

CALIDAD DE ACEITE DE OLIVA USANDO IA - ANÁLISIS DE LA CALIDAD DEL ACEITE DE OLIVA CON APRENDIZAJE SUPERVISADO

P.E. Ban(1)
(1) Industria 4.0 INTI Mendoza, Energías Renovables y Gestión Energética Cuyo.
INTI, Araoz 1511, Lujan de Cuyo, Mendoza, Argentina | tmendoza@inti.gob.ar

1. Resumen del Caso

Realizamos una solución utilizando Inteligencia Artificial para poder determinar la calidad del aceite de oliva, las cuales pueden ser: Aceite de oliva virgen extra, Aceite de oliva virgen y Aceite de oliva lampante.

La clasificación de la calidad está dada por el Consejo Oleícola Internacional tomando como base la acidez del producto que es su primordial índice de pureza.

Se han analizado los índices clásicos de calidad, los datos fueron relevados de distintas regiones de Argentina, usando los análisis realizados en el laboratorio de análisis físico químicos del INTI Regional Mendoza años 2020 al 2024.

En nuestra discusión de los parámetros fisicoquímicos, nos centraremos en aquellos que son legalmente requeridos para la clasificación del aceite de oliva virgen o extra virgen.

PARÁMETROS FÍSICO-QUÍMICOS

- Acidez
- Índice de peróxido
- K232
- K270
- ΔK

Parámetro	Virgen extra AOVE	Virgen	Lampante
Acidez	≤ 0,8	≤ 2,0	> 2
Índice de peróxidos	≤ 20	≤ 20	> 20
K ₂₃₂	≤ 2,50	≤ 2,60	---
K ₂₇₀	≤ 0,22	≤ 0,25	---
ΔK	≤ 0,01	≤ 0,01	---

2. Procesos

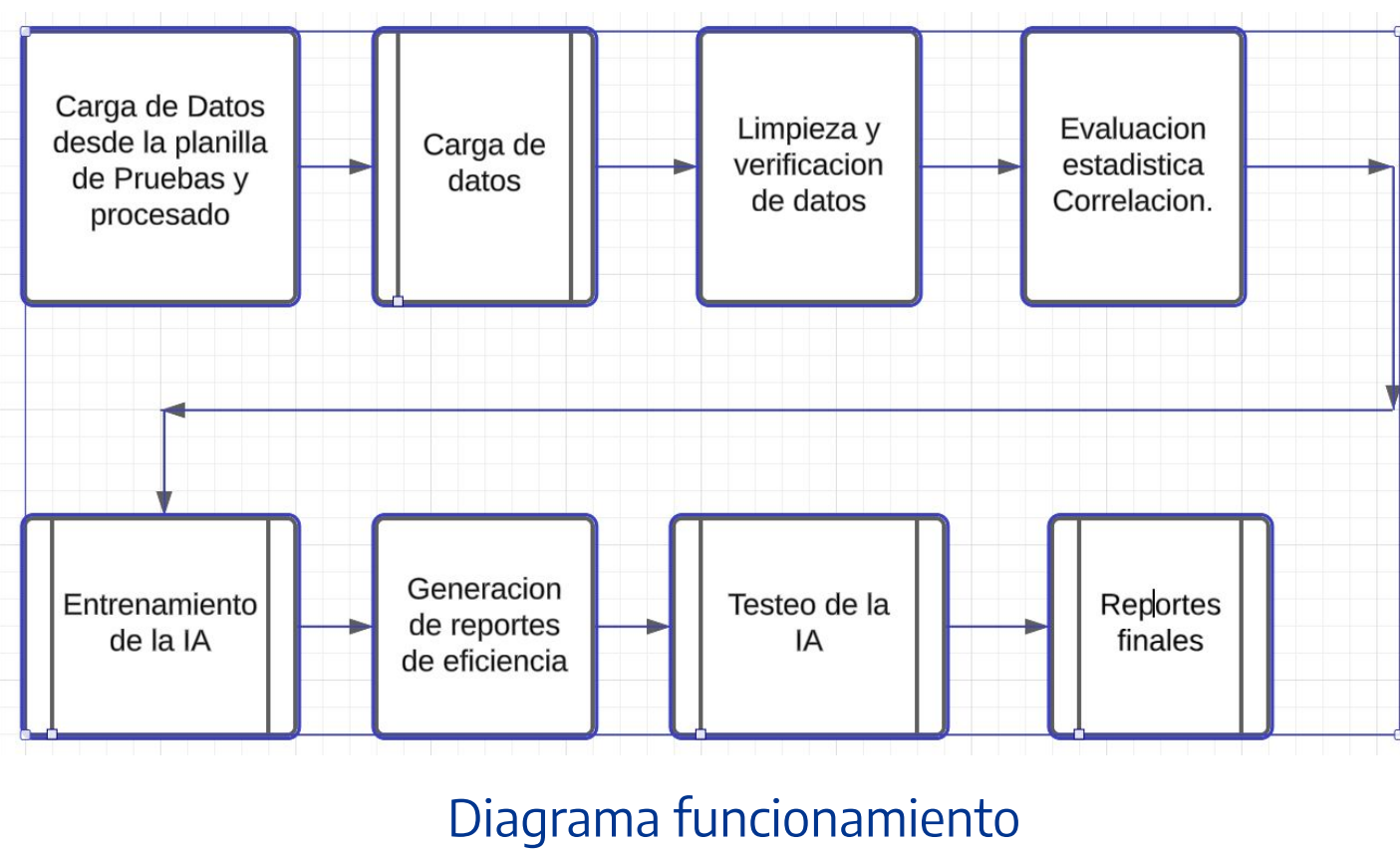
- Modelo y entrenamiento:

Se procesan los datos disponibles para extraer características que nos permitan distinguir entre unas muestras u otras.

En base a esas características o parámetros definimos un modelo de entrenamiento.

Al modelo se le entrena con el primer conjunto de muestras informando del tipo de muestra que procesa en todo momento y él en base a esa información, reestructura su algoritmo (el peso que le otorga a las características usadas) con objeto de poder dar respuesta satisfactoria al 100% de las muestras.

- Validación-Predicción, operaremos con las muestras restantes.



Primero ejecutamos el modelo obtenido en el paso anterior y le pedimos nos informe sobre lo que ocurre con cada muestra.

Después de que nos dé su respuesta, le decimos si ha acertado o no.

Al terminar tendremos un porcentaje de aciertos que nos servirá para determinar si el modelo es válido o no.

3. Herramientas o métodos utilizados

Que utilizamos:

Utilizamos Python como lenguaje de programación.

Scikit-Learn es una de estas librerías gratuitas para Python. Cuenta con algoritmos de clasificación, regresión, clustering y reducción de dimensionalidad. Definimos un número de elementos para entrenamiento, en este caso 75% y el 25% para prueba.

TensorFlow es una biblioteca de código abierto para aprendizaje automático a través de un rango de tareas, capaces de construir y entrenar redes neuronales para detectar y descifrar patrones y correlaciones



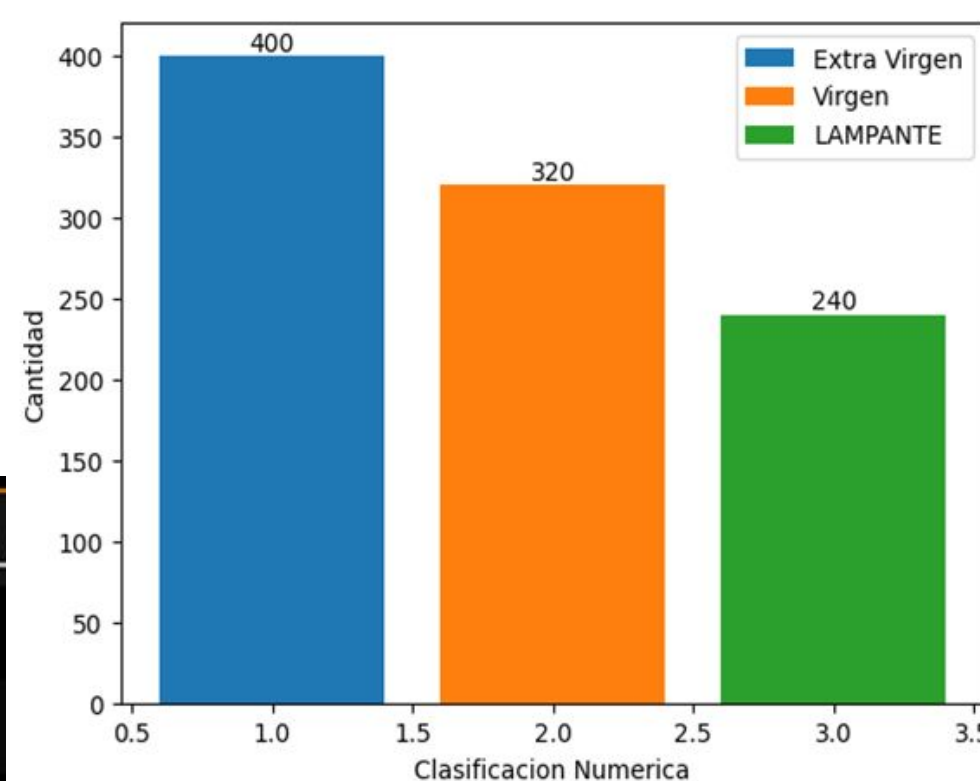
Datos de entrenamiento

Analizamos los datos obtenidos y su distribución en función de la clasificación numérica dada:

- 1.Extra Virgen
- 2.Virgen
- 3.Lampante

Orden	K232	K270	Acidez	I. Peróxidos	Clasif_Num	Quality	Ac.Extra	Ac.Virgen	Ac.
0	1	1.66	0.14	0.20	8.9	1 EXTRA	1	0	0
1	2	1.63	0.12	0.17	6.4	1 EXTRA	1	0	0
2	3	1.74	0.13	0.24	10.4	2 VIRGEN	0	1	0
3	4	3.00	0.90	1.00	15.0	3 LAMPANTE	0	0	1
4	5	1.63	0.12	0.17	6.4	1 EXTRA	1	0	0

Datos de entrenamiento.



4. Resultados alcanzados

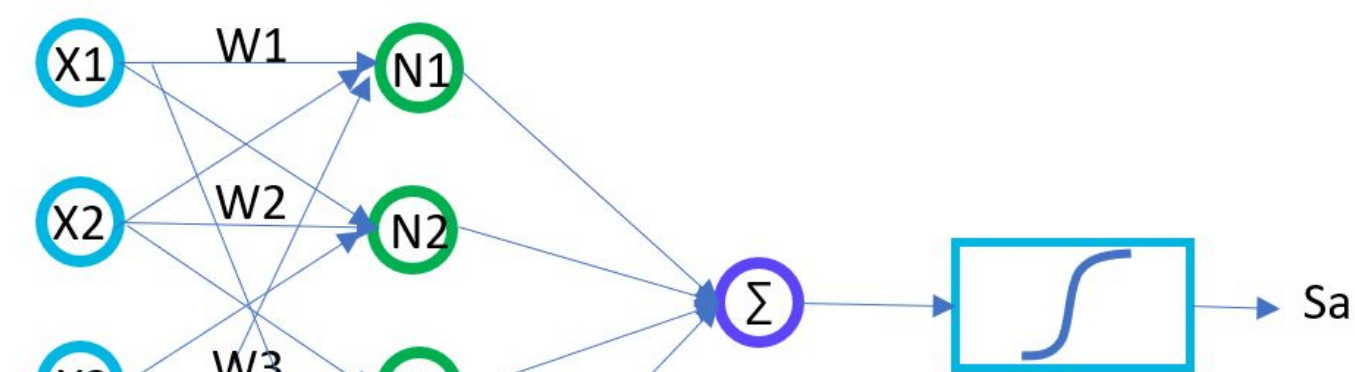
El algoritmo desarrollando responde a los valores en forma correcta. Utilizamos el 75% de los datos obtenidos para entrenar y el 25 para testear y corregir.

Regresión Logística.

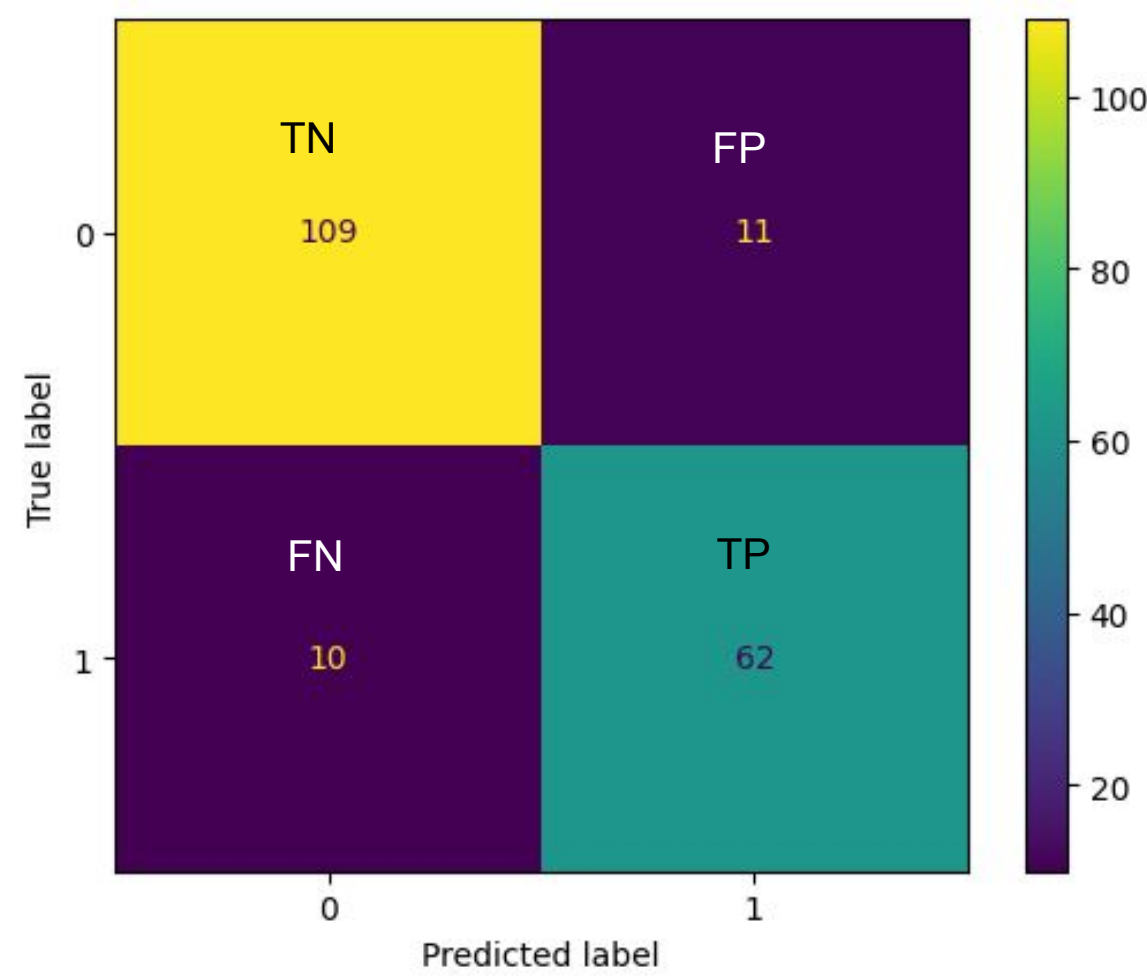
Creación del modelo.

La red neuronal tiene:

- 1 capa de entrada con 4 neuronas.
- 1 capa oculta de 4 neuronas.
- 1 capa salida tipo sigmoid.



Esquema de la red neuronal usada.



Resultados finales.

Aceite Virgen Extra

Precisión: 0.890625 -> 89,06%

Matriz de Confusión:

$\begin{bmatrix} 109 & 11 \\ 10 & 62 \end{bmatrix}$

	precision	recall	f1-score	support
0	0.92	0.91	0.91	120
1	0.85	0.86	0.86	72
accuracy			0.89	192

Como se puede observar el sistema puede detectar el tipo de aceite Virgen Extra, teniendo una efectividad del 89%, pero presenta un valor de falsos negativos de 11 detecciones, que corresponde a una relación del 11% aproximadamente con respecto a las detecciones positivas, esto representa un problema ya que detectamos aceites de calidad inferior como aceite virgen extra.

5. Próximos pasos

- Reforzar con nuevas habilidades la herramienta de IA
- Integrarlo con otras aplicaciones de laboratorio para hacer análisis on-line del tipo de aceite
- Incluir nuevos estudios como detección de oxidación del aceite.
- Aumentar el número de variables a analizar para poder determinar característica como origen o localización del olivar, terruño, certificación de origen.
- Este tipo de soluciones podríamos aplicarlo a diferentes tipos de industria agrícolas y frutícola como frutas finas, vinos, conservas, nueces, extractivas en general, entre otras.