

Manual de Usuario - Causal Inference Shadow Detection

Versión 1.0

1. Introducción

Causal Inference Shadow Detection es un sistema desarrollado en Python con Streamlit, diseñado para simular entornos observacionales, descubrir estructuras causales y realizar inferencia mediante regresión estadística y do-calculus.

Este entorno permite a investigadores y estudiantes explorar la relación causal entre variables como Luz (A), Esfera (B), Superficie (C) y Sombra (Y), con aplicaciones en visión por computador, robótica y física computacional.

La interfaz guía al usuario a través de cuatro etapas clave: generación de datos, descubrimiento causal, comparación con modelo de referencia e inferencia cuantitativa.

2. Requisitos del Sistema

Antes de ejecutar la aplicación, asegúrese de cumplir con los siguientes requisitos:

- Sistema Operativo: Windows 10/11, macOS o Linux (Ubuntu 20.04+ recomendado).
- Python: Versión 3.12 o superior.
- Memoria RAM: Mínimo 4 GB (8 GB recomendado para datasets grandes).
- Espacio en Disco: 500 MB libres.

Dependencias:

Todas las librerías necesarias están listadas en el archivo Requirements.txt y pueden instalarse con el comando:

```
pip install -r Requirements.txt
```

Librerías clave incluidas: streamlit, pandas, numpy, causalinferenc, pgmpy, networkx, matplotlib, scikit-image, opencv-python.

3. Instalación Paso a Paso

Siga estos pasos para configurar y ejecutar la aplicación localmente:

3.1 Descarga del Proyecto

Descargue el repositorio completo desde GitHub o extraiga el archivo ZIP en una carpeta local de su preferencia. El enlace para descarga es: https://github.com/jvelez-dev/Causal_Inference_Shadow_Detection.git

3.2 Crear y Activar Entorno Virtual (Recomendado)

Ejecute los siguientes comandos en su terminal o símbolo del sistema:

```
python -m venv venv
```

Activar en Windows:

```
venv\Scripts\activate
```

Activar en macOS/Linux:

```
source venv/bin/activate
```

3.3 Instalar Dependencias

Con el entorno virtual activado, instale todas las dependencias necesarias:

```
pip install -r Requirements.txt
```

3.4 Ejecutar la Aplicación

Inicie la interfaz de usuario con el siguiente comando:

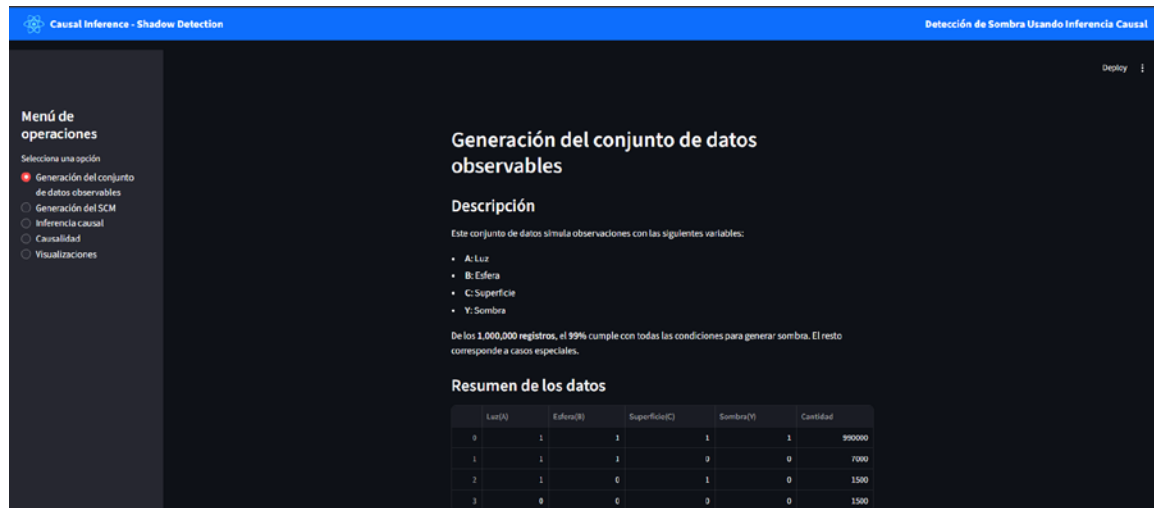
```
streamlit run app/interface.py
```

La aplicación se abrirá automáticamente en su navegador predeterminado en la dirección:
<http://localhost:8501>

4. Uso de la Aplicación

La interfaz está organizada en secciones secuenciales. A continuación, se detalla cada una:

4.1 Generación de Datos



Menú de operaciones

Seleccione una opción

- ☒ Generación del conjunto de datos observables
- ☐ Generación del SCM
- ☐ Inferencia causal
- ☐ Causalidad
- ☐ Visualizaciones

Generación del conjunto de datos observables

Descripción

Este conjunto de datos simula observaciones con las siguientes variables:

- A: Luz
- B: Esfera
- C: Superficie
- Y: Sombra

De los 1,000,000 registros, el 99% cumple con todas las condiciones para generar sombra. El resto corresponde a casos especiales.

Resumen de los datos

Luz(Y)	Esfera(B)	Superficie(C)	Sombra(Y)	Cantidad
0	1	1	1	990000
1	1	1	0	7000
2	1	0	1	1500
3	0	0	0	1500

En esta sección, el sistema genera automáticamente un conjunto de datos observacionales con 1000 registros, simulando las siguientes variables:

- A (Luz): Intensidad de la fuente lumínica.
- B (Esfera): Presencia (1) o ausencia (0) del objeto (variable de tratamiento).
- C (Superficie): Tipo de superficie (reflectante, absorbente, etc.).
- Y (Sombra): Presencia (1) o ausencia (0) de sombra (variable de resultado).

Se muestran:

- Tabla de frecuencias de combinaciones observadas.
- Gráficos de distribución de cada variable.
- Estadísticas descriptivas básicas (medias, desviaciones, etc.).

Este dataset es la base para todos los análisis causales posteriores.

4.2 Descubrimiento Causal

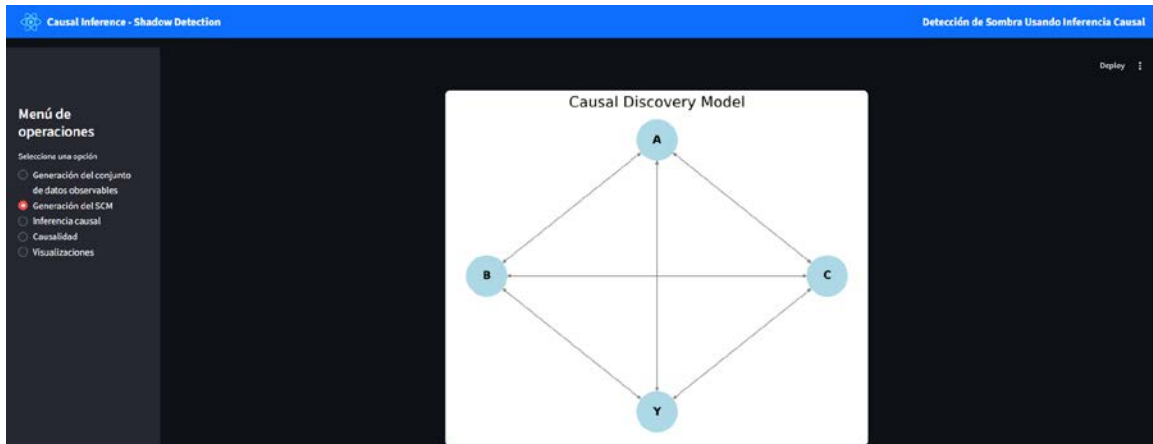


La aplicación implementa el algoritmo PC (Peter-Clark) para inferir la estructura causal subyacente a partir de los datos observados.

Resultados:

- Visualización interactiva de un Grafo Dirigido Acíclico (DAG) que representa las relaciones causales descubiertas.
- Lista de independencias condicionales encontradas (ej: $Y \perp\!\!\!\perp A \mid B, C$).

Este paso permite validar o refutar hipótesis previas sobre el sistema causal (por ejemplo: “La sombra es causada únicamente por la esfera”).



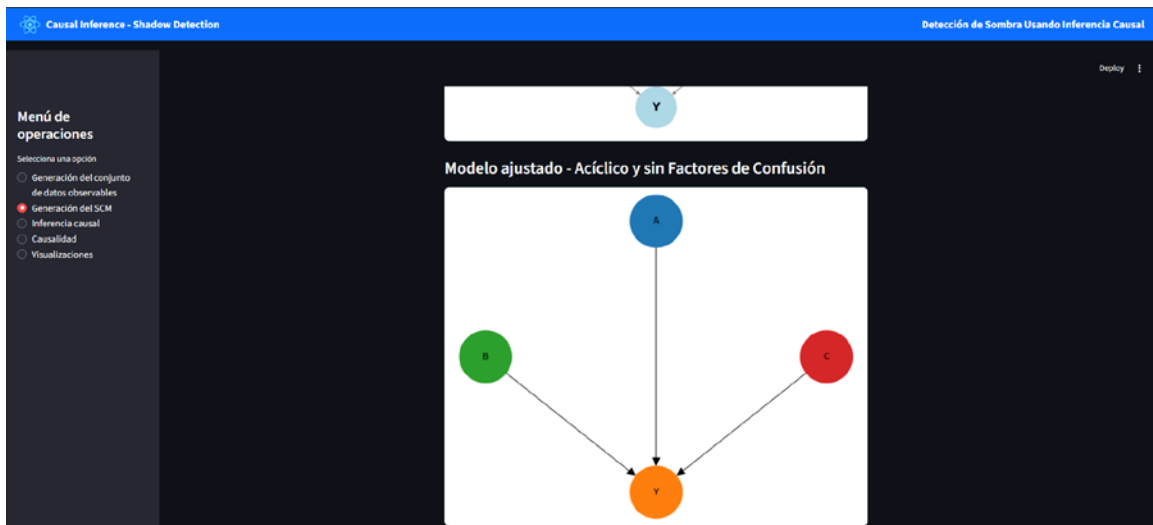
4.3 Modelo Causal Proporcionalado

Se muestra un modelo causal de referencia (ground truth) basado en conocimiento experto, con la siguiente estructura:

- Luz (A) → Superficie (C)
- Luz (A) → Sombra (Y)
- Esfera (B) → Sombra (Y)
- Superficie (C) → Sombra (Y)

El sistema compara este modelo con el descubierto automáticamente, mostrando:

- Diferencias estructurales en el DAG.
- Distribuciones de Probabilidad Condicional (CPDs) aprendidas para cada nodo.



4.4 Inferencia Causal

Existen dos modos complementarios de inferencia:

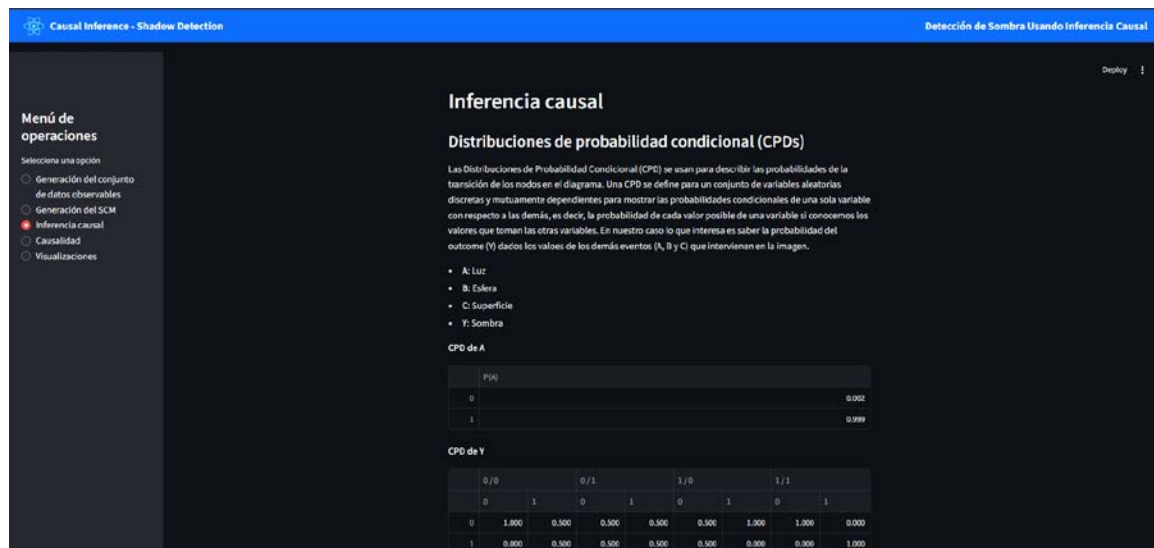
4.4.1 ATE (Average Treatment Effect)

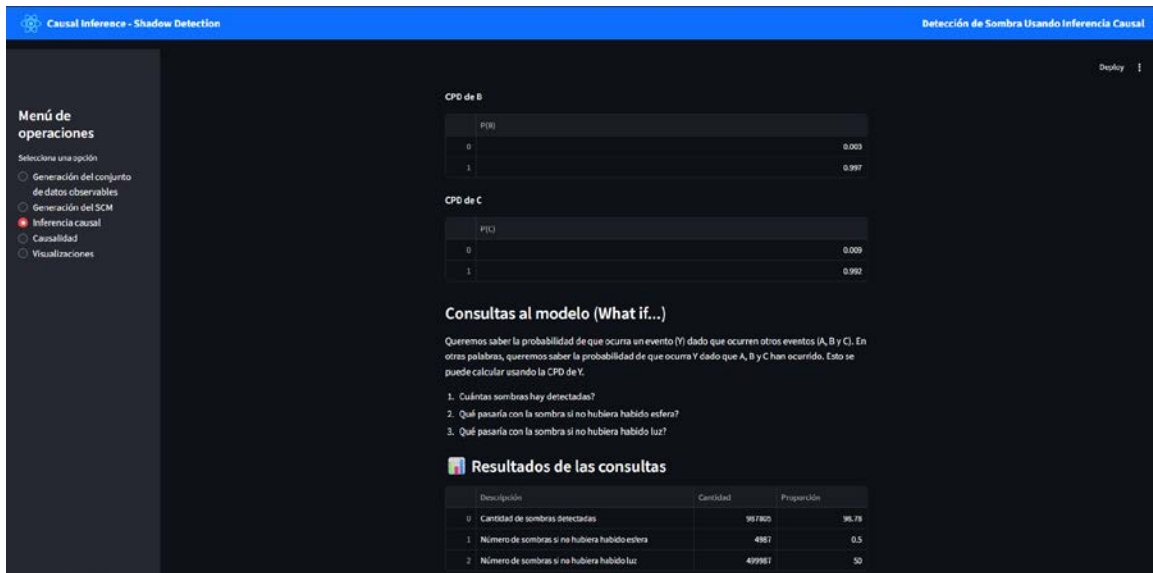
Calculado mediante regresión OLS (Mínimos Cuadrados Ordinarios) con la librería CausalInference.

Métricas disponibles:

- ATE (Average Treatment Effect): Efecto promedio del tratamiento (esfera) sobre el resultado (sombra).
- ATT (Average Treatment Effect on the Treated): Efecto en el grupo que recibió el tratamiento.
- ATC (Average Treatment Effect on the Controls): Efecto en el grupo de control.

Se muestra un resumen estadístico con errores estándar, intervalos de confianza y significancia.





4.4.2 Do-Calculus (Intervenciones)


Implementado con la librería pgmpy mediante eliminación de variables.

Permite realizar consultas contrafactuales del tipo:

- $do(B=0)$: ¿Cuál sería la probabilidad de sombra si no hubiera esfera?
- $do(B=1)$: ¿Cuál sería la probabilidad de sombra si sí hubiera esfera?

Los resultados se presentan en tablas interactivas con probabilidades condicionales completas.

Interpretación: Un ATE positivo y significativo, junto con un salto claro en $P(Y|do(B=1))$ vs $P(Y|do(B=0))$, confirma una relación causal robusta.


Causal Inference - Shadow Detection

DetECCIÓN DE SOMBRA USANDO INFERENCIA CAUSAL

Menú de operaciones

Selecione una opción

- Generación del conjunto de datos observables
- Generación del SCM
- Inferencia causal
- Causalidad**
- Visualizaciones

Causalidad

El escenario más común para aplicar la causalidad corresponde a un grupo de individuos dividido en 2 subgrupos: el primer grupo será el grupo que recibirá un tratamiento y, el segundo grupo, será el grupo de control que no recibirá tratamiento. Así se podrá inferir el efecto causado por el tratamiento. Esto implica que el ejemplo cambiará un poco de orientación:

El individuo será el escenario conformado por luz y superficie (estos siempre serán percibidos) El tratamiento será la esfera El efecto será la sombra

Average Treatment Effect (ATE)


El Average Treatment Effect (ATE) es una medida de causalidad que se utiliza para inferir el efecto causal de un tratamiento en una población. En nuestro caso, el tratamiento será la esfera y el efecto será la sombra. Un ATE positivo indica que la hipótesis planteada cobra fuerza y, en este caso, es altamente probable que la causa de la sombra sea la presencia de la esfera, es decir, que al aplicar el tratamiento B, el efecto obtenido en Y será el esperado.

Métrica	Valor
0 ATE	0.993

Estimaciones de Causalidad (OLS)

Treatment Effect Estimates: OLS					
	Est.	S.e.	z	P> z	[95% Conf. Int.]
ATE	0.993	0.008	118.74	0.000	0.993 0.993

Display

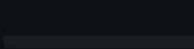

Causal Inference - Shadow Detection

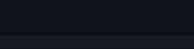
DetECCIÓN de Sombra Usando Inferencia Causal

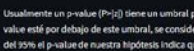
Menú de operaciones

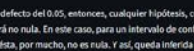
Selecione una opción


- Generación del conjunto de datos observables
- Generación del SCM
- Inferencia causal
- Causalidad**
- Visualizaciones










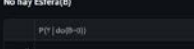


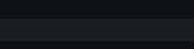


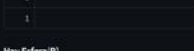


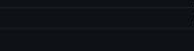





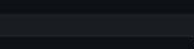


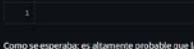


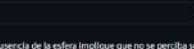


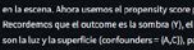


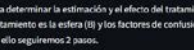
































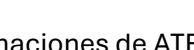








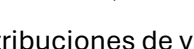






























































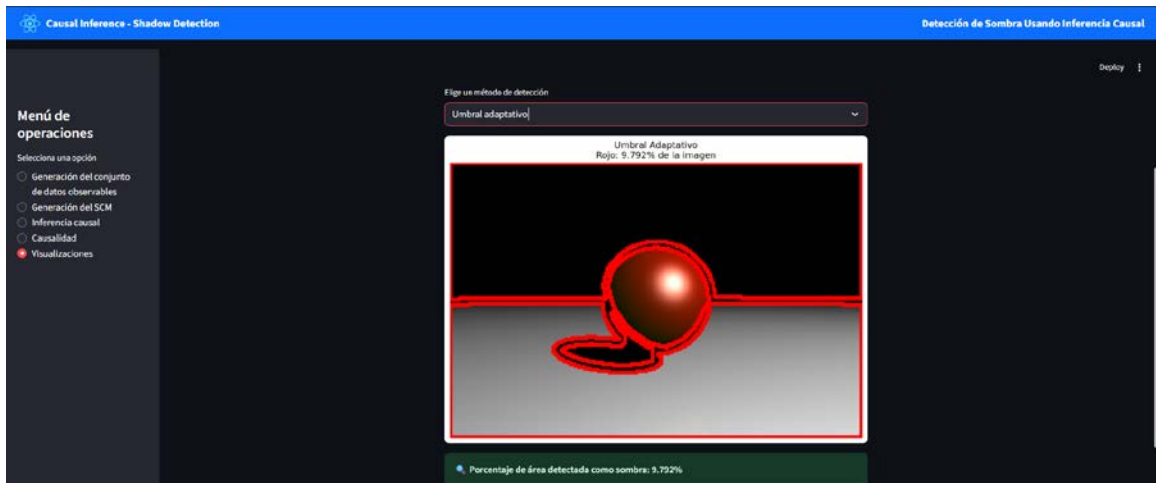
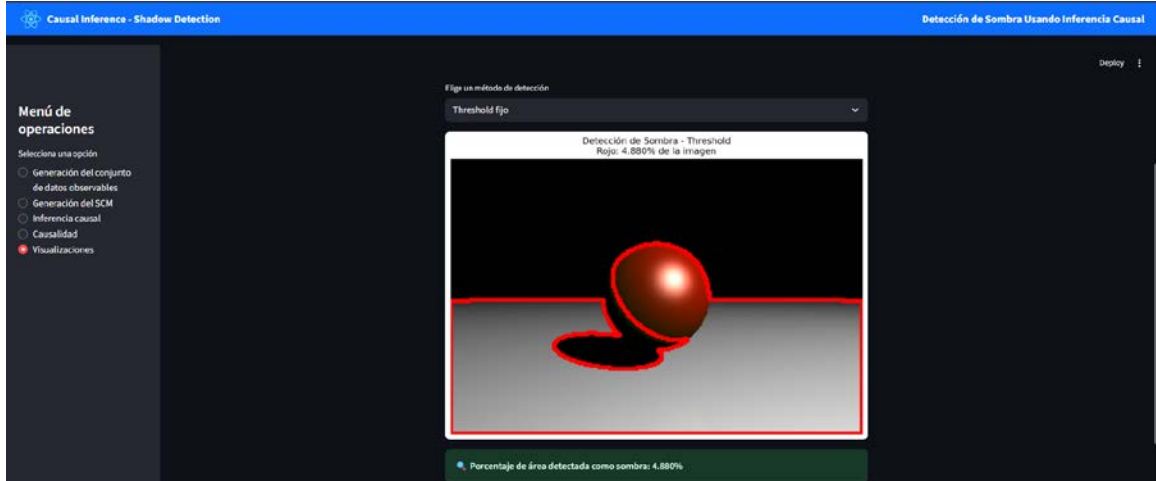


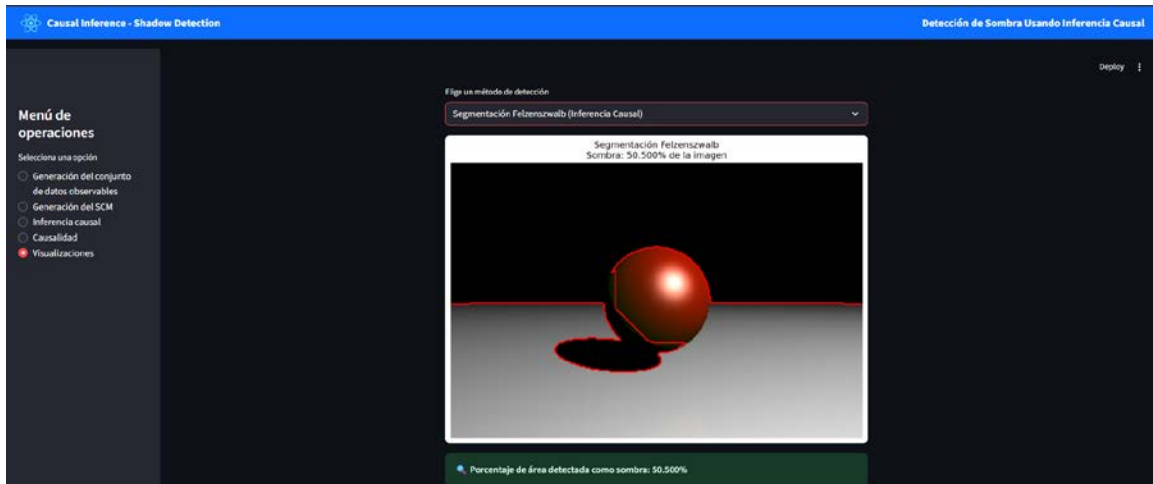


5. Resultados y Visualización

Todos los resultados se presentan de forma clara e interactiva:

- Tablas: Frecuencias, CPDs, estimaciones de ATE/ATT/ATC, resultados de do-calculus.
- Gráficos: DAG interactivo, gráficos de barras, distribuciones de variables.
- Textos explicativos: Interpretación cualitativa de los resultados obtenidos.





6. Pruebas y Validación

La carpeta tests/ incluye pruebas unitarias básicas para validar el funcionamiento del núcleo lógico (core.py).

Ejecución de pruebas:

Abra una terminal en la raíz del proyecto y ejecute:

```
pytest tests/
```

Las pruebas actuales validan:

- Generación correcta de datos.
- Cálculo de CPDs.
- Estimación de efectos causales (ATE, ATT, ATC).
- Funcionamiento de las intervenciones do-calculus.

Se recomienda ejecutar las pruebas tras cualquier modificación del código fuente.

7. Soporte y Solución de Problemas

Problemas Comunes:

- `ModuleNotFoundError`: Ejecute `pip install -r Requirements.txt` nuevamente.
- Error al cargar imagen: Asegúrese de que el formato sea PNG/JPG y no esté corrupto.
- `KeyError: 'att'`: Verifique que está usando `est_via_matching()` si necesita ATT/ATC.
- DAG no se muestra: Instale `graphviz` y asegúrese de tener `dot` en el PATH del sistema.

8. Licencia

Este software es de código abierto y se distribuye bajo la licencia MIT.

Si utiliza este sistema en su investigación, por favor cite:

Causal Inference Shadow Detection (v1.0). Autor: Velez-Bedoya J.I. Año 2025.