

# Readme

[English]

- Description:

This repository contains the files generated during the preparation of my master's final project. This project is about predictive analysis of naval incidents using machine learning algorithms. Please, use project main report (located in 2.ExportedReports folder) as entry point.

- Versions / branches:

This project has been created in two parallel versions with the same methods, contents and conclusions:

- Branch #1: Texts in English and Python as a language in data processing.
- Branch #2: Texts in Spanish and R as a language in data processing.

This repository contains both two branches.

- Directory content:

- 1.Approach: Preparatory information for the development of the project.
- 2.ExportedReports: Main Report and annexes (pdf format)
- 3.DataPreprocess: Notebooks for the Data Preprocessing.
- 4.DataExplore: Notebooks for Data Exploration (EDA)
- 5.DataModel: Notebooks for Machine learning models, comparison and explainability

- Title of the Master's project:

"Machine Learning for safety and efficiency in the naval environment:

Predictive analysis of naval incidents in the US, 2002 – 2015"

- Master's degree tittle:

"Big Data and Data Science applied to economics and business administration and management" (1500 hours)

- University:

UNED (Spain)

- Author:

Oscar Anton

- Date of presentation:

December 2023

- Maintenance:

No extensive code or content maintenance is planned after its release in December 2023. It could be reviewed later if required.

- License:

It is published under the premises of Creative Commons BY-NC-ND 4.0 DEED  
Attribution-NonCommercial-NoDerivatives 4.0 International

[Spanish]

· Descripción:

Este repositorio contiene los archivos generados durante la preparación de mi TFM. Este trabajo consiste en un análisis predictivo de incidentes navales utilizando algoritmos de aprendizaje automático. Por favor, utilice el informe principal del proyecto (ubicado en la carpeta 2.ExportedReports) como punto de entrada.

· Versiones / branches:

Este TFM se ha elaborado en dos versiones paralelas con los mismos métodos, contenidos y conclusiones:

- Versión #1: Textos en inglés y Python como lenguaje en el tratamiento de datos.
- Versión #2: Textos en español y R como lenguaje en el tratamiento de datos.

Este repositorio contiene ambas versiones.

· Contenido de los directorios:

- 1.Approach: Información preparatoria para la elaboración del proyecto.
- 2.ExportedReports: Memoria del TFM y anexos (formato pdf)
- 3.DataPreprocess: Notebooks para el preprocesado de datos.
- 4.DataExplore: Notebooks para Exploración de datos (EDA)
- 5.DataModel: Notebooks para Modelos de machine learning, comparativas y explicabilidad

· Título del TFM:

“Machine Learning para la seguridad y eficiencia en el entorno naval:  
Análisis predictivo de incidentes navales en EEUU, 2002 – 2015”

· Título del máster:

“Big Data y Data Science aplicados a la economía y a la administración y dirección de empresas” (1500 horas)

· Universidad:

UNED (España)

· Autor:

Oscar Antón

· Fecha de presentación:

Diciembre de 2023

· Mantenimiento:

No se ha previsto un mantenimiento exhaustivo de código ni contenido después de su presentación en diciembre de 2023. Podría ser revisado más adelante si alguna circunstancia específica lo requiere.

· Licencia de uso:

Se publica bajo las premisas de Creative Commons BY-NC-ND 4.0 DEED  
Atribución-NoComercial-SinDerivadas 4.0 Internacional