

# Cátedra ENIA

**IA3: Cátedra de Inteligencia Artificial en Aeronáutica y Aeroespacio**

**PT3: Gemelo Digital mediante simulación de sistemas aeronáuticos**

# Gemelos Digitales

## ¿Qué son?



Representación **virtual dinámica** de un **sistema físico** con intercambio **bidireccional** de datos con él.

# Gemelos Digitales

## Primeros acercamientos

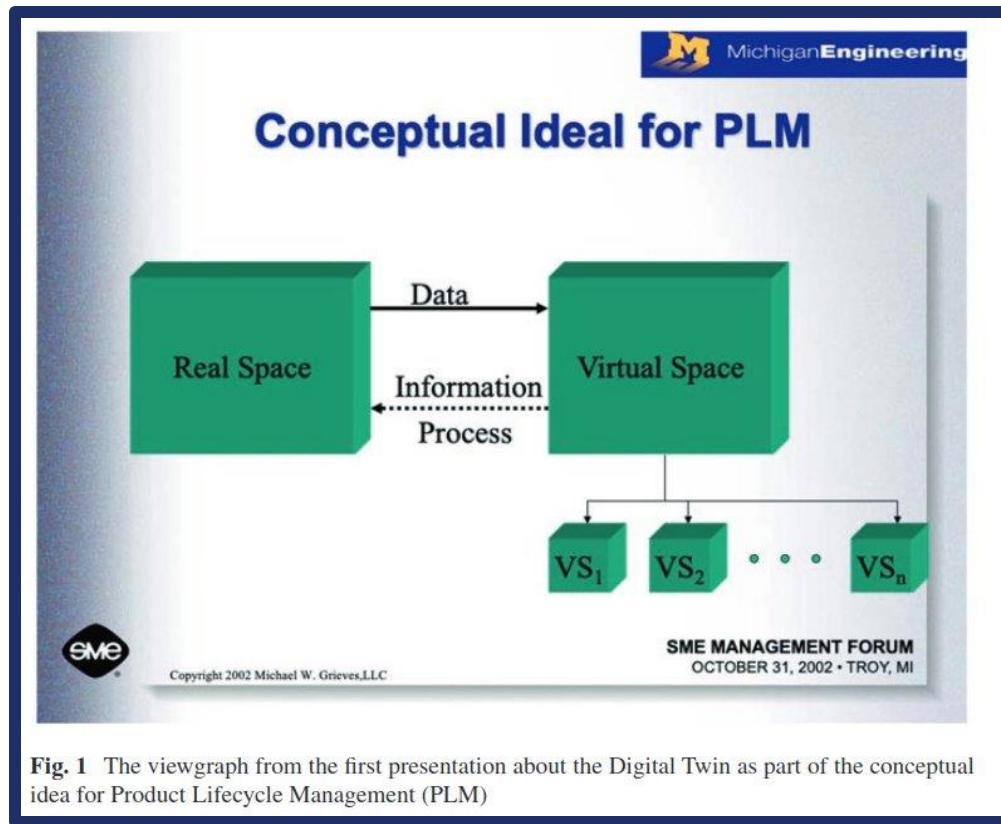


### Proyecto Apollo de la NASA:

- La **NASA** no podía probar los sistemas en el espacio.
- Crearon una **réplica física exacta** de la nave espacial para realizar pruebas en tierra.
- Debido a las similitudes con el concepto de gemelo digital, se considera un **Gemelo Físico**.

# Gemelos Digitales

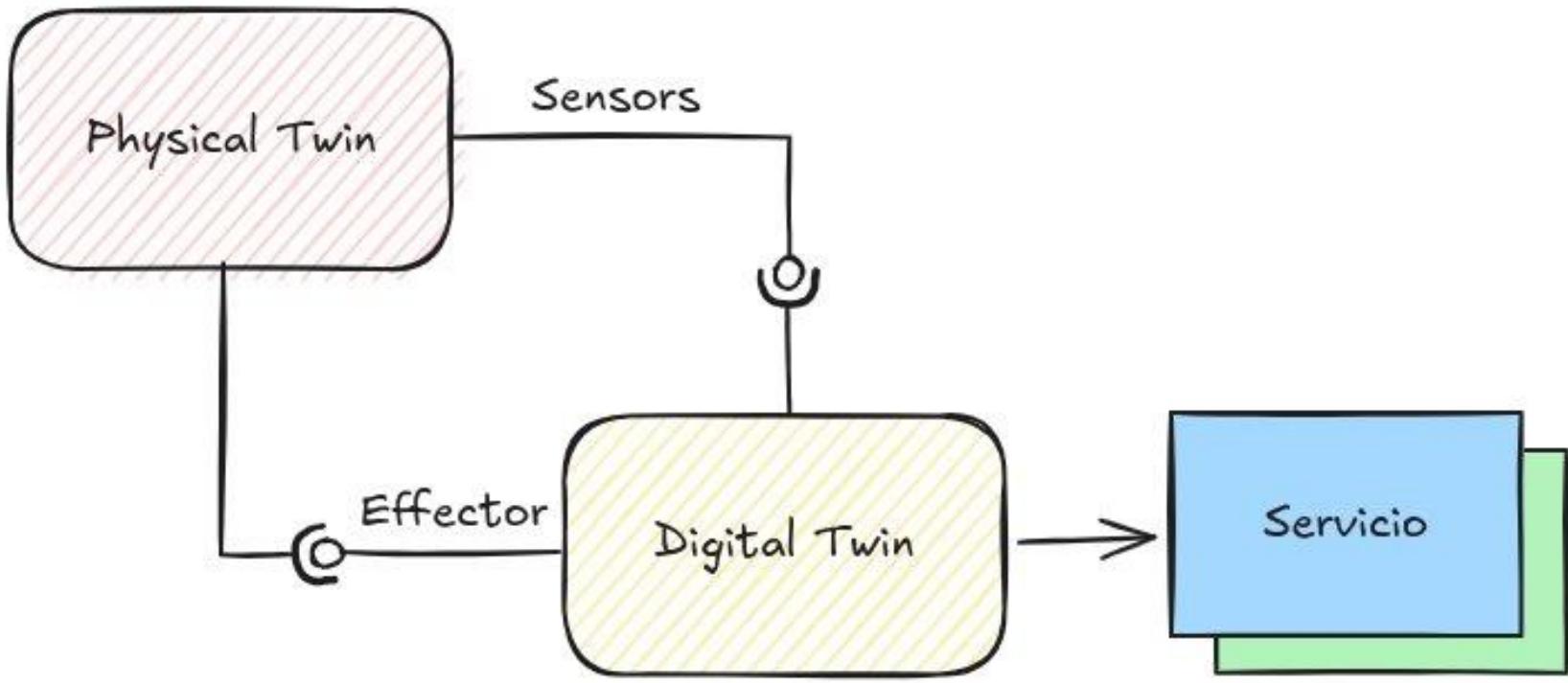
## Formalización



- Michael Grieves formalizó el concepto en 2002 en una conferencia de gestión del ciclo de vida del producto por
- En 2005 le pone nombres como "**Mirrored Spaces**"
- 2006 lo llama "**Digital Twin**" o "**Virtual Doppelganger**".

# Gemelos Digitales

## Definición



# Gemelo Digital en el sector aeroespacial

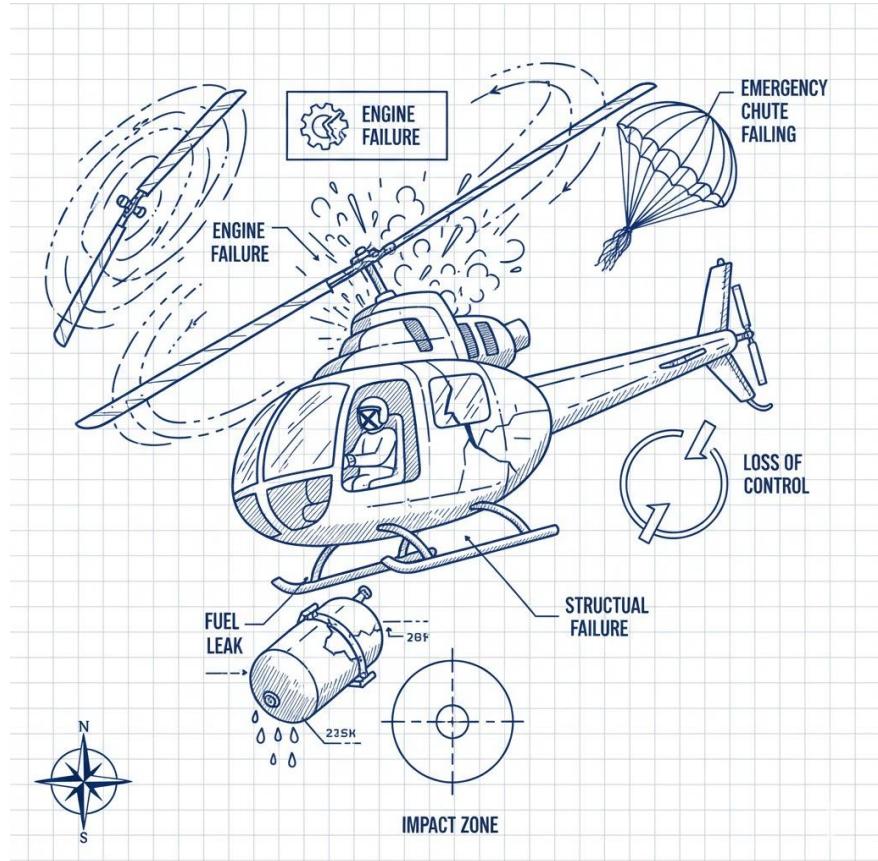
**Simular** miles de diseños

Monitorear en **tiempo real** el desgaste de una pieza para predecir cuando cambiarla

Someter a un **ala de avión** a un **huracán**

Hacer los sistemas más **seguros, eficientes** y **fiables**

# Falta de datos de fallos



- Sistemas **alta fiabilidad** y muchos **controles**
- Datos restringidos por cuestiones de **seguridad, confidencialidad o regulación**
- Sistemas **sin hardware** necesario
- Sistemas que **no registran** los datos históricos
  - Generación de datos sintéticos
  - Técnicas avanzadas (aprendizaje no supervisado)

# Tipos de Modelos simulación

## Modelos basados en Física

- **Funcionamiento:** Construido sobre ecuaciones fundamentales físicas
- **Ventajas:** Alta fidelidad, explicabilidad
- **Desventajas:** Requiere conocimiento del dominio, en ocasiones lento.

## Modelos basados en Datos

- **Funcionamiento:** Aprende patrones estadísticos de datos existentes

**Ventajas:** No se necesita un modelo matemático exacto, datos de ejecución. Velocidad

**Desventajas:** Caja negra, necesidad de muchos datos.

# Modelos basados en física

## MATLAB

### Simulink

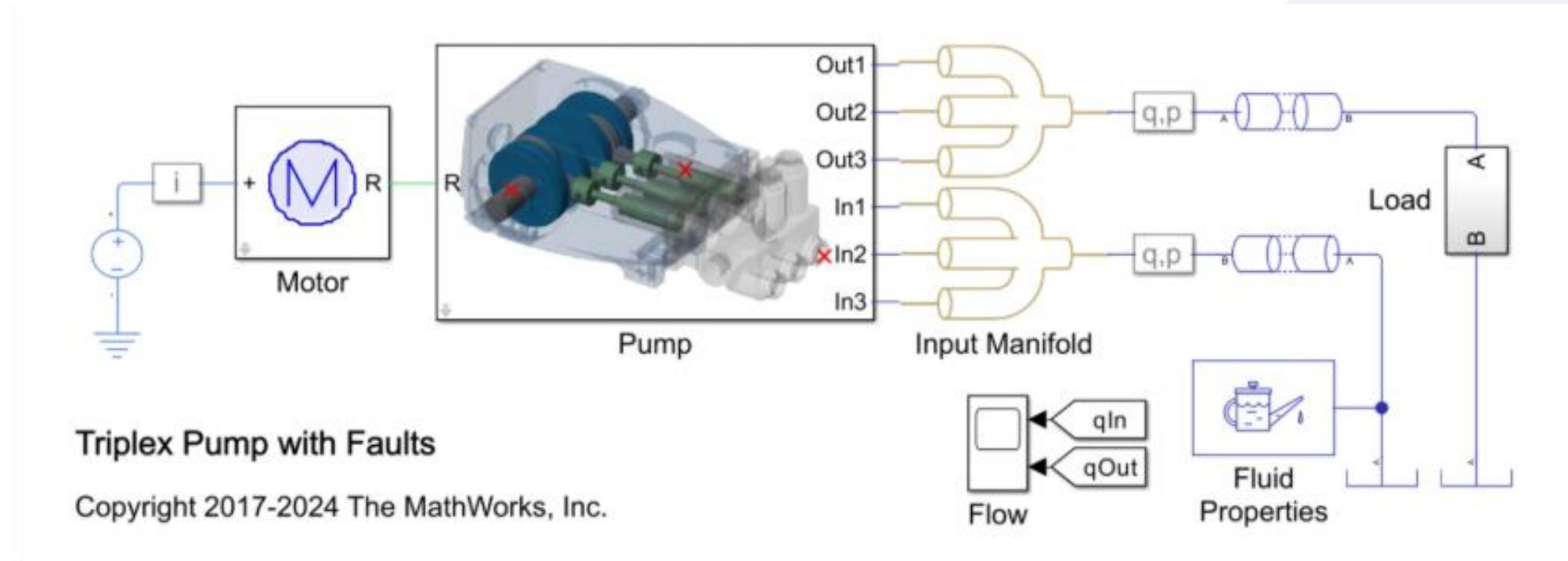
- Herramienta de **simulación** y **diseño** basada en **diagramas de bloques** para modelar, analizar y probar sistemas dinámicos multidominio.

### Simscape

- Permite crear rápidamente modelos de **sistemas físicos** dentro del entorno de Simulink
- Simplifica el modelado de sistemas físicos al permitir trabajar directamente con **componentes predefinidos** (resistencias, motores, válvulas, etc.) y ecuaciones físicas integradas en estos.

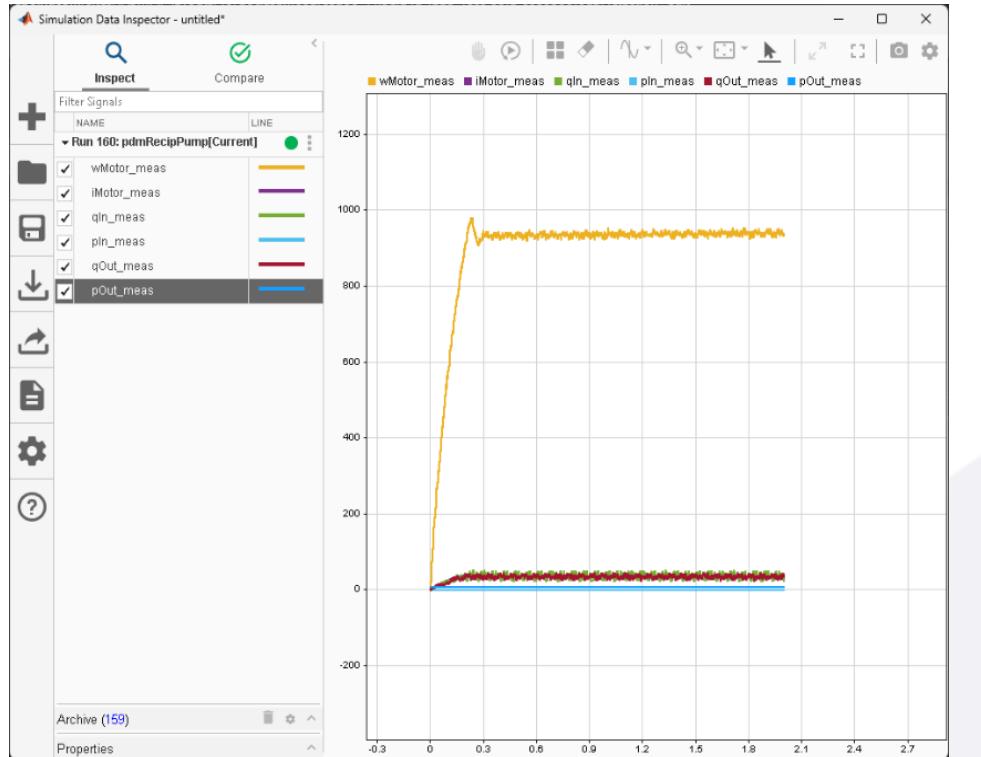
# Modelos basados en física

## Caso de uso bomba hidráulica

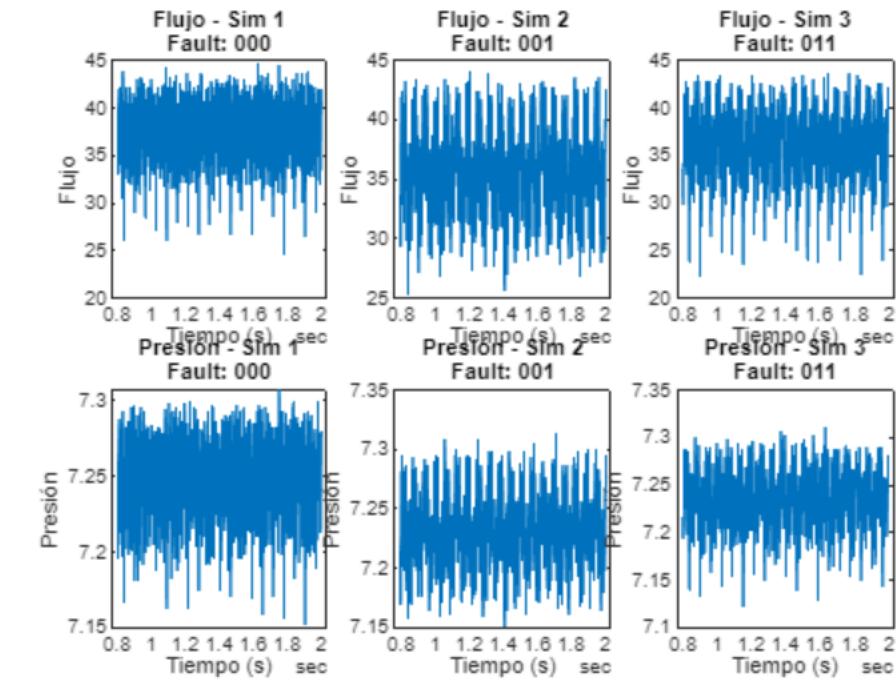


# Modelos basados en física

## Caso de uso bomba hidráulica



Visualización de los resultados procesados con fallos identificados



# Modelos basados en Datos

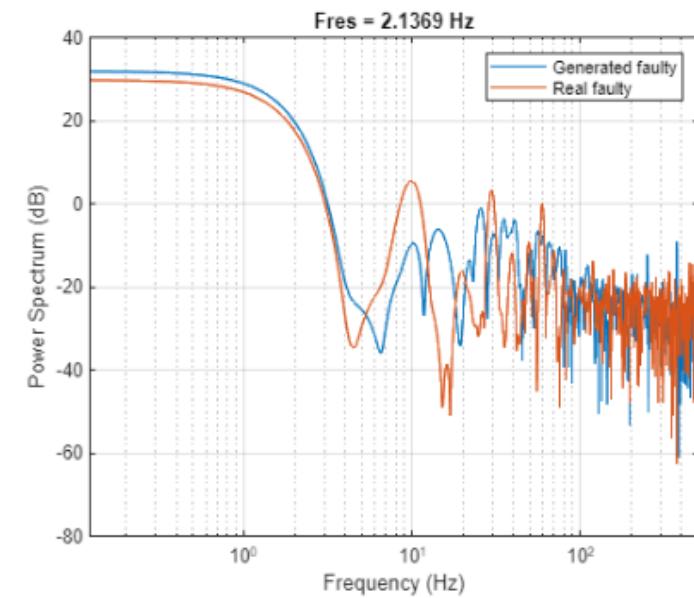
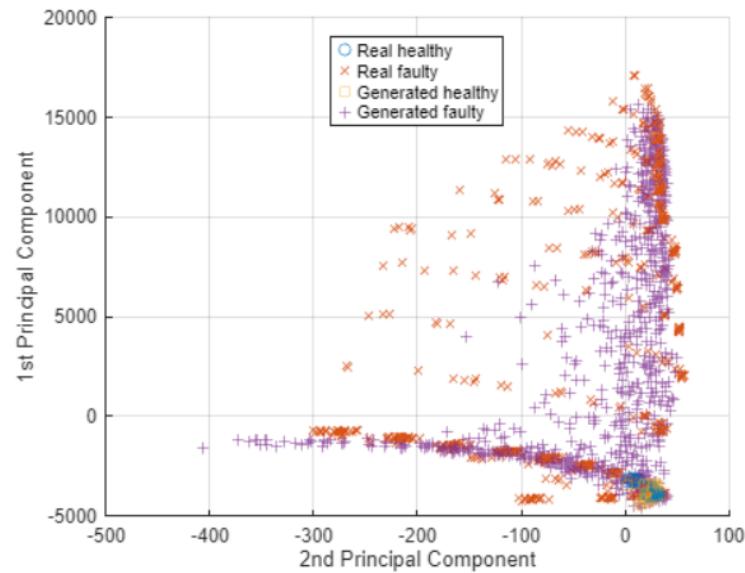
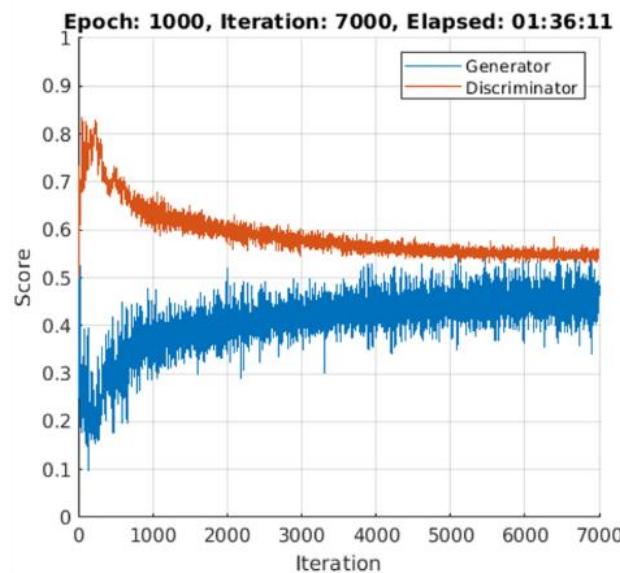
## Caso de uso bomba hidráulica



- Se destila un modelo a partir de los **datos simulados**
- Arquitectura **GAN** (Red Generativa Adversaria)
- Se predice la variable del **flujo de salida**
- La entrada:
  - Ruido
  - Condición:
    - Bomba saludable
    - Fuga
    - Obstrucción
    - Desgaste

# Modelos basados en Datos

## Caso de uso bomba hidráulica



# Resultados

## Caso de uso bomba hidráulica

### Característica

Generación de 2000 señales

Consistencia de Señales

Explicabilidad

Uso Ideal

### Modelo Basado en Física

14h en CPU, 2h GPU

Mantiene la coherencia entre **todas** las señales.

Permite rastrear la **física exacta**; ideal para **validar hipótesis de diseño**.

**No hay restricción de tiempo y cómputo**, o para **validar hipótesis** de diseño.

### Modelo GAN (Data-Driven)

70 segundos

Replica bien **solo las variables objetivo**.

Caja Negra

**Latencia factor crítico**

# Modelos de generación de Datos sintéticos

## Simulación Explícita

- Modelos basados en agentes

## Estadísticos y probabilísticos clásicos

- ARIMA/SARIMA (Series temporales)
- Hidden Markov Models (HMM)
- Gaussian Mixture Models (GMM)
- Cúpulas

## Aprendizaje profundo

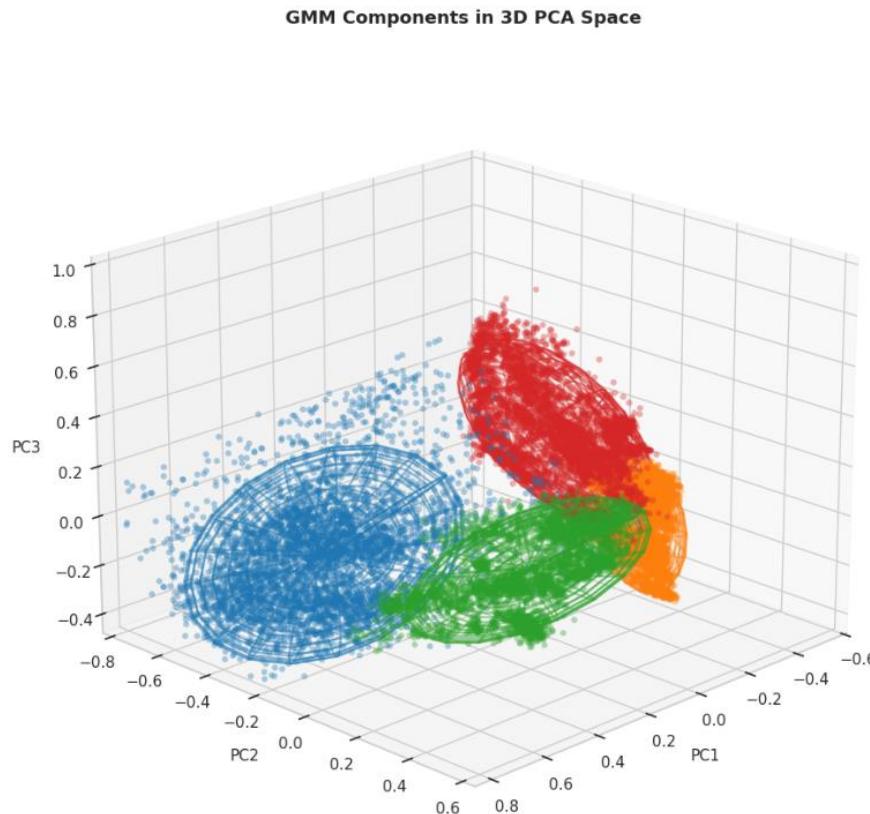
- Variational Autoencoders (VAE)
- Generative Adversarial Networks (GAN)
- Diffusion Models
- RNN/LSTM/GRU/Transformers

## Otras

- SMOTE
- DBA
- Noise Inyection

# Generación de datos sintéticos

## GMM para un autoencoder



PERFORMANCE OF AUTOENCODER WITH MAHALANOBIS DISTANCE SCORING UNDER DIFFERENT GMM AUGMENTATION SCENARIOS

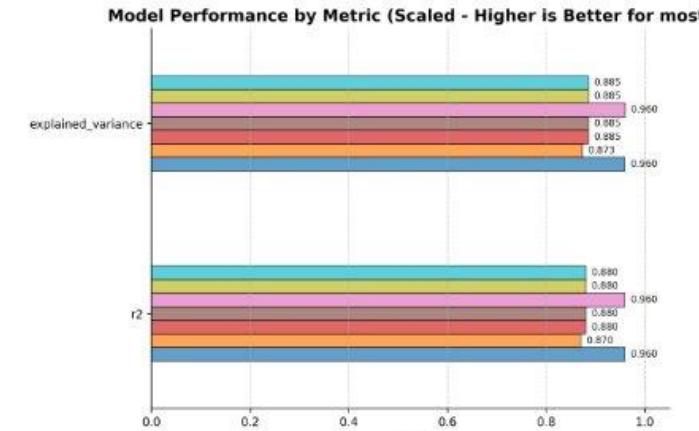
Scenario	Precision	Recall	F1-score	Accuracy
A1 (25% Real)	0.7731	0.6799	0.7235	0.7927
A2.1 (25% + GMM $\times$ 1)	0.8248	0.8206	0.8227	0.8589
A2.2 (25% + GMM $\times$ 2)	0.8346	0.8534	0.8439	0.8740
A4 (10% Real)	0.7585	0.6155	0.6796	0.7684
A5.1 (10% + GMM $\times$ 2)	0.8082	0.8008	0.8045	0.8447
A5.2 (10% + GMM $\times$ 3)	0.8285	0.8312	0.8298	0.8640

# Evaluación de datos

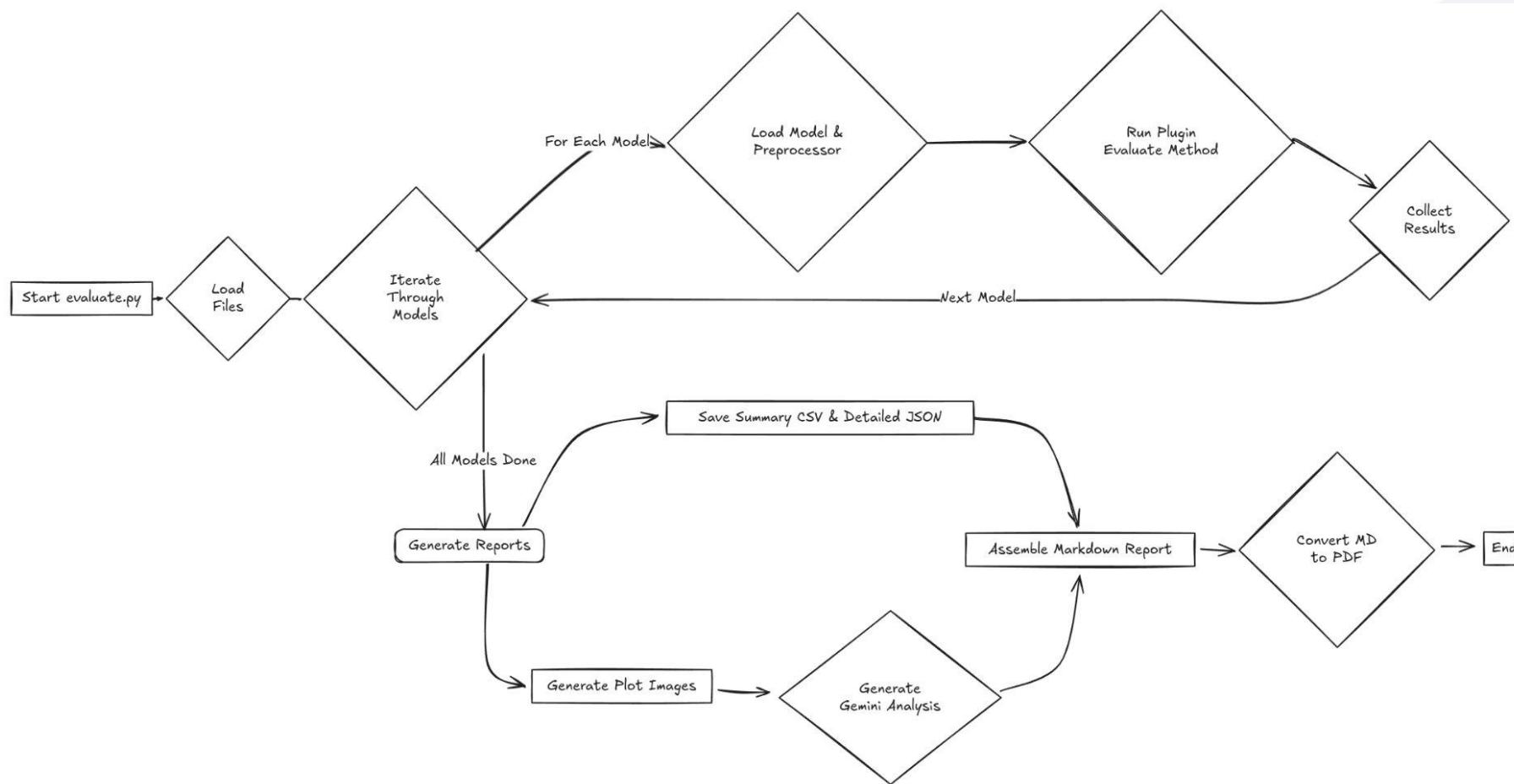
Paquete de **python** para la evaluación automática de modelos

Es capaz de generar un informe detallado utilizando **IA Generativa**, donde se comparan los diferentes modelos para evaluar su rendimiento.

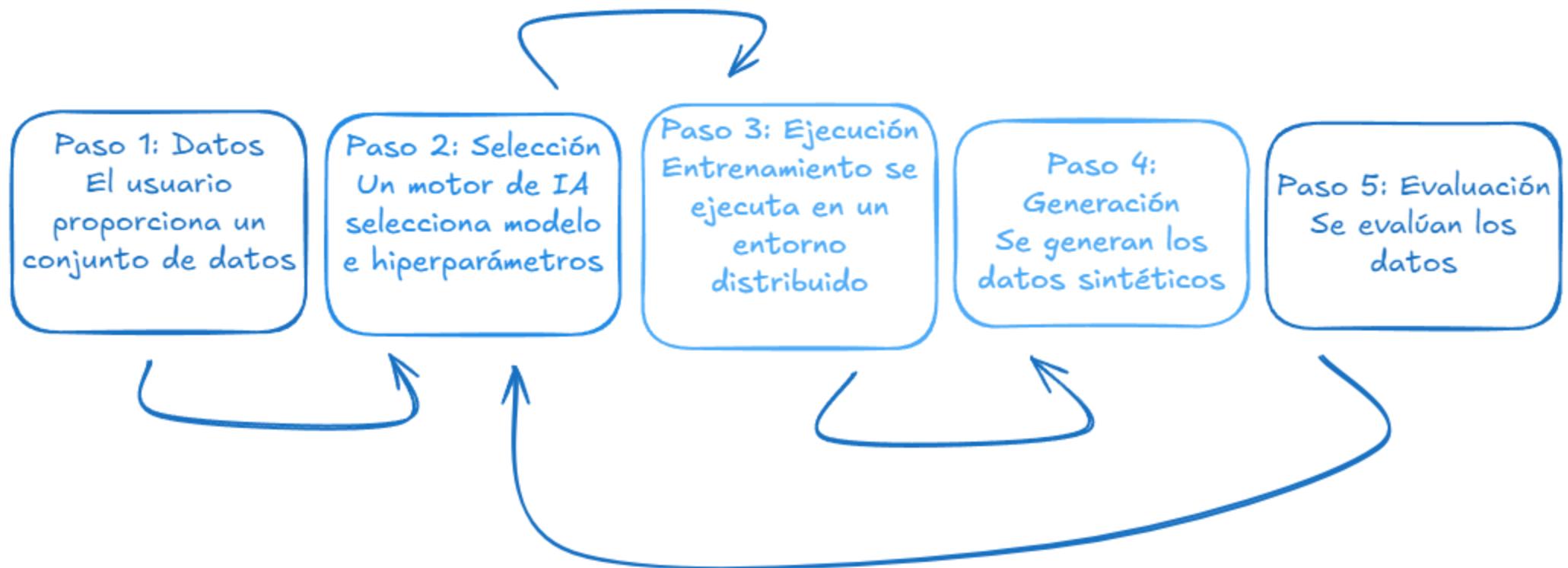
Soporta **Plugins** para modelos con evaluación o métricas personalizadas.



# Evaluación de datos



# Sistema de Generación de Datos Sintéticos



# Muchas Gracias