

Índice general

1. Introducción	3
1.1. Objetivos	3
1.2. Hipótesis	4
1.3. Revisión bibliográfica	4
Bibliografía	7

Capítulo 1

Introducción

destrucción de vastas áreas de bosque, la degradación del suelo pérdida de biodiversidad
s

, además de poner en riesgo infraestructuras y vidas humanas.

Andalucía es la segunda comunidad de España con más terreno forestal, con 4.325.378 ha de suelo forestal que suponen el 49.37 % de su superficie.

En 2022, 15.786,64 hectarias fueron afectadas por el fuego en Andalucía, prácticamente el doble de la media anual de los 10 años anteriores, 8873,68 hectarias [Memoria Plan INFOCA 2022].

En el Plan Estratégico Estatal del Patrimonio Natural y de la Biodiversidad a 2030 [Cita] “identifica las principales presiones y amenazas, entre las cuales destaca el aumento de la superficie forestal afectada por incendios forestales en 2022”.

1.1. Objetivos

El objetivo de esta investigación es construir modelos de *Machine Learning* que permitan predecir incendios forestales en la Comunidad Autónoma de Andalucía. Para abordar dicho objetivo, se realizarán las siguientes tareas:

1. Construir un conjunto de datos que permita la realización de análisis estadísticos y la posterior construcción de modelos de *Machine Learning* (ML) para la predicción de incendios forestales en Andalucía a partir de un estudio previo del problema.
2. Modelizar el riesgo de incendio forestal usando distintos algoritmos de ML y comparar sus resultados
3. Analizar el desempeño de los modelos en la realidad estudiando potenciales casos de interés.

1.2. Hipótesis

1.3. Revisión bibliográfica

Cortez y Morais [1]: En esta investigación se evalúa el rendimiento de distintos modelos de Machine Learning, como Máquinas de Soporte Vectorial, Árboles de Decisión, Regresión Lineal Múltiple, Naive Bayes, Bosques Aleatorios y Redes Neuronales, para predecir el área quemada por los incendios forestales en el Parque Natural de Montesinho, en el norte de Portugal, a partir de información meteorológica. Los mejores resultados se obtienen para el modelo de SVM considerando 4 variables explicativas: temperatura, humedad relativa, precipitaciones y viento.

Jain et al. [3]: El artículo revisa las aplicaciones del Machine Learning (ML) en la ciencia y gestión de incendios forestales, destacando su uso en seis dominios clave: caracterización de combustibles, detección y mapeo de incendios; clima y cambio climático; ocurrencia, susceptibilidad y riesgo de incendios; predicción del comportamiento del fuego; efectos del fuego; y gestión de incendios. Se identifican 300 publicaciones hasta finales de 2019, mostrando el uso frecuente de algoritmos de ML como random forests, MaxEnt, redes neuronales artificiales, árboles de decisión, máquinas de vectores de soporte y algoritmos genéticos. La revisión enfatiza las ventajas y limitaciones de estos métodos y subraya la necesidad de combinar la experiencia en la ciencia del fuego con técnicas avanzadas de ML para poder construir modelos realistas y útiles.

Vilar del Hoyo et al. [6]: Se aplica regresión logística en una cuadrícula de 1×1 km en la Comunidad Autónoma de Madrid, utilizando datos socioeconómicos como variables predictoras para representar los factores antropogénicos relacionados con el riesgo de incendio. También se evalúa otro enfoque basado en la predicción de la densidad de puntos de ignición en una cuadrícula de 10×10 km, utilizando funciones Kernel. La ocurrencia histórica de incendios de 2000 a 2005 se utiliza como variable de respuesta. El rendimiento de los modelos se evalúa con los incendios ocurridos en 2006 y 2007, obteniendo un AUC de 0.70 y 0.67 para ambos modelos, respectivamente.

Liz-López et al. [4]: En este artículo se presenta el nuevo Modelo de Evaluación de Incendios Forestales (WAM, por sus siglas en inglés) que utiliza deep learning para anticipar el impacto de los incendios a partir de información meteorológica satelital y el NDVI en Castilla y León y Andalucía. Las variables respuesta del modelo son el área quemada, el tiempo de control y extinción, y la cantidad de recursos humanos, aéreos y pesados necesarios para la extinción del incendio. Emplea una red convolucional residual que realiza regresiones sobre variables atmosféricas e índice de verdor. El WAM se preentrena con 100,000 ejemplos de datos sin etiquetar y se ajusta con un pequeño conjunto de 445 muestras etiquetadas.

Gutiérrez-Hernández et al. [2] En este trabajo analiza la superficie calcinada por los incendios forestales ocurridos entre 1975 y 2013 en la Comunidad Autónoma de Andalucía en función de 15 variables explicativas: altitud, insolación, pendiente, precipitación invernal, precipitación estival, temperatura estival, velocidad del viento, frecuencia del viento, superficie protegida, superficie de monte público, superficie de usos forestales, distancia a viario, distancia a zonas pobladas, saldo demográfico y saldo ganadero. Se aplica un modelo de Regresión Lineal Múltiple y un Modelo de Regresión Geográficamente Ponderada.

Los mejores resultados se obtienen en el segundo modelo, puesto que permite considerar la estructura de correlaciones espaciales presentes en los datos.

Martínez-Fernández et al. [5] En este artículo se identifican los factores humanos asociados con un mayor riesgo de incendio forestal en España y se analiza la distribución espacial de la aparición de incendios forestales en el país tomando como unidad de estudio el municipio. Se consideran 108 variables, de las cuales se seleccionan 29 tras un primer estudio exploratorio, las cuales se usan para entrenar un modelo de regresión logística para estimar la probabilidad de una alta o baja ocurrencia de incendios. Finalmente, tan solo resultan significativas 13 de las variables consideradas.

Bibliografía

- [1] CORTEZ, PAULO y MORAIS, ANÍBAL DE JESUS RAIMUNDO (2007). «A data mining approach to predict forest fires using meteorological data». <https://hdl.handle.net/1822/8039>.
- [2] GUTIÉRREZ-HERNÁNDEZ, OLIVER; SENCIALES-GONZÁLEZ, J. M. y GARCÍA, LUIS V. (2015). «Los incendios forestales en Andalucía: investigación exploratoria y modelos explicativos».
- [3] JAIN, PIYUSH; COOGAN, SEAN C.P.; SUBRAMANIAN, SRIRAM GANAPATHI; CROWLEY, MARK; TAYLOR, STEVE y FLANNIGAN, MIKE D. (2020). «A review of machine learning applications in wildfire science and management». *Environmental Reviews*, **28(4)**, pp. 478–505. doi: 10.1139/er-2020-0019. <https://doi.org/10.1139/er-2020-0019>.
- [4] LIZ-LÓPEZ, HELENA; HUERTAS-TATO, JAVIER; PÉREZ-ARACIL, JORGE; CASANOVA-MATEO, CARLOS; SANZ-JUSTO, JULIA y CAMACHO, DAVID (2024). «Spain on fire: A novel wildfire risk assessment model based on image satellite processing and atmospheric information». *Knowledge-Based Systems*, **283**, p. 111198. ISSN 0950-7051. doi: 10.1016/j.knosys.2023.111198. <https://www.sciencedirect.com/science/article/pii/S0950705123009486>.
- [5] MARTÍNEZ-FERNÁNDEZ, JESÚS; VEGA-GARCÍA, CRISTINA y CHUVIECO, EMILIO (2009). «Human-caused wildfire risk rating for prevention planning in Spain». doi: 10.1016/j.jenvman.2008.07.005.
- [6] VILAR DEL HOYO, L.; ISABEL, MARTÍN; M.P. y MARTÍNEZ VEGA, F.J. (2011). «Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data». *European Journal of Forest Research*, **130**, pp. 983–996. doi: 10.1007/s10342-011-0488-2.