

Deteção de ignições florestais usando a metodologia *Fuzzy Logic*

Manuel João Pereira

Engenharia Industrial

Novembro, 2020

Deteção de ignições florestais usando a metodologia *Fuzzy Logic*

Manuel João Pereira

Relatório da dissertação apresentada à **Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Bragança**

Para obtenção do Grau de Mestre em

Engenharia Industrial

Área de especialização em Engenharia Mecânica

sob Orientação de:

Prof. Dra. Ana Isabel Pereira

Novembro, 2020

Certifico que li este relatório e que na minha opinião, é adequado no seu conteúdo e forma como demonstrador do trabalho desenvolvido no âmbito da Unidade Curricular de Projeto.

Ana Pereira, Orientadora

Certifico que li este relatório e que na minha opinião, é adequado no seu conteúdo e forma como demonstrador do trabalho desenvolvido no âmbito da Unidade Curricular de Projeto.

Arguente

Aceite para avaliação da Unidade Curricular de Projeto

A Escola Superior de Tecnologia e Gestão não se responsabiliza pelas opiniões expressas neste relatório.

Agradecimentos

Com esta tese é mais um capítulo que concluo e quero agradecer a um conjunto de pessoas que me apoiaram e ao mesmo tempo contribuíram a tornar mais fácil superar esta fase da minha vida. Quero agradecer em especial:

À minha orientadora, Professora Doutora Ana Isabel Pereira pelo seu enorme empenho e dedicação, mas também pelo vasto conhecimento que me transmitiu ao longo dos anos como docente.

À Escola Superior de Tecnologia e Gestão de Bragança, do Instituto Politécnico de Bragança, pela disponibilidade dos equipamentos necessários.

Não podendo deixar de dar um enorme agradecimento aos meus pais que sem o esforço deles nada disto seria possível.

Por fim, só me resta agradecer aos meus amigos pela sua amizade, companheirismo e espírito de entreajuda a qual sempre aumentou ao longo do nosso percurso académico.

Resumo

O principal objetivo desta dissertação é apresentar um sistema de detecção de ignições florestais o qual é baseado num algoritmo de inteligência artificial utilizando a metodologia *Fuzzy Logic*. Este sistema vai ter a capacidade de identificar pequenos focos de incêndio através da análise do desvio dos dados lidos, dando assim um alerta, recorrendo para tal a sensores de chama, humidade e temperatura, localizados nas florestas em estudo. Este método consiste então em técnicas de aprendizagem supervisionada e não supervisionada, para fornecer uma série temporal de valores de perigo e que depende apenas de variáveis climáticas e sensores de chama como dito anteriormente. Posto isto a resposta de alarme será fornecida consoante o erro obtido através da comparação de uma base de dados com uma base de teste com valores lidos em tempo real na floresta.

Palavras-chave: *Fuzzy Logic*; *Machine Learning*; Incêndio Florestal; Sensores

Abstract

The main objective of this dissertation is to present a forest ignition detection system which is based on an artificial intelligence algorithm using the *Fuzzy Logic* methodology. This system will have the ability to identify small fires by analyzing the deviation of the data read, this giving an alert, using flame, humidity and temperature sensors located in the forests under study. This method then consists of supervised and unsupervised learning methods, to provide a time series of hazard values and which depends only on climatic variables and flame sensors as stated above. That said, the alarm response will be provided depending on the error obtained by comparing a database with a test base with values read in real time in the forest.

Keywords: *Fuzzy Logic*; *Machine Learning*; Forest fire; Sensors

Índice

Capítulo 1	1
Introdução.....	1
1.2 Objetivos.....	3
1.3 Estrutura do Relatório.....	4
Capítulo 2	5
Inteligência Artificial em Contexto Florestal	5
2.1 Inteligência Artificial.....	5
2.2 <i>Machine Learning</i>	7
2.2.1 <i>Naive Bayes Classification</i>	10
2.2.2 <i>Decision Tree</i>	10
2.2.3 <i>K-Nearest Neighbor</i>	11
2.2.4 Redes Neurais Artificiais	11
2.2.5 <i>Support Vector Machine</i>	12
2.4 Método <i>Fuzzy Logic</i>	13
2.4.1 Algoritmo <i>Fuzzy Logic</i>	15
2.4.2 Implementação do <i>Fuzzy Logic</i>	18
Capítulo 3	19
Dados.....	19
3.1 Aquisição de Dados	19
3.2 Local de Estudo	24
Capítulo 4	27
Caracterização dos dados.....	27
4.1 Caracterização da base de dados.....	29
4.2 Caracterização da base de teste.....	30

Capítulo 5	31
Resultados obtidos e sua discussão	31
5.1 Teste com valores de base de dados e base de teste iguais	31
5.2 Teste com valores de base de dados e base de teste diferentes.....	34
5.3 Teste com diversas combinações de valores de base de dados e base de teste.....	37
5.3.1 BD - CSIIF	38
5.3.2 BD-SIIF.....	40
5.3.3 BD-CIIF	42
5.4 Ensaio realizado com dados reais	44
5.5 Discussão de Resultados	45
Conclusões e Trabalhos Futuros.....	46
Bibliografia.....	47

Índice de Figuras

Figura 1 - Diagrama de funcionamento de Machine Learning (Da Silva, 2017)	7
Figura 2 - Exemplo de um sistema de reconhecimento machine learning (Soares, 2019).	8
Figura 3 - Exemplo de um sistema de filtração machine learning (Soares, 2019).....	8
Figura 4 - Esquema representativo dos diferentes tipos de Machine Learning (Soares, 2019).....	9
Figura 5 - Representação do grau de utilização dos algoritmos pertencentes ao Machine Learning (Sena, 2020)	12
Figura 6 - Estrutura de um sistema baseado em Fuzzy Logic (Da Silva, 2017)	14
Figura 7 - Gráfico de funções de associação	15
Figura 8 - Gráfico da defuzificação	17
Figura 9 - Sensor de chama	20
Figura 10 - Sensor de humidade do solo	20
Figura 11 - Sensor de humidade e temperatura do ar	21
Figura 12 - Sensor de UV	22
Figura 13 - Interior do Módulo.....	23
Figura 14 - O módulo no local de trabalho.....	23
Figura 15 - Localização geográfica da serra da noqueira	24
Figura 16 - Protótipo da rede implementada (Brito , Pereira, Lima , Castro, & Valente, 2020).....	25
Figura 17 - Dados obtidos do ensaio dos 5 sensores	27
Figura 18 - Valores medidos da base de dados com informação de ignição florestal....	29
Figura 19 - Valores medidos da base de teste com informação de ignição florestal.....	30
Figura 20 - Importância de cada Sensor	32
Figura 21 - Importância de cada combinação de Sensores.....	32
Figura 22 - Erro do esperado em relação ao obtido.....	33
Figura 23 - Gráfico de superfície do risco de incêndio	33
Figura 24 - Importância de cada sensor em ambiente de incêndio.....	34
Figura 25 - Importância de cada combinação de Sensores em ambiente de incêndio ...	35
Figura 26 - Erro do esperado em relação ao obtido em ambiente de incêndio.....	35

Figura 27 - Gráfico de superfície do risco de incêndio	36
Figura 28 - Comparação da aprendizagem e da validação	36
Figura 29 - Gráficos do erro e dispersão respectivamente	38
Figura 30 - Gráficos do erro e dispersão respectivamente	39
Figura 31 - Gráficos do erro e dispersão respectivamente	39
Figura 32 - Gráficos do erro e dispersão respectivamente	40
Figura 33 - Gráficos do erro e dispersão respectivamente	41
Figura 34 - Gráficos do erro e dispersão respectivamente	41
Figura 35 - Gráficos do erro e dispersão respectivamente	42
Figura 36 - Gráficos do erro e dispersão respectivamente	43
Figura 37 - Gráficos do erro e dispersão respectivamente	43
Figura 38 - Gráfico do erro com dados reais	44
Figura 39 - Gráfico dispersão com dados reais	45

Índice de Tabelas

Tabela 1 - Categorias Fuzzy Logic (G. Lewis & J. Denming, 2019).....	15
Tabela 2 - Tabela de regras na sua forma geral.....	16
Tabela 3 - Tabela de regras para o exemplo da temperatura.....	16
Tabela 4 - Tabela de condições impostas generalizada.....	17
Tabela 5 - Tabela de condições impostas	17
Tabela 6 - Valores de média e desvio padrão dos sensores com uma distância de 100cm	28
Tabela 7 - Valores de média e desvio padrão dos sensores com uma distância de 10cm	28

Capítulo 1

Introdução

Na atual sociedade um crescimento sustentável implica monitorizar o nosso ambiente para medir o impacto do ser humano e detetar em simultâneo, alterações relevantes e prejudiciais, tais como, incêndios florestais e a desertificação. Este trabalho de investigação está inserido no promovido pelo Projeto SAFe – Sistema de Alerta Florestal, financiado pela fundação La Caixa e Centro de investigação em Digitalização e Robótica Inteligente (CeDRI) do Instituto Politécnico de Bragança. O projeto SAFe tem como objetivo estudar e analisar os dados (sensores de chama, temperatura e humidade) provenientes de estações sensorizadas montadas em vários pontos florestais estratégicos de forma a conseguir utilizar estes para identificar eficazmente uma ignição florestal num determinado local. Estas estações estão colocadas em locais de difícil visibilidade para os postos de vigia, nomeadamente vales e matas muito densas. Os dados daí provenientes serão transmitidos em tempo real para uma central onde são analisados por algoritmos de *machine learning*.

O fogo é um fenómeno de combustão manifestado em forma de luz, chama e calor. Para que este mesmo possa ocorrer é necessária a presença de três elementos principais, os quais são, oxigênio, combustível e uma fonte de ignição. Este fenómeno sempre fez parte da vida do Homem, desde as aplicações mais simples como o cozinhar e geração de energia, mas em simultâneo também pode resultar em grandes estragos. De acordo com os dados do *Centro de Estatística de Incêndio*, anualmente este fenómeno é responsável pela morte de 70.000 a 80.000 pessoas em todo o mundo (G. Lewis & J. Denning, 2019).

Anteriormente à introdução da tecnologia, a detecção de ignições florestais era realizada por uma simples inspeção visual. Mas já nos tempos pré-guerra se constatou que isto não era o suficiente, quando Walter Jaeger, em 1930, acidentalmente descobriu um sensor capaz de detetar colunas de fumo (A. Sowah, R. Ofoli, N. Krakani, & Fiawoo, 2017). Estes sensores são classificados em fotoelétricos e baseados em ionização. Os detetores fotoelétricos empregam a capacidade de dispersar a luz e detetar pequenas partículas de fumo, enquanto que os detetores baseados na ionização são projetados com uma câmara ionizada, a qual contem um radioisótopo. Na ausência de partículas de fumo, as moléculas de ar ionizadas dentro da câmara possibilitam a passagem de uma pequena corrente elétrica entre os elétrodos, quando existe presença de fumo essa corrente elétrica sofre uma ligeira queda.

Nem todos os aspetos de um incêndio florestal são negativos, os incêndios são essências para manter o equilíbrio nos ecossistemas e para a regulação da vegetação existente nas florestas. No entanto, estes mesmos, tem de ser geridos de forma adequada, caso contrário o resultado pode ser desastroso para o ambiente natural e para as propriedades (Jimenez, et al., 2016).

Um exemplo prático de um sistema implementado está situado no parque nacional Peneda-Gêres, o sistema designado por CICLOPE foi montado no ano de 1998. Este acabou por se mostrar pouco preciso em relação ao número de ocorrência falsas emitidas durante um ano de estudo foram emitidas 578 ocorrências das quais 369 falsas e 209 verdadeiras, este sistema acabou por ser retirado do serviço (Da Silva, 2017).

Na maior parte dos sistemas usados na detecção de ignições florestais, atuam automaticamente com um alerta sonoro através de uma sirene ou uma luz estroboscópica. No entanto, não é o sistema mais conveniente, pois estes mesmos necessitam de um acionamento manual do alarme. Notificação remota de uma ignição foi uma revolução e o seu desenvolvimento tem sido enorme nas últimas décadas, com a rápida adoção de dispositivos domésticos inteligentes adicionando funcionalidades ICT aos (Alves, 2018).

Neste trabalho vamos apresentar o desenvolvimento de um sistema de detecção de ignições florestais baseado no algoritmo *Fuzzy Logic*, com acesso aos dados em tempo real provenientes de sensores de chama, humidade e temperatura.

Como conhecemos a maior parte dos sistemas computacionais, eles somente trabalham com 0 e 1, mas este algoritmo assume e trata todos os valores entre estes. Um exemplo

muito fácil de entender na questão dos incêndios é a probabilidade de existir um incêndio num dia de chuva, ou seja, se um sensor deteta fumo deveria ser atuado um alarme, mas usando *Fuzzy Logic* este mesmo vai analisar os dados provenientes do sensor de humidade relativa do ar antes de dar o alarme de modo a ser uma resposta fiável.

A fiabilidade de um sistema de deteção de ignições florestais depende em grande parte dos sensores usados e da sua capacidade de registar precisamente pequenas discrepâncias de valores. *Fuzzy Logic* providencia uma forma relativamente fácil de processamento de dados vindos de diversas fontes, agrupando-os de forma a conseguir tomar uma decisão (fuzificação). Por exemplo, é possível classificar a temperatura em frio, normal e quente, assumindo para tal uma escala de valores (graus °C neste caso), o que vai permitir ao *Fuzzy* utilizar as regras de lógica e exclusão para tomar uma decisão. Este tipo de método vai ser aplicado para todas as variáveis em estudo (temperatura, humidade e sensor de chama), posto isto serão reunidos todos os dados obtidos pelo algoritmo e é tomada a decisão se existe ou não uma ignição na área identificada (defuzificação) (A. Sowah, R. Ofoli, N. Krakani, & Fiawoo, 2017).

1.2 Objetivos

No presente trabalho, propõe-se um método para identificar rapidamente ignições florestais, de forma a minimizar a propagação e os consequentes danos dos mesmos. Pretendemos uma deteção num período de tempo máximo de 6min, visto que depois deste período a probabilidade de a ignição florestal ter evoluído para incêndio é bastante elevada. Para atingir os objetivos deste trabalho são analisados diversos parâmetros, tais como, a temperatura ambiente, a humidade relativa, e os valores provenientes de um sensor de chama. Os dados são analisados recorrendo ao software MatLab.

Este trabalho possui os seguintes objetivos:

- Escolha definitiva dos parâmetros a utilizar;
- Escolha do melhor algoritmo possível, tanto em termos de resultados como em termos de facilidade de aplicação;
- Caracterização de dados obtidos pelos sensores;
- Análise dos resultados obtidos e posterior comparação com outros métodos.

1.3 Estrutura do Relatório

Este trabalho está dividido em cinco capítulos com seguinte organização:

O primeiro capítulo será de caráter introdutório, tendo por finalidade realizar um enquadramento do estudo, bem como referir os objetivos do trabalho e uma breve descrição sobre a estrutura da dissertação.

O segundo capítulo vai ser mais dedicado as bases teóricas utilizadas, vão ser também mencionadas todas as escolhas efetuados durante o mesmo. Posteriormente será descrito todo o processo de cálculo, incluindo a descrição do algoritmo de *Fuzzy Logic*, a implementação do mesmo e as funções predefinidas utilizadas.

O terceiro capítulo vai incidir somente na obtenção dos dados utilizados, nomeadamente tipos de sensores utilizados e local em estudo

No quarto capítulo procedemos à caracterização dos dados, da base de dados e da base de teste.

No quinto e último capítulo concluímos o estudo demonstrando os resultados obtidos bem como a sua discussão.

Capítulo 2

Inteligência Artificial em Contexto Florestal

Neste capítulo vai ser apresentado de uma forma geral o que é o *Machine Learning* e quais as suas vertentes mais utilizadas. Posteriormente vai ser abordado o estado da arte associado à inteligência artificial na gestão florestal e por fim a apresentação do algoritmo *Fuzzy Logic*.

2.1 Inteligência Artificial

Hoje é comum ouvirmos falar de inteligência artificial através dos diversos meios de comunicação, e de alguma forma existe uma dicotomia entre o que o nosso imaginário preenche em relação à inteligência artificial e o que de facto temos ao dispor para nos auxiliar no nosso mundo. Ainda existem várias diferenças significativas entre, se assim o quisermos designar, as máquinas de carbono que são os humanos e as máquinas de silício que são os computadores (Soares, 2019).

De uma certa forma a inteligência artificial faz parte do nosso quotidiano discretamente, está oculta sob diversas formas, por via tecnológica, escrita, cinema, televisão ou rádio. Neste momento é um ramo que se encontra em constante desenvolvimento, não podendo ainda ser comparada a inteligência artificial com a humana, existem inúmeras diferenças significativas, sendo a consciência a que tem maior expressão (Soares, 2019).

Sendo assim, a que se resume a inteligência artificial? Este processo envolve um conjunto muito alargado de áreas que chegam desde a aprendizagem e a perceção ao mais

específico, como por exemplo jogar xadrez, conduzir um carro numa rua movimentada, prever doenças ou outro tipo qualquer de acontecimento.

Atualmente existem duas vertentes da inteligência artificial, (IA), IA estreita (conhecida por ANI (*Artificial Narrow Intelligence*)) é a mais utilizada neste momento, e a AI genérica também conhecida por AGI (*Artificial General Intelligence*) (Soares, 2019)

- *ANI Artificial Narrow Intelligence*

Artificial Narrow Intelligence (ANI), também conhecida como IA “Fraca”, é a IA que mais existe no mundo hoje. *Narrow IA* é o IA programado para realizar uma única tarefa - seja verificar o tempo, ser capaz de jogar xadrez ou analisar dados brutos para escrever relatórios.

Os sistemas ANI podem atender a uma tarefa em tempo real, mas extraem informações de um conjunto de dados específico. Como resultado, esses sistemas não funcionam fora da tarefa única para a qual foram projetados.

Ao contrário da IA geral, a IA estreita não é consciente ou movida pela emoção como os humanos. A IA estreita opera dentro de uma faixa pré-determinada e pré-definida, mesmo que pareça muito mais sofisticada do que isso.

- *AGI (Artificial General Intelligence)*

Esta vertente está relacionada com a possibilidade do sistema ou máquina inteligente ter a simples capacidade de executar qualquer atividade intelectual que um ser humano seja capaz de executar. Este processo tenta ao máximo copiar o pensamento lógico do ser humano. Pela maioria dos investigadores esta área é ainda considerada mera ficção científica. Devemos também ter em consideração que na generalidade a evolução da IA se tem apoiado no funcionamento do cérebro do ser humano, resultando que os modelos matemáticos, ou algoritmos construídos, acabam por ser representações aproximadas da forma como os seres humanos funcionam (Soares, 2019). Esta vertente da inteligência artificial é a mais promissora nesta área.

2.2 Machine Learning

O *Machine Learning (ML)* é um ramo da inteligência artificial que utiliza a criação de algoritmos e técnicas que visam a extração de padrões e regras computacionais. As técnicas de *Machine Learning* são amplamente aplicadas em reconhecimento de padrões e filtração de dados, utilizando algoritmos computacionais e métodos estatísticos para a extração de respostas (Weiqiang, 2019).

Haykin afirma a existência de inúmeras atividades associadas à aprendizagem e as quais não possuem definição precisa, pois dependem do ponto de vista do observador e do contexto em que se encontram (Soares, 2019).

Neste campo de estudo, Cherkassky e Mulier definem como exemplo das múltiplas definições de aprendizagem um ramo mais dedicado à estatística, em que, aprender é o processo de estimar uma estrutura desconhecida ou dependência de um sistema a partir de um número finito de observações, sendo que de forma geral a aprendizagem abrange três componentes, gerador, sistema e máquina de aprendizagem como representado na Figura 1 (Zhao & K. Bose, 2002).

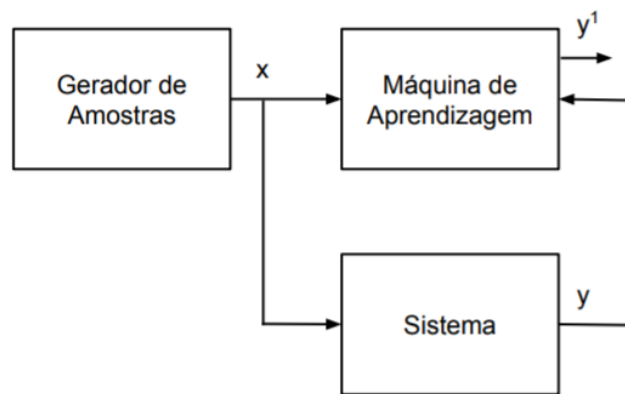


Figura 1 - Diagrama de funcionamento de Machine Learning (Da Silva, 2017)

O gerador de amostras é utilizado para a obtenção de valores de entrada, a partir de uma fonte de interesse externa.

Por sua vez o sistema é responsável pela produção de um valor de saída para cada conjunto de entrada a partir de uma densidade condicional fixada e desconhecida. Por fim, o objetivo do procedimento da máquina de aprendizagem é proceder à criação de funções $f(x, w)$, (Zhao & K. Bose, 2002).

Daqui retiramos que um algoritmo de aprendizagem é um conjunto preestabelecido de regras bem definidas para a solução do problema de aprendizagem, como representado na Figura 2 e 3.

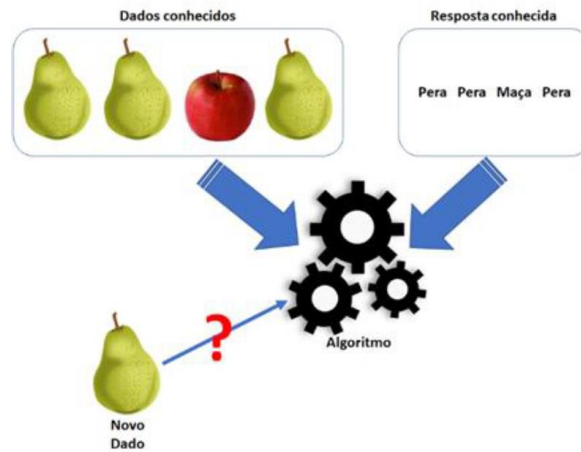


Figura 2 - Exemplo de um sistema de reconhecimento machine learning (Soares, 2019).

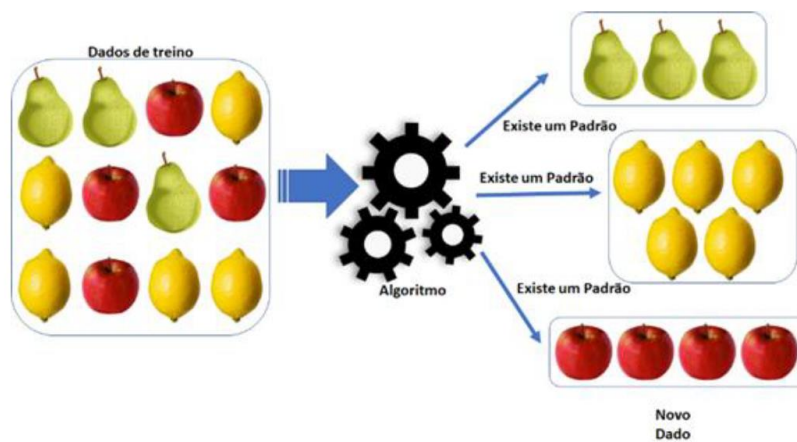


Figura 3 - Exemplo de um sistema de filtração machine learning (Soares, 2019).

Existem diferentes tipos de *Machine Learning* que servem para resolver problemas em vários campos, representados na Figura 4.

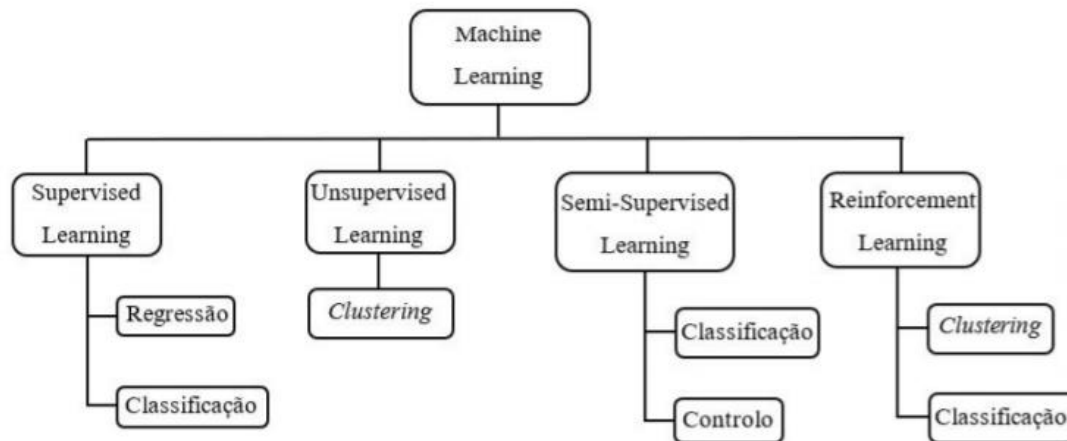


Figura 4 - Esquema representativo dos diferentes tipos de Machine Learning (Soares, 2019).

Supervised Learning - é uma técnica muito semelhante ao processo de aprendizagem do ser humano. É realizado sob treino, em que o algoritmo é constituído por pares de entrada e pares de saída. Divide-se em diferentes técnicas com diferentes algoritmos (Sena, 2020):

- **Regressão:** Regressão Linear, Regressão Lógica e Decision Tree.
- **Classificação:** *Support Vector Machine (SVM)*, *Naive Bayes Classification*, *Decision Tree*, *K-Nearest Neighbor*, Redes Neurais Artificiais e *Fuzzy Logic*.

Unsupervised Learning - é utilizado, geralmente, para investigar as características e pré-processamento dos dados. É um método semelhante a um aluno que apenas resolve os problemas por construção e atributo, mas que não aprende a resolvê-los porque não têm saídas corretas conhecidas, assim os seus dados para treinar o modelo contêm apenas dados de entrada. Divide-se numa técnica com diferentes algoritmos (Sena, 2020):

- **Clustering:** *K-means*, *Hierarchical model* e *Mixture model*.

Reinforcement Learning - é usado, geralmente, quando é necessária uma interação ideal, como controlar as jogadas em jogos, um exemplo é o jogo de xadrez (Sena, 2020).

Divide-se em diferentes algoritmos:

- **Classificação e Clustering:** *Low density Separation, Graph Based method e Heuristic Approach.*

Semi-Supervised Learning- é uma técnica que identifica o melhor classificador de cada informação não rotulada e rotulada. Utilizando informações não rotuladas, que consiste em ocorrências diferentes dos próprios sinais de entrada, consegue-se obter alto desempenho de classificação. Divide-se em diferentes algoritmos (Sena, 2020):

- **Classificação e Controlo:** *Monte Carlo Method, Temporal Difference method e Direct policy Search.*

O *Supervised Learning* é a técnica de *ML* escolhido para elaborar este estudo, como referido anteriormente, é dividido em técnicas de classificação e regressão. As subsecções seguintes descrevem brevemente alguns algoritmos constituintes da técnica de classificação do *Supervised Learning*, uma vez que é o método utilizado ao longo do estudo (Sena, 2020).

2.2.1 Naive Bayes Classification

O método de *Naive Bayes*, é um classificador probabilístico que consegue prever as probabilidades de associação de categorias baseadas no rótulo de determinada classe. Executa uma digitalização de dados e, logo, a classificação é fácil. O *Naive Bayes Classification* pode ser utilizado, por exemplo, para desenvolver um sistema probabilístico de identificação do cancro da mama. Este classificador pode ser utilizado para ajudar um especialista numa decisão com uma alta precisão (Sena, 2020).

2.2.2 Decision Tree

A *Decision Tree* é uma ilustração gráfica de uma decisão exata, que pode ser utilizado para regressão ou classificação. Os nós, as raízes e as decisões de ramificação são os principais componentes deste sistema. A *Decision Tree* pode ser utilizada, quando existem problemas médicos com um elevado número de parâmetros de classificação de um conjunto de dados. As duas principais desvantagens deste algoritmo são o *overfitting*,

que acontece quando existe um elevado número de dados e não existem nós suficientes para espalhá-los. Existem diferentes abordagens que podem ser utilizadas na *Decision Tree*, como o ID3, CYT, C5.0, J48 e o CART que é a abordagem mais utilizada e está estruturado como uma sequência de perguntas, cujas respostas determinam qual será a próxima pergunta, caso existir. Os resultados a essas perguntas aparecem em estrutura de árvore onde as extremidades são os nós terminais (Sena, 2020).

2.2.3 *K-Nearest Neighbor*

A técnica de *K-Nearest Neighbor* é frequentemente usada na classificação das amostras, ao usar esta técnica a distância é calculada para treinar o conjunto de dados para cada K pontos de dados mais próximos a partir do número N de amostras de treino. O princípio de funcionamento do *K-Nearest Neighbor* é baseado numa atribuição de peso para cada ponto de dados conhecido como vizinho. Este algoritmo é dividido em duas etapas, na primeira treina a partir dos dados e na segunda testa em novas instâncias.

É um algoritmo supervisionado e muito lento, visto que demora mais tempo para conseguir obter classificação treinadas, mas é especialmente eficaz para grandes conjuntos de dados (Sena, 2020).

2.2.4 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNA's) são utilizadas na ciência da computação como modelos computacionais inspirados pelo sistema nervoso central. Todos os organismos multicelulares possuem algum tipo de sistema nervoso, cuja complexidade e organização varia de acordo com o tipo de animal (Osório & Bittencout, 2000). Este sistema é responsável por dotar o organismo, através de entradas sensoriais, de informações sobre o estado do ambiente no qual vive e se move. A informação de entrada é processada, comparada com as experiências passadas, e transformada em ações apropriadas ou absorvidas sob a forma de conhecimento. O sistema nervoso pode ser organizado em diferentes níveis: moléculas, sinapses, neurónios, camadas, mapas e sistemas (Osório & Bittencout, 2000). A representação de conhecimentos nas redes neurais, como diz o próprio nome, é fortemente ligada a noção de conexão entre neurónios (elementos processadores de informação) que interagem uns com os outros através das ligações. O

conhecimento de uma rede neural artificial está codificado na estrutura da rede, onde se destacam as conexões (sinapses) entre as unidades (neurónios) que a compõe. Assim, as redes neurais artificiais podem ser definidas como estruturas de processamento (rede) compostas por um número de unidades interconectadas (neurónios artificiais), sendo que cada unidade apresenta um comportamento específico de entrada/saída determinado pela sua função de transferência (Sena, 2020).

2.2.5 Support Vector Machine

O *Support Vector Machine* (SVM) é um conceito na ciência da computação projetado para fornecer suporte às tarefas de *Machine Learning*. É uma ferramenta simples e com um processo de destaque. Originalmente projetada para classificação binária, podendo ser estendida para classificação multi-classe (Sena, 2020).

O modelo SVM representa pontos no espaço bidimensional, com o intuito de encontrar um hiperplano ideal que consiga separar os pontos, em duas classes, caso se trate de uma classificação binária. Com o hiperplano pretende-se minimizar a distância entre os pontos mais próximos em relação a cada uma das classes. Os pontos que se encontrarem mais próximos do hiperplano designam-se por vetores de suporte.

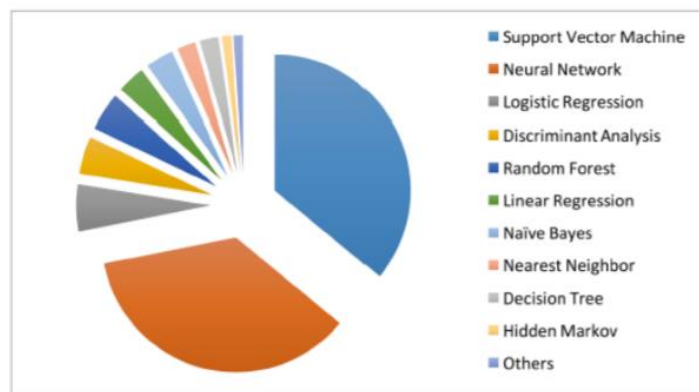


Figura 5 - Representação do grau de utilização dos algoritmos pertencentes ao Machine Learning (Sena, 2020)

A Figura 5 faz referência ao grau de utilização de cada algoritmo referido anteriormente, mostrando que o *Support Vector Machine* e as Redes Neurais Artificiais são os mais utilizados.

2.4 Método *Fuzzy Logic*

Neste subcapítulo vamos então dar introdução ao algoritmo *Fuzzy Logic*, sobre o qual esta dissertação vai incidir, aos seus diversos constituintes e à forma como efetua o seu cálculo. Vamos também explicar o que é necessário para o correto funcionamento do algoritmo e quais os valores de entrada necessários.

O algoritmo *Fuzzy Logic* trabalha com inúmeros valores entre 0 e 1 usando para tal a lógica, este é empregado para lidar com o conceito de verdade parcial, ou seja, pode variar entre ser completamente verdadeiro ou completamente falso (Jimenez, et al., 2016).

- Fuzificador

A fuzificação é o processo de atribuir a entrada numérica de um sistema a conjuntos de fuzzy com um certo grau de associação. Como já foi mencionado este grau de associação pode variar entre 0 e 1, caso seja 0, o valor não pertence ao conjunto, em contrapartida se for 1, o valor pertence completamente ao conjunto. Este valor entre 0 e 1 vai representar a incerteza de pertencer ao conjunto (Zhao & K. Bose, 2002).

- Operadores de fuzzy

Existem quatro operadores utilizados por este sistema, nomeadamente, “*AND*”, “*OR*”, “*NOT*” e por fim o “*IF THEN*”. Cada uma delas desempenha uma função específica, para o nosso caso a mais adequada será então o “*IF THEN*”. Um exemplo prático para esta regra pode ser a seguinte já aplicada aos incêndios:

```
IF temperature IS very cold THEN no fire
IF temperature IS cold THEN possibility of fire
IF temperature IS warm THEN fire
IF temperature IS hot THEN surely fire
```

- Defuzificador

O objetivo da defuzificação é a obtenção de uma variável continua a partir de valores introduzidos, isso seria relativamente fácil caso os valores verdadeiros de saída fossem exatamente os obtidos na fuzificação. Mas como, todos os valores verdadeiros de saída

são calculados de forma independente, na maior parte dos casos eles representam um conjunto de números, ou seja, é necessário decidir um número que corresponda então à melhor resposta para o nosso problema (Da Silva, 2017).

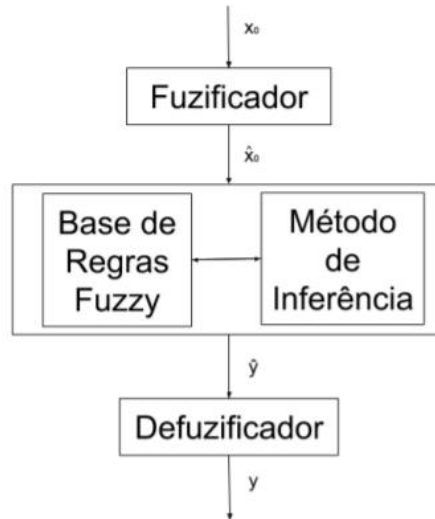


Figura 6 - Estrutura de um sistema baseado em Fuzzy Logic (Da Silva, 2017)

A Figura 6 é constituída por 4 partes, tendo cada uma delas uma determinada função para o correto funcionamento, o fuzificador, onde ocorre a fuzificação, é o processo no qual ocorre a conversão dos valores de entrada do sistema para conjuntos *Fuzzy*, com as respectivas faixas de valores onde estão definidos. Já a base de regras *Fuzzy* providencia uma maneira formal para representar as estratégias, objetivos e informações utilizadas por especialistas na área, através das suas experiências. O especialista auxilia na formulação de um conjunto de regras *Fuzzy*, articulando as associações de entradas/saídas linguísticas. O método de inferência utilizado no sistema baseado em regras *Fuzzy* é o de Mamdani, que utiliza o operador “OU” e os operadores “E” e “ENTÃO” em cada regra individual da base de regras *Fuzzy*. Para um sistema com mais de uma regra, tem-se a contribuição de cada regra individualmente, de forma a se obter a saída *Fuzzy*. Por último o defuzificador onde ocorre a defuzificação é o processo que permite a representação de um conjunto *Fuzzy* por um valor real que melhor o represente, ou seja, através de uma escala é possível dizer se o valor está dentro dos critérios impostos (Da Silva, 2017).

2.4.1 Algoritmo *Fuzzy Logic*

Neste subcapítulo vai ser apresentado a forma como é elaborado um algoritmo utilizando a metodologia *Fuzzy Logic* passo a passo.

Passo 1 – Definir variáveis linguísticas

Variáveis linguísticas (*VL*) são variáveis de entrada e saída, dados fornecidos e resposta obtida respectivamente, na forma de palavras ou frases simples,

$$VL(t) = \{C_1, C_2, C_3, \dots, C_n\},$$

onde $C_i, i = 1, \dots, n$ representa as categorias da variável. Um exemplo de $VL(t)$ associado à temperatura é por exemplo:

$$VL(t) = \{muito - frio, frio, quente, muito - quente, calor\}$$

Cada membro deste conjunto é um termo linguístico e pode cobrir uma parte dos valores gerais de temperatura.

Passo 2 – Construir funções de associação para cada variável

Nesta etapa de *fuzificação* os dados de entrada, ou seja, os dados fornecidos são transformados de números para linguagem *Fuzzy*. Neste caso são divididos em cinco categorias:

Tabela 1 - Categorias Fuzzy Logic (G. Lewis & J. Denning, 2019)

<i>LP</i>	<i>Large Positive</i>
<i>MP</i>	<i>Medium Positive</i>
<i>S</i>	<i>Small</i>
<i>MN</i>	<i>Medium Negative</i>
<i>LN</i>	<i>Large Negative</i>

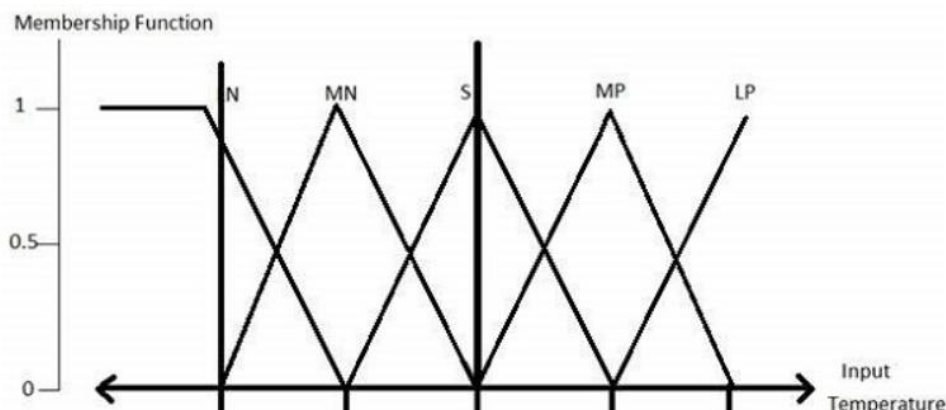


Figura 7 - Gráfico de funções de associação

Com a observação da Tabela 1 e da Figura 7 é possível analisar a escala que foi implementada ao *Fuzzy Logic* para análise dos dados de entrada, este exemplo está de uma forma relacionada com a temperatura, mas o mesmo se efetua para qualquer outro tipo de situação alterando somente a escala e a designação utilizada.

Passo 3 – Construir regras de conhecimento para cada variável

No passo 3 temos como objetivo a construção de uma tabela onde definimos as regras para o *Fuzzy Logic*, ou seja, vamos lhe dizer o que ele deve fazer em cada situação em concreto.

De uma forma geral a tabela tem o seguinte aspeto, tendo em conta as VL(t):

Tabela 2 - Tabela de regras na sua forma geral

Temperatura lida / Parâmetro	C1	C2	C3	C4	C5
C1	S/alteração	Ação 1	Ação 1	Ação 1	Ação 1
C2	Ação 2	S/alteração	Ação 1	Ação 1	Ação 1
C3	Ação 2	Ação 2	S/alteração	Ação 1	Ação 1
C4	Ação 2	Ação 2	Ação 2	S/alteração	Ação 1
C5	Ação 2	Ação 2	Ação 2	Ação 2	S/alteração

Um exemplo prático da Tabela 2 é:

Tabela 3 - Tabela de regras para o exemplo da temperatura

Temperatura lida / Parâmetro	Muito frio	Frio	Quente	Muito quente	Calor
Muito frio	S/alteração	Aquecer	Aquecer	Aquecer	Aquecer
Frio	Arrefecer	S/alteração	Aquecer	Aquecer	Aquecer
Quente	Arrefecer	Arrefecer	S/alteração	Aquecer	Aquecer
Muito quente	Arrefecer	Arrefecer	Arrefecer	S/alteração	Aquecer
Calor	Arrefecer	Arrefecer	Arrefecer	Arrefecer	S/alteração

Com a análise da Tabela 2 e da Tabela 3 podemos observar como são discriminadas as regras, numa determinada zona queremos um efeito, e em outra zona pretendemos o contrário, o que acontece com o exemplo de aquecer e arrefecer.

Seguidamente vamos mostrar as tabelas de condições para a obtenção de resposta por parte do *Fuzzy Logic*.

Tabela 4 - Tabela de condições impostas generalizada

Nº	Condição	Ação
1	IF VL(t)=(C1 OR C2) AND target = C3	Ação 1
2	IF VL(t)=(C3 OR C4) AND target= C3	Ação 2
3	IF (VL(t)=C3) AND (target=C3) THEN	S/Ação

Um exemplo da Tabela 4 é:

Tabela 5 - Tabela de condições impostas

Nº	Condição	Ação
1	IF temperature=(Frio OR Muito frio) AND target = Quente	Aquecer
2	IF temperature=(Quente OR Muito Quente) AND target= Quente	Arrefecer
3	IF (temperature=Quente) AND (target=Quente) THEN	S/Ação

Observando a Tabela 4 e a Tabela 5 conseguimos facilmente entender como definimos as condições para o *Fuzzy Logic*, não passam de condições lógicas formadas por comandos os quais ditam a comparação entre valores.

Passo 4 – Obter o valor de *Fuzzy*

As operações de *Fuzzy* executam uma avaliação de regras. As operações usadas para *OR* e *AND* são Max e Min, respetivamente. Combinando todos os resultados da avaliação é possível formar um resultado.

Passo 5 – Defuzificação

Este processo é realizado segundo as regras de conhecimento

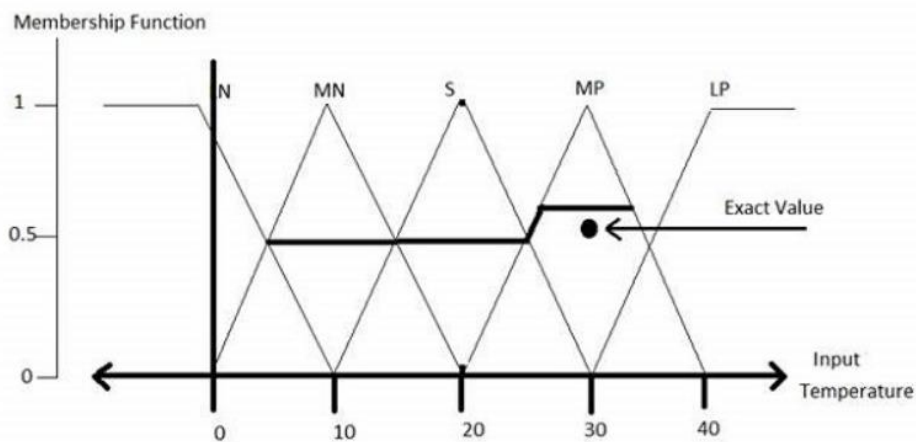


Figura 8 - Gráfico da defuzificação

Na Figura 8 é possível observar o valor exato calculado pelo *Fuzzy Logic* para o exemplo em questão da temperatura ambiente. Tendo em conta os parâmetros analisados obteve-se um *exact value* para a temperatura na casa dos 30°C, ficando assim situado na área de *Medium Positive*. Sendo assim vai responder com uma ação a qual permite chegar à nossa temperatura desejada, seja ela mais ou menos que os 30°C em questão.

2.4.2 Implementação do Fuzzy Logic

A implementação do método *Fuzzy Logic* foi baseada em funções predefinidas do *MatLab* nomeadamente *gensurf*, *anfis*, *evalfis* e *genfis*, as quais são explicadas posteriormente. O código descrito realiza diversas análises aos parâmetros de estudo, não só individualmente como também agrupados, tentando saber qual o parâmetro com mais peso no cálculo, isto é, numa determinada situação pode ser mais importante a humidade do ar do que propriamente a temperatura. Na parte inicial testamos o código meramente com valores dos sensores medidos através do laboratório, sendo que a ideia seria testar com valores provenientes de uma situação real de incêndio. Estes valores de situação real foram obtidos por ensaios laboratoriais realizados no IPB, recorrendo ao material disponível nos laboratórios. No final do relatório vai ser ainda apresentado um ensaio realizado com dados provenientes de leituras feitas com sensores no campus do IPB.

Para o correto funcionamento deste código é necessário importar uma base de dados, a qual é proveniente dos sensores, foi posteriormente convertida em formato *Excel (.xlsx)*, de forma a ser possível importar esta.

Seguidamente vamos então descrever as funções utilizadas, começando pela *gensurf*, a qual tem a função de gerar uma superfície do sistema de inferência *Fuzzy*, seguidamente usou se também a função *anfis* a qual ajusta o sistema recorrendo aos dados provenientes da base de dados, posto estas temos ainda a função *evalfis* a qual avalia a inferência do sistema *Fuzzy* e por fim temos então a função *genfis* que tem como objetivo gerar uma superfície, recorrendo para tal ao resultado obtido pelo sistema *Fuzzy*.

Capítulo 3

Dados

No Capítulo 3 vão ser abordados todos os tópicos relacionados com a obtenção dos dados utilizados para a realização do estudo, como nomeadamente o tipo de sensores e suas características, descrição da caixa completa com todos os sensores e a localização da área de atuação das mesmas.

3.1 Aquisição de Dados

Para a realização deste trabalho foram então selecionados quatro sensores distintos que possuem as seguintes especificações:

Sensor de chama

Este sensor é sensível ao espectro da chama e é ajustável consoante a intensidade a ser registada

Possui as seguintes especificações:

- Espectro de funcionamento: 760nm ~ 1100nm
- Ângulo de funcionamento: 0 - 60 graus
- Voltagem de funcionamento: 3.3V ~ 5.3V
- Temperatura: -25 °C ~ 85 °C
- Dimensões: 27,3 mm * 15,4 mm
- Diâmetro da lente: 2.0m



Figura 9 - Sensor de chama

Como podemos observar pelas especificações vemos que se trata de um sensor bastante compacto (Figura 9), mas em simultâneo possui um intervalo de valores bastante amplo e torna o perfeito para as nossas necessidades, também é necessário mencionar que este é o sensor que vai ter mais incidência sobre a resposta final do *Fuzzy Logic*.

Sensor de humidade do solo

Na Figura 10 é possível observar o sensor de humidade do solo, é constituído por duas peças separadas onde uma entra em contacto com o solo e outro modulo que contem um chip LM393 o qual vai ler e enviar os dados para um microcontrolador.

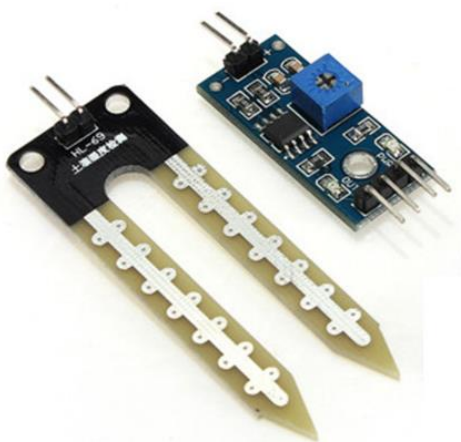


Figura 10 - Sensor de humidade do solo

Possui as seguintes especificações:

- Voltagem de funcionamento: 3.3V ~ 5V
- Sensibilidade ajustável através de um potenciômetro
- Dimensões: 30 mm * 15 mm
- Dimensões da sonda: 60mm * 20mm

Sensor de Temperatura e humidade do ar:

Na Figura 11 é possível observar o sensor de humidade e temperatura do ar, este contém um módulo digital DHT11, o qual vai ser então utilizado para obter os dados da temperatura e da humidade do ar naquele instante.

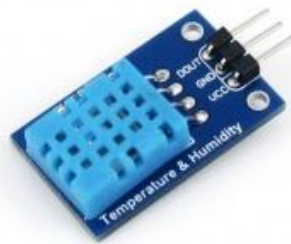


Figura 11 - Sensor de humidade e temperatura do ar

Possui as seguintes especificações:

Temperatura

- Resolução: 1 ° C
- Precisão: ± 2 ° C
- Intervalo de funcionamento: 0 ° C ~ 50 ° C

Humidade

- Resolução: 1% RH
- Precisão: ± 5 % RH (0 ~ 50 ° C)
- Intervalo de funcionamento: 20% RH ~ 90% RH (25 ° C)
- Voltagem de funcionamento: 3.3 V ~ 5,5 V

Como podemos observar a resolução em termos de temperatura varia apenas de 1 em 1 ou seja nunca vamos ter valores irregulares, visto que também não há interesse em tanta precisão para este efeito, quanto à humidade temos uma resolução bastante boa, de um 1%, mas como este dado se torna mais relevante devido à chuva é importante ter um valor de humidade preciso.

Sensor de UV

Este tipo de sensor, representado na Figura 12, é capaz de medir a radiação ultravioleta através de um chip. (Pereira, Mendes, & Pereira , 2019)



Figura 12 - Sensor de UV

Possui as seguintes especificações:

- Voltagem de funcionamento: 3.3V ~ 5V
- Dimensões: 22 mm * 16 mm
- Espectro de funcionamento: 200nm ~ 370nm

Estes sensores vão ser então todos ligados a uma placa Arduino, a qual vai processar e armazenar os dados. Posteriormente a ideia será que os dados sejam enviados para uma central de controlo onde sejam analisados em tempo real todos os valores de diversas caixas. Todos estes módulos sensorizados vão estar colocados num determinado local com o intuito de analisar toda a área existente, não havendo assim espaços vazios onde possa ocorrer uma ignição florestal.

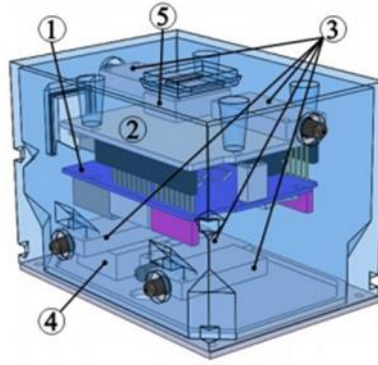


Figura 13 - Interior do Módulo

Sendo os seus componentes representados por:

1. Arduino Uno;
2. Módulo da comunicação LoRa;
3. Sensores de chama;
4. Sensor de humidade e temperatura do ar;
5. Sensor ultravioleta;

Como é possível observar na Figura 13 cada sensor tem o seu lugar predefinido de modo a ter um local de trabalho perfeito. O sensor de humidade do solo será ligado ao solo através de uma extensão de fio.



Figura 14 - O módulo no local de trabalho

O aspeto da mesma no local de trabalho será então o representado na Figura 14. O módulo estará preso a uma árvore, suficientemente alta para não sofrer danos de animais ou outras possíveis causas, onde os sensores terão um alcance mais amplo, isto é, caso esta esteja mais junta ao solo a densidade das matas não permite uma leitura ideal. Vai ser sempre necessário ter em conta a densidade das matas de forma a colocar os sensores com uma orientação onde seja possível analisar a maior área possível, caso contrário será desvantajoso colocar demasiados sensores para uma pequena área.

3.2 Local de Estudo

A aplicação dos módulos de sensores para aquisição de dados será implementada na região de Bragança, na área da Serra da Nogueira, conforme mostrado na Figura 15. Devido às características dessa floresta, espalhar os módulos de sensores em toda a região seria caótico e difícil de entender os dados (Lloret et al., 2009). Portanto, é necessário desenvolver uma estratégia para colocar os módulos. Os pontos escolhidos devem considerar os dados fornecidos pelos órgãos ambientais nacionais.

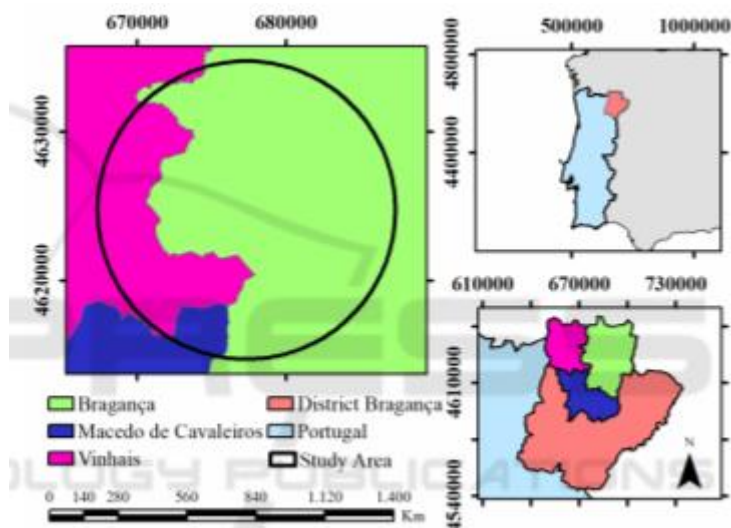


Figura 15 - Localização geográfica da serra da noqueira

Alguns fatores são determinísticos para a escolha desses pontos, como a ocupação do solo, histórico e estimativa de áreas com risco de incêndios, áreas as quais foram queimadas ao longo dos anos, elevação do terreno e densidade florestal, entre outros. Para a análise destes fatores, o software QGIS (QGIS, 2019) é usado como uma base de visualização e os dados geográficos são fornecidos por Copernicus (Copernicus,

2019). Foram identificados os locais de colocação dos módulos sensorizados, sendo estes dados obtidos através de uma metodologia que respeite as características da Área de Estudo de acordo com a faixa de módulos de sensores (Brito , Pereira, Lima , Castro, & Valente, 2020).

Atendendo ao facto de que os sensores se encontram aplicados a vários km's de distância da unidade de controlo, existe a necessidade de transmitir os dados provenientes dos sensores constantemente, necessitando assim de um sistema de comunicação eficiente. Para tal foi adotado o protocolo LoRa (Long Range) de maneira a garantir a conexão entre os sensores e a central. O sistema de comunicação LoRa é uma tecnologia de rede de banda larga, é uma rede de baixo consumo permitindo transmissões de longo alcance, cerca de 10 km em meio rural, permitindo assim a comunicação dos dados dos sensores (Brito , Pereira, Lima , Castro, & Valente, 2020).

O sistema anteriormente descrito terá o seguinte aspeto:

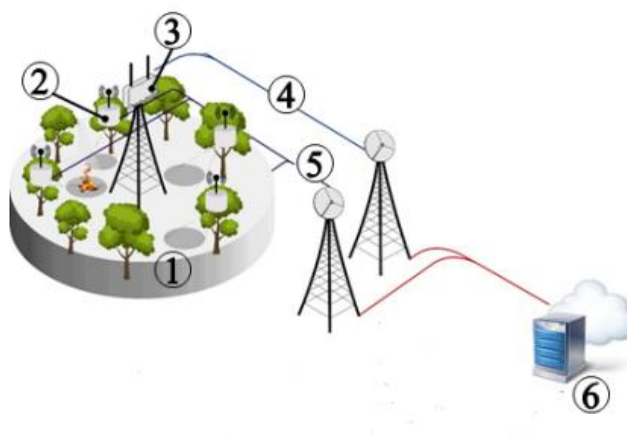


Figura 16 - Protótipo da rede implementada (Brito , Pereira, Lima , Castro, & Valente, 2020)

Legenda:

1. Area em estudo;
2. Modulo de recolha de dados;
3. LoRaWAN Gateway, o dispositivo que recebe os dados de todos os sensores implementados na área em estudo;
4. Comunicação através da internet, que nos facilita a transmissão em direto de dados;
5. Transmissão LoRa;
6. Servidor onde será aplicado o algoritmo de inteligência artificial que nos vai possibilitar a criação dos alertas;

Capítulo 4

Caracterização dos dados

Os dados utilizados para este estudo foram obtidos em ambiente de laboratório através de um ensaio com cinco sensores de chama disponíveis de forma a observar a diferença entre cada um. Com isto pretendia-se escolher o melhor sensor possível, recorrendo para tal ao cálculo da média e do desvio padrão de todos os dados medidos no ensaio.

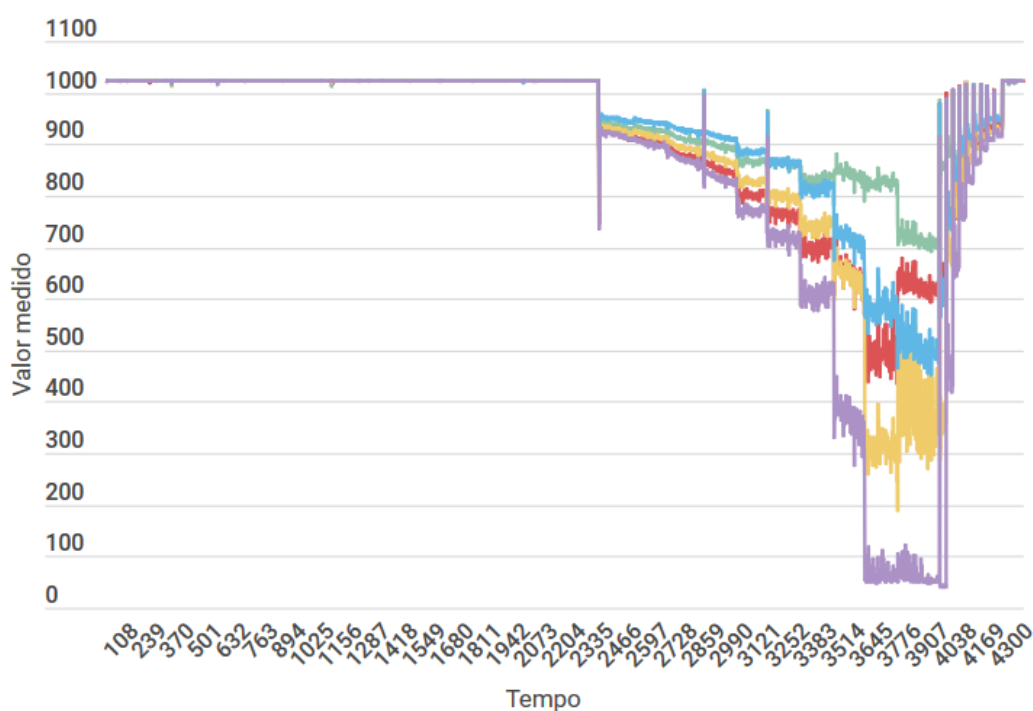


Figura 17 - Dados obtidos do ensaio dos 5 sensores

Na Figura 17 é possível observar os valores obtidos do ensaio realizado com os cinco sensores. Durante a primeira fase do ensaio mantiveram-se os sensores direcionados apenas à luz do dia de forma a observar a oscilação obtida através da incidência solar. No instante 2308 foi acendida uma chama, com uma distância de cerca de 100cm do sensor,

seguidamente em períodos idênticos foi se aproximando de 10 em 10 cm a chama do sensor. Com este procedimento foi possível observar o comportamento do sensor podendo verificar-se a oscilação dos valores, acontecimento importante para o caso, facilitando assim a deteção de um foco de incêndio. Simultaneamente estudado para cada sensor foi também a média e o desvio padrão dos resultados. Foram analisados os resultados obtidos pelos sensores, considerando o protocolo descrito anteriormente, sendo os resultados apresentados na Tabela 6.

Tabela 6 - Valores de média e desvio padrão dos sensores com uma distância de 100cm

Distância de 100cm					
	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5
Média	934.72	916.07	925.43	946.05	914.55
Desvio Padrão	15.44	18.22	16.40	13.20	18.82

Como objetivo principal temos a tarefa de escolher então o sensor ideal para as nossas necessidades, ou seja, encontrar o balanço entre a média e o desvio padrão. Para o problema em si necessitamos de um sensor onde seja bem notório a oscilação de valores para detetar o mais pequeno foco de ignição. O ideal seria então uma média relativamente baixa e um desvio padrão baixo o qual indica que os valores obtidos não oscilam muito, esse facto será essencial de modo a minimizar o impacto da luz solar que incide na lente, pois esta luz nunca vai incidir na mesma direção nem com a mesma intensidade devido à vegetação da qual o sensor está rodeado.

Tabela 7 - Valores de média e desvio padrão dos sensores com uma distância de 10cm

Distância de 10cm					
	Sensor 1	Sensor 2	Sensor 3	Sensor 4	Sensor 5
Média	717.58	629.06	399.00	513.30	64.37
Desvio Padrão	12.13	16.29	66.58	25.59	17.52

Na tabela 7 a chama está a uma distância de 10 cm do sensor. Nesta situação é possível observar uma enorme discrepância que existe entre os cinco sensores. Deste resultado foi então escolhido o sensor número 5 como o melhor para o caso, é visível que existem desvios padrões inferiores como é o caso do sensor 1 e 2 mas devido as médias altíssimas perdem todo o interesse.

4.1 Caracterização da base de dados

A base de dados resume-se a um grupo de dados obtidos em ambiente de laboratório, onde foi inicialmente realizado um ensaio sem focos de ignição, de modo a obter uma situação ideal. Na segunda fase foi realizado durante um determinado período um ensaio com um foco de ignição de forma a ter dados de situação real. Para ambos os ensaios foram utilizados o mesmo número de dados (4893). Na situação da base de dados sem informação de ignição florestal presente a oscilação de valores praticamente não é notória, o valor máximo medido é de 1023 e o valor mínimo é de 1020. Quanto à situação da base de dados com informação de ignição florestal foi feito um gráfico para ilustrar as oscilações dos valores medidos representado na Figura 18.

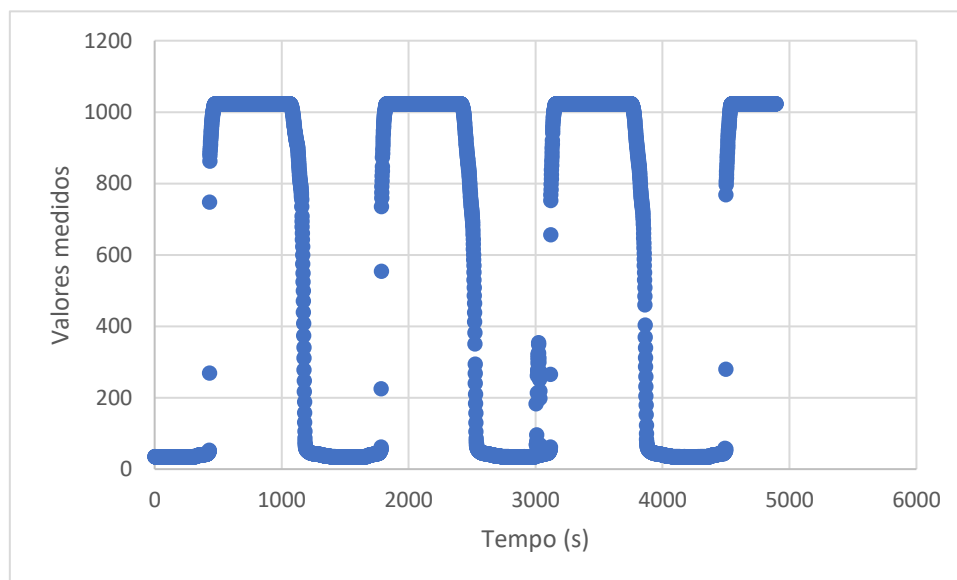


Figura 18 - Valores medidos da base de dados com informação de ignição florestal

4.2 Caracterização da base de teste

O conjunto de dados de teste são dados lidos pelo sensor em ambiente controlado e os quais serão usados para testar o algoritmo.

Para tal feito, como já mencionado anteriormente, foram realizados ensaios no laboratório recorrendo a situações de ignição presente e ignição ausente de forma a simular uma leitura normal de um sensor. Com este procedimento pretendeu-se obter valores de situações semelhantes a uma ignição florestal. Estes valores são analisados pelo algoritmo e comparados com o conjunto de dados de treino. Caso haja uma discrepância será disparado um alarme. Como já mencionado anteriormente para ambos os ensaios foram utilizados o mesmo número de dados (4893).

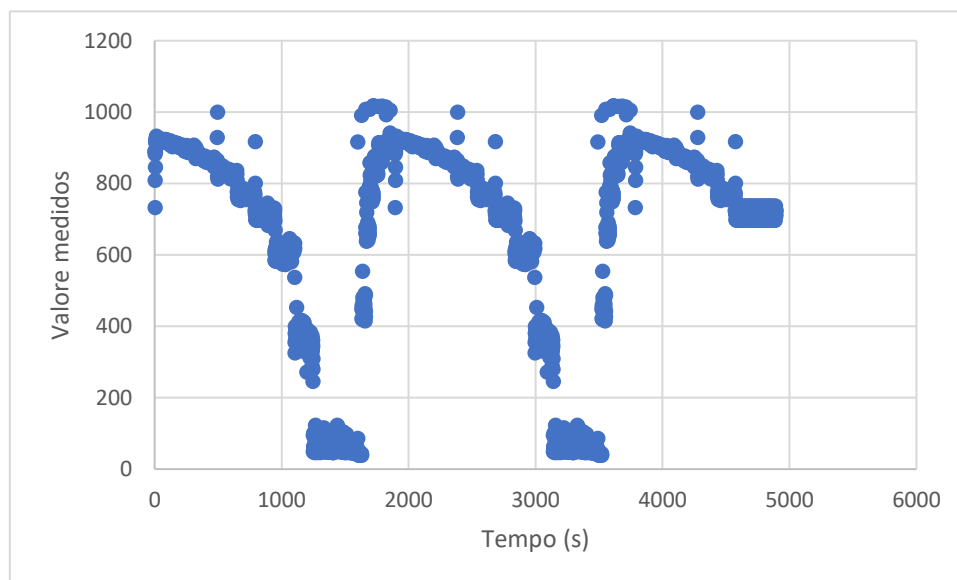


Figura 19 - Valores medidos da base de teste com informação de ignição florestal

É de fácil observação na Figura 19 que o sensor obtém valores muito diferentes dos obtido sem presença de ignições, a queda dos valores lidos é significativa, decrescendo de um máximo de 1018 para um mínimo de 39. É neste momento que se pretende que, em comparação com a base de dados, se obtenha um alarme.

Capítulo 5

Resultados obtidos e sua discussão

Neste capítulo vamos então demonstrar os resultados obtidos nos diversos testes recorrendo para tal aos dados obtidos em ambiente de laboratório como mencionado anteriormente. Para tal tem-se dois conjuntos de valores, nomeadamente a base de dados e a base de teste. Na base de dados vamos ter reunidos os valores com os quais pretendemos demonstrar ao algoritmo uma situação em que não existe incêndio, já na base de teste vamos ter o conjunto de valores onde está compreendido um foco de ignição florestal. Nos gráficos obtidos pelo *MatLab* obtemos um erro que é designado por *RMS Error (Root-Mean-Square-Deviation)*, ou seja, o erro é dado através da raiz quadrada do erro médio.

5.1 Teste com valores de base de dados e base de teste iguais

Na primeira fase de ensaios o código foi testado com os valores obtidos pelos ensaios laboratoriais anteriormente mencionados e realizados. Nesta primeira etapa, para os valores fornecidos, o algoritmo encontrou o parâmetro com mais peso no estudo, devolvendo essa resposta através da Figura 20:

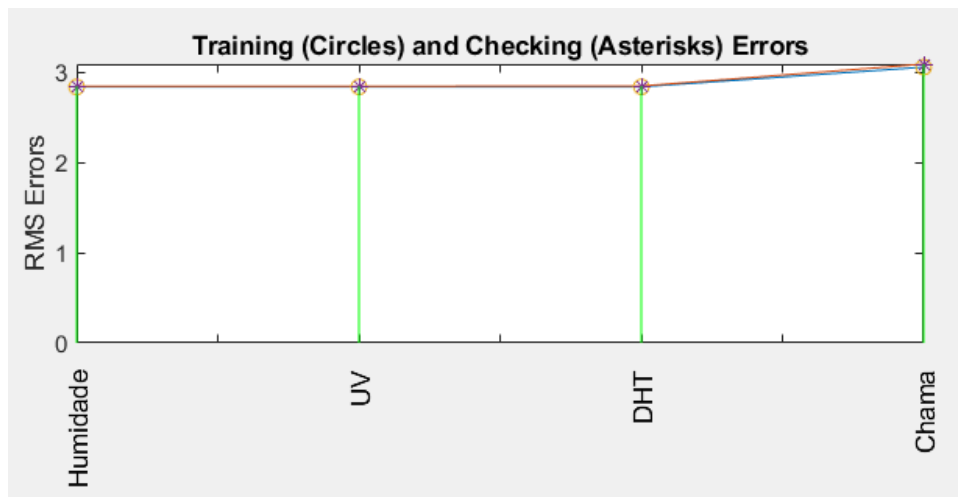


Figura 20 - Importância de cada Sensor

Daqui podemos então retirar que o parâmetro mais importante é o de chama, visto que é o que em todas as combinações apresenta um erro mais elevado.

Posto isto o código vai realizar então todas as combinações possíveis em pares dos diferentes sensores dando assim a combinação mais fiável para uma situação de incêndio:

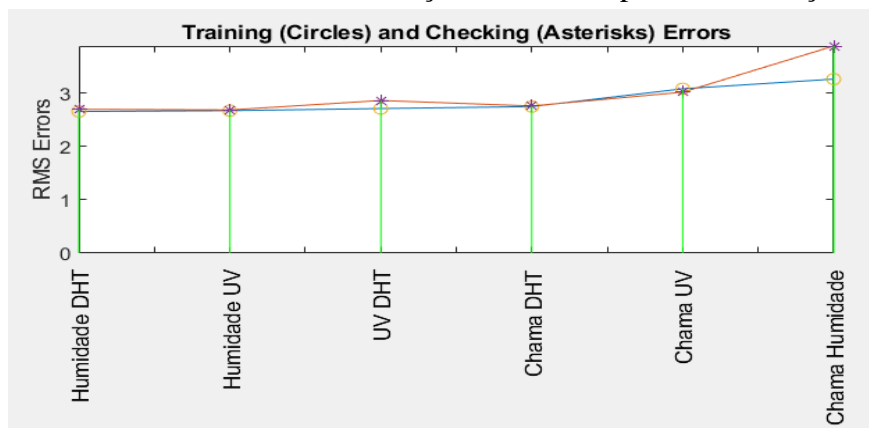


Figura 21 - Importância de cada combinação de Sensores

Observando este gráfico facilmente chegamos à conclusão de que a combinação mais importante será então o sensor de Chama e de Humidade. Já em contrapartida a combinação Humidade/DHT e Humidade/UV são as que devemos ter menos em consideração. O resultado da comparação destas combinações é feita através da análise do gráfico da Figura 21, sendo que a importância de cada sensor vai afetar a combinação no geral.

Em seguida vai ser analisado o erro em função da base teste e da verificação, recorrendo para tal à base de dados, o qual vai ser apresentado na Figura 22:

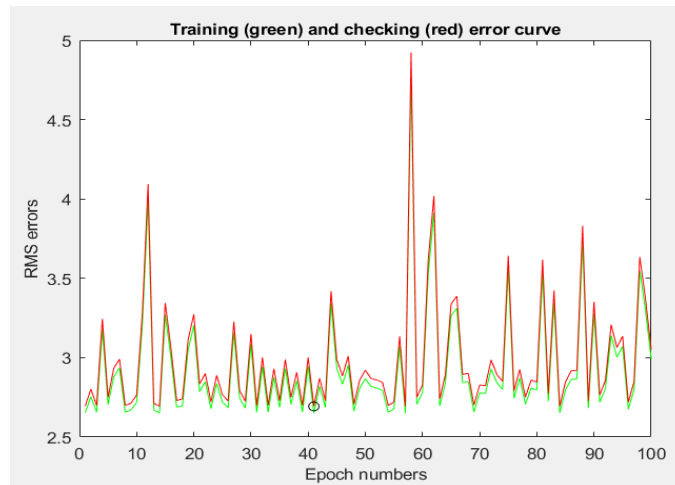


Figura 22 - Erro do esperado em relação ao obtido

Posto esta análise é nos então possível obter um gráfico de superfície o qual demonstra o risco de incêndio em função de dois parâmetros, representado na Figura 23:

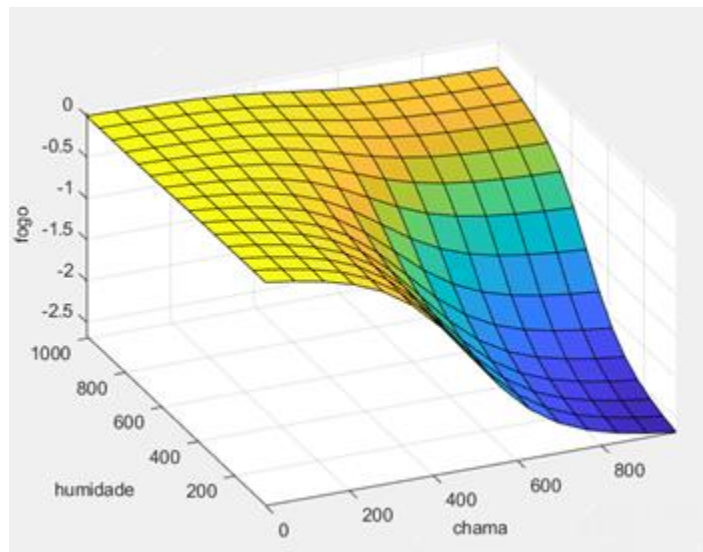


Figura 23 - Gráfico de superfície do risco de incêndio

Através da análise da Figura 23 seria então possível a deteção de uma potencial ignição florestal, pois caso ocorresse a presença de uma ignição florestal esta mesma seria exposta em forma de algo anormal e não linear como se verifica. Os valores representados para a chama e para a humidade são os obtidos pelos sensores, já no eixo do “fogo”, é uma escala calculada por uma função predefinida do *MatLab* fornecendo esta uma indicação sobre uma possível ignição.

5.2 Teste com valores de base de dados e base de teste diferentes

Na parte anterior do processo foram efetuados os diversos ensaios recorrendo sempre a uma base de dados e uma base de teste idêntica, nesta parte vamos fornecer então ao código uma base de dados para aprendizagem onde os parâmetros estão em condições ideais, isto é, não há presença de ignições florestais e em simultâneo uma base de teste para realizar a validação onde existe presença de uma ou várias ignições florestais. Com este procedimento pretendemos observar se o código é ou não capaz de detetar pequenas diferenças nos valores lidos pelos sensores.

Como no estudo anterior inicialmente começou se então por descobrir qual o parâmetro individual com mais importância para o resultado, obtendo o seguinte gráfico:

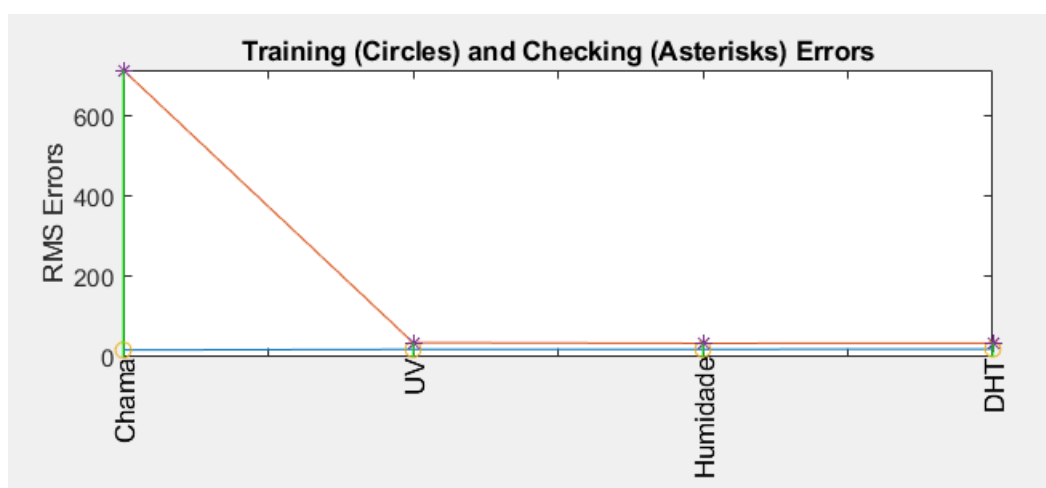


Figura 24 - Importância de cada sensor em ambiente de incêndio

Na Figura 24 ao contrário do sucedido no estudo anterior é de fácil observação que o sensor com mais erro (*RMS*), ou seja, o que tem mais importância é o de chama. A análise da Figura 24 leva de imediato à conclusão que existe uma diferença bastante notória nos valores lidos pelo sensor.

Posto isto comparou-se as combinações de sensores entre si para saber qual a mais critica.

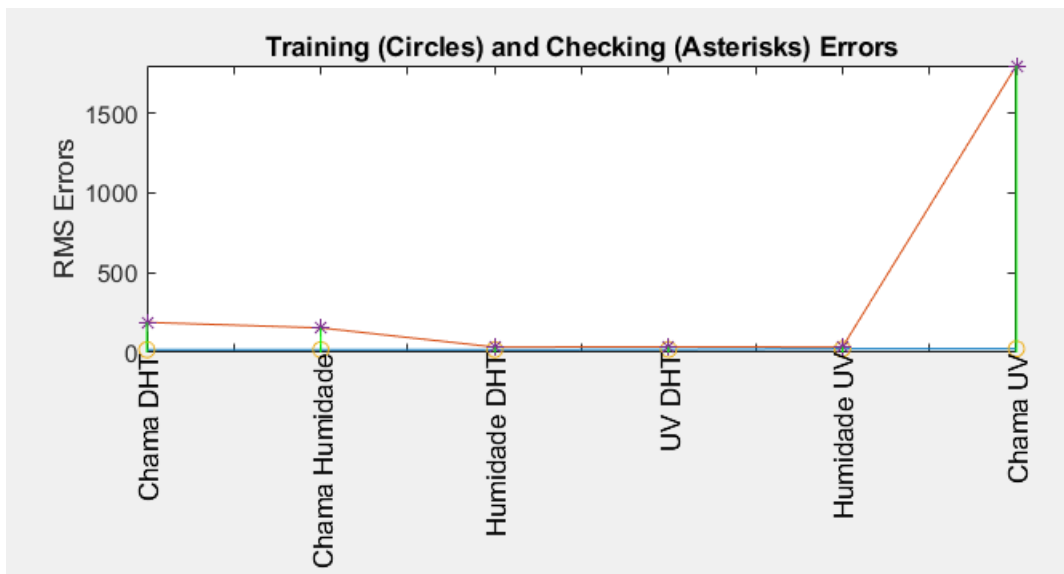


Figura 25 - Importância de cada combinação de Sensores em ambiente de incêndio

Com a análise da Figura 25 ainda se torna mais notório o enorme erro (*RMS*) que existe na combinação do sensor de chama com o UV, ou seja, é esta a combinação que está a afetar mais o código neste momento. Ainda assim é possível observar que todas as combinações onde está presente o fator “chama” tem um erro superior a qualquer uma das outras combinações, nas quais se aproxima bastante de 0.

Posto isto, vamos ver a Figura 26, a qual representa o erro geral que o código calculou em relação aos valores estudados e lidos no momento do incêndio obtendo um resultado ainda mais intenso.

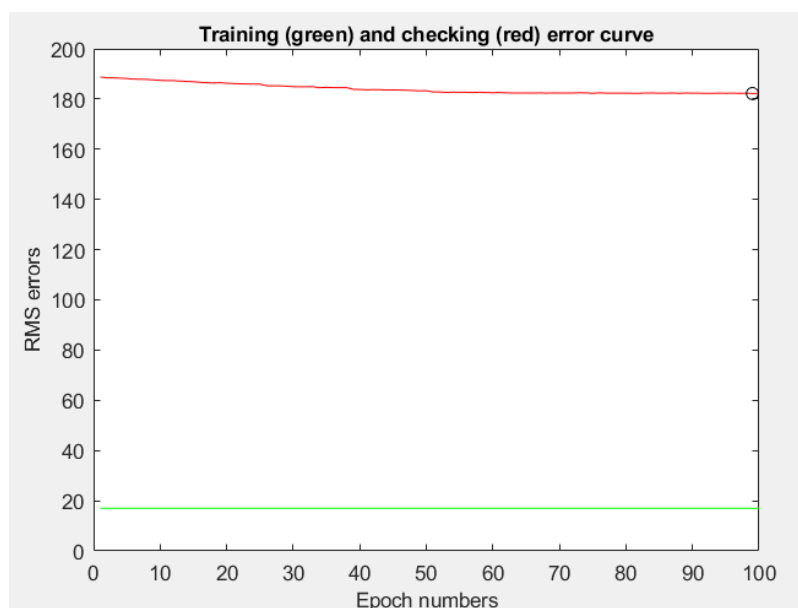


Figura 26 - Erro do esperado em relação ao obtido em ambiente de incêndio

Daqui retiramos logo a conclusão de que os valores não correspondem em nada com o que é normal ou esperado. No estudo realizado anteriormente onde os valores eram idênticos era possível ver que a curva obtida era exatamente igual tendo apenas um ligeiro desvio geral.

Em seguida vamos expor os últimos dois gráficos obtidos, o de superfície que corresponde ao risco de incêndio e por fim o gráfico da comparação dos valores obtidos.

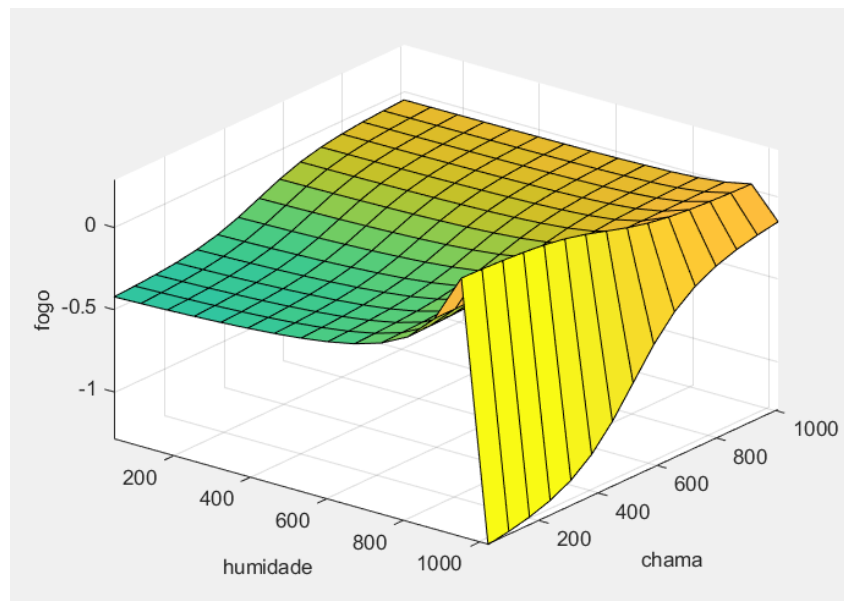


Figura 27 - Gráfico de superfície do risco de incêndio

Na Figura 27 como foi referido no estudo anterior caso se verificasse uma situação de ignição florestal seria de esperar algo de anormal no gráfico, como podemos observar essa situação verifica-se.

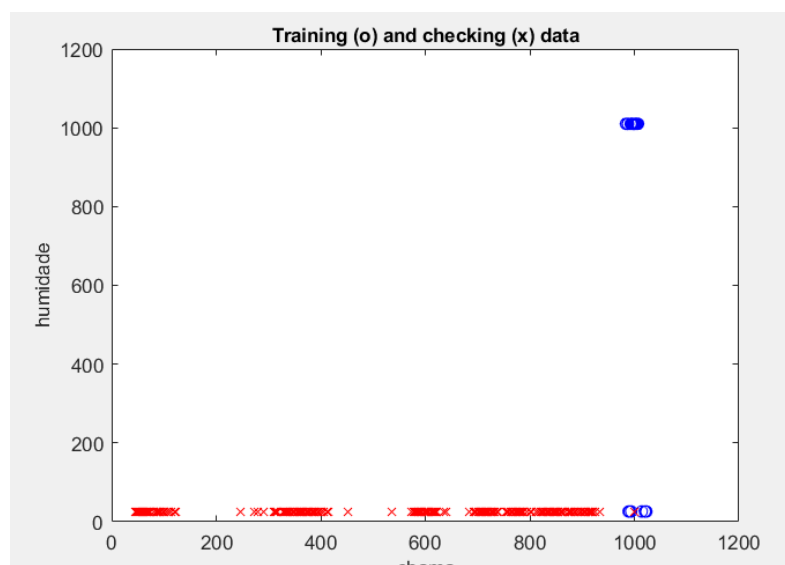


Figura 28 - Comparação da aprendizagem e da validação

Já na Figura 28 não podendo ser deixado de parte, acaba por ser o mais importante, este mesmo é um gráfico de dispersão onde estão assinalados os valores da base de dados e os valores da base de teste, como é possível observar para uma ignição florestal não se verificar seria de esperar que os valores de ambos se encontrassem na mesma zona do gráfico, daí seria possível tirar a conclusão de que não há ignição. Neste caso é claramente visível que estamos perante uma situação de incêndio.

5.3 Teste com diversas combinações de valores de base de dados e base de teste

Neste subcapítulo vamos apresentar os resultados de alguns ensaios feitos nos quais foram alterados constantemente os valores da base de dados e da base de teste, isto é, testou se o funcionamento do código recorrendo a diversas combinações possíveis entre ambas. Vamos demonstrar para cada ensaio o gráfico do erro e da dispersão obtidos, visto serem estes os mais interessantes para analisar.

Foram realizados diversos testes para analisar o comportamento do método *Fuzzy Logic*, para simplificar a notação das diversas bases de dados e de treino usadas optou-se então por designar cada uma da seguinte forma, **BD** designa “Base de Dados” e **BE** “Base de Teste”, variando estas ainda consoante a informação de ignição presente, sendo que **CIIF** significa “Com Informação de Ignição Florestal”, **SIIF** designa “Sem Informação de Ignição Florestal” e por último temos **CSIIF** a qual significa “Com e Sem Informação de Ignição Florestal”.

5.3.1 BD - CSIIF

Nesta secção pretende-se analisar o comportamento do método quando se tem uma base de dados constante, a qual contém dados com e sem informação de ignição florestal, alterando somente a base de teste segundo as hipóteses possíveis e anteriormente apresentadas.

- BT-CSIIF

Nesta secção será usada uma base de dados com e sem informação de ignição florestal com uma base de teste com e sem informação de ignição florestal, sendo possível observar os resultados na Figura 29.

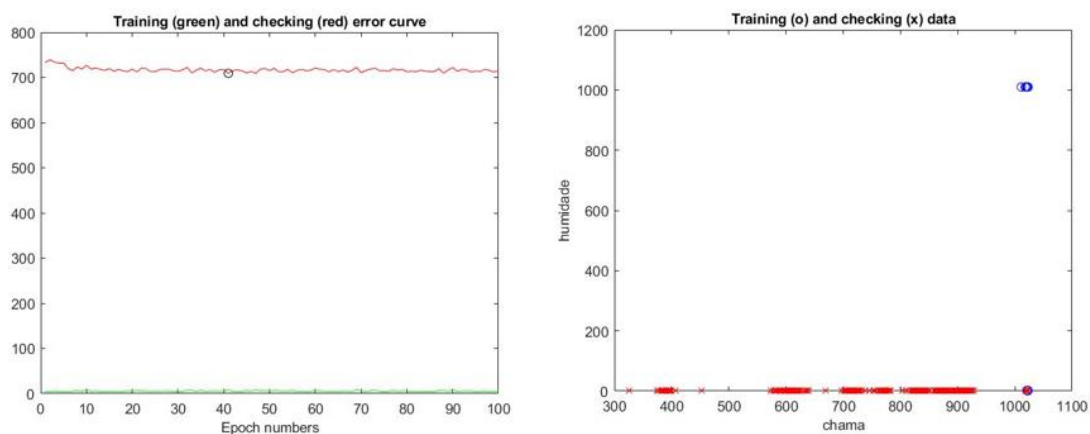


Figura 29 - Gráficos do erro e dispersão respetivamente

Nesta situação é possível observar que o erro *RMS* é bastante notório (eixo y), ou seja, daqui podemos retirar a conclusão de que existe uma ignição nos dados analisados, caso contrário o erro deveria tender para valores na ordem dos 50. Já no gráfico de dispersão é possível observar que os valores da base de dados e da base de teste não coincidem.

- BT-CIIF

Nesta secção será usada uma base de dados com e sem informação de ignição florestal com uma base de teste com informação de ignição florestal, sendo possível observar os resultados na Figura 30.

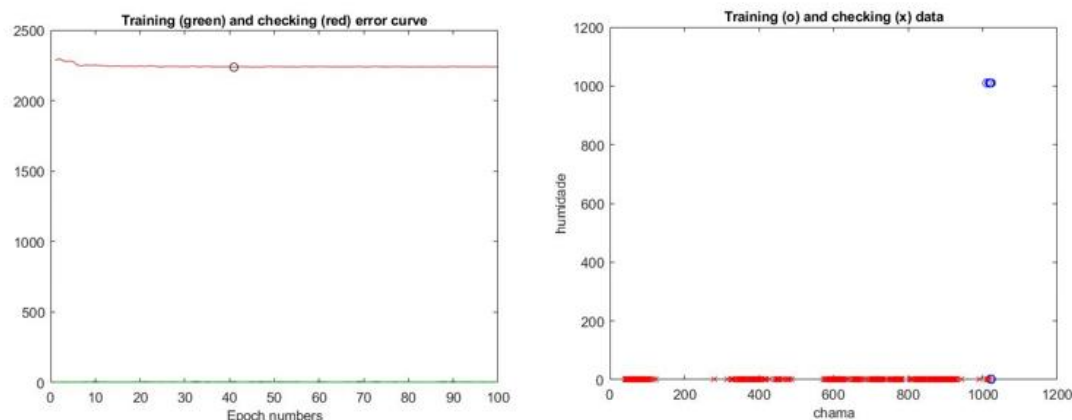


Figura 30 - Gráficos do erro e dispersão respetivamente

Neste ensaio podemos imediatamente observar um erro ainda maior do que anteriormente visto no eixo do y, neste caso estamos com um *RMS* de cerca de 2300 em relação ao 750 do ensaio anterior, o que é perfeitamente normal visto que temos uma base de teste a funcionar somente com valores de ignições florestais, já no gráfico de dispersão não é notória uma grande diferença.

- BT-SIIF

Nesta secção será usada uma base de dados com e sem informação de ignição florestal com uma base de teste sem informação de ignição florestal, sendo possível observar os resultados na Figura 31:

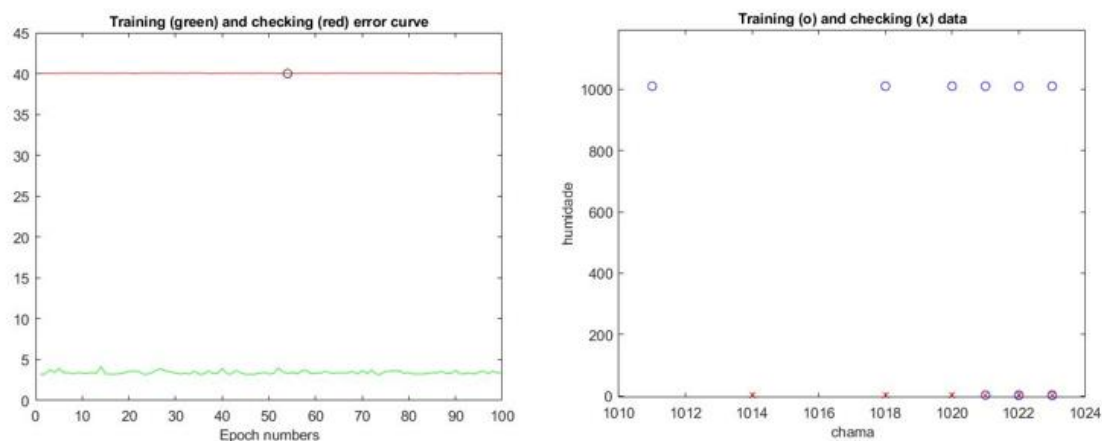


Figura 31 - Gráficos do erro e dispersão respetivamente

Nesta situação de ter uma base de dados com e sem informação de ignição florestal e uma base de teste sem informação de ignição florestal conseguimos facilmente observar que o erro é mínimo, é possível verificar que não existe qualquer informação de ignição. Já no gráfico de dispersão também observamos que existem diversos valores coincidentes, o que nos mostra que não existe uma grande oscilação entre os mesmos.

5.3.2 BD-SIIF

Como no subcapítulo anterior vamos também neste manter uma base de dados fixa, mas sem informação de ignição florestal, ou seja, valores considerados normais, alterando novamente a base de teste em todas as combinações possíveis.

- BT-CSIIF

Nesta secção será usada uma base de dados sem informação de ignição florestal com uma base de teste com e sem informação de ignição florestal, sendo possível observar os resultados na Figura 32:

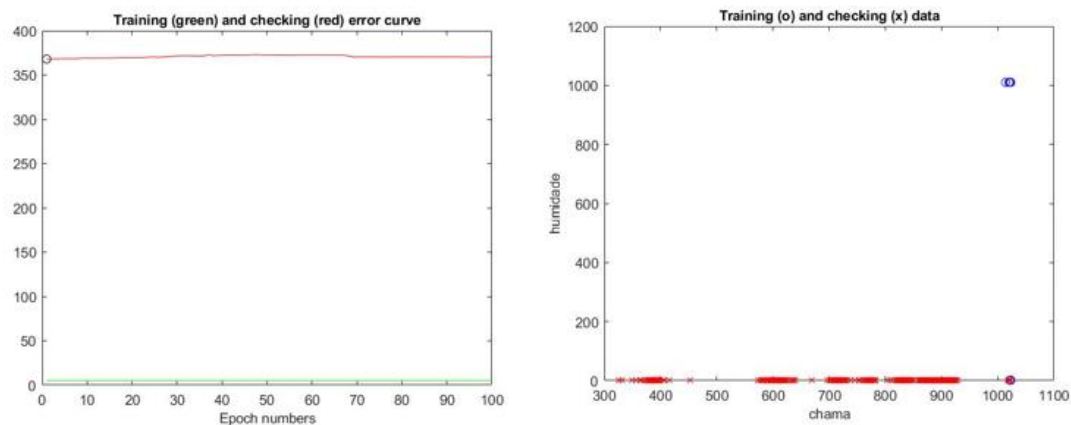


Figura 32 - Gráficos do erro e dispersão respetivamente

Nesta situação em concreto podemos observar que o erro é alto, mas ligeiramente mais baixo que nas situações anteriores, o que era de esperar visto que temos uma base de dados sem informação de ignição florestal. Quando comparamos a mesma com uma base de teste com e sem informação de ignição florestal o erro vai aumentar, sendo mais notório quando se comparar com uma base de teste com informação de ignição florestal como vamos observar no próximo exemplo.

- BT-CIIF

Nesta secção será usada uma base de dados sem informação de ignição florestal com uma base de teste com informação de ignição florestal, sendo possível observar os resultados na Figura 33:

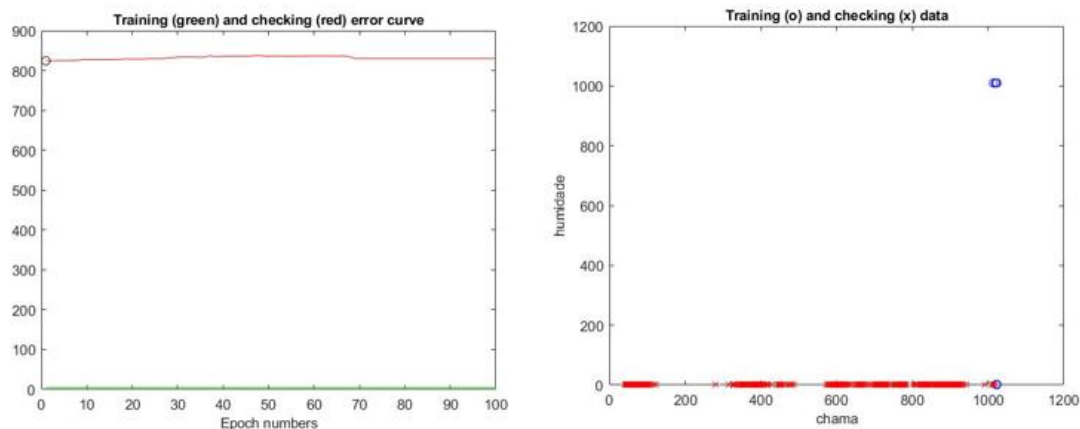


Figura 33 - Gráficos do erro e dispersão respetivamente

Como já referido anteriormente nesta situação o erro é bastante mais elevado, como era de esperar, pois como estamos perante uma base de dados sem informação de ignição florestal e uma base de teste com informação de ignição florestal a discrepância de valores é enorme.

- BT-SIIF

Nesta secção será usada uma base de dados sem informação de ignição florestal com uma base de teste sem informação de ignição florestal, sendo possível observar os resultados na Figura 34:

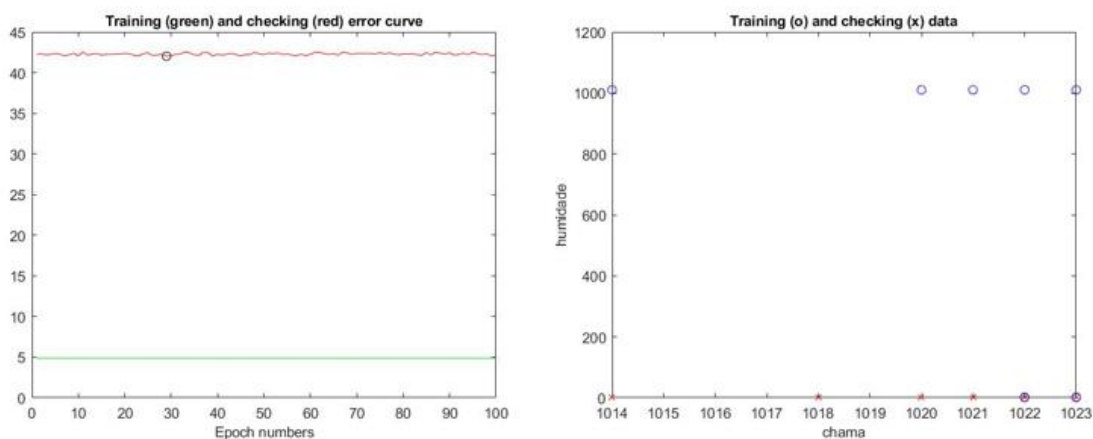


Figura 34 - Gráficos do erro e dispersão respetivamente

Quando estamos perante a situação de valores idênticos é possível observar que o erro é mínimo, fenómeno já visto anteriormente quando existe semelhança de valores, já no gráfico de dispersão também se constata o mesmo facto, os valores encontram-se todos em posições de chama idênticas, variando apenas e só a humidade.

5.3.3 BD-CIIF

Nesta última fase vamos então demonstrar os resultados obtidos recorrendo a uma base de dados com informação de ignição florestal, com isto pretendemos saber o resultado de todas as combinações possíveis tentando assim obter resultados mais viáveis.

- BT-CIIF

Nesta secção será usada uma base de dados com informação de ignição florestal com uma base de teste com e sem informação de ignição florestal, sendo possível observar os resultados na Figura 35:

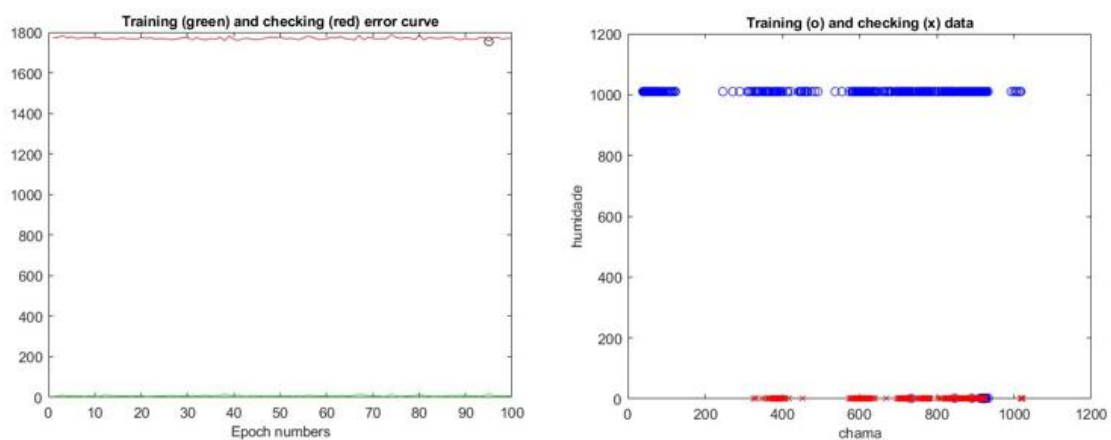


Figura 35 - Gráficos do erro e dispersão respetivamente

Esta situação é bastante diferente do estudo oposto, neste caso o erro é bastante mais elevado e a dispersão dos valores é mais semelhante, estando assim presente novamente o facto das humidades diferentes.

- BT-CIIF

Nesta secção será usada uma base de dados sem informação de ignição florestal com uma base de teste com informação de ignição florestal, sendo possível observar os resultados na Figura 36:

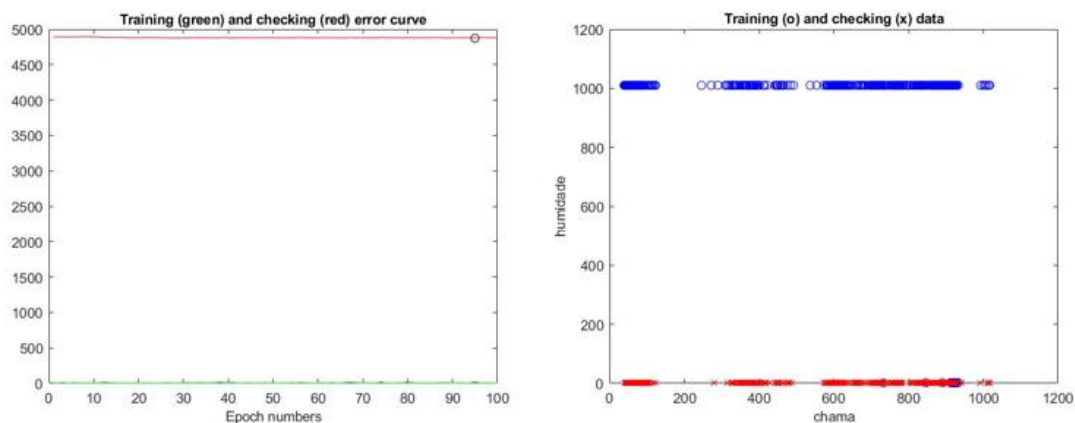


Figura 36 - Gráficos do erro e dispersão respetivamente

Neste caso é possível observar então o gráfico com erro mais elevado de todos os ensaios, uma situação que é devida à base de dados sem informação de ignição florestal e à base de teste com informação de ignição florestal. Analisando o gráfico da dispersão é possível observar que os resultados associados ao sensor de humidade variam cerca de 1000 unidades.

- BT-SIIF

Nesta secção será usada uma base de dados sem informação de ignição florestal com uma base de teste com e sem informação de ignição florestal, sendo possível observar os resultados na Figura 37:

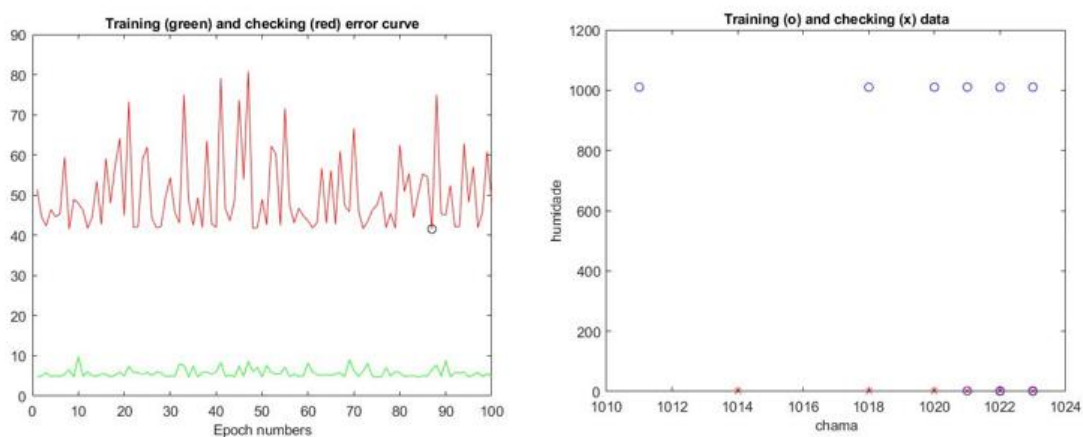


Figura 37 - Gráficos do erro e dispersão respetivamente

O último ensaio não deixa, porém, de ser o mais “estranho” por assim dizer, quando temos uma situação em que os valores são quase coincidentes, mas com alguma variação as linhas de erro da base de dados e da base de teste não são lineares, tendendo então a formar imensos picos, obtendo assim um intervalo de erro de cerca de 35. Já na dispersão observamos que os valores estão aproximadamente iguais, tendo uma escala de incêndio iniciada em 1010 e com final de 1024 limite máximo do sensor.

5.4 Ensaio realizado com dados reais

Para concluir o nosso estudo foi realizado um ensaio recorrendo a dados obtidos num ambiente real, isto é, dados provenientes de sensores colocados ao ar livre no campus do IPB. Os sensores estiveram em funcionamento no dia 03/09/2020 durante 24h, sendo assim possível analisar o comportamento dos valores tanto durante o dia como durante a noite também. Resta mencionar que nenhum destes sensores captou dados provenientes de um foco de incêndio, estes dados serão então comparados entre si de forma a analisar o efeito da incidência da luz solar na lente do sensor de chama.

Na Figura 38 é possível observar o erro obtido do ensaio.



Figura 38 - Gráfico do erro com dados reais

Com a observação detalhada da Figura 38 é possível observar o sucessivo decréscimo do erro, este mesmo é devido ao pôr do sol, ou seja, quando entramos num ambiente mais escuro, o sensor reduz em muito o erro. Nesse ambiente é possível detetar os mais

pequenos focos de incêndio, já durante o dia como o erro é relativamente alto isto torna-se mais complicado pois não vai existir uma oscilação tão notória.

Na Figura 39 podemos observar o gráfico de dispersão do ensaio realizado.



Figura 39 - Gráfico dispersão com dados reais

Com a observação da Figura 39 conseguimos observar a diversidade de valores existente, ou seja, temos situações em que os valores de chama estão em coerência como temos situação em que acontece exatamente o contrário, o mesmo se verifica no caso da humidade. O que é facilmente observável é o facto de a escala da humidade estar compreendida entre 10 e 90, fenómeno devido ao sensor utilizado ler a humidade relativa.

5.5 Discussão de Resultados

Deste estudo foi nos possível concluir que o algoritmo funciona em perfeitas condições, deteta qualquer anomalia por mínima que seja, o qual era o objetivo inicial. O único problema detetado é como já anteriormente mencionado a incidência solar sobre a lente do sensor. Este fenómeno o qual origina falsos valores pode ser corrigido com o simples posicionamento do módulo sensorizado, é necessário colocar o mesmo num local sombrio de forma a impedir a incidência direta do sol. Caso isto seja realizado com sucesso os valores de erro mais elevados são sem dúvida de focos de incêndio.

Em relação aos resultados obtidos em ambiente de laboratório foram bastante úteis para analisar o efeito de diversas combinações de bases de dados e bases de teste, com isto foi

possível observar as variações de erro. Já os restantes resultados onde está contida a humidade, não existe fiabilidade nos mesmos, visto que o sensor está num ambiente fechado e não ao ar livre.

Conclusões e Trabalhos Futuros

Neste trabalho, o sistema de deteção de ignições florestais utilizando *Fuzzy Logic* é um sistema bastante fiável em determinadas condições, pois como se pode constatar a incidência solar torna-se um grande problema para a leitura de dados. Sendo possível observar que quando existe incidência solar sobre os sensores de chama os valores refletem essa observação, emitindo valores similares aquando a presença de chama, daí como anteriormente mencionado que a sua utilização deve única e exclusivamente ser feita em matas densas onde não exista incidência solar constante, zonas sombrias. Em parceria com os Bombeiros Voluntários de Bragança também foi transmitida a informação de que essa seria a principal utilização e também a de mais interesse para o sistema. O principal problema na deteção de ignições em matas muito densas é a dificuldade em a coluna de fumo ser visível, por norma quando esta ascende e o posto de vigia mais próximo a deteta já se encontra num incêndio ativo. Este sistema pode em determinadas situações reduzir o tempo de deteção da ignição por vários minutos, minutos os quais podem ser preciosos e definir se estamos perante um incêndio de grandes dimensões ou de um foco de fácil extinção. Como sugestão de aperfeiçoamento do sistema resta somente a questão solar, caso seja possível obter valores ótimos durante a incidência solar estamos perante um sistema com uma efetividade acima do normal, dispensando grande parte do auxílio humano e facilitando em muito a vida dos bombeiros e de toda a biodiversidade existente no local incidente. Resta então somente deixar a sugestão de um trabalho futuro que seria o desenvolvimento de uma lente a qual seria capaz de reduzir o efeito da incidência solar, mas em simultâneo não obstruir a eficácia do sensor em detetar uma ignição florestal.

Bibliografia

- A. Sowah, R., R. Ofoli, A., N. Krakani, S., & Fiawoo, S. Y. (2017). Hardware Design and Web-Based Communication Modules of a Real-Time Multisensor Fire Detection and Notification System Using *Fuzzy Logic*.
- Alves, J. (2018). Detecção de incêndios florestais com recurso a deep learning e visão computacional.
- Brito, T., Pereira, A. I., Lima, J., Castro, J. P., & Valente, A. (2020). Optimal Sensors Positioning to Detect Forest Fire Ignitions. (SCITEPRESS, Ed.) pp. 411-418.
- Da Silva, I. D. (2017). *Sistema de Alerta de Incêndio Florestal*. Campinas.
- dos Santos, H. G. (2018). *Comparação da performance de algoritmos de machine learning para análise preditiva em saúde pública e medicina*.
- G. Lewis, T., & J. Denning, P. (2019). Learning Machine Learning.
- Hastie, T., Tibshirani, R., & Friedman, J. (2008). The Elements of Statistical Learning.
- Jimenez, S. G., Jurio, A., Pagola, M., De Miguel, L., Barrenechea, E., & Bustince, H. (28 de Setembro de 2016). Forest fire detection: A *Fuzzy* system approach based on overlap indices. *Applied Soft Computing*.
- Khamis, A., Ghani, M., Gan, C., Aras, M., Khamis, M., Sutikno, T., & Zanariah, J. (Setembro de 2017). *Fuzzy Logic* Implementation with MATLAB for PV-Wind Hybrid System.
- Lall, S., & Mathibela, B. (2016). The Application of Artificial Neural Networks for Wildfire Risk Prediction.
- Osório, F., & Bittencout, J. (2000). Sistemas Inteligentes baseados em Redes Neurais Artificiais aplicados ao Processamento de Imagens.
- Pereira, D., Mendes, J., & Pereira, M. (2019). Sensors .
- Riano Vargas, K., Alonso-Echeverri, J. I., Gaona-Garcia, P. A., & Montenegro-Marin, C. E. (2018). *Fuzzy* logic-based model for the prevention of structural fires on urban buildings using mobile devices and IoT technologies.
- Sena, I. P. (2020). Identificação de Tendências no Atendimento Médico.
- Soares, J. (2019).
- Soares, J. (2019). Uma breve viagem pela inteligência artificial. p. 34.
- Soto, M. E. (2012). The identification and assessment of areas at risk of forest fire using *Fuzzy* methodology.
- Trilles, S., Juan, P., Díaz, L., Aragón, P., & Huerta, J. (2013). Integration of Environmental Models in Spatial Data Infrastructures: A Use Case in Wildfire Risk Prediction.

Weiqiang, H. (2019). Machine learning applied to pack classification.

Wilkens, Sascha;. (20 de Junho de 2019). Machine learning in risk measurement:
Gaussian process regression for value-at-risk and expected shortfall.

Zhao, J., & K. Bose, B. (2002). Evaluation of membership functions for *Fuzzy Logic*
controlled.