**EDEM**

**MÁSTER DATA ANALYTICS**

**CURSO 2023/2024**

**Trabajo de Fin de Máster**

|  |
| --- |
| **TÍTULO:** |

|  |
| --- |
| **Alumnos:**  Jorge Domínguez Lozano  Ignacio Reyes Vázquez  Borja Cabo Huélamo  Francisco Tudela Muñoz |

|  |
| --- |
| **Tutor Edem:** Pedro Nieto Pelaez  **Tutor Empresa:** Sara Adam Goig  **Empresa:** Roquette |

Septiembre de 2024



**Resumen**

La contaminación industrial es un problema que ha estado siempre presente en la sociedad moderna y que afecta negativamente tanto al medio ambiente como a la salud pública. Identificar rápidamente las fuentes de las que procede la contaminación es esencial para poder mitigar sus efectos y mejorar la sostenibilidad de los procesos industriales. En nuestro estudio, hemos llevado a cabo una investigación en el uso de técnicas de machine learning para analizar los datos obtenidos de sensores industriales. Mediante el desarrollo de modelos predictivos, hemos logrado identificar patrones y relaciones en los datos que permiten detectar el origen de la contaminación de manera más precisa y oportuna. Este enfoque no solo ayuda a identificar las fuentes de contaminación, sino que también proporciona una herramienta valiosa para la toma de decisiones en tiempo real y la optimización de procesos industriales, con el fin de reducir el impacto ambiental.

**Palabras clave: Contaminación, Machine Learning, Sensores, ??? (Mínimo 4)**

**Abstract**

**Key words:**

**Índice**

**1.Introducción**

**1.1 Objetivos generales**

**2. Metodología de trabajo**

**3.Arquitectura**

**4. Exploración y Preprocesamiento de los Datos**

**4.1 Origen de los Datos**

**4.2 Tratamiento de Valores Ausentes**

**4.3 Codificación de Variables**

**5. Creación de Modelos**

**5.1 Selección de Variables**

**5.2 Regresión Logística**

**5.3 Random Forest**

**6.Análisis del Modelo Ganador**

**7. Conclusiones**

**8.Referencias y Bibliografía**

**Anexos**