

# Técnicas de *Machine Learning* en la Gestión de Negocio

Conferencia de Stata (2021)

Franco A. Mansilla Ibáñez

# Agenda

- Analítica
  - Introducción.
  - Objetivos.
  - *Benchmark*.
  - Medidas de Evaluación.
- Metodologías Crisp-DM
  - Introducción.
  - Etapas.
- Machine Learning
  - ¿Qué es el *Machine Learning*?
  - Sistemas de aprendizaje.
  - Técnicas de *Machine Learning*
- Aplicación en Stata 17

# Analítica

## Introducción

La costumbre: tomar decisiones con base a **criterios expertos** (método Delphi)

Luego comenzó en aplicar **reglas de decisión estáticas**.

*Desde que surgieron los computadores y el almacenamiento de datos, se pudieron aplicar técnicas cuantitativas para validar **hipótesis de negocios**.*

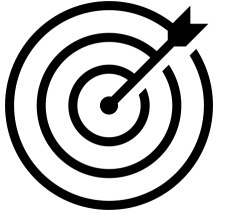
Teniendo historia se pudieron realizar los primeros **modelos estadísticos clásicos**.

En la actualidad se utilizan los modelos de *Machine Learning* para **mejorar el poder predictivo**.

*Los métodos cuantitativos se consideran de carácter táctico en vez de estratégicos.*

# Analítica

## Objetivos



¿Cuál(es) son los **objetivos** de aplicar métodos cuantitativos en los Negocios?

1. Establecer información base necesaria para aplicar reglas de negocios, estrategias y políticas.

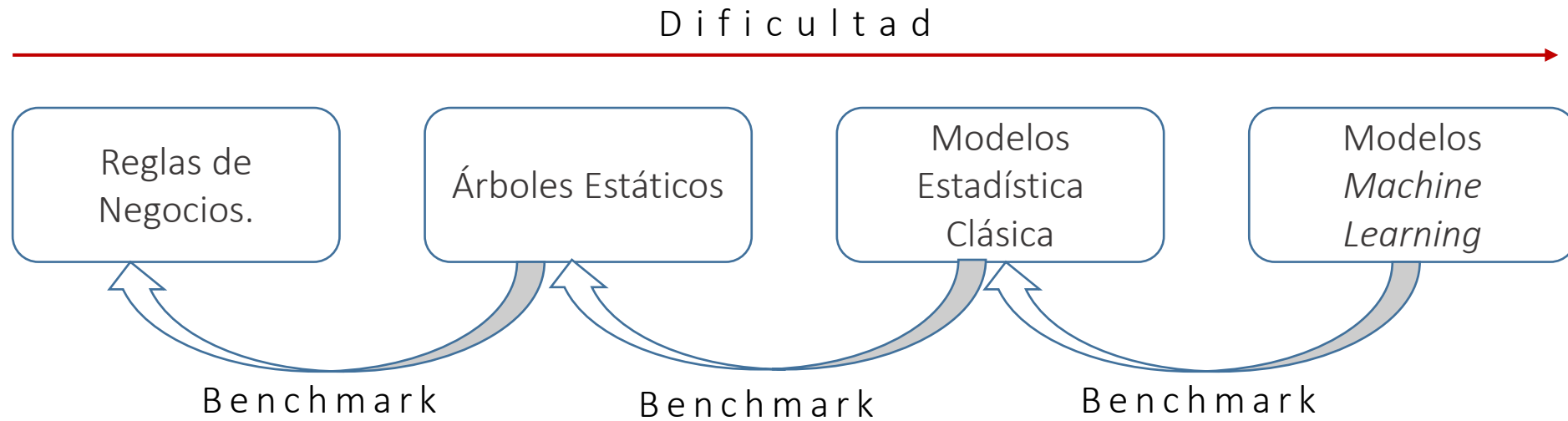
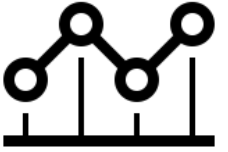
2. Identificar perfiles personalizados a nivel de personas, empresas, sectores, etc.

3. Introducir variedades de información que son intuitivas y/o pocas conocidas

4. Comprobar hipótesis de negocios para establecer estrategias de negocios.

# Analítica

## Benchmark



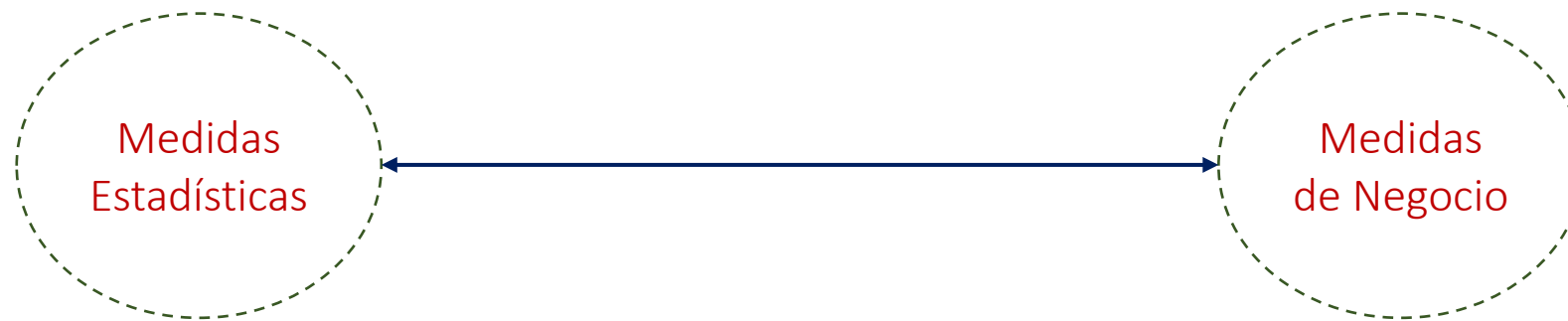
- Ejemplo

# Analítica

## Medidas de Evaluación



- Los métodos cuantitativos en los negocios no pueden implementarse sino se aplican las evaluaciones de negocio.



- Las medidas estadísticas evalúan la **perfección del modelo**, en cambio, las medidas de negocio evalúan la **tolerancia al equivocarse**.
- **Ejemplos**: Modelo de Inferencia y Modelo de Pronóstico.

# Metodología Crisp-DM

## Introducción

- Las técnicas de *Data Science* o *Data Analytics*, que tanto interés despiertan hoy en día, en realidad surgieron en la década de los 90.
- Cuando se usaba el termino KDD (*Knowledge Discovery in Databases*) para referirse al (amplio) concepto de **hallar conocimiento en los datos**.
- CRISP-DM proporciona una descripción **normalizada del ciclo de vida de un proyecto** estándar de análisis de datos.
- Consiste en 6 fases (estándar) + 1 fase (seguimiento).

# Metodología Crisp-DM

## Etapas

1. Entendimiento del Negocio      ¿Cuál es la necesidad?
2. Entendimiento del Dato.      ¿Qué datos tenemos?
3. Preparación de los Datos.      ¿Cómo preparamos nuestros datos?
4. Modelamiento      ¿Cómo modelamos los datos?
5. Evaluación del Modelo.      ¿Qué tan bueno es el modelo?
6. Implementación.      ¿Cuándo/Donde se implementa?
7. Seguimiento      ¿Cuándo/Cómo se hace el seguimiento?



# Machine Learning

## ¿Qué es Machine Learning?

- Es el **metodo de aprendizaje de la Inteligencia Artificial**, que **proporciona a los sistemas de la capacidad de aprender y mejorar automaticamente**.
- El concepto de aprender surge al **identificar patrones complejos**, que en la actualidad ninguno humano sería capaz de lograr.
- En general, los sistemas de aprendizaje tiene que aprender, y la forma de aprender es entregando una base de datos, obligatoriamente, con un **set de características** (PREDICTORAS), pero **no necesariamente** una marca al evento a modelar (TARGET).

# Introducción

## Sistemas de Aprendizaje

- Ejemplo clasico para entender los aprendizajes, es el **perro policial para encontrar sustancias ilícitas**.

Clasificación	Aprendizaje	Descripción
En la que Entrenan	Supervisado	Existe un target que marca el evento en que se pretende modelar: Continuo, Dicotómico o Categórico.
	No-Supervisado	No existe un target que marca el evento en que se pretende modelar
	Semi-Supervisado	Se consideran sistemas hibridos (sup. y no-sup) donde el target tiene una marca incompleta. (Detección de Anomalías es muy usado, me marca lo normal para detectar lo anormal).
	Reforzado	Son sistemas hibridos (no-sup y online) aprendera nuevas conductas cuando le entregamos nueva informacion. Es usado en los sistemas donde tienen que aprender con nueva informacion (perro robot)
Aprenden en el Tiempo	Online	El sistema aprende de forma incremental, sea en pequeños grupos (mini-bach)
	Bach	El sistema no aprende de forma incremental, se entrena utilizando todos los datos disponibles. Se entrena nuevamente con todos los datos disponible.
Realizan Predicciones	Instancias	Modelos que aprenden - conjunto de entrenamiento y luego generaliza para nuevos ejemplares (Modelos de Inferencia de venta o renta)
	Modelos	Modelos que describen un conjunto de datos y se utilizan para realizar predicciones (Modelos de conducto, mediante conductas pasadas predicir conductas futuras).

# Machine Learning

## Técnicas

1. Algoritmos de Regresión.
2. Algoritmos Bayesianos.
3. Algoritmos de Agrupación.
4. Algoritmos Arboles de Decisión.
5. Algoritmos de *Ensemble*.
6. Algoritmos de Redes Neuronales.
7. Algoritmos de Reducción de Dimensión.
8. Algoritmos de Aprendizaje Profundo.