

**ESTRATEGIA COMPUTACIONAL PARA ESTIMAR CUAL ES LA MEJOR RUTA PARA GENERAR EL MEJOR RENDIMIENTO EN LA COCCION DEL ATUN A PARTIR DE DATOS REALES SUMINISTRADOS POR MANUFACTURA DE ATUN, UTILIZANDO ALGORITMOS DE MACHINE LEARNING**

Corporación Universitaria Remington.

Nombre de la facultad: Ingenierías

Nombre del programa académico: Especialización en Dirección de Operaciones y Mejora Continua

Estudiante:

Giovanni Realpe Gonzalez

Tutor: Juan Carlos Briñez de León

Opción de Trabajo de grado Seminario-Diplomado.  
2025.

## Tabla de Contenidos

Resumen	3
Palabras Claves	4
1 Marco Conceptual y Contextual	5
1.1 Especies más Importantes de Atún	5
1.2 Que es Machine Learning?	9
1.3 Etapas en el Desarrollo de un Modelo de Machine Learning	9
2 Pregunta Problema	12
3 Acercamiento a los Datos	12
4 Objetivos	15
4.1 Objetivo general	15
4.2 Objetivos específicos	15
5 Análisis de los Datos	16
5.1 Estadísticas Descriptivas Generales	16
5.1.1 Variables numéricas clave	16
5.1.2 Estadísticas descriptivas que presentar	18
5.1.3 Visualización de variables clave	18
5.1.4 Correlaciones y relaciones relevantes	25
6 Desarrollo de los Modelos y despliegue de resultados	26
6.1 Preparación y análisis de los datos	27
6.2 Modelo de toma de decisiones	28
6.2.1 Código de los Modelos	29
6.2.2 Resultados de los Modelos	30
6.2.3 Análisis de desempeño	30
6.3 Validación del Modelo Entrenado	34
6.3.1 Resultados del Modelo Entrenado	35
7 Conclusiones y Trabajos Futuros	38
7.1 Conclusiones	38
7.1.2 Trabajos Futuros	40
Referencias Bibliográficas	41

## Resumen

En este proyecto al cual se llamo “estrategia computacional para estimar cual es la mejor ruta para generar el mejor rendimiento en la cocción del atún a partir de datos reales suministrados por manufactura de atún, utilizando algoritmos de Machine Learning” sustenta el desarrollo e implementación de un modelo de clasificación sirve para mejorar los procesos de cocción del atún en una de las plantas más grande de procesamiento de atún del Ecuador, ubicada en Posorja. La mayoría de las plantas que procesan atún funcionan con estructuras familiares y con sistemas de información poco robustos, lo que dificulta la toma de decisiones eficientes en los procesos críticos como lo es la cocción del atún, y sus decisiones se basan bajo su experiencia empírica.

El proceso de cocción de atún tiene un impacto directo en el rendimiento del pescado y en la rentabilidad de esta categoría, esto debió a que la materia prima representa en el costo total de fabricación del atún, en las latas en un 52%, en pouch 62% y en los lomos congelados un 80%.

Esta investigación se enfoca en analizar las principales variables y las mas criticas en el proceso de cocción del atún, como son; la especie, talla, peso de la pieza de atún, tempera de corte de las cocinas y el tiempo de cocción, así mismo el destino final del producto (latas, pouch o lomos congelados). Se utilizo la información histórica de este proceso de enero a mayo del 2025, con más de 4.100 registros. Con estos datos recopilados se entrenaron 10 modelos de clasificación con Python, siendo el modelo Decisión Tree y Random Forest con mayor nivel de accuracy del 98% de precisión.

Los resultados de los modelos demostraron una alta correlación entre variables como la talla del atún y tiempo de cocción del atún, así como la temperatura a la cual se corta la cocina o se para el calentamiento y el rendimiento total final. El tener un adecuado dominio de estas variables nos permitirá disminuir la merma y maximizar el rendimiento final del atún después de la cocción, contribuyendo beneficios hasta por 1 millón de dólares mensuales para la categoría de atún.

### **Palabras Claves**

Machine Learning, Cocción de Atún, Reducción de Merma, Optimización de Procesos, Random Forest, Industria Atunera.

## **1. Marco Conceptual y Contextual**

Actualmente en el Ecuador existen varias compañías donde procesan atún, se podría decir que el 90% son compañías familiares, no todas tienen implementado sistema de información robustos, por lo que hace relevante e innovador incluir algoritmos y técnicas Machine Learning, las cuales ofrecen su gran bondad de analizar datos y poderla transformar en información para tomar decisiones antes de iniciar el proceso productivo.

Vamos a enfocarnos en una de las partes más importantes del proceso productivo del atún, que es el preproceso o hasta la cocción del atún, la cual dependen de muchas variables especialmente de las especificaciones que tienen el atún, dentro de las más importantes son la especie, el peso y la talla, dependiendo de estas especificaciones y del destino final que tiene el atún como; si se va a empacar en una lata con liquido de cobertura en agua o en aceite, si se va a empacar en un pouch o si se empaca el lomo limpio en fundas plásticas. Todas estas especificaciones o características influyen en la toma de decisión, tales como la temperatura de cocción, tiempo de cocción y la temperatura de corte de la cocina.

### **1.1 Especies más Importantes de Atún:**

Antes de continuar con otros estudios sobre la cocción del atún, quiero ampliar las características y como es el atún:

**YellowFin o Aleta Amarilla:****Thunnus albacares**

**Dimensiones:** Longitud Máxima: 2 metros, Longitud Normal: 40 cm a 170 cm, Récord Absoluto IGFA: 176,333 Kg.

**Características Morfológicas:** Poseen un cuerpo fusiforme, más estilizado que el atún rojo o el patudo. Su cabeza y ojos son pequeños, la 2da. aleta dorsal y la anal son las más largas de todos los atunes, (van haciéndose más largas con el tiempo). Su hígado no presenta estrías en su superficie ventral. Posee vejiga natatoria. Se observan de 26 a 35 denticulos en el 1er. arco branquial. Sus aletas pectorales suelen sobrepasar el nacimiento de la 2da aleta dorsal, pero no van más allá del final de su base.

**Skipjack o Atún listado:****Katsuwonus Pelamis**

**Dimensiones:** Longitud Máxima: 1 metro, Longitud Normal: 40 a 70 cm, Récord absoluto IGFA: 19 Kg.

**Características Morfológicas:** Se pueden distinguir de las otras especies por la presencia de rayas en la zona ventral. Normalmente 4 o 6 rayas longitudinales muy visibles, que van desde el vientre y los flancos hasta la cola. El lomo del pez es azul oscuro con trazas púrpuras, mientras que los flancos inferiores y el vientre son plateados.

**Bigeye o Patudo:**



**Thunnus obesus**

**Dimensiones:** Longitud Máxima: 1,90 metros, Longitud Normal: 40 a 170 cm, Récord Absoluto IGFA: 170,322 Kg.

**Características Morfológicas:** Se caracterizan por tener un cuerpo muy robusto y ojos bastante grandes. La aleta pectoral alcanza a la 2da. aleta dorsal. El borde de la superficie de la zona ventral del hígado es estriado; siendo el lóbulo central más grande que sus otros dos lóbulos. Posee vejiga natatoria. Se contabilizan entre 23 y 31 denticulos en el 1er. arco branquial.

La temperatura de cocción del atún no se realiza con la temperatura de la cocina, esta temperatura se mide con unas termocuplas que miden la temperatura en la espina del atún no en la parte externa del atún, esto permite garantizar la cocción total de toda la carne del atún.

Preston (1990) indica que la temperatura es un concepto abstracto, el cual se explica por su efecto en las condiciones del medio ambiente, los objetos y sus propiedades, en general se relaciona con el comportamiento de la materia y en la mayoría de los casos define el estado final de esta.

Popularmente la temperatura es relacionada a los conceptos de frío y calor. Algo es más caliente si presente una mayor temperatura, o está más frío si se presenta una disminución en la temperatura. (Universidad de Chile, 2002).

Sin embargo, aunque este concepto es común y aceptable, en la realidad la física define a la temperatura como: “una magnitud escalar relacionada con la energía interna de un sistema termodinámico, definida por el principio cero de la termodinámica, o conocida como la energía sensible, que es la energía asociada a los movimientos de las partículas de ese sistema termodinámico”, esto quiere decir que la temperatura es la cuantificación de la actividad molecular de la materia. (Perez & Muzon, 2011).

La compañía que suministro toda la información del proceso de cocción, tiene un sistema computacional donde se lleva el control de las temperaturas de cada horno o cocina, este proceso se encuentra sistematizado y pueden obtener en tiempo real las temperaturas de cocción de atún, esto permite que se pueda integrar un sistema de Machine Learning y aplicar sistema de predicción antes de iniciar con el proceso de cocción del atún, esto

permite anticiparnos para saber que talla, especie, cocina, tiempo y temperatura de corte utilizar para maximizar el rendimiento del atún, de esta manera entregar a los procesos siguiente de limpieza y empaque mayor cantidad de materia prima cocinada, lo que se traduce a mayor cantidad de producción de latas, pouch o lomos.

Dentro de los datos analizados en la compañía manejan un indicador que por cada punto porcentual de reducción de mermas en toda la cadena de procesamiento de atún impacta positivamente en el estado de resultados de la compañía en aproximadamente un millón de dolores mensual.

### **1.2 Que es Machine Learning?:**

El Machine Learning (o aprendizaje automático) es un campo científico que pertenece al ámbito de la inteligencia artificial. Se basa en permitir que los algoritmos descubran patrones recurrentes en conjuntos de datos para que, de manera autónoma, aprendan y mejoren su rendimiento al realizar una tarea específica o al hacer predicciones.

### **1.2 Etapas en el Desarrollo de un Modelo de Machine Learning:**

Existen cuatro etapas principales en el desarrollo de un modelo de Machine Learning, generalmente supervisadas por un científico de datos, el ML resuelve situaciones por sí solo a partir de un análisis de datos y cuantos más datos tengan mejores resultados, además, para realizar el análisis se utilizan algoritmos que diseñan otros datos según las necesidades (Juárez, G. 2017).

A través de los datos de entrada, ML ejecutar un algoritmo y como resultado, genera más información para el problema (Bishop, C., 2007).

El objetivo de generar más datos se basa en las siguientes técnicas:

- Regresión lineal y polinómica.
- Árboles de decisión.
- Redes neuronales.
- Red bayesiana.
- Cadenas de Markov.

Los pasos que se utilizan en el ML son:

- **Recopilación y preparación de datos:** Se selecciona y organiza un bloque de datos de entrenamiento. Estos datos deben etiquetarse para indicar al modelo qué características identificar, o no etiquetarse para que el modelo detecte patrones por sí mismo. Es importante que los datos estén limpios y organizados para evitar sesgos en el aprendizaje.
- **Selección del algoritmo:** El algoritmo que se escoja depende del tipo de datos y del problema a resolver. Existen diferentes algoritmos que son apropiados para datos numéricos, texto o imagen.
- **Entrenamiento del modelo:** Se ejecuta el algoritmo sobre los datos de entrenamiento y se ajustan variables como pesos y sesgos para aumentar la precisión. El modelo aprende por repetición, comparando los resultados obtenidos

con los esperados, hasta alcanzar un desempeño satisfactorio en la mayoría de los casos.

- **Validación y aplicación del modelo:** El modelo entrenado se implementa para evaluar nuevos datos y realizar predicciones específicas al problema planteado.

Adicional esto permite diseñar maquinas inteligentes para transformar datos en conocimiento, en esta era de la tecnología moderna, contamos con un recurso abundante y de gran valor para la industria alimentaria que son los datos históricos. A mediados del siglo XX, el aprendizaje automático (Machine Learning) surgió como una subdisciplina de la inteligencia artificial, basada en algoritmos de autoaprendizaje que generan conclusiones a partir de datos para crear pronósticos.

En lugar de que un operario humano deba establecer manualmente todas las reglas para garantizar la calidad y eficiencia en la cocción de atún, el aprendizaje automático ofrece una alternativa más eficaz para capturar y procesar datos clave, mejorar gradualmente el rendimiento de los modelos predictivos y tomar decisiones fundamentadas en datos históricos y en tiempo real.

Esto resulta especialmente valioso en la producción de atún empacado en latas, pouch o lomo congelado, donde factores como la temperatura, el tiempo de cocción y la humedad afectan directamente la calidad y la inocuidad del producto final. El aprendizaje automático no solo aumenta la eficiencia en la operación de estos equipos, sino que también juega un papel cada vez más relevante en la investigación y en la automatización de esta fase crítica de la producción.

Gracias al aprendizaje automático, la industria del atún puede garantizar un control de calidad más riguroso, adaptarse a variaciones en la materia prima, reducir desperdicios y garantizar un producto final seguro y estándar para el mercado. De esta manera, los modelos de Machine Learning no solo fortalecen la eficiencia y la seguridad en la cocción de atún, sino que contribuyen a la evolución de una industria cada día más inteligente, sostenible y adaptada a las exigencias de los clientes de los diferentes países.

## 2 Pregunta Problema

¿Una estrategia computacional basada en algoritmos de **Machine Learning (ML)** puede ayudar de manera significativa a **reducir la merma y aumentar el rendimiento** en el proceso de cocción de atún?

## 3 Acercamiento a los Datos

Los datos analizados y utilizado para esta investigación fueron suministrados por una empresa atunera ubicada en Ecuador, específicamente en la Posorja a unos 100 kilómetros de la ciudad de Guayaquil, son expertos en la producción y procesamiento de productos del mar, entre ellos; el procesamiento de Atún, Sardina, Harina para alimento de animales, pelágicos, camarón y Aceite derivados del proceso de Harina. Estos datos son enfocados al primer proceso del atún que es la cocción del mismo, este proceso antecede a la limpieza del atún, donde quitan espinas, piel, cola, cabeza, estos datos corresponden al seguimiento

de diferentes lotes de atún procesados mediante diversas técnicas de cocción durante el periodo considerado a evaluar va desde enero a mayo de 2025 datos reales, el propósito es evaluar la eficiencia y estandarización del proceso de cocción utilizados en la planta atunera, enfocados en la reducción de la merma y el aumento del rendimiento de los diferentes lotes del atún.

La información utilizada para este análisis proviene de registros internos de la empresa, donde se recopilan variables clave para la operación, tales como:

- Turno.
- Libras por pieza de atún.
- Especie.
- Talla.
- Barco.
- Proceso de Eviscerado o Congelado.
- Tiempo de cocción.
- Corte de la temperatura de las cocinas.
- Liquido de cobertura.
- Numero de la cocina.
- Merma %.
- Rendimiento final.
- Peso de las cámaras, peso luego del eviscerado, peso luego de la cocción.

Los datos anteriores fueron procesados y depurados para evaluar la relación entre los diferentes parámetros de cocción y el destino final del atún (latas, pouch o lomos congelados). El análisis se orienta principalmente a determinar cómo las variables en temperatura y duración afectan la reducción de la merma durante la cocción, generando un aumento en el rendimiento final del atún antes de entregarlo al siguiente proceso de limpieza de este.

La información consolidada, se fundamenta en los registros históricos de la empresa y validada por el departamento de control de calidad y contraloría, para los periodos desde enero a mayo de 2025. Esta información es de utilidad como base para diagnosticar el impacto de los diferentes perfiles de cocción en la disminución de mermas, lo que impacta directamente en el aumento del rendimiento del atún, y para proponer mejoras en los protocolos de operación, alcanzando un balance óptimo entre calidad, inocuidad y maximización del rendimiento.

## 4 Objetivos

### 4.1 Objetivo general.

Desarrollar e implementar un modelo basado en Machine Learning que nos permita predecir y optimizar los parámetros críticos de cocción del atún (temperatura, tiempo, especie, talla y otras variables de operación) para reducir la merma y maximizar el rendimiento en el proceso de producción de una planta atunera ubicada en Posorja, Ecuador.

### 4.2 Objetivos específicos.

- Analizar y procesar los datos históricos de la empresa atunera para determinar la relación entre temperatura, tiempo de cocción, especie, talla y otras variables que afecten la generación de la merma y el rendimiento del atún cocido, que permitirá a la compañía utilizar menos materia prima o producir más producto terminado.
- Implementar algoritmos de Machine Learning para predecir la merma y el rendimiento antes de iniciar el proceso de cocción del atún.
- Validar el modelo de Machine Learning en ambiente de operación para evaluar su efectividad en la reducción de mermas y aumento de rendimiento de los diferentes lotes de atún procesados.

- Proponer recomendaciones para la adaptación de parámetros críticos en el protocolo de cocción, garantizando un balance óptimo entre calidad, inocuidad y eficiencia productiva.
- Mejorar los márgenes y utilidad bruta al implementar un modelo Machine Learning que anticipe los rendimientos antes de iniciar el proceso productivo va a generar beneficios importantes para la compañía, un punto porcentual al mes se traduce en un ahorro de 1 millón de dólares.

## 5 Análisis de los Datos

### 5.1 Estadísticas Descriptivas Generales

Se presenta un resumen estadístico de las variables recopiladas que se desarrollaron en la planta atunera en todo el proceso de cocción del atún en la planta de Posorja. Se incluye la descripción de medidas centrales y de dispersión que permiten identificar la distribución y características básicas de los datos.

#### 5.1.1 Variables numéricas clave:

1. **Turnos:** Se trabajan 3 turnos en este proceso de cocción del atún, esta variable se representa en el dataset de la siguiente manera (Primer turno = 10, Segundo turno = 20 y tercer turno = 30).
2. **Libras por pieza de Atún:** El peso que se trabaja del atún inicia desde 2 lbs hasta 50 lbs.

3. **Especie de Atún:** Las principales especies del atún son tres, en la base de datos original estaba en texto, se realiza una homologación para volverlas numéricas (Skipjack = 10, Yellowfin = 20, Bigeye = 30).
4. **Talla del atún:** Los atunes tienen una clasificación de acuerdo con el peso en el mundo atunero en cuatro agrupaciones, en los datos originales están como texto, estos se homologan para volarlo numérico (Pequeña = 10, Mediana = 20, Grande = 30 y Corte = 40). Corte son pescados muy grandes que se deben cortar con cierra para que se vuelva manipulable en el momento de la limpieza por las personas.
5. **Barcos:** La compañía tiene una flota propia, en los datos origen están como texto por su nombre, el cual se tuvo que homologar desde el 1 hasta el barco número 26.
6. **Eviscerado o Congelado:** La compañía de acuerdo con el tamaño del pescado puede tomar dos opciones eviscerar (quitar buche y vísceras) el pescado o pasarlo a la cocina congelado (con el buche y con las vísceras), el cual se homólogo de la siguiente manera; Eviscerado = 10 y Congelado = 20.
7. **Temperatura de Corte de Cocción:** Es el momento que se define no darle más temperatura al atún, estos cortes se miden en grados centígrados, los grados más utilizados son: 60,62,64,68, 70 y 80 grados.
8. **Tiempo de Cocción:** Es el tiempo que el pescado recibe temperatura dentro de las cocinas, los cuales van desde 72 hasta 252 minutos.
9. **Líquido de Cobertura:** Es el líquido con el que finalmente el atún se va a empacar como producto final, dependiendo de este destino se determina los tiempos y corte de

cocción, en los datos originales se tiene como texto, por lo cual se homologan a numéricos: Aceite = 10 y Agua = 20.

**10. Destino Final del Atún en el Producto Terminado:** es el empaque final del atún que se está cocinando, en los datos originales esta como texto, por tal motivo se homologa a numérico; Latas = 10, Pouch = 20, Lomo Congelado = 30.

**11. Numero de Cocina:** Es la numeración de las cocinas existentes por donde puede cocinarse el atún, las cuales van desde el 6,7,11,14,15,15,18,19,20,21 y del 26 a la 31.

**12. Rendimiento Final:** Es el % de atún final frente a la materia prima ingresada inicialmente antes de la cocina y del eviscerado, este % de rendimiento que se manejan históricamente; 70,71 y del 73 al 90% de rendimiento. Como ejemplo: Materia prima ingresada 100 kg y kilos luego del eviscerado y la cocción 90 kgs; merma del 10% y un rendimiento del 90%.

### 5.1.2 Estadísticas descriptivas que presentar:

	Turno	Lbs x Pieza	Especie	Talla	Barco	Evisce_Cong	Temperatura	Corte No.	Tiempo Coccion	Min	Tipo Liquido	Destino Final	N° Cocina	Rendimiento Final
count	4198.000000	4198.000000	4198.000000	4198.000000	4198.000000	4198.000000		4198.000000	4198.000000	4198.000000	4198.000000	4198.000000	4198.000000	4198.000000
mean	20.347785	9.621725	11.924726	23.446879	10.472606	13.156265		64.608385	124.820152	16.067175	12.548833	19.160314	84.336351	
std	8.128293	5.970288	4.662756	7.733709	5.244654	4.648203		3.851985	44.112008	4.885368	5.455814	7.724237	2.441748	
min	10.000000	2.000000	10.000000	10.000000	1.000000	10.000000		60.000000	72.000000	10.000000	10.000000	6.000000	70.000000	
25%	10.000000	5.000000	10.000000	20.000000	7.000000	10.000000		62.000000	87.000000	10.000000	10.000000	14.000000	83.000000	
50%	20.000000	8.000000	10.000000	20.000000	10.000000	10.000000		64.000000	90.000000	20.000000	10.000000	19.000000	85.000000	
75%	30.000000	14.000000	10.000000	30.000000	15.000000	20.000000		64.000000	147.000000	20.000000	10.000000	27.000000	86.000000	
max	30.000000	50.000000	30.000000	40.000000	26.000000	20.000000		80.000000	252.000000	20.000000	30.000000	31.000000	90.000000	

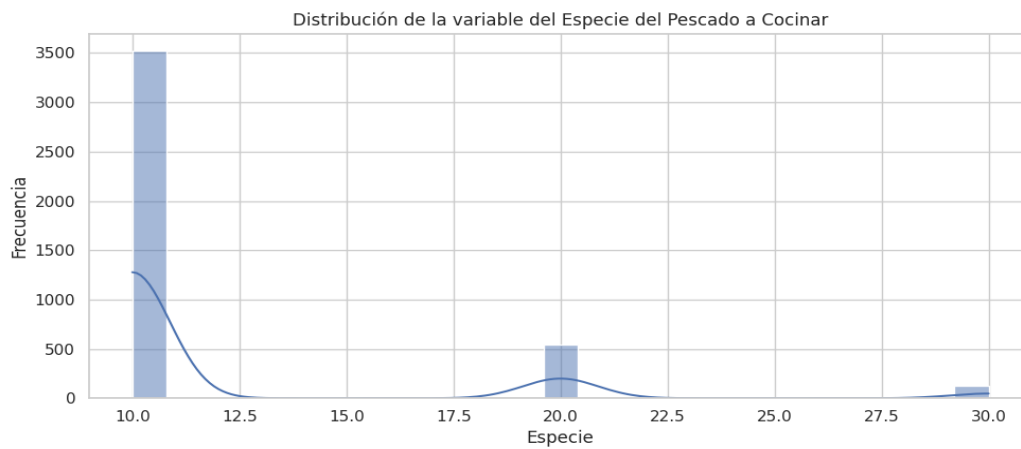
### 5.1.3 Visualización de variables clave:

Se mostrará la distribución de las principales variables resultantes de los datos suministrados por la compañía procesadora de Atún en su proceso de cocción de este, así como el código generado para poder visualizar los gráficos:

- **Especie:**

```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set(style="whitegrid")
plt.figure(figsize=(10, 5))
sns.histplot(Mis_datos['Especie'], bins=25, kde=True)
plt.title("Distribución de la variable del Especie del Pescado a Cocinar")
plt.xlabel("Especie")
plt.ylabel("Frecuencia")
plt.grid(True)
plt.tight_layout()
plt.show()
```

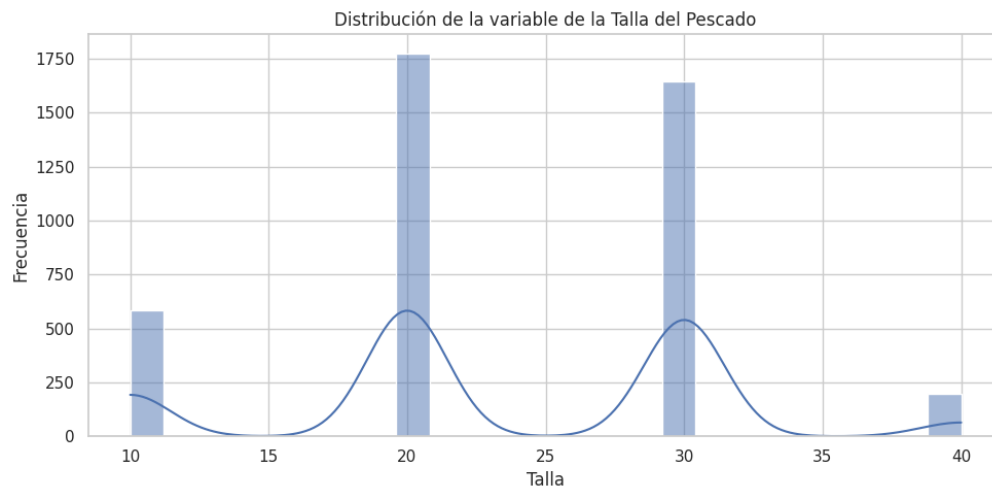


Podemos concluir que la mayor especie que se procesa es el Skipjack, seguida de Yellowfin y por último el Bigeye.

- **Talla:**

```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set(style="whitegrid")
plt.figure(figsize=(10, 5))
sns.histplot(Mis_datos['Talla'], bins=25, kde=True)
plt.title("Distribución de la variable de la Talla del Pescado")
plt.xlabel("Talla")
plt.ylabel("Frecuencia")
plt.grid(True)
plt.tight_layout()
plt.show()
```

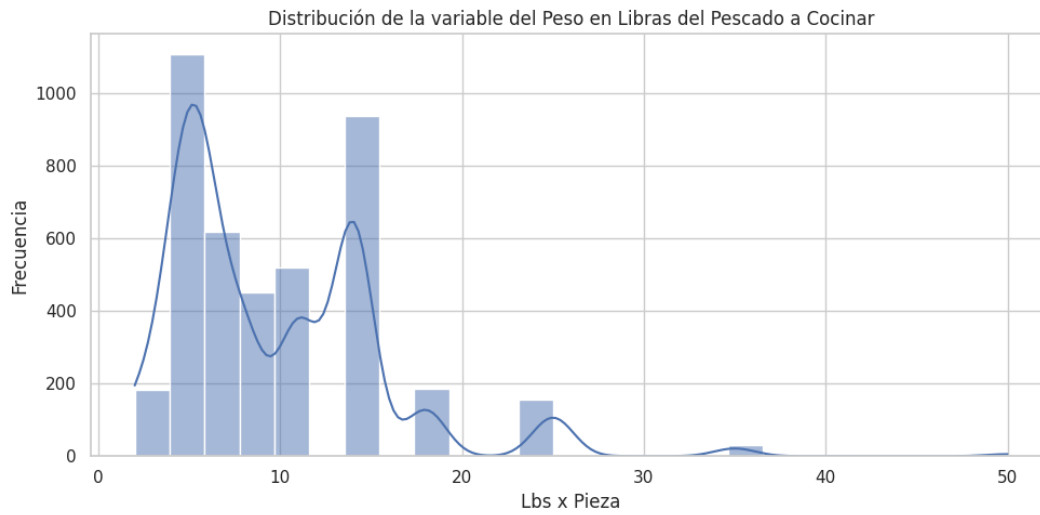


Se puede visualizar que las tallas más predominantes son la Mediana = 20 y la Grande = 30, seguida de la talla Pequeña = 10 y por último la talla para Corte = 40.

- **Libras por Pieza de Atún:**

```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set(style="whitegrid")
plt.figure(figsize=(10, 5))
sns.histplot(Mis_datos['Lbs x Pieza'], bins=25, kde=True)
plt.title("Distribución de la variable del Peso en Libras del Pescado a Cocinar")
plt.xlabel("Lbs x Pieza")
plt.ylabel("Frecuencia")
plt.grid(True)
plt.tight_layout()
plt.show()
```

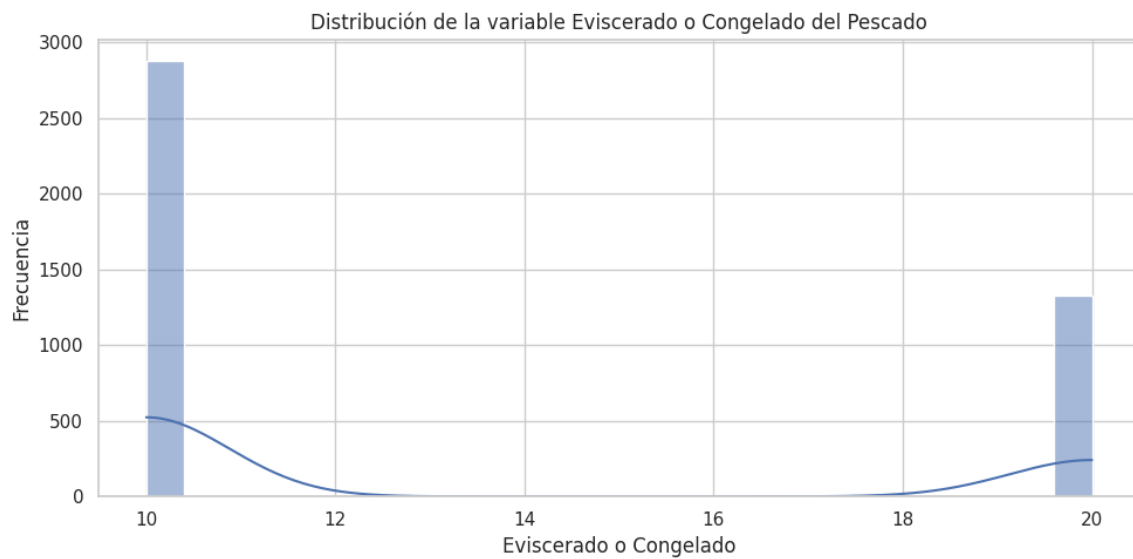


El atún con mayor participación a nivel de lbs está entre 5 y 6 lbs, seguida de 16 lbs.

- **Eviscerado o Congelado:**

```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set(style="whitegrid")
plt.figure(figsize=(10, 5))
sns.histplot(Mis_datos['Evisce_Cong'], bins=25, kde=True)
plt.title("Distribución de la variable Eviscerado o Congelado del Pescado")
plt.xlabel("Evisce_Cong")
plt.ylabel("Frecuencia")
plt.grid(True)
plt.tight_layout()
plt.show()
```

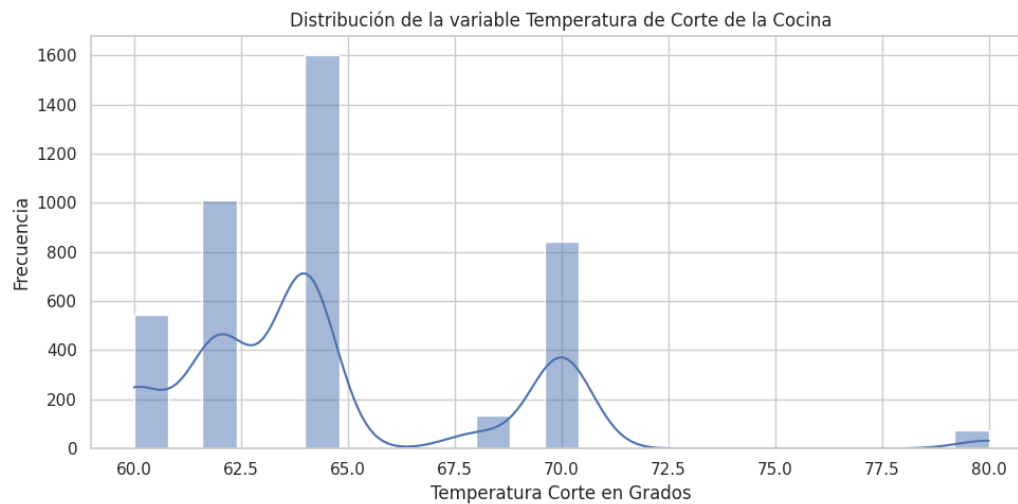


La mayor parte de la información trabajada es atún que se eviscera.

- **Temperatura de Corte:**

```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set(style="whitegrid")
plt.figure(figsize=(10, 5))
sns.histplot(Mis_datos['Temperatura Corte No.'], bins=25, kde=True)
plt.title("Distribución de la variable Temperatura de Corte de la Cocina")
plt.xlabel("Temperatura Corte en Grados")
plt.ylabel("Frecuencia")
plt.grid(True)
plt.tight_layout()
plt.show()
```

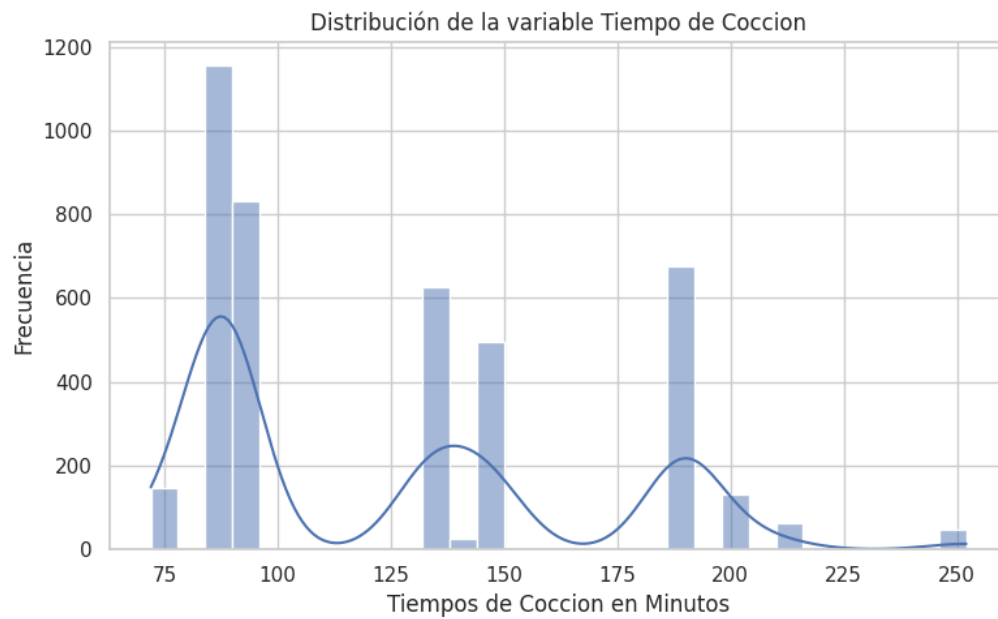


La temperatura de corte de las cocinas para no darle más cocción al atún más utilizada es a 64 grados, seguida de 62 grados.

- **Tiempos de Cocción del Atún:**

```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set(style="whitegrid")
plt.figure(figsize=(8, 5))
sns.histplot(Mis_datos['Tiempo Coccion Min'], bins=30, kde=True)
plt.title("Distribución de la variable Tiempo de Coccion")
plt.xlabel("Tiempos de Coccion en Minutos")
plt.ylabel("Frecuencia")
plt.grid(True)
plt.tight_layout()
plt.show()
```

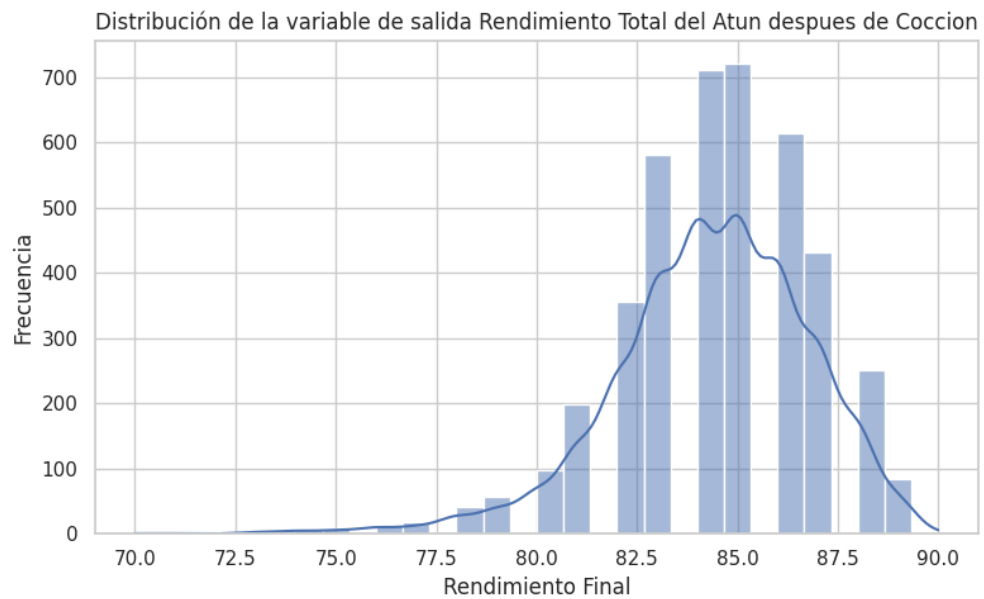


El tiempo de cocción que más se utiliza es de 83 minutos, seguido de 92 minutos.

- **Rendimiento Final:**

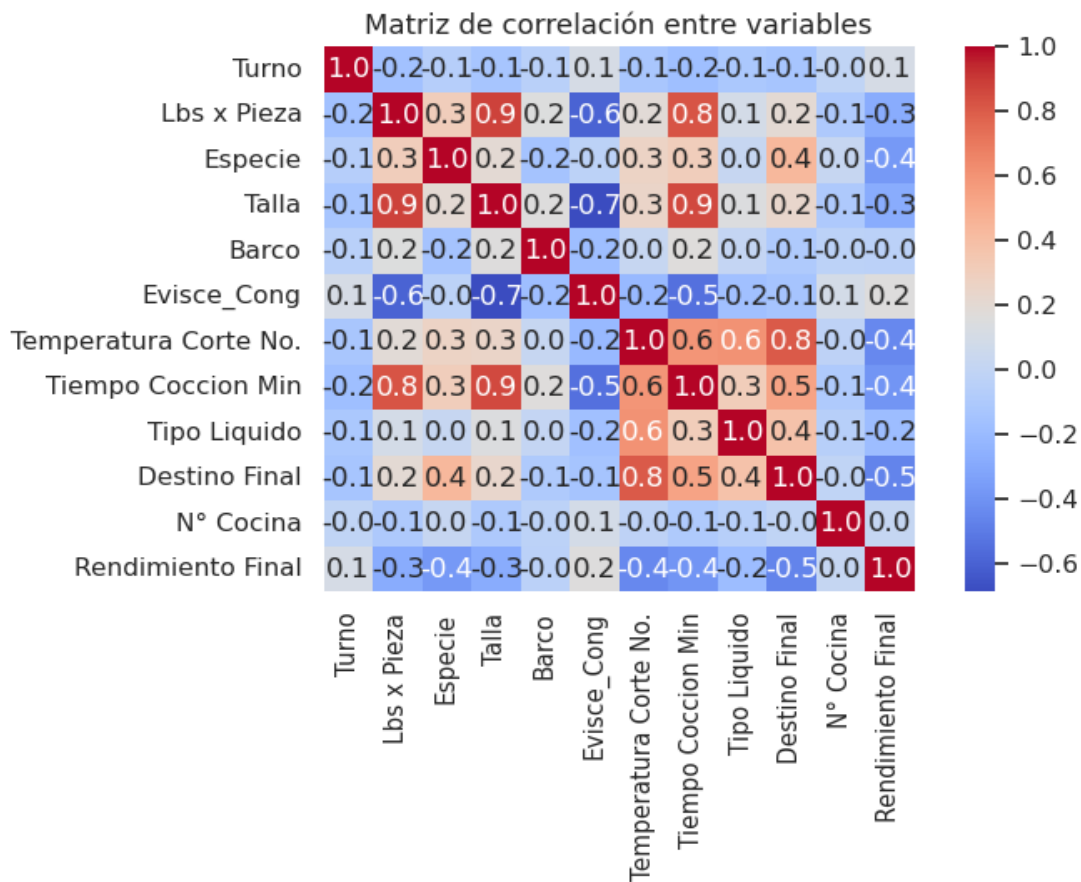
```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set(style="whitegrid")
plt.figure(figsize=(8, 5))
sns.histplot(Mis_datos['Rendimiento Final'], bins=30, kde=True)
plt.title("Distribución de la variable de salida Rendimiento Total del Atun despues de coccion")
plt.xlabel("Rendimiento Final")
plt.ylabel("Frecuencia")
plt.grid(True)
plt.tight_layout()
plt.show()
```



El rendimiento promedio de la operación de cocción del atún esta entre 82 y 83% de rendimiento, seguido de 86% y 87% de rendimiento.

### 5.1.4 Correlaciones y relaciones relevantes:



Entre las correlaciones más representativas están Talla y tiempo de cocción con 0,9, esto significa que entre más grande el pescado mayor tiempo de cocción se le debe dar. Adicional el Peso por Pieza y la Talla del atún presenta una correlación alta del 0,9, lo que significa que, a mayor peso, mayor la talla del atún, como se espera en la clasificación de tallas.

Destino final y Temperatura de Corte, presentan una correlación alta del 0,8, lo que significa que algunos productos o un grupo de productos finales se les debe dar un corte a una temperatura específica.

El eviscerado o congelado con las libras por pieza, tienen una correlación alta del  $-0,6$ , esto quiere decir que el atún a mayor peso en libras tiende a hacer eviscerados y los de menor talla pasan congelados para el proceso de cocción.

Rendimiento final y tiempo de cocción tienen una correlación del  $-0,40$ , lo que significa que a mayor tiempo de cocción menos rendimiento se obtiene, por la deshidratación del atún.

Rendimiento Final y temperatura de corte, tiene una correlación del  $-0,4$ , lo que significa que a mayor temperatura de corte puede generar mayor merma, lo que se traduce a menor rendimiento del atún.

Rendimiento y destino finales, tienen una correlación del  $-0,50$ , lo que significa que dependiendo del destino del atún (latas, pouch o lomo congelado), influye en el rendimiento final.

## **6 Desarrollo de los Modelos y despliegue de resultados**

El algoritmo de clasificación de Python es el más apropiado para predecir el mejor rendimiento en el proceso de cocción del atún de las diferentes especies y tallas, el poder determinar con anticipación el rendimiento esperado del atún, permite tomar correcciones antes de iniciar con el proceso de cocción, adicional se posee un gran volumen de información y de historia.

## 6.1 Preparación y análisis de los datos:

Se trabajo con los datos históricos registrados en preproceso de la planta atunera, las cuales podemos clasificar de la siguiente manera:

- **Variables de materia prima:** tamaño, especie, peso inicial, peso final y talla.
- **Variables de proceso:** temperatura de cocción, duración, presión y temperatura de corte de cocción.
- **Resultado final:** rendimiento final después de la cocción del atún.

Es importante resaltar que se adiciono una columna llamada “Valor Adecuado”, la cual ayuda a los modelos a tomar decisiones, en este caso es 1 cuando el rendimiento es adecuado o cuenta con un rendimiento alto dependiendo de las variables existentes y 0 cuando el rendimiento es bajo, basado en los registros de las diferentes variables de entrada.

Adjunto imagen de datos a utilizar, aparte de 7 registros:

Turno	Lbs x Pieza	Especie	Talla	Barco	Evisce_Cong	Temperatura	Corte No.	Tiempo Coccion Min	Tipo Liquido	Destino Final	N° Cocina	Rendimiento Final	Valor Adecuado
0	10	4	10	10	17	20	62	84	10	10	31	90	1
1	20	4	10	10	15	20	62	84	10	10	20	90	1
2	30	2	10	10	17	20	64	84	20	10	26	89	1
3	10	6	10	20	15	10	62	87	10	10	31	89	1
4	20	5	10	20	6	20	62	87	10	10	21	89	1
5	20	4	10	10	2	20	64	84	20	10	26	89	1
6	10	11	10	30	6	10	64	147	20	10	28	89	1

Adicional se detallan las variables utilizadas y las características de los datos: Datos:

4.198, Columnas 13, Datos Nulos 0, Tipo datos Entero.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4198 entries, 0 to 4197
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Turno                  4198 non-null   int64
1   Lbs x Pieza            4198 non-null   int64
2   Especie                4198 non-null   int64
3   Talla                  4198 non-null   int64
4   Barco                  4198 non-null   int64
5   Evisce_Cong            4198 non-null   int64
6   Temperatura Corte No.  4198 non-null   int64
7   Tiempo Coccion Min     4198 non-null   int64
8   Tipo Liquido           4198 non-null   int64
9   Destino Final          4198 non-null   int64
10  N° Cocina              4198 non-null   int64
11  Rendimiento Final      4198 non-null   int64
12  Valor Adecuado         4198 non-null   int64
dtypes: int64(13)
memory usage: 426.5 KB

```

## 6.2 Modelo de toma de decisiones:

Se utilizo el algoritmo de clasificación con 10 modelos diferentes, los cuales detallo el código y los resultados de este.

La muestra que se va a tomar de los 4.198 fila de datos es del 10% para modelarlo en los 10 modelos a trabajar:

```

import sklearn
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test= train_test_split(X, Y, test_size=0.10, random_state=4500)

```

## 6.2.1 Código de los Modelos:

```

from sklearn.neighbors import KNeighborsClassifier
from sklearn.naive_bayes import GaussianNB
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis, QuadraticDiscriminantAnalysis
from sklearn.tree import DecisionTreeClassifier
from sklearn.svm import SVC
from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier, AdaBoostClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
import warnings
warnings.filterwarnings("ignore")

# Modelo 0: K-Nearest Neighbors
modelo_0 = KNeighborsClassifier(n_neighbors=5)
modelo_0.fit(X_train, Y_train)
Y_pred_0 = modelo_0.predict(X_test)
print("Accuracy KNN:", accuracy_score(Y_test, Y_pred_0))

# Modelo 1: Naive Bayes
modelo_1 = GaussianNB()
modelo_1.fit(X_train, Y_train)
Y_pred_1 = modelo_1.predict(X_test)
print("Accuracy Naive Bayes:", accuracy_score(Y_test, Y_pred_1))

# Modelo 2: Linear Discriminant Analysis
modelo_2 = LinearDiscriminantAnalysis()
modelo_2.fit(X_train, Y_train)
Y_pred_2 = modelo_2.predict(X_test)
print("Accuracy LDA:", accuracy_score(Y_test, Y_pred_2))

# Modelo 3: Quadratic Discriminant Analysis
modelo_3 = QuadraticDiscriminantAnalysis()
modelo_3.fit(X_train, Y_train)
Y_pred_3 = modelo_3.predict(X_test)
print("Accuracy QDA:", accuracy_score(Y_test, Y_pred_3))

# Modelo 4: Decision Tree
modelo_4 = DecisionTreeClassifier()
modelo_4.fit(X_train, Y_train)
Y_pred_4 = modelo_4.predict(X_test)
print("Accuracy Decision Tree:", accuracy_score(Y_test, Y_pred_4))

# Modelo 5: Support Vector Machine
modelo_5 = SVC()
modelo_5.fit(X_train, Y_train)
Y_pred_5 = modelo_5.predict(X_test)
print("Accuracy SVM:", accuracy_score(Y_test, Y_pred_5))

# Modelo 6: Random Forest
modelo_6 = RandomForestClassifier()
modelo_6.fit(X_train, Y_train)
Y_pred_6 = modelo_6.predict(X_test)
print("Accuracy Random Forest:", accuracy_score(Y_test, Y_pred_6))

# Modelo 7: Logistic Regression
modelo_7 = LogisticRegression(max_iter=1000)
modelo_7.fit(X_train, Y_train)
Y_pred_7 = modelo_7.predict(X_test)
print("Accuracy Logistic Regression:", accuracy_score(Y_test, Y_pred_7))

# Modelo 8: Gradient Boosting
modelo_8 = GradientBoostingClassifier()
modelo_8.fit(X_train, Y_train)
Y_pred_8 = modelo_8.predict(X_test)
print("Accuracy Gradient Boosting:", accuracy_score(Y_test, Y_pred_8))

# Modelo 9: AdaBoost
modelo_9 = AdaBoostClassifier()
modelo_9.fit(X_train, Y_train)
Y_pred_9 = modelo_9.predict(X_test)
print("Accuracy AdaBoost:", accuracy_score(Y_test, Y_pred_9))

```

## 6.2.2 Resultados de los Modelos:

Los modelos con mayor accuracy es el modelo numero 4 -Decision Tree- con el 98% de precisión, igual que el modelo numero 6 -Random- Forest con una precisión del 98%, el modelo numero 2 es el menos confiable -Naive Bayes- con una precisión 69%.

```

➔ Accuracy KNN: 0.8261904761904761
Accuracy Naive Bayes: 0.6904761904761905
Accuracy LDA: 0.8214285714285714
Accuracy QDA: 0.7452380952380953
Accuracy Decision Tree: 0.9857142857142858
Accuracy SVM: 0.7595238095238095
Accuracy Random Forest: 0.9833333333333333
Accuracy Logistic Regression: 0.8238095238095238
Accuracy Gradient Boosting: 0.9
Accuracy AdaBoost: 0.8333333333333334

```

## 6.2.3 Análisis de desempeño

Para analizar el desempeño de los mejores modelos se consideró por su precisión el modelo numero 5 Decision Tree y el numero 7 Random Forest. Para el modelo numero 5 utilizamos la matriz de confusión y para el modelo 7 la curva ROC.

Código:

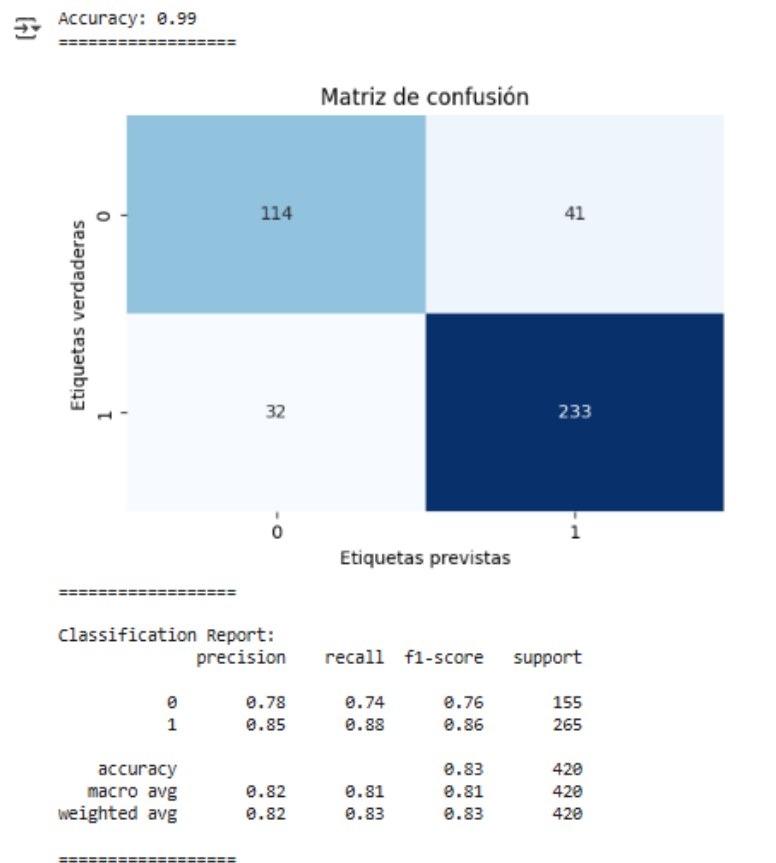
```

# Calcular métricas
accuracy = modelo_4.score(X_test, Y_test) # en esta parte coloco que modelo voy a correr.
cm = confusion_matrix(Y_test, Y_pred_0)
report = classification_report(Y_test, Y_pred_0)

# Calcular la curva ROC y el área bajo la curva (AUC)
Y_pred_prob = modelo_6.predict_proba(X_test)[:, 1]
fpr, tpr, _ = roc_curve(Y_test, Y_pred_prob)
roc_auc = auc(fpr, tpr)

```

## Resultados de la matriz de confusión:



## Interpretacion:

Podemos interpretar de la matriz de Confusion, con una muestra de 420 datos, representan al 10% como dato predeterminado inicialmente:

- Clase 0; 114 datos son verdaderos y 41 son falsos positivos.
- Clase 1; 233 datos son correctos y 32 son falsos negativos.

Con los resultados anteriores, la matriz de Confucion nos determina que el modelo utilizado acierta de manera correcta en sus predicciones en un 88% de los casos, lo que garantiza que se detecta correctamente los lotes de atun que tiene un alto rendimiento.

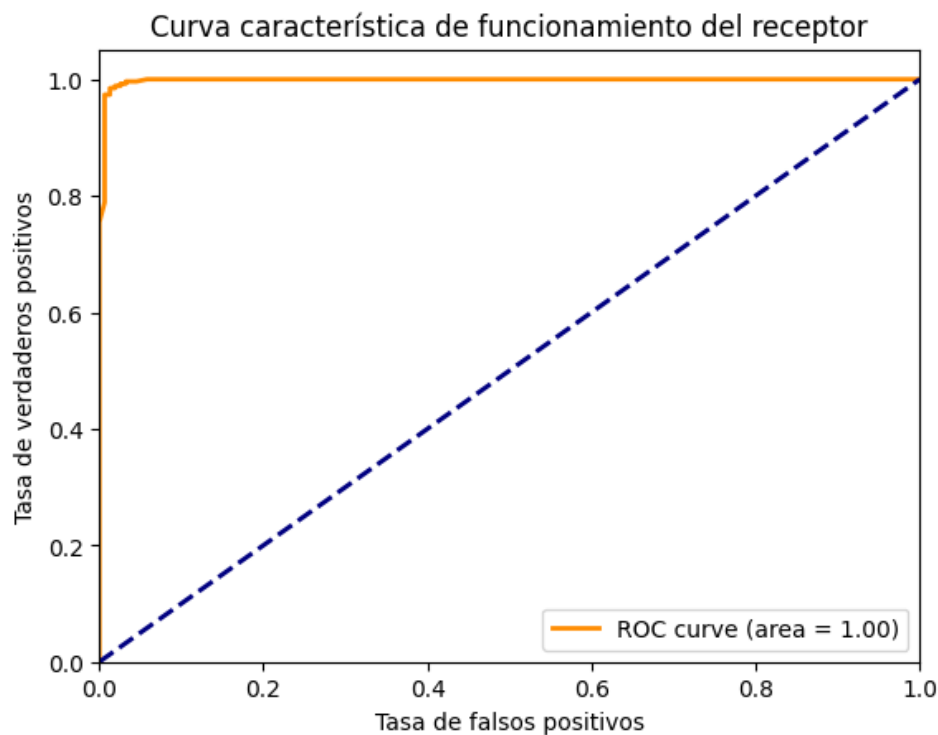
Y el modelo puede acertar los lotes de atun o las piezas de atun que tienen un bajo rendimiento final, el modelo acierta en un 74% de los casos.

Con los resultados del reporte de clasificacion, puede garantizar su funcionamiento para predecir en un 85% los rendimientos altos, se puede concluir que muestra una precision bastante alta, adicional presenta una sensibilidad del 88% de los casos con alto rendimiento del atun.

El modelo puede presentar un grado mayor de fallas para predecir bajos rendimientos con una precision del 78% y con una sensibilidad del 74%, lo que implica que no considera el 26% de los casos que pueden presentar bajos rendimientos.

Lo que podemos interpretar que es modelo seleccionado tiene un mayor alto grado de accuracy para detectar lotes de atun con un mayor rendimiento en la coccion del mismo, pero tambien se debe tener algo de cuidado con las predicciones de los piezas de atun que reflejan bajo rendimiento despues de su coccion. Lo que se recomienda realizar una validacion de los datos historicos suministrados, para mejor la precision del modelo.

### Resultados de la Curva ROC: (Característica Operativa del Receptor)



Con los resultados del modelo 7 Random Forest, se puede concluir que este modelo tiene una eficacia alta para discernir los lotes de atún con alto rendimiento y los que han tenido un bajo rendimiento, el accuracy es muy alto y recomendado para realizar las evaluaciones futuras.

La interpretación anterior se da por el comportamiento en el gráfico de la línea continua curva (naranja), la cual está en gran parte pegada a la parte superior izquierda alineada al 1, lo que indica un alto performance para distinguir entre los lotes de atún con alto y bajo rendimiento, y la tasa de falsos positivos es muy baja.

A pesar del buen resultado es importante validar la información histórica con los especialistas del área productiva que manejan el proceso de cocción del atún, esto permitiría robustecer y mejorar las predicciones futuras.

### 6.3 Validación del Modelo Entrenado

A continuación, se adjunta el código en Python del modelo entrenado para recibir input por parte de los usuarios, con la variables y rango de datos preestablecidos de acuerdo con los datos históricos suministrados por la compañía atunera.

```
import warnings
import numpy as np
from collections import Counter

warnings.filterwarnings("ignore")

# Vector de entrada para un solo sujeto
Target = np.zeros((1, 12)) # una fila y 13 datos

# Inputs con instrucciones dentro del mismo input
Target[0, 0] = int(input("Turno (10 = Turno Uno, 20 = Turno Dos, 30=Turno Tres): "))
Target[0, 1] = int(input("Lbs por Pieza de Atun (desde 2 lbs hasta 50 lbs): "))
Target[0, 2] = int(input("Especie de Atun (SKIP JACK = 10, YELLOW FIN = 20, BIG EYE = 30): "))
Target[0, 3] = int(input("Talla del Atun (Pequeña = 10, Mediana = 20, Grande = 30, Corte = 40): "))
Target[0, 4] = int(input("Barcos del 1 al 26: "))
Target[0, 5] = float(input("Eviserado = 10 y Congelado =20: "))
Target[0, 6] = float(input("Temperatura de Corte de Coccion (60, 62, 64, 68, 70 y 80): "))
Target[0, 7] = float(input("Tiempo de Coccion en Minutos (72,84,87,90,93,132,135,138,147,189,198,210,252): "))
Target[0, 8] = float(input("Tipo de Liquido de Cobertura del Producto Final (Aceite = 10 y Agua = 20): "))
Target[0, 9] = float(input("Destino Final del Atun en el Producto Terminado (Lata = 10, Pouch = 20, Lomo = 30): "))
Target[0, 10] = int(input("Numero de la Cocina a Utilizar (6,7,11,14,15,16,18,19,20,21, del 26 al 31): "))
Target[0, 11] = int(input("Rendimiento Final (70,71, (73 al 90)): "))

# Normalización
Target = scaler.transform(Target)
```

```
# Predicciones
predicciones = [
    modelo_0.predict(Target)[0],
    modelo_1.predict(Target)[0],
    modelo_2.predict(Target)[0],
    modelo_3.predict(Target)[0],
    modelo_4.predict(Target)[0],
    modelo_5.predict(Target)[0],
    modelo_6.predict(Target)[0],
    modelo_7.predict(Target)[0],
    modelo_8.predict(Target)[0],
    modelo_9.predict(Target)[0],
]

# Resultados individuales
nombres_modelos = [
    "KNN", "Bayes", "LDA", "QDA", "Árbol", "SVM",
    "Random Forest", "Logística", "Gradient Boosting", "AdaBoost"
]

for nombre, pred in zip(nombres_modelos, predicciones):
    print(f'Según {nombre}: {'✅ Rendimiento Adecuado' if pred == 1 else '❌ Rendimiento Bajo'})

# Mayoría
conteo = Counter(predicciones)
mayoria = conteo.most_common(1)[0][0]
total = conteo[mayoria]

print("\n--- Decisión por mayoría ---")
print(f"{'✅ Rendimiento Adecuado' if mayoria == 1 else '❌ Rendimiento Bajo'} ({{total}}/10 modelos)")
```

### **6.3.1 Resultados del Modelo Entrenado:**

**Ejercicio Numero 1:** El usuario ingresa los parámetros que se estiman van a procesar durante la jornada laboral, el cual se alinean con lo programado de producción enviado por el equipo de Supply Chain de acuerdo con la demanda. Detallo el flujo del proceso productivo hasta la cocción del atún.

Una vez el equipo de cámaras (almacenamiento del atún congelado a -18 grados centígrados), recibe el plan de producción alistan las especies y tallas que se requieren procesar, esto lo realizan con 6 horas de anticipación, luego pasa al área de descongelado donde el atún pasa de -18 grados a -2 grados, luego pasan al área de eviscerado, y por último pasan a las cocinas, las tallas que no se evisceran pasan directamente a la cocina

El atún se almacena en cámaras en unos tanques inoxidable, estos tanques tienen la trazabilidad del atún que contienen, como; barco que realizo la captura, zona FAO donde fue capturado y datos de las características organolépticas del atún.

Con la información anterior y con la programación enviada por Supply Chain, la cual también lleva la siguiente información; producto terminado a producir, si es en lata, pouch o lomo congelado, cliente, país destino, cantidad a producir, liquido de cobertura (aceite o agua), son el soporte para que las persona de producción que maneja el área de preproceso pueda ingresar en este modelo propuesto, y así poder anticiparse de las variables de tiempo, temperatura de corte pueden llegar a tener un rendimiento alto o no.

Para este primer ejercicio considerando la información ingresada por el usuario, los 10 modelos determinaron que con la información suministrada no se va a lograr tener un buen rendimiento del atún después de la cocción, y el rendimiento final será bajo.

En este caso si no cumple con el rendimiento esperado se deberá mover variables como tiempo y el corte de la temperatura, incluso cambiar de cocina (actualmente se tienen cocinas viejas y nuevas) con diferentes eficiencias en la cocción.

### Datos Ingresados por el Usuario y Resultado:

```

↔ Turno (10 = Turno Uno, 20 = Turno Dos, 30=Turno Tres): 10
Lbs por Pieza de Atun (desde 2 lbs hasta 50 lbs): 2
Especie de Atun (SKIP JACK = 10, YELLOW FIN = 20, BIG EYE = 30): 10
Talla del Atun (Pequeña = 10, Mediana = 20, Grande = 30, Corte = 40): 10
Barcos del 1 al 26: 1
Eviserado = 10 y Congelado =20: 10
Temperatura de Corte de Coccion (60, 62, 64, 68, 70 y 80): 62
Tiempo de Coccion en Minutos (72,84,87,90,93,132,135,138,147,189,198,210,252): 93
Tipo de Liquido de Cobertura del Producto Final (Aceite = 10 y Agua = 20): 10
Destino Final del Atun en el Producto Terminado (Lata = 10, Pouch = 20, Lomo = 30): 10
Numero de la Cocina a Utilizar (6,7,11,14,15,16,18,19,20,21, del 26 al 31)): 11
Rendimiento Final (70,71, (73 al 90)): 73
Según KNN: ✗ Rendimiento Bajo
Según Bayes: ✗ Rendimiento Bajo
Según LDA: ✗ Rendimiento Bajo
Según QDA: ✗ Rendimiento Bajo
Según Árbol: ✗ Rendimiento Bajo
Según SVM: ✗ Rendimiento Bajo
Según Random Forest: ✗ Rendimiento Bajo
Según Logística: ✗ Rendimiento Bajo
Según Gradient Boosting: ✗ Rendimiento Bajo
Según AdaBoost: ✗ Rendimiento Bajo

--- Decisión por mayoría ---
✗ Rendimiento Bajo (10/10 modelos)

```

## Ejercicio Numero 2:

Los resultados de este segundo ejercicio, 9 de 10 modelos de Machine Learning, generaron una predicción viable de obtener un alto rendimiento en el proceso de cocción del atún. Específicamente, en la combinación de la especie Skipjack con una talla mediana de 6 lbs, una temperatura de corte de 64 grados centígrados y un tiempo de cocción de 85 minutos, el cual se correlaciona fuertemente con el rendimiento esperado del 85%.

Esta mejora se puede traducir en aproximadamente \$100.000 al menos de mayor margen y utilidad bruta para la empresa.

Por otro lado, el modelo QDA (Análisis Discriminante), dio como resultado el rechazo de la combinación de datos suministrada por el usuario.

## Datos Ingresados por el Usuario y Resultado:

```

☞ Turno (10 = Turno Uno, 20 = Turno Dos, 30=Turno Tres): 20
Lbs por Pieza de Atun (desde 2 lbs hasta 50 lbs): 6
Especie de Atun (SKIP JACK = 10, YELLOW FIN = 20, BIG EYE = 30): 10
Talla del Atun (Pequeña = 10, Mediana = 20, Grande = 30, Corte = 40): 20
Barcos del 1 al 26: 6
Eviserado = 10 y Congelado =20: 10
Temperatura de Corte de Coccion (60, 62, 64, 68, 70 y 80): 68
Tiempo de Coccion en Minutos (72,84,87,90,93,132,135,138,147,189,198,210,252): 93
Tipo de Liquido de Cobertura del Producto Final (Aceite = 10 y Agua = 20): 10
Destino Final del Atun en el Producto Terminado (Lata = 10, Pouch = 20, Lomo = 30): 10
Numero de la Cocina a Utilizar (6,7,11,14,15,16,18,19,20,21, del 26 al 31)): 19
Rendimiento Final (70,71, (73 al 90)): 89
Según KNN:  Rendimiento Adecuado
Según Bayes:  Rendimiento Adecuado
Según LDA:  Rendimiento Adecuado
Según QDA:  Rendimiento Bajo
Según Árbol:  Rendimiento Adecuado
Según SVM:  Rendimiento Adecuado
Según Random Forest:  Rendimiento Adecuado
Según Logística:  Rendimiento Adecuado
Según Gradient Boosting:  Rendimiento Adecuado
Según AdaBoost:  Rendimiento Adecuado

--- Decisión por mayoría ---
 Rendimiento Adecuado (9/10 modelos)

```

## 7 Conclusiones y Trabajos Futuros

### 7.1 Conclusiones:

- Podemos concluir que soportados por los sistemas de información basado en algoritmos de Machine Learning, tienen un gran potencial para ser una herramienta para considerar en la predicción de impactos futuros en el proceso de cocción del atún, esto gana mayor relevancia considerando que el atún como materia prima tiene un peso en el costo total de fabricación de los productos de enlatados del 52%, en productos pouch 62% y en lomos congelados 80%. Esto implica que estamos enfocándonos en predecir con anticipación los impactos en rendimiento del componente que mas pesa en el costo de producción de atún.
- Los modelos de clasificación desarrollados e implementados, puntualmente hablando del modelo de Decision Tree y Random Forest, evidenciaron un alto nivel de precisión del 98%, lo cual soporta su aplicabilidad para predecir los rendimientos antes de poner en marcha el proceso productivo de cocción del atún. Permitiendo que la compañía tome decisiones y poder corregir proactivamente con base en datos, antes de iniciar con el arranque de la producción.
- Estas modelaciones evidencio una relación fuerte entre variables operativas y el rendimiento del atún después de su cocción, entre las principales correlaciones entre ellas se tienen la Talla del Atún (pequeña, mediana, grande y corte) y el tiempo de cocción, así también como la temperatura de corte de las cocinas y el destino final del producto terminado (Latas, Pouch o Lomos Congelados). Lo que permite realizar

configuraciones de los procedimientos de cocción por especie, talla y destino, aumentando de esta manera la eficiencia del proceso productivo de cocinar el atún.

- Se pudo identificar que las temperaturas de corte de las cocinas combinada con los tiempos de cocción con mucho tiempo, se tiende a conseguir un rendimiento muy bajo de atún, es recomendable e importante contar con un buen control eficiente y dinámico de los parámetros de las variables anteriormente mencionadas, esto se logra gracias a los modelos de Miche Learning analizados en este trabajo.
- La implementación y adherencia en el proceso productivo de estos modelos en el piso de planta pueden significar mejoras grandes en la rentabilidad de la compañía, bajando un punto porcentual de merma representa ahorros estimados de \$1 millón de dólares mensualmente de acuerdo con cálculos realizados por el departamento de Costos de la compañía. En los modelos que se corrieron en este trabajo, se evidenciaron combinaciones optimas que mejoran el margen y utilidad bruta hasta de \$100.000 USD en una jornada de producción.
- Este modelo la gran bondad es que nos anticipa con resultados de rendimientos bajos, sino que también da la oportunidad de realizar varias simulaciones a otros escenarios, ajustando variables como talla del pescado, temperaturas de cocción y tiempos de corte de las cocinas y llegar a un ideal a nivel de resultado final de rendimiento altos. Lo que permite construir el número y no esperar el resultado.

### 7.1.2 Trabajos Futuros

- Para fortalecer los modelos de clasificación de Machine Learning es recomendable incluir al menos uno o dos años de datos históricos, de esta manera robustecer aún más los modelos para predecir los rendimientos del atún en el proceso de cocción.
- Volver más dinámica y de actualización en tiempo real de los datos origen, en columnas adicionales incluir los valores que actualmente se manejan como texto colocar su homólogo en numérico, por ejemplo: Barco Atlantic = 1, Barco Caribe = 2, de esta manera no afectar la data origen de la compañía. Con la homologación en línea se puede desarrollar un vínculo con el modelo y la data real de la producción.

## Referencias Bibliográficas

Corporación universitaria tecnológica de bolívar facultad de ingeniería industrial  
Cartagena d.t. y c.

<https://repositorio.utb.edu.co/server/api/core/bitstreams/fd7ae81d-5388-443e-949a-137dd3a25660/content>

Universidad Politécnica Salesiana, Carrera de ingeniera industrial. Proyecto  
técnico.

<https://dspace.ups.edu.ec/handle/123456789/18712>

Python Machine Learning, Vahid Mirjalili, Sebastian Raschka, segunda edición  
2019. Marcombo, S.A.