

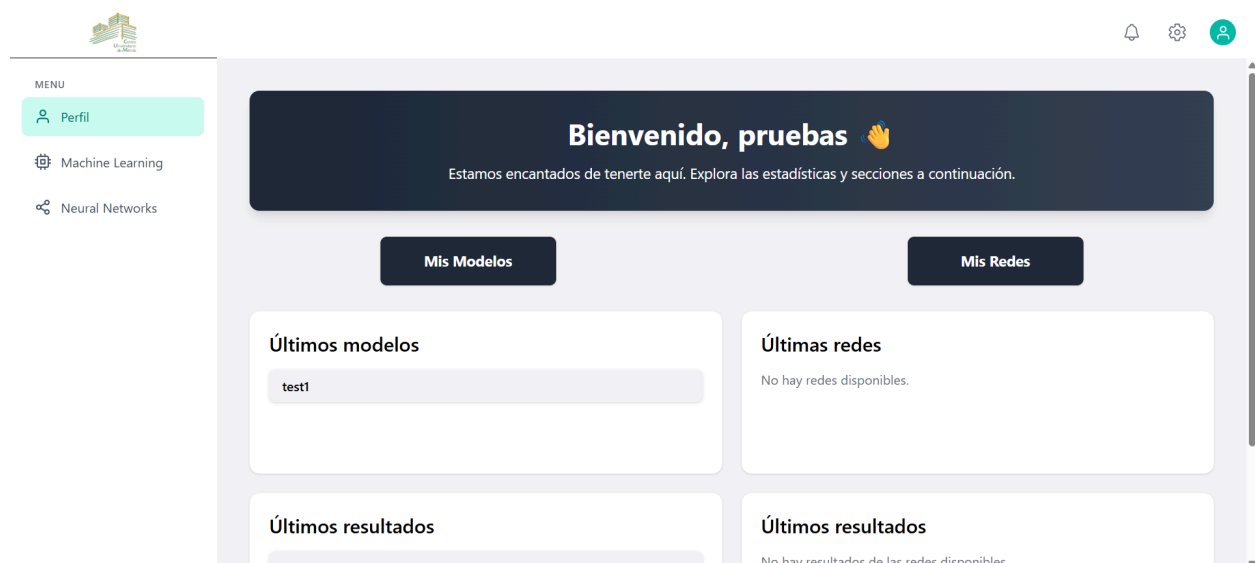
---

# Manual de Usuario: Plataforma de Machine Learning

Bienvenido al módulo de Machine Learning. Esta guía te ayudará a navegar por la plataforma, crear tus propios modelos de Inteligencia Artificial, configurarlos fácilmente aunque no tengas experiencia previa, entrenarlos y descargar los resultados.

## 1. Acceso al Centro de Machine Learning

Al iniciar sesión, llegarás al panel de bienvenida. En el menú lateral izquierdo, haz clic en **Machine Learning**.



Esto te llevará al **Centro de Machine Learning**, donde encontrarás dos opciones principales:

- **Mis Modelos:** Para crear, visualizar y gestionar los modelos que has configurado.
- **Mis Resultados:** Para consultar las métricas y descargar los archivos generados tras entrenar un modelo.



## 2. Creación de un Nuevo Modelo

Para empezar a trabajar, dirígete a la sección **Mis Modelos** y haz clic en el botón oscuro **Crear nuevo modelo**.

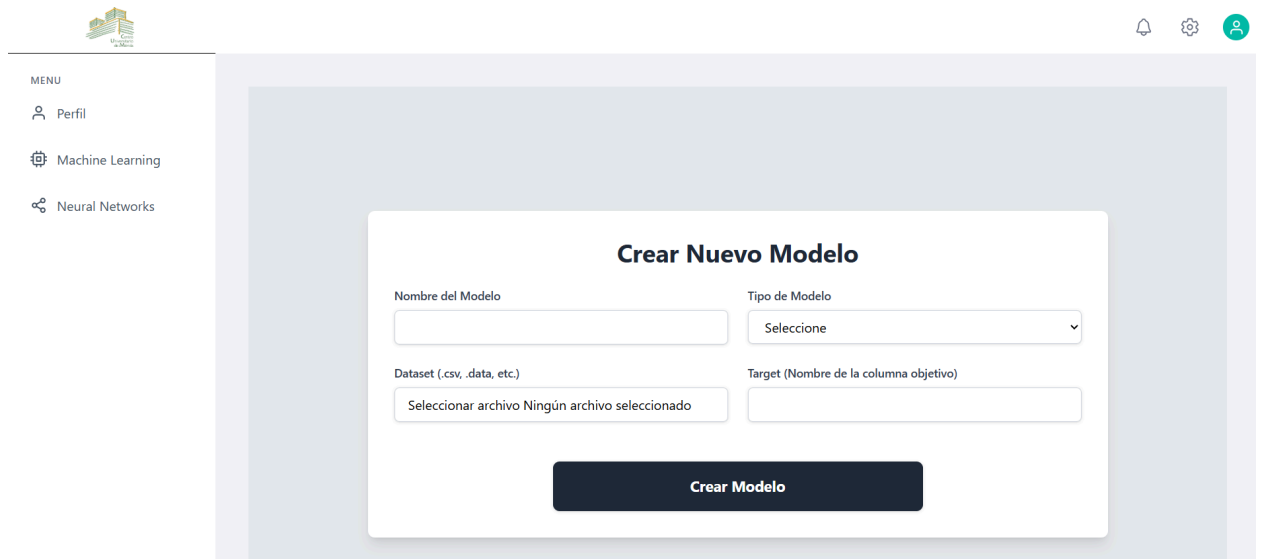


Se abrirá un formulario donde deberás completar la siguiente información básica:

1. **Nombre del Modelo:** Asigna un nombre que te ayude a identificarlo rápidamente (ej. *Mi\_Primer\_Modelo*).
2. **Dataset (.csv, .data, etc.):** Sube el archivo con los datos que utilizarás para enseñar a tu modelo. Preferiblemente en formato .csv y con un máximo de 2GB de tamaño.
3. **Target (Nombre de la columna objetivo):** Escribe el nombre exacto de la columna de tu

archivo que quieres que la IA aprenda a predecir (por ejemplo, *Proteína* o *Precio*).

4. **Tipo de Modelo:** Despliega el menú para elegir el algoritmo que deseas utilizar.

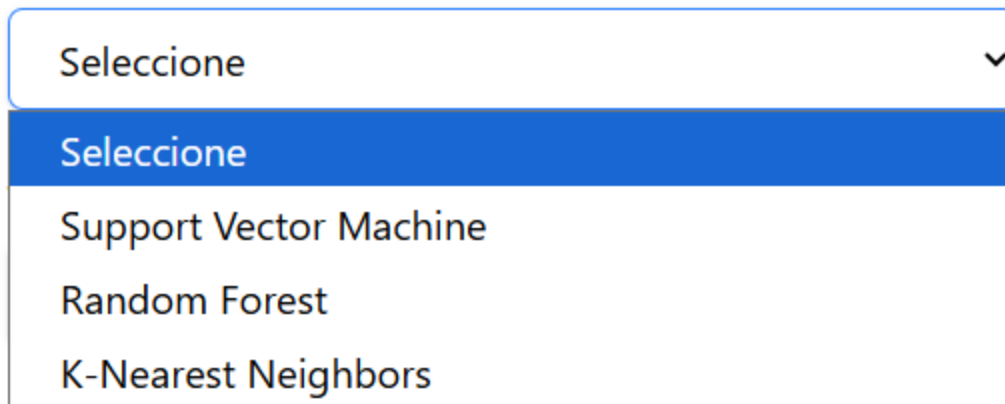


The screenshot shows a web application interface for creating a new model. On the left is a navigation menu with the following items: 'MENU', 'Perfil', 'Machine Learning', and 'Neural Networks'. The main content area is titled 'Crear Nuevo Modelo' and contains a form with the following fields:

- Nombre del Modelo:** A text input field.
- Tipo de Modelo:** A dropdown menu currently showing 'Seleccione'.
- Dataset (.csv, .data, etc.):** A text input field with the placeholder text 'Seleccionar archivo Ningún archivo seleccionado'.
- Target (Nombre de la columna objetivo):** A text input field.

At the bottom of the form is a dark blue button labeled 'Crear Modelo'.

### Tipo de Modelo



The screenshot shows a dropdown menu for selecting a model type. The menu is open, displaying the following options:

- Seleccione (highlighted in blue)
- Support Vector Machine
- Random Forest
- K-Nearest Neighbors

---

## 3. Guía de Algoritmos e Hiperparámetros

Al elegir un Tipo de Modelo, aparecerán opciones específicas llamadas "hiperparámetros". Piensa en ellos como los ajustes que le dicen a la IA cómo debe aprender. Aquí te explicamos qué significa cada uno de forma sencilla:

## A. Support Vector Machine (SVM)

Este modelo intenta trazar fronteras o líneas para separar y clasificar tus datos.

- **Kernel:** Es la forma geométrica que usará la IA para separar los datos.
  - *Linear:* Usa una línea recta. Ideal si tus datos son simples.
  - *Poly (Polinomial), RBF o Sigmoid:* Usan líneas curvas o formas complejas. Útiles cuando los datos están muy mezclados y una línea recta no sirve.
- **C (Parámetro de regularización):** Controla qué tan "estricta" es la IA. Un valor alto hace que la IA intente no cometer ni un solo error al aprender (lo que a veces es malo porque memoriza en lugar de aprender). Un valor bajo es más relajado y permite algunos errores para entender el panorama general.
- **Degree:** (Solo para kernel *Poly*). Define la flexibilidad de la curva. Un número mayor significa una curva con más dobleces.
- **Gamma:** (Solo para *RBF* y *Sigmoid*). Define cuánto alcance tiene la influencia de un solo dato. Si es alto, la IA se fijará mucho en los datos más cercanos; si es bajo, mirará el conjunto más amplio.

### Atributos de SVM

Kernel

C (Parámetro de regularización)

Degree (Solo para kernel 'Poly')

### Atributos de SVM

Kernel

C (Parámetro de regularización)

Gamma (Solo para kernel 'RBF' y 'Sigmoid')

### Atributos de SVM

Kernel

C (Parámetro de regularización)

Gamma (Solo para kernel 'RBF' y 'Sigmoid')

Kernel

C (Parámetro de regularización)

## B. Random Forest (Bosque Aleatorio)

Este modelo crea muchos "árboles de decisión" (como diagramas de flujo) y los hace votar para tomar la decisión final.

- **Número de Estimadores ( $n\_estimators$ ):** Es la cantidad de árboles en tu bosque. A mayor número (ej. 100 o 500), más precisa suele ser la decisión, pero el modelo tardará más en entrenar.
- **Profundidad Máxima ( $max\_depth$ ):** Es el número máximo de preguntas o niveles que puede tener cada árbol. Si el número es muy alto, el árbol se vuelve demasiado complejo; si es muy bajo, puede que no aprenda lo suficiente.

### Atributos de Random Forest

Número de Estimadores ( $n\_estimators$ )

Profundidad Máxima ( $max\_depth$ )

## C. K-Nearest Neighbors (KNN - K Vecinos Más Cercanos)

Este modelo predice el resultado fijándose en los datos que son más similares (sus "vecinos").

- **Número de Vecinos ( $n\_neighbors$ ):** ¿A cuántos vecinos debe preguntar la IA antes de tomar una decisión? Si pones 5, mirará los 5 datos más parecidos para predecir el resultado.
- **Pesos ( $weights$ ):** \* *Uniform*: El voto de todos los vecinos vale lo mismo.
  - *Distance*: Los vecinos que se parecen más (están más cerca) tienen un voto que vale más que los que se parecen menos.
- **P (Parámetro de potencia Minkowski):** Es la forma matemática de medir la distancia entre los vecinos.
  - *1 (Manhattan)*: Mide la distancia en ángulos rectos, como si caminaras por las manzanas de una ciudad.
  - *2 (Euclidean)*: Mide la distancia en línea recta, como volaría un pájaro.

## Atributos de KNN

Pesos (weights)

Número de Vecinos (n\_neighbors)

P (Parámetro de potencia Minkowski)

## 4. Gestión y Entrenamiento del Modelo

Tras crear un modelo y configurarlo, aparecerá en tu listado. Al hacer clic sobre él, accederás a la vista de **Detalles del Modelo**.

En esta pantalla podrás ver un resumen de tu configuración y una **Vista previa del dataset** para comprobar que tus datos se han cargado bien. En la parte inferior tienes tres botones:

- **Editar (Azul):** Para cambiar la configuración si te has equivocado en algún parámetro.
- **Eliminar Modelo (Rojo):** Para borrarlo permanentemente.
- **Entrenar (Morado):** Para que la IA empiece a estudiar tus datos.

The screenshot shows a web application interface. On the left is a navigation menu with 'MENU', 'Perfil', 'Machine Learning', and 'Neural Networks'. The main content area is titled 'Detalles del Modelo' and contains the following information:

- Nombre:** test1
- Tipo:** RF
- Dataset:** genceu\_intacto\_abs\_bruto\_-\_copia.csv
- Target:** Proteina
- Atributos RF:**
  - Número de Estimadores: 505
  - Profundidad Máxima: 2
- Vista previa del dataset:** A table with 10 columns (Proteina, 1000-1008) and 5 rows of data.

Proteina	1000	1001	1002	1003	1004	1005	1006	1007	1008
11,09	0,2112498	0,2110826	0,2107863	0,2104781	0,2101841	0,2098743	0,2095932	0,2092688	0,2090
11,09	0,2065777	0,206457	0,2061544	0,2058338	0,2056196	0,2054762	0,2052483	0,2049988	0,2048
10,13	0,2065429	0,2064923	0,2062648	0,2058975	0,2056305	0,2053114	0,2050088	0,2047121	0,2044
10,13	0,2024987	0,2024834	0,2024386	0,202288	0,2020033	0,201469	0,2011358	0,2007768	0,2004

The screenshot shows a web interface for machine learning. On the left is a menu with 'MENU', 'Perfil', 'Machine Learning', and 'Neural Networks'. The main area displays dataset information: 'Dataset: genuce\_intacto\_abs\_bruto\_-\_copia.csv', 'Target: Proteina', and 'Atributos RF: Número de Estimadores: 505, Profundidad Máxima: 2'. Below this is a 'Vista previa del dataset' table with 10 columns (Proteina, 1000-1008) and 5 rows of numerical data. At the bottom are three buttons: 'Editar' (blue), 'Eliminar Modelo' (red), and 'Entrenar' (purple).

Proteina	1000	1001	1002	1003	1004	1005	1006	1007	1008
11,09	0,2112498	0,2110826	0,2107863	0,2104781	0,2101841	0,2098743	0,2095932	0,2092688	0,2090
11,09	0,2065777	0,206457	0,2061544	0,2058338	0,2056196	0,2054762	0,2052483	0,2049988	0,2048
10,13	0,2065429	0,2064923	0,2062648	0,2058975	0,2056305	0,2053114	0,2050088	0,2047121	0,2044
10,13	0,2024987	0,2024834	0,2024386	0,202288	0,2020033	0,201469	0,2011358	0,2007768	0,2004
10,13	0,2068203	0,2066665	0,2065202	0,2063282	0,2061799	0,2060555	0,2058932	0,2057772	0,2055

## 5. Resultados y Descargas

Al pulsar en Entrenar te redirigirá automáticamente a la pestaña de Mis Resultados que está en el menú principal de Machine Learning, donde podrás ver el progreso de la tarea en tiempo real. Cuando la barra llegue al 100%, tu modelo estará listo.

El sistema generará una tarjeta con los resultados de la ejecución, donde tendrás dos opciones de descarga:

- **Resultados (.csv) (Botón Azul):** Un archivo de Excel/hoja de cálculo con las métricas para que veas qué tan bien lo ha hecho la IA.
- **Modelo (.joblib) (Botón Verde):** El "cerebro" de la IA ya entrenado, empaquetado en un archivo listo para que los programadores lo utilicen en otros sistemas.

MENU

- Perfil
- Machine Learning
- Neural Networks

Entrenamientos en Curso / Recientes

● **Entrenando: test1** EN PROGRESO  
Iniciado hace unos segundos • Task ID: d2be9218...

Iniciando entrenamiento del modelo... 25%

Modelo **test1** 1 resultado

resultados\_test1\_19\_ejecucion

Tipo: RF  
Fecha: 21/1/2026, 13:20:38

↓ Resultados (.csv)

📁 Modelo (joblib)

## 6. ¿Cómo interpretar el archivo de Resultados (.csv)?

Al abrir el archivo `.csv` descargado, verás varias columnas numéricas. La plataforma evalúa la IA haciendo varios "exámenes" (llamados *Folds*). Cada fila es el resultado de uno de esos exámenes. Aquí te explicamos de forma sencilla cómo leer estas métricas:

- **Fold:** Es el número de prueba o "examen". La IA divide tus datos en varios grupos y se evalúa a sí misma repetidas veces para asegurar que no ha tenido suerte en un solo intento.
- **Modelo:** El tipo de algoritmo utilizado (ej. SVM, RF).
- **R2 Score:** Es como el porcentaje de acierto general del modelo a la hora de encontrar patrones. Va de 0 a 1 (o de 0% a 100%). **Cuanto más cerca de 1, MEJOR.** (Por ejemplo, un 0.85 significa que el modelo es bastante bueno entendiendo los datos).
- **MAE (Error Absoluto Medio):** Es el error promedio en las predicciones. Si estás prediciendo precios de casas y el MAE es 10,000, significa que, en promedio, la IA se equivoca por 10,000 dólares (arriba o abajo). **Cuanto más cerca de 0, MEJOR.**
- **RMSE (Raíz del Error Cuadrático Medio):** Es parecido al MAE y se mide en las mismas unidades (ej. dólares), pero esta métrica penaliza de forma muy fuerte los errores grandes. Si el modelo a veces se equivoca por muchísimo dinero, el RMSE se disparará. **Cuanto más cerca de 0, MEJOR.**
- **MSE (Error Cuadrático Medio):** Es la misma idea que el RMSE pero los números están elevados al cuadrado. Suele usarse más por los matemáticos. Al igual que los dos anteriores: **Cuanto más bajo, MEJOR.**

---

## 7. Preguntas Frecuentes (FAQ)

### ¿Qué es un Dataset?

Es simplemente un conjunto de datos estructurados, normalmente en formato de tabla (como un archivo de Excel o .csv). Contiene la información histórica que la IA usará para aprender a encontrar patrones.

### ¿Qué es la columna "Target"?

Es la respuesta a la pregunta que quieres que la IA resuelva. Por ejemplo, si tienes una tabla con datos de casas (metros cuadrados, habitaciones, año de construcción) y quieres que la IA prediga el precio, la columna "Target" será la que contiene los precios.

### No sé de programación, ¿qué modelo debo elegir para empezar?

Si no estás seguro, **Random Forest** es una excelente opción para principiantes. Suele dar muy buenos resultados casi sin necesidad de tocar los hiperparámetros de configuración por defecto.

### Mi entrenamiento lleva mucho tiempo "En progreso", ¿es normal?

Sí. El tiempo de entrenamiento depende de dos cosas: la cantidad de datos que hayas subido (datasets grandes tardan más) y la complejidad del algoritmo y sus hiperparámetros (por ejemplo, un Random Forest con miles de estimadores requerirá más tiempo de procesamiento).

**Me salta en la página un error relacionado con un "token" como " Given token not valid for any token type; code: token\_not\_valid; messages: [object Object]" o "Error al cargar..."**.

Simplemente recarga la página, el servicio necesita refrescar tu token para cerciorarse de que eres tú.

---