Proyecto Práctico

Detección temprana de glaucoma con IA

**Imagen diseñada por ChatGPT\***

Autores:

|  |  |  |
| --- | --- | --- |
| **GRUPO** | **NOMBRE** | **CORREO** |
| **Grupo 3** |  |  |
|  |  |
|  |  |
|  |  |

Tabla de  
contenido

Contexto 1

Objetivo 1

Dataset…………………………………………………………..2

Características principales del dataset**……………………..………………..2**

Explicación de parámetros del dataset**……………………..…………..…4**

**Pasos de desarrollo del modelo en Knime…………....5**

**Anexos……………………………………………..……………..8**

# Contexto Glaucoma

El glaucoma es una enfermedad ocular que afecta al nervio óptico y representa una de las principales causas de ceguera irreversible en el mundo. Según la Organización Mundial de la Salud (OMS), millones de personas viven con glaucoma, muchas de ellas sin saberlo, ya que esta enfermedad es asintomática en sus primeras etapas.

El diagnóstico temprano es crucial para ralentizar la progresión de la enfermedad mediante tratamientos adecuados. Sin embargo, debido a la complejidad y variabilidad de los datos clínicos, la detección temprana sigue siendo un desafío en muchos casos.

Las técnicas de aprendizaje automático (ML) han demostrado ser efectivas en problemas de clasificación médica, al permitir la identificación de patrones en datos complejos que pueden no ser evidentes para el ojo humano. Estas herramientas pueden ayudar a los especialistas a priorizar pacientes y mejorar la eficiencia del diagnóstico.

## Objetivo

El objetivo de este proyecto es construir un modelo de clasificación basado en algoritmos de aprendizaje automático (ML) para predecir el riesgo de glaucoma utilizando un dataset sintético de datos clínicos. El modelo se plantea como una herramienta de apoyo para optimizar el diagnóstico temprano y mejorar la toma de decisiones en entornos médicos.

Para alcanzar este objetivo, se han empleado algoritmos de aprendizaje supervisado como los **árboles de decisión** y los **random forests**. Los árboles de decisión son modelos interpretativosinterpretables que permiten analizar el proceso de clasificación de manera clara, lo que los hace particularmente útiles en contextos médicos, donde la comprensión del razonamiento detrás de las predicciones es crucial. Por otro lado, los random forests, al combinar múltiples árboles de decisión mediante técnicas de ensamble, ofrecen una mayor robustez y precisión al reducir el riesgo de sobreajuste y mejorar la capacidad del modelo para generalizar sobre datos no vistos.

Gracias al uso de estas técnicas, el proyecto busca no solo proporcionar predicciones fiables, sino también ofrecer una herramienta práctica y eficaz para los profesionales de la salud en la lucha contra el glaucoma.

**Dataset[[1]](#footnote-1)**

El dataset utilizado en este proyecto es de origen sintético[[2]](#footnote-2), diseñado para simular datos clínicos relacionados con la salud ocular. Este conjunto de datos contiene información detallada sobre 1,000 pacientes y 12 variables que incluyen características demográficas, mediciones clínicas y datos categóricos.

**Características principales del dataset**

* **Número de registros:** 1,000.
* **Número de variables:** 12.
  + **Variables demográficas:** Edad, Género, Grupo de Edad.
  + **Variables clínicas clave:**

**Explicación**

1. **Intraocular Pressure o Presión intraocular:** Es el indicador más importante en el diagnóstico temprano de glaucoma, se interpreta así:

**Presión muy Baja**: < 7 mlt/mercurio | Hipotonía, no hay tono y fuerza, anomalía severa, ceguera

**Presión Baja**: (de 7 a 9.99) < 10 mlt/mercurio | Sospecha de anomalía

**Presión normal**: de 10 a 21 mlt/mercurio

**Presión alta**: > 21 mlt/mercurio | Hipertensión ocular, sospecha de glaucoma: causa daño al nervio óptico, pérdida de visión

Cuando un paciente tiene diagnóstico de glaucoma, se deben realizar unas metas de presión para llevar la **Presión intraocular** a > 17mlt/mercurio para darle seguridad al ojo de contraer glaucoma.

1. **Pupil Diameter o diámetro de pupila**: Se mide en milímetros (mm) Este parámetro nos dice que una pupila está dilatada cuando > 5,5 mm

Es un tercersegundo indicador que posiblemente hay un glaucoma avanzado.

< 2,5 mm Pupila poco dilatada

2,5 a 5,5 mm de diámetro pupilar normal

 > 5,5 mm Pupila dilatada

**Ejemplo**: Cuando un paciente tiene presión ocular alta, entre mayor sea la presión mayor es la dilatación de la pupila, lo cual representa un factor pronóstico de problema agudo y glaucoma grave 🚨 Urgencia 🚑

1. **Central Corneal Thicke o Grosor corneal central:** Este indicador está relacionado con la medición de la presión intraocular.

Dice sobre el grosor del centro de la córnea, se mide en micras. Este parámetro se puede relacionar con la presión intraocular como factor de riesgo (Una córnea delgada no tiene buena resistencia al flujo del humor acuoso, como no hay resistencia, el líquido genera daños intraoculares y se aumenta la presión intraocular)

Es un tercer indicador para tener en cuenta, debido a que es un factor de riesgo para que se aumente la presión y genere se glaucoma o lo empeore. Se mide en micras:

**Rango bajo**: < 500 micras Factor de riesgo para que el paciente desarrolle alta presión intraocular

**Rango normal**: 500 - 600 micras (0,001mm)

**Rango alto**: > 600 micras Grosor de cornea fuerte

**Ejemplo**: Paciente con 21 mlt/mercurio (límite de presión alta) y un rango bajo de medida del grosor de la córnea,  presentaría dos indicadores de riesgo para glaucoma y alta presión intraocular.

1. **Visual Acuity logMAR  o Agudeza visual (logMAR):** Métrica de calidad visual.

Funciona como un cuarto indicador clave para la detección de glaucoma, en la práctica médica se mide en **Snelle ->** fracción 20/20 (ve claramente a 20 pies) a medida que aumenta el denominador quiere decir que ve menos (20/30)

**Baja visión**: denominador mayor a 20/60

**Mejor visión**:  denominador menor a 20/20

**Ejemplo**: Un paciente que ve a 20/40 significa que lo que una persona saludable debería ver a 40 pies, el paciente lo ve a 20 pies (no ve lo que está lejos, se debe acercar)

La conversión de **Snelle** (Fracción) a decimal es el valor de agudeza visual **logMAR**

**Explicación de parámetros del dataset[[3]](#footnote-3)**

1. **Age** (Edad): Muestra la edad de los pacientes
2. **Gender** (Sexo): No relevante para esta patología
3. **Eye** (Ojo): **OS**: Ojo izquierdo; **OD**: Ojo derecho
4. **Refractive Error**: + info ℹ️
5. **Intraocular Pressure**:  Indicador principal para la sospecha de glaucoma**, el** humor acuoso se produce en el ojo y según influye puede disminuir o aumentar la presión intraocular. “Sirve para nutrir y oxigenar las estructuras del globo ocular que no tienen aporte sanguíneo, como la córnea y el cristalino.”
6. **Central Corneal Thicke**: Dice sobre el grosor del centro de la córnea, se mide en micras. Este parámetro se puede relacionar con la presión intraocular como factor de riesgo (Glaucoma - Una córnea delgada no tiene buena resistencia al flujo del humor acuoso, como no hay resistencia, el líquido genera daños intraoculares y se aumenta la presión )

Segundo indicador Es un factor de riesgo para que se aumente la presión y genere se glaucoma

1. **Pupil Diameter**: Se mide en milímetros (mm) Este parámetro se puede relacionar que una pupila está dilatada cuando > 5,5 mm

Tercer indicador que posiblemente hay un glaucoma avanzado

1. **Visual Acuity logMAR**

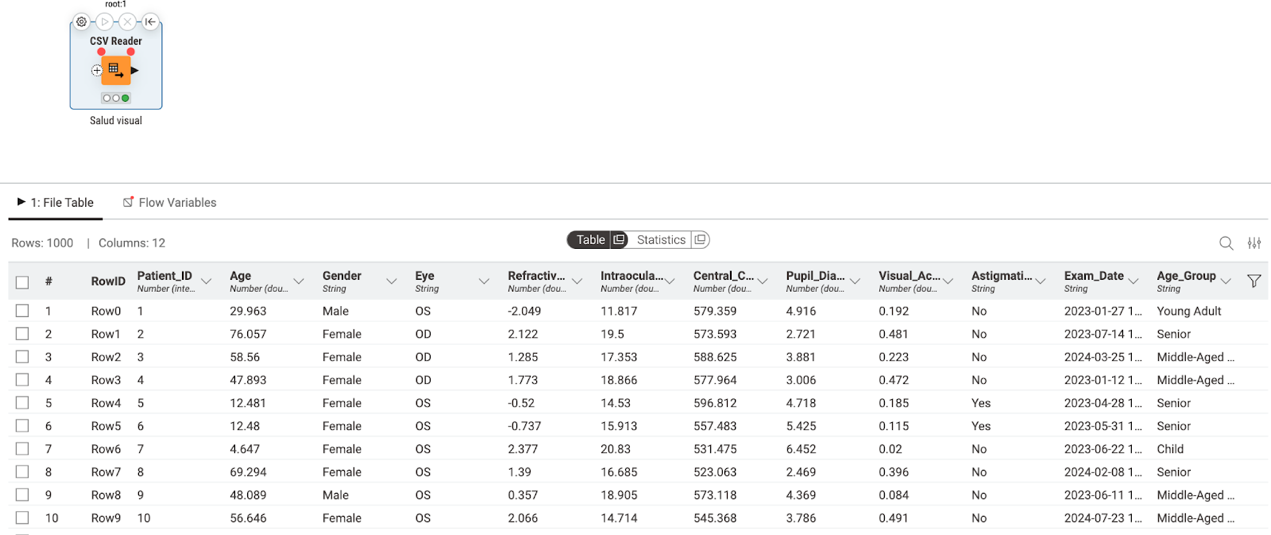
En la práctica médica se mide en **Snelle ->** fracción 20/20 (ve claramente a 20 pies) a medida que aumenta el denominador quiere decir que ve menos (20/30)

1. **Astigmatismo**: No tiene relevancia para deste estudio
2. **Exam Date**: Fecha de realización del examen.

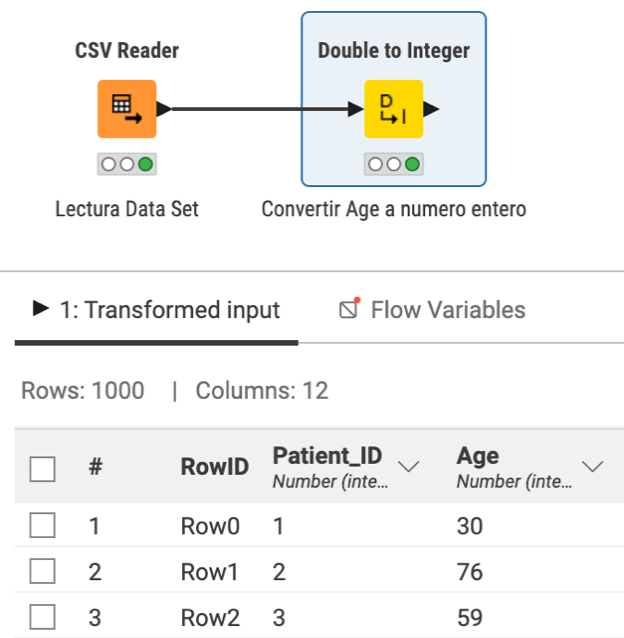
**Age Group** (grupo de edad): Clasificado como: Infante, niño, senior, joven, joven adulto.

**Pasos de desarrollo del modelo en Knime[[4]](#footnote-4)**

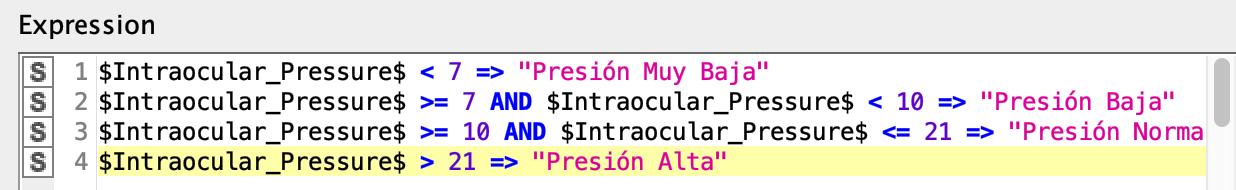
1. Lectura del archivo en un **CSV Reader**



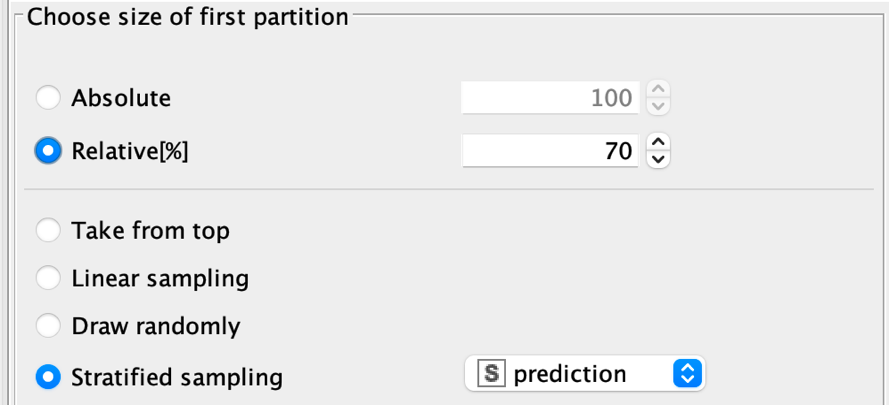
1. Convertir la columna Age de número flotante a entero **Doble to Integer**

****

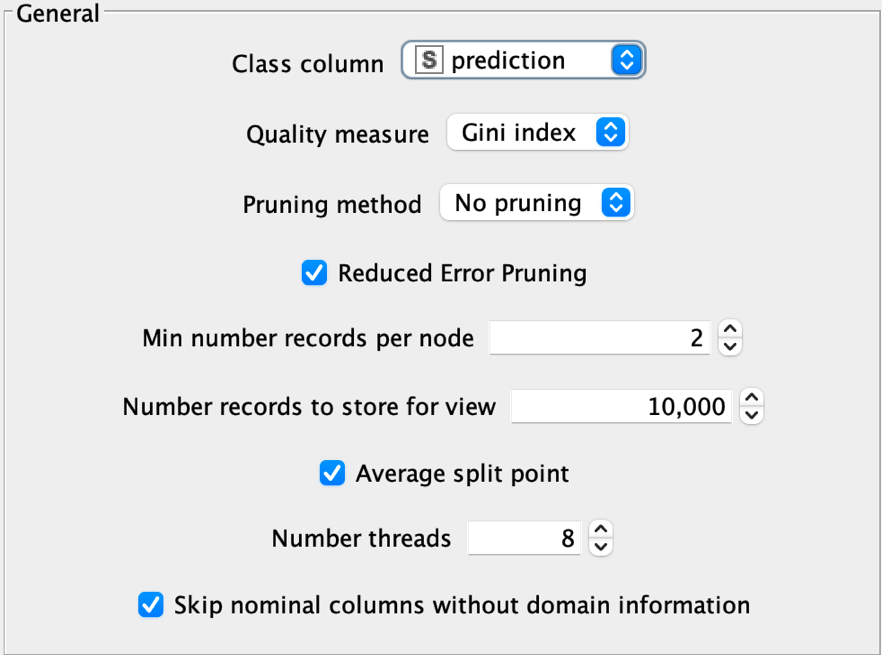
1. Definir criterios del modelo en un **Rule Engine**



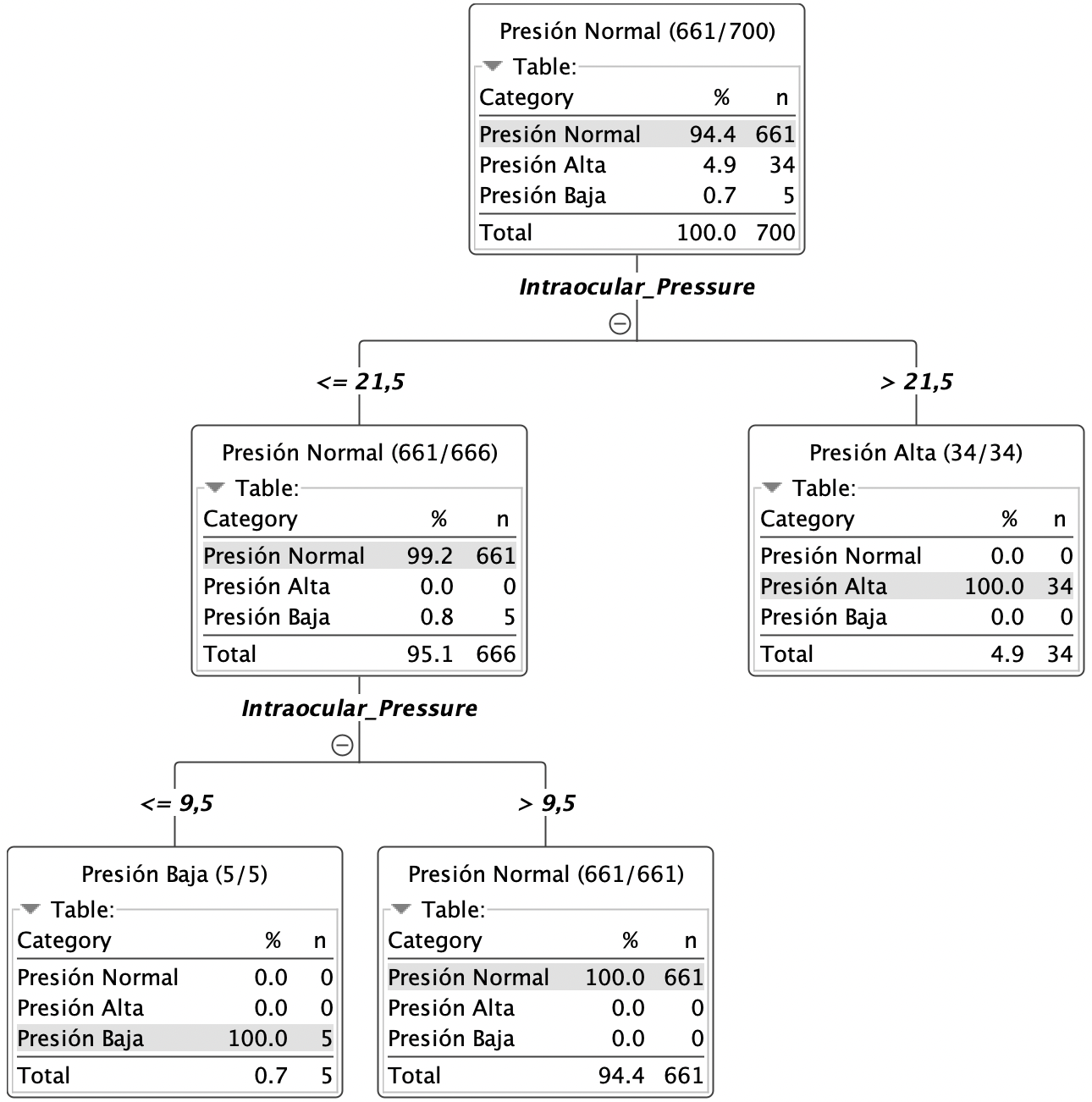
1. Particionar los datos 70% para modelado y 30% para entrenamiento en **Partitioning**

****

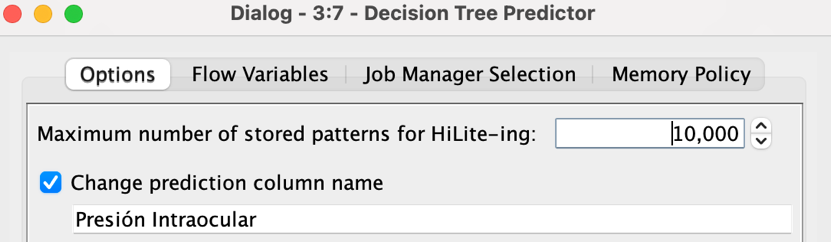
1. Configuración del modelo en **Decision Tree Learner** con Gini index

****

**5.1** Árbol de decisión

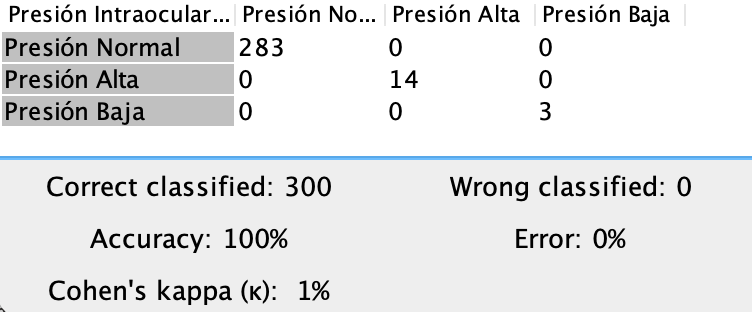
****

1. **Decisión Tree Predictor**

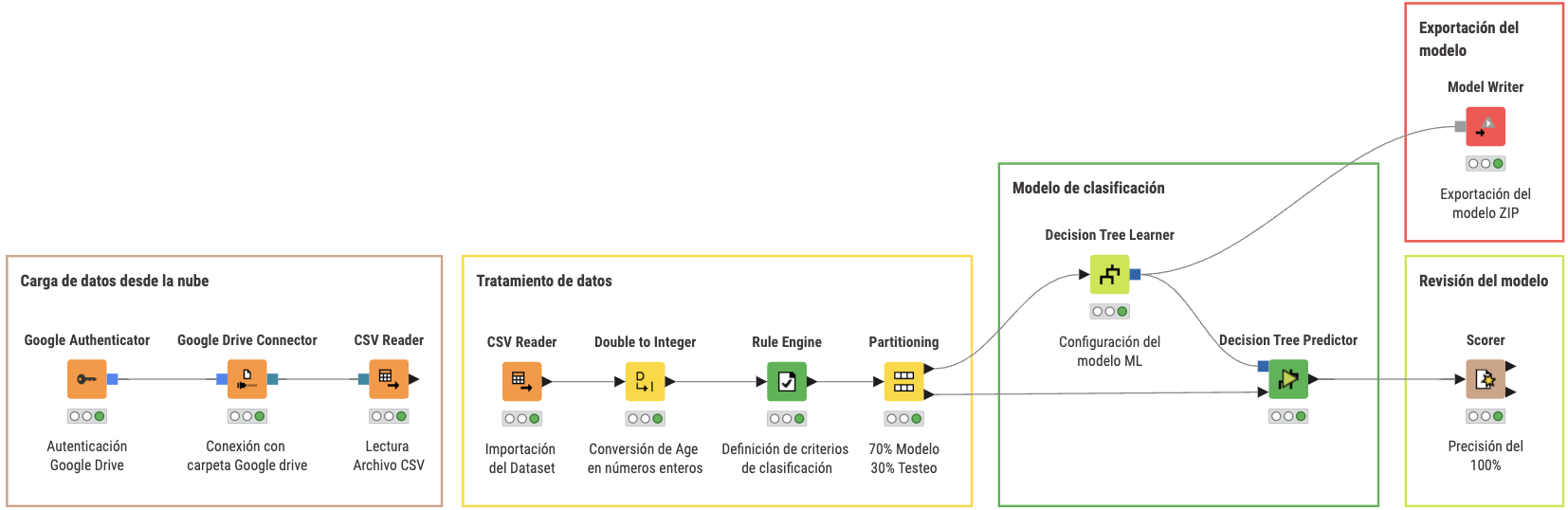


1. Revisión del modelo en **Scorer**

**7.1** Matrix de confusión

****

1. Exportación del modelo en **Model Writer ZIP**

**Captura WorkFlow**

**Anexos**

1. **Referencia del Dataset:**

<https://github.com/doehring-gh/eye-health-dataset-generator>

1. **Link del Dataset:**

https://drive.google.com/drive/folders/1XLj-cYdZCBMJ7q0nt\_PQm-0mtZ7MaMNN?usp=sharing

1. **Video de explicación de parámetros del Dataset:**

<https://youtu.be/TWzVI2IrKQg>

1. **Link Workflow Knime**

<https://hub.knime.com/ialexo7/spaces/Public/Proyecto%20Pr%C3%A1ctico~N91NV2vbae9kr1LC/current-state>

1. **Link Cuaderno de Colab**

1. **Anexo: Referencia del Dataset**  [↑](#footnote-ref-1)
2. **Anexo: Link del Dataset** [↑](#footnote-ref-2)
3. **Anexo:** **Video de explicación de parámetros del Dataset** [↑](#footnote-ref-3)
4. **Anexo: Link Workflow Knime** [↑](#footnote-ref-4)