Acoustic Extinguisher Fire Dataset

Diogo Lemos 202003484
Duarte Lopes up202006408
José Sousa up202006141

Specification of the work to be performed

Title: Analysis of the Acoustic Extinguisher Fire Dataset

Introduction
The goal of this project is to implement a machine learning model that can accurately predict whether a fire extinguisher is in use based on acoustic data. By doing this we are able to improve fire safety by detecting fires earlier due to the knowledge gained.

State of the Dataset

The acoustic fire extinguisher dataset comprises recordings of extinguisher use along with a number of other details, like the extinguisher's size and proximity to a sound wave extinguishing system . With the help of this information, we hope to create a machine learning system that can correctly identify whether a fire extinguisher is being used.

In this dataset we have the presence of various parameters such as Size, Fuel, Distance, Decibel, Airflow, Frequency, Status. With this experimental set-up, 17,442 tests were carried out in total with a similar amount of tests regarded as extinguished and non extinguished.

What to Do

In an earlier phase we started to dig deep in the dataset to find some relevant factors that can lead to a breakthrough in the detection of the use of the fire extinguishing acoustic methods. We decided to begin with the help of python libraries such as pandas, SciPy, Scikit-learn and Seaborn to create tables that help focusing on the status(non-extinguished and extinguished). After the initial tables were created we realised that some factors seem very promising such as the less of a distance the more likely there is an extinguished flame and the higher of an airflow we get a much more likely extinction. The next procedure to be had is to prune the tables for outliers and duplicated data. After all the data is treated we will begin to implement the algorithm we found fitting to allow for the flame extinction status prediction.

Data Pre-Processing

Before applying the prediction algorithms, it was necessary to ‘clean up’ the provided dataset. Using the algorithms without filtering unwanted or useless data could negatively affect the accuracy of the results.

To do so, we mainly focused on:

- Confirm the non-existence of blank and missing values
- Remove duplicate values
- Check if there were any non-numerical values, and convert them into variables that we could work with

Having the dataset ready, we then plotted an heatmap to find out which values were most suitable for the data modelling. After that, we were all set to start applying the supervised learning algorithms.

Algorithms Used

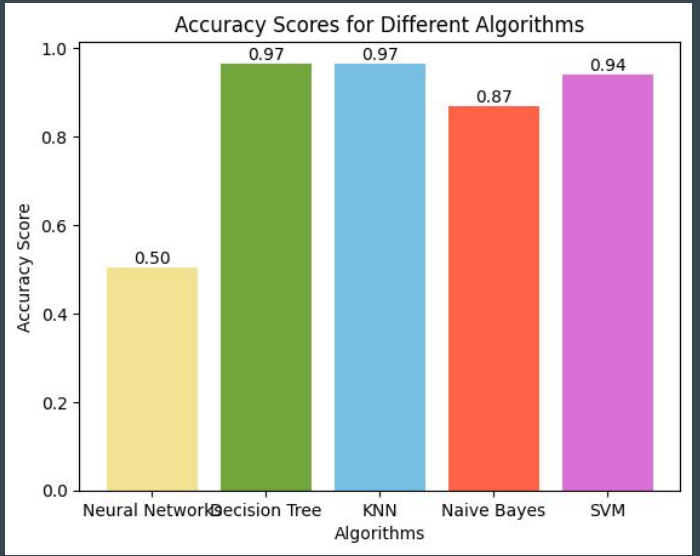
In the second phase of this project we decided that we needed to resort to 5 different algorithms in order to compare them and find the most efficient algorithm for our case of study. The five algorithms we used were the Decision Tree algorithm, K-nearest neighbours algorithm, the Neural Networks, Naive Bayes and finally the SVM.

- Decision Tree: It's a machine learning technique that uses a tree-like structure to make decisions.
- KNN: The K-nearest neighbors algorithm is a machine learning method that classifies data points based on their proximity to other labeled data points in a feature space.
- Neural Networks: Neural networks are a class of machine learning algorithms inspired by the structure and function of biological neural networks, which can learn complex patterns and relationships from data through interconnected layers of artificial neurons.
- Naive Bayes: It is a probabilistic classifier that calculates the probability of a data point belonging to a certain class based on the assumption of independence between features.
- SVM: It is a powerful machine learning method that separates data points into different classes by finding an optimal hyperplane in a high-dimensional feature space.

Results

As we can see from the results we can perceive that three algorithms stand from the rest with two of them reaching the accuracy score of 0.97, the Decision Tree Algorithm and the K-nearest neighbors Algorithm .

The most disappointing result was from the neural network algorithm with a very small 0.50 accuracy score.



Conclusion

The dataset for acoustic fire extinguishers offers a special chance to investigate the possibilities of machine learning for fire detection. Our goal in this research is to create a model that can properly determine from the data if a fire extinguisher is in use and to assess how well different machine learning algorithms perform on this task. With this study we can now determine which are the best algorithms to predict the extinguishing of the fire.