

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**EZEQUIEL NOVAK**

**DESENVOLVIMENTO DE UM PROTÓTIPO DE ESPECTROSCOPIA DE BAIXO  
CUSTO PARA A AVALIAÇÃO DE MATURAÇÃO DE FRUTAS**

**PONTA GROSSA**

**2025**

**EZEQUIEL NOVAK**

**DESENVOLVIMENTO DE UM PROTÓTIPO DE ESPECTROSCOPIA DE BAIXO CUSTO PARA A AVALIAÇÃO DE MATURAÇÃO DE FRUTAS**

**Development of a low-cost spectroscopy prototype for fruit ripening evaluation**

Trabalho de conclusão de curso de graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia Elétrica da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. Sergio Luiz Stevan Junior.

**PONTA GROSSA**

**2025**



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Esta licença permite download e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**EZEQUIEL NOVAK**

**DESENVOLVIMENTO DE UM PROTÓTIPO DE ESPECTROSCOPIA DE BAIXO  
CUSTO PARA A AVALIAÇÃO DE MATURAÇÃO DE FRUTAS**

Trabalho de conclusão de curso de graduação  
apresentado como requisito para obtenção do título de  
Bacharel em Engenharia Elétrica da Universidade  
Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 17/novembro/2025

---

Sergio Luiz Stevan Junior  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Marcella Scoczynski Ribeiro Martins  
Doutorado  
Universidade Tecnológica Federal do Paraná

---

Julio Cesar Guimaraes  
Doutorado  
Universidade Tecnológica Federal do Paraná

**PONTA GROSSA**

**2025**

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, pela força, sabedoria e perseverança, por me guiar nesse caminho e me ajudar a superar todas as dificuldades durante o curso.

Expresso minha gratidão ao meu orientador Prof. Dr. Sergio Luiz Stevan Junior, pelos ensinamentos, pela orientação atenta e disponibilidade, pela sabedoria com que me guiou no desenvolvimento do trabalho, e pela paciência nos momentos difíceis que enfrentei.

Aos meus colegas de sala e de curso pela amizade, pelo companheirismo, e pelo conhecimento e experiências compartilhados.

Aos professores da UTFPR, pelos ensinamentos no decorrer do curso, que serviram como base não só para desenvolver o projeto, mas também para minha formação pessoal e profissional como engenheiro.

À minha família pelo apoio incondicional durante todo o curso, contribuindo para que eu pudesse superar as dificuldades e os momentos adversos. Especialmente à minha mãe, que mesmo doente, me incentivou a seguir em frente a concluir a minha formação.

Aos colegas do CIEBE pela colaboração, pelos conselhos, e toda a parceria no desenvolvimento do trabalho.

Finalmente, expresso a minha gratidão a todos que de alguma forma contribuíram durante esse período tão importante e enriquecedor da minha jornada.

"A imaginação muitas vezes nos leva a mundos que nunca sequer existiram. Mas sem ela não vamos a lugar algum."  
(Carl Sagan, 1980)

## RESUMO

O presente trabalho aborda o desenvolvimento de um protótipo de espectroscopia focado em baixo custo, projetado para avaliação da maturação de frutas, com foco em tomates e bananas. O estudo fundamenta-se na necessidade de alternativas sustentáveis, não destrutivas e de baixo custo para a análise de qualidade de alimentos, diante dos tradicionais processos destrutivos de análise laboratorial, de alto custo e complexidade. O sistema proposto integra os sensores de espectro AS7341 e AS7263, que operam nas faixas do visível (VIS) e do infravermelho próximo (NIR), juntamente com um microcontrolador ESP32 e uma interface IHM para controle e ajustes de parâmetros do dispositivo, e armazenamento de dados em cartão SD. Para comportar o sistema foi projetada e fabricada uma placa de circuito, juntamente com um invólucro plástico em impressão 3D. A metodologia abrangeu o projeto e montagem do equipamento, desenvolvimento de interface gráfica, coleta experimental de dados espectrais e a aplicação de técnica de processamento de dados: a Regressão Linear Múltipla (MLR), a Regressão por Mínimos Quadrados Parciais (PLSR) e a Rede Neural Artificial *Multiplayer Perceptron* (MLP). Os resultados obtidos demonstram a viabilidade do sistema, com desempenho satisfatório na predição dos estágios de maturação dos frutos, sobretudo nos modelos com ambos os sensores espectrais combinados. Dessa maneira, comprovou-se que o protótipo significa uma opção acessível e eficaz, contribuindo para o panorama de tecnologias de espectroscopia portátil e de baixo custo voltadas à agricultura e ao controle de qualidade de alimentos.

**Palavras-chave:** espectroscopia; baixo custo; maturação de frutas; redes neurais artificiais.

## ABSTRACT

This work addresses the development of a low-cost spectroscopy prototype designed for evaluating fruit ripening, focusing on tomatoes and bananas. The study is based on the need for sustainable, non-destructive, and low-cost alternatives for food quality analysis, as opposed to traditional destructive, costly, and complex laboratory methods. The proposed system integrates the AS7341 and AS7263 spectral sensors, which operate in the visible (VIS) and near-infrared (NIR) ranges, along with an ESP32 microcontroller and an HMI interface for device control, parameter adjustment, and data storage on an SD card. To accommodate the system, a circuit board and a 3D-printed plastic enclosure were designed and manufactured. The methodology encompassed equipment design and assembly, graphical interface development, experimental collection of spectral data, and the application of data processing techniques: Multiple Linear Regression (MLR), Partial Least Squares Regression (PLSR), and a Multilayer Perceptron (MLP) Artificial Neural Network. The results demonstrate the feasibility of the system, showing satisfactory performance in predicting fruit ripening stages, especially in models combining both spectral sensors. Thus, the prototype proved to be an accessible and effective option, contributing to the field of portable, low-cost spectroscopy technologies aimed at agriculture and food quality control.

**Keywords:** spectroscopy; low cost; fruit ripening; artificial neural networks.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Propagação de uma onda eletromagnética no espaço.....	19
Figura 2 – O espectro eletromagnético e suas subdivisões .....	20
Figura 3 – Interação entre radiação e matéria .....	22
Figura 4 – Espectro de absorção da clorofila a (Chl a), clorofila b (Chl b) e carotenoides.....	23
Figura 5 – Configuração básica de espectrofotômetros de Transmitância e Reflectância .....	25
Figura 6 – Exemplo de um MLP com duas camadas ocultas .....	29
Figura 7 – Resumo sobre os canais ópticos do AS7341 .....	36
Figura 8 – Responsividade dos canais do AS7341 em relação ao canal F8 .....	36
Figura 9 – Diagrama de blocos funcionais do AS7341 .....	37
Figura 10 – Módulo de prototipagem com o sensor AS7341.....	39
Figura 11 - Responsividade espectral do sensor AS7263 .....	40
Figura 12 - Características Ópticas do sensor AS7263 .....	40
Figura 13 - Diagrama de blocos funcionais do AS7263 .....	41
Figura 14 - Módulo de prototipagem com o sensor AS7263 .....	42
Figura 15 - pinagem do nodeMCU-32S.....	43
Figura 16 - Display LCD touchscreen 2,8" .....	45
Figura 17 - Face inferior do Display com indicação de pinagem .....	45
Figura 18 - módulo conversor DC-DC <i>step-down</i> XL4015 .....	47
Figura 19 - Layout da placa de circuito .....	49
Figura 20 - Modelo 3D do invólucro plástico .....	50
Figura 21 – protótipo finalizado (a) vista interna, (b) vista externa em perspectiva; e (c) vista superior. ....	51
Figura 22 - IHM (a) tela de abertura e (b) tela de inicialização.....	52
Figura 23 – IHM - Tela de meu principal .....	52
Figura 24 - Telas de monitoramento e suas versões, sendo (a) monitoramento em modo contínuo e (b) modo de leitura única.....	53
Figura 25 - Telas de parametrização, sendo (a) de ajuste de ganho, (b) de ajuste dos LEDs, (c) de ajuste das sensibilidades, (d) de ajuste dos intervalos de leitura e (e) de ajuste do modo de leitura.....	55
Figura 26 - Diagrama de funcionamento da interface IHM.....	56
Figura 27 - Fluxograma de procedimento de aquisição de dados.....	58
Figura 28 - Processo de maturação dos tomates em 5 etapas, considerando seu amadurecimento pós colheita de 5 em 5 dias. ....	60

<b>Figura 29 - Processo de maturação das bananas em 4 etapas, considerado um intervalo de 6 dias entre cada. ....</b>	<b>60</b>
<b>Figura 30 - Exemplos posicionados para a amostragem .....</b>	<b>61</b>
<b>Figura 31 - Regiões de coleta de dados .....</b>	<b>62</b>
<b>Figura 32- Organização dos dados no arquivo .csv utilizado nas análises .....</b>	<b>67</b>
<b>Figura 33 - Gráficos de dispersão real x previsto MLP para os tomates, sendo (a) para o sensor VIS, (b) para o sensor NIR, e (c) para ambos .....</b>	<b>72</b>
<b>Figura 34 - Gráficos de dispersão real x previsto MLP para as bananas, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) modelo com ambos os sensores combinados .....</b>	<b>73</b>
<b>Figura 35 - Gráficos de dispersão real x previsto MLR para os tomates, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) com ambos os sensores combinados .....</b>	<b>75</b>
<b>Figura 36 - Gráficos de dispersão real x previsto MLR para as bananas, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) com ambos os sensores combinados .....</b>	<b>76</b>
<b>Figura 37 - Gráficos de dispersão real x previsto PLSR para os tomates, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) modelo com ambos os sensores combinados .....</b>	<b>78</b>
<b>Figura 38 - Gráficos de dispersão real x previsto PLSR para as bananas, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) modelo com ambos os sensores combinados .....</b>	<b>79</b>
<b>Figura 39 - Gráficos dos VIP Scores para os tomates e as bananas, sendo (a) análise dos tomates e (b) análise das bananas.....</b>	<b>81</b>

## LISTA DE TABELAS

Tabela 1 – Sensores espectrais comerciais e principais características .....	35
Tabela 2 - Comparativo entre microcontroladores populares.....	44
Tabela 3 - ajustes padrão de inicialização .....	56
Tabela 4 - Ajustes utilizados no experimento do trabalho .....	59
Tabela 5 - Exemplificação da repetibilidade de dados para o sensor AS7341 .	63
Tabela 6 - Exemplificação da repetibilidade de dados para o sensor AS7263 .	63
Tabela 7 - Demonstrativo de número de neurônios utilizados na camada oculta da MLP .....	68
Tabela 8 - Resultados MLP para os tomates com dados do AS7341 .....	69
Tabela 9 - Resultados MLP para os tomates com dados do AS7263 .....	69
Tabela 10 - Resultados MLP para os tomates com dados de ambos os sensores .....	69
Tabela 11 - Resultados MLP para as bananas com dados do AS7341 .....	70
Tabela 12 - Resultados MLP para as bananas com dados do AS7263.....	70
Tabela 13 - Resultados MLP para as bananas com dados de ambos os sensores.....	70
Tabela 14 - Resultados para a Regressão Linear Múltipla (MLR) .....	74
Tabela 15 - Resultados para a Regressão por Mínimos Quadrados parciais ...	77
Tabela 16 - Tabela comparativa dos resultados .....	82

## LISTA DE ABREVIATURAS E SIGLAS

A/D	<i>Analog-to-Digital</i> (Analógico-Digital)
ADC	<i>Analog-to-Digital Converter</i> (Conversor Analógico-Digital)
BPNN	<i>Backpropagation Neural Network</i> (Rede Neural com Retropropagação)
CIEBE	Centro de Instrumentação em Engenharia Biomédica e Esportiva
CMOS	<i>Complementary Metal-Oxide-Semiconductor</i> (Semicondutor de Óxido de Metal Complementar)
CPU	<i>Central Processing Unit</i> (Unidade Central de Processamento)
DAC	<i>Digital-to-Analog Converter</i> (conversor digital-analógico).
DC-DC	<i>Direct Current to Direct Current</i> (Corrente Contínua para Corrente Contínua)
FIR	<i>Far Infrared</i> (infravermelho Distante)
FWHM	<i>Full Width at Half maximum</i> (largura total à meia altura)
GC-MS	<i>Gas Chromatography–Mass Spectrometry</i> (Cromatografia Gasosa acoplada à Espectrometria de Massas)
GPIO	<i>General Purpose Input/Output</i> (entradas/saídas de uso geral).
IBGE	Instituto Brasileiro de Geografia e Estatística
I <sup>2</sup> C	<i>Inter-Integrated Circuit</i> (Circuito Inter-Integrado)
IHM	Interface Homem Máquina
I/O	<i>I/O – Input/Output</i> (Entrada/Saída)
LCD	<i>Liquid Crystal Display</i> (Display de Cristal Líquido)
LED	<i>Light Emitting Diode</i> (Diodo Emissor de Luz)
MIR	<i>Mid Infrared</i> (Infravermelho Médio)
MLP	<i>Multilayer Perceptron</i>
MLR	<i>Multiple Linear Regression</i> (Regressão Linear Múltipla)
MQO	Mínimos Quadrados Ordinários
MSE	<i>Mean Squared Error</i> (Erro Quadrático Médio)
NIR	<i>Near Infrared</i> (Infravermelho Próximo)
ONU	Organização das Nações Unidas
pH	Potencial Hidrogeniônico
PLSR	<i>Partial Least Squares Regression</i> (Regressão por Mínimos Quadrados Parciais)
PWM	PWM – <i>Pulse Width Modulation</i> (Modulação por Largura de Pulso)
RAM	<i>Random Access Memory</i> (Memória de Acesso Aleatório)
RGB	<i>Red, Green, Blue</i> (Vermelho, Verde, Azul)

RMSE	<i>Root Mean Squared Error</i> (Raiz do Erro Quadrático Médio)
RNAs	Redes Neurais Artificiais
SI	Sistema Internacional de Unidades
SMUX	<i>Synchronous Multiplexer</i> (Multiplexador Síncrono)
TFT	<i>Thin-Film Transistor</i> (Transistor de Película Fina)
TSS	<i>Total Soluble Solids</i> (Sólidos Solúveis Totais)
UART	<i>Universal Asynchronous Receiver/Transmitter</i> (Transceptor Assíncrono Universal)
UTFPR	Universidade Tecnológica Federal do Paraná
VIP	<i>Variable Importance in Projection</i> (Importância da Variável na Projeção)
VIS	<i>Visible Spectrum</i> (Espectro Visível)

## LISTA DE SÍMBOLOS

C	Carbono
H	Hidrogênio
O	Oxigênio
S	Enxofre

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>15</b>
<b>1.1</b>	<b>Objetivos</b> .....	<b>17</b>
1.1.1	Objetivo geral .....	17
1.1.2	Objetivos específicos.....	17
<b>1.2</b>	<b>Justificativa</b> .....	<b>18</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b> .....	<b>19</b>
<b>2.1</b>	<b>O Espectro Eletromagnético</b> .....	<b>19</b>
<b>2.2</b>	<b>Interação da radiação com a matéria</b> .....	<b>21</b>
2.2.1	Espectroscopia na faixa visível.....	22
2.2.2	Espectroscopia na faixa do Infravermelho Próximo (NIR) .....	23
<b>2.3</b>	<b>Fundamentos da espectroscopia: absorção e refletância</b> .....	<b>24</b>
<b>2.4</b>	<b>Funcionamento e componentes de um espectrômetro</b> .....	<b>25</b>
2.4.1	Fonte de luz.....	26
2.4.2	Elemento seletor de comprimento de onda (monocromador).....	26
2.4.3	Detector .....	27
2.4.4	Gerenciamento eletrônico .....	27
<b>2.5</b>	<b>Ferramentas de aprendizado de máquina: uma visão abrangente</b> ..	<b>27</b>
2.5.1	Redes Neurais Artificiais <i>Multilayer Perceptron</i> (MLP) .....	27
2.5.2	Regressão Linear Múltipla (MLR): .....	29
2.5.3	Regressão por Mínimos Quadrados Parciais (PLSR). .....	30
2.5.4	Métricas de avaliação de desempenho .....	32
<b>3</b>	<b>MATERIAIS E MÉTODOS</b> .....	<b>34</b>
<b>3.1</b>	<b>Materiais</b> .....	<b>34</b>
3.1.1	Sensores de espectro AS7341 (VIS-NIR) e AS7263 (NIR) .....	34
<u>3.1.1.1</u>	<u>Sensor AS7341 (VIS–NIR)</u> .....	<u>35</u>
<u>3.1.1.2</u>	<u>Sensor AS7263 (NIR)</u> .....	<u>39</u>
3.1.2	Microcontrolador ESP32.....	42
3.1.3	Display de controle .....	44
3.1.4	LEDs externos e controle de alimentação. ....	46
<b>3.2</b>	<b>Projeto da placa de circuito e invólucro plástico</b> .....	<b>48</b>
<b>3.3</b>	<b>Projeto da interface IHM</b> .....	<b>51</b>
<b>3.4</b>	<b>Metodologia experimental</b> .....	<b>57</b>
<b>3.5</b>	<b>Amostras</b> .....	<b>59</b>

3.6	Processamento de dados .....	64
4	RESULTADOS E DISCUSSÃO .....	67
4.1	Base de dados .....	67
4.2	Métricas e resultados para a <i>Multilayer Perceptron</i> (MLP) .....	68
4.3	Resultados da Regressão Linear Múltipla (MLR) .....	74
4.4	Resultados da Regressão por Mínimos Quadrados Parciais (PLSR) .....	77
4.5	Comparativo entre os diferentes métodos.....	82
5	CONCLUSÃO .....	83
	REFERÊNCIAS.....	85
	APÊNDICE A - Códigos em <i>python</i> dos métodos de análise utilizados 91	
	ANEXO A - Registradores dos sensores espectrais.....	103

## 1 INTRODUÇÃO

O crescimento populacional tem sido um assunto de interesse e preocupação da sociedade como um todo. Segundo a ONU, em seu relatório *Perspectivas da População Mundial 2024*, a expectativa é que a população mundial atual seja de 8,2 bilhões de pessoas, podendo chegar a 10,3 bilhões em 2080. Já no Brasil, segundo o IBGE, estima-se que a população possa chegar a 220 milhões de pessoas em 2041.

Diante desse cenário, surgem preocupações crescentes com relação à segurança alimentar e à capacidade produtiva do setor agrícola. Com o aumento da populacional, cresce também a demanda por alimentos, e conseqüentemente, a demanda por meios produtivos mais eficientes desde o plantio, colheita, passando pelo escoamento da produção e inspeções de qualidade até chegar ao consumidor final, gerando uma preocupação mundial em como fazer o escalonamento de toda a cadeia produtiva de maneira sustentável e eficiente, para que haja uma maior produção e que o desperdício durante o processo produtivo seja o mínimo possível.

Portanto, torna-se essencial aumentar a eficiência dos processos produtivos para garantir a segurança alimentar e minimizar impactos ambientais. Perdas ocorrem em diversas etapas, desde a colheita até o consumo final, sendo necessário adotar estratégias que reduzam essas ineficiências (NASCIMENTO, 2018).

Nesse contexto, é imprescindível a busca por novas tecnologias que possam proporcionar o aumento da eficiência da produção agrícola e pecuária, destacando ferramentas capazes de monitorar a qualidade dos alimentos de forma rápida, precisa e sustentável.

A busca por métodos alternativos e eficientes para o monitoramento e avaliação da qualidade de alimentos no geral é um tema de muita importância, especialmente diante das necessidades de avaliação em larga escala. Segundo Pereira (2017), as técnicas tradicionais utilizadas para análise de alimentos, como a inspeção visual, enfrentam limitações significativas, incluindo subjetividade, custo elevado, e baixo rendimento na análise de grandes volumes.

Dentre as alternativas tecnológicas emergentes, a espectroscopia tem avançado significativamente, capaz de fornecer informações detalhadas sobre a composição química estado físico dos alimentos. Grande parte do conhecimento atual sobre a estrutura da matéria vem das investigações utilizando essa técnica, que fornecem informações sobre a interação da radiação com a matéria, oferecendo

informações detalhadas através dos espectros de emissão e absorção (ALCANTARA JR, 2002).

Na agricultura, são utilizados diversos tipos de espectroscopia para análise de alimentos e culturas, incluindo as mais comuns: espectroscopia gasosa (acoplada à cromatografia gasosa – GC-MS) para de terminação e quantificação de elementos específicos em compostos voláteis, sendo uma técnica de ampla utilidade na identificação de pesticidas em diferentes matrizes agrícolas (PEREIRA et al., 2019); a espectroscopia de infravermelho próximo (NIR), uma técnica rápida e eficiente, para verificar a composição nutricional (EMBRAPA, 2018); já a espectroscopia ultravioleta-visual (UV-VIS), técnica sensível e de baixo custo, empregada na detecção de contaminação por agrotóxicos e adulteração (VALDERRAMA et al., 2014); e por fim, a espectroscopia Raman para identificar doenças em plantas e grãos, e a presença de baixas concentrações de defensivos agrícolas em alimentos (SILVA, 2024).

Mesmo com os avanços no campo da espectroscopia, o mercado ainda carece de equipamentos de baixo custo e de caráter portátil que sejam capazes de fazer o monitoramento da qualidade de alimentos em campo, com eficiência e sustentabilidade.

Diversos autores têm aplicado com sucesso dispositivos simples para a análise de alimentos, principalmente utilizando espectroscopia de NIR. Deshpande et al. (2021), evidenciaram a aplicabilidade de um dispositivo de baixo custo na análise de adulteração do leite, ao detectar adulterações e prever a pureza de amostras com alta precisão, graças à integração com algoritmos de aprendizado de máquina.

Sulistyo et al. (2021), desenvolveram um espectrômetro portátil de NIR baseado em Arduino para estimar, de forma não destrutiva, a qualidade de laranjas da variedade Siamesa. O dispositivo foi capaz de prever com boa precisão propriedades químicas como acidez (pH) e sólidos solúveis totais (TSS), utilizando redes neurais com retro propagação (BPNN).

Na área biomédica, Ionescu e Paşca (2019) propuseram um protótipo para monitoramento não invasivo da glicose no sangue utilizando sensores espectrais AS7263 e AS7262 que utilizam luz visível e Infravermelho, aliados a algoritmos de interpretação cromática e de absorbância. O sistema foi capaz de estimar concentrações de glicose com base em valores RGB e curvas de absorção, demonstrando o potencial da espectroscopia para medições ópticas em tecidos vivos.

Tais estudos demonstram a viabilidade e eficácia de utilização de dispositivos de baixo custo não só na cadeia produtiva de alimentos, mas também em outras áreas, como na área biomédica, reforçando sua relevância em diferentes contextos da sociedade e da tecnologia.

Nesta perspectiva, este trabalho apresenta o desenvolvimento e construção de um protótipo de espectrofotômetro NIR portátil, utilizando uma plataforma microcontrolada em conjunto com sensores espectrais comerciais de baixo custo, com controle via interface IHM, sendo aplicada na análise da maturação de tomates e bananas para obtenção de prova de conceito (PoC), com a coleta de dados espectrais e análise utilizando ferramentas estatísticas de regressão multivariada e análise computacional através da aplicação de ferramentas baseadas em redes neurais artificiais (RNAs).

## **1.1 Objetivos**

### **1.1.1 Objetivo geral**

O objetivo geral deste trabalho de conclusão de curso é o desenvolvimento, teste e validação de um protótipo de baixo custo de espectroscopia NIR, baseada no espectro visível e do infravermelho próximo, com aplicação de prova de conceito para a avaliação e classificação da maturação de frutas baseado em NIR, sendo esta uma forma alternativa promissora para a avaliação da qualidade de alimentos.

### **1.1.2 Objetivos específicos**

Os objetivos específicos deste trabalho incluem:

- Desenvolvimento de um protótipo eletrônico utilizando sensores espectrais comerciais e um microcontrolador de baixo custo;
- Programação e testes para o funcionamento e controle independente, utilizado um interface homem-máquina (IHM);
- Desenvolvimento de um procedimento experimental de coleta de dados com o protótipo, para minimização de erros de medida e parametrização de variáveis relacionadas ao procedimento de coleta;
- Criação de uma base de dados com informações obtidas com o protótipo, de diferentes frutas com diferentes graus de maturação;

- Aplicação de técnicas de classificação de dados por meio de Ferramentas de aprendizado de Máquina, análise e discussão dos resultados.

## 1.2 Justificativa

As análises químicas tradicionais, amplamente difundidas na avaliação da qualidade de alimentos, como a cromatografia gasosa e titulação, são técnicas com importantes limitações: são geralmente destrutivas, requerem preparo de amostras previamente, utilizam reagentes químicos que geram resíduos, necessitando serem tratados, além de necessitarem de um tempo considerável e infraestrutura laboratorial especializada. Essas limitações tornam esses métodos menos atrativos e viáveis para a aplicação em campo e em ambientes produtivos de larga escala, que necessitam de processos ágeis e eficientes (METROHM, 2021).

Com esse cenário, a espectroscopia, especialmente nos comprimentos de onda do infravermelho próximo (NIR), é uma alternativa viável, por ser uma técnica não destrutiva, aplicada diretamente sobre o objeto alvo de análise, sem necessidade de preparo de amostras. Por ser sustentável e não gerar resíduos, essa alternativa permite análises em tempo real, com a possibilidade de ser integrada em sistemas de automação (ALMEIDA, 2015).

Atualmente, a maioria dos equipamentos de espectroscopia disponíveis comercialmente são de uso laboratorial, tendo como principal característica o alto custo de aquisição e manutenção, limitando a sua utilização e disseminação em propriedades rurais, centros de distribuição e cooperativas (EMBRAPA, 2018). Diante desse contexto, o desenvolvimento de dispositivos portáteis de baixo custo, que monitorem algumas poucas, porém efetivas, faixas de comprimento de onda, é uma solução promissora e viável para aplicações preliminares e com medições em campo.

Este trabalho se fundamenta, portanto, pela proposta de um sistema eletrônico compacto microcontrolado, acessível e eficiente, com interface simplificada, capaz de realizar análise espectrais simplificadas, com a aplicação na avaliação da maturação de tomates e bananas. A combinação entre o hardware de baixo custo, a interface IHM de operação desenvolvida, e a análise computacional de dados, faz do projeto um tema relevante no contexto tecnológico, com potencial impacto na melhoria do processo produtivo e de controle de qualidade de alimentos.

## 2 FUNDAMENTAÇÃO TEÓRICA

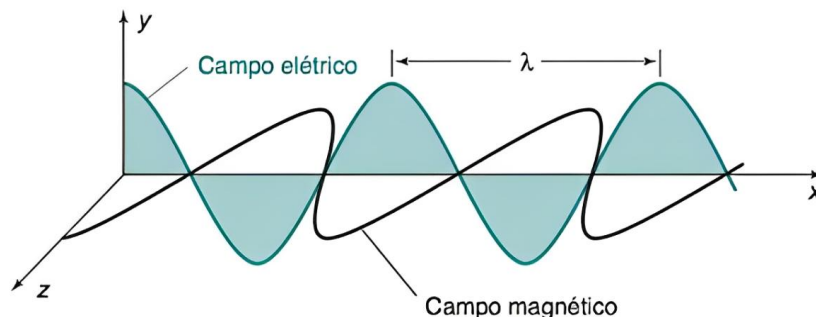
As técnicas espectroscópicas possuem um potencial inegável e diversas áreas de aplicação. Embora um espectrômetro básico seja composto por poucos elementos, estes são de alta sensibilidade. Conceitos complexos de óptica e eletrônica estão envolvidos, e qualquer minúcia tem o potencial de interferir significativamente no resultado final.

Por conta desses pormenores, é essencial compreender as partes que compõem o espectrômetro. Isso permite otimizar a construção do aparelho, garantindo acessibilidade e resultados coerentes. Para isso, este capítulo iniciará com conceitos importantes sobre o espectro eletromagnético e sua interação com a matéria, os quais são os fundamentos da espectroscopia. Serão abordadas também as diferentes aplicações dessa técnica, o funcionamento dos sensores de espectro e as opções comerciais disponíveis. Todos esses conceitos servem de base para definir os componentes de um espectrofotômetro básico, cuja sensibilidade é alta tanto a aspectos construtivos quanto a interferências externas, conforme será explorado no tópico 3.

### 2.1 O Espectro Eletromagnético

A luz é uma onda eletromagnética definida como o conjunto de campos elétrico e magnético que oscilam continuamente no espaço, orientados perpendicularmente um ao outro, conforme ilustrado na figura 01.

**Figura 1 – Propagação de uma onda eletromagnética no espaço**



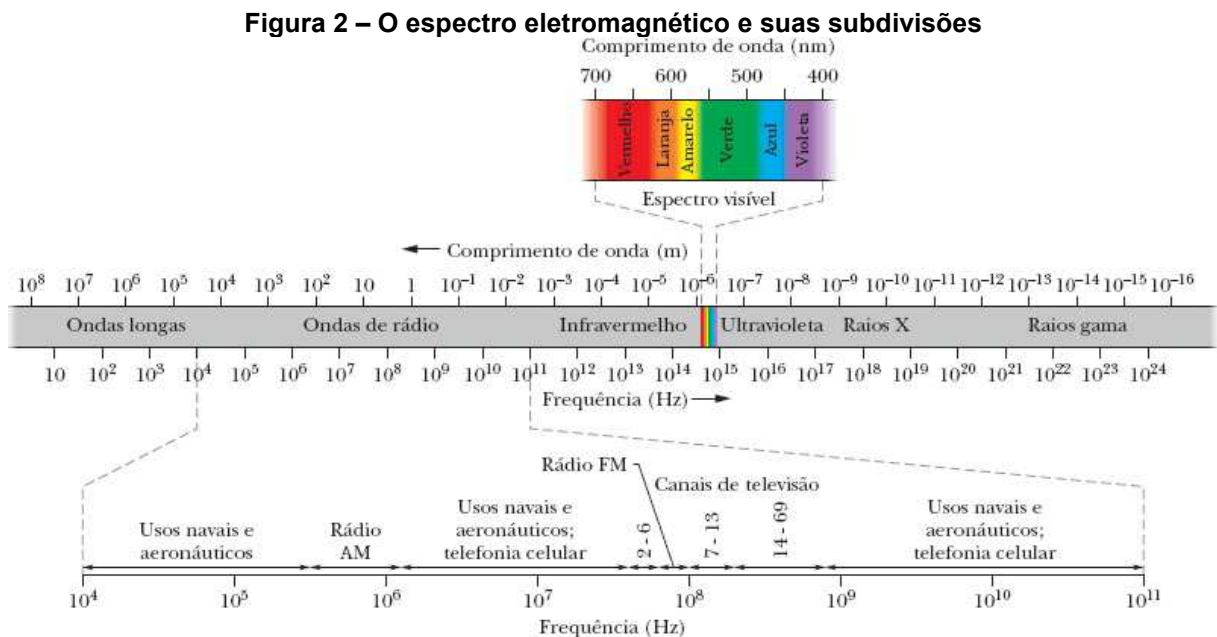
**Fonte: Harris (2012)**

O campo elétrico oscila no plano xy, perpendicularmente, o campo magnético oscila no plano xz. A distância entre as cristas das ondas, que marca um ciclo

completo de oscilação, é o comprimento de onda ( $\lambda$ ). A frequência ( $f$ ) da onda é definida como a quantidade de ciclos completos de oscilação feitos em um segundo, que no Sistema Internacional de Unidades (SI), recebe a unidade de *Hertz* (Hz). A equação (1) define a velocidade de propagação ( $v$ ) de uma onda eletromagnética, como sendo o comprimento de onda multiplicado pela frequência da mesma.

$$v = \lambda \cdot f \quad (1)$$

No SI a unidade de  $v$  é metros por segundo (m/s), sendo que, para luz no vácuo,  $v$  tem o valor de aproximadamente  $2,998 \times 10^8$  m/s (HARRIS, 2012). Essas grandezas são responsáveis por definir em qual região do espectro eletromagnético uma onda eletromagnética está localizada. A combinação entre todas as frequências possíveis forma o espectro eletromagnético (HALLIDAY; RESNICK; WALKER, 2016). As subdivisões do espectro eletromagnético, com ênfase na faixa visível pelo ser humano estão apresentadas na figura 2.



Fonte: Halliday, Resnick e Walker (2016)

As faixas do espectro visível (400 nm a 780 nm) e do infravermelho próximo (de 780 nm a 2500 nm) são especialmente importantes na espectroscopia, por interagirem com diversas substâncias químicas. Essa interação será abordada com mais profundidade no próximo tópico. As diversas bandas do espectro

eletromagnético interagem com a matéria de formas distintas, gerando alterações a nível atômico ou molecular, sendo que, as de baixa energia como a luz visível e infravermelho causam excitação eletrônica ou vibrações moleculares (HARRIS, 2012).

## **2.2 Interação da radiação com a matéria**

A espectroscopia é fundamentalmente definida pela forma como a radiação eletromagnética interage com a matéria. Quando uma onda eletromagnética incide sobre um material, ela pode ser absorvida, transmitida, refletida ou espalhada. O resultado dessa interação é determinado por fatores como a energia da radiação incidente, a estrutura molecular específica do material e a faixa do espectro eletromagnético utilizada (SKOOG et al., 2006).

As abordagens mais amplamente exploradas na espectroscopia são as de absorção e emissão, com particular destaque para as regiões do espectro visível e do infravermelho próximo (NIR). A proeminência dessas regiões se deve, em grande parte, à sua vasta aplicação na análise de alimentos (GOETTEMS, 2017; SILVA, 2021). Uma das grandes vantagens da espectroscopia é sua característica não invasiva (LEITE E PRADO, 2012). Contudo, é essencial selecionar a banda de radiação apropriada para cada análise, a fim de evitar danos à amostra e maximizar a capacidade analítica da técnica.

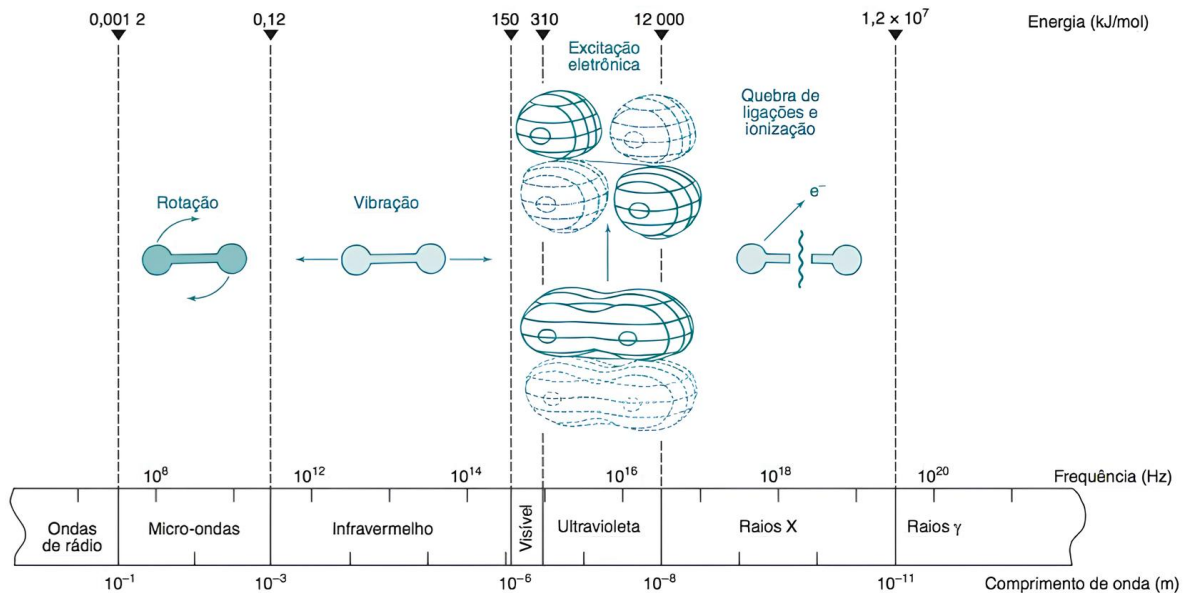
A determinação do comprimento de onda a ser utilizado depende da substância (amostra) que está sendo analisada, pois os comprimentos de onda podem afetar sua absorção (fluorescência). Por essa razão, as regiões do espectro mais frequentemente utilizadas compreendem o visível (400 nm a 700 nm) e o infravermelho próximo (700 nm a habitualmente até 1300 nm, podendo chegar a 2500 nm).

As interações podem ser classificadas em duas categorias principais: ionizantes e não ionizantes (SALICIO E CASTRO). A radiação ionizante, exemplificada pelos raios X e raios gama, possui energia suficiente para causar uma ruptura irreversível na estrutura da matéria. Em contrapartida, a radiação não ionizante provoca apenas desorganizações temporárias, permitindo que a matéria retorne ao seu estado fundamental de menor energia após um período, um fenômeno conhecido como excitação.

A Figura 3 apresenta como a radiação interage com as moléculas nas diferentes frequências, causando quebras de ligações e ionização nas frequências de

raios gama, raios X e ultravioleta, excitações eletrônicas na faixa visível e vibrações e rotações nas faixas do infravermelho e micro-ondas, respectivamente.

**Figura 3 – Interação entre radiação e matéria**



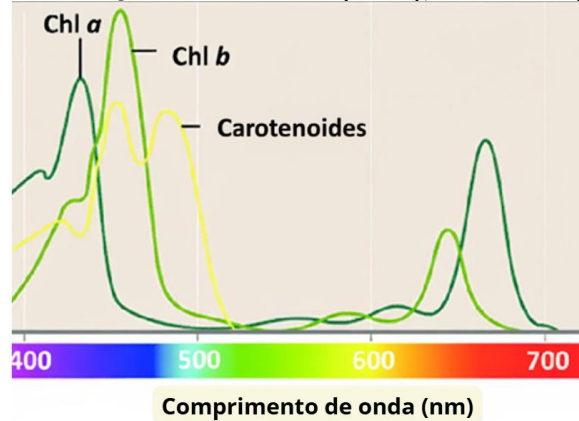
Fonte: Adaptado de Harris (2012)

De acordo com a figura 3, a radiação na faixa de micro-ondas provoca o movimento de rotação das moléculas, já na faixa do infravermelho essa interação causa a vibração das moléculas. Na região visível e do ultravioleta ocorrem as excitações eletrônicas, onde os elétrons saltam para níveis orbitais de maior energia. Nas faixas de alta energia, como o ultravioleta de comprimento de onda curto e a faixa de raios X, a radiação possui energia suficiente para romper ligações químicas (HARRIS, 2012).

### 2.2.1 Espectroscopia na faixa visível

Dentro da faixa do espectro visível, que se estende de aproximadamente 400 nm a 700 nm, diversas substâncias pigmentares presentes nos alimentos exibem picos de absorção característicos. As clorofilas a e b, por exemplo, absorvem fortemente nas regiões do azul (430 nm a 470 nm) e do vermelho (630 nm a 680 nm), resultando na reflexão da cor verde. Já os carotenoides absorvem predominantemente na faixa de 400 nm a 500 nm, o que confere a coloração amarela e alaranjada observada em diversas frutas (TAIZ ET AL., 2017). A Figura 4 ilustra esses espectros de absorção específicos.

**Figura 4 – Espectro de absorção da clorofila a (Chl a), clorofila b (Chl b) e carotenoides**



Fonte: Adaptado de Guidi, Tattini e Landi (2017)

Essa característica de absorção e reflexão dos pigmentos é de grande valia, pois permite avaliar sua presença por meio da análise espectral na faixa visível. Conseqüentemente, torna-se possível realizar a análise do processo de maturação de frutas de maneira não destrutiva, o que representa uma vantagem significativa em diversas aplicações.

### 2.2.2 Espectroscopia na faixa do Infravermelho Próximo (NIR)

Na faixa do infravermelho, temos algumas subdivisões: o infravermelho próximo (NIR), infravermelho médio (MIR) e o infravermelho distante (FIR). Dentro da faixa do NIR, que se estende de 780 nm a 2500 nm, a radiação interage principalmente com as moléculas em nível vibracional e rotacional. Essas interações causam estiramentos e deformações em determinadas ligações químicas, como N-H, O-H, C-H e S-H, presentes em diversos compostos orgânicos (BARRETO, 2020).

Essas excitações originam bandas de absorção do espectro sobrepostas, chamadas de sobretons e bandas de combinação, que são características dessas moléculas heteronucleares, formadas por elementos químicos distintos. Devido a essa característica, os compostos químicos com ligações dessa natureza têm bandas de absorção bem aparentes, diferentemente das moléculas homonucleares, como o O<sub>2</sub> e o N<sub>2</sub>, que não absorvem significativamente dentro dessa faixa do espectro (MAGALHÃES, 2014). Portanto, por meio desse fenômeno, é possível identificar as propriedades físico-químicas das amostras, como teor de água, carboidratos e

proteínas, justificando a escolha do NIR como faixa de interesse nas análises qualitativas e quantitativas da qualidade de alimentos.

O quadro 1 mostra as faixas do espectro e as interações com diferentes ligações químicas.

**Quadro 1 - Faixas de interação do espectro com algumas ligações químicas presentes em compostos orgânicos**

Comprimento de Onda [nm]	Comportamento
2500	Alongamento da ligação S-H
2200-2460	Alongamento da ligação C-H
2000-2200	Alongamento da ligação N-H e O-H
1620-1800	Primeiro sobretom do alongamento da ligação C-H
1400-1600	Primeiro sobretom do alongamento da ligação N-H e O-H
1300-1420	Alongamento da ligação C-H
1100-1225	Segundo sobretom do alongamento da ligação C-H
1020-1060	Alongamento da ligação S=O
950-1100	Segundo sobretom do alongamento da ligação N-H e O-H
850-950	Terceiro sobretom do alongamento da ligação C-H
775-850	Terceiro sobretom do alongamento da ligação N-H
600-700	Alongamento da ligação C-S
450-550	Alongamento da ligação S-S

Fonte: Adaptado de Magalhães (2014)

### 2.3 Fundamentos da espectroscopia: absorção e refletância

A espectroscopia, compreendida a interação entre radiação e matéria, é uma técnica analítica baseada na absorção, emissão e espalhamento da radiação eletromagnética em diferentes faixas de comprimentos de onda, o que permite a identificação da composição química e concentração de substâncias específicas a partir de suas respostas espectrais (SILVA, 2021; TAIZ et al., 2016). As técnicas baseadas em absorção e em refletância são amplamente utilizadas para análises dessa natureza.

Na espectroscopia de absorção, a radiação incide sobre a amostra, atravessando-a. Parte da energia é absorvida de acordo com a composição da amostra, e parte a atravessa. Para realizar a medição do espectro absorvido, primeiramente se detecta a transmitância, que é a parte da radiação que passa pela amostra. A absorbância então é calculada a partir do espectro de transmitância, utilizando os conceitos da Lei de Beer-Lambert. Esta técnica é bastante empregada em análise de soluções e líquidos (HARRIS, 2012).

Por outro lado, na espectroscopia de reflectância, a radiação incidente é parcialmente absorvida e parcialmente refletida pela superfície da amostra, onde a parte refletida é então detectada. Essa forma de emprego é útil para analisar objetos

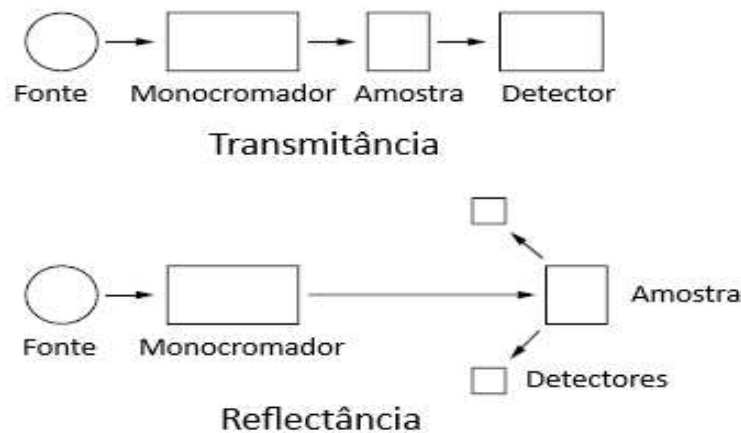
sólidos e opacos, onde não ocorre a transmitância, como, por exemplo, frutas e o solo (MAGALHÃES, 2014).

## 2.4 Funcionamento e componentes de um espectrômetro

Um espectrofotômetro básico possui alguns componentes básico que desempenham funções essenciais para o seu correto funcionamento, possibilitando a obtenção de espectros característicos das amostras analisadas. Os principais componentes são: fonte de radiação, sistema seletor de comprimento de onda (monocromador), detector e sistema de aquisição de dados.

A figura 5 mostra o conceito de construção de espectrofotômetros de transmitância e de reflectância, respectivamente, onde esses conceitos de espectroscopia de absorção e refletância são utilizados. Como o objetivo do trabalho é analisar frutas, o método construtivo de reflectância foi o utilizado, por ser o mais adequado para essa aplicação, de acordo com a bibliografia.

**Figura 5 – Configuração básica de espectrofotômetros de Transmitância e Reflectância**



**Fonte: Adaptado de Burns e Ciurczak (2007)**

A seguir serão descritos a função de cada componente e seus correspondentes utilizados na construção do protótipo do presente trabalho. Os detalhes construtivos do protótipo e as características dos componentes serão mais detalhadas no tópico 3.

#### 2.4.1 Fonte de luz

É a origem da radiação eletromagnética que incide sobre a amostra. Nos espectrofotômetros laboratoriais são utilizadas lâmpadas de tungstênio, halógenas ou de deutério. No entanto, em dispositivos portáteis com foco em baixo custo, é mais adequada a utilização de LEDs emissores em diferentes faixas do espectro, como na região visível e do NIR. LEDs são vantajosos não apenas do ponto de vista do custo, mas também no baixo consumo energético, tamanho reduzido estabilidade de emissão espectral e longa vida útil. É importante conhecer o perfil espectral da fonte e avaliar a intensidade na região de interesse e que ela mantenha estabilidade durante as medições (SKOOG et al., 2017).

#### 2.4.2 Elemento seletor de comprimento de onda (monocromador)

A função do monocromador é isolar uma faixa específica do espectro para análise. Em espectrofotômetros tradicionais, esse elemento corresponde a prismas ou redes de difração que decompõem as diferentes frequências da fonte de radiação, permitindo que apenas alguns determinados comprimentos de onda atinjam a amostra. No presente projeto esse elemento funciona de maneira diferente, onde foram utilizados sensores espectrais de multicanais com filtros de interferência integrados, nesse caso os sensores AS7341 (VIS-NIR) e o AS7263 (NIR), os quais se baseiam em *arrays* de fotodiodos dedicados a frequências espectrais específicas. Esses filtros e fotodiodos que respondem a faixas distintas do espectro substituem a necessidade de dispositivos monocromadores, reduzindo o custo, complexidade e tamanho do equipamento (SKOOG, HOLLER; CROUCH, 2017; SIESLER, 2008).

Os fotodiodos são sensores semicondutores que convertem a intensidade de luz em corrente elétrica. Quando a radiação incide sobre o fotodiodo, ele emite um sinal proporcional à intensidade de luz recebida, apresentando alta sensibilidade, rapidez de resposta e funcionamento eficaz em faixas que abrangem o ultravioleta, visível e infravermelho próximo. Por essas características, é um componente muito disseminado como detector em espectrofotômetros por sua sensibilidade, rapidez de resposta e compatibilidade com diferentes comprimentos de onda (SKOOG et al., 2017).

### 2.4.3 Detector

É o sensor responsável por converter a quantidade de luz recebida nas diferentes faixas de comprimento de onda em um sinal elétrico. No presente trabalho os sensores AS7263 e AS7341 são os detectores. Cada canal óptico dos sensores corresponde a um fotodiodo com resposta calibrada, onde o mesmo converte a intensidade da radiação recebida dentro de uma faixa de comprimentos de onda específica, convertendo em corrente elétrica contínua, que depois é convertida em um sinal digital. Esse sinal digital é então recuperado via interface I<sup>2</sup>C (AMS, 2019).

### 2.4.4 Gerenciamento eletrônico

O gerenciamento dos sensores e a recuperação e exibição dos dados é feito por uma unidade microcontrolada. No escopo do protótipo, esse papel é executado por um microcontrolador ESP32. O mesmo oferece alta capacidade de processamento, diferentes interfaces de comunicação, baixo consumo e comunicação sem fio, além do tamanho reduzido (ESPRESSIF, 2022).

## 2.5 Ferramentas de aprendizado de máquina: uma visão abrangente

A capacidade de prever resultados futuros ou de compreender as relações subjacentes em conjuntos de dados é fundamental em diversas áreas do conhecimento, especialmente em espectroscopia, onde a análise de diferentes faixas do espectro é fundamental para estabelecer relações entre a resposta espectral em cada comprimento de onda e características físico-químicas das amostras. Para tal, ferramentas estatísticas e de inteligência artificial oferecem abordagens distintas, cada qual com suas particularidades e aplicações ideais. Este capítulo se aprofunda em três metodologias proeminentes: a *Multilayer Perceptron* (MLP), a Regressão Linear Múltipla (MLR) e a Regressão por Mínimos Quadrados Parciais (PLSR). Compreender seus fundamentos é crucial para a escolha da metodologia mais adequada em diversas pesquisas científicas.

### 2.5.1 Redes Neurais Artificiais *Multilayer Perceptron* (MLP)

A *Multilayer Perceptron* (MLP) se destaca como um dos pilares das redes neurais artificiais *feedforward*, célebre por sua notável capacidade de modelar relações não lineares complexas entre dados de entrada e saída. Como destaca

Haykin (1999), a arquitetura *feedforward* assegura que a informação flua unidirecionalmente, da entrada para a saída, sem laços de retroalimentação.

Uma MLP é essencialmente uma estrutura de camadas interconectadas de neurônios: uma camada de entrada, pelo menos uma camada oculta e uma camada de saída. Cada conexão entre neurônios possui um peso específico, que é ajustado durante o processo de aprendizado.

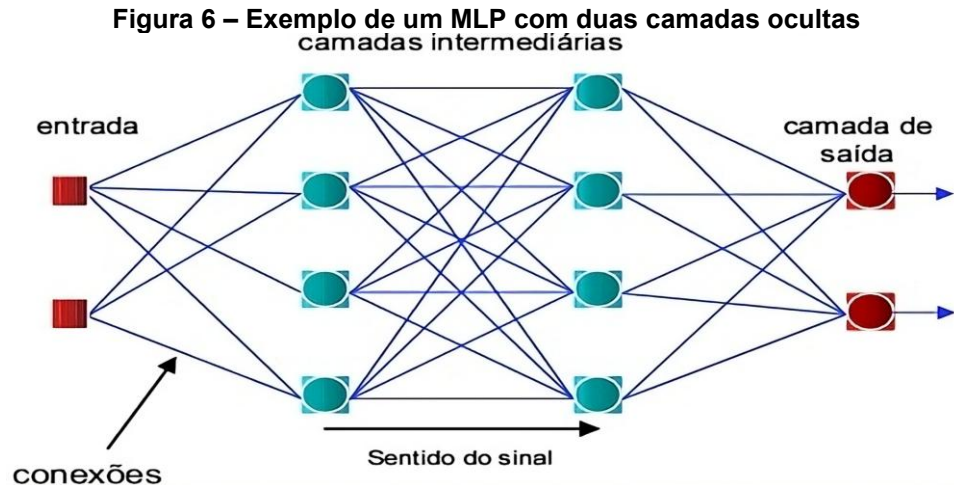
O funcionamento da MLP se desdobra em três etapas principais:

- Propagação Direta (*Forward Propagation*): Os dados de entrada são alimentados à camada inicial. Cada neurônio subsequente calcula uma soma ponderada de suas entradas, que é então transformada por uma função de ativação não linear (como sigmoide, ReLU ou tanh). Esta função é o que confere à rede sua aptidão para aprender padrões complexos e não lineares, conforme a informação avança até a camada de saída.
- Cálculo do Erro: A saída gerada pela rede é comparada com o valor real esperado. A diferença é quantificada por uma função de custo, sendo o erro quadrático médio (MSE) uma das mais empregadas.
- Retropropagação do Erro (*Backpropagation*): O erro calculado é então "propagado" de volta pela rede, ajustando iterativamente os pesos de cada conexão. Este ajuste, que visa minimizar o erro, segue o algoritmo de gradiente descendente, conforme detalhado por (RUMELHART et al., 1986).

Para um neurônio  $j$  em uma camada, a saída  $y_j$  é calculada por:

$$y_j = \phi \sum_{i=1}^n (W_{ji}X_i + B_j) \quad (2)$$

Onde  $X_i$  são as entradas,  $W_{ji}$  são os pesos das conexões,  $B_j$  é o viés (bias) do neurônio  $j$ ,  $n$  é o número de entradas, e  $\phi$  é a função de ativação. A capacidade da MLP de aprender e generalizar a partir de dados a torna uma ferramenta poderosa para uma vasta gama de aplicações preditivas, como podemos observar na figura 7.



Fonte: Vieira (2006)

Além da estrutura da rede, o desempenho de uma MLP depende fortemente do ajuste de seus hiperparâmetros, como a função de ativação (por exemplo, ReLU, sigmoide ou tangente hiperbólica), o número de épocas de treinamento, a divisão dos dados entre treino, validação e teste, o número de camadas ocultas e o número de neurônios por camada. Esses fatores controlam tanto a capacidade de aprendizado da rede quanto sua generalização. Um modelo com muitos parâmetros pode se ajustar perfeitamente aos dados de treinamento, mas falhar na previsão de novos dados, o que é um problema de ajuste conhecido como *overfitting*. Por isso, é fundamental aplicar técnicas de validação cruzada, regularização e testes iterativos para encontrar a configuração mais adequada ao problema (HAYKIN, 2009).

No contexto do presente trabalho, a MLP foi utilizada para prever o estágio de maturação através dos dados dos sensores espectrais em diferentes comprimentos de onda, permitindo uma análise conjunta, e demonstrado como os dados do espectro visível e NIR podem se complementar.

### 2.5.2 Regressão Linear Múltipla (MLR):

A Regressão Linear Múltipla (MLR) é um modelo estatístico robusto, amplamente utilizado para prever o valor de uma variável dependente (resposta) a partir de duas ou mais variáveis independentes (preditoras). Sua premissa fundamental é a existência de uma relação linear entre a variável dependente e as variáveis preditoras, um conceito bem estabelecido na estatística, como descrevem (MONTGOMERY et al.,2012).

A ideia da MLR reside na busca pela melhor linha, ou hiperplano, que se ajusta aos dados. Este ajuste é alcançado minimizando-se a soma dos quadrados dos resíduos, que representam as diferenças entre os valores observados e os previstos pelo modelo. O principal objetivo é estimar os coeficientes que definem essa relação linear.

A equação geral da MLR é expressa como:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_p X_p + \varepsilon \quad (3)$$

Nesta formulação:

- $Y$  simboliza a variável dependente.
- $X_1, X_2, \dots, X_p$  são as  $p$  variáveis independentes.
- $\beta_0$  é o intercepto, indicando o valor de  $Y$  quando todas as variáveis independentes são zero.
- $\beta_1, \beta_2, \dots, \beta_p$  são os coeficientes de regressão, que quantificam a mudança média em  $Y$  para cada unidade de aumento na respectiva variável independente, mantendo as demais constantes.
- $\varepsilon$  é o termo de erro aleatório, que engloba a variância não explicada pelo modelo.

Os coeficientes  $\beta$  são tipicamente estimados através do método dos mínimos quadrados ordinários (MQO), que busca otimizar o ajuste da linha aos pontos de dados. Apesar de sua simplicidade teórica, a MLR é uma ferramenta poderosa para desvendar e quantificar relações lineares em diversos fenômenos.

No presente trabalho, a MLR foi utilizada como método de referência, estimando o dia de maturação do fruto a partir dos dados espectrais, permitindo avaliar o desempenho de modelos mais avançados frente a uma abordagem estatística tradicional.

### 2.5.3 Regressão por Mínimos Quadrados Parciais (PLSR).

A Regressão por Mínimos Quadrados Parciais (PLSR) emerge como uma técnica multivariada sofisticada, que integra elementos da análise de componentes principais (PCA) com a abordagem da regressão múltipla. Sua relevância é particularmente acentuada em cenários de dados desafiadores, como aqueles com

um elevado número de variáveis preditoras altamente correlacionadas (multicolinearidade), ou quando o número de preditores excede o de observações, conforme detalhado por (WOLD ET AL.,2001).

Diferentemente da MLR, que examina as variáveis preditoras individualmente, a PLSR constrói um novo conjunto de variáveis latentes (também conhecidos como componentes latentes ou fatores PLS. Estes componentes são derivados tanto das variáveis preditoras quanto da variável dependente, e são construídos de maneira a maximizar a covariância entre elas. Subsequentemente, uma regressão é conduzida utilizando esses componentes latentes em relação à variável dependente.

A grande vantagem da PLSR reside em sua capacidade de mitigar os problemas de multicolinearidade e alta dimensionalidade. Ela realiza isso ao extrair a informação mais relevante dos dados, transformando-os em um conjunto reduzido de componentes.

Embora a PLSR não possa ser expressa por uma única equação linear direta como a MLR, seu funcionamento pode ser compreendido em termos de decomposição e projeção. O processo busca encontrar pesos  $w_j$  para as variáveis preditoras  $X$  e  $c_j$  para a variável resposta  $Y$ , de modo que maximizem a covariância entre as pontuações (*scores*)  $t_j = X_{jI}$  e  $u_j = Y_{.cJ}$

Em sua essência, a PLSR visa modelar a relação entre as matrizes de variáveis preditoras  $X$  e de variáveis de resposta  $Y$  da seguinte forma:

$$X=TP^T+E \quad (4)$$

$$Y=UQ^T+F \quad (5)$$

$$U=TB+G \quad (6)$$

Nestas equações:

- $T$  é a matriz de pontuações (*scores*) de  $X$  (os componentes latentes).
- $P$  é a matriz de carregamentos (*loadings*) de  $X$ .
- $U$  é a matriz de pontuações de  $Y$ .
- $Q$  é a matriz de carregamentos de  $Y$ .
- $E, F, G$  representam as matrizes de resíduos.
- $B$  é a matriz de coeficientes que estabelece a relação entre  $U$  e  $T$ .

A previsão de  $Y$  a partir de  $X$  é então efetivada por meio dos componentes latentes. A PLSR é um processo iterativo, que constrói esses componentes de forma sequencial, oferecendo uma solução robusta para conjuntos de dados complexos onde a MLR tradicional encontraria dificuldades. Uma métrica importante dentro da utilização do PLSR é a análise dos *VIP Scores* (*Variable Importance in Projection*). Os VIPs avaliam o quanto cada componente é relevante para prever a variável dependente, no geral um  $VIP \geq 1$  indica que o componente possui relevância significativa para o modelo, já valores menores significam que aquele componente não tem relevância significativa no modelo de predição (FARRÉS et al., 2015).

#### 2.5.4 Métricas de avaliação de desempenho

Para avaliar a precisão e a capacidade de generalização dos modelos de regressão, é de suma importância a utilização de métricas estatísticas que possam comparar os valores preditos pelo modelo com os valores reais observados. Nesse trabalho, três dessas métricas foram empregadas na análise da performance dos modelos: o Erro Médio Quadrático (MSE), a Raiz do Erro Médio Quadrático (RMSE) e o Coeficiente de Determinação ( $R^2$ ).

O Erro Quadrático Médio (*Mean Squared Error*) é a média aritmética dos quadrados das diferenças entre os valores reais e os valores previstos. Ao elevar essa diferença ao quadrado, o MSE proporciona uma contribuição mais acentuada dos erros maiores em relação aos menores, sendo uma métrica muito utilizada na análise de desempenho. Sendo assim, o MSE é muito relevante como parâmetro de calibragem dos modelos de aprendizado de máquina durante o treinamento. A fórmula de cálculo do MSE é a seguinte:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

Onde  $n$  é o número de amostras,  $y_i$  é o valor real da  $i$ -ésima amostra e  $\hat{y}_i$  é o valor previsto pelo modelo para a  $i$ -ésima amostra.

Já a Raiz do Erro Médio Quadrático (*Root Mean Squared Error*) é, basicamente, a raiz quadrada do MSE. A principal vantagem de se analisar o RMSE é que o erro é expresso na mesma unidade dimensional da variável de interesse, o

que facilita a interpretação da magnitude do erro médio do modelo estudado. Sua fórmula pode ser expressa como:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

Por último, o Coeficiente de determinação ( $R^2$ ) é uma medida adimensional que indica a proporção da variância total da variável dependente que é explicada pelo modelo de regressão. O valor de  $R^2$  varia de 0 a 1, onde valores mais próximos de 1 indicam um melhor ajuste do modelo aos dados observados. O cálculo é apresentado pela equação a seguir:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (9)$$

Onde  $\bar{y}$  é a média dos valores reais observados, e as demais variáveis correspondem as da fórmula do MSE. Quando comparamos as métricas, o MSE e RMSE quantificam o erro absoluto, enquanto o  $R^2$  indica a qualidade do ajuste (MONTGOMERY et al., 2012).

### 3 MATERIAIS E MÉTODOS

No presente tópico serão abordadas as características construtivas do protótipo, essenciais para compreender o seu funcionamento, desde a escolha dos componentes, passando pelo projeto da placa e das partes mecânicas. Também será abordada o método experimental, detalhando o procedimento de aquisição de dados, bem como o tratamento dos mesmos, passando pelas ferramentas e softwares utilizados. Por meio desse detalhamento, busca-se esclarecer os pormenores sobre as etapas de projeto e construção, elucidando e justificando escolhas no processo de desenvolvimento do protótipo de espectrofotômetro VIS-NIR.

#### 3.1 Materiais

Nesta subseção, serão abordados os materiais utilizados, como os sensores, LEDs, microcontrolador, tela IHM, involucro plástico, e suas principais características elétricas e mecânicas.

##### 3.1.1 Sensores de espectro AS7341 (VIS-NIR) e AS7263 (NIR)

Os sensores de espectro são os componentes principais do espectrofotômetro, a partir da sua escolha se iniciou o desenvolvimento do protótipo, pois todos os demais componentes foram escolhidos e projetados de forma que atendessem a compatibilidade e garantissem o funcionamento adequado do aparelho. Para escolher o sensor, os principais requisitos foram o baixo custo, e a abrangência das faixas do espectro visível e do NIR. Optou-se por módulos compactos ante montagem de *arrays* com fotodiodos, por serem menores e já integrarem filtros óticos, circuitos de controle, amplificação e conversores A/D calibrados. Essas características trazem enorme vantagem, eliminando a necessidade de desenvolvimento de circuitos complexos de leitura e condicionamento de sinais adicionais.

Dentre os modelos comerciais disponíveis, os sensores AS7341 e AS7263 se destacam pelo preço reduzido, abrangências de faixas espectrais visíveis e do NIR, complementaridade de faixas espectrais entre si e fácil conexão e comunicação, via protocolo I<sup>2</sup>C (AMS OSRAM, 2021). A tabela a seguir, traz um comparativo entre alguns modelos de sensores comerciais, destacando a faixa espectral, número de

canais, faixa de preço, protocolo de comunicação e a resolução, considerando a largura total à meia altura FWHM<sup>1</sup> (*Full Width at Half Maximum*).

**Tabela 1 – Sensores espectrais comerciais e principais características**

Sensor	Faixa espectral (nm)	Nº de canais	Protocolo de comunicação	Resolução (FWHM)	Faixa de preço (USD)	Aplicação típica
<b>AS7341</b> (ams OSRAM)	350 – 1000	11	I <sup>2</sup> C	~20 nm	10 – 15	VIS + NIR, análise de cor, pigmentos, aplicações agrícolas
<b>AS7263</b> (ams OSRAM)	610 – 860	6	I <sup>2</sup> C	~20 nm	15 – 20	NIR, detecção de água, açúcares e compostos orgânicos
<b>AS7262</b> (ams OSRAM)	450 – 650	6	I <sup>2</sup> C	~20 nm	15 – 20	VIS, análise de cor e pigmentos
<b>Hamamatsu C12880MA</b>	340 – 850	288	SPI / UART	15 nm	250 – 300	Espectroscopia portátil, biomédica, química
<b>Hamamatsu C12666MA</b>	340 – 780	256	SPI / UART	15 nm	250 – 300	Portátil, pesquisa, qualidade de alimentos
<b>Avantes AvaSpec-Mini</b>	360 – 1100	~2048 (CCD)	USB	< 2 nm	2.000 – 4.000	Pesquisa, análises industriais

**Fonte: Adaptado de dados dos fabricantes (ams OSRAM, Hamamatsu, Avantes, 2024)**

Fazendo um comparativo, percebe-se que os sensores AS7341 e AS7263 se destacam sobretudo no preço acessível, além disso o AS7341 se sobrepõe ao AS7262 pelo número de canais e faixa de frequência coberta estendida ao NIR. Já os sensores Hamamatsu, mais sofisticados e com resolução e número de canais muito superiores, tem um custo muito mais alto para o foco do trabalho. Já o espectrômetro da Avante, é um equipamento laboratorial de alta precisão, e conseqüentemente, alto custo, também incompatível com o âmbito do trabalho.

### 3.1.1.1 Sensor AS7341 (VIS–NIR)

O AS7341 é um sensor multicanal compacto fabricado pela ams OSRAM, possui 11 canais espectrais distribuídos na faixa do espectro visível e NIR. Oito canais operam em frequências visíveis (415, 445, 480, 515, 555, 590, 630 e 680 nm), um canal na faixa NIR (~910 nm), além de canais *Flicker Detection* (Detecção de

<sup>1</sup> A resolução espectral de um sensor ou espectrômetro pode ser expressa em termos de FWHM (*Full Width at Half Maximum*, ou largura total à meia altura). Esse parâmetro indica a largura, em nanômetros, de um pico de resposta espectral medido à metade de sua intensidade máxima. Em espectroscopia, é um indicador importante da resolução espectral do sensor (SKOOG; HOLLER; CROUCH, 2017).

cintilação 50 Hz a 2 kHz) e *Clear* (sem filtro). Os filtros gaussianos presentes em de cada canal, são implementados por tecnologia de interferência óptica em CMOS, com resolução espectral (FWHM) variando de 26 a 52 nm, dependendo do canal (AMS OSRAM, 2021). A figura 7 mostra a relação de canais e comprimentos de onda, e a largura de banda a maia altura de cada canal (FWHM).

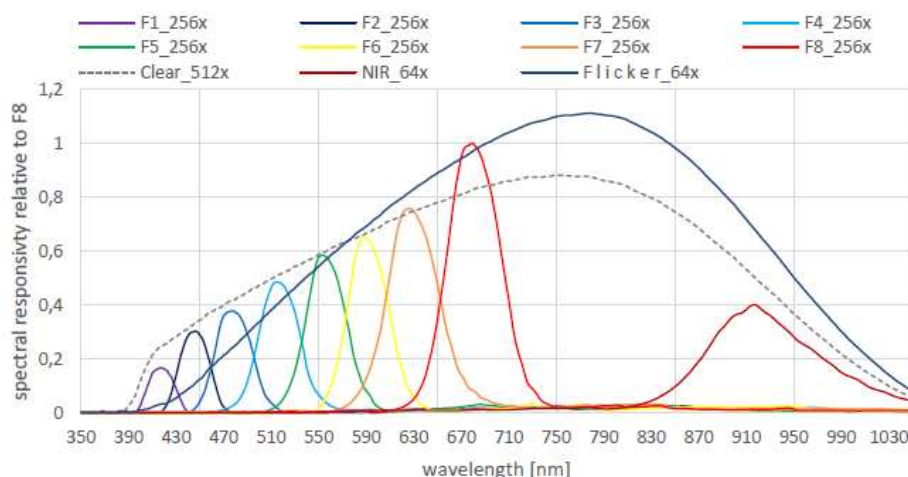
**Figura 7 – Resumo sobre os canais ópticos do AS7341**

Channel	Center Wavelength [nm] typical	Full Width Half Maximum [nm] typical
F1	415	26
F2	445	30
F3	480	36
F4	515	39
F5	555	39
F6	590	40
F7	630	50
F8	680	52
NIR (Near IR)	910	n/a
Clear	Si response/non filtered	n/a
FD (Flicker Detection)	Si response/non filtered	n/a

**Fonte: ams OSRAM (2021)**

Já a figura 8 demonstra graficamente a distribuição de canais e a responsividade espectral típica relativa de cada canal ao canal F8 (680 nm).

**Figura 8 – Responsividade dos canais do AS7341 em relação ao canal F8**



**Fonte: ams OSRAM (2021)**

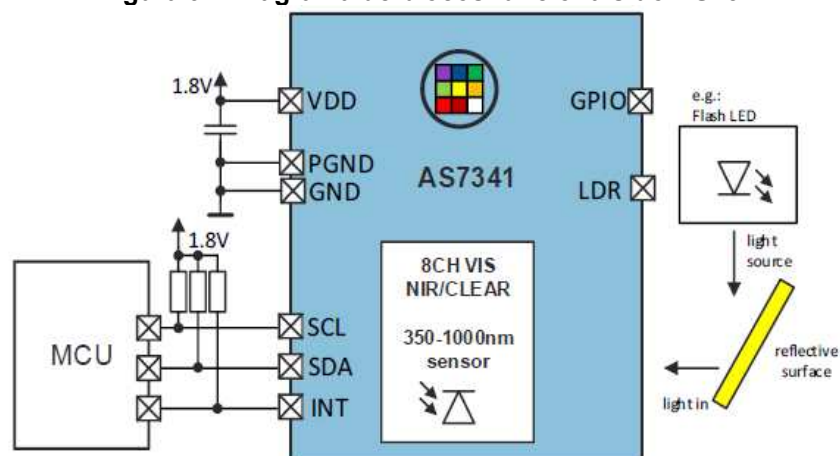
Na figura, o multiplicador na legenda de cada canal refere-se ao ganho aplicado ao sensor (por exemplo “F1\_256x” indica um ganho de 256 vezes aplicado

ao canal F1). Percebe-se que os canais tem respostas distintas para a mesma intensidade de luz recebida, dentro de seu comprimento de onda designado. Essa característica implica na necessidade de normalizar os dados recebidos pelo sensor, de modo que todos os canais sejam comparáveis em uma mesma escala. Para isso, na análise de dados, demonstrada mais adiante, aplica-se a padronização estatística (*Z-score*), que centraliza os valores em torno da média e os divide pelo desvio padrão, garantindo que cada canal contribua de forma equilibrada para a análise espectral e evitando que diferenças de intensidade comprometam os resultados dos modelos.

Sobre as características construtivas, o sensor utiliza uma matriz de fotodiodos 4×4, onde os canais são alocados por meio de um multiplexador interno (SMUX), que direciona os sinais para os 6 Conversores Analógico-Digital (ADC) de 16 bits disponíveis no chip. Esse recurso exige dois ciclos de integração para leitura de todos os canais, pois apenas seis canais podem ser lidos por ciclo (AMS OSRAM, 2021).

Sobre as características elétricas, o sensor opera em 1.8V, possui baixo consumo (entre 210 e 300  $\mu\text{A}$ ), favorecendo aplicações portáteis à bateria. A comunicação é realizada via protocolo I<sup>2</sup>C, com endereço fixo 0x39, onde o chip atua como escravo, respondendo as requisições do microcontrolador conectado. Além disso, o chip possui pino GPIO configurável e pino LDR, designado para o controle de LEDs externos, útil para aplicações de refletância. A figura 9 demonstra o diagrama de blocos funcionais do sensor (AMS OSRAM, 2021).

**Figura 9 – Diagrama de blocos funcionais do AS7341**



Fonte: ams OSRAM (2021)

O AS7341 possibilita alguns ajustes importantes em relação à sua sensibilidade. Os tempos de integração, regem quanto o sensor fica exposto à luz para registrar uma leitura, afetando o valor digital final. Esse tempo é ajustado pelos registradores *ASTEP* (16 bits) e *ATIME* (8 bits), permitindo ajustes de 0 a 999 e 0 a 255, respectivamente. O tempo total de integração é calculado pela fórmula:

$$t_{int} = (ATIME + 1) \times (ASTEP + 1) \times 2,78 \mu s \quad (10)$$

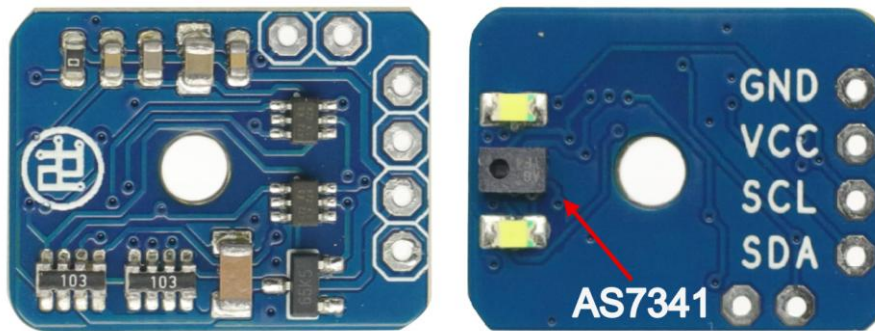
Onde *ATIME* e *ASTEP* são os valores digitais inseridos nos respectivos registradores. O ajuste desses valores deve ser feito de maneira que os valores digitais dos canais mais sensíveis não atinjam a saturação do ADC, e também não devem ter um valor muito baixo, pois podem resultar em sinais pouco significativos ruidosos.

Outro ajuste importante é o parâmetro de ganho *AGAIN*, que oferece vários multiplicadores digitais (0.5×, 1×, 2×, 4×, 8×, 16×, 32×, 64×, 128×, 256×, 512×). Esse ganho amplifica o sinal digital oriundo dos ADCs, aumentando a sensibilidade do sensor (AMS OSRAM, 2021). Também é um que necessita parâmetro de ajuste adequando, para evitar saturação dos ADCs ou gerar sinais pouco significativos e ruidosos.

No Anexo A, está disposta a relação de registradores, endereços e suas respectivas funções do sensor espectral AS7341. No tópico de procedimento de aquisição de dados estão definidos os ajustes adotados e justificativa.

O sensor AS7341 possui versão comercial em um módulo de prototipagem rápida, modelo B0DBQKDV67. Essa placa integra o AS7341, com dois LEDs para iluminação de amostrar em aplicações de refletância. Infelizmente não foi possível encontrar informações sobre a curva de emissão espectral desses leds. Na figura a seguir temos o módulo B0DBQKDV67.

**Figura 10 – Módulo de prototipagem com o sensor AS7341**



Fonte: Adaptado de (MANUALS.PLUS, 2024)

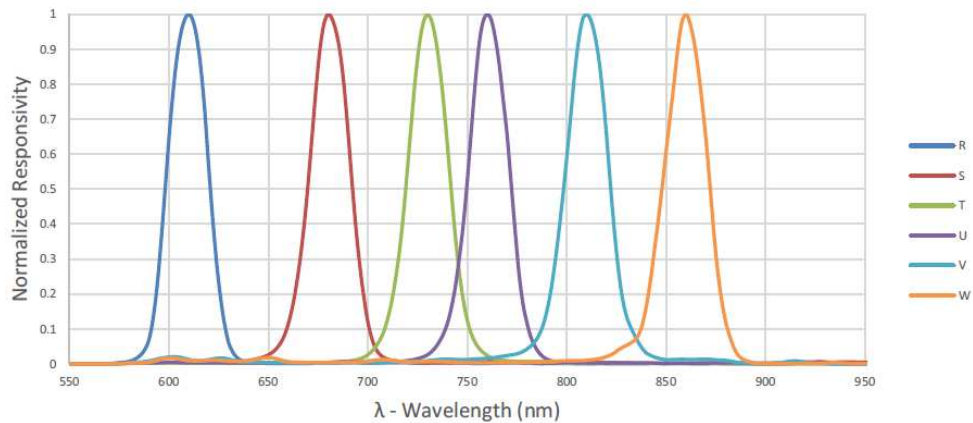
O módulo possui 6 pinos de conexão, descritos a seguir:

- VCC – Pino de alimentação 3.3 a 5 V.
- GND – Pino terra (referência).
- SCL – Pino de clock do protocolo I<sup>2</sup>C.
- SDA – Pino de dados do protocolo I<sup>2</sup>C.
- INT – Pino de interrupção configurável.
- GPIO – Pino I/O de uso geral.

### 3.1.1.2 Sensor AS7263 (NIR)

O sensor AS7263, também fabricado de ams OSRAM, é um espectrômetro compacto multicanal voltado para as faixas de espectro do infravermelho próximo (NIR). Ele possui 6 canais ópticos independentes, centrados nas faixas de 610, 680, 730, 760, 810 e 860 nm. O chip emprega filtros de interferência gaussianos nano-ópticos, de maneira semelhante ao AS7341, com largura total à meia altura (FWHM) de 20 nm para todos os canais (AMS OSRAM, 2021). A figura 10 expressa a responsividade normalizada típica de cada canal do AS7263.

**Figura 11 - Responsividade espectral do sensor AS7263**  
AS7263



Fonte: ams OSRAM (2021)

A figura 12 apresenta um resumo sobre as características ópticas de cada canal do AS7263 apresentadas anteriormente.

**Figura 12 - Características Ópticas do sensor AS7263**

Symbol	Parameter	Test Conditions	Channel (nm)	Min	Typ	Max	Unit
R	Channel R	Incandescent <sup>(2), (4)</sup>	610		35 <sup>(3), (4)</sup>		counts/ ( $\mu\text{W}/\text{cm}^2$ )
S	Channel S	Incandescent <sup>(2), (4)</sup>	680		35 <sup>(3), (4)</sup>		counts/ ( $\mu\text{W}/\text{cm}^2$ )
T	Channel T	Incandescent <sup>(2), (4)</sup>	730		35 <sup>(3), (4)</sup>		counts/ ( $\mu\text{W}/\text{cm}^2$ )
U	Channel U	Incandescent <sup>(2), (4)</sup>	760		35 <sup>(3), (4)</sup>		counts/ ( $\mu\text{W}/\text{cm}^2$ )
V	Channel V	Incandescent <sup>(2), (4)</sup>	810		35 <sup>(3), (4)</sup>		counts/ ( $\mu\text{W}/\text{cm}^2$ )
W	Channel W	Incandescent <sup>(2), (4)</sup>	860		35 <sup>(3), (4)</sup>		counts/ ( $\mu\text{W}/\text{cm}^2$ )
FWHM	Full Width Half Max		20		20		nm
Wacc	Wavelength Accuracy				$\pm 5$		nm

Fonte: ams OSRAM (2021)

Diferentemente do AS7341, o AS7263 conta com um conversor A/D de 16 bits para cada um dos 6 canais espectrais, não necessitando de multiplexador auxiliar para o processo de leitura dos canais. O chip opera com tensão de alimentação de 2,7 a 3,6 V, consumindo até 5 mA em plena operação. A comunicação também é via

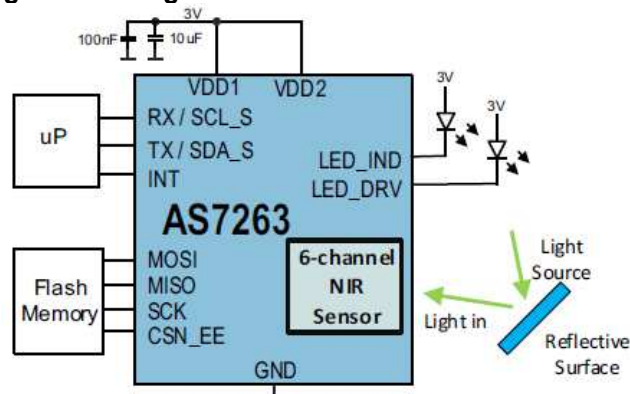
protocolo I<sup>2</sup>C ou UART<sup>2</sup>, proporcionando flexibilidade na integração com microcontroladores (AMS OSRAM, 2021).

Assim como o sensor AS7341, o AS7263 possui parâmetros de ajuste de sensibilidade, que são o tempo de integração e o ganho. O tempo de integração possui faixa de ajuste de 2,8 ms a 714ms. Esse tempo é ajustado pelo registrador INT\_T (0x05) de 8 bits, permitindo a inserção de valores entre 1 e 255. O tempo de integração é calculado pelo formula (11).

$$t_{int} = T_{INT} \times 2,8 \text{ ms} \quad (11)$$

Onde T<sub>INT</sub> é o valor de ajuste do registrador. Já o ganho é ajustado pelo parâmetro GAIN, no registrador Control\_Setup (0x04, bits D4 e D5). Ele permite ajustar valores de 1x, 3.7x, 16x e 64x (AMS OSRAM, 2021). Esse valor corresponde ao multiplicador de amplificação do sinal proveniente do sensor. Assim como no AS7341, é necessário fazer o correto ajuste desses parâmetros para evitar saturação dos ADCs e não ocasionar em sinais de baixa amplitude e ruidosos. No anexo A, também se encontram a lista de registradores do AS7263 e sua descrição. Além da comunicação serial, o sensor possui pinos de interrupção configurável INT, de led indicador de operação LED\_IND e pino de driver de led LED\_DRV para controlar iluminação auxiliar em aplicações de refletância. No Diagrama abaixo estão dispostos os blocos de funcionamento do sensor.

**Figura 13 - Diagrama de blocos funcionais do AS7263**



Fonte: ams OSRAM (2021)

<sup>2</sup> UART (*Universal Asynchronous Receiver/Transmitter*) é um protocolo de comunicação serial assíncrono, que transmite dados bit a bit por meio de linhas de transmissão (TX) e recepção (RX), sem a necessidade de um sinal de clock compartilhado. A sincronização entre transmissor e receptor é realizada a partir de bits de início e parada na transmissão (STALLINGS, 2017).

O módulo comercial utilizado no projeto é o GY-AS7263, sendo uma placa de prototipagem rápida que integra o sensor AS7263. Possui regulação de tensão auxiliar, permitindo alimentação de 3,3 a 5 V, o que facilita a conexão com os principais kits microcontroladores do mercado. O módulo incorpora um LED branco de alta intensidade para aplicações em refletância. No entanto não foi possível obter a especificação desse LED em termos de curva de emissão. A figura a seguir apresenta o módulo GY-AS7263 utilizado.

**Figura 14 - Módulo de prototipagem com o sensor AS7263**



Fonte: adaptado de anúncios comerciais (2025)

O módulo possui 6 pinos de conexão, descritos a seguir:

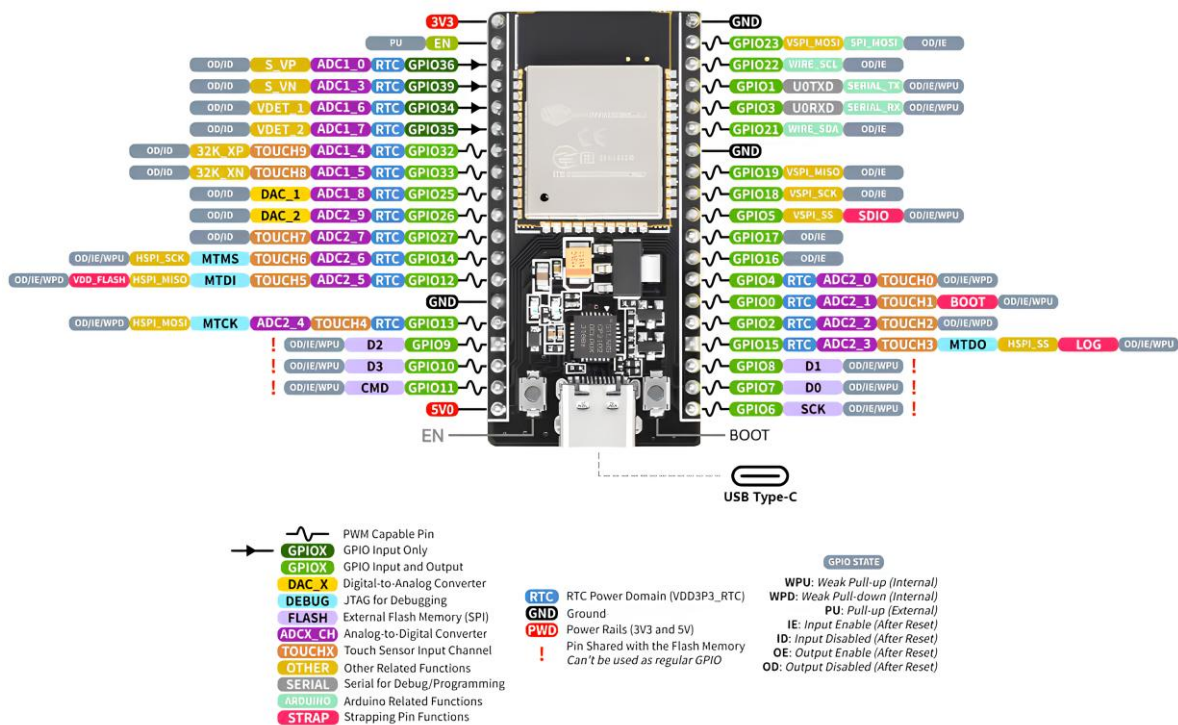
- VIN – Pino de alimentação (3,3 a 5 V).
- GND – Pino terra (referência).
- SCL / SCK – Pino de clock, compatível com o protocolo I<sup>2</sup>C (SCL) ou UART/SPI (SCK).
- SDA / TX – Pino de dados, compatível com o protocolo I<sup>2</sup>C (SDA) ou UART (TX).
- INT – Pino de interrupção configurável.
- RST – Pino de reset do sensor.

### 3.1.2 Microcontrolador ESP32

O microcontrolador escolhido para o projeto foi o ESP32, com a função de gerenciar todo o funcionamento do protótipo. Desenvolvido pela Espressif Systems, combina CPU dual core de 32 Bits, conectividade sem fio via *Bluetooth* e *Wi-Fi* Integradas, conectividade com periféricos via protocolos I<sup>2</sup>C, UART e SPI, além de conversores A/D de 12 bits e saídas de controle com PWM ajustável. O ESP32 opera em até 240 MHz, permitindo maior eficiência na execução de algoritmos mais

complexos de controle e análise em tempo real (ESPRESSIF ,2025). Em termos de memória, o ESP32 dispõe de 520 KBytes de memória RAM estática interna, já em no que se refere a memória Flash, o valor varia dependendo do kit utilizado, sendo mais comum encontra-los com 4 MBytes, como no kit NodeMCU-32S (ELETROGATE,2018). Na imagem a seguir, está disposta a relação de pinos do ESP32 no kit NodeMCU-32S, de 38 pinos, utilizado no protótipo.

Figura 15 - pinagem do nodeMCU-32S



Fonte: ESPBoards (2024)

A figura destaca os diferentes grupos de GPIOs do kit, como os pinos de alimentação (3,3 V a 5V), referência (GND) entradas analógicas com ADCs de 12 bits (ADC1 e ADC2), saídas analógicas baseadas em conversores digitais-analógicos de 8 bits (DAC1 e DAC2), pinos com suporte a PWM ajustável, pinos de comunicação serial (UART, SPI, I<sup>2</sup>C, além de pinos de toque capacitivo (ESPRESSIF, 2025). Essa versatilidade, com várias GPIOs com multifunção configurável, permite a utilização do módulo em aplicações que necessitam de diferentes formas de controle e comunicação, que é o caso do presente trabalho.

Comparando com outros kits disponíveis, como os Kits Arduino Uno e Mega, o ESP32 se sobressai na capacidade de hardware, mantendo uma faixa de preço

semelhante, além de ser um módulo mais compacto. A tabela 2 traz um comparativo entre esses microcontroladores:

Tabela 2 - Comparativo entre microcontroladores populares

Microcontrolador (chip)	Kit / placa	CPU / Clock	Flash / SRAM	GPIOs usuais
<b>ESP32</b>	ESP32 38 pinos	2× Tensilica LX6 @ 240 MHz	520 KB SRAM + flash externa (4MB a 8MB)	~25–32
<b>ATmega328P</b>	Arduino Uno R3	8-bit AVR @ 16 MHz	32 KB Flash / 2 KB SRAM	20
<b>ATmega2560</b>	Arduino Mega 2560	8-bit AVR @ 16 MHz	256 KB Flash / 8 KB SRAM	70

Fonte: Adaptado de Espressif (2025), Arduino (2025)

A maior capacidade de memória e a frequência de CPU permite que o ESP32 rode códigos e bibliotecas mais complexas, essenciais para controlar os sensores e a tela de interface IHM do projeto, o número e flexibilidade das I/Os possibilita uma integração com periféricos de maneira mais simples, facilitando o projeto da placa de circuito e excluindo a necessidade de módulos auxiliares, o que diminui o tamanho e o número de conexões do protótipo.

### 3.1.3 Display de controle

O Display de controle é a ponte que permite o usuário comandar a leitura e ajustar os parâmetros do protótipo. Dentre as opções disponíveis, optou-se pelo display LCD *touchscreen* TFT de 2,8 polegadas. Esse display apresenta cores configuráveis, conta com resolução de 240 x 320 pixels, comunicação via SPI<sup>3</sup>, função de *touchscreen* resistivo e slot para leitura e gravação de cartão *microSD* (LCDWIKI, 2025). A figura 16 apresenta o display.

<sup>3</sup> O protocolo SPI (Serial Peripheral Interface) é um padrão de comunicação serial síncrona que utiliza geralmente quatro linhas: MOSI (Master Out Slave In), MISO (Master In Slave Out), SCLK (Serial Clock) e CS/SS (Chip Select). Nesse protocolo, o dispositivo mestre controla o relógio e inicia as transmissões, enquanto os escravos respondem às requisições, permitindo alta velocidade de comunicação entre microcontroladores e periféricos (WOLF, 2016).

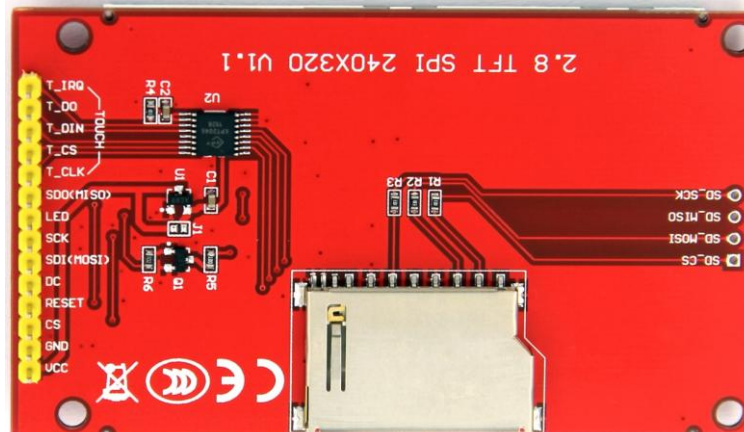
**Figura 16 - Display LCD touchscreen 2,8"**



Fonte: Adaptado de LCDWiki (2025)

Na figura 17 mostra a parte inferior do display, com a organização e descrição da pinagem, além do slot para cartão *microSD*.

**Figura 17 - Face inferior do Display com indicação de pinagem**



Fonte: LCDWiki (2025)

Cada um dos pinos e sua função são descritos a seguir:

- VCC – Alimentação (3,3 V a 5 V).
- GND – referência (*ground*).
- CS – *Chip select* comunicação SPI do display (ativo nível baixo).
- RESET – pino de reset do display.
- DC – Seleção entre modo de comando ou dados. Nível baixo para comandos, nível alto para dados.
- SDI (MOSI) – Sinal de dados SPI, de saída do Mestre (*Master Out Slave In*) para escrita no display.

- Sinal de *clock* SPI do display.
- LED – Controle de luz de fundo do display (0V a 3,3V).
- SDO (MISO) – Saída de dados SPI para o Mestre (*Master In Slave Out*), para leitura do display.
- T\_CLK – Sinal de *clock* SPI da função de toque.
- T\_CS – *Chip select* comunicação SPI da função de toque (ativo nível baixo).
- T\_DIN – Entrada de dados SPI da função de toque.
- T\_DO – Saída de dados SPI da função de toque.
- T\_IRQ – Sinal de interrupção da função de toque, nível baixo quando o toque é detectado.
- SD\_CS – *Chip select* comunicação SPI do módulo de cartão SD (ativo nível baixo).
- SD\_MOSI – Entrada de dados do módulo de cartão SD (*Master Out Slave In*).
- SD\_MISO – Saída de dados do módulo de cartão SD (*Master In Slave Out*).
- SD\_CLK – Sinal de *clock* SPI do módulo de cartão SD.

O display touchscreen apresenta a vantagem da comunicação serial SPI, reduzindo o número de conexões necessárias, comparado a uma tela de comunicação paralela. Além disso utilizar a função de toque na tela para implementar botões de controle do dispositivo, é vantajosa em relação a botões físicos, tanto em redução de hardware necessário, quanto na eliminação de ruídos, muito comuns em botões físicos.

#### 3.1.4 LEDs externos e controle de alimentação.

Além dos sensores com LEDs integrados, foram incluídos no projeto dois LEDs externos adicionais, com o objetivo de aumentar a emissão nos comprimentos de onda do vermelho visível e do infravermelho, que durante os testes foram faixas com pouca resposta espectral dos sensores, já que os leds embarcados, não possuem emissão significativa nessas faixas. O led vermelho difuso de 5 mm emite na faixa de 625 nm, tem tensão de operação de 2.1V, e teve seu circuito projetado

para trabalhar em 20 mA, corrente típica para esse LED (HAMAMATSU, 2025). O LED infravermelho de 5 mm emite na faixa de 940 nm, opera em 1.5 V, e a corrente típica de operação, utilizada no projeto, é de 20 mA (INOLUX CORPORATION, 2019). Ambos são acionados por um circuito *driver* simples de amplificação com transistores npn BC547, onde a corrente dos leds é ajustada de 0 a 20 mA através da interface IHM.

O protótipo inclui sistema de alimentação via baterias, com slots próprios soldados na placa, e com regulador de tensão. As duas baterias utilizadas são recarregáveis, de íon-lítio do tipo 18650, com capacidade nominal de 2600 mAh e tensão padrão de 3.7 V (chegando a 4.2 V com carga máxima) (EEMB, 2010). foram utilizadas em série com o objetivo de fornecer tensão suficiente para alimentar o microcontrolador, o circuito dos LEDs externos, e a tela LCD em 5V. Para análises futuras, foi incluído um sensor de temperatura e umidade AHT10, também com comunicação via protocolo I<sup>2</sup>C, além disso um interruptor liga/desliga foi adicionado para seccionamento da alimentação.

Para garantir que a tensão das baterias em série seja reduzida de maneira eficiente e estável para alimentar esses componentes, e protegê-los, um módulo conversor DC-DC *step-down* foi incluído no projeto. Dentre as opções comerciais disponíveis, foi escolhido o módulo baseado no circuito integrado XL4015.

Esse módulo permite controle independente de tensão constante e corrente constante. Ele opera com tensão de entrada de até 36V, com tensão de saída ajustável de 1,25V a 32V, e corrente máxima de saída de 5A. Pode ter eficiência de até 96%, garantido baixa dissipação de calor (HAREENDRAN, 2021). A figura 18 apresenta o módulo conversor.

**Figura 18 - módulo conversor DC-DC *step-down* XL4015**



**Fonte: Hareendran (2021)**

A corrente máxima de saída e a tensão de saída são ajustados pelos potenciômetros. A escolha do módulo foi baseada na sua eficiência, faixa de tensão correspondente às desejadas, e pelo tamanho compacto. Além disso o ajuste de corrente máxima de saída oferece uma proteção contra curto-circuito e sobrecarga no próprio módulo, protegendo os demais componentes.

### 3.2 Projeto da placa de circuito e invólucro plástico

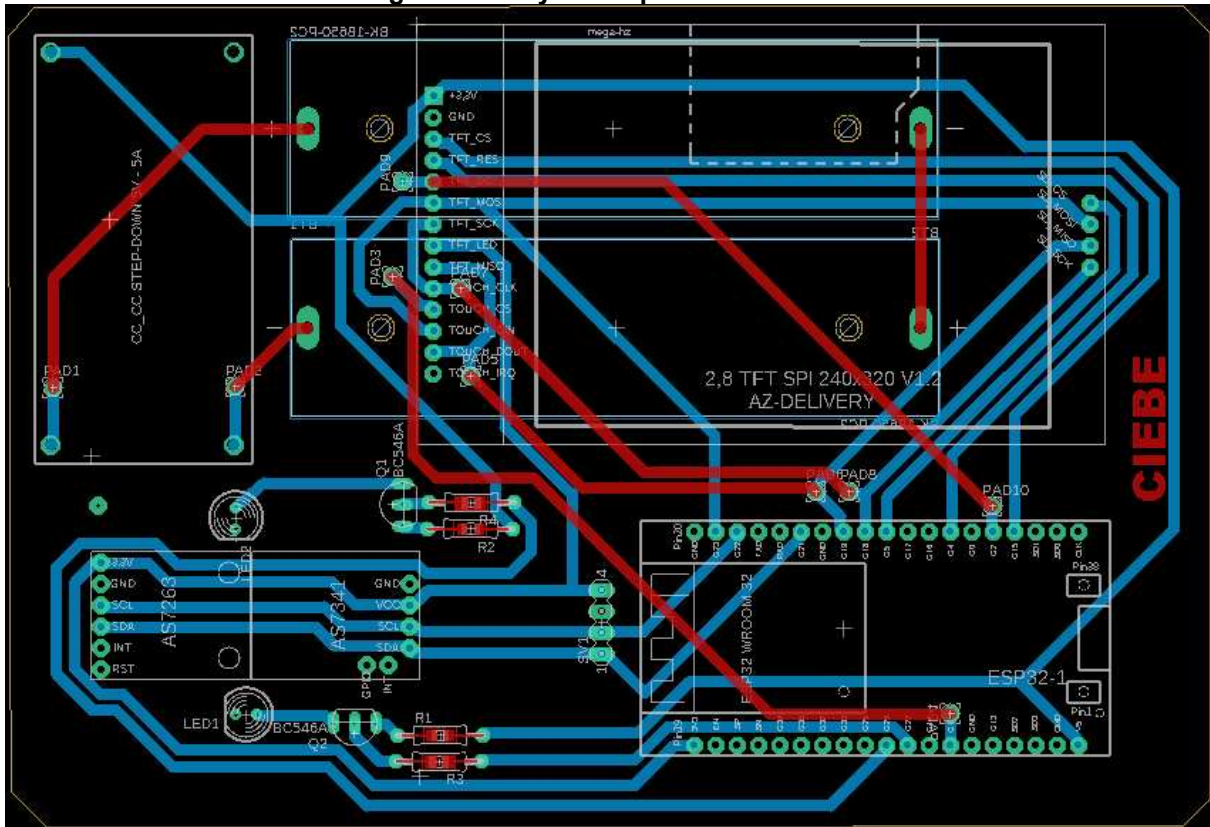
O projeto da placa de circuito impresso e do invólucro plástico é uma etapa crucial no desenvolvimento do protótipo, pois a disposição de componentes de maneira estratégica, e uma estrutura que proporcione medições estáveis e com o mínimo de interferência de luz externa, garantindo redução de interferências e maximizando a confiabilidade dos dados obtidos com o protótipo.

O projeto da placa de circuito impresso foi feito no software Autodesk Fusion, nele foram incluídas as bibliotecas dos componentes utilizados, e elaborados componentes personalizados para os sensores de acordo com as medidas dos módulos. Para elaborar a disposição dos componentes, foram seguidos dois principais critérios:

- **Proximidade dos sensores e Leds externos:** Os sensores espectrais AS7341 e AS7263 foram alinhados de maneira que ambos ficassem o mais próximo possível, permitindo a coleta de dados na mesma região do objeto alvo ao mesmo tempo. Também, os Leds vermelho e infravermelho foram dispostos na lateral dos sensores em lados opostos, angulados em direção à abertura de colocação de amostras, permitindo uma irradiação uniforme de ambos sobre o objeto alvo. Essa disposição permite uma leitura mais eficiente e completa.
- **Otimização de espaço:** os componentes foram dispostos de maneira a otimizar a área total da placa, mantendo assim, a proposta de ser um aparelho compacto.

Para cumprir esses requisitos, a placa foi projetada em dupla face, com a parte inferior abrigando os slots de bateria, e na parte superior os demais componentes. A figura 19 ilustra o layout final da placa, com os componentes utilizados e trilhas.

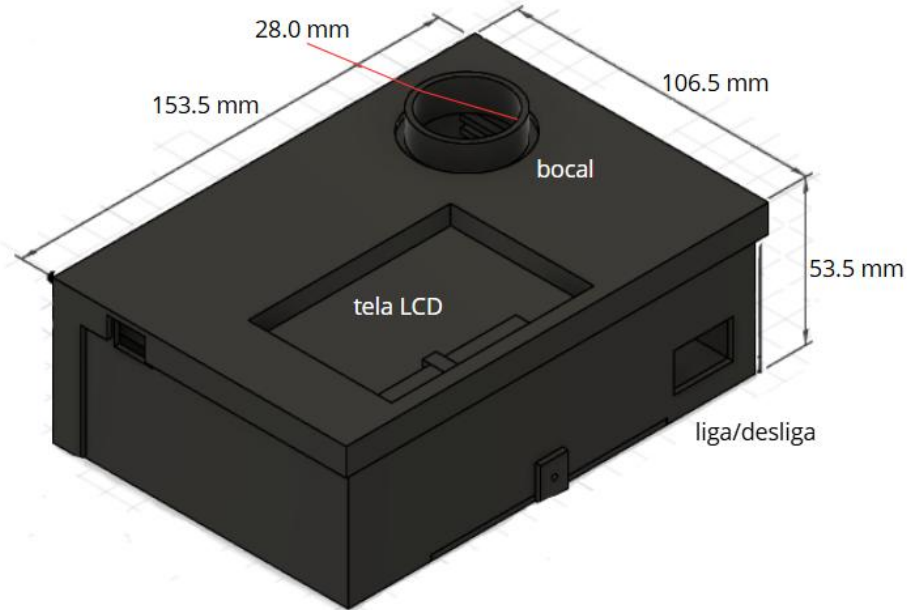
Figura 19 - Layout da placa de circuito



Fonte: Autoria própria (2025)

O invólucro plástico foi projetado em CAD, através do software Autodesk Fusion 360, e fabricado em impressão 3D na cor preta, para englobar todo o sistema, evitando o máximo possível de entrada de luz externa, que interfere na resposta dos sensores. A modelagem e impressão 3D permite alta customização, estrutura leve e robusta. O desenho foi pensado para comportar a placa com um encaixe tipo gaveta, facilitando a montagem. Além disso, foi incluída uma tampa removível na parte inferior, para troca das baterias de lítio, sem a necessidade de desmontar todo o conjunto. O Interruptor liga/desliga foi fixado na lateral da caixa, para fácil acesso. Uma abertura para a entrada de cartão microSD foi deixada na região do display LCD.

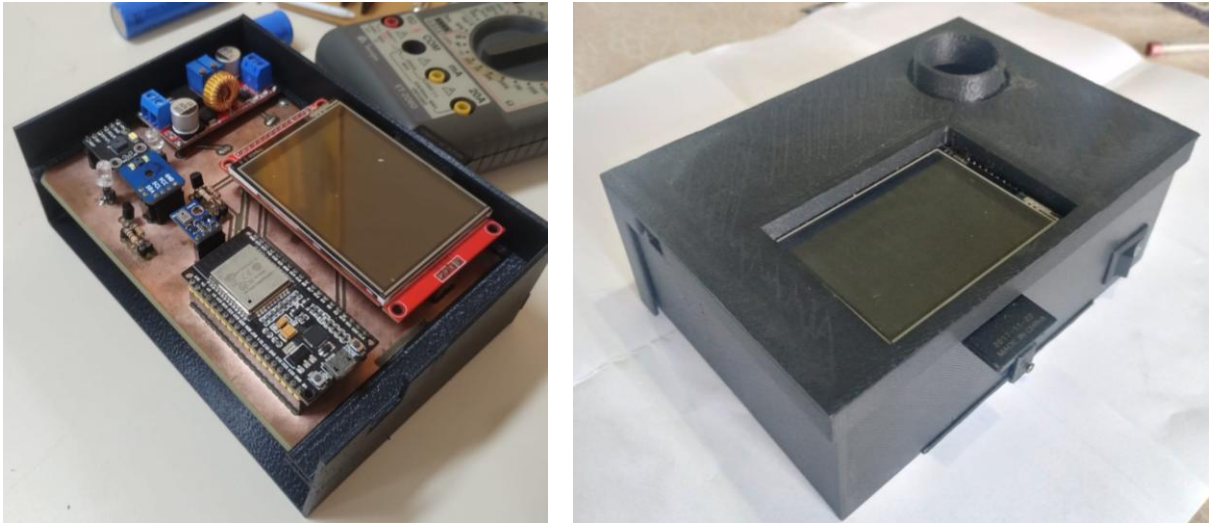
Na parte superior, foi desenhada uma tampa de fechamento com a inclusão da abertura para a tela LCD, e um bocal circular de 28 mm de diâmetro, com ressalto de 8 mm em relação a superfície da tampa, mantendo uma distância de 1,5 cm entre o aro externo do bocal e a superfície dos sensores. Essa distância visa proporcionar um espalhamento da luz dos leds sobre toda a superfície do objeto alvo, ao passo que o ressalto circular ajuda a manter o objeto fixo sobre o bocal, especialmente em objetos esféricos ou cilíndricos, como as frutas. A figura 20 ilustra o modelo 3D do invólucro.

**Figura 20 - Modelo 3D do invólucro plástico**

**Fonte: Autoria própria (2025)**

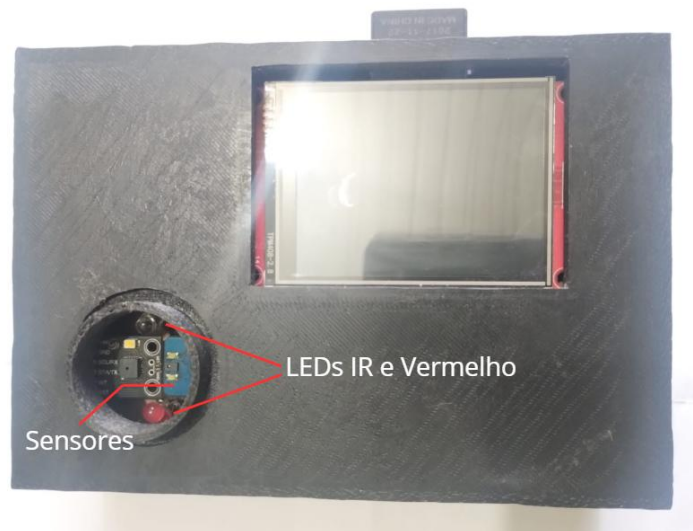
A distância entre o objeto alvo e os sensores é uma variável de controle de grande importância para o sistema, pois sua variação impacta diretamente no espalhamento da luz e, por sua vez, na intensidade detectada pelo sensor, assim mantê-la fixa é fundamental para leituras consistentes. No entanto, no caso das frutas, ainda há pequenas variações dessa distância devido à irregularidade da superfície e do formato, o que acaba afetando o sistema, mesmo que minimamente. A figura 21 mostra o protótipo finalizado, em três perspectivas diferentes, sendo: (a) vista interna das conexões, (b) vista em perspectiva do invólucro, onde pode-se visualizar em uma das laterais o botão de liga/desliga e a entrada para cartão de memórias, enquanto na face tem-se o display e a abertura para posicionamento das amostras a serem analisadas; e em (c) uma vista superior do equipamento, com detalhamento do posicionamento dos leds de iluminação e sensores.

Figura 21 – protótipo finalizado (a) vista interna, (b) vista externa em perspectiva; e (c) vista superior.



(a)

(b)



(c)

Fonte: Autoria própria (2025)

### 3.3 Projeto da interface IHM

A interface IHM é parte fundamental no controle da aquisição de dados e no ajuste dos parâmetros necessários. Ela foi pensada para fácil utilização e controle do dispositivo, sem a necessidade de conectá-lo a um computador. Cada tela, sua função e ajustes possíveis são descritos a seguir:

- **Tela de inicialização:** ao ser ligado o dispositivo exibe o logotipo do CIEBE, que basicamente é uma imagem armazenada no cartão *microSD*, que permite verificar se a comunicação com o módulo de

cartão de memória está funcionando corretamente. Além disso uma mensagem de inicialização também é exibida. As figuras 22 (a) e (b) mostram a imagem e a mensagem respectivamente:

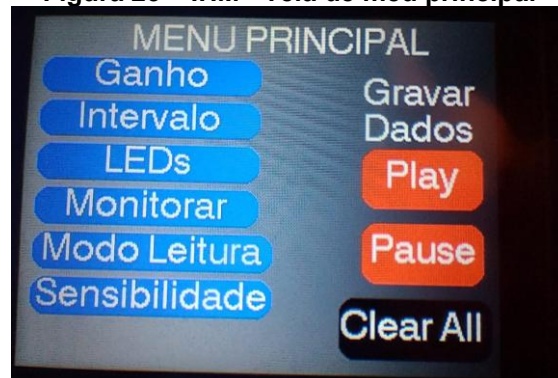
**Figura 22 - IHM (a) tela de abertura e (b) tela de inicialização.**



Fonte: Autoria própria (2025)

- **Menu principal:** após a inicialização, a tela de menu principal é exibida, sendo assim a tela central que permite acessar todas as funcionalidades do espectrofotômetro. A partir desse menu é possível acessar todas as demais telas de ajuste e operação. Nessa mesma tela está o controle da gravação de dados no cartão microSD. O botão *play* inicia a gravação, o botão *pause* interrompe a gravação, e o botão *clear all* apaga todos os dados do arquivo .csv. A figura 23 ilustra essa tela:

**Figura 23 – IHM - Tela de meu principal**



Fonte: Autoria própria (2025)

- **Monitoramento:** Esta é a tela que permite monitorar os dados obtidos. Somente ao acessá-la o dispositivo inicia as leituras, exibindo em tempo real os valores lidos pelos sensores em cada um de seus

respectivos canais espectrais, além dos dados de temperatura e umidade do AHT10. O botão Menu, permite o retorno ao menu principal. A tela possui duas versões uma para o modo de leitura contínuo e uma para o modo de leitura único. A única diferença entre ambas é o botão Amostrador, exibido no modo de leitura único, realiza uma leitura de dados quando apertado. No modo contínuo, o protótipo realiza leituras continuamente, no intervalo de tempo definido na tela de ajuste de intervalo. Já no modo de leitura único, uma leitura é feita somente quando o botão Amostrador é pressionado. A figura 24 exibe em (a) e (b) as telas de monitoramento modo contínuo e modo de leitura única, respectivamente:

**Figura 24 - Telas de monitoramento e suas versões, sendo (a) monitoramento em modo contínuo e (b) modo de leitura única**



Fonte: Autoria própria (2025)

- **Ganho:** permite o ajuste do ganho multiplicativo de ambos os sensores, de acordo com os valores pré-definidos pelas respectivas bibliotecas de controle, como visto na descrição de componentes. Essa configuração amplifica o sinal dos fotodiodos, aumentando a sensibilidade do sistema.
- **LEDs:** Nessa tela é possível ligar/desligar os LEDs integrados aos módulos dos sensores. Adicionalmente, permite controlar a intensidade dos LEDs externos auxiliares (vermelho e Infravermelho) por meio de uma barra de ajuste percentual individual.
- **Sensibilidade:** tela de ajuste dos tempos de integração dos sensores. Ela exibe o tempo de integração de cada sensor em milissegundos, e

os valores de cada registrador que controla essa grandeza dos sensores. O ajuste segue as fórmulas previamente descritas, onde os valores são ajustados por contadores incrementais e decrementais, acionados pelos botões “+” e “-“, acima de cada algarismo. Os contadores foram programados para respeitarem os valores mínimo e máximo de ajuste de cada registrador de tempo de integração.

- **Intervalo:** nessa tela é feito o ajuste do intervalo de tempo entre leituras para o modo de operação contínuo. O tempo é ajustado por meio de dois contadores incrementais/decrementais, um para a unidade de segundos e outro para décimos de segundos, permitindo ajustes de 0,0 a 9,9 segundos. Cabe ressaltar que esse tempo é contado através de um *timer* interno do ESP32, desse modo essa contagem é feita paralelamente ao funcionamento do sistema, sem impedir o funcionamento da rotina principal do microcontrolador. Outro ponto relevante é que o intervalo só é válido se for maior que o tempo total de integração somados de ambos os sensores, pois é o tempo mínimo necessário entre leituras. Se for menor vale o tempo de integração somado dos sensores.
- **Modo de leitura:** Nessa tela é escolhido o modo de leitura do protótipo, quando escolhido o modo contínuo, o dispositivo realizará leituras contínuas no intervalo de tempo definido na tela de intervalo. No modo único, o dispositivo realizará uma leitura toda vez que o botão amostrar da tela de monitoramento for acionado.

A figura 25 (a), (b), (c) e (d) e (e) traz as telas respectivas telas, descritas anteriormente:

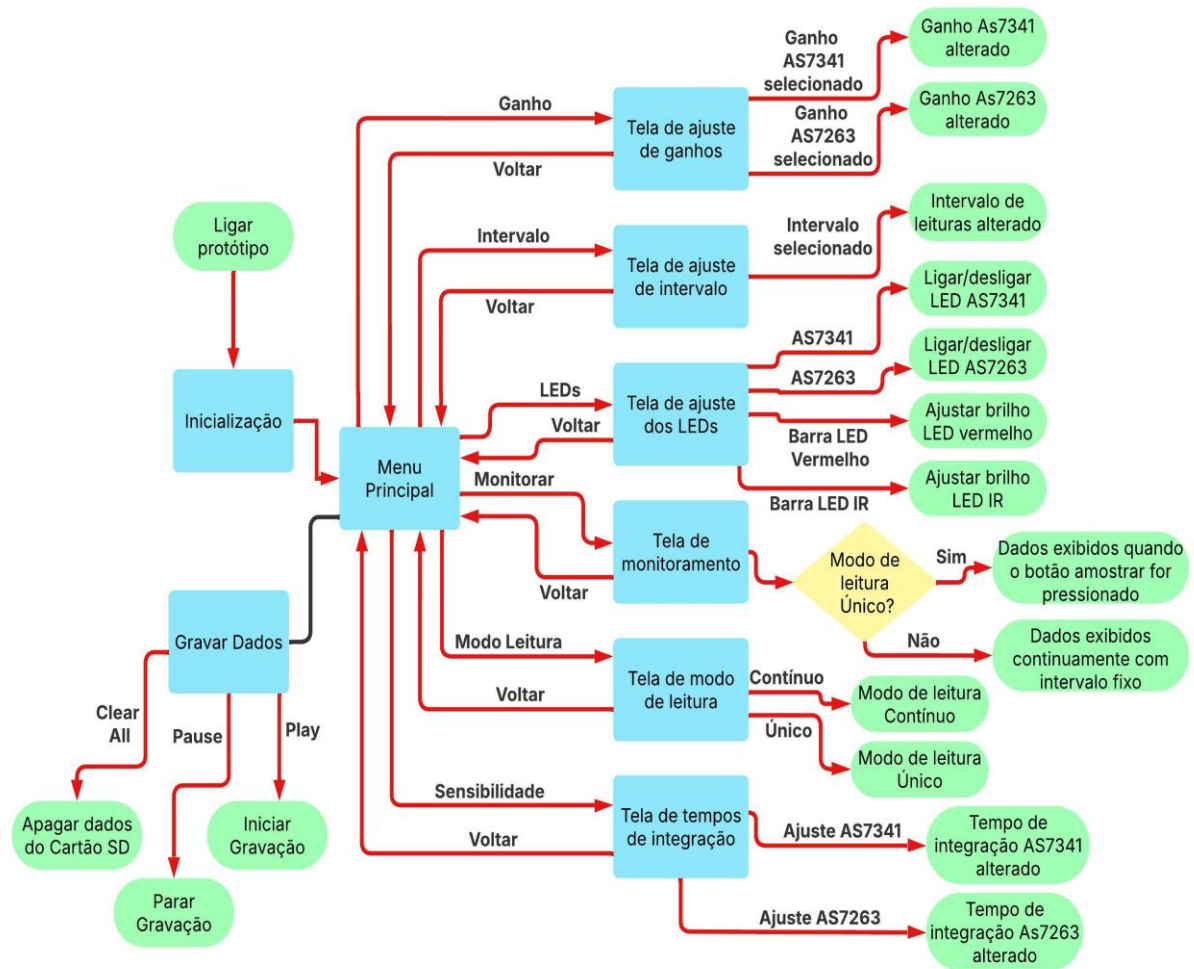
Figura 25 - Telas de parametrização, sendo (a) de ajuste de ganho, (b) de ajuste dos LEDs, (c) de ajuste das sensibilidades, (d) de ajuste dos intervalos de leitura e (e) de ajuste do modo de leitura



Fonte: Autoria própria (2025)

Todas as telas de ajuste contam com um botão voltar, que retorna diretamente para o meu principal. A seguir temos o fluxograma de funcionamento da IHM, indicando cada uma das opções e interconexão entre telas.

Figura 26 - Diagrama de funcionamento da interface IHM



Fonte: Autoria própria (2025)

Cada uma das opções tem um ajuste padrão utilizado para a inicialização, e toda vez que o dispositivo é reiniciado ele volta a configuração padrão. As opções padrão são citadas a seguir:

Tabela 3 - ajustes padrão de inicialização	
Ajuste	Opção padrão
<b>Ganho AS7341</b>	64x
<b>Ganho AS7263</b>	64x
<b>Modo de leitura</b>	Contínuo
<b>LEDs</b>	Todos ligados em 100%
<b>Intervalo</b>	1 segundo
<b>Tempo de integração AS7341</b>	118.9 ms (ASTE <sub>P</sub> = 309 e ATIME = 137)
<b>Tempo de Integração As7263</b>	714 ms (INT_T = 255)

Fonte: Autoria própria (2025)

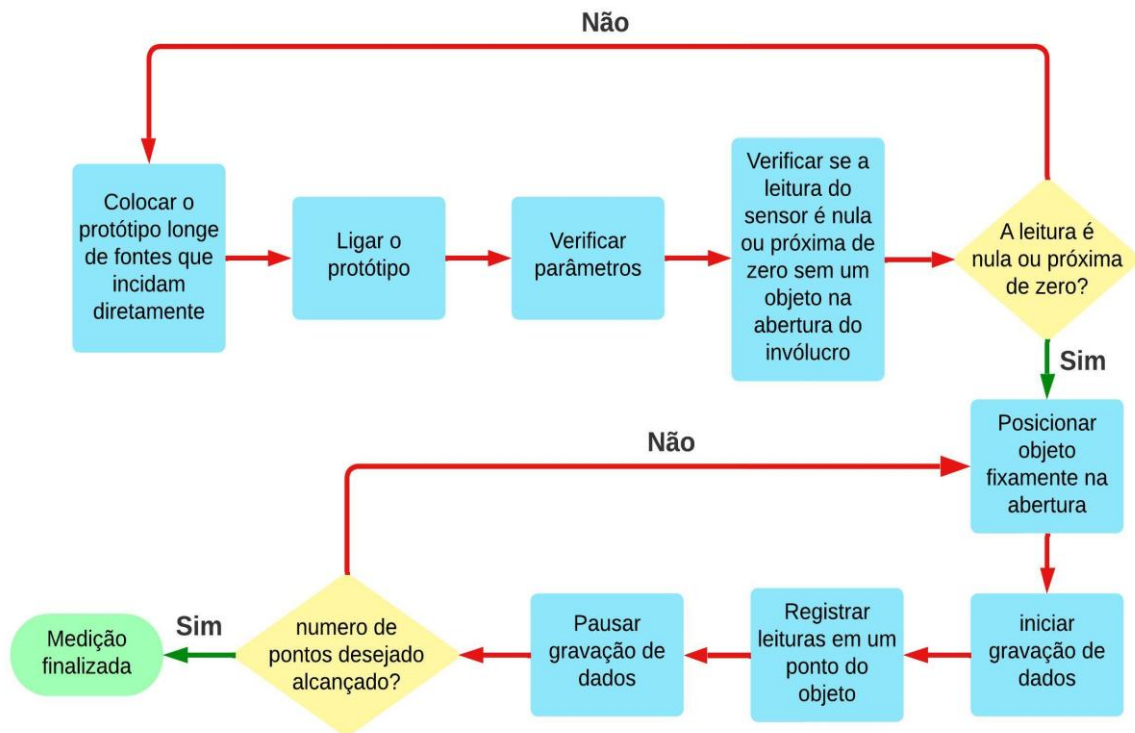
### 3.4 Metodologia experimental

Além do projeto de hardware, a confiabilidade dos dados também depende da padronização do procedimento de coleta de dados com o dispositivo. Para garantir a consistência dos dados, um procedimento experimental padrão de aquisição de dados foi estabelecido, passando pela configuração do protótipo até a coleta dos dados de cada amostra. O passo a passo utilizado é descrito a seguir através destas 8 etapas:

- 1. Ambiente com pouca iluminação externa:** garantir que o local em que o protótipo está posicionado receba pouca iluminação externa, que possa interferir nas medições.
- 2. Inicialização do sistema:** o dispositivo é ligado, a tela de inicialização inicia, e então é verificada se o logo do CIEBE é mostrado corretamente, confirmando o correto funcionamento do módulo *microSD*.
- 3. Verificação dos parâmetros:** são checados todos os parâmetros de ajuste, garantindo que todos estejam na mesma configuração preestabelecida para o experimento. Os parâmetros utilizados na análise do presente trabalho são exibidos no tópico de amostras.
- 4. Verificar interferência de luz externa:** tomar leituras sem objetos no bocal e verificar se as leituras são nulas ou próximas de zero, se não é necessário reposicionar o protótipo em local menos iluminado.
- 5. Posicionamento da amostra.** O objeto alvo é posicionado sobre o bocal fixamente, para que não se mova durante a medição.
- 6. Iniciar gravação de dados.** A gravação dos dados é acionada no menu principal para salvamento no cartão *microSD*.
- 7. Coleta de dados.** O operador navega até a tela de monitoramento, onde é iniciada a coleta de dados temporizada ou manualmente acionada. É coletado o número de amostras pré-determinado no experimento.
- 8. Finalização da coleta.** A gravação é interrompida pelo botão "Pause" no menu principal, e então a amostra é removida do bocal, finalizado o procedimento para aquele ponto de medição.

O protocolo é repetido para cada amostra e ponto de medição sobre a mesma amostra. O diagrama da figura 27 exemplifica o procedimento experimental de aquisição de dados, descrito nas 8 etapas descritas.

**Figura 27 - Fluxograma de procedimento de aquisição de dados**



**Fonte: Autoria própria (2025)**

Para todas as coletas do presente trabalho os seguintes parâmetros do protótipo foram seguidos:

**Tabela 4 - Ajustes utilizados no experimento do trabalho**

<b>Parâmetro</b>	<b>Valor Utilizado</b>	<b>Justificativa</b>
Ganho AS7341	64x	Garantir sensibilidade suficiente e evitar saturação.
Ganho AS7263	64x	Garantir sensibilidade suficiente e evitar saturação.
Tempo de Integração AS7341	118.9 ms	Garantir sensibilidade suficiente e evitar saturação.
Tempo de Integração AS7263	714 ms	Máximo tempo, potencializando a sensibilidade na faixa NIR, já que não ocorre saturação nesse ajuste
LEDs	Todos ligados em 100%	Proporcionar irradiação suficiente em todas as faixas de espectro analisadas
Modo de leitura	Único	Possibilitar maior controle do processo de registro de dados

**Fonte: Autoria própria (2025)**

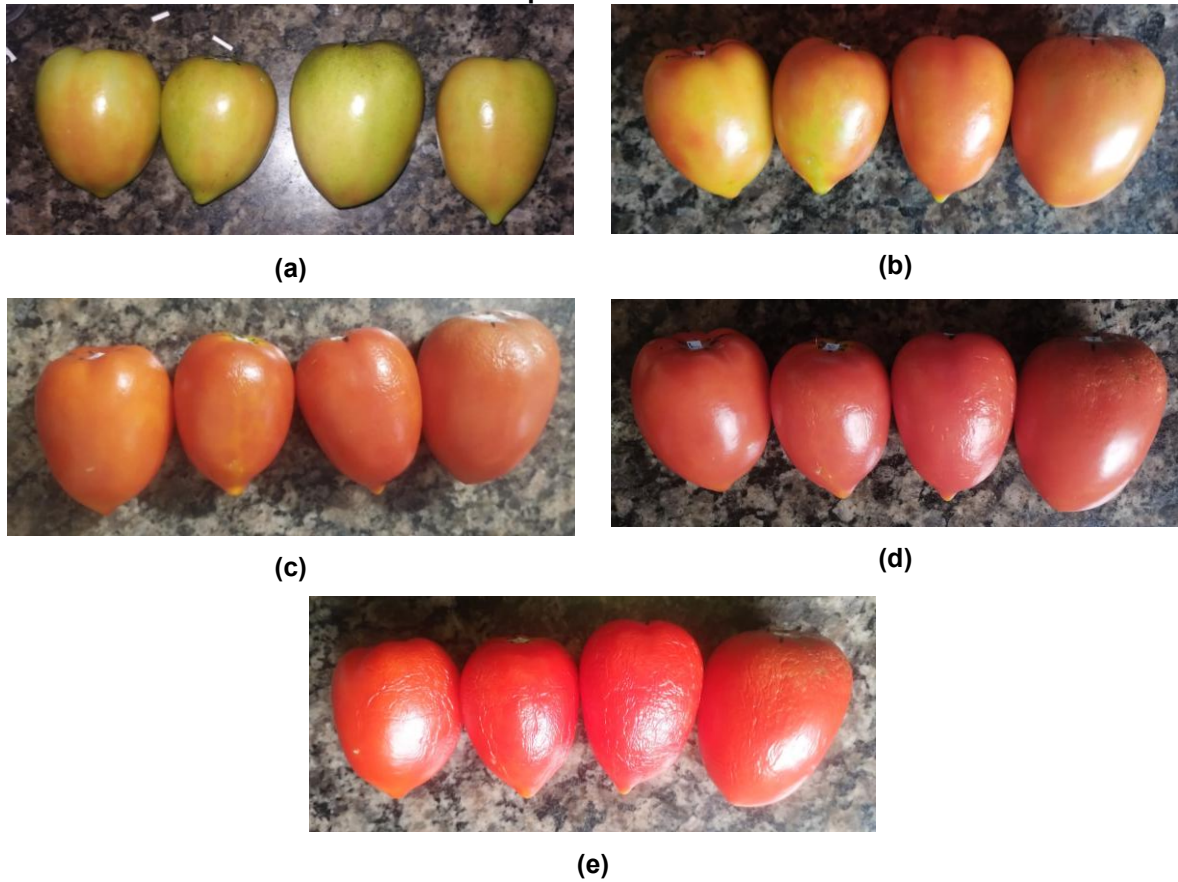
### 3.5 Amostras

Para gerar o banco de dados necessário para futura análise, foram selecionadas duas frutas com processo de maturação diferentes, e de grande popularidade comercial: a banana e o tomate.

Os tomates são do tipo italiano, e as bananas da variedade caturra, todos obtidos de cultivo próprio, de pequena escala e não comercial. Todas as frutas foram colhidas da mesma penca em estágios iniciais de maturação semelhantes. Após a colheita as amostras foram armazenadas em temperatura ambiente, em local arejado e com exposição a iluminação ambiente, protegida do sol. Esses cuidados foram tomados com o intuito de manter uma maturação homogênea entre os frutos. Todos os exemplares foram enumerados para um maior controle no processo de aquisição.

A figura 28 mostra a evolução da maturação dos tomates com intervalos de 5 dias, ou seja: para o primeiro dia de análise (a), para o quinto dia (b), para o décimo dia (c), para o décimo quinto dia (d) e para o vigésimo dia (e).

**Figura 28 - Processo de maturação dos tomates em 5 etapas, considerando seu amadurecimento pós colheita de 5 em 5 dias.**



**Fonte: Autoria própria (2025)**

A figura 29 mostra a evolução do estágio de maturação das bananas, considerando intervalo de 6 dias entre medidas, ou seja: no primeiro (a), sexto (b), décimo segundo (c) e décimo oitavo dia (d).

**Figura 29 - Processo de maturação das bananas em 4 etapas, considerado um intervalo de 6 dias entre cada.**





Fonte: Autoria própria (2025)

Para reduzir a influência de variáveis externas e padronizar as medições todas as coletas de dados foram realizadas as 16h, todos os dias.

No caso dos tomates, as leituras foram distribuídas em quatro quadrantes ao logo da circunferência central do fruto, opostos entre si. Em cada quadrante, foram coletados dados de cinco pontos distintos, totalizando 20 leituras por dia. Considerando os quatro exemplares, foram 80 leituras diárias, durante 20 dias consecutivos.

Já no caso das bananas, o procedimento foi um pouco modificado, devido a superfície com maior irregularidade. As medições foram realizadas em seis regiões distintas, nas duas faces não curvadas da fruta: uma na região central, outra entre o centro e a ponta superior, e outra entre o centro e a ponta inferior. Também foram coletados cinco pontos por região, totalizando 30 leituras diárias por banana, totalizando 120 leituras diárias dos quatro exemplares, ao longo de 18 dias consecutivos. A figura 30 demonstra exemplares de tomate (a) e banana (b) posicionados sobre o dispositivo para o registro de leituras:

Figura 30 - Exemplares posicionados para a amostragem



Fonte: Autoria própria (2025)

A figura 31 (a) e (b) demonstra as regiões de coleta de dados descritas anteriormente:



**Fonte: Autoria própria (2025)**

Para a coleta dos cinco pontos por região, o protótipo foi acionado pelo modo de “Leitura Única”. Assim que posicionada a fruta sobre o bocal, uma única leitura de todos os canais de ambos os sensores foi registrada e armazenada para aquele ponto, repetindo-se o procedimento para todos os pontos. Este procedimento foi definido em leitura única devido à alta repetibilidade das leituras dos sensores, ocasionada pela padronização de geometria de amostragem (distância e iluminação), imposta pelo modelo construtivo do dispositivo. Essa repetibilidade foi notada em testes preliminares, onde foram realizados testes, coletando uma série de leituras de um único ponto, e então verificada a variabilidade dos dados, onde notou-se uma constância alta, com coeficiente de variação, entre média e desvio padrão, abaixo de 0,5%. As tabelas 5 e 6 exemplificam essa repetibilidade e constância de medições em um mesmo ponto, testada em um exemplar de tomate:

**Tabela 5 - Exemplificação da repetibilidade de dados para o sensor AS7341**

leituras	415nm	445nm	480nm	515nm	555nm	590nm	630nm	680nm	910nm
1	922	1540	1367	1557	1905	3929	11084	3518	9751
2	921	1539	1366	1556	1902	3920	11065	3511	9739
3	923	1541	1368	1558	1900	3916	11049	3505	9731
4	922	1541	1367	1557	1905	3926	11075	3516	9739
5	922	1541	1368	1559	1907	3930	11087	3521	9744
6	923	1542	1368	1559	1906	3924	11073	3515	9740
7	923	1542	1369	1560	1908	3928	11086	3520	9746
8	923	1542	1368	1559	1907	3927	11080	3518	9745
9	923	1542	1368	1560	1908	3928	11084	3521	9744
10	922	1541	1367	1558	1906	3926	11079	3519	9742
MÉDIA	922,4	1541,1	1367,6	1558,3	1905,4	3925,4	11076,2	3516,4	9742,1
DES. PADRÃO	0,663	0,943	0,800	1,269	2,458	4,128	11,107	4,779	5,069
DP/MED (%)	0,07%	0,06%	0,06%	0,08%	0,13%	0,11%	0,10%	0,14%	0,05%

**Fonte: Autoria própria (2025)**

**Tabela 6 - Exemplificação da repetibilidade de dados para o sensor AS7263**

leituras	610nm	680nm	730nm	760nm	810nm	860nm
1	9534	3553	2527	1764	9224	9104
2	9522	3544	2528	1763	9213	9103
3	9510	3539	2525	1763	9211	9103
4	9555	3561	2539	1768	9219	9117
5	9541	3561	2535	1768	9216	9118
6	9544	3560	2537	1768	9216	9119
7	9540	3559	2537	1768	9210	9120
8	9519	3557	2531	1766	9209	9117
9	9534	3558	2536	1766	9206	9122
10	9516	3557	2534	1766	9200	9123
MÉDIA	9531,5	3554,9	2532,9	1766	9212,4	9114,6
DES. PADRÃO	13,537	7,148	4,592	1,949	6,468	7,605
DP/MED (%)	0,14%	0,20%	0,18%	0,11%	0,07%	0,08%

**Fonte: Autoria própria (2025)**

Através da análise da tabela evidencia-se a repetibilidade dos dados para um mesmo ponto, para as 10 amostras feitas nos exemplos trazidos, o coeficiente de variação máximo foi de 0,14% para o sensor VIS e 0,20% para o sensor NIR. Essa evidência de estabilidade comprova que uma única leitura por ponto utilizada no procedimento de aquisição de dados é suficiente para a análise, além de otimizar o processo e o tempo total de aquisição de dados do experimento.

### 3.6 Processamento de dados

Após as coletas, os dados foram armazenados em dois arquivos .csv, estruturados como um banco de dados, exibindo os dados de todos os canais espectrais de ambos os sensores, organizados por dia de coleta e exemplar. Devido a diferente responsividade entre canais, explicada anteriormente, foi realizada a normalização de dados antes de análise, utilizando a padronização estatística *Z-score* dada pela fórmula:

$$z = \frac{(x - \mu)}{\sigma} \quad (12)$$

Em que  $z$  é o valor normalizado,  $x$  é o dado original,  $\mu$  é a média dos valores do canal do sensor, e  $\sigma$  é o desvio padrão do mesmo canal. Esse método transforma os dados de cada canal para que sua média seja zero e para que o desvio padrão seja um. Essa normalização impede que os dados de maior amplitude dominem o treinamento dos modelos, garantindo que todos os dados contribuam equilibradamente para a predição (HAYKIN, 1999).

Para a análise dos dados pré-processados e construção de modelos preditivos, foram escolhidas as três técnicas de regressão explicadas anteriormente: A Regressão Linear Múltipla (MLR), como modelo base de referência, a Regressão por Mínimos Quadrados Parciais (PLSR) como modelo intermediário, mais robusto e muito empregado em dados espectrais, e a rede neural artificial *Multilayer Perceptron* (MLP), como método mais complexo e sofisticado, visando melhores resultados, pela sua capacidade de aprender relações não lineares complexas entre dados espectrais e dia de maturação, que os demais métodos não conseguem modelar .

A análise de dados foi executada utilizando códigos em *Python*, dentro do ambiente de desenvolvimento em nuvem *Google Colaboratory*. Esse ambiente permite estruturar o código em *Jupyter Notebooks*, é gratuito e permite o armazenamento de códigos e arquivos em nuvem via *Google Drive*. No apêndice A estão os códigos desenvolvidos para cada um dos métodos.

As bibliotecas utilizadas foram:

- ***Pandas e NumPy***: para operações com conjuntos de dados;

- **Scikit-learn**: para implementação das regressões pelos métodos MLR e PLSR, divisão de dados e cálculo de desempenho dos modelos;
- **TensorFlow** com API **Keras**: para a implementação da rede MLP;
- **Matplotlib**: plotagem de gráficos.

Para garantir uma avaliação imparcial dos modelos, a divisão de dados foi realizada através da função **train\_test\_split** da biblioteca **Scikit-learn**. Essa função aplica um embaralhamento aleatório (**shuffle=True**) antes de realizar a divisão. Esse embaralhamento impede que uma divisão sequencial seja realizada, o que resultaria em resultados inconsistentes para os dados ordenados por tempo, presentes nos arquivos. Ao embaralhar os dados, é garantida que as amostras de cada subdivisão sejam estatisticamente representativas dentro do total de dados, reunindo amostras de cada estágio de maturação para os diferentes exemplares.

Adicionalmente, foi definido o parâmetro **random\_state=42** em todas as operações de divisão de conjuntos de dados, tanto para o código da MLP, quanto para os códigos da MLR e da PLSR. Esse ajuste, mesmo que o embaralhamento de dados seja aleatório em termos de mistura de dados, faz com que o embaralhamento seja exatamente o mesmo toda vez que o código é executado. Esse cuidado garante a reprodutibilidade dos resultados, e uma comparação entre modelos consistente.

Para a MLP, os dados foram divididos em 70% para treino, 15% validação e 15% para teste, foram utilizadas 300 épocas de iteração. A função de ativação utilizada para as camadas ocultas foi a ReLU (*Rectified Linear Unit*), já na camada de saída foi aplicada a função linear, já que se trata de uma aplicação de regressão. Para o treinamento, foi utilizado o otimizador **Adam**, mantendo a taxa de aprendizado padrão de 0,001. A métrica de otimização alvo (função de custo) utilizada foi o erro quadrático médio (*mean squared error* -MSE), adicionalmente foram calculados o Coeficiente de Determinação ( $R^2$ ), que indica o quão bem o modelo se ajusta aos dados, e a Raiz do Erro Quadrático Médio (RMSE), que expressa o erro de predição na mesma unidade da variável resposta (dias) (MONTGOMERY et al., 2012), para comparação direta com os modelos MLR e PLSR. Para o modelo foram testadas diferentes configurações de camadas de neurônios, detalhadas no capítulo 4, buscando uma arquitetura com melhor desempenho para cada conjunto de dados.

Para os métodos de regressão (MLR e PLSR), foi mantida a divisão de dados, para uma comparação consistente entre os métodos. Também foram utilizadas as

métricas de desempenho MSE, RMSE e  $R^2$ , como comparação com a MLP. Para a PLSR foram testados diferentes números de componentes latentes e feita análise de *VIP Scores*, identificando os comprimentos de onda mais influentes no modelo de predição. Os resultados, gráficos, e análise comparativa estão dispostos no próximo Capítulo.

## 4 RESULTADOS E DISCUSSÃO

No presente capítulo serão discutidos os resultados obtidos utilizando os métodos de regressão estatística (MLR e PLSR), comparando com a regressão feita pela rede neural MLP. Além disso uma análise comparativa entre modelos resultados usando os dados de ambos os sensores e de cada um individualmente foi realizada, a fim de demonstrar a complementariedade entre a análise das faixas visíveis do espectro e do infravermelho próximo.

### 4.1 Base de dados

A armazenagem de dados seguiu o procedimento explicado anteriormente, resultando em um arquivo .csv individual e diário para cada exemplar analisado. Ao final do experimento os dados dos arquivos foram reunidos e organizados em um único arquivo .csv, resultando em um arquivo com todos os dados dos tomates e um arquivo com todos os dados das bananas.

O protocolo de coleta, portanto, resultou em um banco de dados vasto e robusto para a análise. Para os tomates, foram 80 leituras diárias, sendo 4 exemplares com 5 leituras por região em 4 regiões distintas, durante 20 dias, totalizando 1600 leituras. Para as bananas foram 120 leituras diárias, sendo 4 exemplares com 5 leituras por região em 6 regiões distintas, durante 18 dias, totalizando 2160 leituras.

Os dados foram organizados contendo uma coluna com o dia de maturação, uma com a identificação do exemplar, e 15 colunas com os dados espectrais de todos os canais. A figura 32 exemplifica essa organização:

**Figura 32- Organização dos dados no arquivo .csv utilizado nas análises**

Exemplar	Dia	415nm	445nm	480nm	515nm	555nm	590nm	630nm	680nm	NIR(910nm)	610nm	680nm_1	730nm	760nm	810nm	860nm
1	1	909	1682	1435	2199	2975	3336	5137	2353	9748	4422	1724	3193	2526	10654	9099
1	1	843	1550	1334	1994	2651	2963	4649	2091	8453	3902	1449	2773	2322	11246	9540
1	1	857	1559	1345	1988	2646	3006	4804	2117	8565	4290	1644	2818	2369	11138	9797
1	1	909	1643	1407	2021	2741	3192	4965	2311	10524	5800	1966	3162	2445	10588	9502
1	1	853	1569	1348	2041	2735	3076	4912	2101	8462	3952	1501	2742	2318	10834	9496
1	1	837	1533	1313	1940	2560	2865	4485	1978	9342	4546	1212	2551	2107	10155	8797
1	1	804	1474	1267	1911	2490	2761	4141	1855	8355	3546	970	2281	1934	9616	8143
1	1	801	1443	1246	1804	2410	2710	4257	1884	9383	4910	1250	2612	2159	10726	9255
1	1	834	1521	1306	1923	2551	2856	4461	1938	9502	4655	1168	2521	2078	10186	8900
1	1	823	1492	1285	1875	2493	2806	4374	1900	9442	4726	1182	2517	2096	10466	9125
1	1	742	1338	1155	1628	2048	2254	3160	1732	7999	3013	959	2174	1774	7866	6439
1	1	747	1364	1173	1668	2103	2322	3250	1710	7743	2950	930	2148	1803	8914	6781

Fonte: Autoria própria (2025)

## 4.2 Métricas e resultados para a *Multilayer Perceptron* (MLP)

Para a MLP foram definidas métricas para a escolha de número de neurônios na camada oculta, baseadas em fórmulas muito presentes na literatura. Para definir o número mínimo de neurônios optou-se pela métrica da metade da soma entre número de entradas e número de saídas, onde as entradas são as faixas espectrais dos sensores, e a saída é o dia de maturação. Como limite superior a métrica de escolha de número de neurônios foi a de o dobro de número de entradas mais um. Definidos esses dois limites, foram então testados valores intermediários de número de neurônios na camada oculta buscando pelo modelo mais otimizado dentro desse intervalo. A tabela 7 mostra o número de neurônios testados considerando a distinção de sensores, e o limite inferior e superior de número de neurônios calculado.

**Tabela 7 - Demonstrativo de número de neurônios utilizados na camada oculta da MLP**

<b>Caso</b>	<b>Entradas (na)</b>	<b>Saída (nc)</b>	<b>Limite Inferior ((na+nc)/2)</b>	<b>Limite Superior (2×na + 1)</b>	<b>Neurônios a testar (1 camada)</b>
Tomates – sensor VIS	9	1	5	19	5,10,15,20,30
Tomates – sensor NIR	6	1	4	13	5,10,15,20,30
Tomates – ambos os sensores	15	1	8	31	5,10,15,20,30
Bananas - sensor VIS	9	1	5	19	5,10,15,20,30
Bananas -sensor NIR	6	1	4	13	5,10,15,20,30
Bananas – Ambos os sensores	15	1	8	31	5,10,15,20,30

**Fonte: Autoria própria (2025)**

A partir dos limites foram definidos os cenários para testes (5, 10, 15, 20 e 30 neurônios na camada oculta para cada cenário). Em alguns casos foram extrapolados os limites, para fim de comparação entre modelos.

Os resultados para cada configuração de neurônios foram divididos e três tabelas para cada fruta: uma com os resultados somente para o sensor AS7341, outra somente para o AS7263, e a última com a análise conjunta dos dados de ambos os sensores. A tabela 8 traz os resultados obtidos com o sensor AS7341 para os tomates, com diferentes números de neurônios na camada oculta, bem como o RMSE e o coeficiente de determinação  $R^2$ , para o estágio de teste e de validação:

**Tabela 8 - Resultados MLP para os tomates com dados do AS7341**

Neurônios na camada	RMSE Validação	R <sup>2</sup> Validação	RMSE Teste	R <sup>2</sup> Teste
oculta	(dias)		(dias)	
5	1,568	0,931	1,661	0,912
10	1,517	0,936	1,595	0,918
15	1,434	0,943	1,552	0,923
20	1,390	0,946	1,459	0,932
30	1,438	0,942	1,530	0,925

Fonte: Autoria própria (2025)

A tabela 9 apresenta os resultados da análise dos tomates com a MLP somente com os dados do AS7263:

**Tabela 9 - Resultados MLP para os tomates com dados do AS7263**

Neurônios na camada	RMSE Validação	R <sup>2</sup> Validação	RMSE Teste	R <sup>2</sup> Teste
oculta	(dias)		(dias)	
5	1,604	0,928	1,869	0,888
10	1,509	0,936	1,761	0,901
15	1,480	0,939	1,698	0,908
20	1,489	0,938	1,735	0,903
30	1,476	0,939	1,692	0,908

Fonte: Autoria própria (2025)

Já a tabela 10 exibe os resultados para da análise de dados dos tomates com os dados de ambos os sensores:

**Tabela 10 - Resultados MLP para os tomates com dados de ambos os sensores**

Neurônios na camada	RMSE Validação	R <sup>2</sup> Validação	RMSE Teste	R <sup>2</sup> Teste
oculta	(dias)		(dias)	
5	1,249	0,956	1,417	0,936
10	1,149	0,963	1,368	0,940
15	1,148	0,963	1,300	0,946
20	1,113	0,965	1,238	0,951
30	1,074	0,968	1,193	0,954

Fonte: Autoria própria (2025)

A seguir são apresentados os resultados obtidos com a MLP utilizando os dados das bananas. A tabela 11 apresenta os resultados obtidos com os dados do sensor AS7341.

**Tabela 11 - Resultados MLP para as bananas com dados do AS7341**

Neurônios na camada oculta	RMSE Validação (dias)	R <sup>2</sup> Validação	RMSE Teste (dias)	R <sup>2</sup> Teste
5	2,078	0,844	2,059	0,839
10	1,903	0,869	1,915	0,861
15	1,876	0,873	1,936	0,858
20	1,790	0,884	1,831	0,873
30	1,785	0,885	1,836	0,872

Fonte: Autoria própria (2025)

A tabela 12 traz os resultados da MLP para as bananas, utilizando somente os dados do sensor AS7263:

**Tabela 12 - Resultados MLP para as bananas com dados do AS7263**

Neurônios na camada oculta	RMSE Validação (dias)	R <sup>2</sup> Validação	RMSE Teste (dias)	R <sup>2</sup> Teste
5	1,983	0,858	1,921	0,860
10	1,842	0,878	1,806	0,876
15	1,809	0,882	1,795	0,878
20	1,832	0,879	1,775	0,880
30	1,825	0,880	1,763	0,882

Fonte: Autoria própria (2025)

Por fim, tabela 13 traz os resultados da MLP para as bananas, utilizando os dados de ambos os sensores:

**Tabela 13 - Resultados MLP para as bananas com dados de ambos os sensores**

Neurônios na camada oculta	RMSE Validação (dias)	R <sup>2</sup> Validação	RMSE Teste (dias)	R <sup>2</sup> Teste
5	1,658	0,901	1,619	0,900
10	1,475	0,921	1,482	0,917
15	1,382	0,931	1,437	0,922
20	1,336	0,936	1,349	0,931
30	1,274	0,941	1,302	0,936

Fonte: Autoria própria (2025)

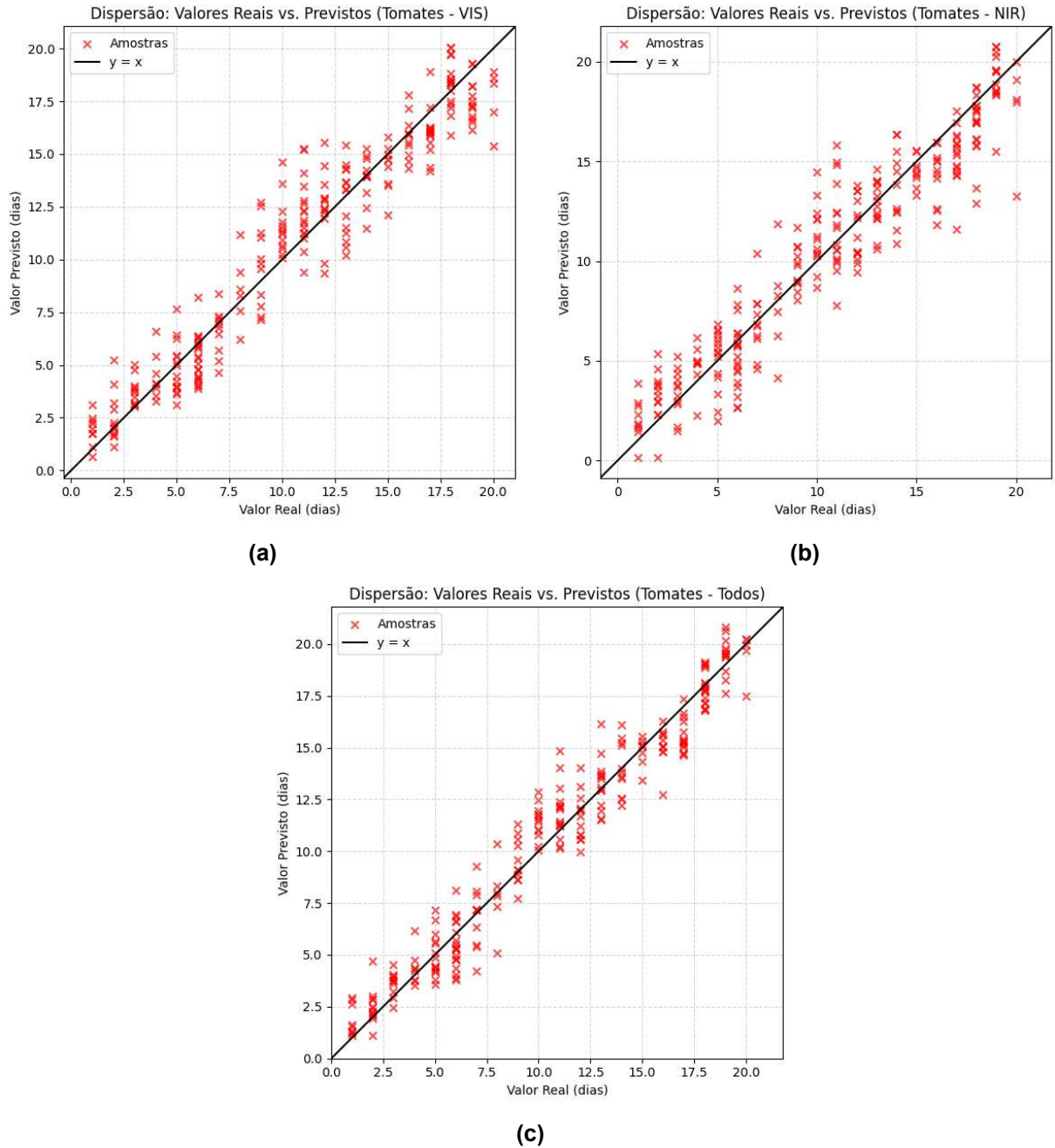
Ao analisar a performance dos modelos, observa-se que em geral os modelos com 30 neurônios acabaram tendo resultados mais satisfatórios na redução do RMSE e com R<sup>2</sup> mais próximo a 1. Na comparação entre resultados de validação e teste

houve pouca discrepância, demonstrando que o ajuste dos modelos foi satisfatório, não ocorrendo *underfitting*, nem *overfitting*.

A principal conclusão, no entanto, é que quando se compara os resultados dos modelos que usam os dados de cada sensor individualmente, com os modelos onde os dados de ambos os sensores são utilizados, percebe-se uma diferença de desempenho considerável. Em todos os cenários a fusão de dados dos sensores VIS e NIR resultou em uma melhoria de desempenho substancial.

Para os tomates, o modelo de dados combinados (30 neurônios), alcançou um  $R^2$  de 0,953, contra 0,932 do modelo do sensor VIS (20 neurônios) e 0,908 do modelo do sensor NIR (30 neurônios). Para as bananas algo semelhante ocorreu: o  $R^2$  subiu de 0,873 (VIS, 20 neurônios) e 0,882 (NIR, 30 neurônios) para 0,936 ao utilizar os dados combinados (30 neurônios). A figura 33 traz um comparativo gráfico entre dia de maturação real x dia de maturação previsto dos tomates para os melhores três modelos para cada conjunto de dados:

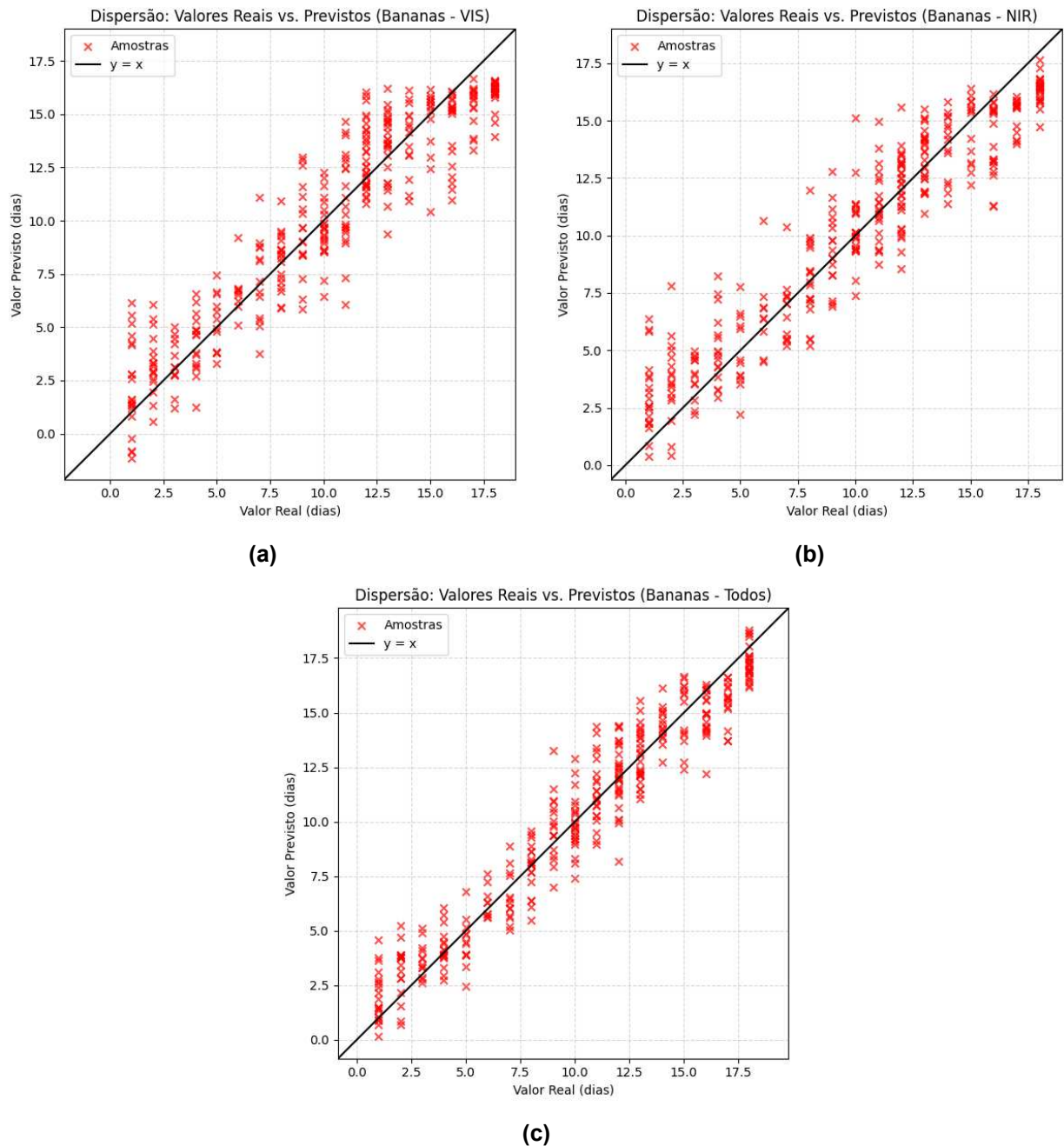
**Figura 33 - Gráficos de dispersão real x previsto MLP para os tomates, sendo (a) para o sensor VIS, (b) para o sensor NIR, e (c) para ambos**



Fonte: Autoria própria (2025)

Já a figura 34 traz os mesmos gráficos de dispersão para as bananas, de cada modelo de melhor desempenho, também divididos em (a) somente sensor VIS, (b) somente sensor NIR e (c) modelo com ambos os sensores combinados:

**Figura 34 - Gráficos de dispersão real x previsto MLP para as bananas, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) modelo com ambos os sensores combinados**



Fonte: Autoria própria (2025)

Pode-se perceber claramente pelos gráficos da figura 32 como a nuvem de dados de teste se densifica mais na linha ideal do modelo de ambos os sensores que nos modelos onde foram analisados separadamente. Esses resultados validam a hipótese de complementariedade dos espectros visível e NIR, capturando informações sobre a mudança de pigmentos visíveis e da composição orgânica, proporcionando um conjunto mais rico que descreve melhor o processo de maturação,

permitindo à rede neural construir um modelo de preditivo substancialmente mais preciso.

### 4.3 Resultados da Regressão Linear Múltipla (MLR)

A Regressão Linear Múltipla (MLR) foi implementada como um método base de referência. Por ser um método mais simples, não necessitando de um processo iterativo complexo, os resultados geralmente apresentam menor acurácia, mas servem como ponto de partida para a avaliação comparativa dos demais métodos, demonstrando o ganho de performance em relação a esse método.

Para manter uma métrica comparativa com os demais métodos, foi feita a mesma divisão de dados, onde utilizou-se os 70% de treinamento para calcular os coeficientes da regressão, e os 15% de teste foram utilizados para calcular a Raiz do Erro Quadrático Médio (RMSE) e o coeficiente de determinação ( $R^2$ ). Também foram realizados ajustes para os três casos utilizados na MLP: utilizando somente os dados de cada sensor individualmente, e também com os dados de ambos os sensores. A tabela 14 expressa os resultados em termos de RMSE e  $R^2$  para cada caso.

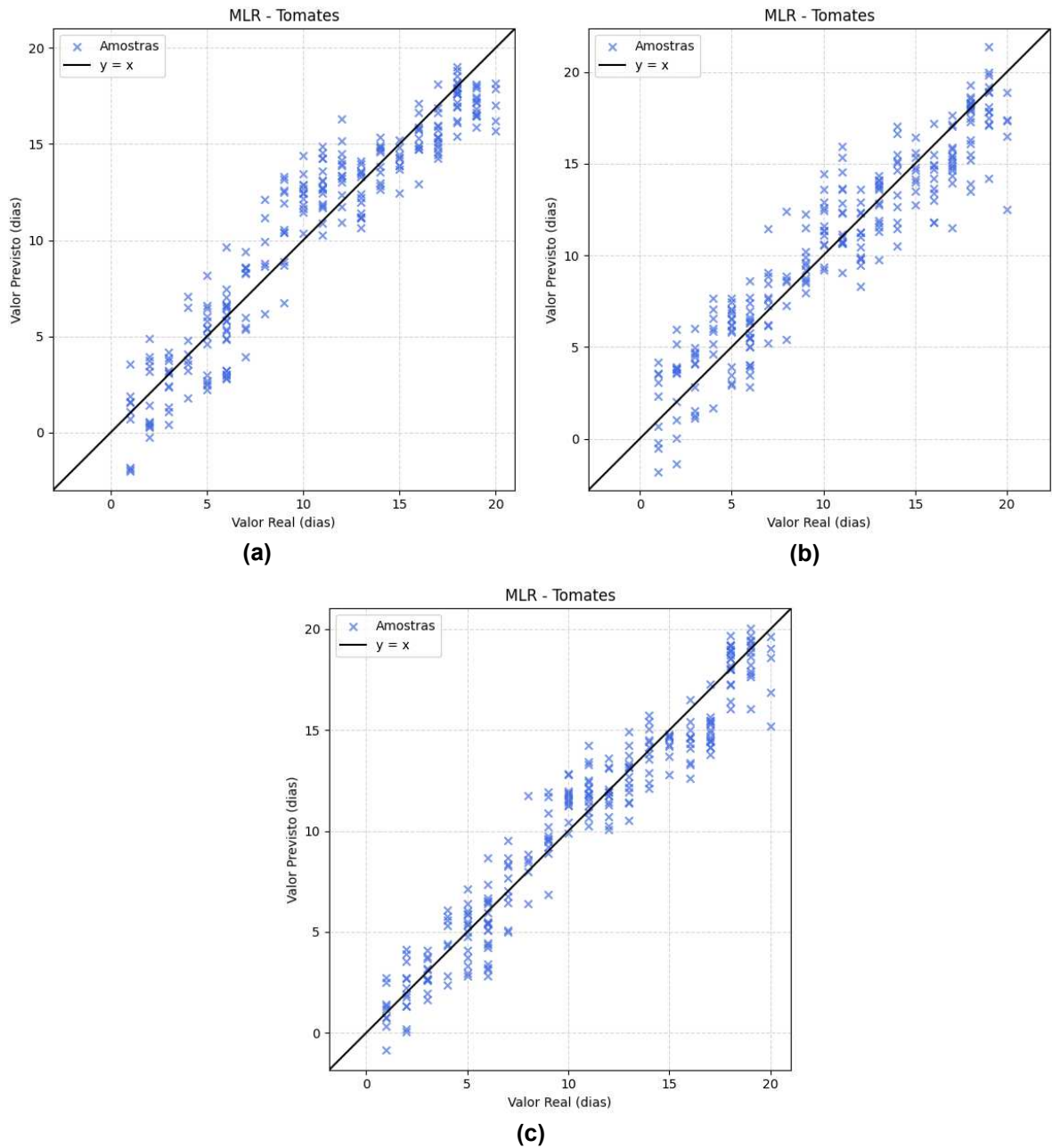
**Tabela 14 - Resultados para a Regressão Linear Múltipla (MLR)**

Frutas	Caso	$R^2$	RMSE (dias)
Tomates	Sensor VIS	0,889	1,858
	Sensor NIR	0,873	1,992
	Ambos	0,932	1,452
Bananas	Sensor VIS	0,863	1,897
	Sensor NIR	0,850	1,990
	Ambos	0,917	1,474

**Fonte: Autoria própria (2025)**

Mesmo no método mais básico, a fusão de dados de sensores mostrou uma melhoria de desempenho considerável do modelo preditivo. Para os tomates, o  $R^2$  no conjunto de teste saltou para 0,932 com a fusão dos dados, demonstrando uma forte correlação linear entre os espectros combinados e os dias de maturação. Para o caso das bananas o resultado foi semelhante. As figuras 34 e 35 complementam a análise, exibindo os gráficos de dia de maturação real x previsto trazendo uma diferenciação visual entre os modelos. A figura 35 (a), (b) e (c) representa os gráficos para os dados do sensor VIS, NIR e ambos os sensores respectivamente, no caso dos tomates:

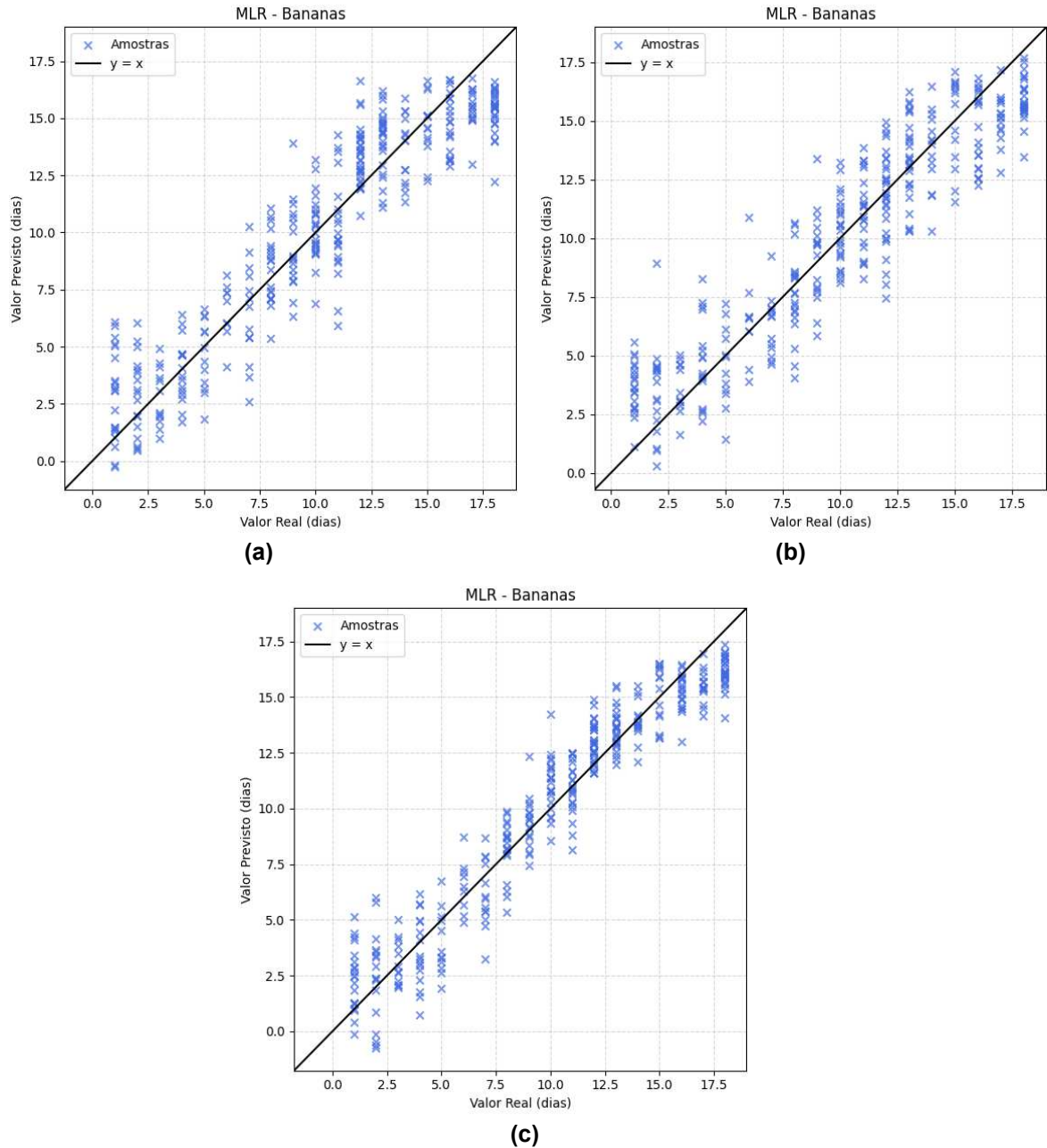
**Figura 35 - Gráficos de dispersão real x previsto MLR para os tomates, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) com ambos os sensores combinados**



**Fonte: Autoria própria (2025)**

A figura 36 (a), (b) e (c) representa os gráficos para os dados do sensor VIS, NIR e ambos os sensores respectivamente, no caso das bananas:

**Figura 36 - Gráficos de dispersão real x previsto MLR para as bananas, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) com ambos os sensores combinados**



**Fonte: Autoria própria (2025)**

Apesar dos bons resultados, o desempenho do MLR é limitado pela sua sensibilidade à multicolinearidade, ou seja, a alta correlação entre as variáveis preditoras (canais espectrais de faixas próximas tendem a variar juntos, uma vez que a reflectância da amostra varia de forma suave e contínua), o que pode tornar os coeficientes do modelo instáveis e menos confiáveis.

#### 4.4 Resultados da Regressão por Mínimos Quadrados Parciais (PLSR)

No contexto deste projeto, o PLSR foi aplicado para correlacionar os espectros obtidos pelos sensores AS7341 e AS7263 com os dias de maturação das frutas, onde foi também verificado quais comprimentos de onda são mais relevantes para prever o estágio de maturação.

A Regressão por Mínimos Quadrados Parciais (PLSR) foi implementada como um método intermediário, sendo uma evolução da MLR, já que ela lida bem com a multicolinearidade dos dados espectrais, através da criação dos componentes latentes. Assim como na MLP, foram testados vários modelos com diferentes números de componentes latentes, para cada caso (nos casos com dados de sensores separados, foram utilizados componentes até o número de canais espectrais do sensor, que é o máximo possível).

Da mesma forma que na MLR, foram utilizados os dados de treino para calcular os coeficientes de cada componente latente, e para fins de comparação, utilizados os dados separados para teste no cálculo do RMSE e do  $R^2$ . A tabela 15 apresenta os resultados para cada um dos casos:

**Tabela 15 - Resultados para a Regressão por Mínimos Quadrados parciais**

Caso	Componentes Latentes	Tomates		Bananas	
		$R^2$	RMSE	$R^2$	RMSE
Sensor	3	0,833	2,282	0,802	2,286
VIS	6	0,881	1,929	0,836	2,076
	9	0,889	1,858	0,863	1,897
Sensor	2	0,845	2,201	0,824	2,153
NIR	4	0,868	2,031	0,847	2,005
	6	0,873	1,992	0,850	1,990
Ambos os	3	0,872	2,000	0,845	2,017
Sensores	6	0,918	1,595	0,886	1,736
	9	0,924	1,544	0,900	1,621
	12	0,928	1,503	0,917	1,476
	15	0,932	1,452	0,917	1,474

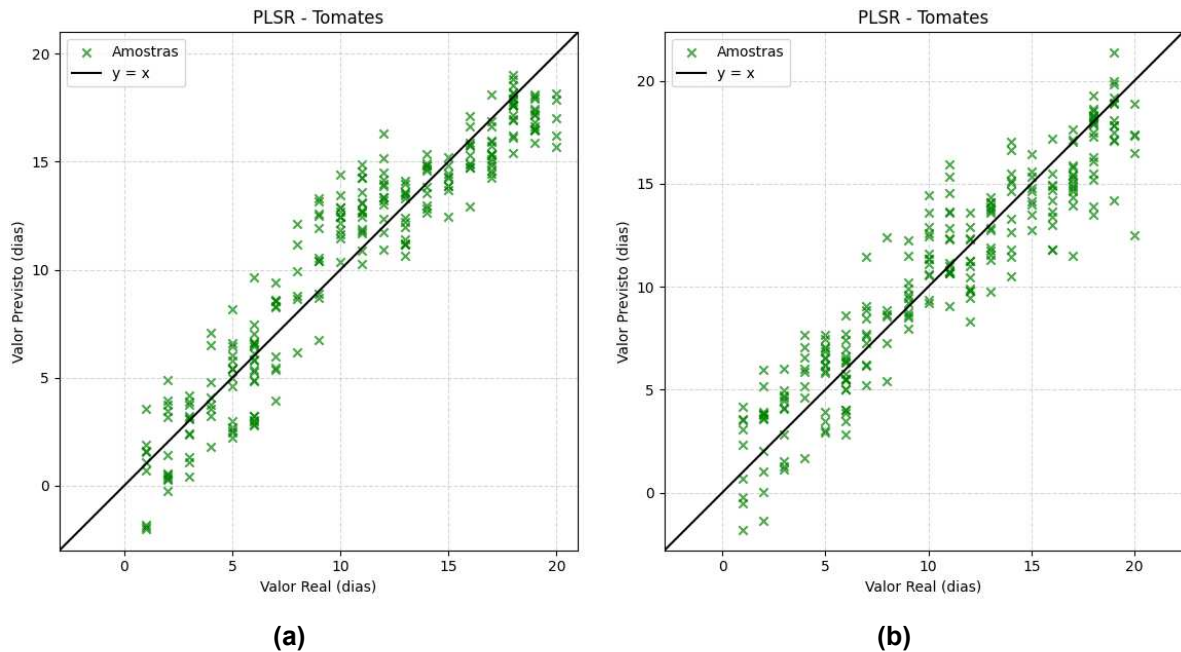
**Fonte: Autoria própria (2025)**

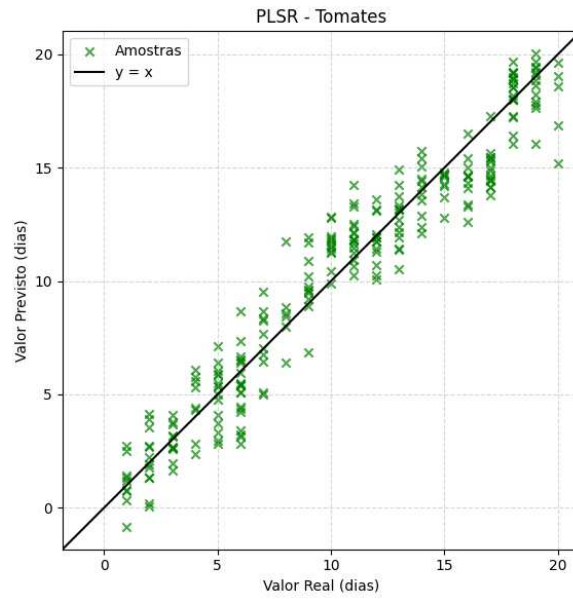
Mais uma vez, a fusão de sensores demonstrou um ganho de desempenho considerável do sistema, em todos os cenários, reforçando a complementariedade da

análise dos espectros visível e do NIR. Para os tomates, o  $R^2$  saltou de 0,889 (VIS, 9 componentes) e 0,873 (NIR, 6 componentes) para 0,932 com (ambos os sensores, 15 componentes). Já para as bananas, o  $R^2$  subiu de 0,863 (VIS, 9 componentes) e 0,850 (NIR, 6 componentes) para 0,917 (ambos os sensores, 12 componentes).

Um ponto importante observado, é que o aumento do número de componentes melhora o desempenho do modelo, porém há uma estagnação quando o número de componentes se aproxima do número total de canais. No caso dos tomates o melhor modelo em termos de  $R^2$  foi o de 15 componentes, mais a diferença para os modelos de 9 e 12 foi bem pequena, algo parecido também ocorre para o caso das bananas. Isso demonstra que os primeiros componentes já capturam a maior parte da informação útil que correlaciona a resposta espectral aos dias de maturação. Após um certo ponto, os componentes adicionais passam a modelar principalmente ruído em vez de padrões relevantes, não contribuindo para uma melhoria significativa do poder preditivo. Na figura 37 estão dispostos os gráficos de dispersão dia real x dia previsto de maturação, para o melhor modelo do sensor VIS (a), sensor NIR (b) e ambos os sensores (c), para os tomates:

**Figura 37 - Gráficos de dispersão real x previsto PLSR para os tomates, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) modelo com ambos os sensores combinados**



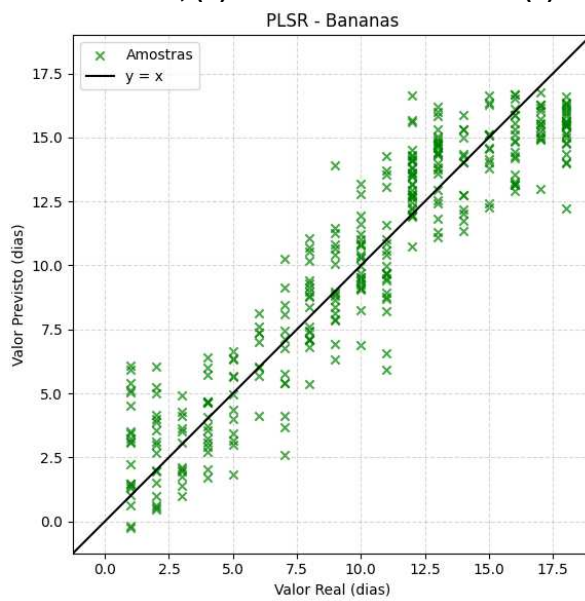


(c)

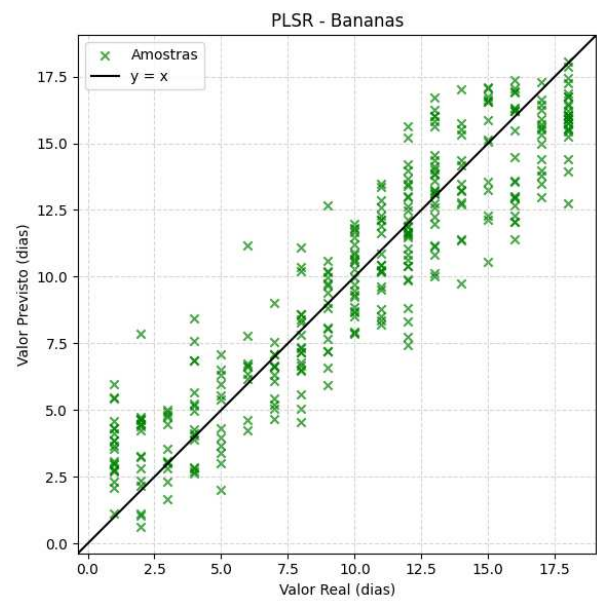
Fonte: Autoria própria (2025)

Na figura 38 estão dispostos os gráficos de dispersão dia real x dia previsto de maturação, para o melhor modelo do sensor VIS (a), sensor NIR (b) e ambos os sensores (c), para as bananas:

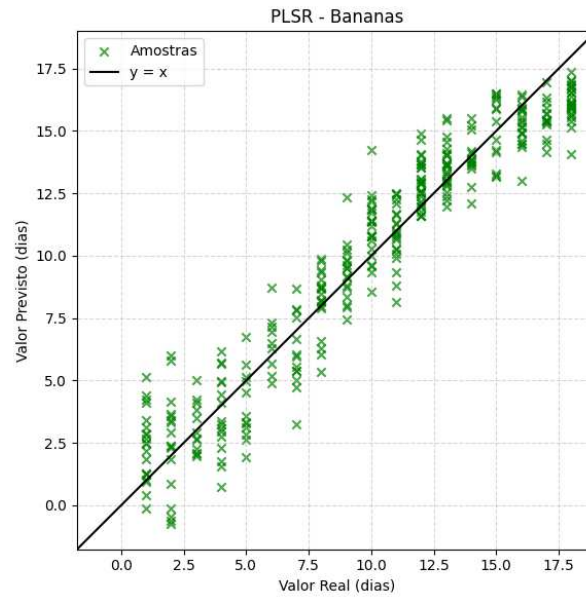
**Figura 38 - Gráficos de dispersão real x previsto PLSR para as bananas, sendo (a) somente sensor VIS, (b) somente sensor NIR e (c) modelo com ambos os sensores combinados**



(a)



(b)



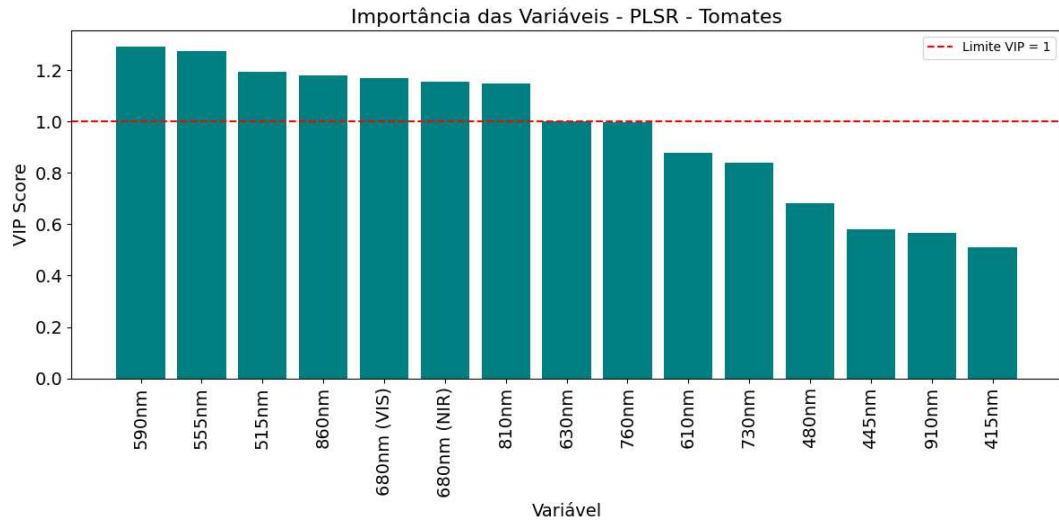
(c)

Fonte: Autoria própria (2025)

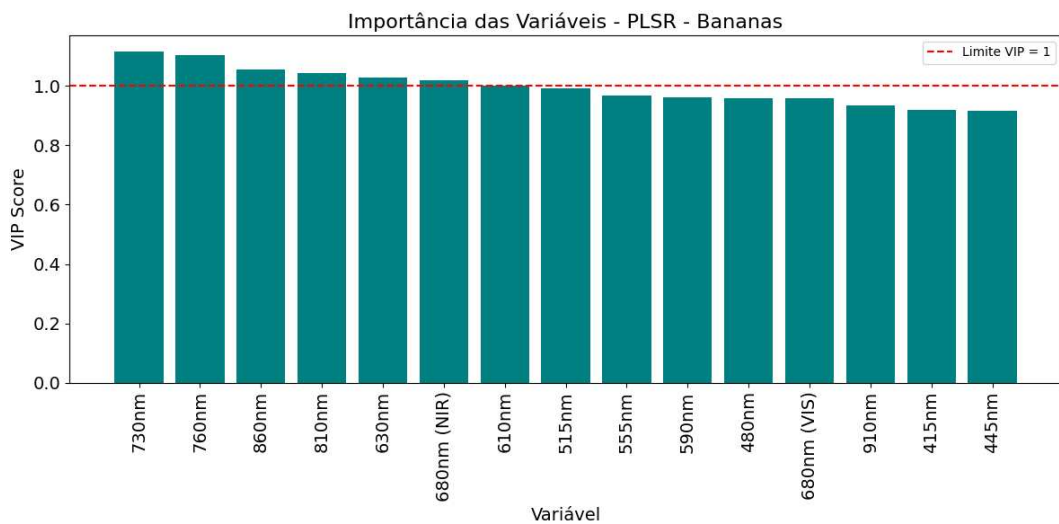
Quando se compara os resultados e gráficos dos resultados com o máximo de componentes, percebe-se que são idênticos com os obtidos com o MLR. Isso se deve ao fato que quanto maior o número de componentes se utiliza na PLSR, mais próximo ela é, matematicamente, do modelo da MLR. Quando se utiliza o número de componentes igual ao número de canais espectrais, então os métodos se igualam. Sendo assim o desempenho superior da PLSR não se dá em questões de métricas como o  $R^2$  e RMSE, mas sim pela sua eficiência: o método consegue obter respostas muito próximas à MLR, com um número menor de componentes que canais espectrais.

A análise dos *VIP scores* traz uma análise de quais canais foram mais relevantes para a construção dos modelos, para cada fruta. A figura 39 mostra a análise para o caso dos tomates (a) e das bananas (b), para os modelos mais eficientes com ambos os sensores.

**Figura 39 - Gráficos dos VIP Scores para os tomates e as bananas, sendo (a) análise dos tomates e (b) análise das bananas.**



(a)



(b)

Fonte: Autoria própria (2025)

Analisando o gráfico, percebe-se que, no caso dos tomates os canais dentro da faixa da verde ao vermelho (515 – 680 nm) e do NIR (810 – 860 nm) tiveram maior relevância, consistente com a degradação da clorofila e síntese de licopeno (carotenoide), além das alterações de composição química registradas pelo NIR. Já para o caso das bananas não houve tanta diferença entre os VIP scores mais relevantes e menos relevantes, se destacando as faixas do NIR (730 – 860 nm). Isso demonstra que o modelo foi capaz de identificar as assinaturas espectrais específicas de cada processo de maturação, corroborando que a fusão de dados de cor (VIS) e composição interna (NIR) é fundamental para a precisão do sistema.

#### 4.5 Comparativo entre os diferentes métodos

Após a apresentação dos resultados de cada método, a presente seção traz uma síntese entre os melhores resultados obtidos em cada um para uma comparação direta. A tabela 16 traz os resultados de melhor desempenho em termos de RMSE e  $R^2$  para os três modelos propostos para cada método:

**Tabela 16 - Tabela comparativa dos resultados**

Caso	Método	Tomates		Bananas	
		RMSE	$R^2$	RMSE	$R^2$
Sensor NIR	MLR	1,992	0,873	1,990	0,850
	PLSR	1,992	0,873	1,990	0,850
	MLP	1,632	0,908	1,763	0,882
Sensor VIS	MLR	1,858	0,889	1,897	0,863
	PLSR	1,858	0,889	1,897	0,863
	MLP	1,459	0,932	1,836	0,872
Ambos os sensores	MLR	1,452	0,932	1,474	0,917
	PLSR	1,452	0,932	1,474	0,917
	MLP	1,193	0,954	1,302	0,936

**Fonte: Autoria própria (2025)**

Fica evidente que a MLP apresentou um desempenho substancialmente superior aos métodos estatísticos tradicionais, em todos os cenários. A acurácia da MLP é atribuída a sua capacidade de lidar com padrões não lineares e relações complexas entre dados, algo que os métodos de regressões lineares não conseguem.

Ao analisar as respostas dos modelos de regressão da MLR e PLSR com melhor desempenho bruto em termos de RMSE e  $R^2$ , encontra-se um fato importante: resultados idênticos atrelados a convergência matemática entre os dois modelos. A principal vantagem do PLSR, no entanto, é a capacidade de atingir um desempenho ótimo, ou muito próximo do máximo, utilizando um número de componentes menor que o número total de canais espectrais. Isso implica em um modelo mais eficiente e robusto, capaz de lidar melhor com ruído e multicolinearidade presente nessa aplicação.

## 5 CONCLUSÃO

O objetivo principal do presente trabalho foi o desenvolvimento de um espectrofotômetro de baixo custo, capaz de medir o espectro de refletância de diferentes faixas do espectro visível e do infravermelho próximo, e retornar dados para a elaboração de modelos preditivos com grau de precisão elevado, no âmbito da Prova de conceito do uso de um dispositivo NIR para avaliação da maturação de frutas.

Para atingir o objetivo foi elaborado um sistema baseado nos sensores de espectro comerciais AS7341 e AS7263, em conjunto com o microcontrolador ESP32, e um display de interface homem-máquina. Esse sistema foi desenvolvido de forma portátil, capaz de gravar dados obtidos com os sensores de espectro em cartão SD, além de possibilitar diversos ajustes de parâmetros e de controle de aquisição de dados.

A partir do protótipo foram levantados extensos conjuntos de dados referentes aos espectros de tomates e bananas durante o processo de maturação, posteriormente analisados, por três modelos distintos de regressão: Regressão Linear Múltipla (MLR), Regressão por Mínimos Quadrados Parciais (PLSR) e uma rede neural artificial do tipo *Multilayer Perceptron* (MLP).

Os resultados obtidos demonstraram o sucesso da proposta, apresentando desempenho satisfatório em todos os modelos testados, evidenciando como a fusão de dados dos sensores da faixa visível e do infravermelho próximo amplia a precisão na predição dos modelos, em comparação ao uso de cada sensor isoladamente. Na comparação de modelos preditivos, a MLP se destacou como modelo mais acurado e eficaz na predição de dia de maturação.

No geral, as análises evidenciaram que um protótipo focado em baixo custo, com poucas faixas de cobertura do espectro, utilizando sensores comercialmente acessíveis, foi capaz de analisar o processo de maturação de frutas de forma eficiente. Sendo assim, o trabalho desenvolvido contribui significativamente para a área de instrumentação aplicada à agricultura apresentando uma alternativa acessível de baixo custo, compacto, portátil e de baixo consumo energético para a utilização no âmbito do controle de qualidade de alimentos.

Como trabalhos futuros, ainda cabe o aprimoramento do sistema óptico, bem como a ampliação de aplicação do dispositivo para análises de maturação, mas também na qualidade de alimentos em geral.



## REFERÊNCIAS

- ALCANTARA JR., Petrus. **Espectroscopia Molecular**. Universidade Federal do Pará, Departamento de Física, março 2002. Disponível em: <https://www.researchgate.net/publication/267222966>. Acesso em: 04 fev. 2025.
- ALMEIDA, Gleiciane Ad Vincula. **Espectroscopia NIR na quantificação de teores nutricionais de alimentos processados**. 2015. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) – Universidade Federal do Acre, Rio Branco, 2015. Disponível em: <https://www.ufac.br/cita/o-programa/dissertacoes/2015/gleiciane-ad-vincula-almeida.pdf>. Acesso em: 06 maio 2025.
- AMS OSRAM. **AS7262 6-Channel Visible Spectral Sensor**. Datasheet, 2021. Disponível em: <https://ams.com/as7262>. Acesso em: 20 agosto 2025.
- AMS OSRAM. **AS7263 6-Channel NIR Spectral Sensor**. Datasheet, 2021. Disponível em: <https://ams.com/as7263>. Acesso em: 20 agosto 2025.
- AMS OSRAM. **AS7341 11-Channel Spectral Sensor**. Datasheet, 2021. Disponível em: <https://ams.com/as7341>. Acesso em: 20 agosto 2025.
- ARDUINO. **Arduino Mega 2560 Rev3**: Datasheet. Disponível em: <https://docs.arduino.cc/resources/datasheets/A000067-datasheet.pdf>. Acesso em: 18 set. 2025.
- ARDUINO. **Arduino Uno R3**: Datasheet. Disponível em: <https://docs.arduino.cc/resources/datasheets/A000066-datasheet.pdf>. Acesso em: 18 set. 2025.
- AVANTES. **AvaSpec-Mini Series Compact Spectrometers**. 2024. Disponível em: <https://www.avantes.com/products/spectrometers/avaspec-mini>. Acesso em: 20 maio 2025.
- BARRETO, Naiane Santos. **Aplicação da técnica de espectroscopia do visível e infravermelho próximo (VIS/NIR) no controle físico-químico da qualidade de vinhos produzidos no submédio do Vale do São Francisco**. 2020. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) – Universidade Federal de Sergipe, São Cristóvão, SE, 2020. Disponível em: [https://ri.ufs.br/bitstream/riufs/13899/2/NAIANE\\_SANTOS\\_BARRETO.pdf](https://ri.ufs.br/bitstream/riufs/13899/2/NAIANE_SANTOS_BARRETO.pdf). Acesso em: 20 jun. 2025.
- BURNS, D. A.; CIURCZAK, E. W. **Handbook of Near-Infrared Analysis**. 3. ed. Boca Raton: CRC Press, 2007.
- CARL SAGAN: 20 frases de Carl Sagan para entender mais sobre a Terra, o espaço e a humanidade. **Exame**, 26 jul. 2017. Disponível em: [https://classic.exame.com/ciencia/20-frases-de-carl-sagan-para-entender-mais-sobre-a-terra-o-espaco-e-ahumanidade/?utm\\_source=copiaecola&utm\\_medium=compartilhamento](https://classic.exame.com/ciencia/20-frases-de-carl-sagan-para-entender-mais-sobre-a-terra-o-espaco-e-ahumanidade/?utm_source=copiaecola&utm_medium=compartilhamento). Acesso em: 10 set. 2025.

CARTER, G. A. **Relationships between leaf pigment concentrations and spectral reflectance in cotton. Remote Sensing of Environment**, v. 47, n. 1, p. 20-28, 1994.  
DAVIES, A. M. C.; WILLIAMS, P. **Near-infrared Spectroscopy: Principles, Instruments, Applications**. Chichester: John Wiley & Sons, 2008.

DESHPANDE, Atharva; DESHPANDE, Shreyash; DHANDE, Shaunak. **NIR Spectroscopy Based Milk Classification and Purity Prediction**. In: 2021 IEEE Pune Section International Conference (PuneCon). MIT ADT University, Pune, Índia, dez. 2021. p. 1-5. DOI: 10.1109/PUNECON52575.2021.9686473. Disponível em: <https://ieeexplore.ieee.org/document/9686473>. Acesso em: 8 dez. 2024.

EEMB CO., LTD. **LIR18650 Battery — Datasheet**. Novembro de 2010. Disponível em: <https://www.ineltro.ch/media/downloads/SAItem/45/45958/36e3e7f3-2049-4adb-a2a7-79c654d92915.pdf>. Acesso em: 25 set. 2025.

ELETROGATE. **Conhecendo o ESP32 – Introdução (1)**. Blog Eletrogate, 2018. Disponível em: <https://blog.eletrogate.com/conhecendo-o-esp32-introducao-1/>. Acesso em: 17 set. 2025.

EMBRAPA. **Espectroscopia no infravermelho próximo (NIRS) como ferramenta analítica em ciência animal**. Embrapa Pecuária Sudeste, 2018. Disponível em: <https://www.alice.cnptia.embrapa.br/bitstream/doc/1106595/1/ID445392018LVespectroscopia.pdf>. Acesso em: 06 maio 2025.

EMBRAPA. **Espectroscopia no infravermelho próximo para avaliar indicadores de qualidade tecnológica e contaminantes em grãos**. 2018. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1106595/espectroscopia-no-infravermelho-proximo-para-avaliar-indicadores-de-qualidade-tecnologica-e-contaminantes-em-graos>. Acesso em: 15 jun. 2025.

ESPBOARDS. **NodeMCU-32S pinout. 2024**. Disponível em: <https://espboards.dev/esp32/nodemcu-32s/>. Acesso em: 17 set. 2025.

ESPRESSIF. **ESP32 Series Datasheet**. 2025. Disponível em: [https://www.espressif.com/sites/default/files/documentation/esp32\\_datasheet\\_en.pdf](https://www.espressif.com/sites/default/files/documentation/esp32_datasheet_en.pdf). Acesso em: 18 set. 2025.

ESPRESSIF. **ESP32 Technical Reference Manual**, 2025. Disponível em: [https://www.espressif.com/sites/default/files/documentation/esp32\\_technical\\_reference\\_manual\\_en.pdf](https://www.espressif.com/sites/default/files/documentation/esp32_technical_reference_manual_en.pdf). Acesso em: 18 set. 2025.

FARRÉS, Mireia; PLATIKANOV, Stefan; TSAKOVSKI, Stefan; TAULER, Romà. **Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation**. Journal of Chemometrics, v. 29, n. 10, p. 528–536, 2015.

FERNANDEZ, Ana Sophia Tovar et al. **Autenticação de orégano (*Origanum vulgare* L.) orgânico utilizando espectroscopia NIR e quimiometria**. Química Nova, v. 43, n. 10, p. 1500–1504, 2020. Disponível em: <https://doi.org/10.21577/0100-4042.20170616>. Acesso em: 06 maio 2025.

GOETTEMS, Elisa Iahn. **Espectroscopia: uma sequência didática com enfoque na astronomia para ensino médio**. 2017. Trabalho de Conclusão de Curso (Licenciatura em Física) – Universidade Federal de Santa Catarina, Florianópolis, 2017. Disponível em: <https://repositorio.ufsc.br/handle/123456789/177854>. Acesso em: 18 maio 2025.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge, MA: MIT Press, 2016.

GUIDI, Lucia; TATTINI, Massimiliano; LANDI, Marco. **Chlorophyll a, b and carotenoids absorbance spectra**. In: How Does Chloroplast Protect Chlorophyll Against Excessive Light? ResearchGate, maio 2017. Disponível em: [https://www.researchgate.net/figure/Chlorophyll-a-b-and-carotenoids-absorbance-spectra\\_fig1\\_317151195](https://www.researchgate.net/figure/Chlorophyll-a-b-and-carotenoids-absorbance-spectra_fig1_317151195). Acesso em: 18 maio 2025.

HALLYDAY, D.; RESNICK, R.; WALKER, J. **Fundamentos de física: óptica e física moderna**. 10. ed. Rio de Janeiro: LTC, 2016.

HAMAMATSU. **C12880MA – Mini-spectrometer**. Datasheet, 2020. Disponível em: <https://www.hamamatsu.com/jp/en/product/type/C12880MA/index.html>. Acesso em: 20 maio 2025.

HAMAMATSU. **C12666MA – Micro-spectrometer**. Datasheet, 2020. Disponível em: <https://www.hamamatsu.com/jp/en/product/type/C12666MA/index.html>. Acesso em: 20 maio 2025.

HAMAMATSU PHOTONICS. **L10363 Series: Red LED. Datasheet**. Disponível em: [https://www.hamamatsu.com/resources/pdf/ssd/l10363\\_series\\_kled1007e.pdf](https://www.hamamatsu.com/resources/pdf/ssd/l10363_series_kled1007e.pdf). Acesso em: 25 set. 2025.

HAREENDRAN, T.K. **DC/DC module gets CV/CC control**. EDN, 10 ago. 2021. Disponível em: <https://www.edn.com/dc-module/>. Acesso em: 26 set. 2025.

HARRIS, D. C. **Análise química quantitativa**. 8. ed. Rio de Janeiro: LTC, 2012.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. New York: Springer, 2009.

HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. 2. ed. Upper Saddle River: Prentice Hall, 1999.

IBGE. **População do país vai parar de crescer em 2041**. Agência de Notícias, 2012. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/41056-populacao-do-pais-vai-parar-de-crescer-em-2041>. Acesso em: 04 fev. 2025.

INOLUX CORPORATION. **INL-5AHIR30: 5 mm Infrared Emitting Diode – Datasheet**. Inolux Corporation, 19 jan. 2019. Disponível em: [https://www.inolux-corp.com/datasheet/IR/Emitter/5mm%20Lamp/INL-5AHIR30\\_V1.0.pdf](https://www.inolux-corp.com/datasheet/IR/Emitter/5mm%20Lamp/INL-5AHIR30_V1.0.pdf). Acesso em: 25 set. 2025.

IONECSU, Marius; PAȘCA, Sever. **Algorithms of Absorbance and Colorimeter for Measuring Blood Glucose**. In: 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE), Bucharest, Romania, 2019. IEEE. DOI: 10.1109/ATEE.2019.8724946. Disponível em: <https://ieeexplore.ieee.org/document/8724946>. Acesso em: 06 maio 2025.

LCDWIKI. **2.8-inch SPI Module ILI9341 (SKU: MSP2807) Specifications and User Manual**. 2025. Disponível em: [https://www.lcdwiki.com/2.8inch\\_SPI\\_Module\\_ILI9341\\_SKU%3AMSP2807](https://www.lcdwiki.com/2.8inch_SPI_Module_ILI9341_SKU%3AMSP2807). Acesso em: 18 set. 2025.

LICHTENTHALER, H. K. **Chlorophylls and carotenoids, pigments of photosynthetic biomembranes**. *Methods in Enzymology*, v. 148, p. 350-381, 1987.

MAGALHÃES, Diogo Faria. **Aplicações da Espectroscopia de Infravermelho Próximo na Monitorização de Processos Farmacêuticos**. 2014. Dissertação (Mestrado em Química Tecnológica) – Faculdade de Ciências, Universidade de Lisboa, Lisboa, 2014.

MANUALS.PLUS. **AS7341 Visible Spectrum Infrared Color Spectrometer Sensor Module User Manual**. 2024. Disponível em: <https://manuals.plus/asin/B0DBQKDV67>. Acesso em: 11 set. 2025.

METROHM. **Determinação rápida e simples da umidade em fertilizantes por espectroscopia NIR**. Metrohm Brasil, 2021. Disponível em: [https://www.metrohm.com/pt\\_br/fique-por-dentro/blog/20-21/easy-moisture-determination-in-fertilizers-by-near-infrared-spec.html](https://www.metrohm.com/pt_br/fique-por-dentro/blog/20-21/easy-moisture-determination-in-fertilizers-by-near-infrared-spec.html). Acesso em: 06 maio 2025.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to Linear Regression Analysis**. 5. ed. Hoboken, NJ: John Wiley & Sons, 2012.

NAÇÕES UNIDAS. **População mundial chegará a 10,3 bilhões em 2080, revela relatório da ONU**. UN News, 01 jul. 2024. Disponível em: <https://news.un.org/pt/story/2024/07/1834411>. Acesso em: 06 maio 2025.

NASCIMENTO, Sílvia Panetta. **Desperdício de alimentos: fator de insegurança alimentar e nutricional**. *Segurança Alimentar e Nutricional*, Campinas, v. 25, n. 1, p. 85-91, jan./abr. 2018. DOI: <http://dx.doi.org/10.20396/san.v25i1.8649917>. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/san/article/view/8649917/17978>. Acesso em: 03 de fev. 2025.

OSBORNE, B. G.; FEARN, T.; HINDLE, P. H. **Practical Near-Infrared Spectroscopy with Applications in Food and Beverage Analysis**. Essex: Longman Scientific & Technical, 1993.

PEREIRA, J. S. et al. **Determinação de resíduos de agrotóxicos em leite bovino empregando método QuEChERS modificado e GC-MS/MS**. *Química Nova*, v. 42, n. 10, p. 1195-1204, 2019. Disponível em: <https://www.scielo.br/j/qn/a/R9jr8SnF3JWQvjRZ834frLn/?lang=pt>. Acesso em: 06 maio 2025.

- PEREIRA, Luiz Fernando dos Santos. **Caracterização do estágio de maturação do mamão papaia golden usando espectroscopia no infravermelho próximo e processamento de imagens RGB digitais**. 2017. Dissertação (Mestrado em Engenharia Agrícola) – Universidade Federal de Viçosa, Viçosa, 2017. Disponível em: <https://repositorio.unicamp.br/Busca/Download?codigoArquivo=461981>. Acesso em: 8 dez. 2024.
- RICHARDSON, A. D.; DUGAN, S. P.; BERLYN, G. P. **An evaluation of the chlorophyll content index in *Rhododendron maximum* L.** Journal of Plant Physiology, v. 159, n. 4, p. 385-392, 2002.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533-536, 1986.
- SILVA, Anderson Victor da. **Deteção de pesticidas em cascas de alimentos por Espectroscopia Raman amplificada por superfície e quimiometria**. 2024. Dissertação (Mestrado em Química Analítica) - Programa de Pós-Graduação em Química, Universidade Federal do Paraná, Curitiba, 2024. Disponível em: <https://acervodigital.ufpr.br/xmlui/handle/1884/94804>. Acesso em: 4 set. 2025.
- SILVA, Larissa Kauly Rosa. **Espectroscopia MIR e NIR, avaliação proteômica e medidas analíticas associadas à quimiometria para discriminação da origem geográfica de queijo Coalho**. 2021. 177 f. Tese (Doutorado em Engenharia e Ciência de Alimentos) – Universidade Estadual do Sudoeste da Bahia, Itapetinga, BA, 2021. Disponível em: <http://www2.uesb.br/ppg/ppgecal/wp-content/uploads/2021/03/LARISSA-KAULY-ROSA-SILVA.docx.pdf>. Acesso em: 8 dez. 2024.
- SIESLER, H. W. **Basic principles of near-infrared spectroscopy**. In: CIURCZAK, E. W.; BURNS, D. A. (ed.). **Handbook of Near-Infrared Analysis**. 3. ed. Boca Raton: CRC Press, 2008. cap. 2.
- SKOOG, Douglas A. et al. **Fundamentos de química analítica**. 8. ed. São Paulo: Pioneira Thomson Learning, 2006.
- SKOOG, D. A.; HOLLER, F. J.; CROUCH, S. R. **Principles of Instrumental Analysis**. 7. ed. Boston: Cengage Learning, 2017.
- STALLINGS, W. **Arquitetura e Organização de Computadores**. 10. ed. São Paulo: Pearson Prentice Hall, 2017.
- SULISTYO, S. B. et al. **Handheld arduino-based near infrared spectrometer for non-destructive quality evaluation of siamese oranges**. IOP Conference Series: Earth and Environmental Science, v. 653, n. 1, p. 012119, 2021. Disponível em: [https://www.researchgate.net/publication/349200385\\_Handheld\\_arduino\\_based\\_near\\_infrared\\_spectrometer\\_for\\_non-destructive\\_quality\\_evaluation\\_of\\_siamese\\_oranges](https://www.researchgate.net/publication/349200385_Handheld_arduino_based_near_infrared_spectrometer_for_non-destructive_quality_evaluation_of_siamese_oranges). Acesso em: 15 jun. 2025.
- TAIZ, Lincoln et al. **Fisiologia e desenvolvimento vegetal**. 6. ed. Porto Alegre: Artmed, 2016.

VALDERRAMA, P. et al. **UV-Vis spectrum fingerprinting and chemometric method in the evaluation of extra virgin olive oil adulteration and fraud..** Journal of Agricultural and Food Chemistry, v. 62, n. 47, p. 11467-11473, 2014. Disponível em: <https://periodicos.utfpr.edu.br/rebrapa/article/view/3447>. Acesso em: 15 jun. 2025.

VIEIRA, Douglas Alexandre Gomes. Rede **Perceptron com Camadas Paralelas (PLP - Parallel Layer Perceptron)**. 2006. Tese (Doutorado em Engenharia Elétrica) – Universidade Federal de Minas Gerais, Belo Horizonte, 2006. Disponível em: <https://www.researchgate.net/publication/237542058>. Acesso em: 26 jun. 2025.

WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. **PLS-regression: a basic tool of chemometrics**. Chemometrics and Intelligent Laboratory Systems, v. 58, n. 2, p. 109-130, 2001.

WOLF, Wayne. **Computers as Components: Principles of Embedded Computing System Design**. 4. ed. San Francisco: Morgan Kaufmann, 2016.

## **APÊNDICE A - Códigos em *python* dos métodos de análise utilizados**

## Código para os métodos MLR e PLSR

### Célula 1 – Importar bibliotecas

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import mean_squared_error, r2_score,
mean_absolute_error
from sklearn.linear_model import LinearRegression
from sklearn.cross_decomposition import PLSRegression
```

### # Carregar dados do Google Drive

```
from google.colab import drive
drive.mount('/content/drive')
```

### Célula 2 – Função de avaliação de modelo (MSE, RMSE, MAE, R<sup>2</sup>)

#### # Definição da função para avaliar o modelo

```
def avaliar_modelo(y_test, y_pred, nome_modelo):
    r2 = r2_score(y_test, y_pred)
    mse = mean_squared_error(y_test, y_pred)
    rmse = np.sqrt(mse)
    mae = mean_absolute_error(y_test, y_pred)
    print(f"{nome_modelo} - R2: {r2:.3f}, MSE: {mse:.3f}, RMSE:
{rmse:.3f}, MAE: {mae:.3f}")

    return r2, mse, rmse, mae
```

### Célula 3 – Função para carregar e preparar os dados

#### # Função para carregar e preparar os dados - Corrigida com listas explícitas de colunas

```
# %%
```

```
def carregar_dados(caminho_csv, sensor='todos'):
```

```
    df = pd.read_csv(caminho_csv)
    df.columns = df.columns.str.strip()
```

#### # Listas explícitas de colunas para cada sensor

```
    colunas_vis_sensor = [
        '415nm', '445nm', '480nm', '515nm', '555nm', '590nm',
        '630nm', '680nm', 'NIR(910nm)'
    ]
    colunas_nir_sensor = [
        '610nm', '680nm_1', '730nm', '760nm', '810nm', '860nm'
```

```

]

if sensor == 'VIS':
    # Seleciona apenas as colunas da lista VIS que existem no
    DataFrame
    colunas_selecionadas = [col for col in colunas_vis_sensor if
col in df.columns]
    X = df[colunas_selecionadas].copy()

elif sensor == 'NIR':
    # Seleciona apenas as colunas da lista NIR que existem no
    DataFrame
    colunas_selecionadas = [col for col in colunas_nir_sensor if
col in df.columns]
    X = df[colunas_selecionadas].copy()

else: # 'todos' - combina as colunas de ambos
    # Combina as listas e seleciona apenas as colunas que existem
    no DataFrame
    colunas_todos_sensor = colunas_vis_sensor + colunas_nir_sensor
    colunas_selecionadas = [col for col in colunas_todos_sensor if
col in df.columns]
    X = df[colunas_selecionadas].copy()

# Garantir que a coluna 'Dia' é sempre o target 'y'
if 'Dia' in df.columns:
    y = df['Dia'].copy()
else:
    # Lidar com o caso onde a coluna 'Dia' não existe
    raise KeyError("A coluna 'Dia' não foi encontrada no arquivo
CSV.")

# Verificação opcional dos nomes das colunas selecionadas
# print(f"Colunas selecionadas para sensor '{sensor}':
{X.columns.tolist()}")

return X, y

# Função para calcular os Vip Scores da PLSR

def calcular_vip(pls_model, X):
    t = pls_model.x_scores_
    w = pls_model.x_weights_
    q = pls_model.y_loadings_

    p, h = w.shape
    s = np.diag(t.T @ t @ q.T @ q).reshape(h, -1)
    Wnorm2 = (w ** 2).sum(axis=0).reshape(h, -1)
    vip = np.zeros((p,))

```

```

for i in range(p):
    weight = np.array([ (w[i, j]**2) / Wnorm2[j] for j in range(h)
])
    vip[i] = np.sqrt(p * (s.T @ weight).squeeze())

return vip

```

#### Célula 4 – Código para MLR (Regressão Linear Múltipla)

```

# Definição da função MLR: Regressão Linear Múltipla
def executar_mlr(X, y, nome='MLR'):
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X)

    # Dividir dados: 70% treino, 15% val, 15% teste

    X_train, X_temp, y_train, y_temp = train_test_split(X_scaled, y,
test_size=0.3, random_state=42)
    X_val, X_test, y_val, y_test = train_test_split(X_temp, y_temp,
test_size=0.5, random_state=42)

    # Treinar modelo MLR
    modelo = LinearRegression()
    modelo.fit(X_train, y_train)

    # Avaliar modelo nos conjuntos de treino, validação e teste
    print(f"--- Avaliação do Modelo {nome} ---")

    # Previsões para avaliação
    y_train_pred = modelo.predict(X_train)
    y_val_pred = modelo.predict(X_val)
    y_test_pred = modelo.predict(X_test)

    # Avaliar e imprimir métricas
    avaliar_modelo(y_train, y_train_pred, f"{nome} - Treino")
    avaliar_modelo(y_val, y_val_pred, f"{nome} - Validação")
    avaliar_modelo(y_test, y_test_pred, f"{nome} - Teste")

    # Imprimir os pesos (coeficientes)
    print(f"\nPesos (coeficientes) finais do modelo {nome}:")
    for feature, weight in zip(X.columns, modelo.coef_):
        print(f" {feature}: {weight:.4f}")

    # Imprimir o intercepto
    print(f" Intercepto: {modelo.intercept_:.4f}")

    print("-" * (len(f"--- Avaliação do Modelo {nome} ---"))) # Linha
separadora

```

```

    return y_test, y_test_pred # Continuar retornando os dados de
teste para plotagem

```

#### Célula 5 – Código para PLSR

```

# Definição da função que executa a PLSR

def executar_plsr(X, y, nome='PLSR', n_componentes=15):
    # Padronizar os dados
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X)

    # Dividir dados: 70% treino, 15% val, 15% teste
    X_train, X_temp, y_train, y_temp = train_test_split(X_scaled, y,
test_size=0.3, random_state=42)
    X_val, X_test, y_val, y_test = train_test_split(X_temp, y_temp,
test_size=0.5, random_state=42)

    # Treinar modelo PLSR
    modelo = PLSRegression(n_components=n_componentes)
    modelo.fit(X_train, y_train)

    # Avaliar modelo nos conjuntos de treino, validação e teste
    print(f"--- Avaliação do Modelo {nome} (Componentes:
{n_componentes}) ---")

    # Previsões para avaliação

    y_train_pred = modelo.predict(X_train).flatten()
    y_val_pred = modelo.predict(X_val).flatten()
    y_test_pred = modelo.predict(X_test).flatten()

    # Avaliar e imprimir métricas
    avaliar_modelo(y_train, y_train_pred, f"{nome} - Treino")
    avaliar_modelo(y_val, y_val_pred, f"{nome} - Validação")
    avaliar_modelo(y_test, y_test_pred, f"{nome} - Teste")

    print("-" * (len(f"--- Avaliação do Modelo {nome} (Componentes:
{n_componentes}) ---"))) # Linha separadora

    # CÁLCULO VIP Scores

    T = modelo.x_scores_
    W = modelo.x_weights_
    Q = modelo.y_loadings_
    p, h = W.shape

```

```

SS = (T**2).sum(axis=0) * (Q**2).sum(axis=0)

soma_SS_W2 = np.sum(SS * (W / np.linalg.norm(W, axis=0))**2,
axis=1)
vip_scores = np.sqrt(p * soma_SS_W2 / np.sum(SS))

# --- MODIFICAÇÃO PARA DISTINGUIR 680nm E RENOMEAR NIR(910nm) ---
nomes_variaveis = X.columns.tolist() # Pega os nomes originais

# Cria uma cópia dos nomes para modificar para o VIP plot/table
nomes_para_vip = nomes_variaveis[:]

# Encontra e renomeia os canais 680nm
if '680nm' in nomes_para_vip:
    index_680nm = nomes_para_vip.index('680nm')
    nomes_para_vip[index_680nm] = '680nm (VIS)'

if '680nm_1' in nomes_para_vip:
    index_680nm_1 = nomes_para_vip.index('680nm_1')
    nomes_para_vip[index_680nm_1] = '680nm (NIR)'

# Encontra e renomeia NIR(910nm)
if 'NIR(910nm)' in nomes_para_vip:
    index_910nm = nomes_para_vip.index('NIR(910nm)')
    nomes_para_vip[index_910nm] = '910nm'

vip_df = pd.DataFrame({'Variável': nomes_para_vip, 'VIP':
vip_scores}) # Usa os nomes modificados
vip_df = vip_df.sort_values(by='VIP', ascending=False)

# Mostrar TODAS as variáveis pelo VIP
print("\nVariáveis pelo VIP (calculado no treino):")
display(vip_df)

# Gráfico dos VIPs - mostrando TODOS os canais
plt.figure(figsize=(12, 6))
plt.bar(vip_df['Variável'], vip_df['VIP'], color='teal')
plt.axhline(1.0, color='red', linestyle='--', label='Limite VIP =
1')

plt.xticks(rotation=90, ha='center', fontsize=14) # Aumenta o
tamanho da fonte dos rótulos do eixo x
plt.yticks(fontsize=14) # Aumenta o tamanho da fonte dos rótulos do
eixo y
plt.ylabel("VIP Score", fontsize=14) # Aumenta o tamanho da fonte
do rótulo do eixo y
plt.xlabel("Variável", fontsize=14) # Adiciona rótulo ao eixo x e
aumenta o tamanho da fonte
plt.title(f"Importância das Variáveis - {nome}", fontsize=16) #
Aumenta o tamanho da fonte do título

```

```

plt.legend()
plt.tight_layout()
plt.show()

# Retornar valores para gráfico real vs predito (usar teste)
return y_test, y_test_pred

# Código para plotagem gráfico Real x Previsto

def plotar_real_vs_predito(y_real, y_pred, titulo="Real vs Previsto",
cor='royalblue'):
    plt.figure(figsize=(6, 6))
    plt.scatter(y_real, y_pred, marker='x', color=cor, alpha=0.7,
label='Amostras')

    # Reta de referência (y = x)

    max_val = max(max(y_real), max(y_pred))
    min_val = min(min(y_real), min(y_pred))
    plt.plot([min_val-1, max_val+1], [min_val-1, max_val+1],
color='black', linestyle='-', label='y = x')

    plt.xlabel("Valor Real (dias)")
    plt.ylabel("Valor Previsto (dias)")
    plt.title(titulo)
    plt.grid(True, linestyle='--', alpha=0.5)
    plt.legend()
    plt.axis('square')
    plt.xlim(min_val-1, max_val+1)
    plt.ylim(min_val-1, max_val+1)
    plt.tight_layout()
    plt.show()

```

#### Célula 6 – Executar para tomates

```

# Caminho do arquivo de Tomates
tomates_csv = '/content/drive/MyDrive/Colab
Notebooks/Tomates_formatado.csv'

# Prompt de carregamento e execução da MLR Tomates
X_tom, y_tom = carregar_dados(tomates_csv, sensor='NIR')
y_real_ml_r_tom, y_pred_ml_r_tom = executar_ml_r(X_tom, y_tom, nome='MLR -
Tomates')
y_real_pls_tom, y_pred_pls_tom = executar_plsr(X_tom, y_tom, nome='PLSR
- Tomates', n_componentes=6)

# Gráficos
plotar_real_vs_predito(y_real_ml_r_tom, y_pred_ml_r_tom, titulo="MLR -
Tomates", cor='royalblue')

```

```
plotar_real_vs_predito(y_real_pls_tom, y_pred_pls_tom, titulo="PLSR - Tomates", cor='green')
```

#### Célula 7 – Executar para bananas

```
# Caminho do arquivo de Bananas
bananas_csv = '/content/drive/MyDrive/Colab Notebooks/Bananas_formatado.csv'

# Prompt de carregamento e execução da MLR Bananas
X_ban, y_ban = carregar_dados(bananas_csv, sensor='Ambos')
y_real_mlr_ban, y_pred_mlr_ban = executar_mlr(X_ban, y_ban, nome='MLR - Bananas')
y_real_pls_ban, y_pred_pls_ban = executar_plsr(X_ban, y_ban, nome='PLSR - Bananas', n_componentes=15)

# Gráficos
plotar_real_vs_predito(y_real_mlr_ban, y_pred_mlr_ban, titulo="MLR - Bananas", cor='royalblue')
plotar_real_vs_predito(y_real_pls_ban, y_pred_pls_ban, titulo="PLSR - Bananas", cor='green')
```

### Código para a MLP

```
# Importar bibliotecas e montar o Google Drive
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import r2_score, mean_squared_error
from keras.models import Sequential
from keras.layers import Dense
from keras.optimizers import Adam

# Montar o Google Drive
from google.colab import drive
drive.mount('/content/drive')

# Função de carregamento dos dados
def carregar_dados(caminho_csv, sensor='todos'):
    df = pd.read_csv(caminho_csv)

    # Remover espaços e limpar nomes de colunas
    df.columns = df.columns.str.strip()

    # Lista de colunas originais
    colunas_originais = df.columns.tolist()
```

```

    if sensor == 'VIS':
        # Selecionar VIS: 415nm 445nm 480nm 515nm 555nm 590nm 630nm
        680nm NIR(910nm)
        colunas_vis = ['415nm', '445nm', '480nm', '515nm', '555nm',
'590nm', '630nm', '680nm', 'NIR(910nm)']
        # Garantir que as colunas existem no DataFrame
        colunas_selecionadas = [col for col in colunas_vis if col in
colunas_originais]
        X = df[colunas_selecionadas]
    elif sensor == 'NIR':
        # Selecionar NIR: 610nm 680nm_1 730nm 760nm 810nm 860nm
        colunas_nir = ['610nm', '680nm_1', '730nm', '760nm', '810nm',
'860nm']
        # Garantir que as colunas existem no DataFrame
        colunas_selecionadas = [col for col in colunas_nir if col in
colunas_originais]
        X = df[colunas_selecionadas]
    else:
        # Todas as colunas espectrais (exceto Exemplar e Dia)
        colunas_espectrais = [col for col in colunas_originais if col
not in ['Exemplar', 'Dia']]
        X = df[colunas_espectrais]

    y = df['Dia']
    print(f"Colunas selecionadas para {sensor}: {X.columns.tolist()}")

    return X, y

# Ajuste dos parâmetros da rede
epocas = 300
percentual_treino = 0.7
percentual_val_teste = 0.3
camadas_ocultas = [20]
funcao_ativacao = 'relu'
taxa_aprendizado = 0.001

# Função para treinar e avaliar o modelo MLP
def treinar_e_avaliar(X, y, descricao):
    # 1. DIVIDIR OS DADOS PRIMEIRO
    X_train, X_temp, y_train, y_temp = train_test_split(X, y,
test_size=percentual_val_teste, random_state=42)
    X_val, X_test, y_val, y_test = train_test_split(X_temp, y_temp,
test_size=0.5, random_state=42)

    # 2. CRIAR E TREINAR O SCALER APENAS COM OS DADOS DE TREINO
    scaler = StandardScaler()
    X_train_scaled = scaler.fit_transform(X_train)

```

```

# 3. USAR O MESMO SCALER (JÁ TREINADO) PARA TRANSFORMAR A VALIDAÇÃO
E O TESTE
X_val_scaled = scaler.transform(X_val)
X_test_scaled = scaler.transform(X_test)

# Criar a rede
model = Sequential()
model.add(Dense(camadas_ocultas[0],
input_shape=(X_train_scaled.shape[1],), activation=funcao_ativacao))
if len(camadas_ocultas) > 1:
    for n_neuronios in camadas_ocultas[1:]:
        model.add(Dense(n_neuronios, activation=funcao_ativacao))
model.add(Dense(1, activation='linear'))

optimizer = Adam(learning_rate=taxa_aprendizado)
model.compile(optimizer=optimizer, loss='mse', metrics=['mae'])

# Treinar (usando os dados escalados)
historico = model.fit(X_train_scaled, y_train,
validation_data=(X_val_scaled, y_val), epochs=epocas, verbose=1)

# Avaliar no TESTE (usando os dados escalados)
y_pred = model.predict(X_test_scaled).flatten()
mse_teste = mean_squared_error(y_test, y_pred)
rmse_teste = np.sqrt(mse_teste)
mae_teste = np.mean(np.abs(y_test - y_pred))
r2_teste = r2_score(y_test, y_pred)

# Avaliar no TREINO (usando os dados escalados)
y_pred_treino = model.predict(X_train_scaled).flatten()
mse_treino = mean_squared_error(y_train, y_pred_treino)
rmse_treino = np.sqrt(mse_treino)
mae_treino = np.mean(np.abs(y_train - y_pred_treino))
r2_treino = r2_score(y_train, y_pred_treino)

# Avaliar na VALIDAÇÃO (usando os dados escalados)
y_pred_val = model.predict(X_val_scaled).flatten()
mse_val = mean_squared_error(y_val, y_pred_val)
rmse_val = np.sqrt(mse_val)
mae_val = np.mean(np.abs(y_val - y_pred_val))
r2_val = r2_score(y_val, y_pred_val)

# Imprimir resultados
print(f"\n{descricao} - Resultados:")
print(f"  Treino:      R² = {r2_treino:.3f}, MSE = {mse_treino:.3f},
RMSE = {rmse_treino:.3f}, MAE = {mae_treino:.3f}")
print(f"  Validação:  R² = {r2_val:.3f}, MSE = {mse_val:.3f}, RMSE
= {rmse_val:.3f}, MAE = {mae_val:.3f}")

```

```
print(f" Teste:      R² = {r2_teste:.3f}, MSE = {mse_teste:.3f},
RMSE = {rmse_teste:.3f}, MAE = {mae_teste:.3f}")
```

```
# Gráficos de histórico
```

```
plt.figure(figsize=(12,5))
plt.subplot(1,2,1)
plt.plot(historico.history['loss'], label='Treino - MSE')
plt.plot(historico.history['val_loss'], label='Validação - MSE')
plt.xlabel('Épocas')
plt.ylabel('MSE')
plt.legend()
```

```
plt.subplot(1,2,2)
plt.plot(historico.history['mae'], label='Treino - MAE')
plt.plot(historico.history['val_mae'], label='Validação - MAE')
plt.xlabel('Épocas')
plt.ylabel('MAE')
plt.legend()
```

```
plt.tight_layout()
plt.show()
```

```
# Chamar a função para plotar o gráfico de dispersão
```

```
plotar_real_vs_predito(y_test, y_pred, titulo=f'Dispersão: Valores
Reais vs. Previstos ({descricao})', cor='red')
```

```
def plotar_real_vs_predito(y_real, y_pred, titulo="Real vs Previsto",
cor='royalblue'):
```

```
    plt.figure(figsize=(6, 6))
    plt.scatter(y_real, y_pred, marker='x', color=cor, alpha=0.7,
label='Amostras')
```

```
# Reta de referência (y = x)
```

```
max_val = max(max(y_real), max(y_pred))
min_val = min(min(y_real), min(y_pred))
plt.plot([min_val-1, max_val+1], [min_val-1, max_val+1],
color='black', linestyle='--', label='y = x')
```

```
plt.xlabel("Valor Real (dias)")
plt.ylabel("Valor Previsto (dias)")
plt.title(titulo)
plt.grid(True, linestyle='--', alpha=0.5)
plt.legend()
plt.axis('square')
plt.xlim(min_val-1, max_val+1)
plt.ylim(min_val-1, max_val+1)
plt.tight_layout()
plt.show()
```

```
# Análise TOMATES - VIS (AS7341)
caminho_tomates = '/content/drive/MyDrive/Colab
Notebooks/Tomates_formatado.csv'
X_vis, y_vis = carregar_dados(caminho_tomates, sensor='VIS')
treinar_e_avaliar(X_vis, y_vis, 'Tomates - VIS')

# Análise TOMATES - NIR (AS7263)
X_nir, y_nir = carregar_dados(caminho_tomates, sensor='NIR')
treinar_e_avaliar(X_nir, y_nir, 'Tomates - NIR')

# Análise TOMATES - Todos os dados combinados
caminho_tomates = '/content/drive/MyDrive/Colab
Notebooks/Tomates_formatado.csv'
X_todos, y_todos = carregar_dados(caminho_tomates, sensor='todos')
treinar_e_avaliar(X_todos, y_todos, 'Tomates - Todos')

# Análise BANANAS - VIS (AS7341)
caminho_bananas = '/content/drive/MyDrive/Colab
Notebooks/Bananas_formatado.csv'
X_vis, y_vis = carregar_dados(caminho_bananas, sensor='VIS')
treinar_e_avaliar(X_vis, y_vis, 'Bananas - VIS')

# Análise BANANAS - NIR (AS7263)
X_nir, y_nir = carregar_dados(caminho_bananas, sensor='NIR')
treinar_e_avaliar(X_nir, y_nir, 'Bananas - NIR')

# Análise BANANAS - Todos os dados combinados
X_todos, y_todos = carregar_dados(caminho_bananas, sensor='todos')
treinar_e_avaliar(X_todos, y_todos, 'Bananas - Todos')
```

## **ANEXO A - Registradores dos sensores espectrais**

## Registadores do AS7341

Addr	Name	<D7>	<D6>	<D5>	<D4>	<D3>	<D2>	<D1>	<D0>
0x60	ASTATUS	ASAT_STATUS					AGAIN_STATUS [3:0]		
0x61	CH0_DATA					CH0_DATA_L [7:0]			
0x62						CH0_DATA_H [7:0]			
0x63	ITIME					ITIME_L [7:0]			
0x64						ITIME_M [7:0]			
0x65						ITIME_H [7:0]			
0x66	CH1_DATA					CH1_DATA_L [7:0]			
0x67						CH1_DATA_H [7:0]			
0x68	CH2_DATA					CH2_DATA_L [7:0]			
0x69						CH2_DATA_H [7:0]			
0x6A	CH3_DATA					CH3_DATA_L [7:0]			
0x6B						CH3_DATA_H [7:0]			
0x6C	CH4_DATA					CH4_DATA_L [7:0]			
0x6D						CH4_DATA_H [7:0]			
0x6E	CH5_DATA					CH5_DATA_L [7:0]			
0x6F						CH5_DATA_H [7:0]			
0x70	CONFIG					LED_SEL	INT_SEL	INT_MODE[1:0]	
0x71	STAT					WAIT_SYNC			READY
0x72	EDGE					SYNC_EDGE [7:0]			
Addr	Name	<D7>	<D6>	<D5>	<D4>	<D3>	<D2>	<D1>	<D0>
0x73	GPIO							PD_GPIO	PD_INT
0x74	LED	LED_ACT			LED_DRIVE [8:0]				
0x80	ENABLE		FDEN		SMUXEN	WEN		SP_EN	PON
0x81	ATIME					ATIME [7:0]			
0x83	WTIME					WTIME [7:0]			
0x84	SP_TH_L					SP_TH_L_LSB [7:0]			
0x85						SP_TH_L_MSB [7:0]			
0x86	SP_TH_H					SP_TH_H_LSB [7:0]			
0x87						SP_TH_H_MSB [7:0]			
0x90	AUXID					AUXID [7:0]			
0x91	REVID					REVID [7:0]			
0x92	ID					ID [7:0]			
0x93	STATUS	ASAT				AINT	FINT	CINT	SINT
0x94	ASTATUS	ASAT_STATUS				AGAIN_STATUS [3:0]			
0x95	CH0_DATA					CH0_DATA_L [7:0]			
0x96						CH0_DATA_H [7:0]			
0x97	CH1_DATA					CH1_DATA_L [7:0]			
0x98						CH1_DATA_H [7:0]			

0x99					CH2_DATA_L [7:0]				
0x9A	CH2_DATA				CH2_DATA_H [7:0]				
0x9B					CH3_DATA_L [7:0]				
0x9C	CH3_DATA				CH3_DATA_H [7:0]				
0x9D					CH4_DATA_L [7:0]				
0x9E	CH4_DATA				CH4_DATA_H [7:0]				
0x9F					CH5_DATA_L [7:0]				
0xA0	CH5_DATA				CH5_DATA_H [7:0]				
0xA3	STATUS 2		AVALID		ASAT_DIG	ASAT_ANA		FDSAT_ANA	FDSAT_DIG
0xA4	STATUS 3			INT_SP_H	INT_SP_L				
0xA6	STATUS 5					SINT_FD			
0xA7	STATUS 6	FIFO_OV		OVTEMP	FD_TRIG		SP_TRIG	SAI_ACT	INT_BUSY
0xA9	CFG 0			LOW_POWER		REG_BANK	WLONG		
0xAA	CFG 1						AGAIN[4:0]		
0xAC	CFG 3				SAI				
0xAF	CFG 6				SMUX_CMD[4:3]				
0xB1	CFG 8	FIFO_TH [7:6]				FD_AGC	SP_AGC		
0xB2	CFG 9		SIEN_FD		SIEN_SMUX				
0xB3	CFG 10	AGC_H [7:6]		AGC_L [7:6]				FD_PERS [2:0]	
0xB5	CFG 12							SP_TH_CH [2:0]	
0xBD	PERS							APERS [3:0]	
0xBE	GPIO 2					GPIO_INV	GPIO_IN	GPIO_OUT	GPIO_IN
0xCA						ASTEP [7:0]			
0xCB	ASTEP					ASTEP [15:8]			
0xCF	AGC_GAIN_MAX			AGC_FD_GAIN_MAX [7:4]				AGC_AGAIN_MAX [3:0]	
0xD6	AZ_CONFIG					AT_NTH_ITERATION [7:0]			
0xD8	FD_TIME 1					FD_TIME [7:0]			
0xDA	FD_TIME 2			FD_GAIN [7:3]				FD_TIME [10:8]	
0xD7	FD_CFG0	FD_FIFO							
0xDB	FD_STATUS			FD_VALID	FD_SAT	FD_120HZ_VALID	FD_100Hz_VALID	FD_120Hz	FD_100Hz
0xF9	INTENAB	ASIEN				SP_IEN	FIEN	CIEN	SIEN
0xFA	CONTROL						AZ_SP_MAN	FIFO_CLR	CLEAR_SAI_ACT
0xFC	FIFO_MAP					FIFO_WRITE_CH5_DATA – FIFO_WRITE_CH0_DATA [6:1]			ASTATUS
0xFD	FIFO_LVL	FIFO_LVL [7:0]							
0xFE		FDATA [7:0]							
0xFF	FDATA								FDATA [15:8]

Fonte: ams OSRAM (2021)

### Registadores do AS7263

Addr	Name	<D7>	<D6>	<D5>	<D4>	<D3>	<D2>	<D1>	<D0>
<b>Version Registers</b>									
0x00: 0x01	HW_Version	Hardware Version							
0x02: 0x03	FW_Version	Firmware Version							
<b>Control Registers</b>									
0x04	Control_Setup	RST	INT	GAIN	Bank	DATA_RDY	RSVD		
0x05	INT_T	Integration Time							
0x06	Device_Temp	Device Temperature							
0x07	LED_Control	RSVD	ICL_DRV	LED_DRV	ICL_IND	LED_IND			
<b>Sensor Raw Data Registers</b>									
0x08	R_High	Channel R High Data Byte							
0x09	R_Low	Channel R Low Data Byte							
0x0A	S_High	Channel S High Data Byte							
0x0B	S_Low	Channel S Low Data Byte							
0x0C	T_High	Channel T High Data Byte							
0x0D	T_Low	Channel T Low Data Byte							
0x0E	U_High	Channel U High Data Byte							
0x0F	U_Low	Channel U Low Data Byte							
0x10	V_High	Channel V High Data Byte							
0x11	V_Low	Channel V Low Data Byte							
0x12	W_High	Channel W High Data Byte							
0x13	W_Low	Channel W Low Data Byte							
<b>Sensor Calibrated Data Registers</b>									
0x14: 0x17	R_Cal	Channel R Calibrated Data (float)							
0x18: 0x1B	S_Cal	Channel S Calibrated Data (float)							
0x1C: 0x1F	T_Cal	Channel T Calibrated Data (float)							
0x20: 0x23	U_Cal	Channel U Calibrated Data (float)							
0x24: 0x27	V_Cal	Channel V Calibrated Data (float)							
0x28: 0x2B	W_Cal	Channel W Calibrated Data (float)							

Fonte: ams OSRAM (2021)