



Pedro Tiago Relvão Figueiredo

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

Coimbra, outubro de 2023



Pedro Tiago Relvão Figueiredo

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

Trabalho de Projeto submetido ao Instituto Superior de Contabilidade e Administração de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de **Mestre em Análise de Dados e Sistemas de Apoio à Decisão** realizado sob a orientação do Professor Doutor André Rodrigues e coorientador Professor Doutor Kiril Bahcevandziev.

Coimbra, outubro de 2023

TERMO DE RESPONSABILIDADE

Declaro ser o autor deste projeto, que constitui um trabalho original e inédito, que nunca foi submetido a outra Instituição de ensino superior para obtenção de um grau académico ou outra habilitação. Atesto ainda que todas as citações estão devidamente identificadas e que tenho consciência de que o plágio constitui uma grave falta de ética, que poderá resultar na anulação do projeto.

AGRADECIMENTOS

Queria agradecer primeiro à minha família. À minha mãe Helena Relvão, por ter me dado o mundo. Ao meu pai, Paulo Figueiredo por ter ajudado. E ao meu irmão, Ivo Figueiredo por ter crescido com ele e aprendendo muito. Sem vocês não estava aqui, e por isso estou grato. À minha querida avó, Dona Conceição, por estar sempre a abençoar-me, mesmo nos dias em que eu não a vejo.

Um agradecimento especial aos meus orientadores cuja orientação, conhecimento e colaboração foram fundamentais para o desenvolvimento deste projeto.

Ao pessoal do gabinete de informática por me receber sempre bem e ter disponibilizado um computador para o projeto, mesmo que essa componente não tenha vido ao encontro do meu projeto.

À minha namorada Jéssica, a quem eu amo muito e ter-me deixado ter instalado o sensor no seu apartamento. Agora já podes fechar a janela (foi só durante o verão).

Um obrigado ao Paulo Ribeiro do Observatório Geofísico e Astronómico da Universidade de Coimbra, que partilhou os dados climáticos históricos comigo. A sua partilha teve um impacto positivo no progresso do projeto.

A todos os meus amigos e familiares que eu gosto muito, também estou grato por partilhar o meu tempo convosco.

RESUMO

Este projeto teve como objetivo principal desenvolver um sistema de recomendação de datas de plantação de culturas hortícolas, com a intenção de auxiliar os agricultores na tomada de decisões, tendo em conta as alterações climáticas, com a finalidade de auxiliar os agricultores a tirarem mais proveito da produtividade das colheitas.

Para atingir este objetivo, foram exploradas técnicas de deep learning, com destaque para o modelo Long Short Term Memory (LSTM), para prever as temperaturas e avaliá-las face aos intervalos ótimos de cada fase de crescimento para cada cultura.

O sistema de previsão foi implementado numa aplicação web que fornece recomendações de datas de plantação com base na temperatura média diária. Além disso, a aplicação oferece uma visualização prática dos dados climáticos e do solo mais favoráveis para um crescimento saudável das culturas.

Comparando os resultados do sistema com calendários agrícolas, observou-se que o modelo oferece uma visão mais dinâmica, adaptável e concreta. O sistema é capaz de considerar as alterações climáticas como um fator importante para a recomendação de uma data de plantação. A realização deste projeto oferece uma ferramenta robusta para agricultores otimizarem as suas práticas agrícolas.

Palavras-chave: Inteligência Artificial, *Deep Learning*, Alterações Climáticas, Culturas Hortícolas, Parâmetros Climáticos, Intervalos Ótimos de Temperatura, Sistemas de Recomendações, Recomendações Algorítmicas, Datas de Plantação.

ABSTRACT

The main aim of this project was to develop a system for recommending planting dates for horticultural crops, with the aim of helping farmers to make decisions taking climate change into account, to help farmers get more out of their crop yields.

To achieve this goal, deep learning techniques were explored, with emphasis on the Long Short Term Memory (LSTM) model, to predict temperatures and evaluate them against the optimal intervals for each growth phase for each crop.

The forecasting system was implemented in a web application that provides recommendations for planting dates based on the average daily temperature. In addition, the application offers a practical visualization of the most favourable climate and soil data for healthy crop growth.

Comparing the results of the system with agricultural calendars, it was observed that the model offers a more dynamic, adaptable, and concrete vision. The system can take climate change into account as an important factor in recommending a planting date. This project offers a robust tool for farmers to optimize their agricultural practices.

Keywords: Artificial Intelligence, Deep Learning, Climate Change, Horticultural Crops, Climate Parameters, Optimal Temperature Intervals, Algorithmic Recommendations, Recommendation System.

ÍNDICE GERAL

1	INTRODUÇÃO	1
1.1.	Contexto	1
1.2.	Motivação	3
1.3.	Objetivos	3
1.4.	Organização do Documento	4
2	ESTADO DA ARTE	5
2.1	Alterações Climáticas e o Impacto na Agricultura	5
2.2	Impacto das Temperaturas no Desenvolvimento de uma Cultura	6
2.3	Inteligência Artificial - Enquadramento	7
2.4	Redes Neurais e Deep Learning	9
2.5	Inteligência Artificial na Agricultura	10
2.6	Trabalhos Relacionados	11
3	METODOLOGIA	18
3.1	Tecnologias	18
3.1.1	Django	18
3.1.2	Plotly	19
3.1.3	Python para Análise de Dados	20
3.1.4	Tensorflow	20
3.1.5	PostgreSQL	20
3.1.6	LoRaWAN	21
3.1.7	TTN (The Things Network)	21
3.1.8	ChatGPT	22

3.1.9	GitHub.....	22
3.1.10	Webhooks.....	23
3.1.11	Visual Paradigm	23
3.1.12	Google Colab.....	23
3.1.13	Postman	23
3.1.14	Ngrok	24
3.2	Hardware	24
3.2.1	Sensor.....	24
3.2.2	Placa Gráfica para Treino dos Modelos.....	25
3.2.3	Servidor Local e Computador Pessoal.....	25
3.3	Métrica de avaliação dos modelos.....	26
3.4	Eliminação de <i>Outliers</i>	27
3.5	Metodologia para a Escolha do Melhor Modelo Baseados em Deep Learning	27
3.5.1	Escolher um modelo base.....	28
3.5.2	Escolher um modelo de referência.	29
3.5.3	Arquiteturas Usadas.....	30
3.5.4	Janelas Temporais, Input e Output, Windows e Horizons	31
3.5.5	Seleção das Variáveis de Entrada.....	32
3.5.6	Escolha dos Hiperparâmetros.....	33
3.5.7	Layers nos Modelos de Deep Learning	33
3.6	Implementação de um Modelo	34
3.6.1	Construir o Modelo.....	34
3.6.2	Compilar o Modelo.....	34
3.6.3	Treinar o Modelo	35

3.6.4	Avaliar o Modelo.....	35
4	Conjunto de Dados.....	36
4.1	Recolha dos Dados.....	36
4.1.1	Dados Climáticos.....	36
4.1.2	Dados das colheitas das Culturas na Região de Coimbra	37
4.1.3	Dados das culturas hortícolas	38
4.2	Processamento dos dados	38
4.2.1	Junção dos Dados Diários Climáticos	38
4.2.2	Limpeza dos Dados.....	39
4.2.3	Adicionar Dados	40
4.2.4	Divisão dos Dados	41
4.3	Análise Exploratória dos Dados	42
4.3.1	Temperatura e Precipitação.....	42
4.3.2	Tendências da Temperatura e da Precipitação	44
4.3.3	Análise das Características Ambientais da Cultura da Batata.....	46
5	Desenvolvimento	55
5.1	Caraterísticas do sistema.....	55
5.2	Tipos dos Utilizadores	57
5.3	Requisitos Funcionais	58
5.4	Requisitos Não Funcionais.....	75
5.5	Design do User Interface.....	76
5.4.	Arquitetura do Sistema.....	80
5.6	Aplicação Web	81
5.6.1	Aplicações no Django	81

5.6.2	Autenticação, Login e Registro	82
5.6.3	Painel de Administrador e Gestão de Dados	84
5.6.4	Visualização de Dados com Plotly e Bootstrap no Frontend	86
5.7	Restantes informações acerca das culturas.....	89
5.8	Integração do Modelo com a Aplicação Web	90
5.8.1	Armazenamento do modelo	90
5.8.2	Preparação dos dados para o modelo.....	90
5.8.3	Implementação do Algoritmo de Recomendação da Data de Plantação	91
5.9	Integração com dispositivos IoT.....	94
6	Resultados e Discussão.....	96
6.1	Avaliação dos Modelos	96
6.1.1	Resultados dos Modelos com diferentes arquiteturas.....	96
6.1.2	Resultados dos Modelos com Diferentes Janelas de Temporais	97
6.1.3	Resultados dos Modelos com Variáveis	99
6.1.4	Resultados dos Modelos com Layers e Hiperparâmetros Diferentes	100
6.2	Testes da Aplicação	101
6.2.1	Utilização do Ngrok para Testar a Aplicação Web	101
6.2.2	Uso do Postman para Testar os Algoritmos	102
6.3	Resultado das Recomendações de Datas de Plantação	103
6.4	Limitações	108
6.4.1	Modelo de Previsão Único	108
6.4.2	Tamanho da window.....	108
6.4.3	Utilização de um sensor como representativo uma Quinta Agrícola	108
7	Conclusão e Trabalho Futuro.....	110

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

7.1	Conclusão	110
7.2	Trabalho Futuro	110
	REFERÊNCIAS BIBLIOGRÁFICAS.....	113
	ANEXOS.....	118

ÍNDICE DE FIGURAS

<i>Figura 1 - Inteligência Artificial e as suas Subcategorias</i>	8
<i>Figura 2 - Rede Neural Artificial</i>	9
<i>Figura 3 - Arquitetura do Django</i>	19
<i>Figura 4 - Bibliotecas de Python para Análise de Dados</i>	20
<i>Figura 5 - Arquitetura do TTN. Fonte: thethingsnetwork.org</i>	21
<i>Figura 6- Sensor Dragino S-31</i>	24
<i>Figura 7 - Placa Gráfica NVIDIA A100</i>	25
<i>Figura 8 - Previsão do Persistence Model aplicada aos dados da Temperatura</i>	29
<i>Figura 9 – Localização do Sensor</i>	37
<i>Figura 10 - Conjunto de Dados para Treinar e Testar Modelos</i>	41
<i>Figura 11 - Média da Temperatura Máxima e Mínima por Mês em Coimbra</i>	43
<i>Figura 12 - Temperatura Média Mensal em Coimbra</i>	44
<i>Figura 13 - Regressão Linear da Temperatura Média Agrupada de 30 em 30 anos</i>	45
<i>Figura 14 - Regressão Linear da Precipitação Média Agrupada de 30 em 30 anos</i>	46
<i>Figura 15 - Gráfico de Colheita Anual da Batata em Coimbra</i>	47
<i>Figura 16 - Intervalo das Temperaturas Ótimas para cada Fase da Batata</i>	49
<i>Figura 17 - Intervalos Ótimos face à Temperatura Média Mensal da Batata</i>	50
<i>Figura 18 - Registo das Temperaturas Médias Diárias face aos Intervalos da Batata</i>	53
<i>Figura 19 - Diagrama de Pacotes</i>	55
<i>Figura 20 - Diagrama de Caso de Uso de Utilizador sem Conta</i>	59
<i>Figura 21 - Diagrama Caso de Uso Administrador</i>	63
<i>Figura 22 - Diagrama de Caso de Uso do Utilizador Registado</i>	68
<i>Figura 23 – Mockup Página Inicial</i>	77

<i>Figura 24 - Página 1 do Dashboard.....</i>	<i>78</i>
<i>Figura 25 - Página 2 do Dashboard.....</i>	<i>79</i>
<i>Figura 26 - Arquitetura do Sistema.....</i>	<i>81</i>
<i>Figura 27 - Página de Registrar Utilizador.....</i>	<i>83</i>
<i>Figura 28 - Página de Login de Utilizador.....</i>	<i>83</i>
<i>Figura 29 - Página Inicial do Administrador.....</i>	<i>85</i>
<i>Figura 30 - Página do Administrador para Inerir Dados Novos das Culturas.....</i>	<i>85</i>
<i>Figura 31 - Gráfico de Temperaturas Ótimas e Gráfico de Época de Plantação.....</i>	<i>87</i>
<i>Figura 32 - Gráfico de Precipitação (esquerda) e Gráfico do pH do Solo (direita).....</i>	<i>87</i>
<i>Figura 33 - Gráfico dos Resultados do Algoritmo.....</i>	<i>87</i>
<i>Figura 34 - Gráfico da Temperatura Média Máxima e Mínima Diária.....</i>	<i>88</i>
<i>Figura 35 - Gráfico da Previsão da Temperatura.....</i>	<i>88</i>
<i>Figura 36 - Informações gerais na página da cultura.....</i>	<i>89</i>
<i>Figura 37 - Página da lista de culturas.....</i>	<i>90</i>
<i>Figura 38 - Diagrama de Atividades do Algoritmo de Recomendação.....</i>	<i>93</i>
<i>Figura 39 - Resultados do Teste GET.....</i>	<i>102</i>
<i>Figura 47 - Diagrama de Classes da Cultura.....</i>	<i>119</i>
<i>Figura 48 - Época de Plantação para a Batata.....</i>	<i>120</i>
<i>Figura 49 - Recomendação do Algoritmo para a Batata.....</i>	<i>120</i>
<i>Figura 50 - Época de Plantação para o Tomate.....</i>	<i>121</i>
<i>Figura 51 - Recomendação do Algoritmo para o Tomate.....</i>	<i>121</i>
<i>Figura 52 - Época de Plantação para o Feijão Verde. Fonte.....</i>	<i>122</i>
<i>Figura 53 - Recomendação do Algoritmo para o Feijão Verde.....</i>	<i>122</i>

ÍNDICE DE TABELAS

<i>Tabela 1 - Tabela referindo a temperatura e a humidade do ar, pH e insolação.....</i>	<i>13</i>
<i>Tabela 2 - Dados do EcoCrop referente à batata e cenoura</i>	<i>15</i>
<i>Tabela 3 - Temperaturas Cardinais para a Batata (°C).....</i>	<i>48</i>
<i>Tabela 4 - Resultados do MAE para o Mês de fevereiro como o de plantação</i>	<i>51</i>
<i>Tabela 5 - Resultados do MAE para o Mês de março como o de plantação</i>	<i>51</i>
<i>Tabela 6 - Resultados do MAE para o Mês de abril como o de plantação</i>	<i>52</i>
<i>Tabela 7 - Resultados do MAE em relação à Temperatura Média Diária</i>	<i>54</i>
<i>Tabela 8 - Atores e os seus respetivos Casos de Uso</i>	<i>57</i>
<i>Tabela 9 - Cenário de Registrar Utilizador.....</i>	<i>60</i>
<i>Tabela 10 – Requisitos Funcionais do Cenário Registrar Utilizador.....</i>	<i>60</i>
<i>Tabela 11 - Cenário Pesquisar Culturas</i>	<i>61</i>
<i>Tabela 12 - Requisitos Funcionais do Cenário Pesquisar Culturas</i>	<i>61</i>
<i>Tabela 13 - Cenário Visualizar Dados e Recomendações de Culturas</i>	<i>62</i>
<i>Tabela 14 - Requisitos Funcionais do Cenário Visualizar Dados</i>	<i>62</i>
<i>Tabela 15 - Cenário Adicionar Cultura.....</i>	<i>64</i>
<i>Tabela 16 - Requisitos do Cenário Adicionar Cultura</i>	<i>64</i>
<i>Tabela 17 - Cenário Adicionar Variáveis Ótimas da Cultura</i>	<i>65</i>
<i>Tabela 18 - Requisitos do Cenário Adicionar Variáveis Ótimas</i>	<i>65</i>
<i>Tabela 19 - Cenário Editar Cultura</i>	<i>66</i>
<i>Tabela 20 - Requisitos Funcionais do Cenário Editar Cultura</i>	<i>66</i>
<i>Tabela 21 - Cenário Eliminar Culturas.....</i>	<i>67</i>
<i>Tabela 22 - Requisitos Funcionais do Cenário Eliminar Cultura</i>	<i>67</i>

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

<i>Tabela 23 - Cenário Fazer Login</i>	<i>69</i>
<i>Tabela 24 – Requisitos Funcionais do Cenário Fazer Login</i>	<i>69</i>
<i>Tabela 25 - Cenário Fazer Logout</i>	<i>69</i>
<i>Tabela 26 - Requisitos Funcionais do Cenário Logout</i>	<i>70</i>
<i>Tabela 27 - Cenário Editar Perfil.....</i>	<i>70</i>
<i>Tabela 28 - Requisitos Funcionais do Cenário Editar Perfil.....</i>	<i>71</i>
<i>Tabela 29 - Cenário Gerir Dados da Quinta</i>	<i>71</i>
<i>Tabela 30 - Requisitos Funcionais do Cenário Gerir Dados da Quinta</i>	<i>72</i>
<i>Tabela 31 - Cenário Alterar Senha.....</i>	<i>72</i>
<i>Tabela 32 - Requisitos Funcionais do Cenário Alterar Senha</i>	<i>73</i>
<i>Tabela 33 -Cenário Visualizar Dados da Quinta.....</i>	<i>73</i>
<i>Tabela 34 - Requisitos do Cenário Visualizar Dados da Quinta</i>	<i>74</i>
<i>Tabela 35 - Cenário Visualizar Dados de Previsão de Temperatura</i>	<i>74</i>
<i>Tabela 36 - Requisitos Funcionais do Cenário Visualizar Dados de Previsão</i>	<i>74</i>
<i>Tabela 37 - Requisitos Não Funcionais.....</i>	<i>75</i>
<i>Tabela 38 - Resultado dos Modelos de Janelas Temporais Diferentes</i>	<i>98</i>
<i>Tabela 39 - Resultado dos Modelos com Variáveis Diferentes</i>	<i>99</i>
<i>Tabela 40 - Resultado dos Modelos com Layers e Hiperparâmetros Diferentes.....</i>	<i>100</i>
<i>Tabela 41 - Heatmap de Plantação para o Melão</i>	<i>104</i>
<i>Tabela 42 - Resultados da comparação da recomendação do livro com a do sistema.....</i>	<i>105</i>
<i>Tabela 43 - Resultado do Melhor Dia de Plantação</i>	<i>107</i>

Lista de abreviaturas, acrónimos e siglas

IA - Artificial Intelligence

ML - Machine Learning

DL - Deep Learning

CNN - Convolutional Neural Networks

RNN - Recurrent Neural Networks

IoT - Internet of Things

TTN - The Things Network

IQR - Amplitude Interquartil

UML - Unified Modeling Language

API - Application Programming Interface

URL - Uniform Resource Locator

CRUD - Create, Read, Update, and Delete

MAE - Mean Absolute Error

IPMA - Instituto Português do Mar e Atmosfera

JSON - JavaScript Object Notation

RO - Ribatejo e Oeste

EDM – Entre Douro e Minho

BL – Beira Litoral

1 INTRODUÇÃO

1.1. Contexto

As alterações climáticas estão-se a tornar cada vez mais alarmantes e evidentes no que respeita ao seu impacto no meio ambiente e nas nossas vidas. Quer sejam as temperaturas mais elevadas do que o normal, ondas de calor cada vez mais graves e longas em várias zonas da Europa ou inundações alarmantes na Ásia Ocidental. Estes acontecimentos imprevisíveis estão a moldar o clima em que vivemos. No caso de Portugal, é a ocorrência de secas severas cada vez mais graves e prolongadas, com a Liga para a Proteção da Natureza a indicar que as secas serão o “novo normal” no nosso país [1].

Enquanto Portugal e a zona mediterrânea passaram pela pior seca dos últimos 90 anos [2] em 2022, devido à falta de precipitação durante o inverno e primavera, ao mesmo tempo, a Bulgária estava a passar por um dos verões mais frios da sua história [3]. Estamos perante um ambiente climático instável, aliás de acordo com o Copernicus Climate Change Service o ano de 2019 na Europa foi o mais quente alguma vez registado, registando 1.24°C acima da temperatura média anual na Europa [4]. Há pouco tempo, no dia 6 de julho de 2023, foi registado o dia mais quente de sempre no planeta Terra, devido ao fenómeno climático El Niño [5]. E para acrescentar que o último dia de Verão, no passado dia 21 de setembro, confirmou-se que este Verão foi o mais quente de sempre no hemisfério norte [6].

Ao falarmos de secas, é inevitável não falar da sua consequência mais perigosa, que é a escassez de água. À medida que escrevo este documento, Portugal está ainda a sofrer as consequências da seca histórica registada neste verão, com o anúncio da suspensão do uso de água nas albufeiras, para fins de produção de energia hidroelétrica [7]. Os rios, as barragens e as albufeiras encontram-se em capacidade mínima de água e os agricultores estão muito preocupados com essa situação.

Numa investigação efetuada pela Fundação Calouste Gulbenkian em 2021, conclui-se que 75% da água usada em Portugal, é dedicada à agricultura. Um número que contrasta com a média da União Europeia (24%) e que chega a ser superior à média mundial (69%), no entanto, alinhado com os países mediterrânicos como a Espanha e a Grécia. Sendo Portugal um país mediterrâneo, a sua agricultura é caracterizada como de regadio, ao contrário de países como o Reino Unido ou a Alemanha em que a sua agricultura é regada maioritariamente pela chuva [8].

De acordo com a Agência Portuguesa do Ambiente desses 75% da água usada, mais de um terço é desperdiçado [9]. Os responsáveis por este desperdício são os sistemas de regadio antiquados e a escolha e a má planificação das espécies das culturas (inapropriadas para as condições do cultivo). Um dos casos mais notórios em Portugal é a plantação de abacateiros no Algarve. O abacate, pelas suas origens tropicais, é um fruto que necessita de muita água, mas o Algarve, pelas suas propriedades geográficas, é uma região com pouca chuva em relação ao resto do país. A região ocupada pela produção de abacates no Algarve já ultrapassa 7% de todo o território algarvio [10].

Numa perspetiva de negócio, os agricultores estão à procura de novos métodos para melhorar a eficiência da produção de culturas hortícolas, de modo a reduzir os custos e aumentar os lucros. Por isso o agricultor tem de tomar a decisão mais acertada acerca do que plantar no seu terreno e quando o plantar. Enquanto no passado a tomada de decisão era baseada na experiência ou na intuição, atualmente existem dados suficientes para se desenvolver um modelo inteligente que seja capaz de se adaptar rapidamente às alterações climáticas e que permitam indicar, para uma de cultura, quando é mais adequado plantar num determinado terreno considerando, para além dos elementos climáticos, também as características do solo.

Portugal é um país em que a maioria da prática agrícola extensiva envolve as culturas tradicionais de sequeiro. A agricultura de sequeiro é o cultivo sem rega em regiões onde a precipitação anual é inferior a 500mm por metro quadrado. Depende de técnicas de cultivo específicas, que permitem um uso eficaz e eficiente da humidade do solo. As plantas mais cultivadas e que ocupam a maior área são as oliveiras, as vinhas e os sobreiros [11].

1.2. Motivação

Estes problemas, motivaram-me para desenvolver um projeto, com o apoio da IA (inteligência artificial), no âmbito da promoção de uma prática agrícola mais sustentável e produtiva. Seria algo na área dos sistemas de recomendação, em concreto um sistema de recomendação para indicar a melhor altura para plantação de uma determinada cultura. Todos os anos existem variações dos parâmetros climáticos das diferentes estações do ano em relação ao esperado, umas mais secas, outras mais chuvosas, por isso não faz sentido estar sempre a plantar as mesmas hortícolas no mesmo intervalo de tempo. O objetivo principal deste projeto é auxiliar o agricultor, quer seja em pequena ou grande escala, a tomar a melhor decisão para garantir uma produção e colheita de qualidade, prevenir o desperdício do recurso mais importante do nosso planeta, a água, e ao mesmo tempo aumentar o desempenho económico desta indústria.

Se o clima em Portugal tem vindo a mudar nos últimos anos, não faz sentido que os agricultores mantenham as tradições relativas à época de plantação.

1.3. Objetivos

Portugal passou no verão de 2022 pela pior seca dos últimos 90 anos. Esta seca começou a ser prevista no início deste ano quando o Instituto Português do Mar e da Atmosfera efetuou a pior previsão relativamente ao mês de janeiro. Em fevereiro previu-se que 90% dos terrenos em Portugal estariam a passar por fases de seca severa ou extrema [12]. Ou seja, no início do ano já havia fortes indicações que o ano de 2022 ia ser extremamente seco.

Mas o oposto também pode acontecer. No futuro pode haver meses mais prolongados com climas mais frios e chuvosos. Neste caso, o sistema desenvolvido irá recomendar épocas de plantação diferentes, em que as culturas estarão melhor adaptadas a estas alterações climáticas.

As circunstâncias climáticas e a sua variabilidade são dos fatores mais importantes para decidir se uma cultura é sustentável ou não [13].

Apesar de Portugal ser um país pequeno, o clima e a sua característica pedológica variam de zona para zona. Por isso é pertinente incluir dados referentes à zona de Coimbra, como fatores importantes para a escolha da hortícola mais adequada.

Por isso o objetivo com este projeto é criar uma aplicação web, que permite auxiliar os agricultores na escolha da melhor data de plantação para uma variedade de culturas. Com este sistema os agricultores vão, com o suporte de técnicas IA, ter um auxiliar precioso que conjuga no processo de recomendação, as características das culturas e a meteorologia, procurando evitar o impacto que as alterações climáticas possam ter no desenvolvimento de uma cultura.

1.4. Organização do Documento

O documento segue uma estrutura formal, que é utilizada na elaboração de relatórios de mestrado. Abaixo são mencionados cada capítulo e uma breve descrição de cada um.

Introdução - Uma visão geral do tema de investigação e do seu contexto.

Estado da Arte - Uma análise da investigação existente relevante para o tema da tese.

Metodologia - Detalhes sobre os métodos e instrumentos de investigação utilizados.

Conjunto de Dados- Apresentação e análise dos dados recolhidos.

Desenvolvimento – Explicação dos métodos e das plataformas desenvolvidas.

Resultados e Discussão - Apresentação dos resultados da investigação, a sua interpretação e discussão dos resultados no contexto da literatura existente.

Conclusão - Síntese da investigação e sugestão do trabalho futuro.

2 ESTADO DA ARTE

2.1 Alterações Climáticas e o Impacto na Agricultura

As implicações das alterações climáticas na agricultura mundial tornaram-se cada vez mais evidentes nos últimos anos. À medida que o planeta sofre alterações na temperatura, nos padrões de precipitação, e na frequência de eventos climáticos extremos, o setor agrícola enfrenta desafios, exigindo adaptações inovadoras e urgentes..

Nos países em desenvolvimento, onde a agricultura é a base da economia, tem-se observado que o aumento das temperaturas afeta negativamente o rendimento das culturas [14]. Estas regiões estão a assistir a uma redução da produtividade agrícola, tornando as pequenas e grandes explorações agrícolas cada vez mais vulneráveis. Esta vulnerabilidade não se limita apenas ao rendimento das culturas. A saúde do solo, um componente crítico para a produtividade agrícola, também está ameaçada. Verificou-se que as alterações climáticas têm impacto na população microbiana do solo, o que, por sua vez, afeta as suas atividades enzimáticas [15]. Estes microrganismos desempenham um papel indispensável no ciclo dos nutrientes, na decomposição da matéria orgânica e na manutenção da estrutura do solo.

As alterações climáticas antropogénicas reduziram a produtividade total dos fatores agrícolas a nível mundial em cerca de 21% desde 1961, um abrandamento que equivale à perda dos últimos 7 anos de crescimento da produtividade. O efeito é substancialmente mais grave em regiões mais quentes, como África, América Latina e Caraíbas [16].

Além disso, as alterações climáticas têm implicações mais vastas para a utilização dos solos na agricultura. À medida que a população mundial continua a crescer, a procura de alimentos aumenta. Contudo, as alterações climáticas colocam desafios significativos em termos de disponibilidade de terras e da sua aptidão para a agricultura [17]. Esta situação é ainda mais complicada pelas mudanças na distribuição e no comportamento das pragas de insetos e de nemátodos devido às alterações climáticas [18]. Algumas zonas

estão a sofrer pressões crescentes de pragas, o que tem um impacto direto na saúde e no rendimento das culturas.

É muito importante salientar que qualquer modelo que preveja um decréscimo nas colheitas nas terras agrícolas por causa das alterações climáticas globais, viola um princípio económico básico segundo o qual os valores das colheitas das terras agrícolas devem ser positivos ou iguais a zero [19].

Para reduzir os efeitos das alterações climáticas, existe uma constante pressão para descobrir soluções inovadoras para mitigar o seu impacto. Uma dessas abordagens é a utilização de bio estimulantes [20]. Estas substâncias, conhecidas por melhorarem o crescimento e a saúde das plantas, oferecem uma contramedida potencial para alguns dos impactos negativos das alterações climáticas nas culturas.

É para salientar, que a comunidade mundial procura compreender e enfrentar os desafios colocados pelas alterações climáticas na agricultura, a investigação e a inovação continuam a ser fundamentais. Os estudos e as análises neste domínio são cruciais para a definição de estratégias eficazes de adaptação e atenuação para o futuro.

2.2 Impacto das Temperaturas no Desenvolvimento de uma Cultura

Um aumento da temperatura média sazonal pode reduzir o cultivo de muitas culturas e, por conseguinte, reduzir o rendimento final. Nas zonas onde as temperaturas já estão próximas dos máximos fisiológicos das culturas, o aquecimento terá um impacto mais imediato nos rendimentos [21].

Apesar do aumento da concentração de CO₂, devido às alterações climáticas, que conduz a um maior crescimento e produtividade das plantas devido ao aumento da fotossíntese, o aumento da temperatura compensa este efeito, uma vez que conduz a um aumento da taxa de respiração e evapotranspiração das culturas, a uma maior infestação de pragas e doenças, a uma alteração da flora infestante e a uma redução da duração das culturas [16].

A temperatura desempenha um papel fundamental e é dos indicadores mais importantes no crescimento e desenvolvimento das culturas. As variações de temperatura podem ter um impacto significativo nos processos fisiológicos e biológicos das plantas, como a fotossíntese, a respiração e as atividades metabólicas. Por exemplo, culturas como o milho e brócolos têm limites específicos de temperatura inferior, superior e ótima para a sua fase de crescimento vegetativo. Qualquer desvio destes limites pode levar ao stress, causando potencialmente uma perturbação no crescimento, uma redução do rendimento ou mesmo a morte da planta [22].

Na China, um estudo sobre a produção de trigo de inverno revelou que a temperatura média ótima e os limiares hídricos eram de 7,3°C e 569 mm, respetivamente. Verificou-se que um aumento da temperatura era desfavorável para a produção de trigo de inverno, ao passo que o aumento do abastecimento de água era benéfico [23].

À medida que as temperaturas globais continuam a aumentar, torna-se fundamental compreender a relação entre a temperatura e a produtividade das culturas. As estratégias de adaptação, como a alteração das datas de plantação, a seleção de variedades de culturas que se adaptam melhor e o recurso a práticas agrícolas avançadas, serão cruciais para garantir a segurança alimentar face às alterações climáticas [24].

2.3 Inteligência Artificial - Enquadramento

A IA é um ramo da ciência da computação que visa criar sistemas capazes de realizar tarefas que ainda requerem, inteligência humana. Estas tarefas incluem, reconhecimento de padrões, aprendizagem, planeamento e tomada de decisões [25]. Desde a sua conceção, a IA tem sido objeto de intensa pesquisa e desenvolvimento, levando à criação de algoritmos e sistemas cada vez mais avançados.

O *machine learning* (ML) é uma subcategoria da IA que se concentra em desenvolver sistemas que podem aprender a partir de dados (Figura 1). Em vez de serem explicitamente programados para realizar uma tarefa, os sistemas de ML são treinados usando grandes conjuntos de dados e algoritmos para fazer previsões ou tomar decisões sem serem explicitamente programados para realizar a tarefa [26].

Dentro do domínio do ML, o deep learning (DL) é uma técnica que utiliza redes neurais com múltiplas camadas (*layers*) para analisar vários níveis de representação e abstração dos dados. O DL tem demonstrado um desempenho notável em tarefas complexas, como reconhecimento de imagem e processamento de linguagem natural, superando outras técnicas de ML em muitos casos [27].

Em resumo, enquanto a IA abrange uma vasta gama de algoritmos e técnicas que imitam a inteligência humana, o ML e o DL são subdomínios especializados que se concentram em aprender a partir de dados. No caso do DL acresce a utilização de modelos com múltiplas camadas para melhor lidar com a complexidade dos dados.

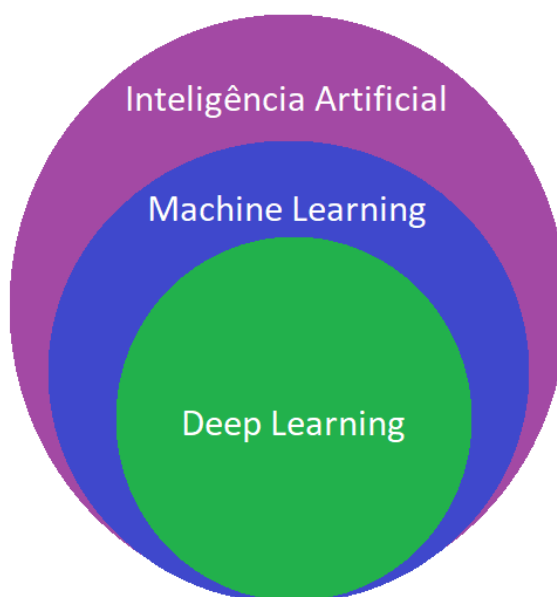


Figura 1 - Inteligência Artificial e as suas Subcategorias

2.4 Redes Neurais e Deep Learning

As redes neurais tentam simular o comportamento do cérebro humano, permitindo-lhe aprender com grandes quantidades de dados.

A DL revolucionou vários domínios da IA, incluindo a visão computacional, o reconhecimento da fala e o processamento de linguagem natural. A principal vantagem dos modelos de DL é a sua capacidade de aprender automaticamente características a partir de dados, sem necessitar de extração manual de características [28].

Uma rede neural típica, como se pode visualizar na Figura 2, é constituída por uma *input layer* (verde), uma ou mais *layers* ocultas (azul) e uma *output layer* (amarelo). Cada *layer* é constituída por *nodes* ou neurónios que estão interligados. A força das interligações é ajustada durante o treino através de um processo designado por retro propagação [29]. Embora uma rede neural com uma única *layer* possa efetuar previsões aproximadas, as *layers* ocultas adicionais podem ajudar a aperfeiçoar essas previsões [27]. As redes neuronais profundas têm frequentemente milhões de parâmetros, o que torna a sua formação computacionalmente intensiva. Também requerem uma quantidade significativa de dados para generalizar corretamente [30].

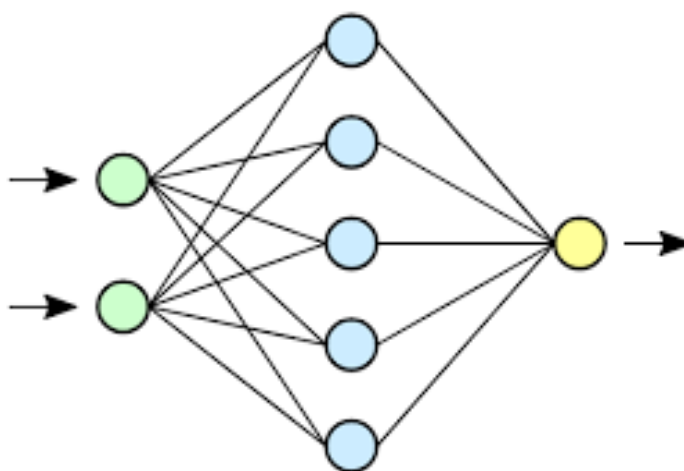


Figura 2 - Rede Neural Artificial

DL tem sido aplicada com sucesso em vários sectores, no reconhecimento de imagens as *convolutional neural networks* (CNN) alcançaram um desempenho destacável em tarefas como a classificação de imagens [31]. Para o processamento de linguagem

natural as *recurrent neural networks* (RNN) são amplamente utilizados em tarefas como a tradução automática e a análise de sentimentos [32]. Os modelos baseados em DL também reduziram significativamente a taxa de erro nos sistemas de reconhecimento de voz [33].

Os modelos baseados em DL tem a capacidade de extrair automaticamente características dos dados, trouxe avanços significativos em vários domínios da IA. À medida que o poder computacional aumenta e mais dados ficam disponíveis, as potenciais aplicações da DL continuarão a crescer.

2.5 Inteligência Artificial na Agricultura

A agricultura, uma das indústrias mais antigas do mundo, está a passar por uma transformação. Com a integração de tecnologias modernas, especialmente a IA, o sector agrícola está a reinventar as suas práticas para responder às exigências de uma população mundial em crescimento e aos desafios colocados pelas alterações climáticas. A IA, com a sua capacidade de analisar grandes quantidades de dados e fazer previsões, está na frente desta modernização [34].

A agricultura de precisão envolve a utilização de tecnologia para melhorar a eficiência e a produtividade das operações agrícolas. As ferramentas baseadas em IA podem analisar dados de imagens de satélite, sensores no campo e previsões meteorológicas para fornecer aos agricultores informações utilizáveis. Isto permite uma rega, fertilização e controlo de pragas mais precisos, garantindo um crescimento ótimo das culturas e minimizando o desperdício de recursos [35].

Os modelos de IA podem ser treinados para reconhecer sintomas de doenças ou pragas nas culturas através da análise de imagens capturadas por *drones* ou câmaras. A deteção precoce permite que os agricultores tomem medidas preventivas antes que uma doença se espalhe, garantindo colheitas mais saudáveis e reduzindo a necessidade de intervenções químicas [36].

Ao analisar dados históricos, condições atuais das culturas e fatores ambientais, os algoritmos de IA podem prever o rendimento das culturas com elevada precisão. As previsões de rendimento exatas ajudam os agricultores a tomarem decisões corretas sobre a colheita e a comercialização, garantindo uma melhor rentabilidade [37].

Os sistemas baseados na IA podem monitorizar a saúde dos solos, a utilização da água e o impacto ambiental das práticas agrícolas. Ao otimizar a utilização dos recursos e ao promover práticas sustentáveis, a IA contribui para a conservação do ambiente e para a viabilidade dos terrenos agrícolas a longo prazo [38].

A integração da IA na agricultura promete um futuro em que a agricultura será mais eficiente, sustentável e capaz de responder às necessidades globais. À medida que a tecnologia continua a avançar, a relação entre a IA e a agricultura aprofundar-se-á, abrindo caminho a soluções inovadoras para os desafios alimentares e ambientais do mundo.

2.6 Trabalhos Relacionados

Um artigo publicado no Indian Journal Science and Technology [39], explica como foi efetuado um sistema de recomendação para cultivo que utiliza informações das características do solo e do clima. O artigo, atribuiu uma especial atenção à alteração do clima e o seu potencial impacto nas futuras plantações. Este projeto é designado pelos autores como Climate-Smart Agriculture. A estratégia seguida foca-se em melhorar a qualidade e a quantidade da colheita. A integração das novas técnicas de análise de dados para o estabelecimento de plantações é algo ainda pouco abordado. Esta integração é uma aposta prometedora no desenvolvimento dos países em que a agricultura tenha um papel importante na economia e pode ajudar a desenvolver outras aplicações como a previsão de colheitas, a identificação de doenças nas culturas ou até a gestão e o apoio no uso de fertilizantes. Também poderá ajudar o governo a criar leis para limitar a prática de plantações de hortícolas que não sejam sustentáveis e para manter o solo saudável e fértil. A técnica utilizada foi as redes neurais artificiais. A recomendação da cultura é baseada em três parâmetros, propriedades do solo, características das culturas hortícolas e fatores

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

climáticos. O sistema classifica a recomendação como devidamente adequada, moderadamente adequada, pouco adequada e não adequada. Este projeto também avaliou técnicas de previsão baseadas em árvores de decisão, mas a rede neural artificial obteve uma melhor performance. No estudo foram analisadas apenas 4 culturas, milho, arroz, cana-de-açúcar e milhete. Estas plantas correspondem às culturas mais cultivadas na Índia. O estudo conclui que numa região específica, o milho e o açúcar são as culturas mais adequadas para cultivar.

No Instituto Tecnológico do Sri Lanka, foi desenvolvido o primeiro sistema de recomendação de culturas no país [39]. O Sri Lanka é um país pequeno com terrenos agrícolas reduzidos, em comparação à sua vizinha, a Índia. Por isso a previsão correta dos cultivos pode ser crucial para o país. Novamente estamos perante um sistema que inclui dispositivos IoT para a aquisição de dados climáticos e do solo. Este sistema considerou todos os fatores principais que são essenciais para o crescimento de uma planta como por exemplo a temperatura, a humidade, o pH do solo, a luz e a humidade do solo. Estes dados em conjunto com outras fontes como os do Departamento da Agricultura do Sri Lanka, permitiram construir uma tabela de culturas em que estão indicados os fatores ambientais ótimos para o bom crescimento das plantas (Tabela 1).

Tabela 1 - Tabela referindo a temperatura e a humidade do ar, pH e insolação. Fonte: (Bandara et al. 2020)

Cultura	Temperatura	Humidade	pH	Índice UV	Humidade do Solo
Feijão	26	63	7	8	63
Alface	23.8	60	6.5	8	60
Cenoura	26	98	6.5	8	98
Couve	29.4	95	7.1	5	95
Beterraba	30	99	6.5	6.5	6.5

Foi nesta tabela que se basearam para criarem o modelo para aumentar a precisão da recomendação. O estudo foi efetuado numa pequena quinta, onde se encontravam os sensores instalados. Após a recolha e o processamento dos dados, aplicaram duas técnicas de aprendizagem automática, a Naive Bayes e a Support Vector Machine. O algoritmo de recomendação analisa quais as melhores culturas para plantar comparando a tabela acima com os valores obtidos pelos sensores. O modelo usou os seguintes dados da quinta: temperatura, humidade, pH do solo, Índice UV e humidade do solo. Como saída, apresenta a cultura recomendada que melhor se adequa às características da quinta. O modelo apresentou uma precisão de 90% e a cultura recomendada foi o feijão verde.

Na Tailândia, no Departamento de Engenharia da Universidade de Kasetsart [40], foi proposto um sistema de recomendação de vários tipos de arroz, com base em regras, para ajudar os agricultores a escolherem a variedade de arroz mais adequada para as suas quintas. O sistema está dividido em duas partes: o sistema de recomendação de vários tipos de arroz e um calendário personalizado para ajudar os agricultores a planear as suas atividades durante o desenvolvimento da cultura recomendada. O arroz é das culturas mais consumidas no mundo, principalmente nos países asiáticos. Na Tailândia são

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

cultivadas mais de 200 variedades de arroz. Cada tipo de arroz tem propriedades diferentes, e isso reflete diferentes adaptações aos diferentes meios ambientes. As variáveis relativas ao arroz, que os autores tiveram em consideração foram, tolerância à seca do solo, nível da sensibilidade de fotoperíodo, pH do solo, salinidade, variância da temperatura, período de crescimento, tolerância de doenças e uso de pesticidas. O estudo inicialmente, categorizou o arroz em 5 categorias, arroz de planície, arroz flutuante, arroz de planalto, arroz inundado em águas profundas e arroz crescido em altitudes superiores a 700 metros.

Para este estudo foram analisadas 11 variedades de arroz. Foi através de uma base de dados fornecida pelo Departamento do Arroz da Tailândia que conseguiram juntar os fatores característicos de cada tipo de arroz para serem avaliados. Caso o sistema recomende mais que um tipo de arroz, a recomendação irá ser ordenada começando pelo tipo de arroz que terá um maior retorno monetário por hectare. Após os modelos estarem treinados, foi aplicado um conjunto de regras, com base nas características da variedade do arroz e a sua respetiva zona geográfica. O sistema efetuou as recomendações e para validar as mesmas, os autores efetuaram 50 inquéritos aos produtores de arroz. Este inquérito tinha como objetivo comparar as recomendações dos especialistas com as recomendações do sistema. O resultado da eficácia das recomendações do sistema foi de 82%.

Neste projeto também é mencionado um calendário personalizado para a cultura. O calendário de culturas fornece uma cronologia das atividades agrícolas, a partir do primeiro dia de plantação. A sua estrutura inclui vários parâmetros como fases de crescimento, para esta categoria indicam-se as várias fases de crescimento do arroz, como a germinação, o filhamento, a floração e o enchimento do grão. Para o próximo parâmetro, temos a alimentação das culturas, para cada fase de crescimento, o calendário oferece orientações sobre a alimentação adequada das plantas ou a fertilização necessária. No parâmetro risco de pragas e doenças, o calendário destaca os riscos potenciais de pragas e doenças em cada fase de crescimento. Fornece medidas preventivas e opções de tratamento. O último parâmetro é o período de colheita, o calendário especifica a data ou o período ideal para a colheita. Se houver um risco elevado de perda de rendimento

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

durante o período sugerido, o sistema também recomendará uma data de plantação alternativa com um risco de perda mais reduzido.

Na Dinamarca, num estudo publicado em abril de 2021 [41], os autores realçam a importância da seleção da cultura adequada para um terreno agrícola, com o objetivo de promover a sustentabilidade e as boas práticas agrícolas. Também é revelado que, no país, este tema ainda não tinha sido abordado utilizando ferramentas de aprendizagem automática. Com este estudo os autores pretendiam confirmar se as culturas que os agricultores atualmente cultivam, são os mais adequados para os seus terrenos. Primeiro, foram mapeadas as sustentabilidades de 41 culturas através do modelo ECOCROP (Ecological Crop Requirements). O modelo ECOCROP é uma base de dados usada para verificar as características fisiológicas. Na Tabela 2, são apresentadas duas culturas, com a indicação dos valores mais adequados para praticar o seu cultivo.

Tabela 2 - Dados do EcoCrop referente à batata e cenoura

Crop	Potato	Carrot
Growing season (days)		
Minimum	90	40
Maximum	160	150
Temperature (°C)		
Killing	-1	-1
Minimum, range	7	3
Minimum, optimal	15	15
Maximum, optimal	25	24
Maximum, range	30	30
Precipitation (mm)		
Minimum, range	250	400
Minimum, optimal	500	600
Maximum, optimal	800	1200
Maximum, range	2000	4000
Soil texture		
Light	0.5	0.5
Medium	1	1
Heavy	0.5	0.5
Organic	1	1
Soil drainage		
Insufficient drainage	0	0
Well-drained	1	1
Soil pH		
Minimum, range	4.2	4.2
Minimum, optimal	5.0	5.8
Maximum, optimal	6.2	6.8
Maximum, range	8.5	8.7

O modelo ECOCROP permite estudar as características das culturas com base na temperatura, precipitação, textura do solo, drenagem do solo e pH do solo. Depois, este modelo foi comparado com outro de nome Maxent. O modelo Maxent basearam-se em vários factores, incluindo variáveis climáticas, mas também inclui variáveis socioeconómicas. Por exemplo, para as cenouras, oito das dez variáveis mais importantes eram variáveis climáticas. A distância a cidades com população superior a 10.000 habitantes foi uma variável socioeconómica importante. Os autores também avaliaram o uso histórico do terreno agrícola em causa, por exemplo se um terreno já era usado ou não para plantar a cultura em questão. Esta análise permite perceber se o mapeamento recomendado pelo modelo ECOCROP e o Maxwent estavam a par da realidade.

Os resultados foram os seguintes, nas cenouras, o modelo ECOCROP foram as que mais se aproximaram dos dados históricos de utilização dos solos, enquanto para as batatas, o modelo Maxent foram as que mais se aproximaram dos dados históricos. A utilização histórica das terras para o ano de 1896 estava maioritariamente de acordo com os resultados do modelo ECOCROP. A maioria das freguesias com uma presença significativa de cenouras situava-se em zonas com solos argilosos no leste da Dinamarca. Em contraste, as áreas com elevada aptidão identificadas pelo modelo Maxent tinham uma baixa produção de cenouras no passado.

Um estudo efetuado com o apoio do Universidade Politécnico de Madrid [42], visava de alertar para a falta de conhecimento de como as alterações climáticas podem influenciar a produção de culturas na Europa. Foram selecionadas 13 culturas para este estudo, mas focaram-se mais tarde em apenas 5, o feijão, o tremçoço-azul, tremçoço-branco, quinoa e a ervilha. A análise em questão, utiliza dados climáticos dos últimos 50 anos e dados de cenário futuros até 2050. Novamente é usado o modelo ECOCROP, apresentado anteriormente, para avaliar a sustentabilidade de um meio ambiente para uma cultura específica. Os autores designaram o valor de sustentabilidade (variável de saída), como 0 (insustentável) ou 1 (sustentável). Aplicaram os dados climáticos atuais, que obtidos da WorldClim 2, ao modelo ECOCROP e adicionaram dados de 30 projeções diferentes, de modelos climáticos que eram capazes de prever o clima até 2050. Para medir o impacto na sustentabilidade das culturas, caso o clima atual e futuro sofresse alterações, usaram

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

vários cenários climáticos futuros. Os resultados sugerem que a atual distribuição destas culturas é influenciada pela aptidão climática. As alterações climáticas terão impactos variados na aptidão das culturas. O Norte da Europa poderá tornar-se mais favorável para estas culturas, enquanto o Sul da Europa poderá enfrentar desafios, especialmente para culturas como o grão-de-bico e a lentilha. O estudo sublinha a necessidade de estratégias de investigação e de melhoramento genético para aumentar a tolerância de certas culturas às alterações climáticas previstas. Sugere também que o planeamento da produção e as políticas agrícolas devem ter em consideração estes impactos para apoiar a resiliência da produção de culturas ricas em proteínas.

É possível avaliar que este tema é relevante para um futuro autossustentável, mas ainda está pouco aprofundado. Técnicas de aprendizagem automática em *Smart Farming*, são conceitos relativamente novos [13]. Sendo assim, é importante fazer mais investigações acerca deste tema, especialmente no nosso país onde existe muita necessidade de apoiar os agricultores a lidarem com os impactos das alterações climáticas. Devo referir que após esta pesquisa ainda não encontrei casos de uso de sistemas de recomendação inteligentes na agricultura em Portugal. Todos os artigos realçam a importância no impacto que as alterações climáticas irão ter na sustentabilidade das culturas, a uma escala global. A maioria dos casos de estudo são originais da Ásia, pois é um continente em que os países dependem mais da agricultura e dos pequenos agricultores. Os dados mais usados são as temperaturas do ar e do solo, a humidade do ar e do solo, o pH do solo e as precipitações. O estudo efetuado na Europa prevê que o sul da Europa irá ser a zona mais afetada para a sustentabilidade das culturas onde Portugal está incluído [42].

3 METODOLOGIA

3.1 Tecnologias

O avanço tecnológico tem proporcionado uma variedade de ferramentas e plataformas que facilitam a análise de dados e o desenvolvimento de aplicações web. No contexto deste trabalho, diversas tecnologias foram utilizadas para garantir a eficiência e precisão dos resultados pretendidos. Este capítulo vai apresentar e descrever as tecnologias selecionadas ao longo da pesquisa. Além das ferramentas tecnológicas, é essencial destacar as técnicas específicas de análise de dados e do desenvolvimento do modelo que foram adotadas.

3.1.1 Django

Django é um *framework web* de alto nível escrito em Python que permite desenvolver aplicações web de forma rápida e eficiente (Figura 3). Ele fornece muitos componentes reutilizáveis e uma arquitetura robusta para facilitar o desenvolvimento. Como por exemplo o *django admin*, que é um painel de administrador onde se pode gerir algumas funcionalidades importantes do sistema através de uma interface. Outra funcionalidade já implementada no sistema é um método de autenticação, para efetuar *logins*.

O Django adota o padrão de projeto Model-View-Controller (MVC), mas é frequentemente referido como Model-View-Template (MVT) devido à sua abordagem distinta. Esta *framework* é amplamente reconhecida pela rapidez no desenvolvimento, a segurança e a capacidade de escalabilidade. A arquitetura do Django (Figura 4) é composto por três componentes principais.

Model - Representa a estrutura de dados da aplicação, definindo a organização das bases de dados, incluindo tabelas, relações e índices. Uma característica notável é o seu ORM integrado, que permite interagir com a base de dados usando apenas Python, eliminando a necessidade de escrever código SQL.

View - Responsável por controlar as interações do utilizador com a aplicação. Processa as solicitações, aplica a lógica de negócios e retorna as respostas adequadas.

Template - Determina a apresentação dos dados ao utilizador. É a camada que define a estrutura e o design das páginas web, permitindo a inserção dinâmica de conteúdo em páginas HTML.

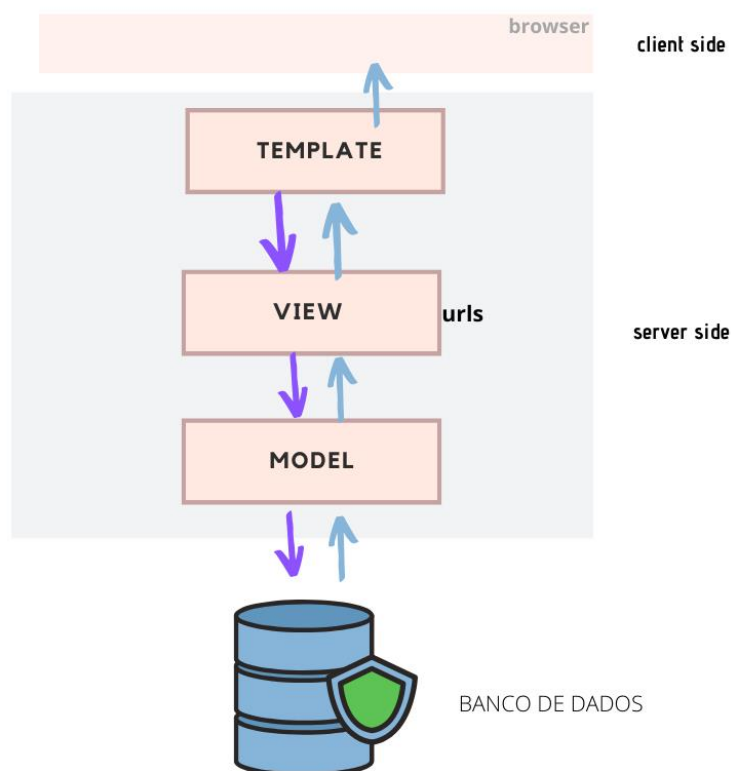


Figura 3 - Arquitetura do Django

3.1.2 Plotly

Plotly é uma biblioteca gráfica interativa para Python que permite criar visualizações de dados interativas e atraentes. É amplamente utilizado para criar gráficos interativos e *dashboards*. O Plotly já vem com algumas funcionalidades práticas implementadas, como dar *zoom* nos gráficos ou fazer o download em formato *png*. Esta biblioteca permitiu-me construir os gráficos que podem ser visualizados no website.

3.1.3 Python para Análise de Dados

Python é a linguagem de programação mais usada para efetuar a análise de dados e também na ciência dos dados. As bibliotecas que se utilizou foram o Pandas, que é usado para manipulação e análise de dados, o NumPy para computação numérica, o Matplotlib para visualização de dados e o scikit-learn para preparar os dados para a sua modelação após o seu processamento



Figura 4 - Bibliotecas de Python para Análise de Dados

3.1.4 Tensorflow

TensorFlow é uma biblioteca de código aberto para *machine learning* e *deep learning* desenvolvida pelo Google. É amplamente utilizado para a criação e treinamento de modelos de *deep learning*. Os meus modelos de previsão foram construídos com esta ferramenta.

3.1.5 PostgreSQL

PostgreSQL é um sistema de gestão de base de dados relacional de código aberto. É conhecido pela sua confiabilidade, escalabilidade e recursos avançados de gestão de dados. Esta base de dados é onde eu guardo todos os dados da minha aplicação *web*.

3.1.6 LoRaWAN

LoRaWAN é uma tecnologia de comunicação sem fios de baixa potência projetada para redes de IoT de longo alcance. É usada para ligar dispositivos IoT à Internet. Este tipo de comunicação permite enviar os dados do sensor para uma *gateway* até uma distância de 10km. Uma das suas vantagens é operar em bandas de comunicação que não requerem licenciamento. O que significa que qualquer indivíduo pode montar a sua infraestrutura de rede sem ter de pagar a um operador de telecomunicações.

3.1.7 TTN (The Things Network)

TTN é uma plataforma de rede de IoT que fornece conectividade para dispositivos IoT usando a tecnologia LoRaWAN. É usado para comunicação de dispositivos IoT em redes de longo alcance. É através do TTN que é possível visualizar a chegada dos dados enviados pelos sensores. Além disso, o TTN permite-nos optar por uma rede pública, ampliando as possibilidades de conectividade e tornando o custo gratuito. A arquitetura do TTN (Figura 5) divide-se em 4 componentes.

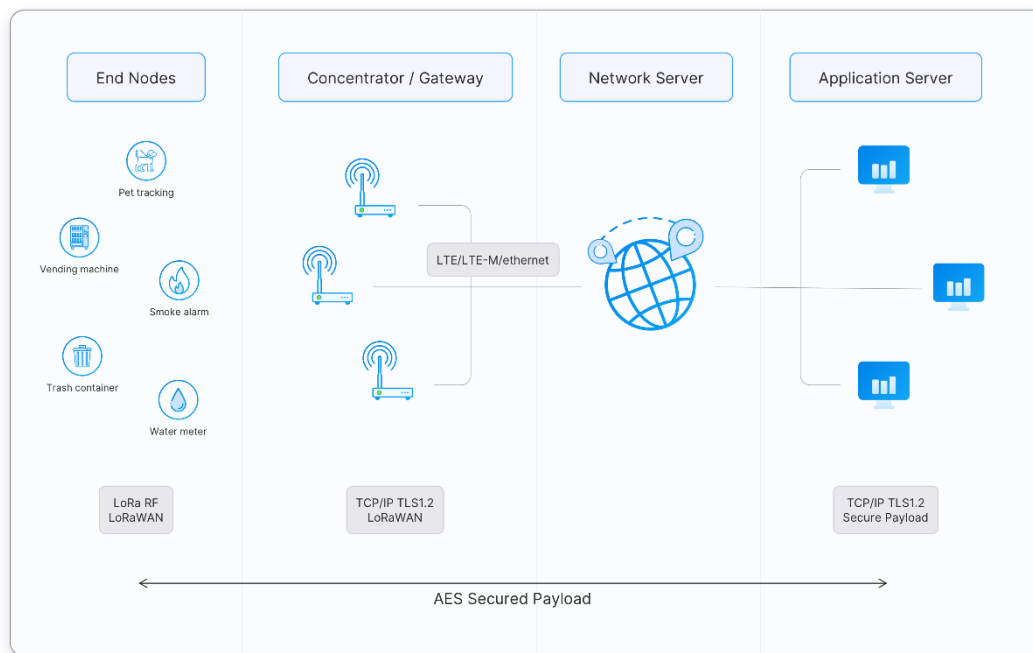


Figura 5 - Arquitetura do TTN. Fonte: thethingsnetwork.org

End Nodes - Estes são os sensores que recolhem e enviam dados. No diagrama, temos exemplos como rastreadores de animais de estimação, máquinas de venda automática, alarmes de incêndio e medidores de água. Eles utilizam a tecnologia LoRaWAN para transmitir dados devido ao seu baixo consumo de energia e longo alcance.

Gateway - Funciona como uma ponte entre os sensores e a Internet. Ele capta os sinais LoRaWAN dos sensores e converte em pacotes de dados que são então enviados via Internet. O diagrama mostra que a comunicação entre os *gateways* e o servidor de rede pode ser realizada através de várias tecnologias, como LTE, LTE-M e Ethernet.

Network Server – É responsável por receber os pacotes de dados dos *gateways*, processa-os e encaminha-os para o servidor de aplicação correspondente. Além disso, ele gere a alocação de canais e taxas de dados para otimizar a capacidade da rede e a vida útil da bateria dos sensores. A comunicação entre o servidor de rede e os servidores de aplicação é segura, geralmente usando TCP/IP com TLS1.2.

Application Server - É aqui que os dados são transformados em informações úteis para o utilizador ou sistema final.

3.1.8 ChatGPT

ChatGPT é uma ferramenta de IA dedicada à linguagem natural que pode ser usada para responder a perguntas e fornecer assistência no desenvolvimento de software. Para este projeto, foi usado para *debugging* de código e resolução de problemas relacionados com a programação.

3.1.9 GitHub

GitHub é uma plataforma de desenvolvimento colaborativo que oferece controlo de versão de código-fonte, gestão de projetos, colaboração e hospedagem de código. É amplamente utilizada para colaboração no desenvolvimento de software.

3.1.10 Webhooks

Webhooks são *callbacks* HTTP que permitem que uma aplicação notifique outra sobre eventos, operando em tempo real. São usados para automatizar a comunicação entre sistemas. É através de webhooks que os dados da TTN chegam à plataforma Django.

3.1.11 Visual Paradigm

O Visual Paradigm é uma ferramenta de modelação de software que suporta várias notações de modelação, incluindo *Unified Modeling Language* (UML). Foi usado para criar diagramas de classe, diagramas casos de uso, diagramas de atividade e outros tipos de modelos que auxiliam no planeamento do projeto e desenvolvimento do software.

3.1.12 Google Colab

Google Colab é uma plataforma em *cloud* que permite a execução de *notebooks* na linguagem Python diretamente no navegador. É amplamente utilizado para treinar modelos quer ML quer DL e também muito utilizado na área análise de dados. A análise de dados e desenvolvimento dos modelos foram feitos nesta plataforma.

3.1.13 Postman

O Postman é uma plataforma de colaboração para desenvolvimento de APIs que simplifica o seu processo de teste e documentação. Permite que os programadores criem, compartilhem, testem e documentem APIs por meio de uma interface amigável. O Postman foi usado para automatizar testes de API e enviar solicitações HTTP personalizadas.

3.1.14 Ngrok

Ngrok é uma ferramenta que permite criar túneis seguros para exposição de servidores locais à internet. O Ngrok gera URLs públicas temporárias para seus servidores locais, facilitando o compartilhamento e o teste de aplicações em diferentes dispositivos e ambientes de rede. É utilizado para testar aplicações web e APIs em desenvolvimento que estão em execução em ambientes locais, como é o caso de *webhooks*.

3.2 Hardware

3.2.1 Sensor

O sensor Dragino S31, utilizado neste projeto, é um dispositivo IoT desenvolvido pela empresa chinesa Dragino, projetado para recolher variáveis ambientais como temperatura e a humidade do ar (Figura 6). Equipado com comunicação via LoRaWAN, permite transmissões de longo alcance e baixo consumo de energia, tornando-o ideal para aplicações agrícolas em áreas rurais ou remotas. Sua durabilidade e design compacto garantem resistência a condições adversas, enquanto sua precisão assegura dados confiáveis para alimentar a base de dados.



Figura 6- Sensor Dragino S-31

Em relação à *gateway*, não é possível identificar o modelo, pois o acesso a esta informação não é público, pelo proprietário do dispositivo.

3.2.2 Placa Gráfica para Treino dos Modelos

A A100 GPU é uma placa gráfica desenvolvida pela NVIDIA, especificamente projetada para aplicações de alta performance em computação, incluindo DL. A utilidade deste dispositivo foi para acelerar a duração do treino do modelo usadas. Para DL eles permitem uma multiplicação de matrizes



Figura 7 - Placa Gráfica NVIDIA A100

extremamente rápida, o que é fundamental para muitas operações em redes neurais. A A100 é facilmente integrada com *frameworks* populares de DL, como TensorFlow e PyTorch [43]. À medida que avançamos no desenvolvimento do modelo de previsão, a sua complexidade aumentava, tornando a duração do treino cada vez mais lenta. Através do Google Colab foi possível alugar esta placa gráfica.

3.2.3 Servidor Local e Computador Pessoal

O servidor local utilizado é um computador pessoal, que possui as seguintes especificações:

Sistema Operacional - Microsoft Windows 10 Home

Processador - AMD Ryzen 5 2600 Six-Core Processor, com uma velocidade de clock de 3400 MHz, 6 núcleos físicos e 12 processadores lógicos.

Memória RAM instalada - 16.0 GB

Placa de vídeo - AMD Radeon 6600 XT

Essas especificações são importantes para o funcionamento do projeto, uma vez que o hardware do servidor local tem de ter um bom desempenho, pois para fazer previsões com modelos de DL, é necessário um bom componente gráfico.

3.3 Métrica de avaliação para a Recomendação da Data de Plantação e Avaliação dos Modelos

No domínio da machine learning e da ciência dos dados, as métricas de avaliação desempenham um papel fundamental na avaliação do desempenho dos modelos. Estas métricas fornecem medidas quantitativas que refletem o grau de precisão das previsões de um modelo com os resultados reais. Ao compreender estas métricas, os profissionais podem tomar decisões informadas sobre a seleção, afinação e implementação de modelos.

No contexto da recomendação da data de plantação irei explicar com exemplos visuais no capítulo 4.3.3. usando a batata como exemplo.

A importância das métricas de avaliação são, a comparação de modelos. As métricas de avaliação permitem a comparação de diferentes modelos na mesma tarefa, ajudando na seleção do modelo com melhor desempenho. A melhoria do modelo também ajuda a compreender onde um modelo está a ter um desempenho inferior. Com a métrica de avaliação, consigo perceber se uma nova variável de entrada teve um impacto positivo nas previsões ou negativo.

Uma das métricas fundamentais utilizadas, especialmente em tarefas de regressão, é o erro médio absoluto (MAE). O MAE mede o valor médio dos erros entre os valores previstos e reais. É calculado como a média das diferenças absolutas entre os valores previstos e reais.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

- n é o número total de observações.
- y_i é o valor real.
- \hat{y}_i é o valor previsto pelo modelo.

A sua interpretação é muito simples, quanto menor é o valor de MAE, melhor é o resultado do modelo, enquanto um MAE maior indica uma performance mais fraca. Um MAE de 0 indica previsões perfeitas.

3.4 Eliminação de *Outliers*

Dentro do processo de limpeza e preparação dos dados, a identificação e eliminação de *outliers* é fundamental para assegurar a qualidade dos dados alimentando os modelos de ML. Os *outliers*, que são valores atípicos em relação à distribuição normal dos dados, podem interferir e levar a modelos menos precisos.

Neste projeto, a amplitude interquartil (IQR) foi adotada como a principal metodologia para a deteção e eliminação de *outliers*. O IQR é uma técnica de estatística que representa a diferença entre o terceiro quartil (Q3) e o primeiro quartil (Q1) dos dados. Os *outliers* são frequentemente definidos como observações que caem abaixo de $Q1 - 1.5 \times IQR$ ou acima de $Q3 + 1.5 \times IQR$.

3.5 Metodologia para a Escolha do Melhor Modelo Baseados em Deep Learning

No domínio da agricultura, a previsão precisa da temperatura média diária é necessário para determinar as datas ideais de plantação face às suas características fisiológicas. Através de técnicas de *DL*, é possível desenvolver modelos que podem prever com precisão as temperaturas futuras, auxiliando assim os agricultores nas suas decisões. Para implementar o modelo foram seguidas 4 fases:

Escolha da arquitetura do modelo – Começa-se por escolher um modelo base (pior performance) e um modelo de referência, reconhecido por ter um bom desempenho (melhor performance). De seguida escolhe-se uma lista das arquiteturas simples do Tensorflow e avalia-se o modelo. A arquitetura que têm o melhor resultado passa para a fase seguinte.

Determinação da janela temporal - Após a arquitetura estiver definida, determina-se a janela de tempo adequada. A janela de tempo influencia diretamente a forma como os dados são fornecidos ao modelo, e uma escolha adequada pode melhorar significativamente a performance do modelo.

Seleção das variáveis de entrada - Com a arquitetura e a janela de tempo definidas, podem-se começar a selecionar as variáveis mais relevantes, que acrescentam valor à previsão.

Escolha dos hiperparâmetros e configuração das *Layers* do Modelo - Com uma base sólida estabelecida, podem-se ajustar e otimizar as *layers* do modelo e os seus hiperparâmetros. Isto inclui decidir sobre o número de *layers*, o número de neurónios por cada *layer*, as funções de ativação, o *learning rate* entre outros parâmetros.

3.5.1 Escolher um modelo base

Em ML, o modelo base é um modelo simples utilizado como ponto de referência para comparar o desempenho de modelos mais complexos. A ideia é definir um padrão mínimo ou uma referência em relação à qual outros modelos são avaliados. Se um modelo sofisticado não conseguir superar o desempenho do modelo base, isso sugere que o modelo pode não ser eficaz ou que pode necessitar de uma afinação adicional.

Um modelo de base pode ser tão simples como:

- Prever a classe mais frequente em problemas de classificação.
- Utilizar a média ou a mediana como previsões em problemas de regressão.
- Utilizar métodos estatísticos simples ou regras heurísticas.

Para este projeto irá ser utilizado o *persistence model* como modelo base. Na previsão de séries temporais, o *persistence model* repete o último resultado para todas as previsões. Para se perceber melhor o modelo na Figura 8 está a ilustração dos resultados do modelo.

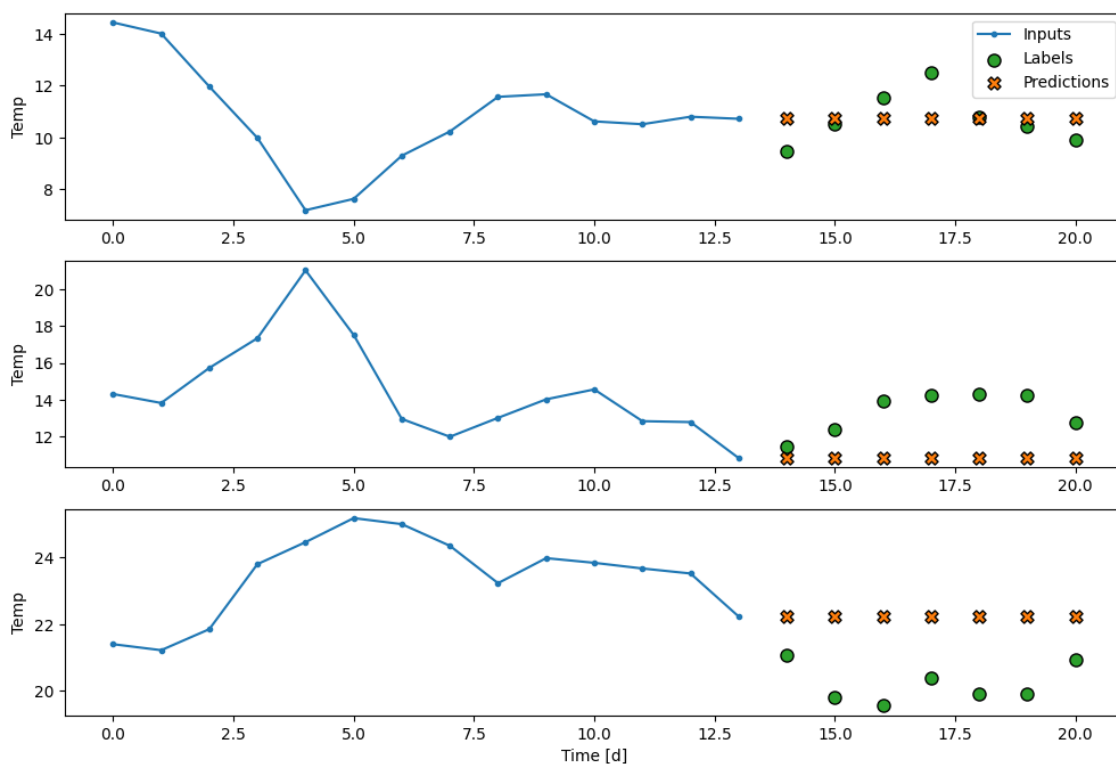


Figura 8 - Previsão do Persistence Model aplicada aos dados da Temperatura

3.5.2 Escolher um modelo de referência.

O modelo N-BEATS, tal como apresentado no documento de investigação intitulado "N-BEATS: Neural Basis Expansion Analysis for Interpretable Time Series Forecasting" de Dmitri Carpov, Boris N. Oreshkin, Nicolas Chapados e Yoshua Bengio, oferece uma arquitetura neural de DL adaptada à previsão pontual de séries temporais univariadas. Este modelo foi concebido para ser interpretável e aplicável numa vasta gama de domínios, tornando-o uma escolha ideal para validar outros modelos de previsão. O foco principal do N-BEATS é abordar o problema da previsão pontual de séries temporais univariadas usando DL [44].

O N-BEATS demonstrou um desempenho de topo em vários conjuntos de dados bem conhecidos, incluindo o M3, o M4 e o TOURISM. Melhorou a precisão das previsões em 11% em relação a outro modelo de referência e em 3% em relação ao vencedor do concurso M4 do ano anterior [44].

O modelo mostra que o DL puro, sem componentes específicos de séries temporais, pode superar as abordagens estatísticas estabelecidas na área da previsão.

O modelo N-BEATS aplica-se a séries univariadas, ou seja, apenas pode utilizar a temperatura como a única variável para prever a temperatura.

O modelo N-BEATS, com a sua arquitetura profunda, serve como uma excelente ferramenta para validar outros modelos de previsão. Ao comparar as métricas de desempenho e compreender os padrões subjacentes, podemos assegurar-nos que os modelos a serem desenvolvidos estão no caminho certo, e possibilitar fazer os ajustes necessários para obter os melhores resultados.

Variáveis de entrada - Temperatura média diária

Variáveis de saída - Temperatura média diária

3.5.3 Arquiteturas Usadas

As redes neurais revolucionaram o campo de IA e do ML, oferecendo uma vasta gama de arquiteturas adaptadas a diferentes tarefas. Este capítulo aborda três tipos importantes de redes neurais, modelo linear, modelo *dense* e rede de memória de curto prazo longa (LSTM).

3.5.3.1 Modelo Linear

Um modelo linear, muitas vezes referido como modelo de regressão linear, é uma das formas mais simples de redes neuronais. Pressupõe uma relação linear entre a entrada e a saída. A sua estrutura consiste numa única *layer* e num único neurónio. O neurónio multiplica a variável de entrada com um peso, adiciona uma polarização e produz uma saída. A sua aplicação é adequada para tarefas em que a relação entre a entrada e a saída é linear, tais como problemas de regressão simples [26].

3.5.3.2 *Modelo Dense*

As redes densas, também conhecidas como redes totalmente ligadas, são o tipo mais comum de arquitetura de redes neuronais. Cada neurónio de uma *layer* está ligado a todos os neurónios da camada anterior e das camadas seguintes. Amplamente utilizados em tarefas que vão desde a classificação de imagens até ao processamento de linguagem natural [28].

3.5.3.3 *Modelo LSTM*

Os LSTMs são um tipo de RNN (Rede Neuronal Recorrente) concebida para reconhecer padrões ao longo de intervalos de tempo. Contém células de memória que armazenam informações por longos períodos. Cada célula tem três portas: porta de entrada, porta de esquecimento e porta de saída. Capaz de aprender dependências de longo prazo, o que o torna adequado para sequências e séries temporais de dados [45]. Amplamente utilizado na previsão de séries temporais.

3.5.4 **Janelas Temporais, Input e Output, Windows e Horizons**

No contexto de séries temporais do TensorFlow, uma *window* refere-se a um conjunto de dados usados para treinar um modelo. Por exemplo, se queremos prever a temperatura para o dia da amanhã usando os dados dos últimos 7 dias, a *window* será igual a 7. O *horizon* refere-se ao número de etapas no futuro que se quer prever. Usando o exemplo anterior, se queremos prever apenas a temperatura do próximo dia, então o *horizon* será igual a 1. Se quiséssemos prever a temperatura dos próximos 30 dias, o *horizon* seria igual a 30.

3.5.4.1 Sliding Windows

Este é um método para criar conjuntos de treino e teste em séries temporais. *Sliding Windows*, traduz-se por janela deslizante, a janela de dados move-se uma etapa de cada vez. Por exemplo, se tivermos dados de 10 dias e estivermos a usar uma *window* de 7 dias com um *horizon* de 1 dia, primeiro usaria-se os dias 1-7 para prever o dia 8, depois os dias 2-8 para prever o dia 9, e assim por diante.

3.5.4.2 Expanding Window (Janela Expansiva)

Neste método, o tamanho da *window* aumenta gradualmente, enquanto o *horizon* permanece constante ou também pode crescer. Por exemplo, podemos começar a treinar com dados dos dias 1-7 para prever o dia 8, depois utilizaríamos os dias 1-8 para prever o dia 9, depois para prever o dia 10 utilizaríamos os dias 1-9 e assim por diante. Isto é útil quando os dados mais antigos ainda são relevantes para previsões futuras.

3.5.5 Seleção das Variáveis de Entrada

Refere-se ao processo de selecionar o conjunto de variáveis que irá construir o modelo. A ideia base é escolher as características que contribuem mais significativamente para a previsão da temperatura, eliminando as variáveis que podem ser redundantes ou irrelevantes. Isto não só pode melhorar a performance do modelo, como também pode reduzir o tempo de treino e tornar o modelo mais interpretável. O critério de seleção de variáveis é o seguinte, após o desenvolvimento dos modelos univariáveis, acrescenta-se uma nova variável antes de treinar o modelo, mantendo a arquitetura e os hiperparâmetros iguais. Se a performance do modelo melhorar, então a nova variável acrescenta valor ao modelo. Caso contrário não irá ser considerada. No capítulo 4.2.3. nas adições de novas variáveis, está divulgado quais as variáveis de entrada que serão utilizadas.

3.5.6 Escolha dos Hiperparâmetros

Estes são os parâmetros que não são aprendidos durante o treino do modelo, mas são definidos antes do início do treino. A sua otimização é fundamental, pois diferentes valores podem resultar em diferentes performances do modelo. Exemplos comuns de hiperparâmetros incluem a taxa de aprendizagem (*learning rate*), o número de épocas (*number of epochs*), o tamanho do lote (*batch size*) e a taxa de abandono (*dropout rate*), entre outros. A escolha adequada dos hiperparâmetros é muitas vezes realizada através de experimentações aleatórias, onde diferentes combinações de hiperparâmetros são testadas para encontrar a que oferece a melhor performance.

3.5.7 Layers nos Modelos de Deep Learning

Um passo crucial é configurar *layers* do modelo. Estas *layers* determinam a profundidade e a complexidade do modelo e têm um impacto direto na sua capacidade de capturar padrões nos dados.

As *layers* em modelos de aprendizagem profunda permitem que o modelo aprenda representações hierárquicas dos dados. Os *layers* iniciais geralmente capturam características de baixo nível, enquanto *layers* mais profundas capturam características de alto nível e abstrações mais complexas. A adição de mais *layers* pode aumentar a capacidade do modelo, permitindo-lhe aprender padrões mais complexos. A configuração das *layers* envolve vários parâmetros.

Número de Layers - Determina a profundidade do modelo. Modelos mais profundos podem capturar padrões mais complexos, mas também são mais propensos a *overfitting* e requerem mais dados para treino. Enquanto os modelos univariáveis que eu implementei no início deste capítulo são compostos por 2 *layers*, um modelo já bem aprofundado pode facilmente ultrapassar as 100 *layers*.

Número de Neurónios por Layer - Define a largura de cada *layer*. *Layers* maiores podem aprender mais características, mas também aumentam a complexidade computacional.

Tipo de Layer - Dependendo da arquitetura escolhida, pode-se optar por diferentes tipos de *layers*, como *layers* densas (*fully connected*), *layers* convolucionais, *layers* LSTM, entre outras. É esta configuração que altera a arquitetura de um modelo. Se estamos perante *layers* LSTM então o modelo é um LSTM.

Função de Ativação: Define a transformação aplicada às saídas de cada neurónio. Funções comuns incluem *ReLU*, *sigmoid* e *tanh*.

3.6 Implementação de um Modelo

O propósito desta implementação serve para responder à primeira questão, que é achar qual a melhor arquitetura para detalhar o meu modelo.

A implementação de um modelo no Tensorflow passa sempre por estes 4 procedimentos.

3.6.1 Construir o Modelo

A primeira etapa envolve a especificação da arquitetura ou estrutura do modelo. Isto inclui a determinação do tipo de modelo, por exemplo, regressão linear ou rede neural. A arquitetura é definida na *layer* de *input*.

3.6.2 Compilar o Modelo

Uma vez definida a arquitetura do modelo, este tem de ser compilado. Esta etapa envolve a especificação do otimizador que ajusta os pesos do modelo, a função de perda que mede os erros do modelo e as métricas a monitorizar durante o treino, que neste caso é o MAE.

3.6.3 Treinar o Modelo

Com o modelo compilado, a próxima etapa é treiná-lo usando dados. Isto envolve alimentar o modelo com dados de entrada e os valores-alvo correspondentes, especificando o número de *epochs* (iterações sobre todo o conjunto de dados) e definindo o tamanho do *batch*.

3.6.4 Avaliar o Modelo

Após a formação, é essencial avaliar o desempenho do modelo em dados não vistos ou de teste. Isto fornece uma indicação do desempenho do modelo em dados do mundo real.

4 Conjunto de Dados

4.1 Recolha dos Dados

4.1.1 Dados Climáticos

Os dados diários climáticos utilizados neste trabalho foram colhidos da estação meteorológica de Santa Clara, localizada no Observatório Geofísico e Astronómico da Universidade de Coimbra. Os dados vão desde 1866 até 2019. A série disponível é bastante heterogénea e encontra-se dispersa em vários formatos. Para os períodos entre 1926 e 1941 e de 1996 a 2004, não existem dados diários disponíveis. O conjunto de dados contém muitas variáveis climáticas, como por exemplo temperatura do solo, velocidade do vento, humidade do solo, entre outros. Os períodos temporais em que existem dados são os seguintes:

1. 1866 a 1925.
2. 1941 a 1996.
3. 2005 a 2017.
4. 2018 a 2019 (dados horários)

Por outro lado, para possibilitar a visualização de dados em tempo real e efetuar futuras previsões, eram necessários dados mais recentes. Para tal, foi instalado um sensor no exterior de um prédio situado na Quinta da Maia em Coimbra a cerca de 80 metros em cima do nível do mar. O sensor é o já referido no capítulo 3.2. Com este equipamento foi possível obter dados recentes com uma periodicidade de 15 minutos. Na Figura 9 apresenta-se a localização do equipamento.

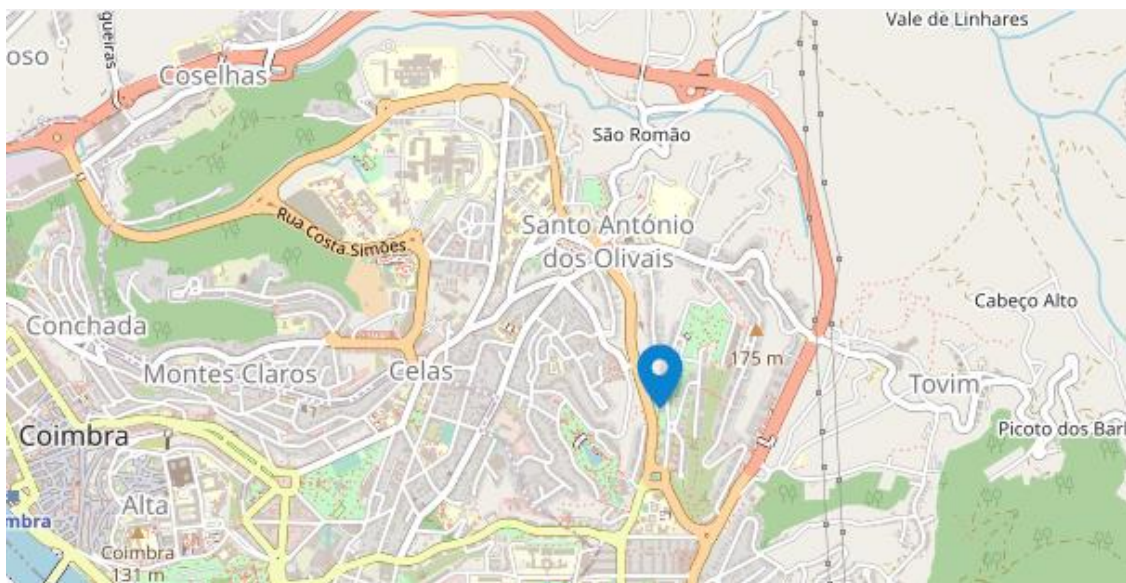


Figura 9 – Localização do Sensor

Os dados climáticos mensais, de 1866 a 1955, foram obtidos a partir do IPMA e são compostos por duas variáveis a data e a temperatura.

4.1.2 Dados das colheitas das Culturas na Região de Coimbra

O INE oferece uma vasta gama de informações estatísticas desde o ano de 1944. No entanto, nos primeiros anos de recolha dos dados relacionados com a produção das culturas na região de Coimbra, a superfície dos terrenos cultivados não foi medida em hectares, mas sim em termos de produção por quinta. Sem o conhecimento da dimensão das quintas, esses dados iniciais não fornecem uma medida precisa da produtividade, portanto, para o propósito desta análise, foram excluídos. No ano de 1969, a superfície dos terrenos cultivados começou a ser medida em hectares, permitindo assim uma análise mais precisa da produtividade das colheitas, expressa em quilogramas por hectare. As culturas temporárias incluídas neste estudo são: trigo, milho, centeio, arroz, aveia, cevada, fava, feijão, grão e batata. Para cada cultura, a produtividade por hectare foi calculada para a região de Coimbra.

Vale a pena referir que a organização dos dados por região foi interrompida em 1989. A partir desse ano, a informação passou a ser agregada a nível nacional, o que implica que os dados após esta data não podem ser comparados diretamente com os dados anteriores a 1989. Recomendamos, portanto, cautela ao interpretar os resultados que incluem dados após este ano [46].

4.1.3 Dados das culturas hortícolas

Para mapear o comportamento das culturas hortícolas foi utilizado o livro chamado Manual de Culturas Hortícolas do Professor e Engenheiro Agrónomico Domingos Almeida [47]. Neste livro cada cultura é descrita detalhadamente, para cada cultura existe o capítulo “Exigências Climáticas”, no qual o autor aborda as condições climáticas mais favoráveis para essa cultura. Outros dados necessários e que não estavam disponíveis no livro, foram obtidos através do ECOCROP.

4.2 Processamento dos dados

4.2.1 Junção dos Dados Diários Climáticos

Os dados da estação meteorológica de Santa Clara, estavam publicados de forma desorganizada e em formato Excel. Cada ficheiro representava um ano e cada folha de Excel representava um mês. Por isso foi necessário automatizar o processo da junção dos dados.

Inicialmente juntaram-se os ficheiros que tinham uma estrutura idêntica através da biblioteca Pandas. Obteve-se uma *dataframe* composta por todos os anos contidos dentro do ficheiro zip (*zip file*). Após ter todos os ficheiros dentro de um zip file, foi desenvolvido uma função para extrair os dados da temperatura para todos os ficheiros e para todas as folhas de cálculo dentro de cada ficheiro.

4.2.2 Limpeza dos Dados

O processo de limpeza de dados é crucial para o projeto, pois dados imprecisos ou inconsistentes, como leituras de temperatura incorretas, podem levar a previsões menos precisas, afetando a precisão das recomendações de datas para cultivo. A limpeza dos dados garante a confiabilidade dos dados usadas para orientar os agricultores na tomada de decisões.

Neste capítulo começou-se por identificar valores em falta. Todos os anos que continham mais de 20% das datas ou dos valores em falta foram eliminados do *dataset*. Em ML, para treinar os modelos, valores em falta irão originar erros neste processo. Por isso este passo é sempre obrigatório.

Depois foram identificados os valores em falta na coluna da temperatura. Existem dois tipos de valores em falta, os valores *null*, ou seja, valores em que as células estão em branco, ou então se o valor for igual a 99.9. Este tipo de valor foi adicionado pela própria estação meteorológica, caso não houvesse registos. Ao substituir os valores 99.9 por *null*, aplicou-se o método de interpolação linear para preencher os valores em falta, visto que é um método muito utilizado para o preenchimento de temperaturas em falta.

Por último, só faltava remover os *outliers*, um dos métodos mais eficazes para a sua identificação foi visualizar o gráfico de linhas das temperaturas. Para remover os *outliers* considerou-se dois métodos, a amplitude interquartil e o z-score. Como o z-score identifica *outliers* considerados valores plausíveis decidiu-se utilizar a amplitude interquartil como método de remover os *outliers*, para todos os conjuntos de dados.

4.2.3 Adicionar Dados

Para otimizar o desempenho do modelo de previsão, foram introduzidas novas variáveis que capturam padrões sazonais e fenômenos climáticos específicos que serão abordados com mais detalhe no capítulo 6.4.2.

Variável do Cosseno - A primeira adição foi a introdução de uma variável cíclica para capturar a sazonalidade anual. Esta variável foi derivada da data, convertendo-a para o dia do ano e , em seguida, aplicando funções trigonométricas para criar uma representação cíclica. O uso do cosseno permite que o modelo reconheça padrões sazonais, como a variação de temperatura ao longo do ano. A equação utilizada foi a seguinte:

$$\text{Variável do Cosseno} = \cos\left(\frac{d \times 2\pi}{365.2425}\right)$$

- d representa o dia do ano e tem de ser maior ou igual a 1 ou menor ou igual a 365.

Variável do El Niño - O fenômeno El Niño tem um impacto significativo nas condições climáticas globais, influenciando padrões de temperatura e precipitação. Para incorporar esta informação, foi criada uma variável que mapeia anos específicos para a intensidade do El Niño observada nesses anos. A intensidade foi categorizada em três níveis: "Fracamente intensidade", "Intensidade moderada" e "Forte intensidade", que foram posteriormente convertidos em valores ordinais. Mas após a análise de resultados no capítulo 6.4.2., decidiu-se utilizar apenas a intensidade forte, pois era a variável que melhorava a performance do modelo. Com isto a variável passou a ser binária, ou ocorreu o fenômeno El Niño ou não. Os valores com intensidade moderada e de fraca intensidade foram substituídos por 0. Esta variável permite que o modelo considere o impacto potencial do El Niño nas condições climáticas.

A introdução destas variáveis visa melhorar a capacidade do modelo de capturar padrões nos dados que não seriam evidentes apenas com a variável original. A combinação de informações sazonais e fenômenos climáticos específicos pode proporcionar previsões mais precisas e robustas.

4.2.4 Divisão dos Dados

Após um conjunto de dados passar pela limpeza, o primeiro passo a fazer antes de passar-se para a modelação dos dados é a separação em conjuntos para treino e para testes (Figura 10). Foi utilizado usando o método mais comum, 80% para treino e 20% para testes, usando a função *train_test_split* da biblioteca *sklearn*. Como estamos perante um conjunto de dados de uma série temporal, a separação dos dados não pode ser feita de forma aleatória, mas sim de forma contínua, pois devido à natureza dos dados, uma série temporal tem de ser treinado a ter em consideração a ordenação temporal.

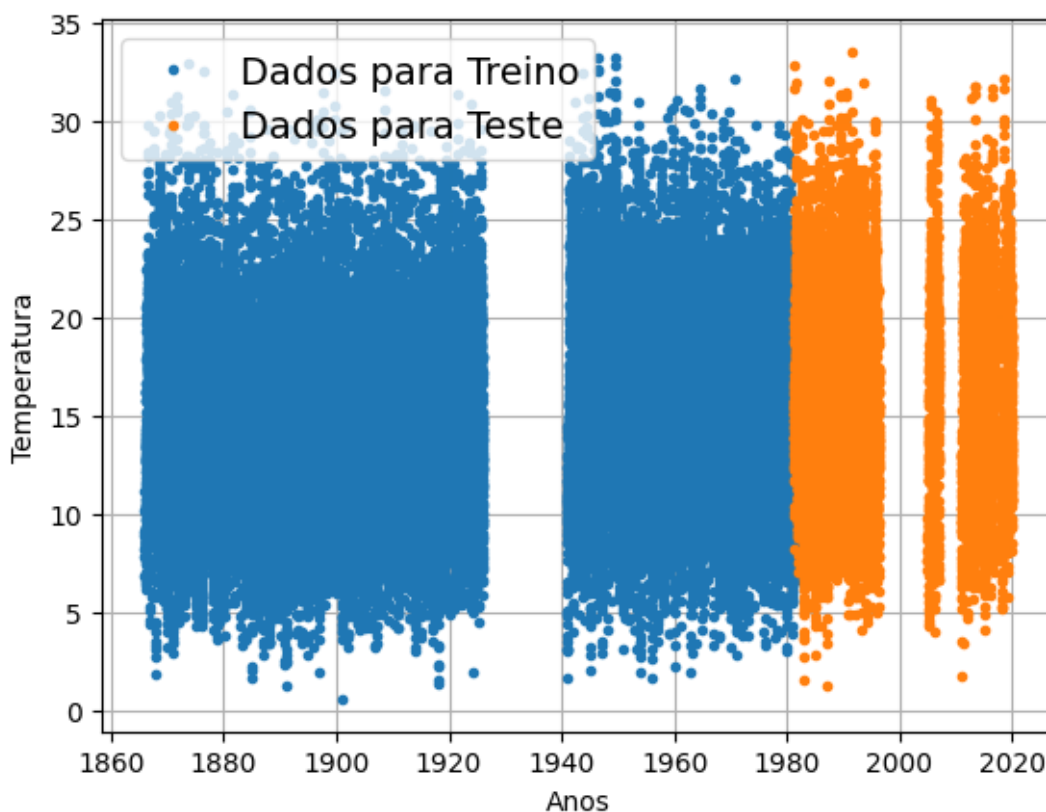


Figura 10 - Conjunto de Dados para Treinar e Testar Modelos

O conjunto de dados para treino é composto por 36620 dias e o conjunto de dados para teste é composto por 9155 dias. É de notar que os anos em falta não causam problemas para a modelação dos dados porque no conjunto de dados não existem valores em branco.

4.3 Análise Exploratória dos Dados

4.3.1 Temperatura e Precipitação.

Coimbra está localizada na Região Centro de Portugal, na sub-região da Beira Litoral. Esta região é caracterizada como um clima mediterrâneo. Os verões são quentes e secos enquanto os invernos são mais chuvosos, amenos e húmidos. Normalmente o clima deste tipo está localizado 30° a 50° de latitude norte e Coimbra localiza-se a aproximadamente nos 40°N.

O clima de Coimbra também é influenciado pela sua proximidade ao oceano. A cidade situa-se a menos de 40km do oceano atlântico. Esta influência marítima pode trazer alguma estabilidade térmica, atenuando as amplitudes térmicas entre o dia e a noite. Além disso, contribui para uma maior humidade relativa do ar e para ventos mais moderados.

Os Verões em Coimbra são quentes, com temperaturas máximas médias que variam entre 25 e 29°C (Figura 11 e 12). Durante esse período, é comum haver dias soalheiros e céu claro. Em alguns casos, pode ocorrer ondas de calor, elevando as temperaturas acima dos 40°C.

Os Invernos em Coimbra são amenos, com temperaturas mínimas médias entre 5° e 8°C. No entanto, as temperaturas não costumam diminuir muito abaixo dos valores mencionados, proporcionando um inverno relativamente suave.

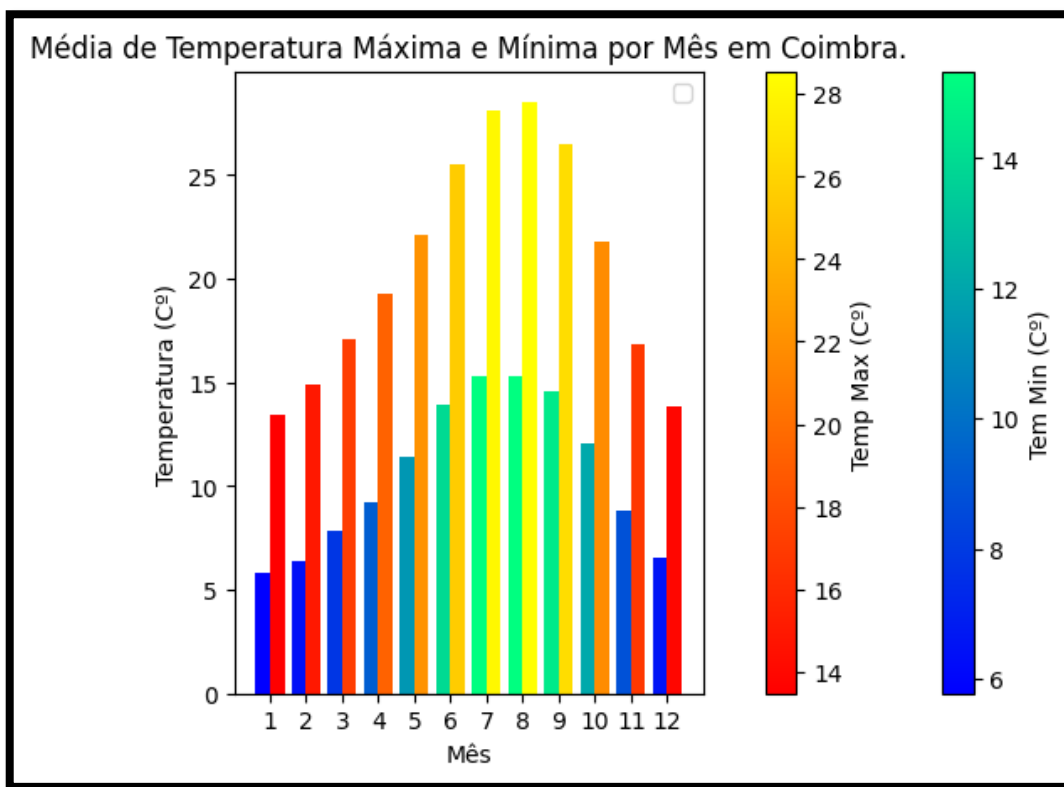


Figura 11 - Média da Temperatura Máxima e Mínima por Mês em Coimbra

Em relação às temperaturas médias são mais frias no inverno a rondar os 10°C e o pico é atingido no verão, que atinge pouco acima dos 20°C.

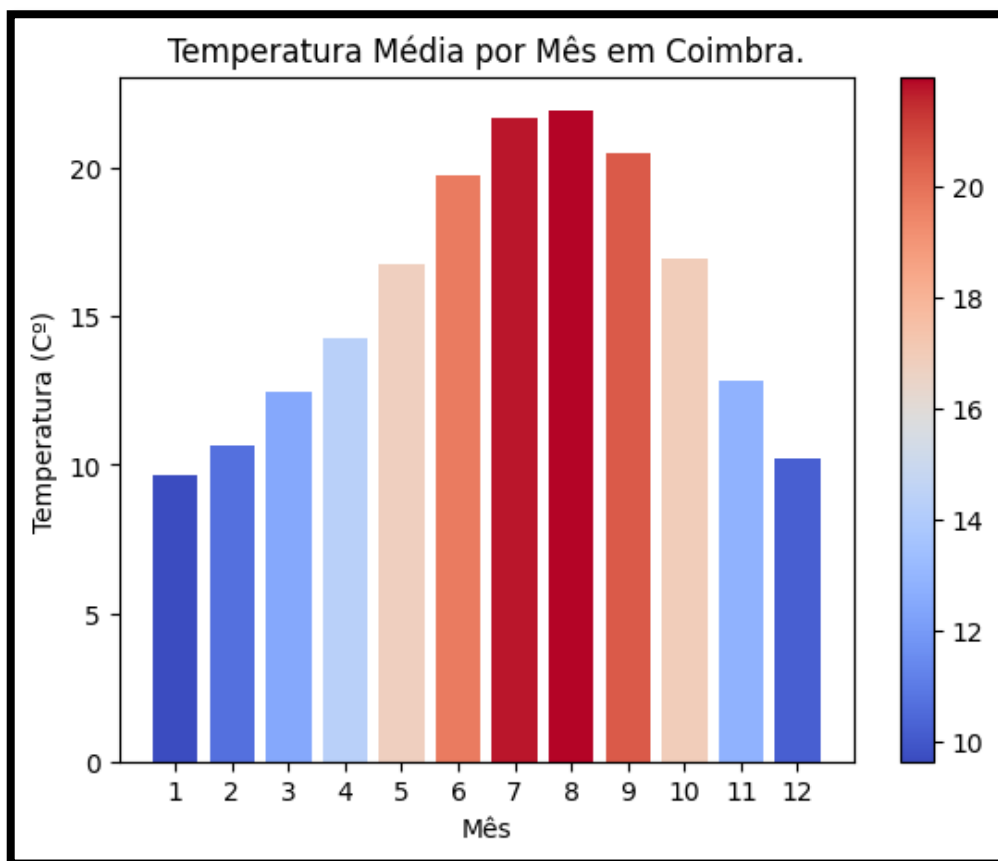


Figura 12 - Temperatura Média Mensal em Coimbra

4.3.2 Tendências da Temperatura e da Precipitação

Ao analisarmos a Figura 13, acerca da tendência da temperatura, reparamos numa ligeira tendência positiva.

O coeficiente da equação da regressão linear é igual a 0,2171, que indica uma relação positiva entre o tempo e a temperatura média. Isso significa que, em média, a cada 30 anos, espera-se um aumento de aproximadamente 0,2171 graus Celsius na temperatura média. O valor R-quadrado de 0,8419 indica que aproximadamente 84.2% da variabilidade na temperatura média pode ser explicada pela variável de tempo utilizada neste modelo de regressão linear. Isto sugere que o tempo é um fator muito relevante para prever a temperatura média, pois a maior parte da variabilidade é capturada pelo modelo.

Estes resultados indicam uma tendência de aumento na temperatura média ao longo do tempo, com um coeficiente positivo e um valor de R-quadrado relativamente alto.

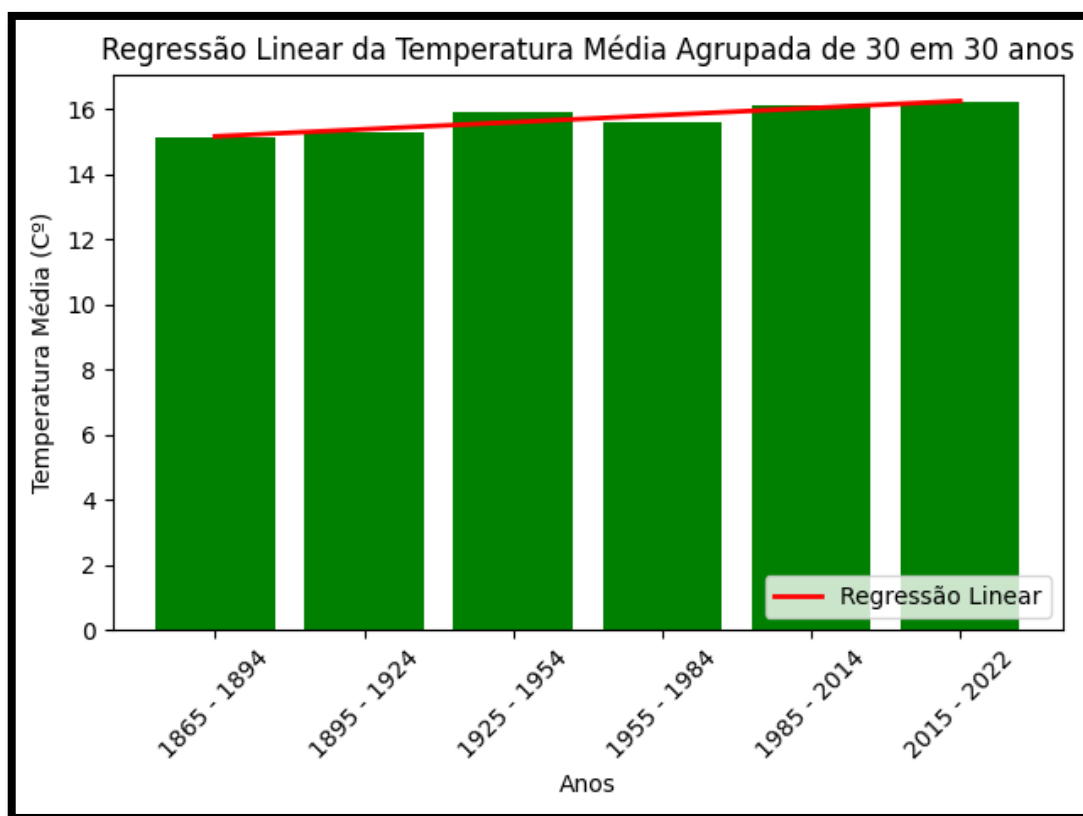


Figura 13 - Regressão Linear da Temperatura Média Agrupada de 30 em 30 anos

No caso da precipitação (Figura 14), o coeficiente da reta de regressão linear é de -16,1, o que indica uma relação negativa entre o tempo e a precipitação média. Isso significa que, em média, novamente de 30 em 30 anos espera-se uma diminuição de aproximadamente 16,1 (mm) na precipitação média. O valor R-quadrado de 0,3352 indica que aproximadamente 33,5% da variabilidade na precipitação média pode ser explicada pela variável de tempo utilizada neste modelo de regressão linear. Isso sugere que o tempo é um fator relevante, mas outros fatores podem influenciar a precipitação média além do tempo.

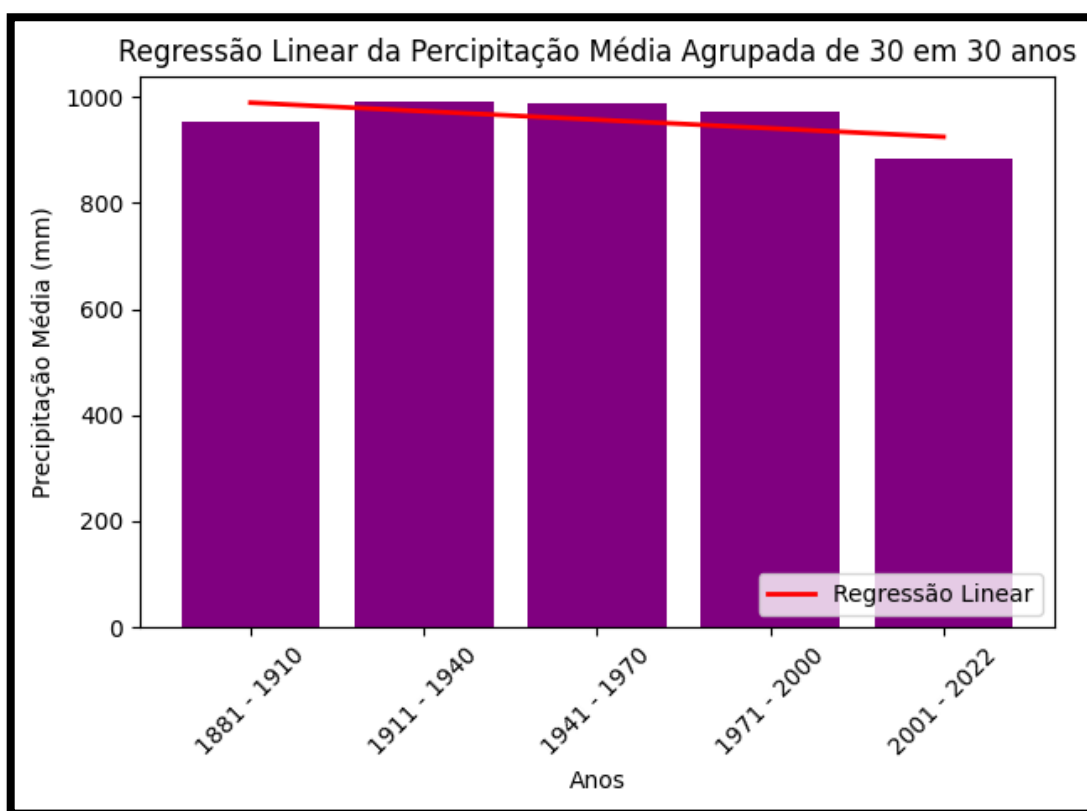


Figura 14 - Regressão Linear da Precipitação Média Agrupada de 30 em 30 anos

4.3.3 Análise das Características Ambientais da Cultura da Batata.

Para a análise dos intervalos ótimos de uma cultura, foi utilizada a batata como o exemplo para a metodologia da recomendação de data de plantação. A batata foi utilizada por dois motivos. Primeiro é uma das culturas mais cultivadas e muito consumida pelos portugueses. Segundo, dos dados das colheitas que foram recolhidos, a batata era aquela que continha dados mais uniformes e daí que a sua escolha irá permitir obter resultados imparciais.

Este segundo aspeto é muito importante nesta análise, pois um dos maiores problemas que foram detetados ao analisar os dados das culturas eram que à medida que o tempo passava a produção (kilos por hectare) também aumentava. Este fenómeno ocorreu por causa do progresso na mecanização da agricultura em Portugal, que ocorreu no final dos anos 80. Devido ao aumento na produtividade das colheitas por causa destes

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

fatores, relacionar a produtividade das mesmas com fatores climáticos tornou-se muito difícil.

A batata era a única cultura de uma lista de 11, que não apresentava um aumento constante durante o percurso de 20 anos (Figura 15). Por isso foi a cultura escolhida para examinar o impacto das condições climáticas.

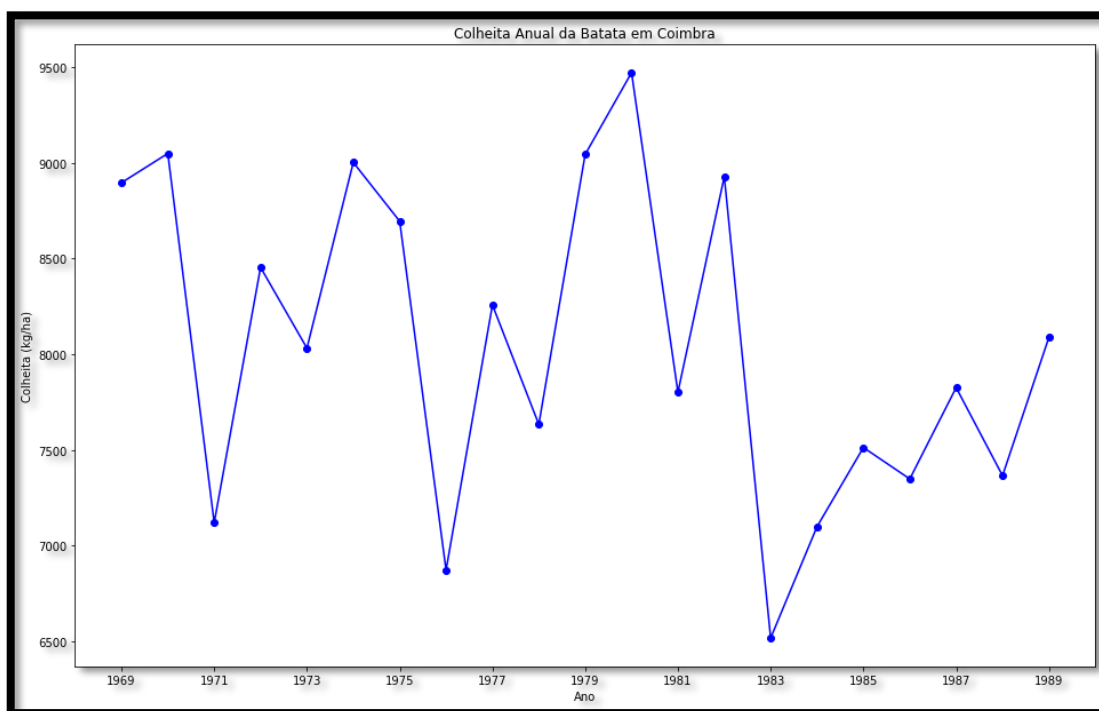


Figura 15 - Gráfico de Colheita Anual da Batata em Coimbra

Tabela 3 - Temperaturas Cardinais para a Batata (°C)

Temperatura Letal mínima	-2 a -4
Vegetação Mínima	7
Vegetação Ótima	15-25
Vegetação Máxima	30-35
Abrolhamento Mínima	2-5
Abrolhamento Ótima	18-20
Tuberização Mínima	7
Tuberização Ótima	Em dias curtos: 20-25 Em dias longos: 12
Tuberização Máxima	30

A metodologia utilizada consiste no cálculo do menor valor possível do MAE em relação ao intervalo das temperaturas ótimas. Para atingir este objetivo, o primeiro passo foi construir o gráfico das temperaturas ótimas da batata. A Figura 16 assume que o tipo de batata seja tardio, em relação à duração do ciclo, o que significa que a média do ciclo destes tipos de batatas dure 150 dias.

As fases de crescimento podem-se agrupar em 3 fases. O abrolhamento (fase inicial), a vegetação (fase intermédia) e a tuberização (fase final). O tempo aproximado para cada fase, para a batata tardia, é 30 dias para o abrolhamento, 90 dias para a vegetação e 30 dias para a tuberização. Cada fase de crescimento têm o seu próprio intervalo ótimo com base da Tabela 3. O gráfico da Figura 16 foi criado utilizando a referida tabela.



Figura 16 - Intervalo das Temperaturas Ótimas para cada Fase de Crescimento da Batata

Os 5 melhores anos e os 5 piores anos em termos de produtividade anual foram identificados e agrupados. Os dados da temperatura média mensal associados a esses anos foram somados e depois calculou-se a média. Assim ficou-se com duas novas variáveis que se podem visualizar no gráfico da Figura 17. A verde estão as temperaturas mensais médias registadas nos 5 anos mais produtivos e a vermelho as temperaturas mensais médias para os 5 piores anos. O propósito foi comparar as temperaturas observadas nos melhores anos com as dos piores anos e perceber se a proximidade dos valores registados das temperaturas aos intervalos ótimos reflete uma maior produtividade da cultura.

Como a temperatura é das variáveis que mais influenciam o comportamento das culturas, ir-se-á examinar o comportamento das mesmas face à produção da batata.

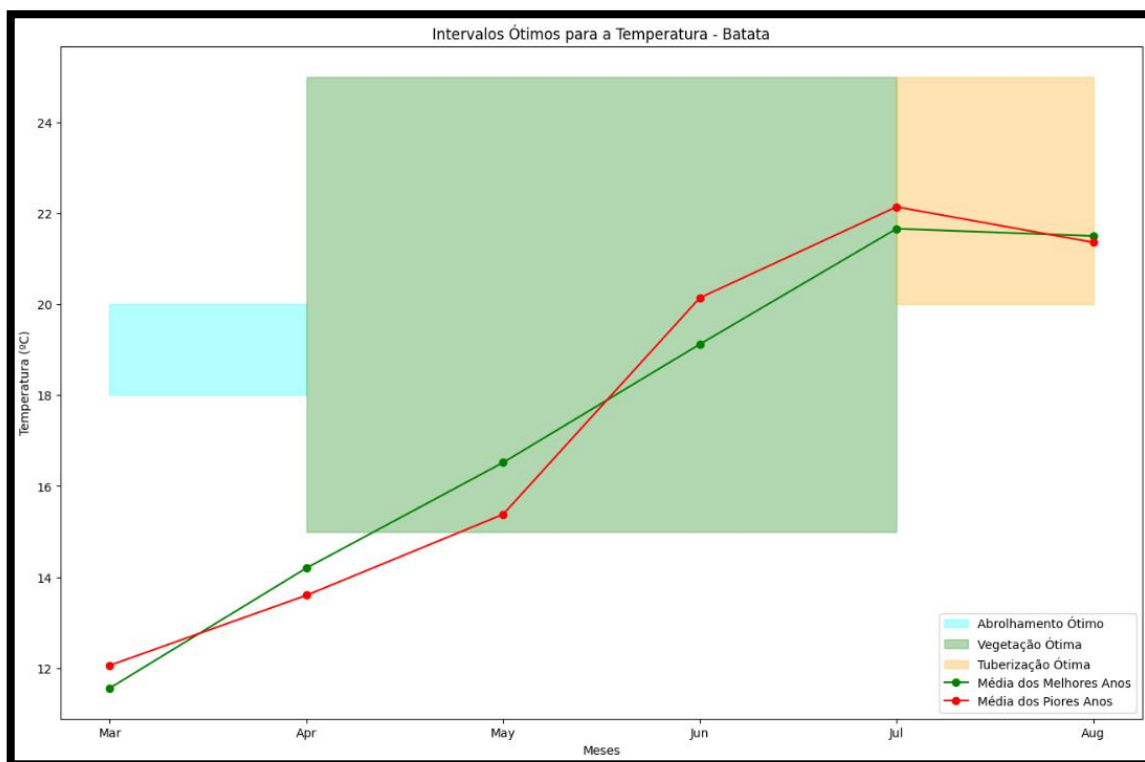


Figura 17 - Intervalos Ótimos face à Temperatura Média Mensal da Batata face à Produção

Como a época de plantação da batata começa nos meses de fevereiro e vai até a abril, o mesmo tipo de gráfico foi efetuado para os restantes meses. Ao calcular a distância da temperatura média mensal em relação aos intervalos, ou seja, o MAE, conseguiu-se avaliar se existe uma relação entre as temperaturas registadas e os intervalos ótimos.

Como quanto menor for o valor do MAE, melhor é o resultado, decidiu-se preencher a célula do menor valor a verde, e do valor maior registado a vermelho. O MAE_Melhor representa o MAE dos 5 melhores anos de produtividade, o MAE_Pior representa os piores anos.

Os resultados para o dia de plantação foram o seguinte.

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

Tabela 4 - Resultados do MAE para o Mês de fevereiro como o de plantação

Mês	MAE_Melhor	MAE_Pior
Fev	7,3	7,5
Mar	6,44	5,94
Abr	0,8	1,4
Mai	0	0
Jun	0	0
Total	<u>14,54</u>	14,84
Media	<u>2,908</u>	2,968

Tabela 5 - Resultados do MAE para o Mês de março como o de plantação

Mês	MAE_Melhor	MAE_Pior
Fev	6,44	5,94
Mar	3,8	4,4
Abr	0	0
Mai	0	0
Jun	0	0
Total	<u>10,24</u>	10,34
Media	<u>2,908</u>	2,968

Tabela 6 - Resultados do MAE para o Mês de abril como o de plantação

Mês	MAE_Melhor	MAE_Pior
Abr	3,8	4,4
Mai	1,48	2,62
Jun	0	0
Jul	0	0
Ago	0	0
Total	<u>5,28</u>	7,02
Media	<u>1,056</u>	1,404

Observando as tabelas 4, 5 e 6, podemos constatar que em todos os meses dos respetivos anos quando se observou um aumento da produtividade, registaram-se temperaturas mais próximas aos intervalos ótimos para a cultura da batata.

Após a análise dos intervalos ótimos da batata em relação às temperaturas médias mensais, foi realizada uma análise em relação às temperaturas médias diárias. Os objetivos foram os seguintes, ver se as aproximações da temperatura média diária aos intervalos ótimos da batata refletem uma melhor colheita dos produtores, medida em kilos por hectare. O processo foi o mesmo que foi feito para a comparação entre as temperaturas médias mensais, dos anos com melhores colheitas e dos com piores colheitas, em relação aos intervalos ótimos das temperaturas da batata (Figura 18).

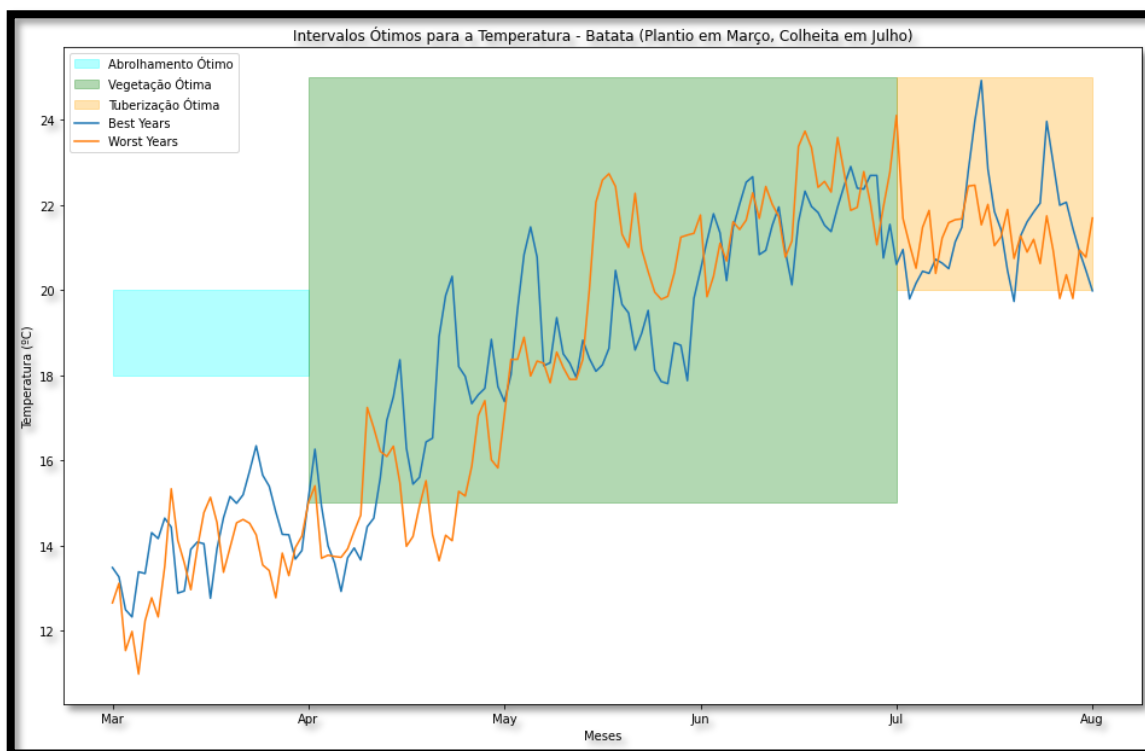


Figura 18 - Registo das Temperaturas Médias Diárias face aos Intervalos Ótimos da Batata

O primeiro dia de cada mês foi o usado para a plantação. Ou seja, na primeira linha o dia de plantação foi 1 de fevereiro, na segunda linha 1 de março e na terceira 1 de abril. Usando a temperatura média diária para comparar com os intervalos ótimos para cada fase de crescimento, obtivemos os resultados da Tabela 7.

Tabela 7 - Resultados do MAE em relação à Temperatura Média Diária

Dia de Plantação	Soma do MAE_Melhor	Soma do MAE_Pior	Média do MAE_Melhor	Média do MAE_Pior
1 de fevereiro	241	244.26	1.58	1.6
1 de março	128.29	150.46	0.84	0.98
1 de abril	68.1	92.32	<u>0.44</u>	0.6
Total	<u>437.39</u>	487.04	<u>2.86</u>	3.18
Média	<u>145.80</u>	162.35	<u>0.953</u>	1.06

Ao avaliar as tabelas verifica-se que nos anos onde houve melhores colheitas, também se verificaram temperaturas mais próximas dos intervalos ótimos, como evidenciado nas tabelas 4, 5, 6 e 7 deste capítulo. Isto verificou-se para todos as épocas de plantação, quer seja para as temperaturas médias mensais ou para as temperaturas médias diárias.

Também se consegue observar que destas 3 datas de plantação a que apresenta um MAE menor médio refere-se ao dia 1 de abril. Neste caso a melhor data de plantação entre os três seria 1 de abril.

O sistema de recomendação, recomenda a data que apresenta valor médio mais baixo do MAE, tendo em conta todos os dias de plantação do ano.

5 Desenvolvimento

5.1 Características do sistema

O sistema vai ter funcionalidades diferentes dependendo do tipo de utilizador (utilizador não registado, utilizador registado e administrador).

Dentro do sistema, são de salientar as seguintes funcionalidades (Figura 19):

- Visualização de Dados
- Gestão do Utilizador
- Recomendação da Data de Plantação
- Gestão da Base de Dados
- Gestão dos Dados
- Modelo de Inteligência Artificial

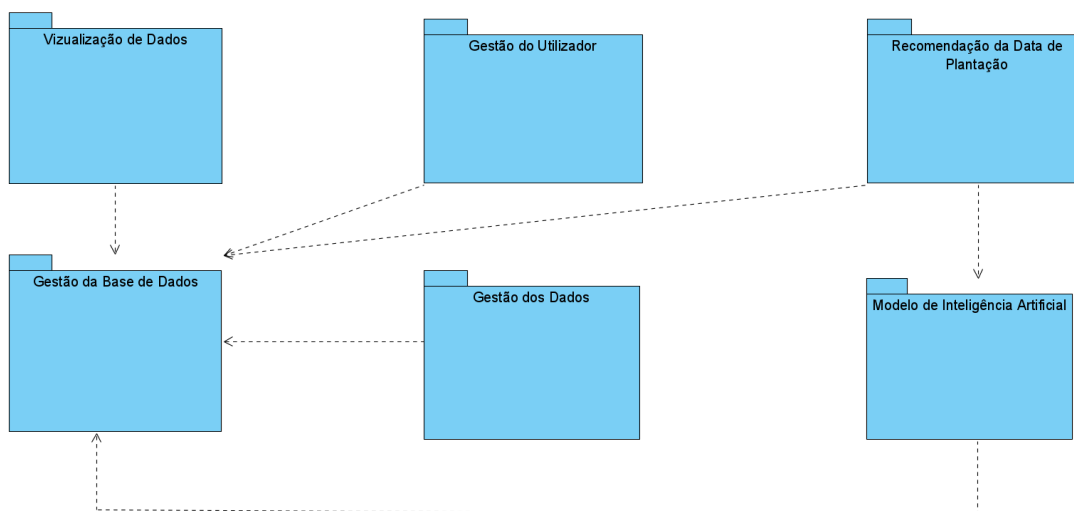


Figura 19 - Diagrama de Pacotes

Gestão de utilizadores - Inclui funcionalidades relacionadas com o registo de utilizadores, a autenticação e a gestão de perfis. Através destas funcionalidades, os utilizadores podem registar, efetuar *login*, *logout* e editar os seus dados pessoais. Garantindo a segurança e a facilidade de acesso, Gestão de Utilizadores serve como ponto de entrada para os utilizadores interagirem com as outras funcionalidades da aplicação.

Recomendação de datas de plantação – É responsável por fornecer aos utilizadores recomendações de datas para a plantação de culturas. Utiliza o modelo de IA para previsões e cruza-os com informações armazenadas sobre as culturas para gerar recomendações. Ao considerar fatores como as temperaturas previstas e as condições ideais de plantação, este pacote garante que os agricultores recebem as melhores sugestões para as suas plantações.

Visualização de Dados - A visualização é fundamental para tornar os dados compreensíveis. Oferece aos utilizadores um *dashboard* interativo onde se podem ver informações detalhadas sobre cada cultura, incluindo as condições ótimas de crescimento. Com funcionalidades como a pesquisa, os utilizadores podem navegar facilmente através de uma vasta quantidade de dados sobre as culturas, tomando decisões informadas com base em dados visuais. Também é responsável pela visualização dos dados climáticos dos sensores.

Gestão de dados – Assegura a integridade, a atualidade e a disponibilidade dos dados. Desde a sincronização dos dados de temperatura de fontes externas até à execução de operações CRUD (*Create, Read, Update and Delete*) sobre a informação das culturas.

Modelos de IA - É onde residem os modelos de previsão da aplicação. Estas funcionalidades são responsáveis pelo processamento das variáveis de entrada do modelo, execução de previsões e apresentação de resultados que podem ser utilizados por outras funcionalidades.

Base de dados - Atuando como a memória da aplicação, é responsável pelo armazenamento e recuperação de todos os dados essenciais. Quer se trate de perfis de utilizador, de informações detalhadas sobre as culturas ou de dados históricos de temperatura, garante ainda que os dados são armazenados de forma segura e que podem ser acedidos de forma eficiente.

5.2 Tipos dos Utilizadores

O sistema é suportado por três tipos de utilizadores, administrador, utilizador sem conta e utilizador registado. O administrador é um utilizador privilegiado e tem controlo sobre todo o sistema. É o responsável pela manutenção do sistema e das suas funcionalidades. Existem outros dois tipos de utilizador. Os utilizadores não registados e os utilizadores registados. Os utilizadores que acedem à plataforma sem se registarem ou iniciarem sessão têm um acesso limitado, centrado essencialmente na visualização de informações. Já os utilizadores que se encontram registados no sistema têm acesso a um conjunto adicional de funcionalidades. Na Tabela 8 estão apresentados os casos de uso a que cada tipo de utilizador tem acesso.

Tabela 8 - Atores e os seus respetivos Casos de Uso

Casos de Uso	Tipo de Utilizador		
	Administrador	Utilizador (não registado)	Utilizador (registado)
Registar	X	X	
Login	X		X
Recuperar Senha	X		X
Alterar Dados da Conta	X		X
Logout	X		X
Pesquisar Culturas	X	X	X
Visualizar Dados de Culturas	X	X	X
Visualizar Dados de Quintas	X		X
Visualizar Recomendações de Data de Plantação	X	X	X
Adicionar Cultura	X		

Adicionar Variáveis Ótimas	X		
Editar Culturas	X		
Eliminar Culturas	X		
Visualizar Previsões da Temperatura	X		X
Visualizar Dados Climáticos	X		X

5.3 Requisitos Funcionais

No desenvolvimento de software, é fundamental compreender e documentar os requisitos. Os requisitos funcionais desempenham um papel fundamental, uma vez que definem as ações e processos que um sistema ou aplicação deve executar. Para representar e comunicar visualmente esses requisitos, os diagramas UML serão utilizados. Este capítulo aborda a essência dos requisitos funcionais e a importância dos diagramas UML para facilitar o seu entendimento.

Os diagramas de casos de uso são ferramentas essenciais na engenharia de software, utilizadas para capturar os requisitos funcionais de um sistema. Eles ajudam a visualizar as funcionalidades de um sistema e as interações entre diferentes atores e o sistema.

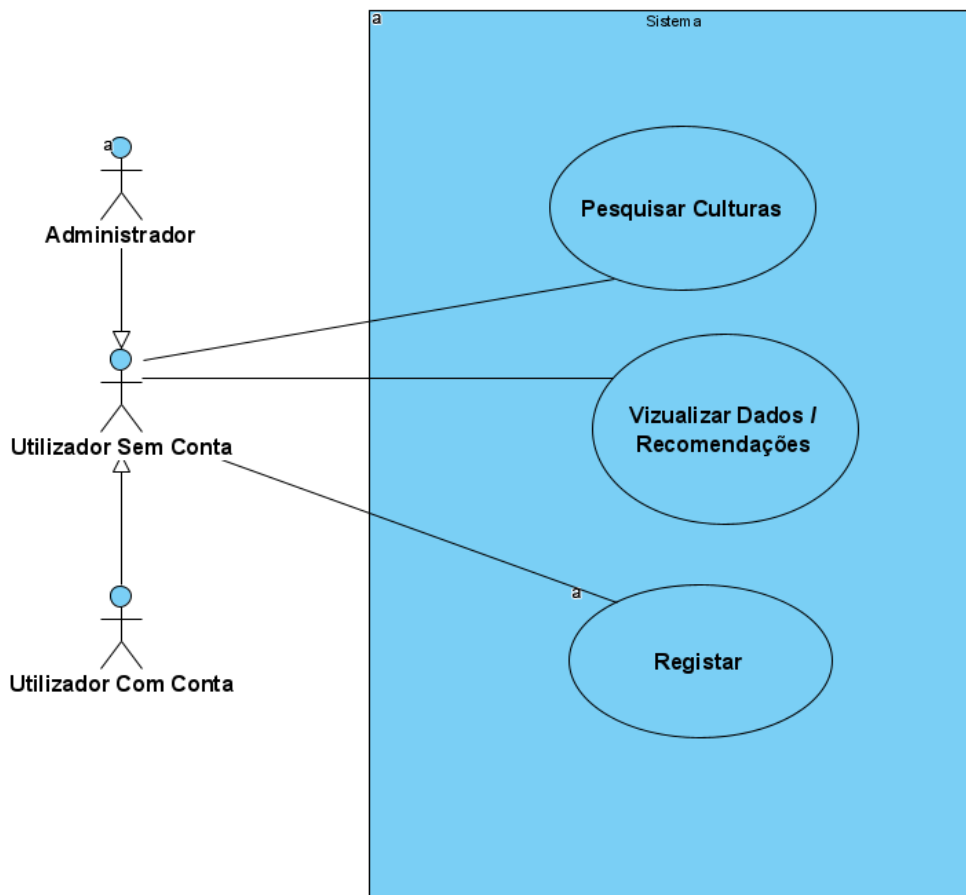


Figura 20 - Diagrama de Caso de Uso de Utilizador sem Conta

Este diagrama (Figura 20) demonstra que qualquer um dos utilizadores pode pesquisar diferentes culturas, visualizar dados e recomendações relacionadas, e que também têm a capacidade de registar-se na plataforma.

Tabela 9 - Cenário de Registrar Utilizador

Registrar Utilizador (Cenário Principal)	
Pré-Condição	O utilizador pretender efetuar o registo no sistema.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona a opção “Registrar”. 2. O utilizador deve inserir e submeter um e-mail, a password e a confirmação da password. 3. O sistema confirma se os dados estão dentro das normas. 4. Os dados preenchidos são enviados para a base de dados do sistema.
Pós-Condição	O utilizador está registado na plataforma com sucesso.
Cenário Secundário	Dados únicos já existentes.
Pós-Condição do Cenário Secundário	Autenticação negada. O sistema alerta o utilizador que já existe alguém com os mesmos dados inseridos.

Tabela 10 – Requisitos Funcionais do Cenário Registrar Utilizador

Requisitos	Descrição
RF1	É exibido um formulário para preenchimento com os dados do utilizador.
RF2	Cada formulário de utilizador deve conter um e-mail único.
RF3	A senha do utilizador deve ter no mínimo 8 caracteres.
RF4	O sistema encripta as senhas.

Tabela 11 - Cenário Pesquisar Culturas

Pesquisar Culturas (Cenário Principal)	
Pré-Condição	O utilizador pretende pesquisar informações sobre culturas específicas.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona a opção “Pesquisar Culturas”. 2. O utilizador insere o nome da cultura que deseja pesquisar. 3. O sistema procura na base de dados as correspondências relacionadas à pesquisa. 4. Os resultados da pesquisa são exibidos para o utilizador.
Pós-Condição	O utilizador visualiza as informações das culturas relacionadas à sua pesquisa.
Cenário Secundário	Nenhuma correspondência encontrada para a pesquisa.
Pós-Condição do Cenário Secundário	O sistema informa o utilizador que não foram encontrados resultados.

Tabela 12 - Requisitos Funcionais do Cenário Pesquisar Culturas

Requisitos	Descrição
RF5	É exibido um campo de pesquisa para a inserção do nome da cultura.
RF6	O sistema deve ser capaz de pesquisar na base de dados com base no texto inserido.
RF7	Os resultados da pesquisa devem ser exibidos de forma clara e organizada.
RF8	Em caso de não existirem resultados, uma mensagem adequada deve ser mostrada.

Tabela 13 - Cenário Visualizar Dados e Recomendações de Culturas

Visualizar Dados e Recomendações de Culturas (Cenário Principal)	
Pré-Condição	O utilizador está localizado na página de resultados da pesquisa.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona uma cultura da lista ou dos resultados da pesquisa. 2. O sistema recolhe os detalhes completos da cultura selecionada da base de dados. 3. As informações detalhadas da cultura são exibidas para o utilizador.
Pós-Condição	O utilizador visualiza as informações detalhadas da cultura selecionada.
Cenário Secundário	Detalhes da cultura não disponíveis.

Tabela 14 - Requisitos Funcionais do Cenário Visualizar Dados e Recomendações de Culturas

Requisitos	Descrição
RF9	O sistema deve exibir uma lista de culturas ou resultados de pesquisa para seleção.
RF10	Ao selecionar uma cultura, o sistema deve recuperar os seus detalhes completos.
RF11	As informações da cultura devem ser exibidas de forma clara e organizada.
RF12	Em caso de falta de detalhes, uma mensagem adequada deve ser mostrada.
RF13	Após a seleção da cultura o algoritmo de recomendação da data é ativado para calcular a data de plantação

O Administrador está em todos os três casos de uso, pois ele tem a capacidade de realizar todas as ações no sistema. Pode gerir as informações das culturas, como adicionar novas culturas, editar dados existentes, e eliminar culturas do sistema. Além disso, o administrador também pode adicionar variáveis que são ótimas para o crescimento e saúde das culturas.

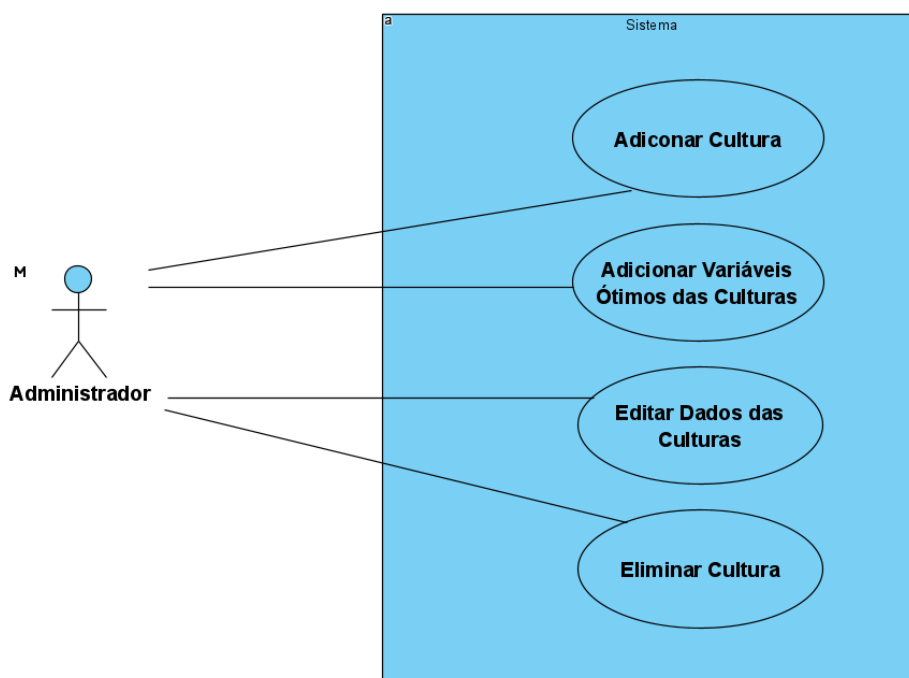


Figura 21 - Diagrama Caso de Uso Administrador

Tabela 15 - Cenário Adicionar Cultura

Adicionar Cultura (Cenário Principal)	
Pré-Condição	O administrador pretende adicionar uma nova cultura ao sistema.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o administrador seleciona a opção “Adicionar Cultura”. 2. É exibido um formulário para o administrador inserir os detalhes da nova cultura. 3. O administrador preenche o formulário e submete os dados. 4. O sistema valida os dados e guarda os na base de dados.
Pós-Condição	A nova cultura é adicionada com sucesso ao sistema.
Cenário Secundário	Dados inseridos são inválidos ou incompletos.
Pós-Condição do Cenário Secundário	Adição negada. O sistema alerta o administrador sobre os erros.

Tabela 16 - Requisitos do Cenário Adicionar Cultura

Requisitos	Descrição
RF14	É exibido um formulário para inserção dos detalhes da nova cultura.
RF15	O sistema deve validar os dados inseridos para garantir sua integridade.
RF16	A cultura adicionada deve ser armazenada na base de dados.
RF17	O sistema deve fornecer feedback sobre o sucesso ou falha da operação.

Tabela 17 - Cenário Adicionar Variáveis Ótimas da Cultura

Adicionar Variáveis Ótimas da Cultura (Cenário Principal)	
Pré-Condição	A cultura já está na base de dados.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o administrador seleciona a opção “Adicionar Variáveis Ótimas”. 2. É exibido um formulário para o administrador inserir as variáveis ótimas da cultura. 3. O administrador preenche o formulário e submete os dados. 4. O sistema valida e adiciona as variáveis à cultura correspondente na base de dados.
Pós-Condição	As variáveis ótimas são adicionadas com sucesso à cultura.
Cenário Secundário	Dados inseridos são inválidos ou incompletos.
Pós-Condição do Cenário Secundário	Adição negada. O sistema alerta o administrador sobre os erros.

Tabela 18 - Requisitos do Cenário Adicionar Variáveis Ótimas

Requisitos	Descrição
RF18	É exibido um formulário para inserção das variáveis ótimas da cultura.
RF19	O sistema deve validar os dados inseridos para garantir sua integridade.
RF20	As variáveis ótimas adicionadas devem ser armazenadas na base de dados associadas à cultura correspondente.
RF21	O sistema deve fornecer feedback sobre o sucesso ou falha da operação.

Tabela 19 - Cenário Editar Cultura

Editar Dados das Culturas (Cenário Principal)	
Pré-Condição	O administrador pretende editar os detalhes de uma cultura existente.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o administrador seleciona a cultura que deseja editar. 2. É exibido um formulário preenchido com os detalhes atuais da cultura. 3. O administrador modifica os detalhes conforme seja necessário e submete as alterações. 4. O sistema valida e atualiza os detalhes da cultura na base de dados.
Pós-Condição	Os detalhes da cultura são atualizados com sucesso.
Cenário Secundário	Dados modificados são inválidos ou incompletos.
Pós-Condição do Cenário Secundário	Edição negada. O sistema alerta o administrador sobre os erros.

Tabela 20 - Requisitos Funcionais do Cenário Editar Cultura

Requisitos	Descrição
RF22	O sistema deve exibir um formulário preenchido com os detalhes atuais da cultura selecionada.
RF23	O sistema deve validar os dados modificados para garantir sua integridade.
RF24	Os detalhes modificados devem ser atualizados na base de dados.
RF25	O sistema deve fornecer feedback sobre o sucesso ou falha da operação.

Tabela 21 - Cenário Eliminar Culturas

Eliminar Culturas (Cenário Principal)	
Pré-Condição	O administrador pretende eliminar uma cultura do sistema.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o administrador seleciona a cultura que deseja eliminar. 2. O sistema solicita confirmação para a eliminação. 3. O administrador confirma a eliminação. 4. A cultura é removida da base de dados.
Pós-Condição	A cultura é eliminada com sucesso do sistema.
Cenário Secundário	O administrador cancela a eliminação.
Pós-Condição do Cenário Secundário	A cultura permanece intacta no sistema.

Tabela 22 - Requisitos Funcionais do Cenário Eliminar Cultura

Requisitos	Descrição
RF26	O sistema deve solicitar confirmação antes de eliminar uma cultura.
RF27	Após confirmação, a cultura deve ser removida da base de dados.
RF28	O sistema deve fornecer feedback sobre o sucesso ou falha da operação.

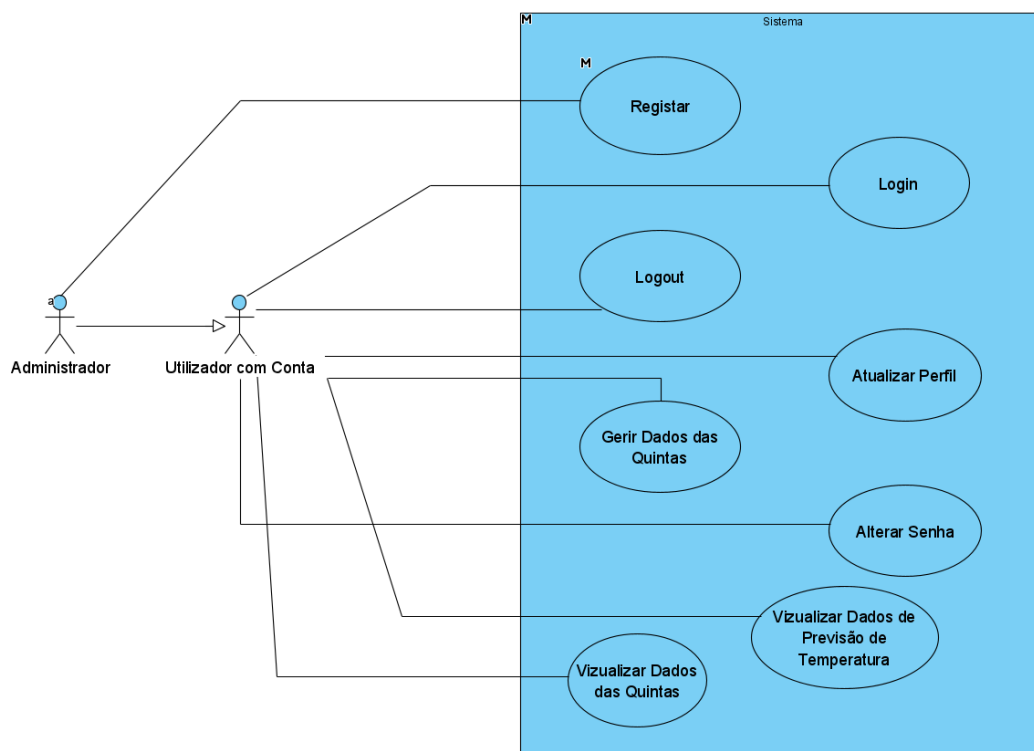


Figura 22 - Diagrama de Caso de Uso do Utilizador Registrado

O diagrama de caso de uso de um utilizador (Figura 22) ilustra as funcionalidades do utilizador principal. As funcionalidades ou casos de uso incluem registar, login, atualizar perfil, gerir dados das quintas, alterar senha e visualizar dados de previsão da temperatura. Cada um desses casos de uso representa uma operação ou ação específica que os atores podem realizar no sistema. Todos estes casos são partilhados com o administrador.

Tabela 23 - Cenário Fazer Login

Fazer Login (Cenário Principal)	
Pré-Condição	O utilizador tem conta validada no sistema.
Descrição	<ol style="list-style-type: none"> 1. O utilizador acede ao software. 2. É solicitada a introdução do e-mail e a password. 3. O utilizador insere os dados válidos e seleciona a opção “Entrar”. 4. O sistema valida os dados inseridos.
Pós-Condição	O utilizador acede à aplicação.
Cenário Secundário	O utilizador insere os dados incorretos.
Pós-Condição do Cenário Secundário	Autenticação negada. O sistema alerta o utilizador da inserção de dados inválidos.

Tabela 24 – Requisitos Funcionais do Cenário Fazer Login

Requisitos	Descrição
RF29	O sistema deve fornecer uma opção clara para iniciar a sessão.
RF30	O sistema deve validar as credenciais fornecidas considerando o que está guardado na base de dados.
RF31	O sistema deve fornecer feedback sobre o sucesso ou falha da autenticação.
RF32	O sistema redireciona para a página pessoal do utilizador

Tabela 25 - Cenário Fazer Logout

Fazer Logout (Cenário Principal)	
Pré-Condição	O utilizador está autenticado e pretende terminar a sua sessão.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona a opção “Logout”. 2. O sistema termina a sessão do utilizador.

	3. O utilizador é redirecionado para a página inicial.
Pós-Condição	O utilizador acede à aplicação.

Tabela 26 - Requisitos Funcionais do Cenário Logout

Requisitos	Descrição
RF33	O sistema deve fornecer uma opção clara para terminar a sessão.
RF34	O sistema deve terminar a sessão do utilizador e garantir que os dados da sessão sejam limpos.
RF35	O utilizador deve ser redirecionado para a página inicial após terminar a sessão.

Tabela 27 - Cenário Editar Perfil

Atualizar Perfil (Cenário Principal)	
Pré-Condição	O utilizador pretende atualizar os detalhes do seu perfil.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona a opção “Atualizar Perfil”. 2. É exibido um formulário preenchido com os detalhes atuais do perfil do utilizador. 3. O utilizador modifica os detalhes conforme necessário e submete as alterações. 4. O sistema valida e atualiza os detalhes do perfil na base de dados.
Pós-Condição	Os detalhes do perfil do utilizador são atualizados com sucesso.
Cenário Secundário	Dados modificados são inválidos ou incompletos.
Pós-Condição do Cenário Secundário	Atualização negada. O sistema informa o utilizador sobre os erros.

Tabela 28 - Requisitos Funcionais do Cenário Editar Perfil

Requisitos	Descrição
RF36	O sistema deve exibir um formulário preenchido com os detalhes atuais do perfil do utilizador.
RF37	O sistema deve validar os dados modificados para garantir sua integridade.
RF38	Os detalhes modificados devem ser atualizados na base de dados.
RF39	O sistema deve fornecer feedback sobre o sucesso ou falha da atualização.

Tabela 29 - Cenário Gerir Dados da Quinta

Gerir Dados da Quinta (Cenário Principal)	
Pré-Condição	O utilizador pretende gerir os dados da sua quinta.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona a opção “Gerir Dados da Quinta”. 2. É exibido um painel com os dados atuais da quinta. 3. O utilizador pode adicionar, modificar ou eliminar dados conforme seja necessário. 4. As alterações são guardadas na base de dados.
Pós-Condição	Os dados da quinta são geridos com sucesso.
Cenário Secundário	Dados inseridos/modificados são inválidos ou incompletos.
Pós-Condição do Cenário Secundário	Operação negada. O sistema informa o utilizador sobre os erros.

Tabela 30 - Requisitos Funcionais do Cenário Gerir Dados da Quinta

Requisitos	Descrição
RF40	O sistema deve exibir um painel com os dados atuais da quinta.
RF41	O utilizador deve ter opções para adicionar, editar ou eliminar dados da quinta.
RF42	O sistema deve validar os dados inseridos/editados para garantir sua integridade.
RF43	As alterações devem ser guardadas na base de dados.
RF44	O sistema deve fornecer feedback sobre o sucesso ou falha da operação.

Tabela 31 - Cenário Alterar Senha

Alterar Password (Cenário Principal)	
Pré-Condição	O utilizador pretende alterar a sua senha.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona a opção “Alterar Senha”. 2. É exibido um formulário para inserir a senha atual e a nova senha. 3. O utilizador preenche o formulário e submete a alteração. 4. O sistema valida a senha atual e, se estiver correta, atualiza a senha na base de dados.
Pós-Condição	A senha do utilizador é alterada com sucesso.
Cenário Secundário	Senha atual incorreta ou nova senha inválida.
Pós-Condição do Cenário Secundário	Alteração negada. O sistema informa o utilizador sobre os erros.

Tabela 32 - Requisitos Funcionais do Cenário Alterar Senha

Requisitos	Descrição
RF45	O sistema deve fornecer campos para inserção da senha atual e da nova senha.
RF46	O sistema deve validar a senha atual consultado o que está guardado na base de dados.
RF47	A nova senha deve ser validada para garantir sua integridade e segurança.
RF48	Se a senha atual estiver correta e a nova senha for válida, a senha deve ser atualizada na base de dados.
RF49	O sistema deve fornecer feedback sobre o sucesso ou falha da alteração.

Tabela 33 - Cenário Visualizar Dados da Quinta

Visualizar Dados da Quinta (Cenário Principal)	
Pré-Condição	O utilizador pretende visualizar os dados da sua quinta
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona a opção “Visualizar Dados da Quinta”. 2. O sistema recupera e exibe os dados da quinta do utilizador. 3. O utilizador pode visualizar e analisar os dados apresentados.
Pós-Condição	O utilizador visualiza os dados da sua quinta com sucesso.

Tabela 34 - Requisitos do Cenário Visualizar Dados da Quinta

Requisitos	Descrição
RF50	O sistema deve recuperar os dados da quinta associados ao utilizador.
RF51	Os dados da quinta devem ser exibidos de forma clara e organizada.

Tabela 35 - Cenário Visualizar Dados de Previsão de Temperatura

Visualizar Dados de Previsão de Temperatura (Cenário Principal)	
Pré-Condição	O utilizador pretende visualizar a previsão de temperatura para a sua localização.
Descrição	<ol style="list-style-type: none"> 1. O caso começa quando o utilizador seleciona a opção “Visualizar Dados da Temperatura”. 2. O sistema recupera e exibe a previsão de temperatura para a localização da quinta do utilizador. 3. O utilizador pode visualizar e analisar a previsão apresentada.
Pós-Condição	O utilizador visualiza a previsão de temperatura com sucesso.
Cenário Secundário	Previsão de temperatura não disponível.
Pós-Condição do Cenário Secundário	O sistema informa o utilizador que a previsão de temperatura não está disponível no momento.

Tabela 36 - Requisitos Funcionais do Cenário Visualizar Dados de Previsão

Requisitos	Descrição
RF52	O sistema deve recuperar a previsão de temperatura para a localização da quinta associada ao utilizador.
RF53	A previsão de temperatura deve ser exibida de forma clara e organizada.
RF54	O utilizador deve ser capaz de visualizar e analisar a previsão apresentada.

5.4 Requisitos Não Funcionais

Aqui estão os tipos de requisitos não funcionais que se pode esperar na plataforma:

Desempenho (D) - Esses requisitos dizem respeito ao desempenho do sistema, incluindo velocidade de resposta, tempo de carregamento de páginas e taxa de transferência de dados.

Segurança (S) - Requisitos de segurança abrangem a proteção de dados, autenticação, autorização e medidas de segurança gerais.

Usabilidade (U) - Esses requisitos se relacionam à facilidade de uso do sistema, incluindo a interface do utilizador e a acessibilidade.

Confiabilidade (C) - Requisitos de confiabilidade envolvem a disponibilidade do sistema, recuperação e tolerância a falhas.

Interoperabilidade (I) - Requisitos de interoperabilidade especificam como o sistema deve interagir com outras aplicações ou sistemas.

Tabela 37 - Requisitos Não Funcionais

Requisitos	Tipo de requisito	Descrição
RNF1	D	O sistema consegue suportar até 10000 utilizadores em simultâneo.
RNF2	C	O sistema deve ser capaz de identificar uma anomalia e reportar o problema aos administradores no máximo de 5 segundos.
RNF3	S	Proteção dos dados pessoais dos utilizadores conforme o RGPD.
RNF4	C	O sistema deverá fazer backups diários às bases de dados para prevenir perdas de dados.
RNF5	C	O sistema estará sempre disponível 24h por dia.

RNF6	D	O sistema deve ser atualizado para manter sempre um bom desempenho.
RNF7	C	O sistema poderá entrar em manutenção quando existir erros ou anomalias não previstas.
RNF8	U	O sistema deve ser testado para garantir que ele execute como pretendido, antes de ser exposto ao público.
RNF9	C	O sistema envia um keep alive para que os sensores uma vez por hora.
RNF10	D	O sistema deve ser capaz de lidar com até 1000 solicitações de utilizadores simultaneamente sem uma degradação significativa no desempenho.
RNF11	C	Caso um dos sensores não responde ao keep live, o sistema notifica o administrador.
RNF13	U	A interface do utilizador deve ser intuitiva e compatível com vários formatos de tela (telemóvel, tablet, computador...).
RNF14	I	O sistema deve ser capaz de receber dados meteorológicos de diferentes fontes de terceiros usando padrões de protocolo comuns.

5.5 Design do User Interface

O design desejado da plataforma é suposto ser o mais simples e intuitivo possível, composto por uma página inicial apenas com uma barra de pesquisa para o utilizador inserir a cultura desejada e por um *side* menu que redireciona para outras páginas acerca do projeto. Na figura 23 está apresentado o *mockup* da página inicial.

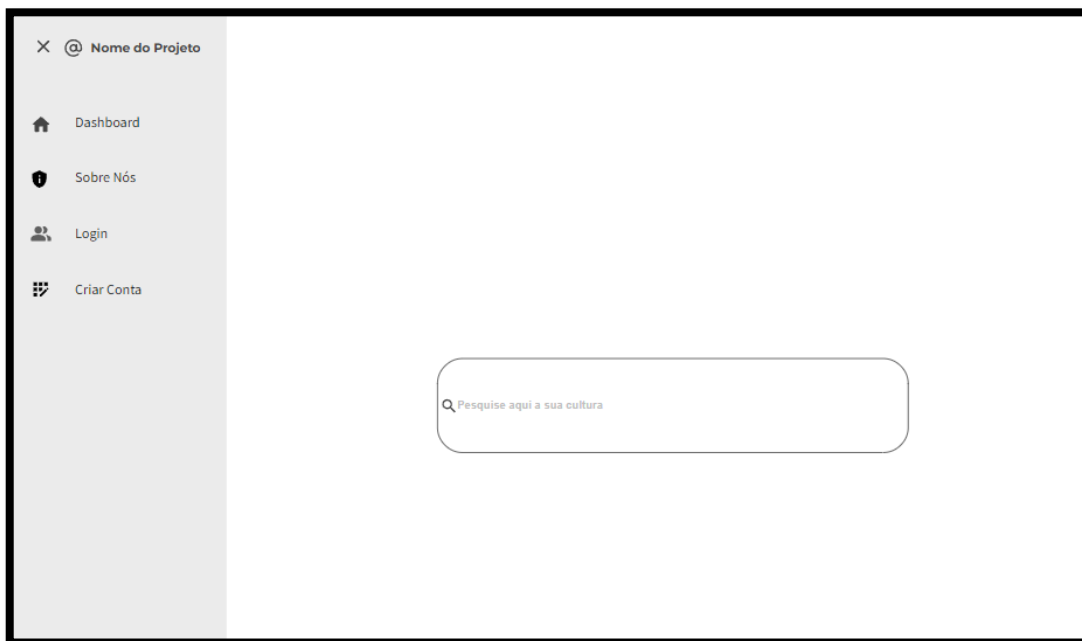


Figura 23 – Mockup Página Inicial

A página principal da plataforma é o dashboard (Figura 24) onde se pode visualizar toda a informação sobre as culturas, disponíveis no sistema. Composto por gráficos e outros tipos de dados e informação. Nela será possível verificar características comuns para cada cultura. Gráficos das temperaturas ótimas, informação geral sobre a cultura, condições do solo mais favoráveis entre outros. Esta é a página com mais funcionalidades do sistema.

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

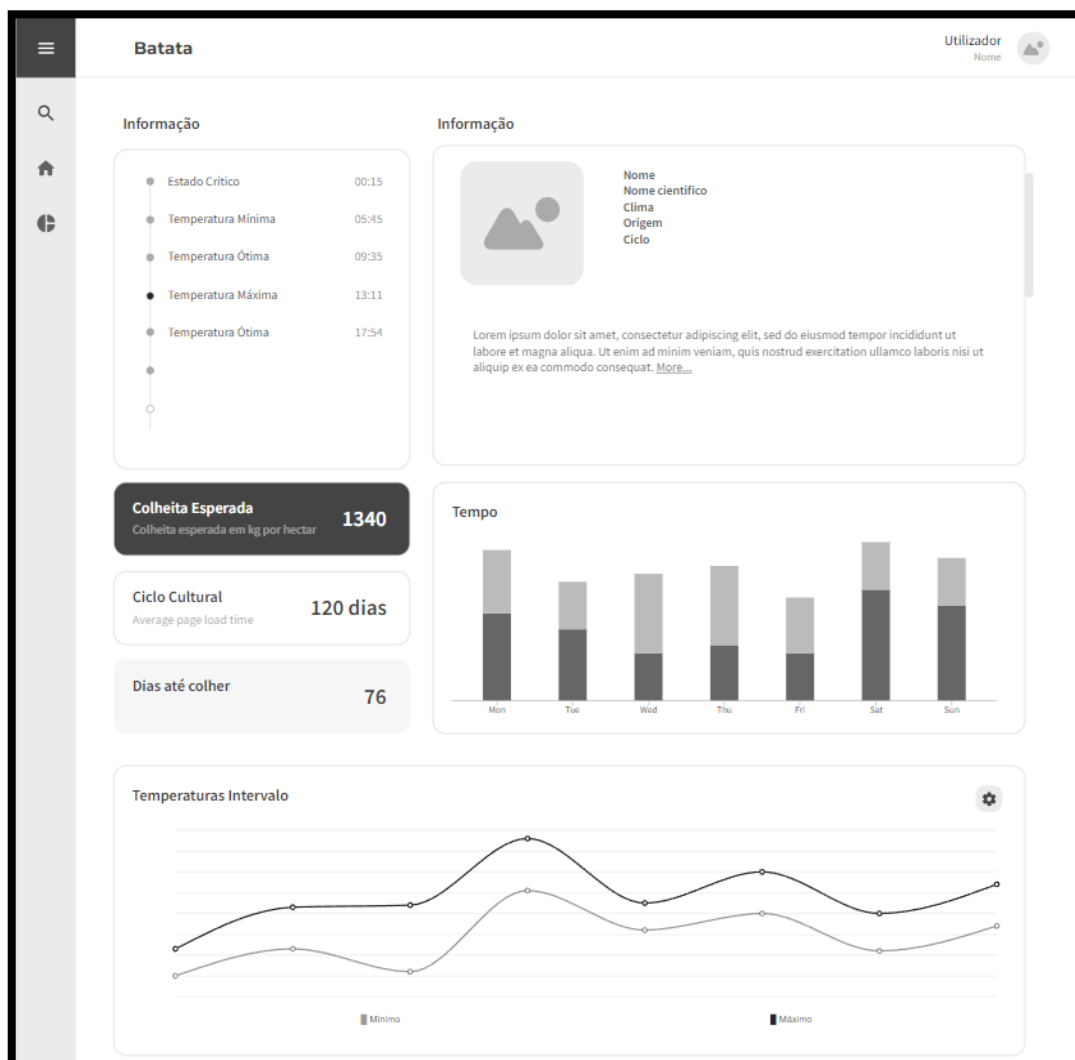


Figura 24 - Página 1 do Dashboard

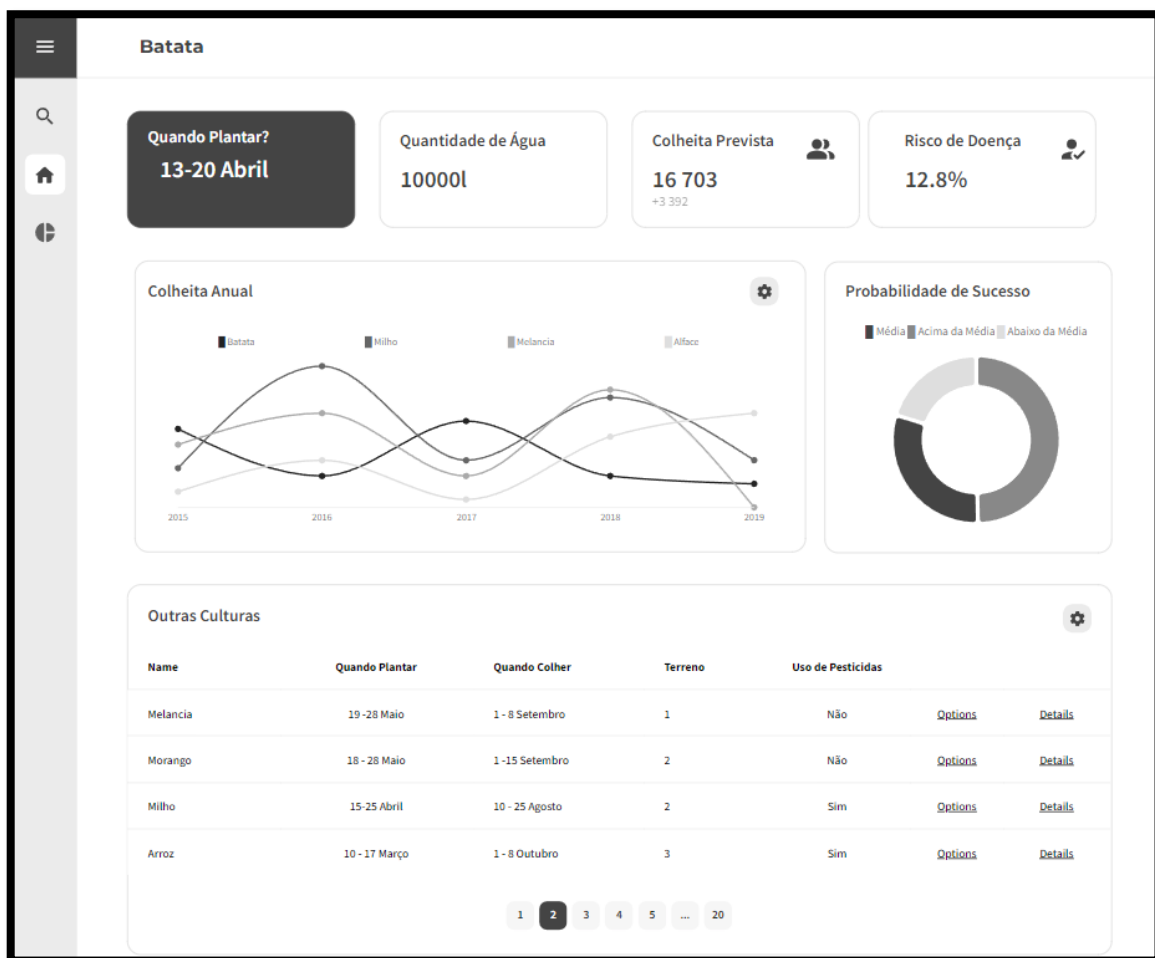


Figura 25 - Página 2 do Dashboard

5.4. Arquitetura do Sistema

O Sistema de Recomendação de Culturas foi projetado com uma arquitetura escalável, permitindo uma fácil manutenção e expansão futura. A arquitetura adotada é baseada em camadas, onde cada camada tem uma responsabilidade específica, garantindo a separação de preocupações e facilitando a integração de novos componentes ou a substituição de componentes existentes. O sistema divide-se por 4 grupos, *frontend*, *backend*, *IoT* e *ML*.

O *frontend* foi desenvolvido usando a framework Django. Esta é a *interface* principal através da qual os utilizadores interagem com o sistema. Ela apresenta visualizações de dados geradas com a biblioteca Plotly para uma representação gráfica clara e interativa das informações.

O backend suporta a lógica principal da aplicação e também foi desenvolvido usando o Django. Ele é responsável pelas solicitações do utilizador, processa os dados e interage com a base de dados e outros serviços.

No IoT estão representados todos os dispositivos e plataformas que coletam e transmitem dados em tempo real para o sistema. Inclui os sensores que coletam dados e a plataforma The Things Network, que facilita a transmissão desses dados para o sistema. The Things Network envia dados através de webhooks para a base de dados.

Por fim, temos a secção de *ML*. É dedicado ao desenvolvimento, treinamento e implementação de modelos de DL. Uma vez treinados, os modelos são armazenados e geridos pelo Django Rest Framework, tornando-os acessíveis para a camada do *backend*. Caso no futuro queiramos desenvolver uma aplicação *mobile*, ou disponibilizar a uma entidade as previsões, usar um RestAPI como o Django Rest Framework, facilita essa integração. Na figura abaixo pode-se visualizar a arquitetura do sistema.

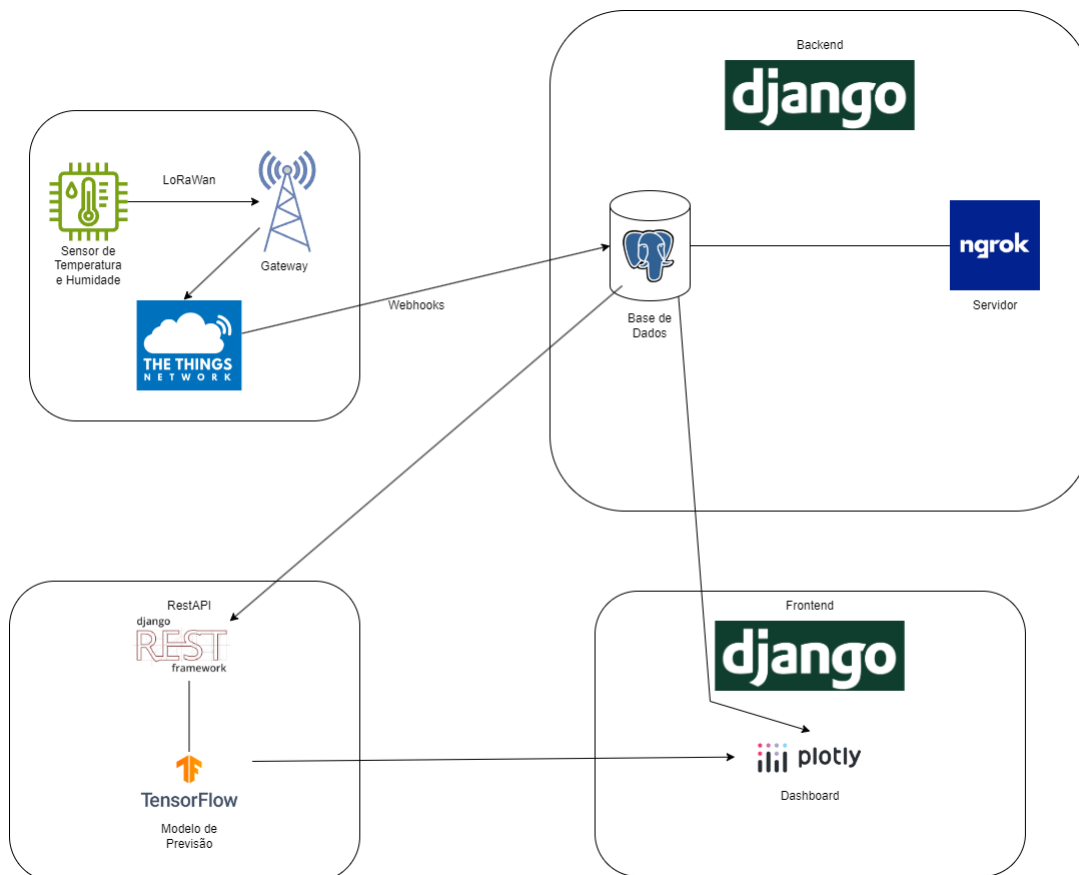


Figura 26 - Arquitetura do Sistema

5.6 Aplicação Web

5.6.1 Aplicações no Django

No contexto do Django, uma aplicação é uma forma de organizar funcionalidades relacionadas dentro de um projeto. Cada aplicação é projetada para fazer uma coisa. Esta abordagem modular permite que os desenvolvedores estruturam seus projetos de forma clara e lógica, facilitando a manutenção, a escalabilidade e a reutilização do código. Cada aplicação no Django é, em essência, um módulo autónomo. Ou seja, significa que ela contém tudo o que é necessário para essa funcionalidade específica, desde modelos de base de dados, *views*, *templates* e até ficheiros estáticos. Esta encapsulação garante que

cada aplicação possa ser desenvolvida e testada de forma independente. O meu projeto é composto por 4 aplicações.

Accounts - Esta aplicação é responsável pela gestão de utilizadores, autenticação e autorização. Ela contém modelos para utilizadores, grupos e permissões, bem como *views* e *templates* para o login, o registo e a gestão de perfis.

Climate - Esta aplicação é dedicada à recolha, processamento e visualização de dados climáticos. Ela interage com APIs externas (como a API do IPMA), sensores e outros sistemas para obter dados de temperatura.

Plants - Esta aplicação foca-se na gestão e visualização de informações relacionadas com as culturas. Contém detalhes sobre diferentes culturas, suas necessidades climáticas, épocas de plantação e outras informações relevantes.

Location - Esta aplicação lida com tudo relacionado à localização da quinta.

5.6.2 Autenticação, Login e Registo

A autenticação é uma parte crucial de muitas aplicações web, garantindo que os utilizadores sejam quem dizem ser e tenham acesso apenas aos recursos apropriados. O Django oferece um sistema de autenticação integrado que facilita a implementação de processos de login, registo e gestão de utilizadores. Isso inclui modelos e formulários para utilizadores, grupos, permissões e muito mais. Isso significa que não é necessário construir um sistema de autenticação do zero, economizando tempo e garantindo segurança.

O processo de registo permite que novos utilizadores criem uma conta na aplicação. O Django fornece formulários prontos para registo, mas eles podem ser personalizados conforme necessário. Neste projeto, pretendia-se simplificar o mais simples possível para tornar a experiência de registar uma conta, um processo rápido e fácil, por isso apenas se solicita aos novos utilizadores três inputs. O e-mail, a senha e a confirmação da senha. Para o login são apenas necessários o email e a senha.

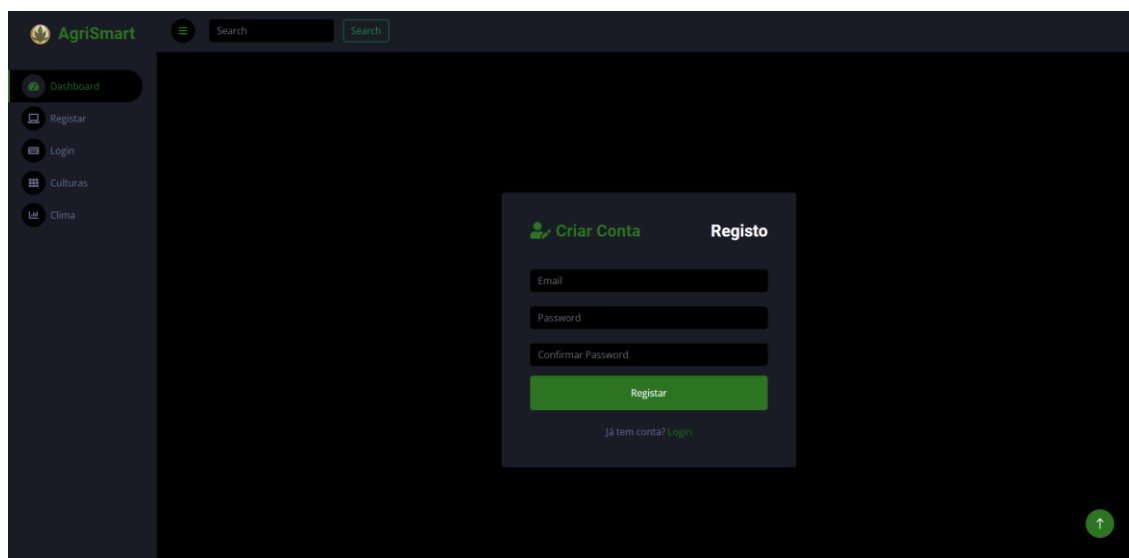


Figura 27 - Página de Registrar Utilizador

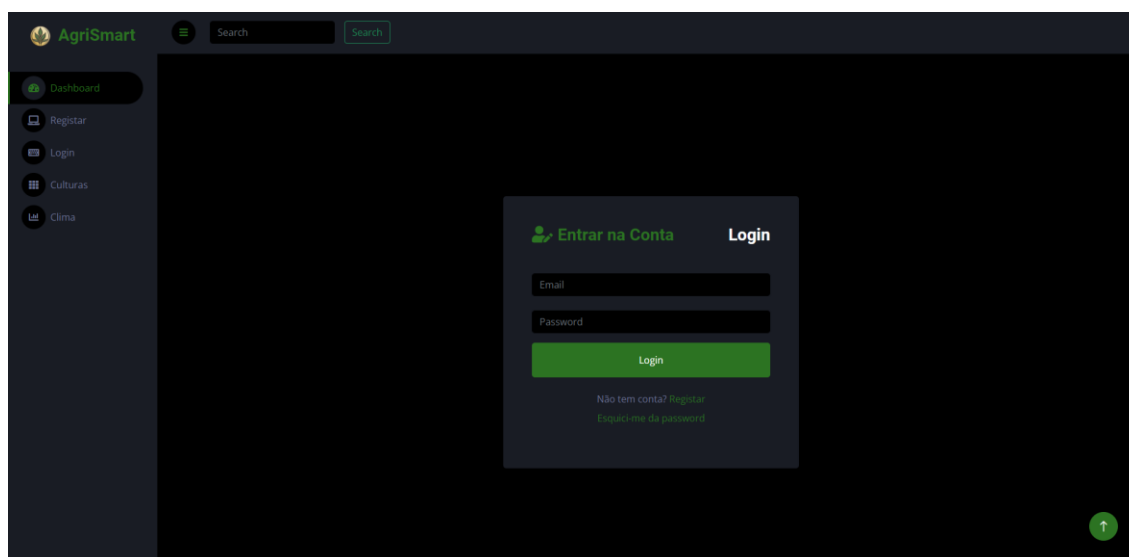


Figura 28 - Página de Login de Utilizador

Em relação à segurança o Django também oferece proteção contra muitos ataques comuns, como ataques de força bruta e CSRF (*Cross-Site Request Forgery*). As senhas dos utilizadores também são armazenadas de forma segura usando *hashing*.

5.6.3 Painel de Administrador e Gestão de Dados

O Django é amplamente reconhecido pela sua *interface* administrativa integrada. Esta *interface* fornece uma maneira rápida e eficiente para os administradores gerirem o conteúdo do site sem a necessidade de desenvolver um painel administrativo personalizado a partir do zero. Vamos explorar como o painel administrativo do Django pode ser utilizado para operações CRUD relacionadas aos dados das culturas.

Operações CRUD para Dados das Culturas:

Criar (*Create*) - Adicionar novas informações sobre culturas é simples. No painel administrativo, o administrador pode preencher um formulário com os detalhes da cultura, como nome, época de plantação, requisitos de temperatura e muito mais. Uma vez preenchido, os dados são armazenados na base de dados.

Ler (*Read*) - O painel administrativo oferece uma visão dos dados das culturas armazenados na base de dados. Esta visão permite uma rápida navegação, pesquisa e filtragem, facilitando a localização de informações específicas.

Atualizar (*Update*) - Modificar informações existentes é igualmente simples. Ao selecionar uma cultura específica, o administrador é apresentado com um formulário preenchido com os dados atuais da cultura. Após fazer as alterações necessárias, os dados atualizados são guardados na base de dados.

Deletar (*Delete*) - Se necessário, o administrador pode remover informações de culturas da base de dados. Uma confirmação é geralmente solicitada para garantir que os dados não sejam excluídos acidentalmente.

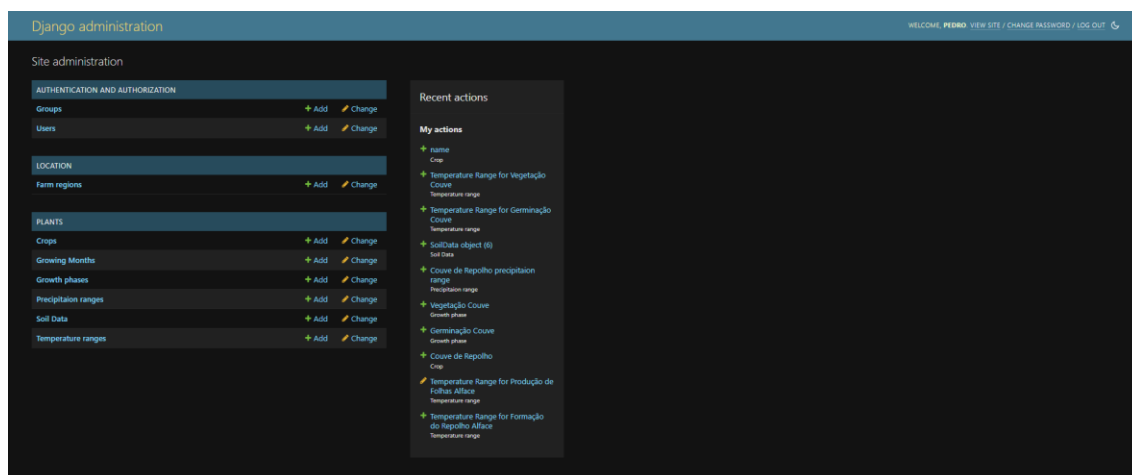


Figura 29 - Página Inicial do Administrador

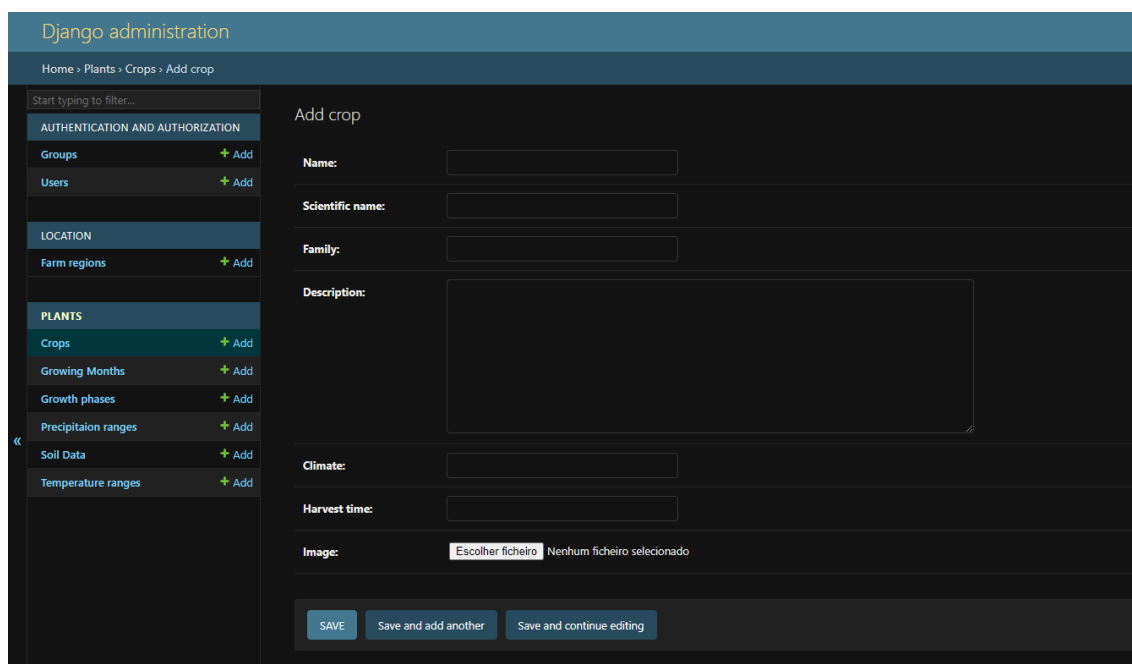


Figura 30 - Página do Administrador para Inerir Dados Novos das Culturas

5.6.4 Visualização de Dados com Plotly e Bootstrap no Frontend

A visualização de dados é uma ferramenta poderosa para transmitir informações complexas de forma intuitiva e compreensível. No contexto da aplicação, a visualização de dados desempenha um papel crucial ao fornecer *insights* sobre as condições ideais para o cultivo de culturas. Seguidamente irá ser apresentada a forma como o Plotly e o Bootstrap foram utilizados para criar gráficos interativos e uma interface de utilizador responsiva.

Com o Plotly, foi desenvolvido um conjunto de gráficos para apresentar uma descrição geral das variáveis ótimas das culturas. Os gráficos desenvolvidos, organizados por tema, foram os seguintes:

Temperaturas Ótimas - Um gráfico que mostra as temperaturas ideais para o cultivo de diferentes culturas e também a opção de visualização dos intervalos absolutos das temperaturas de cada cultura (Figura 31).

Meses de Plantação - Um *heatmap* que visualiza os melhores meses para plantar e colher diferentes culturas (Figura 31).

Dados de Precipitação - *Box plots* que representam os valores de precipitação, ajudando os agricultores a entender as necessidades de água das culturas (Figura 32).

Dados do Solo - Gráficos de *bullet* e *gauge* que fornecem dados sobre as condições ideais do solo para diferentes culturas (Figura 32).

Resultados da Recomendação – Um *heatmap* com o resultado do MAE para todos os dias diferentes de plantação. Também pode ser visualizado em formato de gráfico de linhas (Figura 33).

Temperaturas do Ar – Gráficos de linhas para representar a temperatura média diária, temperatura máxima, temperatura mínima e a previsão da temperatura (Figura 34).

Abaixo estão apresentados os gráficos na página para as culturas.

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

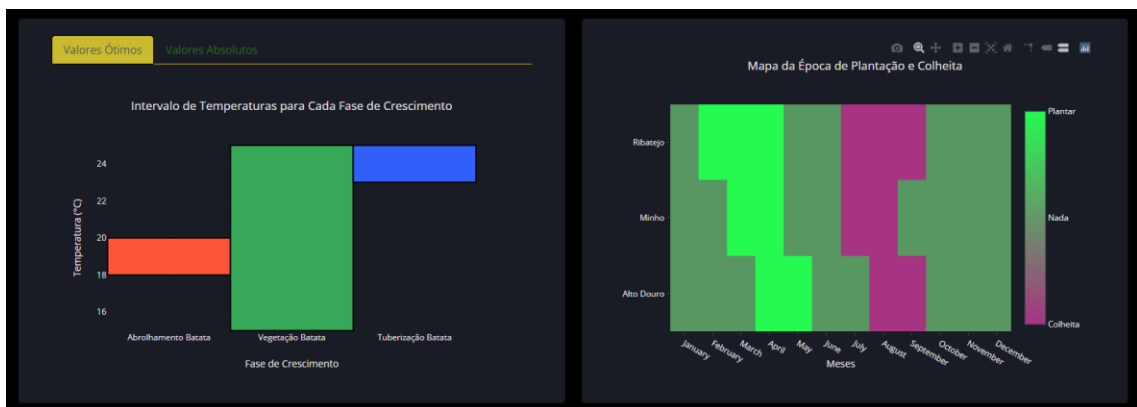


Figura 31 - Gráfico de Temperaturas Ótimas (Esquerda) e Gráfico de Época de Plantação (direita)



Figura 32 - Gráfico de Precipitação (esquerda) e Gráfico do pH do Solo (direita)

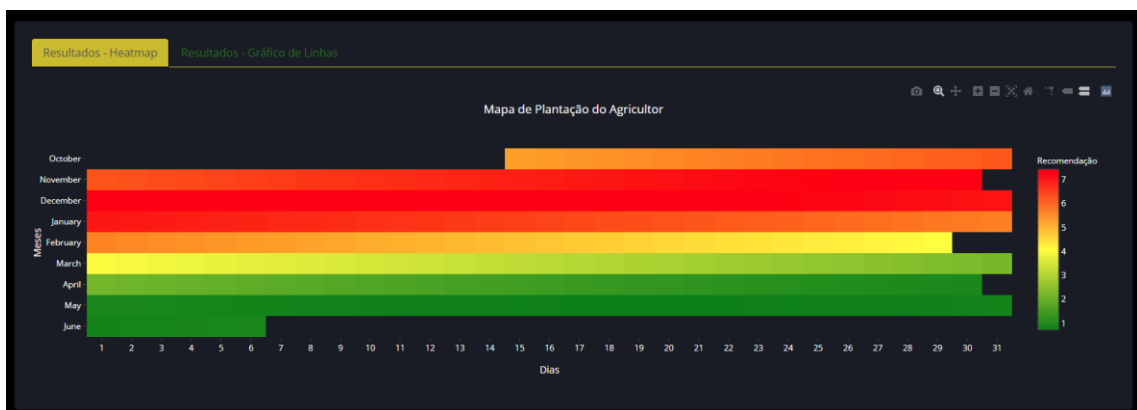


Figura 33 - Gráfico dos Resultados do Algoritmo

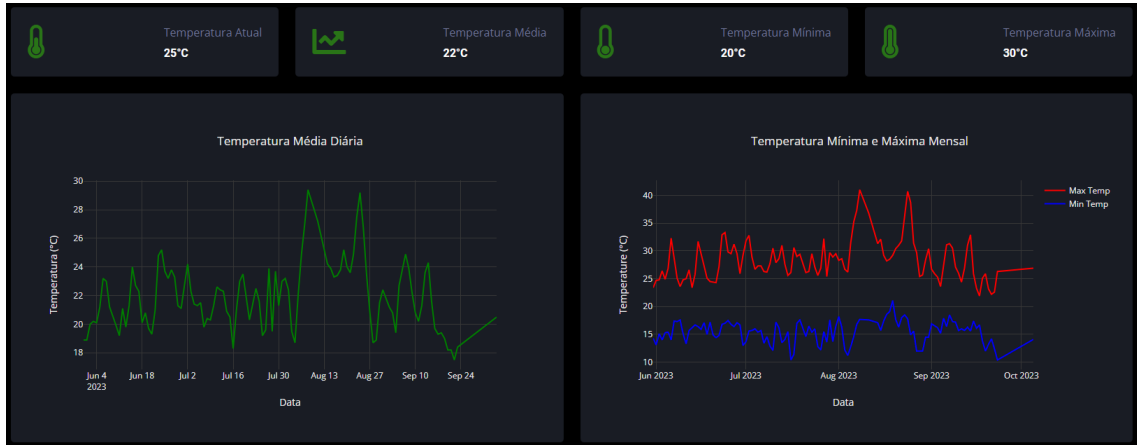


Figura 34 - Gráfico da Temperatura Média Diária e Gráfico da Temperatura Máxima e Mínima (direita)

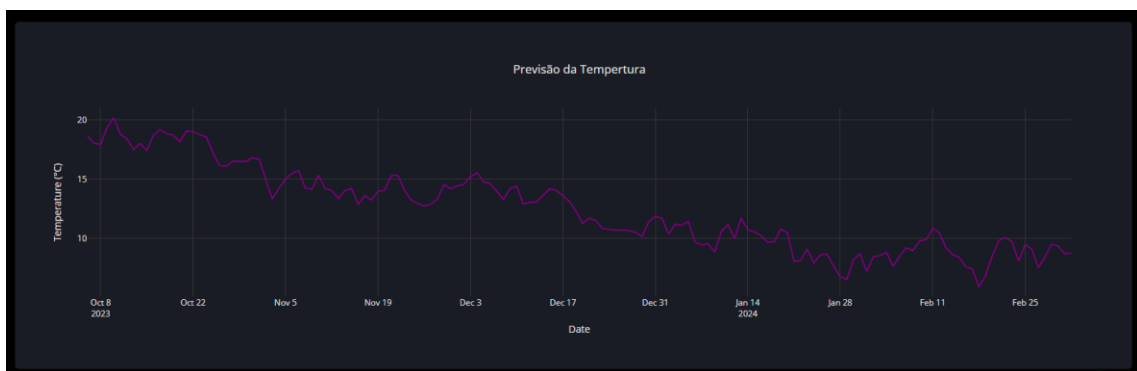


Figura 35 - Gráfico da Previsão da Temperatura

Os dados para alimentar estes gráficos são importados no arquivo `view.py`, que é responsável por processar os dados e prepará-los para visualização. Os *layouts* dos gráficos são escritos em JavaScript, permitindo uma experiência de utilizador dinâmica e envolvente.

5.7 Restantes informações acerca das culturas

Existem outras informações relevantes relacionadas com as culturas, disponíveis na aplicação web. Estas informações complementam a experiência do utilizador ao fornecer dados detalhados sobre as diferentes culturas, incluindo informações descritivas, imagens, famílias botânicas, nomes científicos tipo de clima mais favorável e a duração do seu ciclo de crescimento. Na figura 36 podemos visualizar informações na página web sobre a batata.



Figura 36 - Informações gerais na página da cultura

Além das informações individuais de cada cultura, o sistema apresenta uma página dedicada que lista todas as culturas disponíveis. Esta página oferece uma visão geral completa de todas as culturas que os utilizadores podem explorar. Cada cultura é apresentada com o seu nome vulgar, o seu nome científico e um link que redireciona para a página da respetiva cultura. Os utilizadores podem navegar facilmente pela lista completa de culturas (Figura 37) para encontrar aquelas que mais se enquadram aos seus interesses.

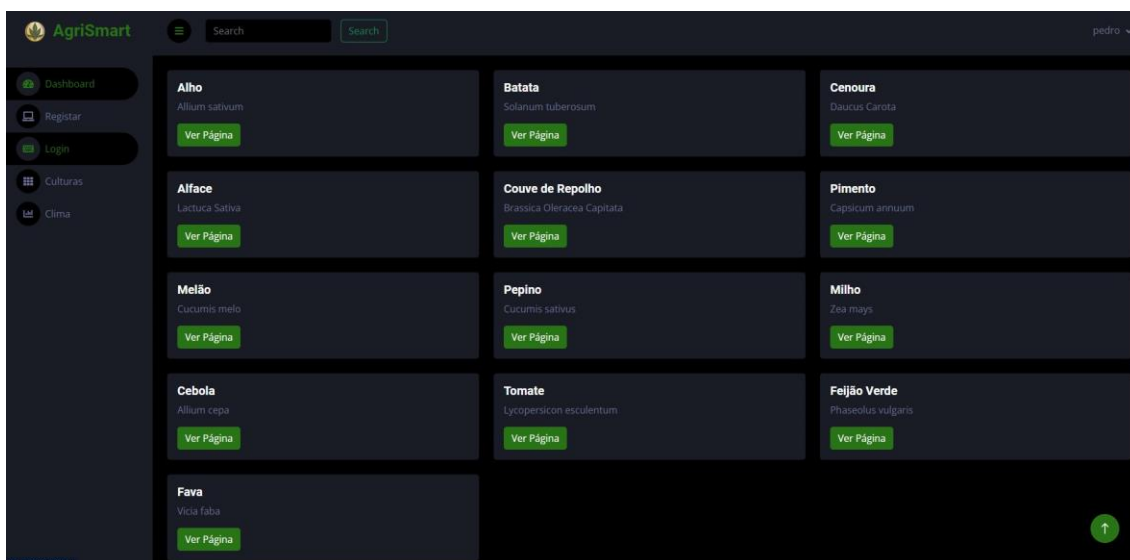


Figura 37 - Página da lista de culturas

5.8 Integração do Modelo com a Aplicação Web

5.8.1 Armazenamento do modelo

Antes de integrar o modelo na aplicação web, foi necessário exportá-lo, após o treino, no formato .h5. Este formato é amplamente reconhecido e suportado por bibliotecas de DL como o TensorFlow. Com o Django Rest Framework já instalado no projeto Django, dispúnhamos da infraestrutura necessária para iniciar a implementação do modelo. Para armazenar o modelo exportado, criou-se uma pasta específica dentro da estrutura de diretórios do projeto Django.

5.8.2 Preparação dos dados para o modelo.

Triggers na Base de Dados – Foram programados *triggers* na base de dados para automatizar a criação de uma tabela destinada às variáveis de entrada do modelo. Estes *triggers* são ativados sempre que o *script* recolhe dados do IPMA e os armazena. Assim, a temperatura média diária é automaticamente copiada para a tabela de entrada do modelo. Adicionalmente, o dia em formato cosseno e a variável que classifica a intensidade do El Niño são também inseridos através de *triggers*.

Processamento no Backend do Django - No *backend*, garantiu-se que a estrutura dos dados enviados ao modelo corresponde exatamente igual à estrutura das variáveis de entrada, que foram usadas durante o treino do modelo. Com a linguagem Python, automatizei um script, que converte os dados dos últimos 60 dias e organiza as variáveis na ordem correta. No final converte-se para num *array* bidimensional. Sempre que existe um pedido de previsão no servidor, este script é ativado na *view.py* da respetiva aplicação.

Serialização dos Dados - A serialização, no contexto do Django Rest Framework, envolve a conversão de objetos complexos num formato facilmente transmissível, como o JSON. Após a serialização, os dados são enviados para a API responsável pelas previsões. Esta API, por sua vez, utiliza o modelo previamente treinado para gerar previsões. Os resultados são então deserializados, convertendo-os de volta para um formato utilizável pelo *frontend*.

Finalmente, após a obtenção das previsões, os resultados são enviados de volta ao Django, que os encaminha para o *frontend*, permitindo uma visualização clara e interativa das previsões geradas pelo modelo.

5.8.3 Implementação do Algoritmo de Recomendação da Data de Plantação

A determinação da data ótima de plantação é uma tarefa complexa que requer uma compreensão profunda do problema em questão. Inicialmente, enfrentou-se o dilema de decidir entre criar um modelo que previsse diretamente a data ou um que previsse a temperatura, recorrendo posteriormente a um julgamento para determinar a data. A escolha recaiu sobre a segunda opção, considerando-a mais flexível e adaptável.

A funcionalidade principal da plataforma é determinar a melhor data de plantação para diferentes culturas. Para isso, criou-se um *script* que calcula o MAE entre as temperaturas previstas e os intervalos de temperatura ótimos para cada fase de crescimento das culturas. Este script tem em consideração a duração da cultura em dias paralelamente com a duração da fase vegetativa. Com base no MAE médio menor, o sistema recomenda a data mais adequada para plantação.

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

O primeiro passo envolveu a obtenção de dados de temperatura. Embora os sensores fossem uma opção inicial, a precisão dos dados do IPMA provou ser superior. Esta escolha garante que o sistema esteja sempre atualizado com os dados mais recentes (e precisos) disponíveis.

Uma vez que os dados de temperatura são recolhidos, o sistema recorre ao modelo para prever as temperaturas dos próximos dias ou semanas. Estas previsões são fundamentais para antecipar as condições climáticas e, conseqüentemente, para determinar o momento ideal de plantação.

O sistema consulta a base de dados para identificar os intervalos de temperatura ideais para cada tipo de cultura. Estes intervalos refletem as condições que favorecem o crescimento saudável das plantas. Com os intervalos de temperatura ideais e as previsões selecionados, o sistema calcula o MAE. Este valor quantifica o quão próximas as previsões estão dos intervalos ideais, servindo como indicador da qualidade da recomendação.

Para facilitar a interpretação, os dados e previsões são convertidos para um formato de data, permitindo uma visualização organizada e intuitiva ao longo do tempo.

Com base em toda a análise realizada, o sistema formula recomendações sobre as datas ideais de plantação. Estas recomendações são apresentadas ao utilizador através da interface da aplicação, proporcionando uma ferramenta valiosa para a tomada de decisões informadas.

Na figura 38 abaixo está ilustrado o diagrama de atividades da recomendação da data de plantação.

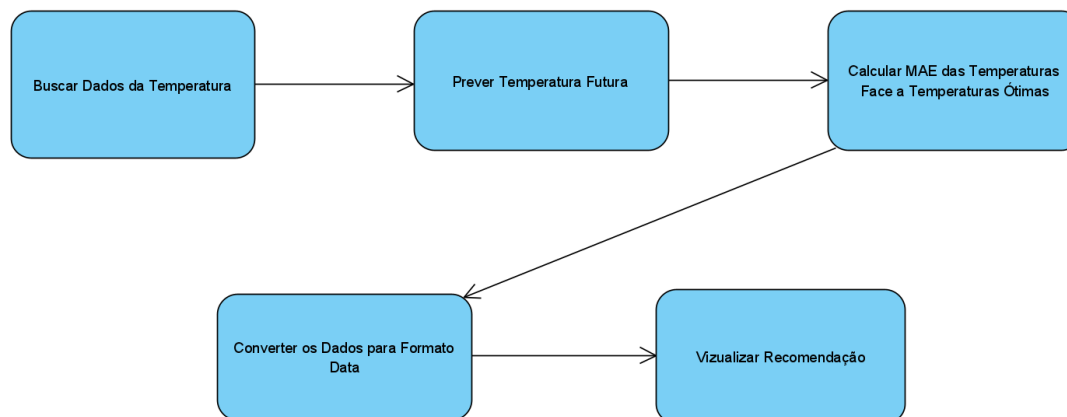


Figura 38 - Diagrama de Atividades do Algoritmo de Recomendação

O algoritmo que implementa a recomendação funciona de acordo com o processo seguinte.

Nome da função: calculate_temperature_mae.

Objetivo: Calcular o MAE para todo o ciclo de crescimento de cada cultura e determinar as melhores datas de plantação.

Entradas: Lista de culturas e lista de temperaturas previstas.

Processo

1. Para cada cultura, são obtidas as fases de crescimento e a duração total do ciclo.
2. As temperaturas previstas são analisadas em janelas correspondentes à duração do ciclo de crescimento.
3. O MAE é calculado para cada janela, comparando as temperaturas previstas com os intervalos ótimos.
4. Os resultados são armazenados num dicionário, incluindo a melhor data e as 30 melhores datas para plantação.

Saída: Dicionário com os valores de MAE e as datas recomendadas para plantação.

5.9 Integração com dispositivos IoT

A rede LoRaWAN foi selecionada pela sua capacidade de transmitir dados a longas distâncias com baixo consumo de energia. O The Things Network (TTN) foi utilizado como a plataforma principal para gerir a comunicação entre o sensor Dragino e a aplicação web. O TTN não só facilita a integração de dispositivos LoRaWAN, mas também oferece funcionalidades robustas para processar e encaminhar os dados recebidos. Como o sensor Dragino S-31 já estava na lista do repositório de sensores do TTN, o seu registo na plataforma foi bastante fácil. No formulário de registo do sensor são pedidas 3 chaves únicas que vêm na caixa do sensor e a configuração fica finalizada, após a inserção das mesmas.

O sensor começa logo a registar dados no TTN, de forma automática, quando estiver ao alcance de uma *gateway* pública. Os dados que são transmitidos pelo sensor são a data, a temperatura do ar e a humidade do ar.

Para garantir que os dados capturados pelo sensor Dragino fossem transmitidos em tempo real para a aplicação web, foi implementada uma solução baseada em *webhooks*. O TTN permite a configuração de *webhooks*, que são acionados sempre que novos dados são recebidos. Assim, sempre que o sensor Dragino envia dados, o TTN processa esses dados e, através do *webhook*, envia-os diretamente para a aplicação *web*.

Para processar os dados recebidos pelo *webhook*, foi implementada a função *webhook_endpoint* no backend do Django. Esta função é responsável por:

Nome da função: `webhook_endpoint`.

Objetivo: Processar e armazenar dados do sensor vindos via POST request, como identificação do dispositivo, temperatura e humidade.

Entradas: Uma requisição HTTP, esperando que essa requisição tenha um método POST e que o corpo da requisição tenha um formato JSON específico.

Processo

1. Verifica se o método da requisição é 'POST'.
2. Caso seja 'POST', lê o corpo da requisição e decodifica o JSON.
3. Extrai o ID do dispositivo.
4. Extrai o valor da temperatura e da humidade.
5. Armazena os dados extraídos na base de dados.
- 6.

Saída 1: Se tudo correr certo, envia uma mensagem de sucesso em formato JSON

Saída 2: Se o método da requisição não for 'POST', responde com uma mensagem de erro em formato JSON.

Dada a necessidade de ter um servidor para receber os dados transmitidos pelo *webhook*, optou-se pelo uso do Ngrok. O Ngrok é uma ferramenta que cria túneis seguros para o *localhost*, permitindo que aplicações locais sejam expostas ao mundo exterior. Assim, mesmo sem um servidor dedicado, foi possível receber e processar os dados em tempo real na aplicação *web*.

6 Resultados e Discussão

6.1 Avaliação dos Modelos

6.1.1 Resultados dos Modelos com diferentes arquiteturas (modelos univariáveis).

O primeiro passo é aplicar os modelos para prever a temperatura, somente com base no histórico da própria temperatura. Decidiu-se prever a temperatura dos próximos 30 dias utilizando a temperatura dos últimos 7.

Tabela 37 - Resultados dos Modelos Univariáveis

Nome	Window	Horizon	Resultado (MAE)
Baseline	7	30	5.9712
Linear	7	30	3.7408
Dense	7	30	3.6496
LSTM	7	30	<u>2.6178</u>
N-BEATS	7	30	<u>1.3415</u>

A seleção da arquitetura adequada para um modelo de ML é uma etapa crítica que pode determinar o sucesso ou fracasso do projeto. Diversas arquiteturas foram testadas para identificar a que melhor se adapta aos dados e ao problema em questão.

O modelo de base, com um MAE de 5,9712, serviu como um ponto de partida para a avaliação.

Os modelos linear e denso, obtiveram resultados idênticos, ambos com um MAE de 3,6496 e 3,7408 respetivamente. A superação destes modelos face ao modelo base é um indicativo positivo de que as arquiteturas selecionadas foram capazes de aprender e adaptar-se aos padrões presentes nos dados.

No entanto, foi o modelo LSTM que se destacou significativamente, alcançando um MAE de 2,6178. Esta arquitetura, projetada especificamente para lidar com sequências e dados de séries temporais, demonstrou a sua eficácia ao capturar dependências temporais nos dados.

O modelo N-BEATS, com um bom MAE de 1,3, superou todas as outras arquiteturas testadas. Embora este modelo tenha mostrado um desempenho superior, é importante notar que ele tem limitações quanto ao número de variáveis que pode processar durante o treino. A sua implementação foi interrompida devido a problemas técnicos que nós encontramos. Por estes motivos foi escolhido a arquitetura LSTM.

Em conclusão, o modelo LSTM foi identificado como a melhor arquitetura entre as quais foram desenvolvidas para este conjunto de dados e problema. A sua capacidade de lidar eficazmente com séries temporais foi evidente nos resultados obtidos. Com a arquitetura selecionada, a próxima etapa envolve a incorporação de variáveis adicionais para tentar melhorar o desempenho do modelo. Será interessante observar como o modelo LSTM se compara ao N-BEATS quando essas variáveis são introduzidas, dadas as restrições do último.

6.1.2 Resultados dos Modelos com Diferentes Janelas de Temporais

A arquitetura LSTM foi selecionada como a mais adequada para o conjunto de dados em questão. Para otimizar ainda mais o desempenho do modelo, foram testadas diferentes configurações de entrada e de saída. A entrada refere-se ao número de dias utilizados para alimentar o modelo e fazer uma previsão, enquanto a saída é o número de dias que o modelo tenta prever. A saída foi mantida constante em 365 dias, representando um ano, enquanto a entrada variou. No capítulo 6.4.1 explica-se o porquê de estar a prever 365 dias.

Os resultados obtidos para o modelo LSTM foram os seguintes.

Tabela 38 - Resultado dos Modelos de Janelas Temporais Diferentes

Nome	Window	Horizon	Resultado (MAE)
LSTM	1	365	3.6647
	7	365	3.4896
	15	365	3.3492
	30	365	3.0485
	60	365	<u>2.6588</u>

A partir dos resultados, observamos que o modelo com uma entrada de 1 dia e uma saída de 365 dias obteve o maior MAE de 3.6647.

Ao aumentar o número de dias de entrada de 1 para 60, observamos uma tendência geral de redução no erro. O modelo com uma entrada de 60 dias e uma saída de 365 dias obteve um MAE de 2.65588, o que indica que fornecer mais dados históricos ao modelo pode ajudar a melhorar a precisão das previsões para um período de um ano.

Em conclusão, a seleção adequada da configuração de entrada e de saída é crucial para otimizar o desempenho do modelo LSTM. A análise sugere que fornecer mais dados históricos ao modelo pode ser benéfico para previsões de longo prazo.

No entanto, é importante notar que ao tentar prever com mais de 60 dias começava a ter-se um resultado muito fora do normal. Isto pode ser devido a *overfitting*, onde o modelo se torna demasiado ajustado aos dados de treino e perde a capacidade de generalizar bem para dados não vistos. Alternativamente, pode ser que a adição de mais dias de entrada introduza ruído ou informações irrelevantes que perturbem o modelo.

6.1.3 Resultados dos Modelos com Variáveis

O modelo base agora considerando é o que utiliza uma entrada de 60 dias sem qualquer variável adicional, serve como ponto de referência e apresenta um MAE de 2,65886.

Tabela 39 - Resultado dos Modelos com Variáveis Diferentes

Nome	El_nino	Função Cosseno	Resultado (MAE)
LSTM_60_365	Todas as Intensidades	Não	2.6785
	Só Intensidades Médias e Altas	Não	2.7042
	Só Intensidades Altas	Não	2.6212
	Não	Sim	2.3833
	Só Intesidades Altas	Sim	<u>2.3808</u>

Ao incorporar a variável El Niño, observamos um ligeiro aumento no erro, alcançando 2,6785. Isto sugere que, embora a variável El Niño possa ter relevância, a sua inclusão completa não melhora o desempenho do modelo neste contexto.

Quando removemos os eventos El Niño de fraca intensidade no modelo, o erro aumenta marginalmente para 2,7042. No entanto, ao remover adicionalmente os eventos de intensidade média no modelo, o erro diminui para 2,62127. Isto indica que os eventos El Niño de forte intensidade têm um impacto mais significativo nas previsões, enquanto os de intensidade fraca e média podem introduzir algum ruído.

O modelo com a variável do cosseno alcançou um erro de 2,38333. Esta diferença foi a maior registada, o que significa que as previsões melhoram substancialmente com esta nova variável de entrada. Surpreendentemente, ao adicionar a variável El Niño a esta combinação no modelo apresentou um MAE praticamente inalterado de 2,3808. Mesmo assim é uma melhoria.

A inclusão da variável El Niño, quando considerando eventos do El Niño de forte intensidade, é benéfico para o modelo, mas os resultados são ainda melhores quando a variável cosseno é introduzida.

6.1.4 Resultados dos Modelos com Layers e Hiperparâmetros Diferentes

O modelo base agora é o que utiliza uma entrada de 60 dias com as duas variáveis adicionais, serve como nosso ponto de referência e apresenta um MAE de 2,3808.

Tabela 40 - Resultado dos Modelos com Layers Diferentes e Hiperparâmetros Diferentes

Nome	Mudança nas Layers ou Hiperparametros	Resultado (MAE)
LSTM_60_365_elnino_cos	Segunda Layer com 64 Neurons	4.4425
	Adicionar Dropout 0.3	2.3809
	Aumentar Neurons para 256	2.3762
	Aumentar Neurons para 512	2.3810
	Adicionar Batch Normalization Layer	2.3786
	256 Neurons + Drop Layer + Batch Normalization	2.3711
	Neurons 512 + Drop Layer 0.2	<u>2.3604</u>

Começando por adicionar uma segunda *layer* LSTM com 64 neurónios resultou num MAE mais alto em comparação com a configuração original, indicando que a adição da segunda *layer* não melhorou o desempenho do modelo. Sempre que se adicionou uma ou mais *layers* os resultados pioraram.

A adição de uma *layer* de *dropout* com uma taxa de 0,3 resultou em praticamente o mesmo resultado.

Aumentar o número de neurónios na *layer* LSTM para 256 também levou a uma melhoria muito reduzida no MAE em relação à configuração original. Aliás qualquer melhoria apresentada nos resultados é bastante pequena.

Combinando ambas as camadas, *dropout* e normalização em lote, no modelo com 512 neurónios, observamos a maior melhoria, com um erro de 2,3604. Esta combinação parece oferecer uma regularização robusta, otimizando o desempenho do modelo.

Mesmo assim, por mais pequena que seja a melhoria do MAE, ela é sempre bem-vinda, pois o modelo está a prever melhor nestas configurações. O que apresentou o melhor resultado foi o modelo com 512 neurónios e uma *dropout layer* com uma taxa de 0,2. Este é o meu modelo final.

6.2 Testes da Aplicação

A fase de testes é fundamental em qualquer projeto de desenvolvimento, pois garante que o sistema funciona conforme o esperado e identifica possíveis falhas ou oportunidade de introduzir melhorias. No contexto deste projeto, foram utilizadas ferramentas específicas para testar diferentes componentes da solução.

6.2.1 Utilização do Ngrok para Testar a Aplicação Web

No desenvolvimento da aplicação *web*, o Ngrok desempenhou um papel crucial ao proporcionar um meio de testar a aplicação num ambiente que simula a operação real em produção.

Ao executar o Ngrok, foi gerado um URL público que redirecionava para a aplicação local. Isso permitiu testar a aplicação em diferentes dispositivos e redes, simulando o acesso de utilizadores reais. Além disso, o Ngrok oferece um painel de controlo onde é possível visualizar o tráfego HTTP, facilitando a identificação e a correção de possíveis problemas.

6.2.2 Uso do Postman para Testar os Algoritmos

O Postman é uma plataforma popular para testar, desenvolver e documentar APIs. Foi utilizado para testar os algoritmos desenvolvidos, especialmente para garantir que estavam a responder corretamente às solicitações GET e POST.

Teste GET - Através do Postman, foi possível enviar solicitações GET para os *endpoints* definidos, verificando se as respostas retornadas estavam de acordo com o esperado. Este teste ajudou a confirmar que os algoritmos estavam a tratar os dados corretamente Figura.

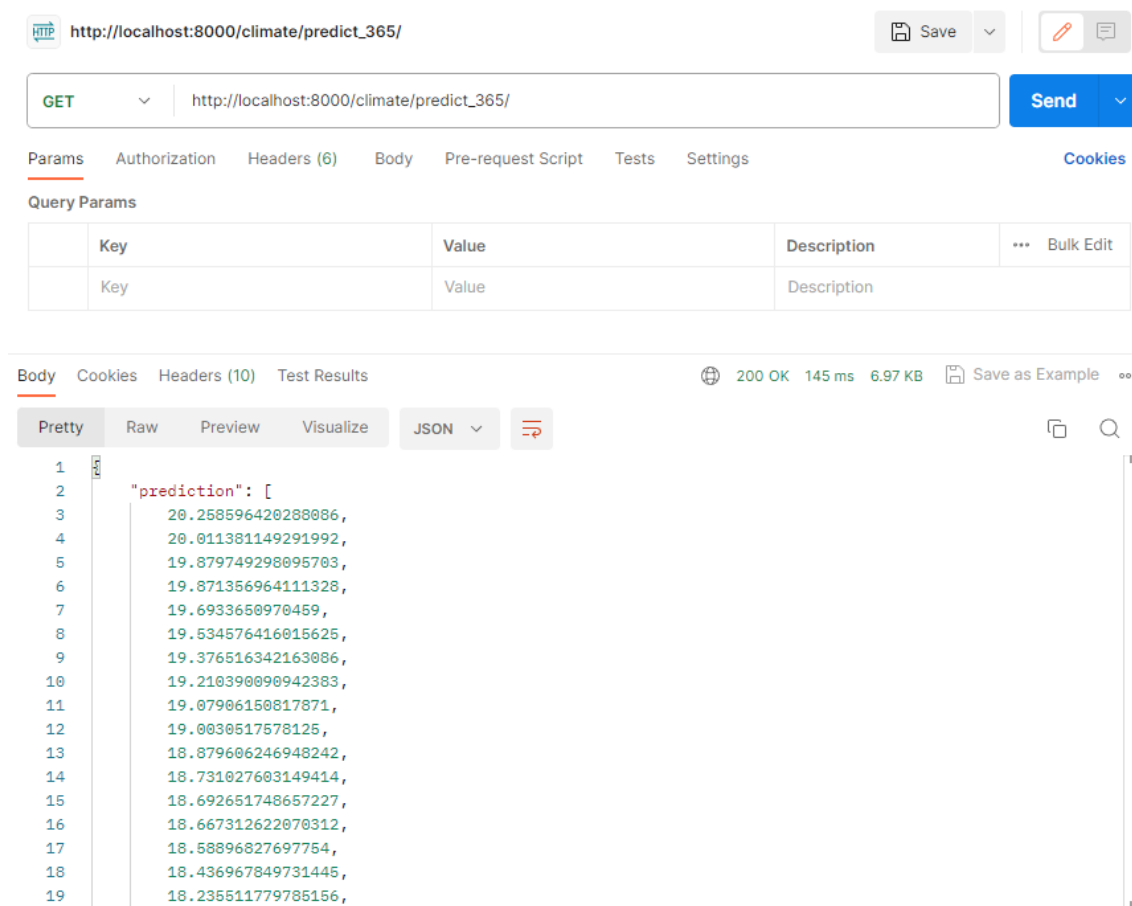


Figura 39 - Resultados do Teste GET

Teste POST - O teste POST foi essencial para garantir que os algoritmos pudessem receber dados, processá-los e retornar uma resposta adequada. No Postman, foi possível definir os dados de entrada, enviar os dados e analisar a resposta do algoritmo.

O Postman também oferece recursos como a criação de ambientes, o que facilitou a simulação de diferentes cenários de teste. Além disso, a capacidade de salvar solicitações e organizar em coleções permitiu uma abordagem sistemática e repetível para os testes.

Em conclusão, a combinação do Ngrok e do Postman proporcionou uma abordagem robusta e eficaz para testar a aplicação web e os algoritmos, garantindo que o sistema estivesse pronto para ser lançado num ambiente de produção.

6.3 Resultado das Recomendações de Datas de Plantação

Durante a fase de teste, uma das abordagens adotadas foi comparar os *heatmaps* gerados pelo modelo desenvolvido com os calendários do livro que serviu de base para este projeto. Estes calendários agrícolas, amplamente utilizados por agricultores, fornecem uma visão geral das melhores épocas para plantação, crescimento e colheita de diversas culturas.

O calendário agrícola do livro recomenda a plantação agrupados por quinzenas e separadas por região agrícola. Replicou-se o *heatmap* para o melão na Tabela 41. As colunas representam os meses do ano, de janeiro (J) a dezembro (D), enquanto as linhas representam diferentes regiões agrícolas, designadas aqui como Entre Douro e Minho (EDM), Ribatejo e Oeste (RO) e Algarve (A). Para cada região, há duas operações listadas, plantar e colher.

Tabela 41 - Heatmap de Plantação para o Melão

Região	Operação	J	F	M	A	M	J	J	A	S	O	N	D
EDM	Plantar												
	Colher												
RO	Plantar												
	Colher												
A	Plantar												
	Colher												

O livro sugere intervalos quinzenais para plantação, enquanto o sistema recomenda datas específicas. Para facilitar a comparação, adotou-se a estratégia de considerar os melhores dias recomendados pelo sistema até igualar o número total de dias sugeridos pelos calendários tradicionais. Por exemplo, se um calendário tradicional sugere plantar uma cultura durante o mês de março e de abril, considera-se isto como um total de 60 dias. Assim, compara-se os 60 dias (entre 1 de março a 30 de abril) com os 60 melhores dias recomendados pelo nosso sistema (60 datas com o menor valor médio de MAE). Se existir datas em comum então é sinal que existe uma concordância entre as recomendações do sistema e os calendários genéricos. Os resultados estão apresentados na Tabela 42.

Tabela 42 - Resultados da comparação da recomendação do livro com a do sistema

Cultura	Recomendação dos dias (livro)	Dias em comum (meu sistema)	Regiões utilizadas	Percentagem de eficácia
Tomate	150	135	RO e EDM	90%
Pimento	150	113	RO e EDM	75.33%
Melão	75	0	RO e EDM	0%
Pepino	150	20	RO e EDM	13.33%
Batata	60	22	RO e EDM	36.67%
Feijão verde	30	15	BL	50%

O sistema apresenta uma semelhança com o calendário do livro para a cultura do tomate e do pimento, com uma eficácia de 90% e de 75.33%. Isto indica que, para estas culturas, as recomendações do sistema são consistentes com as do calendário do livro, sugerindo que o sistema é capaz de prever com precisão os melhores momentos para plantação dessa cultura.

Para a cultura do melão, não foi registada nenhuma data em comum entre as recomendações do sistema e as do calendário do livro, resultando numa eficácia de 0%. Assim como o pepino, o sistema teve uma baixa eficácia de 13.33% para a cultura do pepino. Isso indica uma discordância significativa entre as recomendações do sistema e as do calendário do livro para essa cultura. Mas ao mesmo tempo não quer dizer que a recomendação esteja errada.

A eficácia para a cultura da batata também é relativamente baixa, com 36.67%.

Para a cultura do feijão verde, o sistema apresenta uma eficácia de 50%, o que indica um resultado razoável com o calendário do livro. Embora não seja tão alta quanto a eficácia do tomate, ainda representa uma melhoria em relação a outras culturas.

Só o feijão verde é que tinha uma sugestão de épocas de plantação para a região da Beira Litoral (BL), para as outras culturas, tive de usar as regiões agrárias vizinhas.

Não foi possível avaliar mais culturas, pois no livro, nem todas as culturas têm um *heatmap* de auxílio à época de plantação.

Ao comparar os *heatmaps* do modelo com os calendários, observou-se resultados mistos. Estas discrepâncias podem ser atribuídas à capacidade do modelo considerar variáveis e condições específicas que os calendários genéricos não abordam, pois estes são baseados na experiência e intuição. Isto pode explicar as discrepâncias observadas em algumas culturas, como o melão e o pepino.

Sendo os calendários agrícolas baseados em tendências observadas ao longo de muitos anos, eles não têm a capacidade de adaptar-se rapidamente a mudanças ou anomalias climáticas, ao passo que o modelo desenvolvido, sendo baseado em ML, pode ajustar-se e adaptar-se a novos dados conforme são recebidos.

É importante medir a aptidão das culturas que tiveram uma eficácia menor, pois estas culturas, devido às alterações climáticas podem estar no limite da sua tolerância. Como referido anteriormente no artigo no capítulo 2 [21], se as temperaturas registadas já se encontram próximos dos máximos fisiológicos, as alterações climáticas terão um impacto mais forte no seu rendimento.

Por fim temos os resultados do melhor dia para plantar para cada cultura no sistema. Esta é uma vantagem em relação ao calendário do livro, que apenas recomenda a quinzena ideal.

Tabela 43 - Resultado do Melhor Dia de Plantação

Cultura	Dia
Milho	21 de julho
Batata	26 de maio
Alface	5 de janeiro
Cenoura	12 de agosto
Tomate	22 de junho
Feijão Verde	20 de agosto
Alho	2 de junho
Couve	31 de outubro
Pimento	10 de julho
Melão	6 de julho
Pepino	31 de agosto

Em conclusão, os resultados variam de acordo com a cultura, com eficácias altas (tomate e pimento) e baixas (melão e pepino). Isto sugere que o sistema pode ser eficaz para algumas culturas, mas pode exigir ajustes ou melhorias para outras. É importante considerar a possibilidade de ajustar os parâmetros do modelo ou recolher mais dados específicos para melhorar a eficácia geral do sistema em prever os melhores momentos para a plantação de diferentes culturas.

Enquanto os *heatmaps* gerados pelo modelo oferecem uma visão mais dinâmica e adaptável das condições agrícolas, os calendários agrícolas genéricos fornecem uma base sólida e comprovada para a tomada de decisões. A combinação de ambos pode oferecer aos agricultores uma ferramenta para otimizar suas operações agrícolas.

6.4 Limitações

Ao longo do desenvolvimento deste projeto, foram identificadas várias limitações que são importantes destacar para uma compreensão completa do sistema e das suas capacidades.

6.4.1 Modelo de Previsão Único

A primeira limitação encontrada foi a incapacidade de criar um modelo de previsão específico para cada cultura. Idealmente, cada cultura, com o seu ciclo de crescimento distinto, deveria ter um modelo de previsão personalizado. Por exemplo, para uma batata com um ciclo de 150 dias, seria mais adequado ter uma previsão para apenas 150 dias. No entanto, devido a desafios técnicos e de implementação, optou-se por desenvolver um modelo que previsse os 365 dias, independentemente da cultura em questão.

6.4.2 Tamanho da window

Outro problema que surgiu foi aumentar o tamanho da *window*. Nos resultados da tabela 38, percebe-se que quanto maior é o tamanho da *window*, mais eficaz são as previsões. Mas, no entanto, quando se aumentava o número da *window*, os resultados dos modelos, durante a fase de treino, começavam a desviar-se por um motivo alheio.

6.4.3 Utilização de um único sensor como representativo uma Quinta Agrícola

A ideia de que um único sensor pudesse representar uma quinta inteira e estivesse associado a uma conta de utilizador apresenta desafios. Além disso, ao apenas utilizar um sensor, não consigo validar se a implementação está bem configurada para outros utilizadores, pois apenas possuo um utilizador. Por isso a utilidade do sensor no sistema é indicar a temperatura atual, a temperatura máxima, a temperatura média e temperatura mínima do dia atual. Também foi implementado a visualização dos dados da humidade recolhidos pelo sensor.

Plataforma de Apoio à Decisão de Plantação de Culturas Hortícolas em Coimbra

Em conclusão, é fundamental reconhecer e tratar estas limitações em futuras versões do sistema de modo a garantir um sistema mais flexível, que utilize uma representação mais completa da realidade de uma quinta e que se adapte melhor às diferentes culturas.

7 Conclusão e Trabalho Futuro

7.1 Conclusão

Podemos concluir que os objetivos principais do sistema de recomendação foram alcançados e satisfatórios.

O projeto demonstra potencial para auxiliar os agricultores na tomada de decisões relacionadas com a identificação da melhor época de cultivo. Isto é especialmente relevante num contexto de mudanças climáticas, onde a adaptação às condições climáticas é crucial para a agricultura sustentável.

A aplicação de técnicas de DL, com a arquitetura LSTM, provou ser eficaz na previsão de datas de plantação com base em dados históricos de temperatura.

A comparação entre os resultados do modelo e os calendários agrícolas do livro mostrou que o modelo oferece uma visão mais dinâmica e adaptável das condições agrícolas. Isso é importante, pois o modelo pode ajustar-se rapidamente a mudanças ou anomalias climáticas. O facto de a recomendação ser uma data específica também acrescenta mais valor à recomendação em vez de recomendar um certo período de dias.

7.2 Trabalho Futuro

Um projeto como o aqui apresentado tem um conjunto de desafios em aberto, entre os que consideramos mais promissores incluem-se dotar o sistema das seguintes capacidades:

Modelos específicos para culturas - Com base nas limitações identificadas, um dos principais focos para o futuro será desenvolver modelos de previsão específicos para cada cultura. Isso permitirá previsões mais precisas, alinhadas com o ciclo de vida de cada planta.

Suporte a múltiplos sensores - Em vez de depender de um único sensor para representar uma quinta inteira, serão suportados mais sensores. Isto permitirá que o sistema suporte quintas agrícolas de pequena e grande dimensão.

Recomendações baseadas nas características do solo - Uma extensão natural deste trabalho seria desenvolver um algoritmo que, com base nas características do solo, recomende as culturas mais adequadas para plantação. Isso ajudaria os agricultores a tomar decisões informadas e maximizar a produtividade.

Feedback contínuo - À medida que novos dados são recolhidos, o sistema será projetado para aprender continuamente e ajustar-se, garantindo que as previsões e recomendações permaneçam relevantes e cada vez mais precisas.

Enquanto este trabalho estabelece uma base sólida, o horizonte é vasto e promissor. Com a contínua evolução da tecnologia e a crescente necessidade de soluções agrícolas sustentáveis, estamos confiantes de que as inovações futuras neste domínio trarão benefícios significativos para agricultores e consumidores em todo o mundo.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] “Seca Será ‘o Novo Normal’ Em Portugal Devido Às Alterações Climáticas e a Erros Na Agricultura, Alerta a Liga Para a Proteção Da Natureza,” *Agência Luso*, 2022.
- [2] “Ano Mais Seco Desde 1931 Está a Prejudicar Produção de Cereais Em Portugal,” *Euronews*, 2022.
- [3] dnesh, “<https://www.dnesh.bg/stranata/2022/07/13/jega-li-rekordno-studen-iuli-unas.536759>,” DNESH, 2022. [Online].
- [4] “<https://apambiente.pt/prevencao-e-gestao-de-riscos/secas>,” Agência Portuguesa do Ambiente, 2022. [Online].
- [5] Expresso, “<https://expresso.pt/sociedade/2023-07-07-Recorde-do-dia-mais-quente-de-sempre-na-Terra-foi-quebrado-tres-vezes-esta-semana-continuara-a-temperatura-a-subir--fb8e5bab>,” Expresso, 2023. [Online].
- [6] Lusa, “<https://www.tsf.pt/mundo/este-foi-o-verao-mais-quente-de-sempre-alguma-vez-registado-no-hemisferio-norte-16976778.html>,” TSF, 2023. [Online].
- [7] V. Rural, “<https://www.vidarural.pt/destaques/recursos-hidricos-15-albufeiras/>,” Vida Rural, 2022. [Online].
- [8] p. d. c. R. O. I. F. D. e. C. C. The Consumer Intelligence Lab, “O Uso da Água Em Portugal,” 2020.
- [9] “<https://www.confagri.pt/seca-regadio-usa-75-da-agua-portugal-desperdica-um-terco-revela-apa>,” CONFAGRI, 2022. [Online].

- [10] ProCiv, “<http://www.prociv.pt/pt-pt/RISCOSPREV/RISCOSNAT/SECAS/Paginas/default.aspx>,” 2022. [Online].
- [11] F. Bada, “<https://www.worldatlas.com/articles/what-are-the-major-natural-resources-of-portugal.html>,” WorldAtlas, 2018. [Online].
- [12] N. Guedes, “<https://cnnportugal.iol.pt/ipma/seca-extrema/percentagem-de-territorio-portugues-em-seca-severa-ou-extrema-duplicou-em-quinze-dias/20281231/621375000cf21847f0ac8812>,” CNN, 2022. [Online].
- [13] S. L. G. C. V. a. M. J. B. Wolfert, “Big Data in Smart Farming – A Review.” *Agricultural Systems*, 2017.
- [14] R. Mendelsohn, “The impact of climate change on agriculture in developing countries. *Journal of Natural Resources*,” 2009.
- [15] G. K. M. & K. P. Malhi, “Impact of climate change on agriculture and its mitigation strategies: A review. *Sustainability*,” 2021.
- [16] T. R. A. C. M. C. R. G. C. a. D. B. L. Ariel Ortiz-Bobea, “Anthropogenic climate change has slowed global agricultural productivity growth,” 2021.
- [17] N. Arora, “Climate change impact on food production and its sustainable solutions. *Environmental Sustainability*,” 2019.
- [18] S. Z. M. Ž. I. L. V. & L. D. Skendžić, “The impact of climate change on agricultural insect pests. *Insects*,” 2021.
- [19] R. Darwin, “The Impact of Global Warming on Agriculture: A Ricardian Analysis: Comment,” 1999.
- [20] D. Del Buono, “Can biostimulants be used to mitigate the effect of anthropogenic climate change? It is time to respond. *Science of the Total Environment*,” 2021.

- [21] A. Mahato, “Climate Change and its Impact on Agriculture,” 2014.
- [22] J. L. H. a. J, “Temperature extremes: Effect on plant growth and development,” 2015.
- [23] Y. S. J. W. Z. Z. S. M. K. J. Z. P. N Huang, “Climatic threshold of crop production and climate change adaptation: A case of winter wheat production in China,” 2022.
- [24] W. J. Sacks, “Crop planting dates: an analysis of global patterns,” 2010.
- [25] S. a. N. P. Russell, “Artificial Intelligence: A Modern Approach.,” 2010.
- [26] M. S. a. C. M. Bishop, Pattern Recognition and Machine Learning, 2006.
- [27] Y. B. Y. & H. G. LeCun, “Deep learning. Nature, 521(7553), 436-444.”.
- [28] I. B. Y. & C. A. Goodfellow, “Deep Learning. MIT Press,” 2016.
- [29] D. E. H. G. E. & W. R. J. Rumelhart, “Learning representations by back-propagating errors,” 1986.
- [30] C. B. S. H. M. R. B. & V. O. Zhang, “Understanding deep learning requires rethinking generalization.,” 2017.
- [31] A. S. I. & H. G. E. Krizhevsky, “Imagenet classification with deep convolutional neural networks.,” 2012.
- [32] A. S. N. P. N. U. J. J. L. G. A. N. .. & P. I. Vaswani, “Attention is all you need. Advances in neural information processing systems,” 2017.
- [33] G. D. L. Y. D. D. G. E. M. A. R. J. N. .. & K. B. Hinton, “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups.,” 2012.

- [34] K. G. B. P. M. D. P. S. & B. D. Liakos, “Machine learning in agriculture: A review.,” 2018.
- [35] N. W. M. & W. N. Zhang, “Precision agriculture—a worldwide overview. Computers and electronics in agriculture,” 2002.
- [36] A. & P.-B. F. X. Kamilaris, “Deep learning in agriculture: A survey. Computers and Electronics in Agriculture,” 2018.
- [37] S. & W. L. Khaki, “Crop yield prediction using deep neural networks. Frontiers in Plant Science,” 2019.
- [38] C. & K. J. M. Zhang, “The application of small unmanned aerial systems for precision agriculture: a review.,” 2012.
- [39] P. Bandara, Crop Recommendation System, 2020.
- [40] A. R. A. a. H. C. Kawtrakul, Development of an Expert System for Personalized Crop Planning, 2015.
- [41] A. B. Møller, Can We Use Machine Learning for Agricultural Land Suitability Assessment?, 2021.
- [42] R. C. V.-O. a. J. v. E. Manners, ProteinRich Legume and Pseudo-Cereal Crop Suitability under Present and Future European Climates, 2020.
- [43] “<https://www.nvidia.com/en-eu/data-center/a100/>,” NVIDIA. [Online].
- [44] B. N. C. e. Y. B. Dmitri Carпов, N-BEATS: NEURAL BASIS EXPANSION ANALYSIS FOR INTERPRETABLE TIME SERIES FORECASTING, 2019.
- [45] S. & S. J. Hochreiter, Long short-term memory. Neural computation, 1997.

- [46] INE,
“https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpagenumber=1&PUBLICACOEScolecao=107660&PUBLICACOESTipoo=ea&selTab=tab0,” INE. [Online].
- [47] D. Almeida, “Manual de Culturas Hortícolas,” Editorial Presença, 2006.

ANEXOS

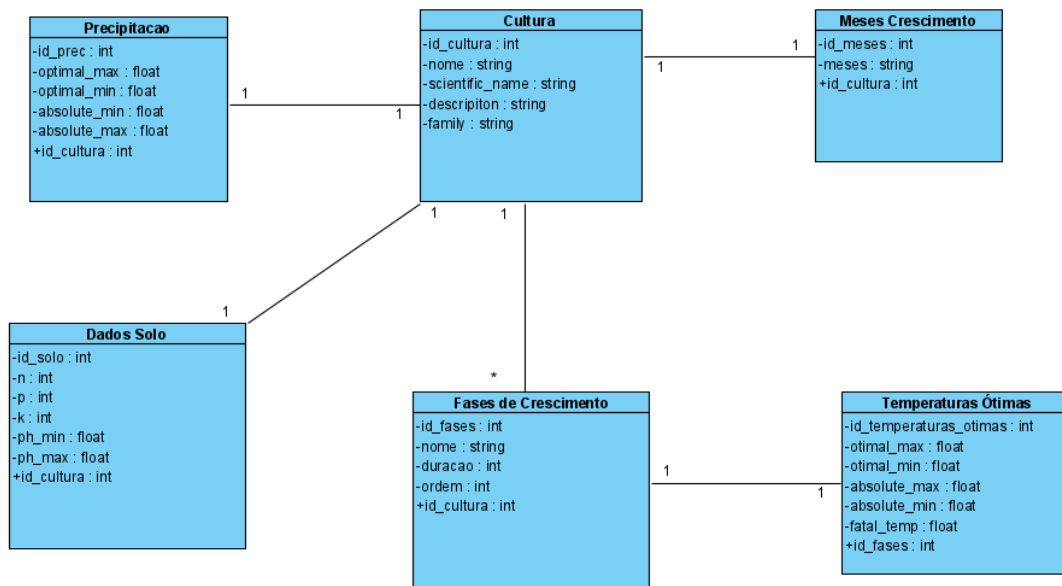


Figura 40 - Diagrama de Classes da Cultura

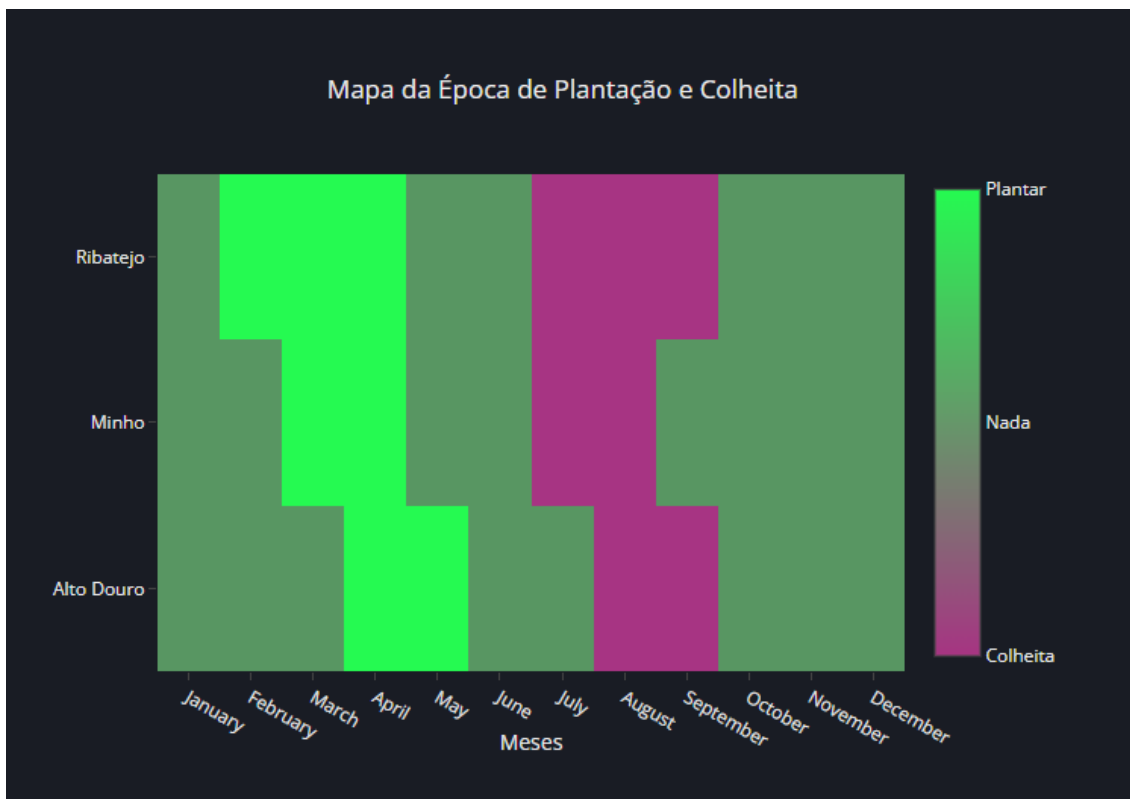


Figura 41 - Época de Plantação para a Batata. Fonte: Manual de Culturas Hortícolas

II

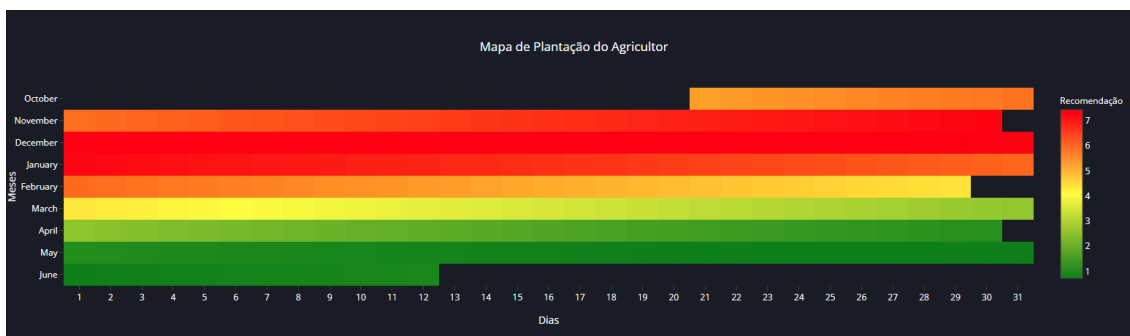


Figura 42 - Recomendação do Algoritmo para a Batata

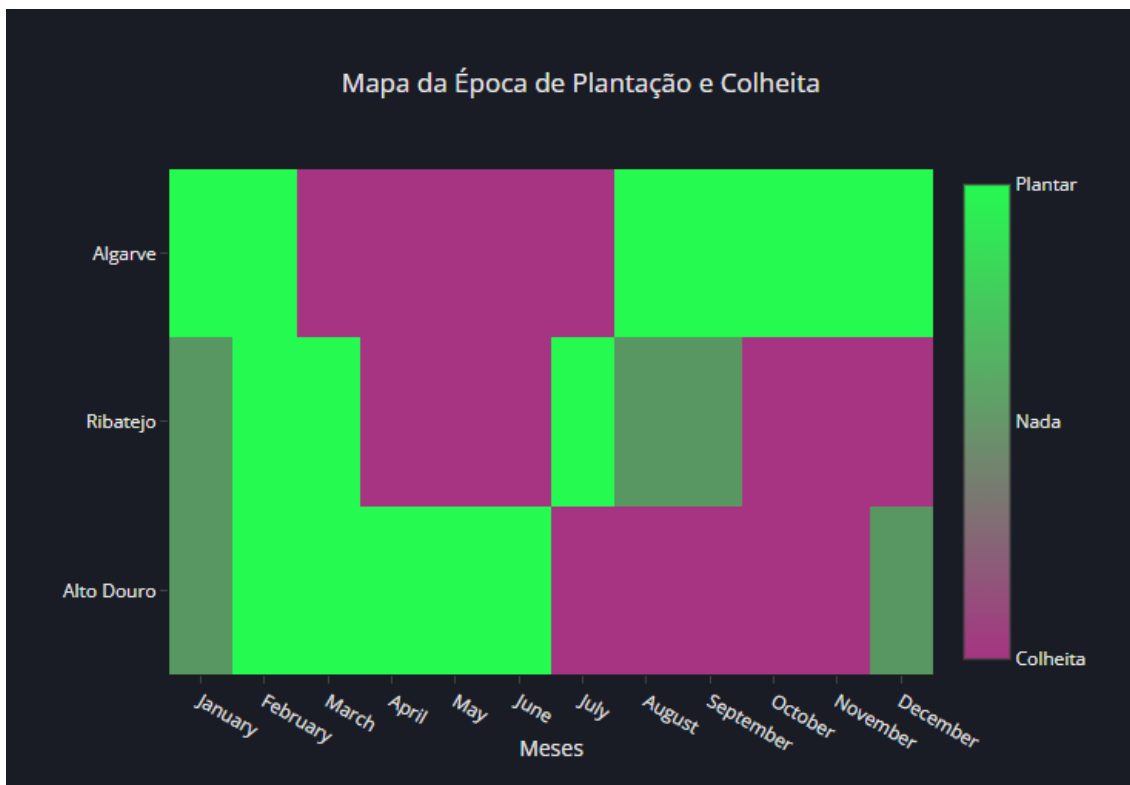


Figura 43 - Época de Plantação para o Tomate. Fonte: Manual de Culturas Hortícolas

II

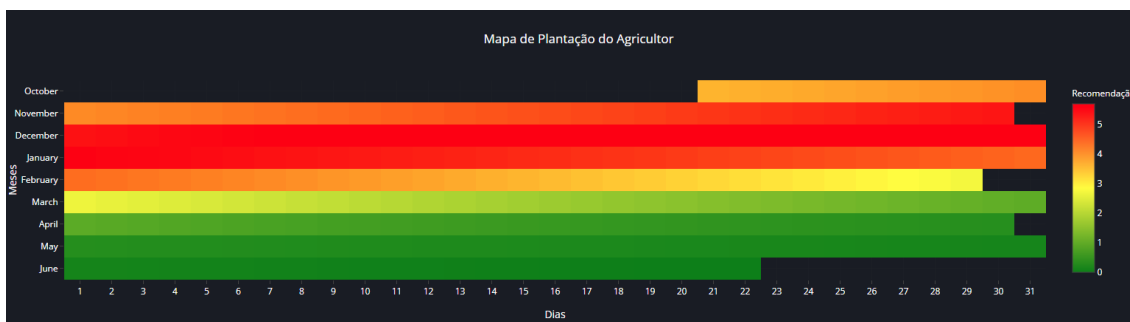


Figura 44 - Recomendação do Algoritmo para o Tomate

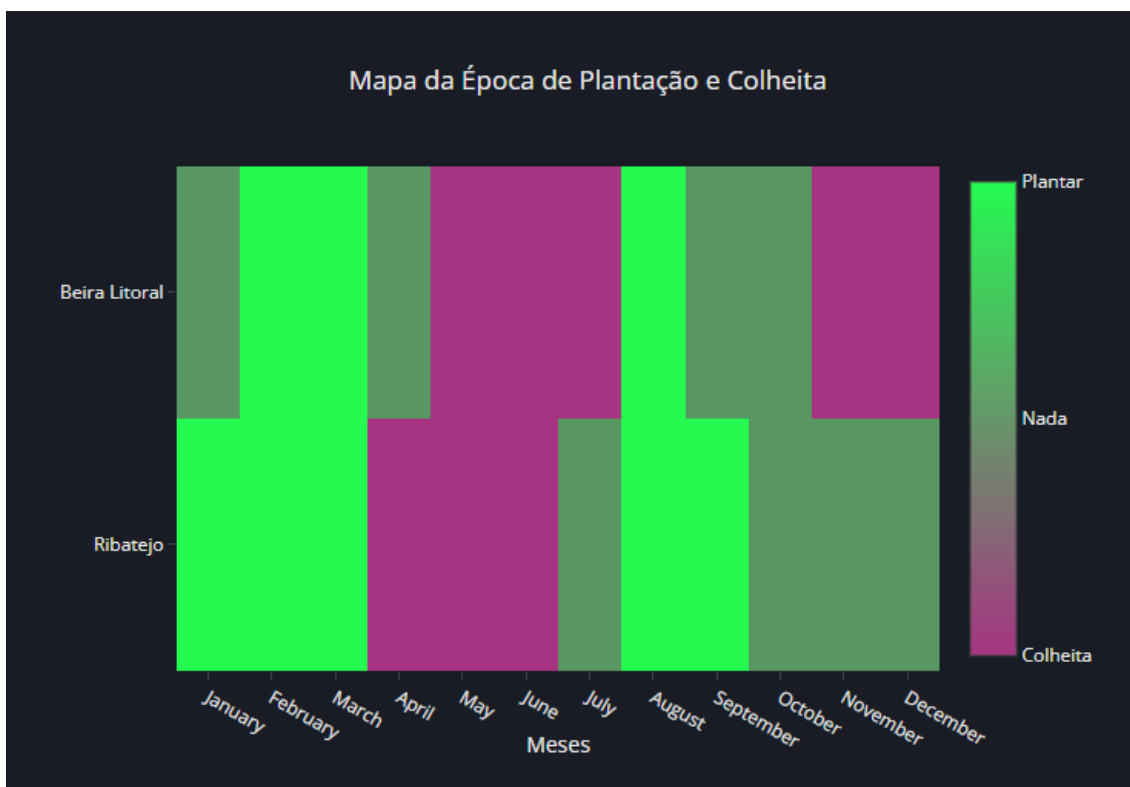


Figura 45 - Época de Plantação para o Feijão Verde. Fonte: Manual de Culturas Hortícolas II

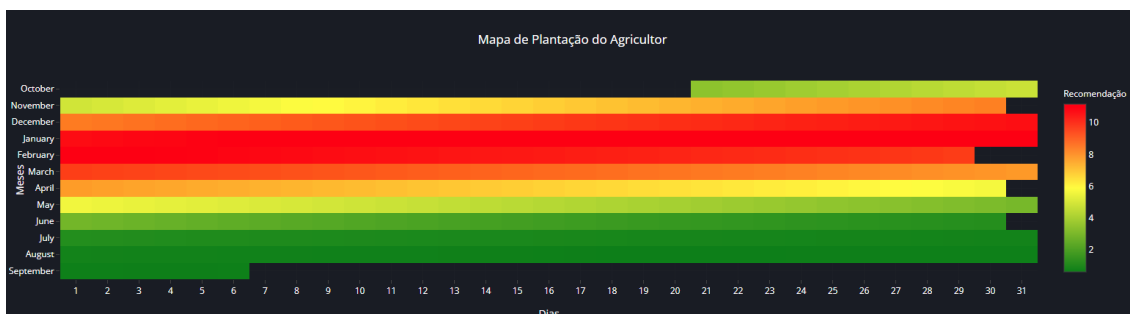


Figura 46 - Recomendação do Algoritmo para o Feijão Verde

