# 📚 Documentación Técnica: NanoMof: Prescriptor Edafológico

## I. Introducción

El “**NanoMof: Prescriptor Edafológico”** es una plataforma de Decisión Predictiva (MDP) desarrollada por NanoMof para optimizar la dosificación de enmiendas de Biochar. El objetivo primario es transformar la aplicación de dosis empíricas a una **prescripción precisa y cuantificable** basada en la ciencia de datos y la experiencia agronómica histórica.

La aplicación utiliza un flujo de trabajo iterativo que permite el refinamiento continuo del modelo de predicción a partir de la ingestión de nuevos datos validados.

## II. Modelo Predictivo Central

### Algoritmo: Extreme Gradient Boosting (XGBoost)

El motor predictivo de la aplicación es el algoritmo **XGBoost**, un marco de *boosting* que implementa el aprendizaje automático de árboles de decisión en gradiente. Este modelo es altamente eficiente y escalable, y ha sido seleccionado por su capacidad para:

1. **Manejar Relaciones No Lineales:** Es ideal para modelar la compleja interacción edáfica donde la respuesta de la Dosis Efectiva no es linealmente proporcional a las variables del suelo.
2. **Alta Precisión:** Estudios de referencia en ciencia edáfica demuestran que XGBoost puede alcanzar coeficientes de determinación (R2) superiores a 0.90 en la predicción de eficiencia de adsorción y respuesta de enmiendas.
3. **Resistencia al Sobreajuste:** Su regularización inherente (L1 y L2) ayuda a prevenir el sobreajuste, asegurando que el modelo generalice bien a nuevos lotes de suelo no vistos previamente.

### Features de Entrenamiento (Variables de Entrada)

El modelo está entrenado para predecir la Dosis Efectiva (y) a partir de dos *features* edáficas fundamentales (X):

|  |  |  |  |
| --- | --- | --- | --- |
| **Feature** | **Descripción** | **Unidades** | **Rol en el Modelo** |
| **ph** | Potencial de Hidrógeno del suelo. | 0-14 | Indicador primario de toxicidad y de disponibilidad de nutrientes. |
| **mo** | Contenido de Materia Orgánica. | % | Indicador de la capacidad de intercambio catiónico (CIC) inicial y *buffer* nutricional. |

### Variable Objetivo (Output)

|  |  |  |
| --- | --- | --- |
| **Variable** | **Descripción** | **Unidades** |
| **dosis\_efectiva** | La cantidad de Biochar necesaria para lograr el mejoramiento buscado (ej. optimizar la retención hídrica o aumentar la CIC). | Toneladas por Hectárea (T/Ha) |

## III. Flujo de Trabajo y Módulos de la Aplicación

La aplicación se estructura en dos módulos interdependientes alojados en Streamlit.

### Módulo 1: 📂 Carga tus Datos Reales (Entrenamiento y Calibración)

Este módulo es el punto de ingestión de conocimiento.

1. **Ingesta de Datos:** El usuario carga un archivo CSV (delimitado por punto y coma ; y codificado en latin1) que contiene la Base de Datos Maestra de NanoMof.
2. **Preprocesamiento:** Se verifica la integridad estructural de las columnas (ph, mo, dosis\_efectiva). No se aplica escalado de *características* en esta versión para mantener la interpretabilidad directa de los rangos agronómicos.
3. **Ajuste del Modelo (model.fit):** Se entrena el modelo XGBoost utilizando los datos cargados.
4. **Validación:** Se calcula y muestra el coeficiente de determinación (R2) para evaluar el desempeño del modelo con los datos de entrenamiento. El modelo entrenado se almacena en memoria para su uso inmediato en la fase de predicción.

### Módulo 2: 🤖 Modo Simulación (Predicción y Consulta)

Este es el módulo de producción para la toma de decisiones.

1. **Inputs del Usuario:** Se solicitan los valores de pH y MO del suelo objetivo (el nuevo lote a tratar).
2. **Predicción (model.predict):** El sistema utiliza el último algoritmo XGBoost entrenado (generado en el Módulo 1) para calcular la Dosis Efectiva (T/Ha).
3. **Output:** Se presenta el resultado de la dosis recomendada, basado en el conocimiento acumulado.

**Nota:** La precisión de la Dosis de Prescripción está directamente correlacionada con la calidad y cantidad de datos cargados en el Módulo 1.

***NanoMof 2025 | HV Martínez-Tejada***