Instituto Superior de Agronomia ULisboa

Master's in Green Data Science 2023-2024

Practical Machine Learning/Aprendizagem Automática Aplicada

Instructor: Manuel Campagnolo

Students: Alícia Eva Lopes Gouveia (27900) & Sofia Margarida Matias Rodrigues (27899)

Forecasting Forest Fires project

Problem Statement

Forest fires play a major role in Portugal's environment, their impact leads to the destruction of forest and other negative effects surrounding biodiversity. Forests provide a vast of direct and indirect benefits to humankind therefore, disasters like forest fires have a major impact. In this *Forecast Forest Fires* project, we aim to develop a method for early fire detection that can help decision makers plan mitigation and extinguishing tactics and help authorities and emergency services to be alert and prepare accordingly.

To achieve this, we plan to use machine learning techniques where the main goal will be the prediction of the number of incidents in a given period and location.

Challenges

Developing this project, we predict some of the challenges we'll be facing are related to the data quality and completeness, variety of factors that fires depend on, temporal and spatial variability, feature selection, seasonality and trends as well as model selection and validation.

Dataset

In the project will be using tabular data, where we will work with a dataset that represents the fire occurrences registered between the years of 2011 and 2020 in Portugal. The dataset that will be used is called "Registos_Incendios_SGIF_2011_2020", that was obtained through the ICNF website (https://www.icnf.pt/florestas/gfr/gfrgestaoinformacao/estatisticas). This dataset is composed of 41 columns and 177129 entries. For the project we will only select the columns with data that are relevant for our goal. We also need to handle missing values and prepare the data accordingly.

We'll also use as resource, studies regarding this field such as:

Predict forest fires using machine learning (neuraldesigner.com)

Predictive modeling of wildfires: A new dataset and machine learning approach - ScienceDirect

Detection of forest fire using deep convolutional neural networks with transfer learning approach - ScienceDirect

Forest fire and smoke detection using deep learning-based learning without forgetting | Fire Ecology | Full Text (springeropen.com)

Method or Algorithm:

We'll explore some existing models, used for problems similiar to our project such as: Time Series Models (ARIMA, SARIMA), Random Forest, Gradient Boosting, or neural networks like LSTM.

We'll implement the model which will fit our goal best.

Evaluation:

The results will be evaluated by performing metrics (Mean Absolute Error, Root Mean Squared Error or Mean Absolute Percentage Error), cross validation and visual inspection. To compare the results, we can use statistical tests, residual analysis, seasonal decomposition or feature importance.