

# API de Recomendación de Tratamiento

## Brain–Cancer AI Pipeline

Repositorio: <https://github.com/serpentegg-forge>

13 de junio de 2025

### Resumen

Esta documentación describe la API REST `/predict` del sistema de recomendación de tratamiento para tumores cerebrales. La API combina un modelo de visión por computadora (EfficientNet) para clasificar la condición del tumor y un modelo XGBoost para sugerir el tratamiento óptimo, recibe como entrada una imagen de resonancia (MRI) y datos clínicos, y devuelve predicciones con sus probabilidades.

## Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Requisitos e Instalación</b>	<b>2</b>
<b>3. Arranque del Servidor</b>	<b>2</b>
<b>4. Esquema del Endpoint</b>	<b>2</b>
4.1. POST <code>/predict</code> . . . . .	2
4.2. Respuesta Exitosa (200 OK) . . . . .	3
4.3. Códigos de Estado . . . . .	3
<b>5. Ejemplo de Uso con curl</b>	<b>3</b>
<b>6. Flujo Interno del Pipeline</b>	<b>3</b>
<b>7. Licencia y Contribución</b>	<b>3</b>

## 1. Introducción

El **Brain–Cancer AI Pipeline** expone un servicio HTTP que:

- Acepta una imagen MRI (formato JPG/PNG).
- Recibe variables clínicas: edad, sexo y nota de historia clínica.
- Devuelve el código y etiqueta de la condición tumoral, el código y etiqueta del tratamiento recomendado, y las probabilidades asociadas.

## 2. Requisitos e Instalación

- **Python 3.12+**
- **Dependencias** (instalar con `pip install -r requirements.txt`):
  - fastapi
  - uvicorn
  - pydantic
  - torch, torchvision, timm
  - xgboost
  - joblib
  - Pillow
- Clonar el repositorio:

```
git clone https://github.com/serpentegg-forge/brain-cancer-ai-pipeline.git
cd brain-cancer-ai-pipeline
```

## 3. Arranque del Servidor

En la raíz del proyecto:

```
py -3.12 -m uvicorn api.main:app --reload \
  --host 127.0.0.1 --port 8000
```

Una vez iniciado, la API estará disponible en <http://127.0.0.1:8000>.

## 4. Esquema del Endpoint

### 4.1. POST /predict

- **URL:** <http://<host>:<puerto>/predict>
- **Método:** POST
- **Encabezados:**
  - Content-Type: multipart/form-data
- **Parámetros de formulario:**

```
image file JPG/PNG (requerido)
age int (0age<120) (requerido)
```

`sex str {M, F}` (requerido)

`clinical_note str` (mínimo 10 caracteres; debe seguir la sintaxis de los registros clínicos usados en entrenamiento para que el preprocesador interno extraiga correctamente síntomas, intensidad y duración: describir síntomas tras “experiencing” o “suffering from” seguido de la duración y opcionalmente “described as <intensidad>in nature”).

**Ejemplo:** Patient experiencing recurrent seizures and blurred vision for the past few months, described as severe in nature.

## 4.2. Respuesta Exitosa (200 OK)

```
{
  "condition_code": 1,
  "condition_label": "Brain Glioma",
  "treatment_code": 3,
  "treatment_label": "surgery",
  "probabilities": [0.12, 0.05, 0.70, 0.13]
}
```

## 4.3. Códigos de Estado

**200 OK:** Predicción exitosa.

**415 Unsupported Media Type:** Si `image` no es JPG/PNG.

**422 Unprocessable Entity:** Error de validación de datos (edad, sexo, nota).

**500 Internal Server Error:** Error en el servidor o en la inferencia.

## 5. Ejemplo de Uso con curl

```
curl -X POST "http://127.0.0.1:8000/predict" \
-F "image=@ruta/a/brain_glioma_0001.jpg" \
-F "age=72" \
-F "sex=F" \
-F "clinical_note=Paciente_con_convulsiones_recurrentes_y_vision_borrosa"
```

## 6. Flujo Interno del Pipeline

1. La imagen es procesada por *EfficientNet* para inferir `condition_code`.
2. La nota clínica se parsea para extraer síntomas, intensidad y duración.
3. Se construye un `DataFrame` con edad, sexo, `condition_code` y síntomas.
4. Se aplica un `ColumnTransformer` pre-entrenado.
5. El modelo XGBoost ponderado predice `treatment_code` y sus probabilidades.

## 7. Licencia y Contribución

Este proyecto está bajo licencia MIT. Para contribuciones, abre un *issue* o *pull request* en GitHub:

<https://github.com/serpentegg-forge/brain-cancer-ai-pipeline>