1 2 9 0

UNIVERSIDADE Ð
COIMBRA

Nuno Pires

INTELLIGENT SYSTEM FOR LOCALISING AND MONITORING FOREST FIRES

Dissertation in the context of the Master in Informatics Engineering, specialization in Information Systems, advised by Professor Alberto Cardoso and Professor Jacinto Estima and presented to the Department of Informatics Engineering of the Faculty of Sciences and Technology of the University of Coimbra.

January 2024

DEPARTAMENTO DE
ENGENHARIA INFORMÁTICA
FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE D
COIMBRA

Nuno Pires

Intelligent system for localising and monitoring forest fires

Dissertation in the context of the Master in Informatics Engineering, specialization in Information Systems, advised by Professor Alberto Cardoso and Professor Jacinto Estima and presented to the Department of Informatics Engineering of the Faculty of Sciences and Technology of the University of Coimbra.

January 2024

DEPARTAMENTO DE
ENGENHARIA INFORMÁTICA
FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE
COIMBRA

Nuno Pires

Sistema inteligente para localização e monitorização de incêndios florestais

Dissertação no âmbito do Mestrado em Engenharia Informática, especialização em Sistemas de Informação, orientada pelo Professor Alberto Cardoso e Professor Jacinto Estima e apresentada ao Departamento de Engenharia Informática da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Janeiro 2024

Abstract

Fire can have disastrous consequences. Decision-support systems play a central role in dealing with forest fires. Its early warning capacity and real-world impact help to protect forests, species, and communities from wildfire.

The presented work proposes a system for forecasting and monitoring forest fires using multiple data sources. Data fusion, aggregation, and enhancement techniques are also mentioned.

The main purpose of the system is to provide important information for emergency decision-making, such as the geolocation, severity, and temporal evolution of a wildfire. It will employ statistical and machine learning methodologies to predict and determine fire occurrence, susceptibility, and risk.

Finally, the system, with the help of data visualisation tools, will show findings and insights.

The document also presents current approaches and obstacles to forest fire prediction, as well as the suggested methodology and analysis of risk.

Keywords

Decision support system, Fire management, Fire forecasting, Machine learning, Spatial and temporal prediction

Resumo

Os incêndios podem ter consequências desastrosas. Os sistemas de apoio à decisão desempenham um papel central na luta contra os incêndios florestais. As suas capacidades de alerta e o seu impacto no mundo real ajudam a proteger as florestas, as espécies e as comunidades.

O trabalho apresentado propõe um sistema de previsão e monitorização de incêndios florestais que utiliza fontes diversas de dados. Onde são utilizadas técnicas de fusão, agregação e melhoramento de dados.

O principal objetivo do sistema é fornecer informações importantes para a tomada de decisões de emergência, tais como a geolocalização, a gravidade e a evolução temporal de um incêndio florestal. O sistema empregará metodologias estatísticas e de aprendizagem automática para prever e determinar a ocorrência, a suscetibilidade e o risco de incêndio.

Finalmente, com a ajuda de ferramentas de visualização de dados, o sistema será capaz de apresentar informações e resultados.

No documento também são analisadas as abordagens actuais e os obstáculos à previsão de incêndios florestais, bem como a metodologia sugerida e a análise de risco.

Palavras-Chave

Sistema de apoio à decisão, Gestão de incêndios, Previsão de incêndios, Aprendizagem automática, Previsão espacial e temporal

Contents

List of Figures

List of Tables

Chapter 1

Experiences

1.1 Samples

1.2 Uncovering important variables

1.2.1 Feature importance with RandomForestClassifier

The following experiments were designed using the RandomForestClassifier algorithm (Pedregosa et al., 2011). The goal was to obtain important features that would be able to explain the occurrence of wildfires. These tests were conducted on the three samples previously mentioned.

Firstly, for each sample, a wildfire occurrence was set 3 hours before the wildfire occurrence and 1 hour after, and then the most important variables were calculated. Their complete hourly records were used. With a random state of 445621151 for RandomForestClassifier, the tables 1.1 and 1.2 show the results for each year.

Table 1.1: Variable importance for 2015 and 2019

2015		2019	
Variable	Importance	Variable	Importance
soil_moisture_100_to_255cm	0.122	wind_speed_100m	0.072
DMC	0.057	soil_temperature_100_to_255cm	0.065
ISI	0.053	BUI	0.062
BUI	0.051	soil_temperature_7_to_28cm	0.056
wind_speed_100m	0.042	soil_moisture_100_to_255cm	0.053

Table 1.2: Variable importance for 2022

2022	
Variable	Importance
DMC	0.127
BUI	0.111
DC	0.087
ISI	0.054
FWI	0.052

The last 24 rows of each sample were assigned as 1 in the boolean scale of wildfire occurrence, while the rest were assigned as 0. With the RandomForestClassifier algorithm, it was possible to obtain the data shown in Table 1.3 depicting the five most important features according to this method. The two least important features are *rain* and *precipitation*. Although *rain* was considered one of the least important features, it cannot be discarded because it is a variable in *FWI* component analysis, and three components from *FWI* are among the most important features that appear in table 1.3.

Table 1.3: Variable importance for hourly values

Variable	Importance
DC	0.091
DMC	0.075
soil_moisture_100_to_255cm	0.073
BUI	0.067
surface_pressure	0.067

Table 1.4 displays the five most important features, according to a daily average value. For each variable, a mean value was calculated, and it was assigned to the last day as a boolean value for wildfire 1, depicting an occurrence of a wildfire. This method showed little improvement in relation to the one shown in table 1.3. The two least important features considered were *precipitation* and *dew_point_2m*.

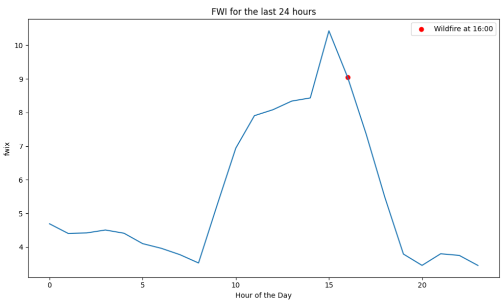
Table 1.4: Variable importance with daily mean method

Variable	Importance
soil_moisture_100_to_255cm	0.100
soil_moisture_28_to_100cm	0.090
terrestrial_radiation	0.075
wind_speed_100m	0.072
terrestrial_radiation_instant	0.054

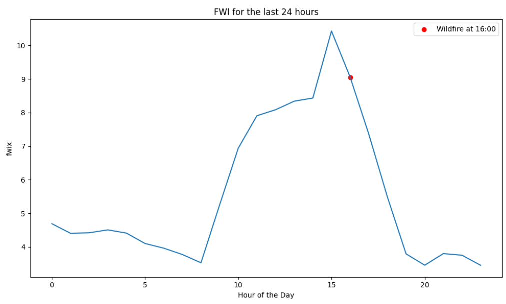
Another experiment was conducted by selecting a time frame between 9:00 and 20:00 hours. Taking a glace at figure 1.1, the value of *FWI* starts to increase around 8 or 9 in the morning, has its highest value around 15, and then it starts to decrease. After around 20 hours, it reaches its lowest point in the samples from the years 2015 and 2019.

Table 1.5 shows soil moisture as the most important feature, like in Table 1.3. There was also another experiment where the time frame was selected according to the hour of the wildfire.

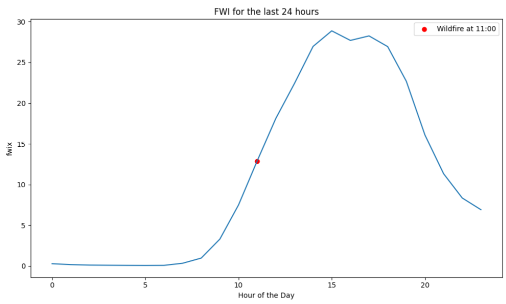
Figure 1.1: Hourly FWI value for the day of wildfire occurence



(a) 2015



(b) 2019



(c) 2022

Table 1.5: Variable importance with *FWI* time frame method

Variable	Importance
soil_moisture_100_to_255cm	0.072
BUI	0.067
DC	0.059
soil_moisture_7_to_28cm	0.057
DMC	0.054

A time frame three hours after and before the hour of wildfire (table 1.6) was selected. It yielded the same variables as table 1.5 but in a different order.

Table 1.6: Variable importance with 3-hours time frame method

Variable	Importance
DC	0.073
DMC	0.055
soil_moisture_100_to_255cm	0.046
soil_moisture_28_to_100cm	0.045
BUI	0.043

The authors from (Wang et al., 2023) set that the maximum time frame for wildfire weather variable analysis is 16 days. Given that the sample from 2022 only contains hourly data for 14 days, the table 1.7 displays the most important features taken 14 days prior to the wildfire occurrence from all three samples. Like the experiment displayed in table 1.4, a daily average was calculated for each day.

Table 1.7: Variable importance 14-days prior time frame method

Variable	Importance
soil_temperature_28_to_100cm	0.123
wind_speed_10m	0.105
apparent_temperature	0.087
terrestrial_radiation	0.072
soil_moisture_100_to_255cm	0.049

A last experiment (table 1.8) was also conducted with an hourly method. The last 360 rows from each sample were selected without taking in averages, it was set as a boolen variable for wildfire occurrence all rows 3 hours prior and 2 hours after the wildfire.

Table 1.8: Variable importance 14-days prior hourly time frame method

Variable	Importance
sunshine_duration	0.125
apparent_temperature	0.111
soil_temperature_0_to_7cm	0.107
temperature_2m	0.107
terrestrial_radiation	0.100

References

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Wang, Z., He, B., Chen, R., and Fan, C. (2023). Improving wildfire danger assessment using time series features of weather and fuel in the great xing'an mountain region, china. *Forests*, 14(5):986.