a taxa NOMINAL e a dedução correspondente, identificando em qual intervalo de rendimento a empresa se enquadra. Ao encontrar a faixa da taxa NOMINAL aplicam-se os percentuais da Tabela 2 a uma fórmula e se encontra a taxa real.

Tabela 2 - Incidência de tributos conforme taxa real

Faixas	Percentual de Repartição dos Tributos					
	IRPJ	CSLL	COFINS	PIS/PASEP	CPP	ICMS
1ª Faixa	5,50%	3,50%	12,74%	2,76%	41,50%	34,00%
2ª Faixa	5,50%	3,50%	12,74%	2,76%	41,50%	34,00%

3ª Faixa	5,50%	3,50%	12,74%	2,76%	42,00%	33,50%
4ª Faixa	5,50%	3,50%	12,74%	2,76%	42,00%	33,50%
5ª Faixa	5,50%	3,50%	12,74%	2,76%	42,00%	33,50%
6ª Faixa	13,50%	10,00%	28,27%	6,13%	42,10%	–

Nesse regime especial, as empresas têm uma carga tributária simplificada e unificada, que pode incluir contribuições para a seguridade social a cargo da pessoa jurídica, incluindo:

- FGTS (Fundo de Garantia do Tempo de Serviço)²⁵: a alíquota de contribuição é de 8% do salário bruto do empregado.
- 13º salário²⁶: correspondente a 1/12 do salário por mês trabalhado.
- Férias: A remuneração das férias é acrescida de 1/3 conforme o artigo 7º, inciso XVII, da Constituição Federal.
- Contribuições previdenciárias: a alíquota total varia de acordo com a atividade da empresa e o enquadramento tributário, mas para o cálculo sobre férias, FGTS e Descanso Semanal Remunerado (DSR), utiliza-se uma alíquota média de 7,93%.
- FGTS/Provisão de multa para rescisão: Além dos 8% mensais, as empresas devem depositar 4% para a provisão de multa rescisória do FGTS, em caso de demissão sem justa causa.

Esses valores são consistentes com as práticas padrão de encargos trabalhistas, custos fixos de impostos e luz que fazem parte do cálculo de funcionamento de toda empresa, ainda que pequena comerciante de plantas ornamentais.

²⁵ O FGTS é regulamentado pela Lei nº 8.036/1990 e pelo Decreto nº 99.684/1990.

²⁶ O pagamento do 13º salário é garantido pela Lei nº 4.090/1962, com a alíquota de 8,33%.

Sunpatiens

A Sunpatiens, uma espécie híbrida de impatiens originária do Japão. Criada a partir de uma espécie de impatiens da Nova Guiné, a Sunpatiens é o resultado de um melhoramento genético realizado pela empresa Sakata. Esta planta, que pode atingir uma altura de até 50 centímetros, foi desenvolvida para ser mais resistente a uma variedade de ambientes e climas. Sua robustez e a capacidade de florescer durante longos períodos ao longo do ano a tornam uma escolha popular para aqueles que desejam manter seus jardins coloridos e vibrantes, como exemplo da Figura 2.

Figura 2 - Sunpatiens já com sua floração, no momento de entrega do fornecedor para o comerciante



A Sunpatiens é comercializada em três séries distintas, cada uma adaptada a diferentes condições e preferências de cultivo:

- Sunpatiens Compact: Esta variedade é ideal para locais que recebem bastante sol, calor e umidade, sendo perfeita para cestos suspensos, jardins verticais e vasos colocados em janelas.
- Sunpatiens Spreading: Com suas flores brancas delicadas, é ótima para áreas ensolaradas, onde pode se espalhar graciosamente.
- Sunpatiens Vigorous: Conhecida por sua resistência a chuvas, ventos e geadas, é a escolha certa para cobrir grandes espaços e criar um impacto visual marcante.

Para um cultivo bem-sucedido, a Sunpatiens requer um equilíbrio entre exposição ao sol e sombra, um solo bem drenado e rico em húmus, e um espaçamento adequado entre as plantas para promover um crescimento saudável²⁷. Embora o cultivo a partir de sementes possa ser desafiador, suas mudas são altamente comercializáveis, oferecendo uma opção acessível para jardineiros e paisagistas²⁸. A combinação de sua popularidade, facilidade de cultivo e preço acessível faz da Sunpatiens uma escolha ideal para este estudo, onde buscamos entender não apenas suas características botânicas, mas também seu potencial de mercado e as implicações de seu cultivo e comercialização.

A comercialização dessa espécie no distrito federal é feita por meio de mudas que giram em torno de 10 a 15 centímetros em sua fase inicial e já na fase adulta com 50 centímetros. Para este estudo, as referências de tamanhos seguirão conforme o considerado pelos lojistas: pequeno porte seriam as de 15 a 17 centímetros; médio porte considera-se as de 18 centímetros até os 30 centímetros; grande porte dos 31 até os 50 centímetros.

Com observação dentro de empresas de pequeno porte de comercialização de plantas, foi possível constatar que a forma comum de aquisição se dá por meio de

²⁷ HASSANI, N. How to Grow and Care for SunPatiens. **The Spruce Make your best home**, 2024. Disponível dia <https://www.thespruce.com/sunpatiens-growing-guide-5202267>. Acessado dia 20 de fevereiro de 2024, as 20:47.

²⁸ SAKADA SEED AMERICA. SunPatiens Culture Guide. **Kientzler**, 2024. Disponível dia <https://www.kientzler.eu/north-america/culture-information/Sakata/SunPatiens-Culture-Guide.pdf>. Acessado dia 20/02, as 20:50.

caixarias para as menores plantas e pote para as de tamanho médio e grande. Essa aquisição e classificação dos portes é corroborada por estudos que indicam a importância do tamanho na comercialização de plantas ornamentais²⁹.

Diversos fatores que para o consumidor são apenas visuais como a cor da folha e sua aparência, a vitalidade da espécie e a floração, também são fatores que estudiosos utilizam para criar sistemas de cultivo que ao final resultem no fornecimento de material vegetal de alta qualidade³⁰ para os mercadores mais exigentes.

Cada caixaria de plantas pequenas contém um total de 15 mudas em sacos individuais. Para fins de estudo e coleta de dados, foram selecionadas 4 caixas aleatórias e estabelecido um prazo de observação de 90 dias, sendo 60 dias para observação do desenvolvimento com fatores diferenciados para as 4 caixas e 30 dias para observação da venda após o desenvolvimento da planta nos dias anteriores.

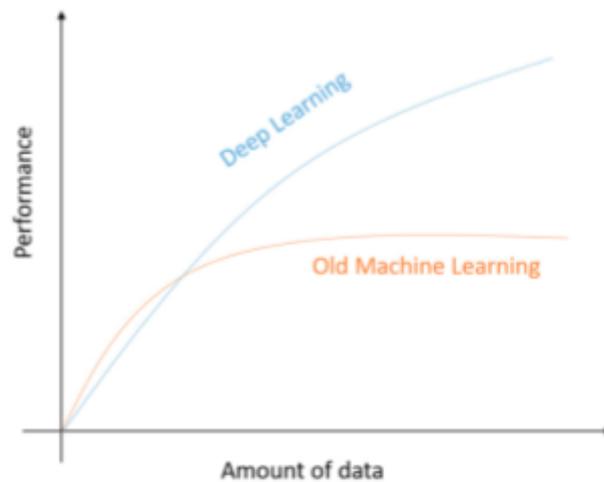
Aprendizado de máquina por redes neurais

Deep Learning, ou **Aprendizado Profundo**, é um conjunto de algoritmos em aprendizado de máquina que tenta modelar abstrações de alto nível em dados usando arquiteturas compostas de várias transformações não-lineares. Redes de Deep Learning são uma subclasse de redes neurais artificiais onde são construídas com várias camadas, cada uma adicionando um nível de complexidade ou abstração. A Figura 1 dá uma noção do diferencial de performance dos métodos anteriores de aprendizado de máquina.

²⁹ DATTA, S. K. Breeding of ornamentals: success and technological status. **The Nucleus**, California, 25 Agosto 2022. 107-28. p. 110.

³⁰ CARDOSO, J.; VENDRAME, W. Innovation in Propagation and Cultivation of Ornamental Plants. **Environmental horticulture Department**, Gainesville, 2 march 2022. p. 229.

Figura 3 – diferença de performance entre os modelos de aprendizado de máquina e aprendizado profundo



Fonte - imagem original do trabalho acadêmico: A State of art survey on Deep Learning Theory

Essa complexidade inerente às redes de Deep Learning as torna particularmente sensíveis à natureza da base de dados utilizada para treinamento e validação. A performance de cada espécie de rede de aprendizagem pode variar significativamente dependendo de vários fatores, tais como:

- **Composição da Base de Dados:** A proporção de variáveis numéricas em relação às textuais pode afetar a escolha do modelo ideal.
- **Presença de Valores Nulos:** Redes diferentes têm capacidades distintas de lidar com dados incompletos ou ausentes.
- **Extensão da Base de Dados:** Redes mais complexas, como as de Deep Learning, tendem a performar melhor com grandes volumes de dados, pois bases de dados menores aumentam o risco de **overfitting**³¹.

Portanto, ao selecionar uma rede de aprendizado para uma aplicação específica, é essencial considerar esses aspectos para otimizar a performance do modelo, inclusive no que tange a probabilidade a acontecimentos³². Pois mesmo com que seja necessário por exemplo uma implementação de dados artificiais, a depender

³¹ ALOM, Z.; TAHA, T. M.; YAKOPCIC, C. E. A. A State-of-the-Art Survey on Deep Learning Theory and Architectures. **Electronics**, Basel, 31 janeiro 2019. 292-358. P. 300.

³² BUDUMA, N.; BUDUMA, N. P. J. **Fundamentals of Deep Learning**. Kindle. ed. Sebastopol: O’Reilly Media, 2022. P. 459

da quantidade, é necessário repensar no modelo ao invés de pensar na implementação de dados.

Redes Neurais Recorrentes (RNN)

As **Redes Neurais Recorrentes (RNNs)** são uma classe de redes neurais que se destacam no trabalho com dados sequenciais³³. Ao contrário das redes neurais tradicionais, que tratam as entradas como independentes, as RNNs capturam dependências temporais mantendo um estado interno que é atualizado a cada etapa temporal. Essa capacidade as torna adequadas para uma variedade de aplicações, como:

- **Processamento de Linguagem Natural (PLN)**, onde são usadas para tradução automática, análise de sentimentos e modelagem de linguagem.
- **Análise de Séries Temporais**, como previsão de demanda, análise financeira e monitoramento de sensores.
- **Reconhecimento de Fala**, incluindo transcrição de áudio para texto e comandos de voz.
- **Gerenciamento de Diálogos**, como em chatbots e assistentes virtuais.

Particularmente úteis para identificar padrões e tendências em séries temporais, elas podem ser treinadas para sugerir ajustes de preços dinâmicos, personalizar promoções e detectar anomalias no comportamento de compra.

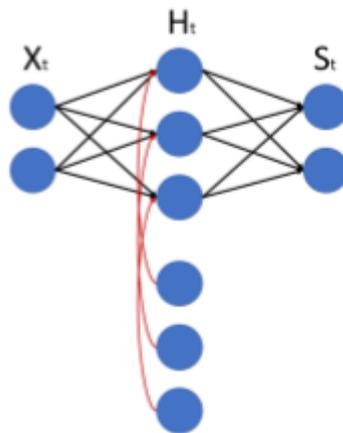
Utilizando a arquitetura dessa rede que normalmente é desenvolvida com auxílio de Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU).³⁴, poderia ser aplicada para determinar a melhor data de venda ou quando aplicar uma promoção da seguinte maneira, ilustrada pela Figura 4³⁵:

³³ EKPA, A. K.; IMO, I. U. Recurrent Neural Network (RNN)-Based Injection Substation Load Prediction And Forecasting. **Journal of Multidisciplinary Engineering Science and Technology (JMEST)**, 1 janeiro 2024. 16899-906. p. 16900

³⁴ DERI SISWARA, A. M. S. A. H. W. Classification Modeling with RNN-Based, Random Forest, and XGBoost for Imbalanced Data: A Case of Early Crash Detection in ASEAN-5 Stock Markets. **Statistics and Data Sciences Department, Faculty of Mathematics and Natural Sciences**, Indonesia, 12 junho 2024.

³⁵ Os nomes atribuídos são "X" para a entrada, "H" para a camada oculta e "S" para a saída. Observa-se que o índice "t" é adicionado a cada um desses neurônios para indicar a dependência temporal na rede.

Figura 4 - funcionamento da RNN



Fonte: Didática Tech – como funcionam as redes neurais recorrentes rnn

- **Entrada de Dados (X_t):** A camada de entrada receberia seqüências de dados históricos de vendas, incluindo informações como datas de vendas anteriores, volumes de vendas, promoções realizadas, feriados e eventos especiais.
- **Processamento Temporal (H_t):** A camada oculta da RNN processaria esses dados sequenciais. Devido às conexões recorrentes, a RNN é capaz de manter um estado interno que reflete as informações dos passos temporais anteriores, permitindo que a rede utilize o contexto das vendas passadas para fazer previsões.
- **Saída de Previsões (S_t):** A camada de saída geraria previsões baseadas no processamento da camada oculta, como prever volumes de vendas para datas futuras ou avaliar o impacto potencial de diferentes estratégias promocionais.

Funcionamento Detalhado:

- **Aprendizado de Padrões:** A RNN aprenderia padrões nos dados, como períodos de alta demanda ou o sucesso de promoções anteriores.
- **Previsão de Tendências:** Com base no aprendizado, a RNN poderia prever tendências futuras, como um aumento esperado nas vendas em determinadas épocas do ano.
- **Otimização de Promoções:** A rede poderia sugerir a melhor data para iniciar uma promoção, maximizando a visibilidade e as vendas.
- **Adaptação Dinâmica:** A RNN se adaptaria a novos dados à medida que fossem coletados, tornando as previsões e recomendações mais precisas com o tempo.

Para treinar essa RNN, seria necessário um conjunto de dados abrangente e bem preparado, além de um processo de validação e teste para garantir que as previsões sejam confiáveis e aplicáveis na prática. Sendo necessário coletar e preparar dados históricos de vendas, treinar o modelo para prever vendas futuras e identificar os melhores momentos para promoções. Embora as RNNs sejam poderosas, outros métodos e modelos podem ser considerados, dependendo da complexidade e natureza dos dados disponíveis.

Ao incorporar Redes Neurais Recorrentes (RNNs) em modelos preditivos, estes se tornam capazes de reter informações de entradas anteriores e utilizá-las como parâmetros para influenciar decisões em entradas futuras. Essa característica das RNNs habilita o processamento de informações ao longo do tempo e o aprendizado de sequências, o que é essencial para tarefas como reconhecimento, reprodução de sequências ou associação e predição baseada no tempo³⁶.

Redes Neurais de Memória de Longo e Curto prazo (LSTM)

Como apresentado antes, RNN armazenam feedbacks de conexões passadas recente para influenciar futuras entradas, simulando uma espécie de memória de curto prazo na rede neural, e a rede neural LSTM tem uma estrutura similar³⁷. Para mitigar os problemas de desaparecimento do gradiente e explosão que podem ocorrer em Redes Neurais Recorrentes (RNNs) padrão, Hochreiter et al. propuseram a Long Short-Term Memory (LSTM), uma arquitetura capaz de lidar com escalas temporais extensas e preservar os erros ao longo do tempo³⁸. Posteriormente, Gers et al. aprimoraram a LSTM com a introdução de uma "porta de esquecimento", possibilitando que as células

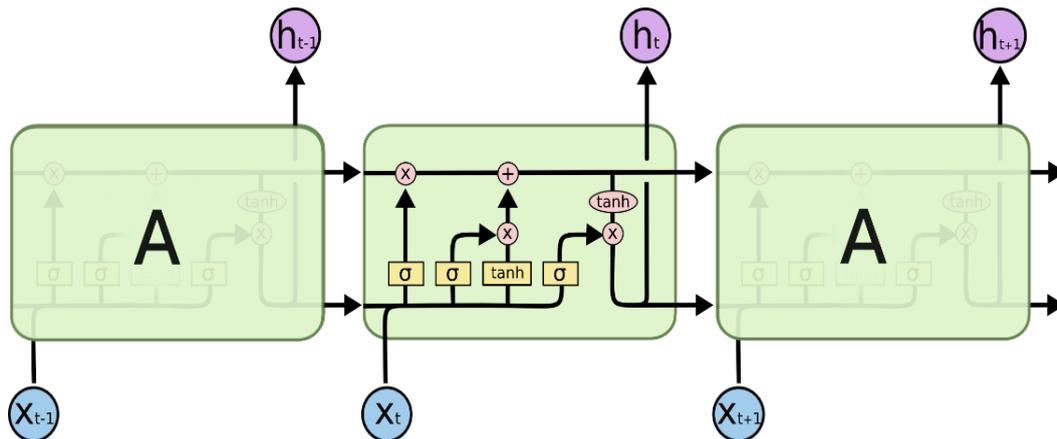
³⁶ BULLINARIA, J. A. Biological Neurons and Neural Networks, Artificial Neurons. **Neural Computation: Lecture 2**, 2015. Disponível dia https://www.cs.bham.ac.uk/~pjt/NC/l2_JB.pdf. Acessado dia 10 de junho de 2024, as 00:04.

³⁷ CHEN, N. Exploring the development and application of LSTM variants. **Proceedings of the 4th International Conference on Signal Processing and Machine Learning. Southwest Jiaotong University, 999 Xian Road**, Sichuan, 5 novembro 2021. 7227–7259. P. 7230

³⁸ HOCHREITER, S. . & S. J. Long Short-Term Memory Spiking Networks and Their Applications. **Neural computation**, 1 novembro 1997. 1735-1780. P. 1750.

da rede resetassem suas informações internas conforme necessário, melhorando assim a gestão da memória³⁹. A Figura 5 traz uma ideia de funcionamento dessa rede.

Figura 5 - Ilustração do funcionamento de uma rede LSTM de perpetuação dos estados ocultos ($h_{(t-1)}$) ao longo do tempo, sinal de entrada (x_t) alimenta os retângulos amarelos que indicam uma camada da rede neural com dois diferentes tipos de funções de ativação: sigmoide (σ) e de tangente hiperbólica (\tanh). Os vetores em preto indicam o fluxo percorrido pelo os sinais. o círculos na coloração rosa indica operações aritméticas com resultados das camadas da rede neural, aplicando operações de soma (+), multiplicação (\times) e tangente hiperbólica (\tanh).



Fonte. colah.github.io

A estrutura da LSTM incorpora o conceito de acumulação, modificação e incorporação dos valores dos estados ocultos anteriores em uma unidade de memória ao decorrer do tempo. A continuidade desses valores é representada na figura 04 pelo vetor horizontal que transita no topo do diagrama, possibilitando o trânsito da informação com a opção de manter o conteúdo inalterado durante o processamento dos sinais na rede neural.

$$f_t = \sigma(w_f \times [h_{t-1}, x_t] + b_f) \quad (6)$$

$$i_t = \sigma(w_i \times [h_{t-1}, x_t] + b_i) \quad (7)$$

$$c_t = \tanh(w_c \times [h_{t-1}, x_t] + b_c) \quad (8)$$

Para possibilitar a remoção e adição de valores ao estado oculto, a arquitetura LSTM emprega uma estrutura conhecida como Gates. Conforme descrito por Olah

³⁹ GERS FA, S. J. C. F. Learning to forget: continual prediction with LSTM. **Neural Comput**, 12 outubro 2000. 2451-2471. P. 2460.

(2015), esses Gates consistem em uma camada neural que utiliza funções de ativação do tipo sigmoide (σ) ou tangente hiperbólica (\tanh)⁴⁰. O primeiro Gate, caracterizado pela função sigmoide na figura 4, é denominado "Porta de Esquecimento". Quando a saída dessa função é zero, ela elimina os valores previamente armazenados no estado oculto (h_{t-1}) ao multiplicá-los pelo valor nulo. O segundo Gate determina a adição de novos valores ao estado oculto (h_{t-1}), apresentando duas camadas conforme ilustrado na figura 4, com a operação $(i_t \times c_t) + h_{t-1}$. A primeira camada com função de ativação sigmoide decide se os novos valores candidatos, gerados pela segunda camada com função de ativação tangente hiperbólica, serão incorporados ao novo estado oculto (h_{t-1}). Se a ativação sigmoide resultar novamente em zero, novos valores não serão adicionados a h_t , pois o produto será nulo e, conseqüentemente, não haverá alterações após a soma com h_{t-1} no final do processo.

$$o_t = \sigma(w_0 \times [h_{t-1}, x_t] + b_0) \quad (9)$$

$$h_t = o_t \times \tanh(ct) \quad (10)$$

Ao final do processo, a rede neural toma uma decisão com base na entrada x_t , considerando os resultados obtidos pelas Portas de Esquecimento e de Inserção. O sistema então dispõe de um estado oculto final h_{t-1} que é incorporado na decisão. A camada neural responsável pela decisão utiliza a função de ativação sigmoide (σ), a mesma empregada anteriormente nas estruturas das Gates. O sinal resultante é multiplicado pelo produto da função tangente hiperbólica aplicada ao valor h_{t-1} , conforme indicado na função 10, produzindo assim um valor de saída h_t que é influenciado pelas saídas anteriores e, por sua vez, influenciará a saída subsequente h_{t+1} .

⁴⁰ OLAH, C. Understanding LSTM Networks. **Github**, 2015. Disponível dia <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Acesso dia 27 agosto 2024, as 14:54.

Random Forest

Dentre as possibilidades de aplicação de algoritmos de aprendizado de máquina e redes neurais, a Random Forest é sempre considerada uma opção viável devido a sua capacidade de geração de diferentes níveis e árvores de decisão⁴¹. Em suma, consiste em um algoritmo que cria uma infinidade de árvores de decisão durante a fase de treinamento, utilizando diferentes partições da base de dados com o objetivo de reduzir o overfitting e melhorar a performance de predição⁴².

Para garantir que cada árvore de decisão no conjunto traga uma perspectiva única, Random Forest emprega a seleção aleatória de características⁴³. Essa aleatoriedade garante que cada árvore se concentre em diferentes aspectos dos dados, promovendo um conjunto diversificado de preditores dentro do conjunto.

Cada conjunto se apresenta de forma diferente, por conta da técnica de bagging, envolvendo a criação de múltiplas amostras bootstrap do conjunto de dados original, permitindo que as instâncias sejam amostradas com reposição⁴⁴. Isso resulta em diferentes subconjuntos de dados para cada árvore de decisão, introduzindo variabilidade no processo de treinamento e tornando o modelo mais robusto.

Após criado essa variabilidade de modelos de dados, o próprio algoritmo examina o problema para realizar a tarefa de previsão final em um mecanismo interno de tomada de decisão equilibrada e objetiva⁴⁵. Essa abordagem ajuda a evitar que o modelo se concentre demais nos dados de treinamento, tornando-o menos propenso ao overfitting⁴⁶.

Random Forest lida com grandes volumes de dados como um explorador experiente descobrindo quais características são mais importantes. Ao treinar cada

⁴¹ RAMIREZ, C. A. APLICACIÓN DEL MACHINE LEARNING EN AGRICULTURA DE PRECISIÓN. **Revista CINTEX**, Manizales, 11 novembro 2020. 14-27. P. 20.

⁴² WALSH, I.; FISHMAN, D.; AL, E. Recommendations for machine learning validation in biology. **Aristotle University of Thessaloniki**, 2020. Disponível dia https://www.researchgate.net/publication/342547853_Recommendations_for_machine_learning_validation_in_biology. Acessado dia 10 de julho de 2024, as 07:52.

⁴³ LOUPPE, G. Understanding Random Forests: From Theory to Practice. **ARXIV**, 2014. Disponível dia <https://arxiv.org/abs/1407.7502>. Acesso dia 15 maio 2024, as 17:47.

⁴⁴ BREIMAN, L. Random Forests. **Machine Learning**, 01 outubro 2001. 5-32. P. 7.

⁴⁵ BREIMAN, L, Ibidem.

⁴⁶ HO, T. K. Random decision forests. **Proceedings of 3rd International Conference on Document Analysis and Recognition**, Montreal, 14 agosto 1995. 278-282. P. 280

árvore de decisão, ele também reserva um grupo secreto de casos (out-of-bag) para testes. Essa validação integrada garante que o modelo não apenas se saia bem no treinamento, mas também tenha um bom desempenho em novos desafios.

Ela se adapta a situações com valores ausentes, fazendo previsões usando as informações disponíveis. Essa abordagem paralela aproveita o poder da tecnologia moderna, tornando todo o processo mais rápido e eficiente para lidar com projetos em larga escala.

MÉTODO

Formação de preços

Custos da empresa

Para que uma empresa possa funcionar corretamente, ela precisa arcar com diversos custos que envolvem desde a sua estrutura física até os seus recursos humanos. Esses custos podem ser classificados em fixos, variáveis e semivariáveis, dependendo da sua relação com o volume de produção ou de vendas da empresa.

O objeto de estudo deste trabalho são empresas de pequeno porte que se enquadram dentro do regime tributário do simples nacional. Nesse espectro de observação temos que considerar os seguintes pontos de formação de valores:

Simple nacional
IPTU
IPVA
Luz
Água e esgoto
Funcionários
Contador

Internet e telefone
Risco de perda
Taxa de débito
Taxa de crédito
Adubo

Como estudo, todas as referências que serão utilizadas serão das pequenas empresas, que seguem o regime do simples nacional. Muito embora a empresa (floricultura) que concordou em ajudar com a observação dos espécimes vegetais não tenha concordado com a divulgação de dados fiscais, é possível estabelecer padrões para realização dos cálculos necessários para este projeto de iniciação científica.

Com esses dados, podemos então realizar o cálculo, com base na legislação e no chamado custo real:

- Taxa Real = $(RBT12 * TAXA - PD) / RBT12$.

Para a estrutura física da loja, vamos considerar que a despesa é apenas de uma proprietária do ambiente, ou seja, não temos o aluguel para descontar, apenas IPTU dentro das regulamentações e tabela fixada pelo governo do Distrito Federal⁴⁷ e tendo como base a localização de um dos polos de comercialização de plantas no Distrito Federal.

Com relação a água, o preço será estimado através de metragem cúbica e o preço cobrado pela empresa prestadora do serviço de água e esgoto do Distrito Federal. O próximo passo seria o cálculo do consumo de energia elétrica que será utilizado também os cálculos presentes em conta de luz fornecida pela concessionária de energia local: NEO Energia. Sua composição do preço final é calculada conforme o gasto de Kw/h adicionado de um percentual de iluminação pública e do pagamento de três impostos: ICMS, PIS, COFINS.

⁴⁷ GOVERNO DO DISTRITO FEDERAL. LEI Nº 7.368, DE 26 DE DEZEMBRO DE 2023. Anexo I. **Diário Oficial do Distrito Federal**, 2023. Disponível dia https://www.dodf.df.gov.br/index/visualizar-arquivo/?pasta=2023%7C12_Dezembro%7CDODF%20241%2027-12-2023%20SUPLEMENTO%20B%7C&arquivo=DODF%20241%2027-12-2023%20SUPLEMENTO%20B.pdf. Acessado dia 12 de março de 2024, as 16:35

Lógico que há outras despesas fixas relativas ao espaço físico: compras de material de limpeza e materiais de manutenção (ferramentas, materiais elétricos, hidráulicos, entre outros). Contudo, por serem de uma variação definida pela metragem quadrada da loja e qualidade de serviço prestado, não seria possível adicionar aos cálculos deste trabalho. Finalizando a parte das despesas fixas relativas ao estabelecimento em si e a própria empresa (adotante do Simples Nacional) podemos partir para as despesas semivariáveis, que seria os funcionários. A mão de obra é uma despesa semivariável no quesito de quantidade.

Relativo a isto, segundo informações do próprio Sebrae⁴⁸, que por sua vez se baseia na LC 139/11 uma empresa de pequeno porte pode ter de 10 a 49 funcionários. E aqui calcularemos 11 funcionários: um fixo e ao menos 10 variáveis. O funcionário fixo seria o profissional responsável pela contabilidade e os demais seriam responsáveis pela gerência, atendimento ao público e serviços gerais.

Nem todos podem contar com a presença de um advogado constante na empresa, tampouco técnicos para os equipamentos eletrônicos (câmeras, computadores, roteadores, geladeira, freezer), mas obrigatoriamente um contador precisa ser constante e se insere nos salários administrativos a serem contabilizados para as despesas.

Os contadores possuem um conselho profissional para regulamentar tanto seu trabalho, quanto ajudar nas questões de piso salarial em dissídios coletivos que depois de firmados tem força de lei.

Outro custo a ser adicionado é justamente a questão da internet e telefone. Muito embora passe despercebido na maioria dos cálculos de custos de um produto ou empresa, hoje uma empresa não sobrevive sem ao menos um plano fixo (telefone fixo e internet) e um móvel para entrar em contato e fazer mesmo suas propagandas para sua empresa.

Nessa parte, será necessário considerar apenas os valores dos planos mais vendidos, segundo as próprias prestadoras de serviço telefônico. Considerando como

⁴⁸ SEBRAE. EPP: entenda o que é uma empresa de pequeno porte. **EMPREENDEDORISMO | EMPREENDEDOR,** 2021. Disponível dia <https://sebrae.com.br/sites/PortalSebrae/ufs/ac/artigos/epp-entenda-o-que-e-uma-empresa-de-pequeno-porte.305fd6ab067d9710VgnVCM100000d701210aRCRD>. Acesso dia 20 março 2024, as 13:30.

ponto de referência a própria empresa que concedeu os materiais de observação deste trabalho, temos duas empresas: uma prestadora de banda larga que seria a Tim e uma outra prestadora de serviços telefônicos, que seria a Claro.

Outro custo que demanda atenção é das máquinas de cartões de crédito e débito. Mesmo considerando a dificuldade para o cálculo desse gasto, não pode ser deixado de lado considerando que este é um dos principais meios de pagamento utilizado por todos.

Apesar de não ser comum o comentário ou a consideração do item que será agora mencionado, é de suma importância nos dias de hoje em que temos uma necessidade de realizar entregas. Uma empresa precisa considerar a existência de um meio de transporte para realizar entregas e outras utilidades essenciais para a empresa.

Considerando uma empresa de pequeno porte, não seria de bom uso os carros de passeio, da categoria hatch ou sedan. O mais comum que se encontram nas empresas do distrito federal são picapes médias e pequenas, já usadas ao menos com 5 anos de utilização. Para cotar o quanto seria o gasto necessário para incluir nas despesas fixas, é necessário considerar os cálculos do IPVA relativos ao preço do carro no Distrito Federal.

Formação de preços – Sunpatiens

A Sunpatiens, uma espécie de planta comercial conhecida por sua resistência e beleza, no entanto, determinar o preço ideal para essa planta pode ser um desafio. Através da observação do seu desenvolvimento e decaimento, foi possível determinar que do estado inicial de recebimento das plantas até seu decaimento total, a floração dura um período total de 15 dias.

E contando com o tempo de exposição e venda da planta, a floração é um fator determinante para que a aquisição do produto pelo consumidor final ocorra. Por isso, durante todo o período de observação (com e sem fertilizantes) foi possível determinar

que uma planta, após a sua perda total de flores só recupera sua floração após 8 semanas em condições ideais de água e exposição solar (sol pleno - considerando os meses de novembro de 2023 a fevereiro de 2024).

Também foi possível determinar que para a recuperação da floração por meio de adubação, é possível em um período de 30 dias, considerando a aplicação de adubo com o dobro da dose indicada pelo fabricante em duas doses divididas em 15 dias. Por conta desse tempo, o indicado é, para manter a floração da planta durante todos os 25 dias sugeridos para sua venda, iniciar a aplicação do adubo no final da primeira semana e a segunda aplicação ser feita no 21 dia (caso ainda não tenha sido vendida).

Isso induz uma alteração no tempo de venda em razão da alteração do custo diário da planta que será acionado posteriormente nos cálculos e tratado pelo código de aprendizado de máquina, baseado nos dados inseridos na planilha de observação.

Seletiva das plantas e manutenção

Respeitando o ciclo biológico de cada planta, entram dentro dos fatores de observação a umidade, a luz solar obrigatoriamente. Além disso foram considerados também a qualidade da terra e ainda a qualidade dos principais componentes estruturais de plantas com flores: a vitalidade do seu caule, suas folhas, as flores e seu tamanho.

As condições de solo, água, luz e temperatura influenciam no desenvolvimento e produtividade das plantas, muitas expostas à alta intensidade luminosa e, por mutações, adquirem mecanismos morfológicos e fisiológicos que possibilitam sua sobrevivência⁴⁹.

A observação das plantas foi conduzida pelo próprio pesquisador, Leão Pereira Neto, em local residencial, Brasília DF, lago norte entre os períodos de novembro de 2023 a fevereiro de 2024, com altitude de 1100m, o clima da região, segundo a

⁴⁹ DARDENGO, J. D. F. et al. ANÁLISE DA INFLUÊNCIA LUMINOSA NOS ASPECTOS ANATÔMICOS DE FOLHAS DE *Theobroma speciosum* WILLD EX SPRENG. (MALVACEAE). **Ciência Florestal**, Santa Maria, 27 julho 2017. 843-851. P. 847.

classificação de Koppen_geiger⁵⁰, situa-se em uma faixa de transição entre o subtipo Cfa mesotérmico úmido sem estiagem ou pequena estiagem e o subtipo Aw tropical úmido, com estação chuvosa e quente e de precipitação pluviométrica média do ano de 1430mm⁵¹.

Com observação dentro de empresas de pequeno porte de comercialização de plantas, foi possível constatar que a forma comum de aquisição se dá por meio de caixarias para as menores plantas e pote para as de tamanho médio e grande.

Para a quantidade de água fora estabelecido uma quantidade igualitária de 40 ml por saco de planta durante os meses de menor umidade do ar e aproximadamente 27ml para os meses de maior umidade do ar. A medida foi feita através de um copo de medida e a rega das plantas feita de forma uniforme para maior controle. Contudo devo ressaltar que a dinâmica de uma loja de plantas não permite esse tipo de controle, sendo possível somente partir de uma medida geral de 600 ml para cada caixaria para os dias menos úmidos e de 400 ml de água para os dias mais úmidos.

Para a quantidade de exposição ao sol, as plantas foram expostas de maneira uniforme durante os 60 dias. Duas caixas receberam luz plena (durante todo o dia) e outras duas caixas receberam metade do dia de luz, outra metade de luz indireta (somente claridade).

De acordo com Boeger e Wisniewski⁵², as folhas de sol apresentam características xeromórficas⁵³, oriundas de uma maior intensidade luminosa, menor disponibilidade de água e nutrientes no solo ou mesmo pela baixa temperatura⁵⁴. As

⁵⁰ UNIVERSITY OF MELBOURNE. Classificação climática de Köppen- Geiger. **Academia.EDU**, 2024. Disponível dia https://www.academia.edu/36723493/Classifica%C3%A7%C3%A3o_clim%C3%A1tica_de_K%C3%B6ppen_Geiger_Classifica%C3%A7%C3%A3o_clim%C3%A1tica_de_K%C3%B6ppen_Geiger?auto=download. Acessado dia 15 de julho de 2024, as 09:42

⁵¹ INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. Centro de Previsão de Tempo e Estudos Climáticos. **CPTEC, Arquivos do Laboratório de Meteorologia e Climatologia, vinculados ao CPTEC-INPE**, 2024. Disponível dia <https://www.cptec.inpe.br/>. Acessado dia 1 julho 2024, as 8:24.

⁵² BOEGER, M. R.; WISNIEWSKI, C. Comparação da morfologia foliar de espécies arbóreas de três estádios sucessionais distintos de floresta ombrófila densa (Floresta Atlântica) no Sul do Brasil. **Revista Brasil Botanica**, Curitiba, 26 março 2003. 61-72. P. 65

⁵³ BRANT, R. S. et al. Adaptações fisiológicas e anatômicas de *Melissa officinalis* L. (Lamiaceae) cultivadas sob malhas termo refletoras em diferentes intensidades luminosas. **Rev. Bras. Pl. Med.**, Botucatu, 22 abril 2011. 467-474. P. 470.

⁵⁴ GOMES, R. et al. Amazon chicory: growing at full sunlight or under shade? **Horticultura Brasileira**, Capanema, 1 abril 2023. 1-6. . p. 2

características observadas nessas folhas são a presença de abundante cutícula, paredes celulares mais espessas, endoderme mais espessa, além de se verificar folhas pequenas e compactas, parênquima paliçádico mais desenvolvido que o parênquima lacunoso, grande frequência de estômatos, e algumas vezes, células menores e abundante esclerênquima. Geralmente, estas características não são observadas em folhas de sombra por receberem uma menor intensidade luminosa⁵⁵.

Com relação a qualidade da terra, por falta de equipamentos não foi realizada nenhuma medição de nutrientes ou composição mais detalhada. Partiu-se então da composição padrão comercializada que seria uma mistura entre terra vermelha (também chamada de terra vegetal, que nada mais é do que a terra comum encontrada na região, sem adição de nenhum composto químico ou orgânico para seu enriquecimento de nutrientes) e compostos orgânicos vegetais (adubo vegetal que inclui materiais de decomposição como cascas de árvores, cascas de legumes, a exemplo de palha de arroz, que são curtidos por um tempo junto com a terra vegetal que a torna negra, e comercialmente adquire o nome de substrato).

A partir dessa composição inicial a qualidade da terra foi medida da seguinte forma: úmida ou seca. Esse tipo de qualificação de terra colabora para saber o estado da planta se está recebendo água em excesso ou carente de água. Fora essa medida, muito embora não tenha sido feita nenhuma medição, houve uma nova divisão com relação às caixas. As duas caixas que receberam sol pleno receberam uma dose adicional de adubo NPK, de subdivisão proporcional, para favorecer a floração.

Por fim, é necessário deixar claro que as caixas foram expostas a condições diferentes e cada uma levou a uma pontuação final que foram levadas ao tempo de venda. A variação da pontuação está entre 0-22 e cada planta recebeu um número que varia de 1 a 60, sendo que cada uma das caixas continha 15 plantas. Sendo assim, as plantas 1 a 15 pertenciam a caixa 1; plantas 16-30 pertenciam a caixa 2; plantas 31-45 estavam contidas da caixa 3 e as plantas 45-60 permaneceram contidas na caixa 4.

⁵⁵ CUTTER, E. G. **Plant Anatomy**. 2. ed. Londres: Edward Arnold L:td., 1978. P. 136.

Observação e divisão das espécies

Todas as plantas foram recebidas com flores, e a transformação do estado inicial foi objeto de análise. O tempo de venda das plantas com flor foi de aproximadamente duas semanas para cada caixa, enquanto o tempo de venda das plantas sem flor variou de quatro a cinco semanas.

Respeitando o ciclo biológico de cada planta, os fatores de água, iluminação e solo foram cuidadosamente ajustados para cada caixa, resultando em resultados diferenciados. A qualidade da terra foi uma variável chave, com duas caixas recebendo adubação à base de NPK para enriquecer o solo. Além disso, a vitalidade dos componentes estruturais das plantas, como o caule, folhas, flores e tamanho, foram monitorados para garantir o desenvolvimento saudável e a atratividade comercial das Sunpatiens.

Durante o período estabelecido, quatro caixas foram observadas, cada uma contendo 15 plantas. As condições de cultivo e cuidados foram meticulosamente registradas para cada caixa, gerando resultados variados:

- Caixa 1: Contém exclusivamente plantas de grande porte, com 18 cm de altura. Essas plantas receberam 600 ml de água e 6 horas de sol diariamente por dois meses. O tempo médio de cuidado foi de 5-10 minutos por dia. Ao final dos 60 dias de observação, a Caixa 1 teve o menor tempo de vida, com uma degradação exacerbada nos últimos 15 dias devido à falta de sol e água, resultando em perda total.
- Caixas 2 e 3: Ambas contêm plantas de médio porte, com 13 cm de altura, e seguiram o mesmo regime de rega e cuidados da Caixa 1, mas com 6 horas de sol diariamente. A Caixa 2 não teve floração devido à falta de luz solar, mas as plantas permaneceram vivas, com degradação de algumas folhas e perda de 15% das plantas. A Caixa 3, recebendo luz plena e hidratação adequada, teve perda total da floração inicial em 7 dias, mas recuperou parte da floração após 15 dias e novamente após mais 15 dias com a aplicação de adubo.
- Caixa 4: Composta por plantas de pequeno porte, com 10 cm de altura, recebeu 12 horas de sol por dia e foi a caixa que melhor se desenvolveu, com a melhor floração.

Durante o estudo, foi observado um padrão de floração das plantas, com flores surgindo a cada 15 dias, variando conforme a exposição à luz do sol e a suplementação de nutrientes no solo. Todas as plantas foram recebidas com flores, e o estado inicial foi mantido durante a observação. O tempo de venda das plantas com flor foi de aproximadamente duas semanas para cada caixa, enquanto o tempo de venda das plantas sem flor variou de três a quatro semanas.

A qualidade da terra foi uma variável chave, com duas caixas recebendo adubação à base de NPK. A vitalidade dos componentes estruturais das plantas, como o caule, folhas, flores e tamanho, foram monitorados para garantir o desenvolvimento saudável e a atratividade comercial das Sunpatiens.

Embora cada caixa tenha seu desenvolvimento individualizado, o único padrão detectado foi o período de venda, que ocorre concomitantemente com a presença de flores; sendo imprescindível para o sucesso comercial, já que as unidades sem flores não obtiveram sucesso na venda, permanecendo na loja e gerando mais custos.

Todos os dados relativos ao tempo de observação, crescimento ou decaimento da planta, água, exposição ao sol, aplicação de adubo ou não foram inserido em planilha de dados, dentro do programa Microsoft Excel, em formato de arquivo XLSX

Codificação

Desenvolver um sistema de predição de valores para comercialização de plantas baseado no seu tempo de vida vai além da noção de custos de operação de uma empresa. Como deve-se especular fatores, por exemplo o fato da aparência influencia muito na venda, principalmente para plantas ornamentais. A composição do visual da espécie de estudo (qualidade visual do vigor da planta, qualidade da folha e da flor) é um fator variável que será utilizado para o estudo.

A combinação de técnicas a serem utilizadas para implementação da ferramenta serão as redes neurais profundas, *Deep Learning*. Os passos objetivos para se alcançar esse objetivo primordial são:

- Primeira etapa - Coleta de dados - escolha dos objetos e dados que serão utilizados para a formação da informação requerido por um cliente ou pelo próprio cientista de dados.
- Segunda etapa - Base de dados - definir as relações, tratamentos e parâmetros para que as metas de acurácia para o sistema de predição de valores por Inteligência Artificial funcionem adequadamente;
- Terceira etapa - Identificar os processos relacionados a venda das plantas para garantir que o resultado da inteligência artificial seja o mais semelhante possível com o ocorrido durante a captura dos dados;
- Quarta etapa - Se possível e necessário, incrementar o conjunto de dados de observação das plantas para utilização no sistema de inferência rede neurais profundas;
- Quinta etapa - Modelar uma rede neural profunda a fim de realizar o aprendizado do padrão reconhecido dos momentos de venda e assim poder gerar novas predições para as novas espécimes que entrem na loja;
- Sexta etapa - Aplicar Testes, a fim de alcançar resultados tanto quantitativos como também qualitativos em relação a qualidade da ferramenta para a venda das plantas.

Cada programador, possui total liberdade para desenvolver seu código, utilizando-se de qualquer estrutura sem pré-definição e um único objetivo de atender uma necessidade pré-determinada. Após entender o problema e identificar o objetivo do projeto que foi identificado a necessidade de uma construção de código que perpassa por todos os pontos esclarecidos anteriormente: sendo o cálculo de custos empresariais, armazenagem dos dados de observação, tratamento e por fim aplicação de algoritmos de aprendizado de máquina

O processo de obtenção dos dados não passou por um caminho de automação, sendo os mesmos coletados individualmente para cada espécime de planta e aos poucos foram sendo armazenados dentro de um único arquivo .xlsx, salvo no formato padrão do programa Excel. Foi um processo que não requisitou autorização de acesso e tampouco conversão, sendo armazenados em dois locais distintos para minimizar o risco de perda do arquivo, devidamente protegidos por senha. Finalizando assim a primeira etapa

Com os dados coletados, o código em si foi desenvolvido tratando primeiramente a parte de cálculo dos custos da empresa baseado no método descrito anteriormente, através de funções nativas de cálculo e interação com o usuário. Dando

início a segunda etapa, passou-se para a importação de bibliotecas necessárias para tratamento da base de dados e desenvolvimento das redes neurais.

Cada uma dessas bibliotecas utilizadas apresentou uma importância para o desenvolvimento do código, oferecendo ferramentas para manipulação, limpeza e análise de dados⁵⁶. Suas utilizações mais especificamente foram destinadas a formatação da planilha para corrigir alguns erros presentes nos dados, preenchimento e substituição de valores nulos através da seleção de subconjunto de dados (presentes em cada coluna desejada).

Módulos pertencentes a esta biblioteca foram utilizados para determinar os valores presentes em cada coluna e assim ser possível determinar a substituição por valores numéricos que normalmente são melhores tratados por códigos de aprendizado de máquina e para tratar os dados temporais em uma determinada faixa de tempo selecionada que foram úteis para prosseguir com a aplicação das partes seguintes do código.

Completando o conjunto de tratamento, na terceira etapa passou-se a construção de gráficos de correlação entre as diversas variáveis e assim determinar ou confirmar a classe alvo, por resultado das correlações e os dados de observação. Facilitando assim o indicativo de que modelos de aprendizado de máquina poderiam ser utilizados.

Corrigindo todos os dados faltantes, na quarta etapa partiu da implementação da base de dados com novos pontos de observação baseado nos valores de venda, dias da semana e qualidade da planta. Assim seria possível realizar uma implementação na próxima etapa de previsões de venda e dias de promoção. Finalizando essa parte de tratamento dos dados o próximo passo foi a arquitetura do modelos de aprendizado de máquina

No desenvolvimento de modelos de Machine Learning, a seleção e importação de bibliotecas adequadas são fundamentais para o sucesso do projeto. Com as bibliotecas corretas podemos além de fazer a divisão dos dados para teste e

⁵⁶ PYDATA. PANDAS - Python Data Analysis Library. **Pandas**, 2024. Disponível dia <https://pandas.pydata.org/about/>. Acesso dia 15 julho 2024, as 19:34.

treinamento, podemos aplicar os hiperparâmetros necessários para evitar overfitting ou underfitting e aprimorar a capacidade de predição de valores.

Criar um conjunto de testes é simples e ao mesmo tempo importante. Com a biblioteca adequada, que está no código utilizado para esta pesquisa, e devidamente disponível em plataforma pública.⁵⁷ Foi realizada uma separação entre os dados de treino e de teste numa proporção de 20% para este e 80% para o treino. O pré-processamento dos dados por um processo de normalização dos dados, frequentemente realizado com o MinMaxScaler, que também é chamado de normalização, onde os valores são ajustados e redimensionados para ficar dentro do intervalo de 0 a 1; subtraindo o valor mínimo e dividindo pela diferença entre o valor máximo e mínimo:

$$X_std = (X - X.min(axis=0)) / (X.max(axis=0) - X.min(axis=0))$$
$$X_scaled = X_std * (max - min) + min$$

Como boa parte das características únicas detinham normalmente 4 pontos ou variações. Isso passou a ser essencial para o desenvolvimento do código. No que diz respeito à construção de modelos de Deep Learning, a biblioteca keras se destaca por sua simplicidade e eficiência. Principalmente a escolha da API Sequencial que permite criar redes neurais de forma fácil, criando camada por camada.

Com Sequential, a metodologia de criação foi criar camadas de redes com 50 neurônios. A primeira camada para capturar as dependências temporais, uma segunda sem neurônios para evitar *overfitting*, uma terceira para extrair os padrões mais complexos de entrada, com outros 50 neurônios, uma terceira camada para uma regularização adicional e por fim uma última camada de conexão com a primeira unidade para gerar a previsão final do modelo.

O modelo foi compilado utilizando o otimizador Adam devido à sua eficiência computacional e uso reduzido de memória, tornando mais facilmente aplicável outros computadores, além de utilizar as médias móveis para obter uma estimativa mais

⁵⁷ Biblioteca publica destinada a armezanar o codigo implementado nesta pesquisa - <https://github.com/leaop/Graduation/tree/main/5%20semestre/PIC>

adequada⁵⁸. Métricas como *MeanSquaredError* também foram aplicadas para avaliar o desempenho do modelo durante o treinamento.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Essa métrica - média de erro quadrático (MSE) - com fórmula matemática acima é usada para perda de regressão calculando a medida da diferença quadrada entre as observações reais e previstas, considerando principalmente a magnitude média do erro.

Na sexta e última etapa, ficou mais fácil identificar a importância de cada uma dessas bibliotecas na pipeline de Machine Learning, desde o pré-processamento e análise de dados até a construção e avaliação de modelos. A escolha dessas bibliotecas é guiada pela necessidade de realizar tarefas específicas de forma eficiente e eficaz no contexto do problema de Machine Learning em questão e aplicar os testes ao final para aprimoramento. A integração dessas ferramentas contribui para a criação de um fluxo de trabalho robusto e confiável, essencial para o desenvolvimento de soluções baseadas em dados a partir de 3 modelos de aprendizado de máquina, cada um realizado testes através de métricas que serão apresentadas em momento posterior.

RESULTADOS E DISCUSSÃO

Com relação ao resultado, temos que apresentar cálculos e observações de três diferentes áreas que resultam na formação dos tópicos anteriores da metodologia de trabalho: formação dos custos da empresa; custos de manutenção de cada espécime

⁵⁸ INSTITUTO EUROPEU DE CERTIFICAÇÃO DE TECNOLOGIAS DE INFORMAÇÃO - EITCI ASBL. Como o otimizador Adam otimiza o modelo de rede neural? **EITCA**, 2023. Disponível em: <https://pt.eitca.org/intelig%C3%A7%C3%A3o-artificial/eitca-ai-dl-f-aprendizado-profundo-com-tensorflow/fluxo-tensor/modelo-de-rede-neural/revis%C3%A3o-de-exame-modelo-de-rede-neural/como-o-otimizador-adam-otimiza-o-modelo-de-rede-neural/>. Acesso em: 4 julho 2024.

de Sunpatiens para sugestão do preço final de venda e resultados obtidos pela aplicação do código de aprendizado de máquina.

Dos cálculos empresariais

Para facilitar os cálculos, foi considerado que a empresa modelo (ficta) possui um lucro bruto anual de R\$ 1.200.000,00 (um milhão e duzentos mil reais), que equivaleria a uma renda bruta de R\$ 100.000,00 (cem mil reais) mensal. Dentro desse espectro e com base nos ditames legais, esta empresa se situaria na 4ª faixa, com uma taxa de 10,70% e uma dedução de R\$ 22.500,00, conforme a seguinte metodologia de resolução:

- Taxa Real = $(RBT12 * TAXA - PD) / RBT12$
- $(R\$ 1.200.000,00 \times 10,70\%) - R\$ 22.500,00 / 1.200.000,00$
- $(128.400,00 - 22.500,00) / 1.200.000,00$
- $105.900,00 / 1.200.000,00 = 0,0882$ ou 8,22%
- Assim, para esse tipo de empresa, a taxa real é de 8,22%

Assim, para esse tipo de empresa, a taxa real é de 8,22%. Nesse caso, o ICMS representa 33,50% da distribuição dos tributos. Então, calculamos: 33,50% de 8,22 = 2,7587%, que é a alíquota correspondente ao ICMS.

O próximo valor a ser calculado sob a mesma metodologia de tributação federal foi o imposto sobre o local da empresa: temos uma média de valores do IPTU para cálculo dessa despesa do local físico: ao total temos um valor aproximado de R\$ 18.000,00 (dezoito mil reais) para essa despesa, o que equivaleria a 18% do lucro bruto, destinado para a manutenção do espaço físico⁵⁹. Ao particionar durante o ano-calendário, temos um equivalente a R\$ 1.500,00 (um mil e quinhentos reais).

Com relação aos custos com a água, não foi possível obter diretamente o valor padrão ou média de valores pagos pela empresa. Por cada conta de água também contar dados sensíveis, protegidos pela LGPD, foi realizado pesquisa de valores com base em estimativas fornecidas pela empresa. Um **caminhão-pipa** no **Distrito Federal** pode variar dependendo de vários fatores, como a capacidade do caminhão e a

⁵⁹ Ao contrário do simples nacional que se torna um gasto mensal, esse valor do IPTU é anual.

empresa que oferece o serviço, a título de exemplo temos aqui uma estimativa de valores:

Uma empresa localizada em Taguatinga, Brasília, oferece serviços de caminhão-pipa. Eles têm caminhões com capacidades variadas e preços diferentes⁶⁰. Em contação feita com a empresa, por exemplo, um caminhão-pipa com 10.000 litros de capacidade está à venda por R\$ 75.000⁶¹.

Para um cálculo mais fidedigno serão utilizadas as tarifas publicadas pela empresa prestadora de serviço de água e esgoto do Distrito Federal. Considerando que uma empresa de pequeno porte, utiliza-se ao menos de três caminhões pipa de 60.000 L (sessenta mil litros), podemos considerar que, pela CAESB, estamos diante da tarifa 5 de valores. Pelo cálculo fornecido temos que multiplicar o consumo dentro da faixa 5 pela alíquota e somar a os resultados obtidos à tarifa fixa, obtendo assim o valor do serviço de abastecimento de água. Então, teremos um gasto mensal equivalente a R\$1635,00 (um mil seiscentos e trinta e cinco reais). Então temos que o gasto com água equivale a 1,635% aproximadamente da despesa mensal fixa da empresa.

Finalizado essa parte, o próximo dado a ser calculado ainda com estimativas foram os custos com despesas de energia elétrica dentro dos moldes tarifários da empresa local.

A tarifa de energia corresponde ao valor cobrado por unidade consumidora (medido em R\$/kWh) e representa o preço da energia elétrica consumida. A definição dessas tarifas⁶² é de responsabilidade da ANEEL (Agência Nacional de Energia Elétrica)⁶³, que atua em conformidade com as leis e os contratos de concessão estabelecidos com as empresas distribuidoras em diferentes regiões do país.

⁶⁰ [Sobre | Braságuas Caminhão pipa \(brasagua.srv.br\)](#), acessado dia 12 de março de 2024, as 16:40

⁶¹ Essas são somente estimativas de preços e podem variar. É sempre recomendável entrar em contato diretamente com as empresas para obter informações atualizadas e detalhadas sobre os preços dos caminhões-pipa.

⁶² Há uma distinção entre preço e tarifa. A tarifa corresponde ao valor cobrado pela prestação dos serviços da distribuidora de energia, enquanto o preço é a combinação da tarifa com os tributos. O valor final da conta é obtido somando a tarifa aos tributos.

⁶³ ANEEL. Custo da energia que chega aos consumidores. **Agência Nacional de Energia Elétrica**, 2022. Disponível dia
<https://www.gov.br/aneel/pt-br/assuntos/tarifas/entenda-a-tarifa/custo-da-energia-que-chega-aos-consumidores>. Acesso dia 15 fevereiro 2024, as 23:48.

Segundo a própria operadora de energia elétrica⁶⁴ a composição dos custos repassado para a empresa são compostos por: custo de geração de energia, da sua transmissão, encargos setoriais, de distribuição e por fim os tributos.

Nesse molde de custos tarifários e juntamente com o consumo, a empresa apresenta um consumo consistente de R\$680,00 (seiscentos e oitenta reais), com poucas variações para mais ou menos. O que equivaleria a um total de 33,79% do seu valor equivalente a taxas e impostos. O total aproximado consumido do kw/h foi de 464 Kw/h.

Sendo assim, temos nos cálculos um total de R\$ 450,21 reais de preço efetivo pago pelo Kw/h, 12,15% de contribuição de iluminação pública, 20% de ICMS 0,29% de PIS e 1,35 de COFINS. Mas para fins de consideração da despesa mensal mínima, temos que a despesa com luz equivale a 0,68% das despesas mensais aplicadas ao lucro bruto da empresa.

A partir do valor bruto e do gasto de kw/h, podemos estimar o gasto diário da conta de luz, dividindo o gasto pelos 30 dias corridos do mês, chegando a aproximadamente um valor de R\$ 0,97 (noventa e sete centavos) por hora, aproximadamente R\$ 0,02 (dois centavos) por minuto. Considerando que cada caixa leva uma média mensal de 3,23 minutos, o total gasto por caixa é equivalente a R\$ 0,07 (sete centavos).

No regime do Simples Nacional, amplamente adotado por micro e pequenas empresas, o custo de um funcionário é calculado levando em consideração alíquotas mais suaves e isenções de determinados encargos. Por exemplo, para um funcionário com salário-base de R\$ 1.500,00 (um mil e quinhentos reais), o cálculo seria o seguinte:

Salário	R\$ 1.500,00
8% de FGTS	R\$ 120,00
Provisão mensal	R\$ 210,00
Vale transporte	R\$ 220,00

⁶⁴ NEO ENERGIA. Composição Tarifária. **Entenda sua tarifa**, 2024. Disponível dia <https://www.neoenergia.com/web/brasil/sua-casa/composicao-tarifaria>. Acesso dia 15 fevereiro 2024, 23:55.

Total	R\$ 2.050,00
Desconto do INSS e vale transporte	R\$ 235,00
Custo efetivo mensal	R\$ 1.815,00

Sendo assim temos o seguinte: cada um dos 10 funcionários custa R\$ 1.815,00 (um mil oitocentos e quinze reais). O que no final seria um equivalente a 18,15% por mês do faturamento bruto.

Considerando o resultado de quanto um funcionário gasta custa para a empresa do simples nacional, podemos considerar agora os cálculos proporcionais, considerando uma jornada de trabalho de 44 horas semanais.

Para determinar o custo diário de um funcionário, é necessário primeiro calcular seu salário por hora. Dado que a jornada semanal é de 44 horas distribuídas ao longo de seis dias consecutivos, totalizando 176 horas por mês, podemos dividir o salário mensal pelo número total de horas trabalhadas para obter o salário por hora. Supondo que o funcionário trabalhe todos os dias do mês sem folgas, o custo diário pode ser calculado multiplicando-se o salário por hora pelo número de horas em um dia. Assim, considerando um salário mensal de R\$ 1.815 (um mil oitocentos e quinze reais) e uma jornada de trabalho de 24 horas por dia, o custo diário desse funcionário seria aproximadamente R\$ 247,44 reais.

Para calcular o custo da hora trabalhada, podemos usar a mesma fórmula usada anteriormente para calcular o salário por hora. Neste caso, como a jornada diária é de 8 horas, podemos calcular o salário por hora dividindo o salário mensal pelo número de horas trabalhadas por mês.

- Salário por hora = Salário mensal / (jornada diária * dias úteis por mês)
- Salário por hora = 1815 reais / (8 horas/dia * 22 dias/mês) ≈ 10,34 reais/hora

Portanto, o custo por hora trabalhada seria aproximadamente R\$ 10,34 (dez reais e trinta e quatro centavos). Sendo assim, temos o custo por minuto trabalhado chega a R\$ 0,1723 centavos (dezessete centavos). Esse cálculo de minutos é necessário pois o cuidado com as caixas das Sunpatiens leva cerca de alguns minutos e não horas.

Ainda tratando de despesas com funcionários, o próximo funcionário, ainda que não diretamente envolvido com a manutenção e venda das plantas, é o contador da empresa. Segundo pesquisas feitas⁶⁵, o piso salarial de um contador na cidade de Brasília é de R\$ 4.559,75 (quatro mil quinhentos e cinquenta e nove reais e setenta e cinco centavos), contando todas as despesas com encargos sociais, assim como o custo de funcionários da empresa.

Ao levar em consideração o faturamento bruto, temos a despesa fixa com o contador equivalente a 4,56% do bruto mensal. Isso sendo fracionado para dias, horas e minutos, este custo administrativo em si, equivale a uma despesa de R\$ 28,50 (vinte e oito reais e cinquenta centavos) por hora trabalhada, saindo o minuto trabalhado por R\$ 0,48 (quarenta e oito centavos).

Resta ainda o cálculo de duas despesas diretamente ligadas à comunicação, marketing e venda: telefone, internet e taxas das operadoras de cartão de crédito. O pacote da TIM para serviços de banda larga está custando um valor de R\$ 138,94 (cento e trinta e oito reais e noventa e quatro centavos) e para as linhas móveis a operadora Claro com o plano mais vendido está o equivalente a R\$ 64,99 (sessenta e quatro reais e noventa e nove centavos). Essas despesas juntas ao mês equivalem a um total de R\$ 203,93 (duzentos e três reais e noventa e três centavos), o que daria um custo diário de R\$ 6,80 (seis reais e oitenta centavos).

Comparar as taxas das máquinas de cartão pode ser uma tarefa desafiadora. As marcas frequentemente oferecem promoções que podem obscurecer o entendimento de suas tarifas padrão. Uma abordagem útil é simular uma transação de R\$ 100 para determinar o montante real que será depositado em sua conta.

Ao decidir qual máquina de cartão é mais adequada para seu MEI, leve em consideração quatro tipos de taxas:

- Taxa de débito.
- Taxa de crédito.
- Taxa de crédito parcelado.
- Taxa de antecipação, caso opte por receber os valores antes do prazo.

⁶⁵ DISSÍDIOS SALARIAIS. Salário de Contador - Reajuste, dissídio e piso salarial 2024. **Dissídio**, 2024. Disponível dia <https://dissidio.com.br/salario/contador/>. Acessado dia 20 de abril de 2024, as 16:56.

Marcas com políticas de tarifação mais transparentes tendem a ser mais fáceis de comparar durante a compra e de administrar no cotidiano. Tomemos como exemplo as taxas oferecidas pela Maquininha de Cartão MEI Fácil:

- Taxa de débito: 0,99% (a mais vantajosa no mercado para MEIs).
- Taxa de crédito antecipado (depósito em até 1 dia útil): 4,29%.
- Taxa de crédito parcelado antecipado (depósito em até 1 dia útil): 4,29% + 1,60% por parcela.
- Taxa de crédito (depósito em 30 dias): 3,11%.
- Taxa de crédito parcelado (depósito em 30 dias): 3,90% por parcela.

Pesquisar amplamente é crucial para garantir a escolha da máquina com as menores taxas para o CNPJ MEI⁶⁶. E essas percentagens têm que ser adicionadas ao final do preço somada a outro fator que seria o imposto de comercialização.

Finalizando os custos da empresa, o último valor que entra para determinação do custo diário da empresa com as plantas é o pagamento sobre o veículo de transporte da empresa. A alíquota do IPVA para automóveis é de 3,50% do valor de mercado disponibilizado na tabela FIPE. Considerando um exemplo, uma Saveiro Trendline 1.6 T.Flex 8V de 2017, pela tabela está custando no mercado R\$ 57.236,00 (cinquenta e sete mil duzentos e trinta e seis reais). Sendo assim, o IPVA seria de R\$ 2.005,66 (dois mil e cinco reais e sessenta e seis centavos). O que representa um gasto fixo de R\$ 167,14 (cento e sessenta e sete reais e quatorze centavos) por mês, equivalente a 0,17% do gasto mensal.

Sendo assim, totalizando tudo, temos:

Simple nacional	8,22%
IPTU	1,5%
IPVA	0,17%
Luz	0,68%
Água e esgoto	1,635%

⁶⁶ NEON. Máquina de cartão para MEI: o que considerar na hora de ter a sua! **Mei Fácil por Neon**, 2024. Disponível dia <https://blog.meifacil.com/sou-microempreendedor-individual/maquina-de-cartao-para-mei/#:~:text=Para%20te%20ajudar%20nesse%20exemplo%2C%20vamos%20falar%20sobre,%28recebe%20em%2030%20dias%29%20%E2%80%93%203%2C90%25%20por%20parcela>. Acessado dia 12 de maio de 2024, as 09:10

10 Funcionários	18,15%
Contador	4,56%
Internet e telefone	0,20%
Risco de perda	10%
Taxa de débito	0,99%
Taxa de crédito	3,11% + 1,60% por parcelamento
Adubo	0,00168%

Isso ao todo equivale a um total de 56,553% de gastos fixos. Se você tem gastos fixos que correspondem a 56,55% da sua renda bruta ao longo de um mês, isso significa que essa porcentagem é distribuída ao longo dos dias do mês. Assumindo um mês com 30 dias para simplificar, seus gastos diários representam aproximadamente 1,8851% da sua renda bruta. Essa porcentagem representa o quanto você gasta a cada dia em relação à sua renda total.

A composição do preço de venda deve incluir necessariamente esses gastos e a lógica de futuros cálculos devem também seguir essa porcentagem. Uma caixa de Sunpatiens chega a um preço de R\$ 25 (vinte e cinco reais), contendo 15 unidades. O total de despesas deve ser passado para a caixa de mudas e seus custos diários calculados. Cada muda tem um custo inicial de R\$2,61 (dois reais e sessenta e um centavos) - seria os (preço de custo + custos mensais) / quantidade de unidades. E são vendidas a um preço de R\$4,00 (quatro reais). Considerando o custo diário dessa planta de R\$0,037 reais.

Para determinar em quantos dias a planta pode ser vendida a um preço de R\$4,00 para obter algum lucro, precisamos calcular o lucro obtido por dia e, em seguida, encontrar o número de dias necessários para que o lucro total seja superior a zero.

Vamos começar calculando o lucro por dia:

- Preço de venda da planta: R\$ 4,00
- Custo inicial da planta: R\$ 2,61
- Custo diário por 1,8851% do custo inicial

Portanto, o lucro por dia é:

- Lucro por dia = Preço de venda - Custo inicial - Custo diário
- Lucro por dia = R\$ 4,00 - R\$ 2,61 - (1,8851% * R\$ 2,38)

Vamos calcular o custo diário primeiro:

- Custo diário = 1,8851% de R\$ 2,61
- Custo diário \approx 0,0492 reais

Agora, podemos calcular o lucro por dia:

- Lucro por dia \approx R\$ 4,00 - R\$ 2,61 - R\$ 0,0492
- Lucro por dia \approx R\$ 1,582454726

Portanto, considerando o valor final de venda, descontado os gastos com o cartão de crédito ou de débito, levaria aproximadamente 25 dias para que o valor acumulado chegasse a R\$3,82 (três reais e oitenta e dois centavos). Sem considerar o desenvolvimento ou decaimento da planta, que mudará o preço a seguir por conta da inserção do adubo e do decaimento das flores que ocorre em exatamente 15 dias (perda total de todas as flores).

Preço de venda

Um pote de adubo chega a um preço de R\$ 10 (dez reais) contendo 400g. A indicação de uso para o caso de preparação junto com a água é de 15 gramas por litro de água. Considerando que foram aplicadas em duas caixas, o total utilizado foi 33 gramas por aplicação. Para a planta responder ao estímulo, foram necessárias 2 aplicações. Totalizando um total de 66 gramas do produto.

Se considerar a porção ideal de para cada planta de 40ml de água, o produto terá sua perda, pois 400ml de água não serão utilizados. Esse valor de perda deve então ser passado para as considerações de cálculo de gastos. Considerando que cada caixaria de plantas contém um total de 15 plantas, somando o valor da água destinada (R\$ 0,02), temos um gasto de R\$ 0,07 (sete centavos) – composto basicamente pelos dois centavos da água, mais cinco centavos do adubo.

Para calcular a porcentagem de gasto que você tem com este funcionário por minuto trabalhado, precisamos primeiro determinar quantos minutos ele trabalha por

mês. Ele trabalha 8 horas por dia, então em 6 dias por semana, ele trabalha $6 * 8 = 48$ horas por semana. Como há 4 semanas em um mês em média, o número total de horas trabalhadas por mês seria 192 horas.

Como cada hora tem 60 minutos, o número total de minutos trabalhados por mês seria 11.520 minutos. Agora, para calcular a porcentagem de gasto por minuto trabalhado, dividimos o salário mensal do funcionário pelo número total de minutos trabalhados por mês e então multiplicamos por 100 para obter a porcentagem

- Porcentagem de gasto por minuto = $(\text{Salário mensal} / \text{Número total de minutos trabalhados por mês}) * 100$
- Porcentagem de gasto por minuto = $(1815 / 11520) * 100 \approx 15,75\%$

Portanto, você teria aproximadamente 15,75% de gasto com este funcionário por minuto trabalhado. Sendo assim, temos o custo por minuto trabalhado chega a R\$ 0,15 (quinze centavos). Esse cálculo de minutos é necessário pois o cuidado com as caixas das Sunpatiens leva cerca de alguns minutos e não horas. Durante os dias de aplicação do adubo o tempo do funcionário para cuidar das plantas é alterado de 3 minutos em média para 10 minutos.

Então para o sétimo dia e o vigésimo primeiro dia eu tenho um custo adicional diário passa dos R\$ 0,0492 (quatro centavos) para R\$ 0,12 (doze centavos) contando com o adubo e R\$1,05 (um real e cinco centavos) para o funcionário, totalizando nos dias de aplicação um total de R\$1,17 reais. Mas o cálculo não termina por aqui, pois o que ocorre é que cada funcionário no dia de aplicação gasta um total de 10 minutos para preparar a dose para a caixa com 15 unidades, nos dias normais, o tempo de cuidado seria de aproximadamente 2 minutos e meio. Sendo assim, ainda que consideremos um dispêndio de 2 minutos arredondados, excepcionalmente teria-se um acréscimo de 8 minutos por caixa no dia.

A partir de então, temos $R\$1,05 * 8$, que seria igual a R\$ 8,04 (oito reais e quatro centavos), repartido para cada planta altera-se o custo diário, para o dia de aplicação para R\$ 0,68 (sessenta e oito centavos). Para determinar em quantos dias o valor inicial de R\$ 2,61 (dois reais e sessenta e um centavos) chegaria a R\$ 3,82 (três reais e oitenta e dois centavos), considerando aumentos diferenciados nos dias 7 e 21, primeiro calculamos o total dos aumentos nos dias 7 e 21, que resultaram em R\$ 1,36

(um real e trinta e seis centavos). Em seguida, subtraímos esse aumento total do aumento necessário para atingir o preço final, resultando em um aumento total restante de -R\$0,15 (quinze centavos). Como o aumento total restante é negativo, significa que os aumentos nos dias 7 e 21 já cobrem o aumento necessário para atingir R\$ 3,82 (três reais e oitenta e dois centavos). Portanto, o valor inicial de R\$ 2,61 (dois reais e sessenta e um centavos) chegaria a R\$ 3,82 (três reais e oitenta e dois centavos) em aproximadamente 11 dias, chegando ao 21 dia de aplicação a um custo de aproximadamente R\$ 4,90 (quatro reais e noventa centavos).

Machine Learning

A pesquisa buscou a aplicação de técnicas de redes neurais profundas presentes na área de comercialização de plantas, para se tornar capaz de processar os dados e determinar dias melhores e quais as características levar em consideração para aplicação de promoções sob o valor de venda de forma a estimular o comércio antes da perda total do produto ou de seu custo de manutenção for maior do que o preço de venda, levando a um prejuízo.

Os modelos criados foram possíveis de serem perfeitamente executados após o devido tratamento dos dados conforme código em apenso. Ao total foram elaborados 3 aprendizados de máquina: uma Árvore de decisão; uma RNN e uma Random Forest, conforme tabela 3.

Tabela 3 - modelo de rede neural simples construído para este projeto

Layer (type)	Output Shape	Parameter
Simple_rnn_2	None, 60, 50	3250
Dropout_2	None, 60, 50	0
Simple_rnn_3	None, 50	5050

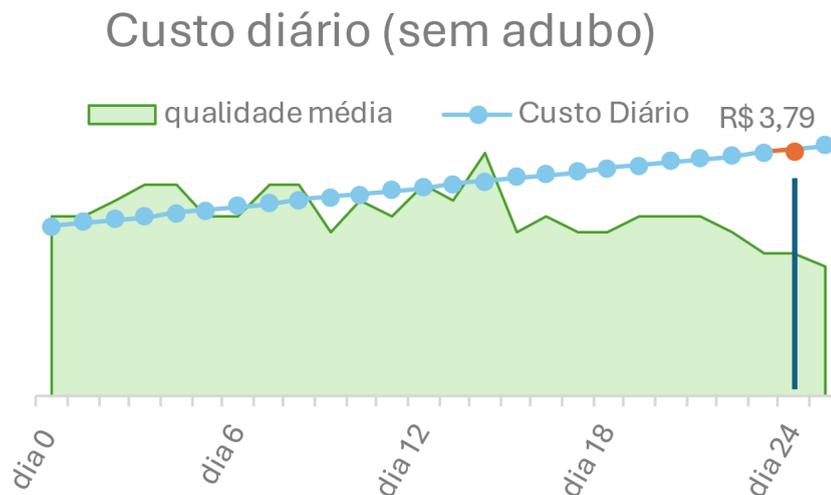
Dropout_3	None, 50	0
Dense_1	None, 1	51

Fonte: código de autoria própria

A árvore de decisão, com seu modelo transparente (considerando sua fácil visualização e interpretação) em uma acurácia de 98%, apontou que os melhores pontos de foco para se seguir no tratamento e melhora foram as características de flores perfeitas, nos momentos em que acabaram de desabrochar seus novos brotos. Sendo assim, os melhores ganhos econômicos seriam no dia de entrada, que contém maior quantidade de flores perfeitas e após o estímulo da floração, seriam os novos momentos de venda e expectativa de ganhos.

Através do gráfico da Figura 6, podemos deduzir as informações anteriores:

Figura 6 - Crescimento diário do custo de uma unidade de planta e sua variação de qualidade

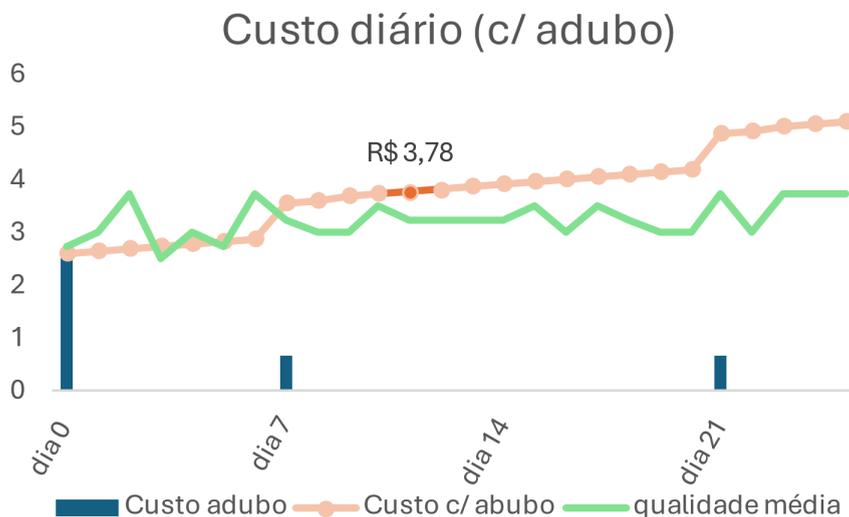


Fonte: base de dados da pesquisa

Ocorre que, nos primeiros dias o custo diário é bem inferior ao preço sugerido de venda, assim como a qualidade é superior; representando o melhor momento com relação aos ganhos. Nos demais dias, a qualidade vai flutuando entre melhor e pior, mas com um ritmo de queda, enquanto o custo só aumentando, até chegar ao seu ponto em que no vigésimo quarto dia seu custo é superior ao preço de venda. E pela qualidade da planta, sua probabilidade de venda é ínfima.

Como demonstrado pelo gráfico da Figura 7, o custo aumenta exponencialmente pelo detalhe que do custo de aplicação do adubo. A aplicação das doses ocorrerá no sétimo dia e no vigésimo primeiro dia. Isso garantiu na primeira aplicação uma recuperação da qualidade da planta que já estava em queda e uma manutenção da qualidade na segunda dose. Mas diferente do que ocorre no custo sem adubação, no décimo dia já temos o ponto de curva em que se não ocorrer a venda, a manutenção da planta ultrapassará o preço praticado, configurando um prejuízo para a empresa.

Figura 7 - Crescimento do custo diário de uma unidade de planta, adicionada de aplicação de adubo e sua variação de qualidade média



Fonte: base de dados da pesquisa

Essas observações tanto da árvore, quanto dos gráficos, foram confirmadas pela observação dos dias de venda. Fatores que foram implementados nos modelos posteriores.

O uso da RNN se deu pelas características de independência condicional das variáveis de entrada e saída. Essa rede poderia detectar relações mais complexas que não foram visualizadas através da árvore de decisão e como dito anteriormente, é uma rede com bastante utilidade para análise de séries temporais. Muito embora seja uma rede adequada para os propósitos deste trabalho, ocorreram algumas divergências.

A análise das métricas de desempenho do modelo de previsão de vendas revela uma performance moderadamente satisfatória.

Figura 8 - métricas obtidas por aplicação dos códigos de RNN

Métrica	Valor
Mean Absolute Error (MAE)	0.616
Mean Squared Error (MSE)	0.878
Root Mean Squared Error (RMSE)	0.937
Coefficient of Determination (R ²)	0.627

Fonte: código de aprendizado de máquina desenvolvido para esta pesquisa

O MAE de 0.616 sugere que, em média, o modelo desvia dessa margem ao estimar o valor de venda, indicando uma proximidade razoável entre as previsões e os valores reais. Por outro lado, o Mean Squared Error (MSE) e o Root Mean Squared Error (RMSE), com valores de 0.878 e 0.937 respectivamente, apontam para a existência de alguns erros significativos nas previsões, o que pode ser um sinal de que o modelo está penalizando mais fortemente os desvios maiores.

O *Coefficient of Determination* (R²), que se situa em 0.627, mostra que o modelo é capaz de explicar aproximadamente 62.74% da variabilidade dos dados de venda. Embora isso demonstre uma capacidade moderada do modelo em capturar a dinâmica das vendas, também indica que existem outros fatores não contemplados pelo modelo que estão influenciando os resultados. Em suma, as métricas refletem um modelo com um bom ponto de partida para previsões, mas que ainda há espaço para melhorias, seja no ajuste fino dos parâmetros, seja na inclusão de outras variáveis explicativas que possam enriquecer a capacidade preditiva do sistema.

Essas divergências ressaltadas pelas métricas estão embasadas pelo fato da base de dados ser de pequena extensão para o próprio algoritmo. Embora seja capaz

de ser bem aplicado ao estudo de plantas⁶⁷, inclusive para detecção de doenças⁶⁸, a sua melhor performance ocorre com base de dados extensos, com uso de imagens e outras medições que não foram possíveis de serem coletadas e produzidas.

Esses fatores levaram a implementação do último algoritmo de aprendizado de máquina que foi o Random Forest. Com base no resultado dos outros dois mecanismos anteriores, foi realizado um último tratamento da base de dados e selecionado como alvo de determinação o ponto de venda. Assim, com a devida separação entre treino e teste, a acurácia chegou a 99,4%.

Figura 9 - tabela comparativa das métricas entre dois algoritmos de aprendizado de máquina diferentes

Métrica	RNN	R. F
Mean Absolute Error (MAE)	0.616	0.0056
Mean Squared Error (MSE)	0.878	0.0056
Root Mean Squared Error (RMSE)	0.937	0.0745
Coefficient of Determination (R ²)	0.627	0.9200

Fonte: código de aprendizado de máquina desenvolvido para esta pesquisa

O resultado demonstrado na Figura 9 mostra a melhora significativa nas métricas obtidas – sendo significativamente mais preciso e confiável pois explica melhor a variância de dados – e ainda demonstrou que a primeira semana é de suma importância para a comercialização do produto. Com maior probabilidade de venda nas caixas 1 e 4, a maior chance de venda se dá durante a primeira semana, apresentando a venda enquanto o custo diário se mantém entre R\$ 2,68 (dois reais e sessenta e oito centavos) e R\$ 3,20 (três reais e vinte centavos).

⁶⁷ AHN, J. Y. et al. Evaluating Time-Series Prediction of Temperature, Relative Humidity, and CO2 in the Greenhouse with Transformer-Based and RNN-Based Models. **Department of Integrative Biological Sciences and Industry**, Sejong, 16 fevereiro 2024. 2-15. P. 5.

⁶⁸ SIVALINGAM, V. et al. A Novel Approach for Plant Leaf Disease Predictions Using Recurrent Neural Network RNN Classification Method. **Journal of Advanced Research in Applied Sciences and Engineering Technology**, 21 abril 2023. 327-338. P. 330.

CONSIDERAÇÕES FINAIS

A presente pesquisa focou na espécie Sunpatiens, conhecida por sua vivacidade e resistência, e como ela se enquadra nesse mercado de varejo de plantas em transformação e crescimento. Com um padrão relativamente consistente de floração ao longo do tempo e sob diversas condições climáticas, a Sunpatiens se mostrou a espécie perfeita para o início dos estudos de desenvolvimento e aprimoramento do algoritmo aqui apresentado.

Os custos empresariais associados ao cultivo e comercialização da espécie selecionada foram avaliados considerando os investimentos em tecnologia disponível e inovação para manter a qualidade e competitividade no mercado e para o propósito preditivo foi possível alcançar medições e previsões do preço de venda e custo desta espécie, podendo ser expandida para outras futuramente para novas espécies, com integração de tecnologias de análise de solo e até mesmo uma expansão na escala de estudos para implementação de outras redes neurais que comportem o aumento exponencial da base de dados.

É importante notar que, embora as empresas que adotam o Simples Nacional tenham facilidades para pagamento de todas as despesas, esses custos não podem ser subestimados. Mesmo empresas de pequeno porte, se não tiverem uma boa circulação de seus produtos, podem enfrentar muitas perdas e despesas além dos lucros. Como muitas dessas empresas não levam em consideração um estudo sobre o seu fluxo de mercadorias e balanço entre lucro e prejuízo utilizando-se de métricas e conhecimentos da ciência de dados.

No caso da espécie Sunpatiens, por exemplo, os custos diários superaram o já na primeira semana os 50% de lucro sobre o produto, finalizando o oitavo dia com perda de 75% do valor da venda somente para cobrir as despesas com uma unidade da planta. Período este (semana) em que também ocorre a perda da qualidade da floração, e em alguns casos, a perda da floração inteira, o que reduz quase que totalmente a possibilidade de realizar uma venda.

Como o período de floração ocorre em um intervalo quinzenal com a devida aplicação de adubo, o ideal pelos dados obtidos seria a venda dentro da primeira

semana, já que na segunda floração o custo seria superior ao preço de venda até então praticado. Fato este confirmado também pelo algoritmo de Random Forest, que indica que mesmo que não se realize uma promoção no final da primeira semana, há uma necessidade de realização de promoção no decorrer das próximas semanas para poder circular o produto e eliminar os gastos da empresa.

Os resultados obtidos com o modelo proposto, apresentou duas sugestões para a empresa que cedeu as plantas para observação: a primeira delas foi a alteração do preço final de venda das plantas. No momento de início da pesquisa até seu desenvolvimento do aprendizado de máquina, o preço calculado foi o praticado no comércio local, que era inferior ao sugestionado. Após os resultados obtidos, a empresa optou por promover a alteração do valor.

Essa alteração de valor tem a capacidade de promover um lucro de 2% ainda que a venda ocorra na segunda semana de aplicação do adubo, caso a floração não ocorra de maneira natural. Ocorrendo uma segunda floração naturalmente, o lucro seria entre 25 a 28% após a terceira semana de exposição

A segunda sugestão de tomada de decisão com relação ao controle de estoque, de forma eficiente e lucrativa apontando quais os dias de maior probabilidade de venda, quais os fatores que mais influenciaram na venda e o momento em que ocorre a venda. Com os cálculos apresentados e gráficos também foi possível determinar uma sugestão de promoção entre o final da primeira semana e o final da segunda, já que mesmo com a redução de valores, o lucro permanece acima dos 40%.

REFERÊNCIAS

AHN, J. Y. et al. Evaluating Time-Series Prediction of Temperature, Relative Humidity, and CO₂ in the Greenhouse with Transformer-Based and RNN-Based Models.

Department of Integrative Biological Sciences and Industry, Sejong, 16 fevereiro 2024. 2-15.

ALOM, Z.; TAHA, T. M.; YAKOPCIC, C. E. A. A State-of-the-Art Survey on Deep Learning Theory and Architectures. **Electronics**, Basel, 31 janeiro 2019. 292-358.

ANDRADE, R.; PALLUETO, A. W. Aspectos institucionais da formação de preços. **Revista Nova Economica**, Minas Gerais, 11 outubro 2019. 565-90.

ANEEL. Custo da energia que chega aos consumidores. **Agência Nacional de Energia Elétrica**, 2022. Disponível em: <<https://www.gov.br/aneel/pt-br/assuntos/tarifas/entenda-a-tarifa/custo-da-energia-que-chega-aos-consumidores>>. Acesso em: 15 fevereiro 2024.

ANTONIK, L. R. **Empreendedorismo**. Kindle. ed. São Paulo: Alta Books, 2023.

BANCO DO BRASIL. Inteligência artificial no agronegócio. **Blog BB**, 2021. Disponível em: <<https://blog.bb.com.br/inteligencia-artificial-agronegocio/>>. Acesso em: 13 abril 2023.

BOEGER, M. R.; WISNIEWSKI, C. Comparação da morfologia foliar de espécies arbóreas de três estádios sucessionais distintos de floresta ombrófila densa (Floresta Atlântica) no Sul do Brasil. **Revista Brasil Botanica**, Curitiba, 26 março 2003. 61-72.

BRANT, R. S. et al. Adaptações fisiológicas e anatômicas de *Melissa officinalis* L. (Lamiaceae) cultivadas sob malhas termorrefletoras em diferentes intensidades luminosas. **Rev. Bras. Pl. Med.**, Botucatu, 22 abril 2011. 467-474.

BRASIL. Lei Complementar Nº 123, de 14 de dezembro de 2006. **Estatuto Nacional da Microempresa e da Empresa de Pequeno Porte**, Brasília, 2006. Disponível em: <https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp123.htm>. Acesso em: 17 janeiro 2024.

BREIMAN, L. Random Forests. **Machine Learning**, 01 outubro 2001. 5-32.

BUDUMA, N.; BUDUMA, N. P. J. **Fundamentals of Deep Learning**. Kindle. ed. Sebastopol: O'Reilly Media, 2022.

BULLINARIA, J. A. Biological Neurons and Neural Networks, Artificial Neurons. **Neural Computation: Lecture 2**, 2015. Disponível em: <https://www.cs.bham.ac.uk/~pjt/NC/l2_JB.pdf>. Acesso em: 10 junho 2024.

CARAMORI, I. DF tem o mesmo número de empreendedores e servidores públicos. **Metrópolis**, 2019. Disponível em: <<https://www.metropoles.com/distrito-federal/economia-df/df-tem-o-mesmo-numero-de-empreendedores-e-servidores-publicos>>. Acesso em: 27 julho 2024.

CARDOSO, J.; VENDRAME, W. Innovation in Propagation and Cultivation of Ornamental Plants. **Environmental horticulture Department**, Gainesville, 2 march 2022. 2-4.

CARON, B. et al. EFICIÊNCIA DO USO DA RADIAÇÃO SOLAR POR PLANTAS *Ilex paraguariensis* A. ST. HIL. CULTIVADAS SOB SOMBREAMENTO E A PLENO SOL. **Ciência Florestal**, Santa Maria, 2 abril 2014. 257-265.

CHEN, N. Exploring the development and application of LSTM variants. **Proceedings of the 4th International Conference on Signal Processing and Machine Learning. Southwest Jiaotong University, 999 Xian Road**, Sichuan, 5 novembro 2021. 7227–7259.

COSTA, A. H.; COZMAN, F. G. O futuro da pesquisa em inteligência artificial. **Dossiê Inteligência Artificial na Pesquisa Científica**, São Paulo, 13 abril 2024. 133-46.

COSTA, V. C. Inteligência artificial identifica plantas doentes simulando processo cerebral. **Embrapa**, 2023. Disponível em: <<https://www.embrapa.br/busca-de-noticias/-/noticia/78204383/inteligencia-artificial-identifica-plantas-doentes-simulando-processo-cerebral>>. Acesso em: 13 março 2024.

CUTTER, E. G. **Plant Anatomy**. 2. ed. Londres: Edward Arnold L:td., 1978.

DARDENGO, J. D. F. et al. ANÁLISE DA INFLUÊNCIA LUMINOSA NOS ASPECTOS ANATÔMICOS DE FOLHAS DE *Theobroma speciosum* WILLD EX SPRENG. (MALVACEAE). **Ciência Florestal**, Santa Maria, 27 julho 2017. 843-851.

DATTA, S. K. Breeding of ornamentals: success and technological status. **The Nucleus**, California, 25 Agosto 2022. 107-28.

DERI SISWARA, A. M. S. A. H. W. Classification Modeling with RNN-Based, Random Forest, and XGBoost for Imbalanced Data: A Case of Early Crash Detection in ASEAN-5 Stock Markets. **Statistics and Data Sciences Department, Faculty of Mathematics and Natural Sciences**, Indonesia, 12 junho 2024.

DISSÍDIOS SALARIAIS. Salário de Contador - Reajuste, dissídio e piso salarial 2024. **Dissídio**, 2024. Disponível em: <<https://dissidio.com.br/salario/contador/>>. Acesso em: 20 abril 2024.

DUBOIS, A.; KUPLA, L.; SOUZA, L. E. D. **Gestão de Custos e Formação de Preços - Conceitos, Modelos e Ferramentas**. 4ª. ed. São Paulo: Atlas, 2024.

EKPA, A. K.; IMOYI, I. U. Recurrent Neural Network (RNN)-Based Injection Substation Load Prediction And Forecasting. **Journal of Multidisciplinary Engineering Science and Technology (JMEST)**, 1 janeiro 2024. 16899-906.

FREDENBURGH, J. How the Covid-19 pandemic hit the cut-flower chain. **BBC**, 2020. Disponível em: <<https://www.bbc.com/future/ bespoke/made-on-earth/how-the-covid-19-pandemic-hit-the-cut-flower-chain.html>>. Acesso em: 18 fevereiro 2024.

GERS FA, S. J. C. F. Learning to forget: continual prediction with LSTM. **Neural Comput**, 12 outubro 2000. 2451-2471.

GOMES, R. et al. Amazon chicory: growing at full sunlight or under shade? **Horticultura Brasileira**, Capanema, 1 abril 2023. 1-6.

GOVERNO DO DISTRITO FEDERAL. LEI Nº 7.368, DE 26 DE DEZEMBRO DE 2023. Anexo I. **Diário Oficial do Distrito Federal**, 2023. Disponível em: <https://www.dodf.df.gov.br/index/visualizar-arquivo/?pasta=2023%7C12_Dezembro%7CDODF%20241%2027-12-2023%20SUPLEMENTO%20B%7C&arquivo=DODF%20241%2027-12-2023%20SUPLEMENTO%20B.pdf>. Acesso em: 12 março 2024.

HASSANI, N. How to Grow and Care for SunPatiens. **The Spruce Make your best home**, 2024. Disponível em: <<https://www.thespruce.com/sunpatiens-growing-guide-5202267>>. Acesso em: 20 fevereiro 2024.

HO, T. K. Random decision forests. **Proceedings of 3rd International Conference on Document Analysis and Recognition**, Montreal, 14 agosto 1995. 278-282.

HOCHREITER, S. . &. S. J. Long Short-Term Memory Spiking Networks and Their Applications. **Neural computation**, 1 novembro 1997. 1735-1780.

IBM. O que é ciência de dados? **Tópicos Ciência de dados**, 2024. Disponível em: <<https://www.ibm.com/br-pt/topics/data-science>>. Acesso em: 22 julho 2024.

INSTITUTO EUROPEU DE CERTIFICAÇÃO DE TECNOLOGIAS DE INFORMAÇÃO - EITCI ASBL. Como o otimizador Adam otimiza o modelo de rede neural? **EITCA**, 2023. Disponível em: <<https://pt.eitca.org/intelig%C3%AAncia-artificial/eitc-ai-dltf-aprendizado-profundo-com-tensorflow/fluxo-tensor/modelo-de-rede-neural/revis%C3%A3o-de-exame-modelo-de-rede-neural/como-o-otimizador-adam-otimiza-o-modelo-de-rede-neural/>>. Acesso em: 4 julho 2024.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. Centro de Previsão de Tempo e Estudos Climáticos. **CPTEC, Arquivos do Laboratório de Meteorologia e Climatologia, vinculados ao CPTEC-INPE**, 2024. Disponível em: <<https://www.cptec.inpe.br/>>. Acesso em: 1 julho 2024.

ISLAM, M. F. et al. Comparative study of sunflower varieties in Bangladesh's coastal region. **International Journal of Multidisciplinary Perspectives**, Bangladesh, 31 dezembro 2023. 110-17.

JORNAL NACIONAL. Pandemia faz a venda de plantas e flores aumentar em todo país. **Globo - G1**, 2021. Disponível em: <<https://g1.globo.com/jornal-nacional/noticia/2021/01/29/pandemia-faz-a-venda-de-plantas-e-flores-aumentar-em-todo-pais.ghtml>>. Acesso em: 22 julho 2024.

LIMA, T. et al. Mudança organizacional: o impacto do home office no comportamento das servidoras e estagiárias da diretoria do Ministério Público de Contas. **Revista de Gestão e Secretariado**, São Paulo, 14 junho 2023. 9033-9051.

LOUPPE, G. Understanding Random Forests: From Theory to Practice. **ARXIV**, 2014. Disponível em: <<https://arxiv.org/abs/1407.7502>>. Acesso em: 15 maio 2024.

NEO ENERGIA. Composição Tarifária. **Entenda sua tarifa**, 2024. Disponível em: <<https://www.neoenergia.com/web/brasil/sua-casa/composicao-tarifaria>>. Acesso em: 15 fevereiro 2024.

NEON. Máquina de cartão para MEI: o que considerar na hora de ter a sua! **Mei Fácil por Neon**, 2024. Disponível em: <<https://blog.meifacil.com/sou-microempreendedor-individual/maquina-de-cartao-para-mei/#:~:text=Para%20te%20ajudar%20nesse%20exemplo%2C%20vamos%20falar%20sobre,%28recebe%20em%2030%20dias%29%20%E2%80%93%203%2C90%25%20por%20parcela>>. Acesso em: 12 maio 2024.

OBA, L. X. et al. Planting fertilization and Metarhizium anisopliae inoculation in the initial growth of sugarcane. **Revista de Agricultura Neotropical**, Viçosa, 30 novembro 2023. 2-5.

OLAH, C. Understanding LSTM Networks. **Github**, 2015. Disponível em: <<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>>. Acesso em: 27 agosto 2024.

PYDATA. PANDAS - Python Data Analysis Library. **Pandas**, 2024. Disponível em: <<https://pandas.pydata.org/about/>>. Acesso em: 15 julho 2024.

RAMIREZ, C. A. APLICACIÓN DEL MACHINE LEARNING EN AGRICULTURA DE PRECISIÓN. **Revista CINTEX**, Manizales, 11 novembro 2020. 14-27.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological Review**, Cornell, junho 1958. 386-408.

SAKADA SEED AMERICA. SunPatients Culture Guide. **Kientzler**, 2024. Disponível em: <<https://www.kientzler.eu/north-america/culture-information/Sakata/SunPatients-Culture-Guide.pdf>>. Acesso em: 20 fevereiro 2024.

SCIKIT-LEARN DEVELOPERS. sklearn.preprocessing. **Scikit Learn**, 2024. Disponível em: <<https://scikit-learn.org/stable/api/sklearn.preprocessing.html>>. Acesso em: 15 julho 2024.

SEBRAE. EPP: entenda o que é uma empresa de pequeno porte. **EMPREENDEDORISMO | EMPREENDEDOR**, 2021. Disponível em: <<https://sebrae.com.br/sites/PortalSebrae/ufs/ac/artigos/epp-entenda-o-que-e-uma-empresa-de-pequeno-porte,305fd6ab067d9710VgnVCM100000d701210aRCRD>>. Acesso em: 20 março 2024.

SEGURA, M. Inteligência artificial aplicada a negócios. **Revista Inteligência Competitiva**, v. 8, n. 3, p. 101-110, jul./set. 2018., Guarulhos, jul/set 2018. 101-110.

SIVALINGAM, V. et al. A Novel Approach for Plant Leaf Disease Predictions Using Recurrent Neural Network RNN Classification Method. **Journal of Advanced Research in Applied Sciences and Engineering Technology**, 21 abril 2023. 327-338.

UNIVERSITY OF MELBOURNE. Classificação climática de Köppen- Geiger. **Academia.EDU**, 2024. Disponível em: <https://www.academia.edu/36723493/Classifica%C3%A7%C3%A3o_clim%C3%A1tica_de_K%C3%B6ppen_Geiger_Classifica%C3%A7%C3%A3o_clim%C3%A1tica_de_K%C3%B6ppen_Geiger?auto=download>. Acesso em: 15 julho 2024.

W3 SCHOOLS. NumPy Introduction. **W3 Schools**, 2024. Disponível em: <https://www.w3schools.com/python/numpy/numpy_intro.asp>. Acesso em: 14 julho 2024.

WALSH I, F. D. G.-G. D. T. T. P. G.; GROUP, E. M. L. F.; HARROW J, P. F. T. S. DOME: recommendations for supervised machine learning validation in biology. **Nat Methods**, 18 outubro 2021. 1122-1127.

WALSH, I.; FISHMAN, D.; AL, E. Recommendations for machine learning validation in biology. **Aristotle University of Thessaloniki**, 2020. Disponível em: <https://www.researchgate.net/publication/342547853_Recommendations_for_machine_learning_validation_in_biology>. Acesso em: 10 julho 2024.

WASKOM, M. An introduction to seaborn. **Seaborn**, 2024. Disponível em: <<https://seaborn.pydata.org/tutorial/introduction>>. Acesso em: 14 julho 2024.

WASKOM, M. seaborn: statistical data visualization. **Seaborn**, 2024. Disponível em: <<https://seaborn.pydata.org/>>. Acesso em: 14 julho 2024.