



# **La Minería de Datos en la Construcción de Modelos de Soporte para la Toma de Decisiones**

Viterbo H. Berberena González

Doctor en Ciencias Técnicas

Director de Minería de Datos de Pearson S.A. de C.V.

# Objetivos

- Presentar el concepto de minería de datos como estrategia de modelación matemática.
- Definir la metodología de la minería de datos.
- Explicar, de forma muy general, los alcances de algunos proyectos genéricos de minería de datos.
- Comentar sobre algunos casos de éxito en México y las nuevas propuestas.

# Enfoque limitado de algunos modelos de soporte actuales

- Dirigidos a explicar el comportamiento de la variable objetivo.
- No toman en consideración las variables independientes o muy pocas de éstas.
- Los modelos matemáticos descriptivos:
  - Modelos de Markov.
  - Modelos de series de tiempo.
  - Modelos de líneas de espera, etc.
- Los modelos de investigación de mercados:
  - Síntomas, no el patrón real de comportamiento del mercado.
  - Modelos cuantitativos.
  - Modelos cualitativos.

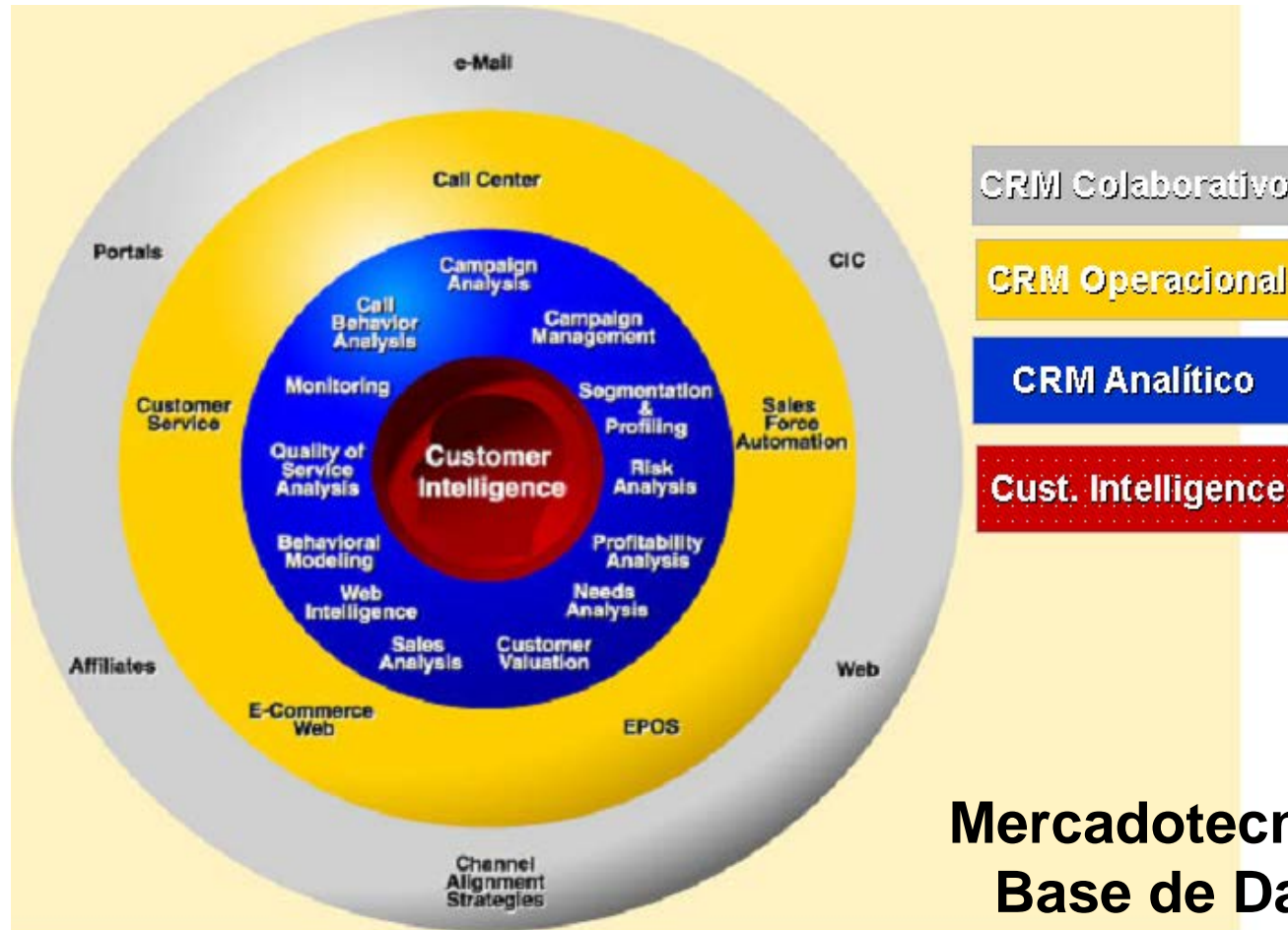
# Características fundamentales de los modelos de minería de datos

- Estáticos.
  - Estabilización de variables de comportamiento dinámico.
- Preponderantemente estocásticos.
  - Variables aleatorias.
- Variables independientes.
  - Discretas (cualitativas o no métricas).
    - Nominales o clasificatorias
    - Ordinales o de rangos.
  - Continuas (cuantitativas o métricas).
    - Intervalos.
    - Proporciones.
- Variables dependientes (objetivos o de respuestas).
  - Discretas (nominales (binarias) y ordinales).
  - Continuas (métricas).

# Premisas de la investigación de mercados en bases de datos

La necesidad transformar los urgentes problemas de negocios en soluciones estratégicas y ventajas competitivas sostenibles, a través de una potente fusión del hardware, el software, las bases de datos masivas y oportunistas de las empresas e instituciones y el conocimiento técnico de especialistas, analistas y consultores.

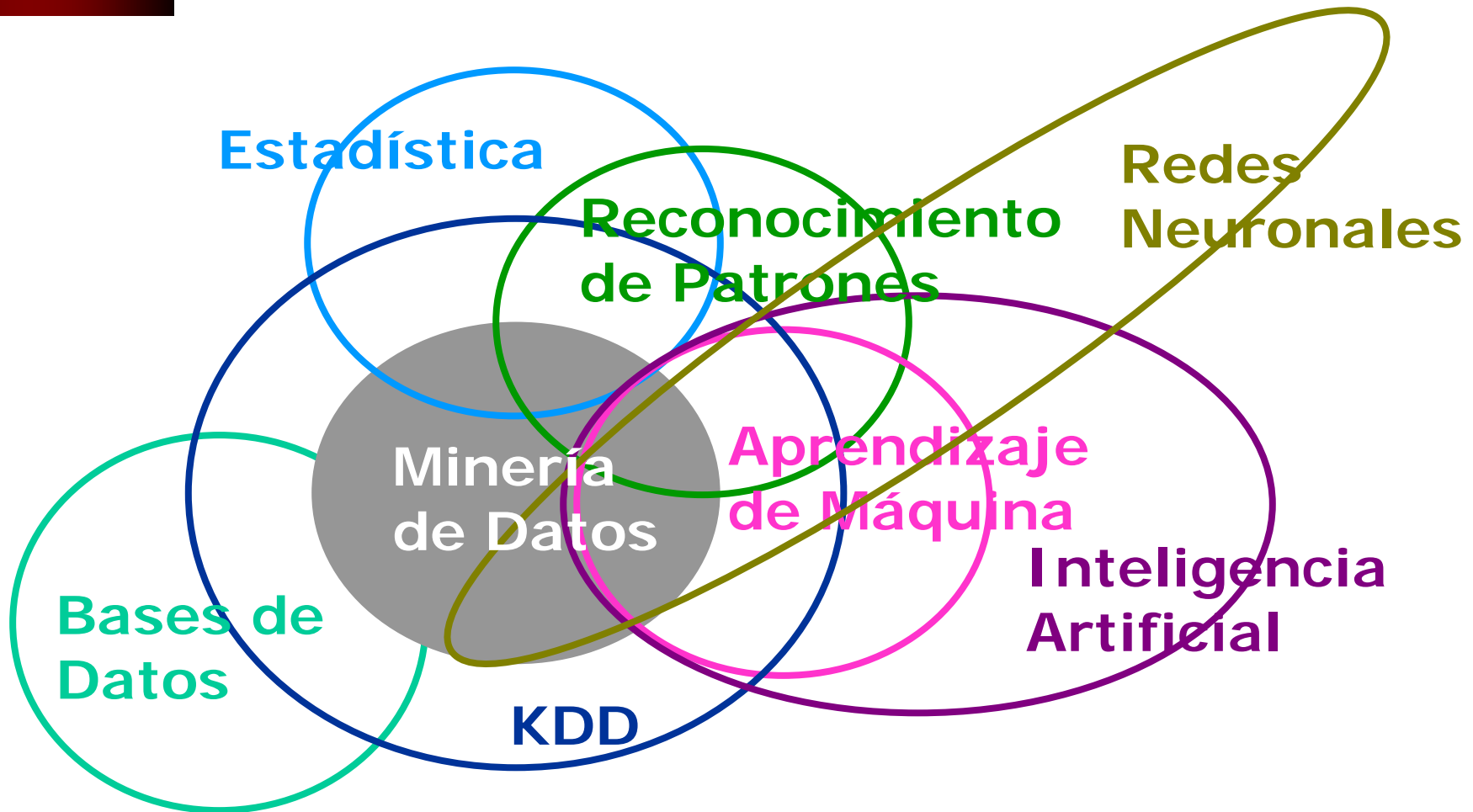
# Relación con CRM Analítico



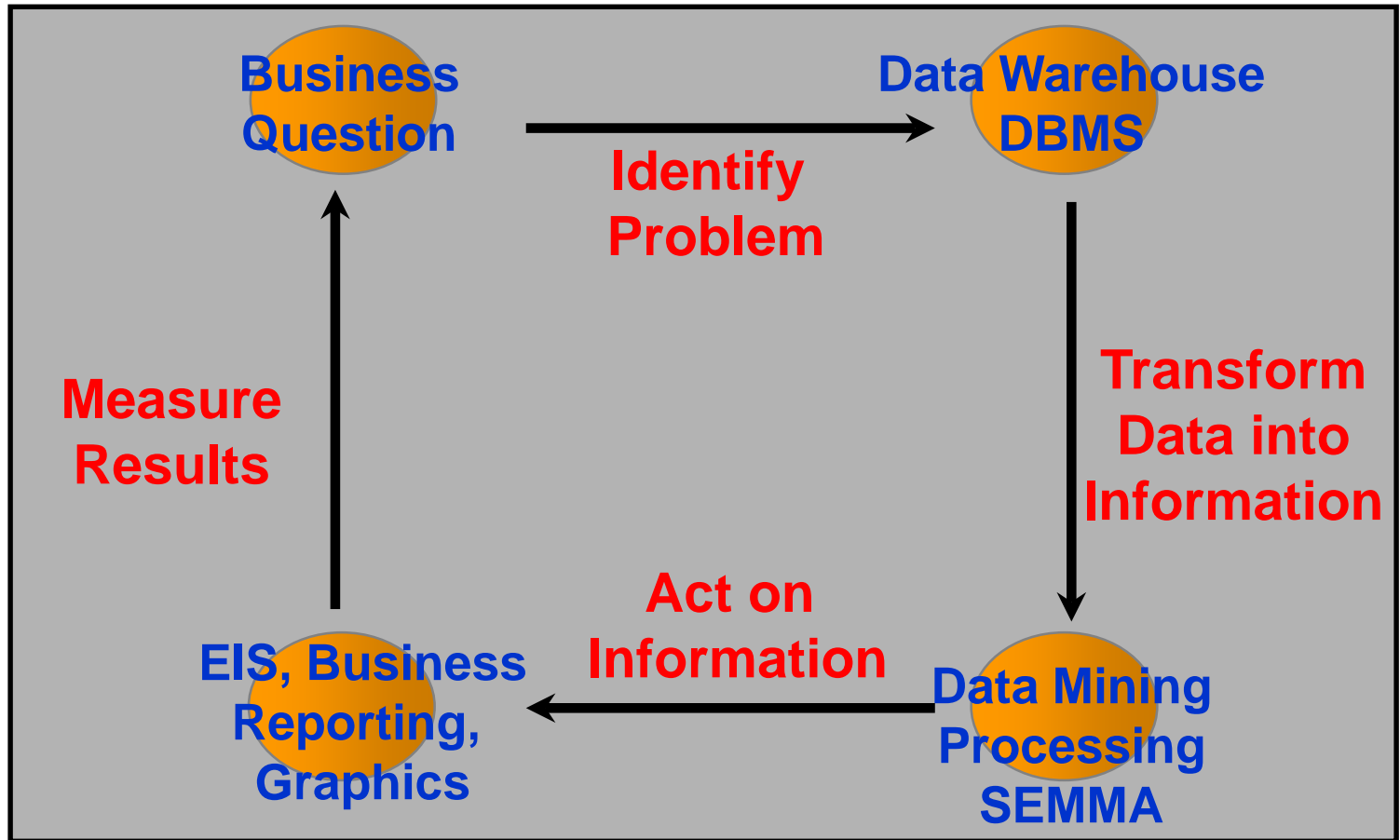
# La Minería de Datos

La minería de datos es el proceso iterativo (cíclico) de selección, exploración y modelación de grandes volúmenes de datos, con el fin de revelar patrones de comportamiento antes desconocidos, y así obtener una ventaja competitiva para la organización.

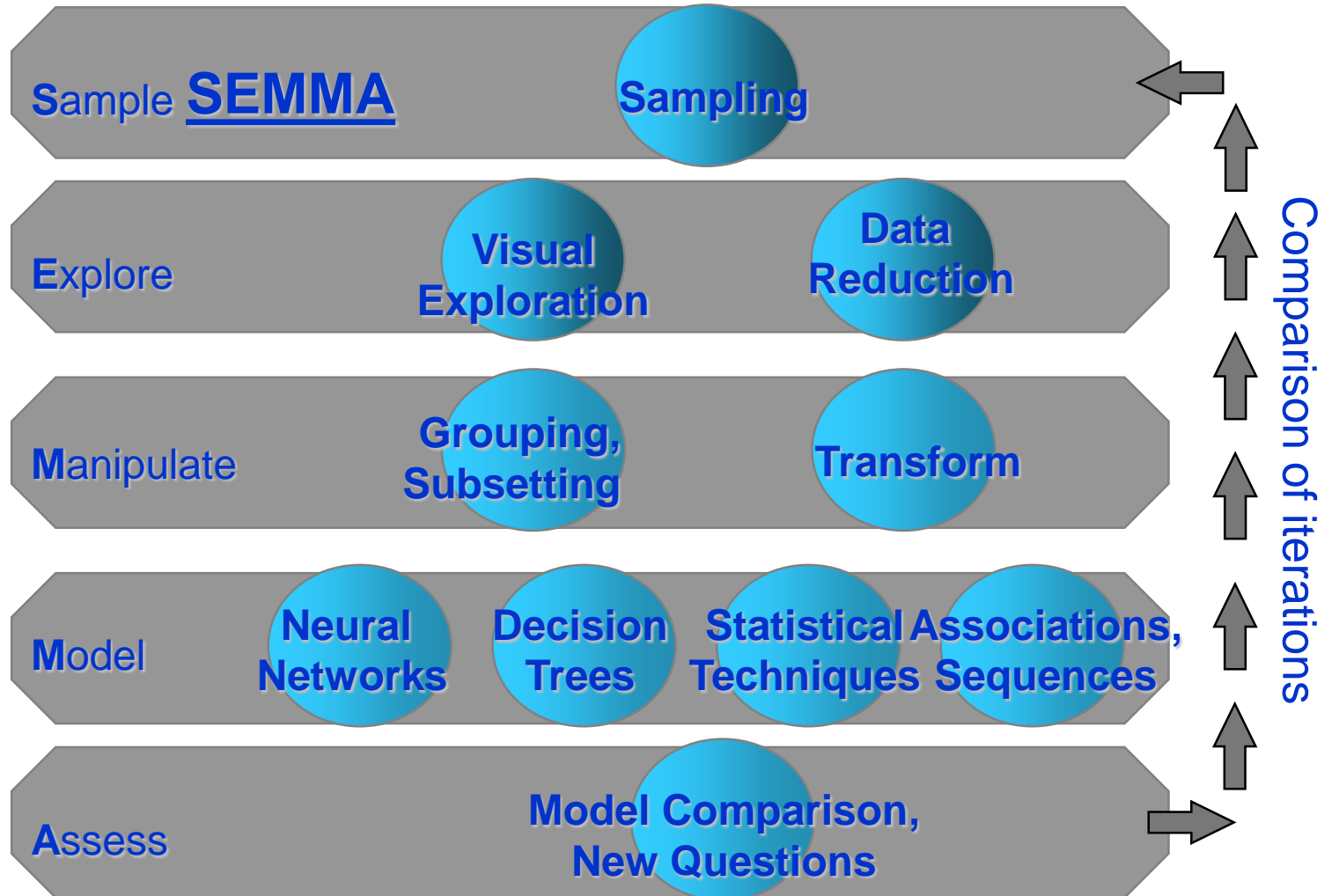
# Enfoque Multidisciplinario



# El Proceso de Minería de Datos



# La Metodología de Investigación



# Algunos proyectos genéricos de minería de datos

- Modelos de lealtad de clientes.
  - Customer attrition.
  - Churn analysis.
  
- Modelos de respuesta a las campañas de mercadeo.
  - Response analysis.
  - Profitability analysis.
  
- Modelos de ventas cruzadas.
  - Basket analysis.
  - Association analysis.
  - Sequence analysis.

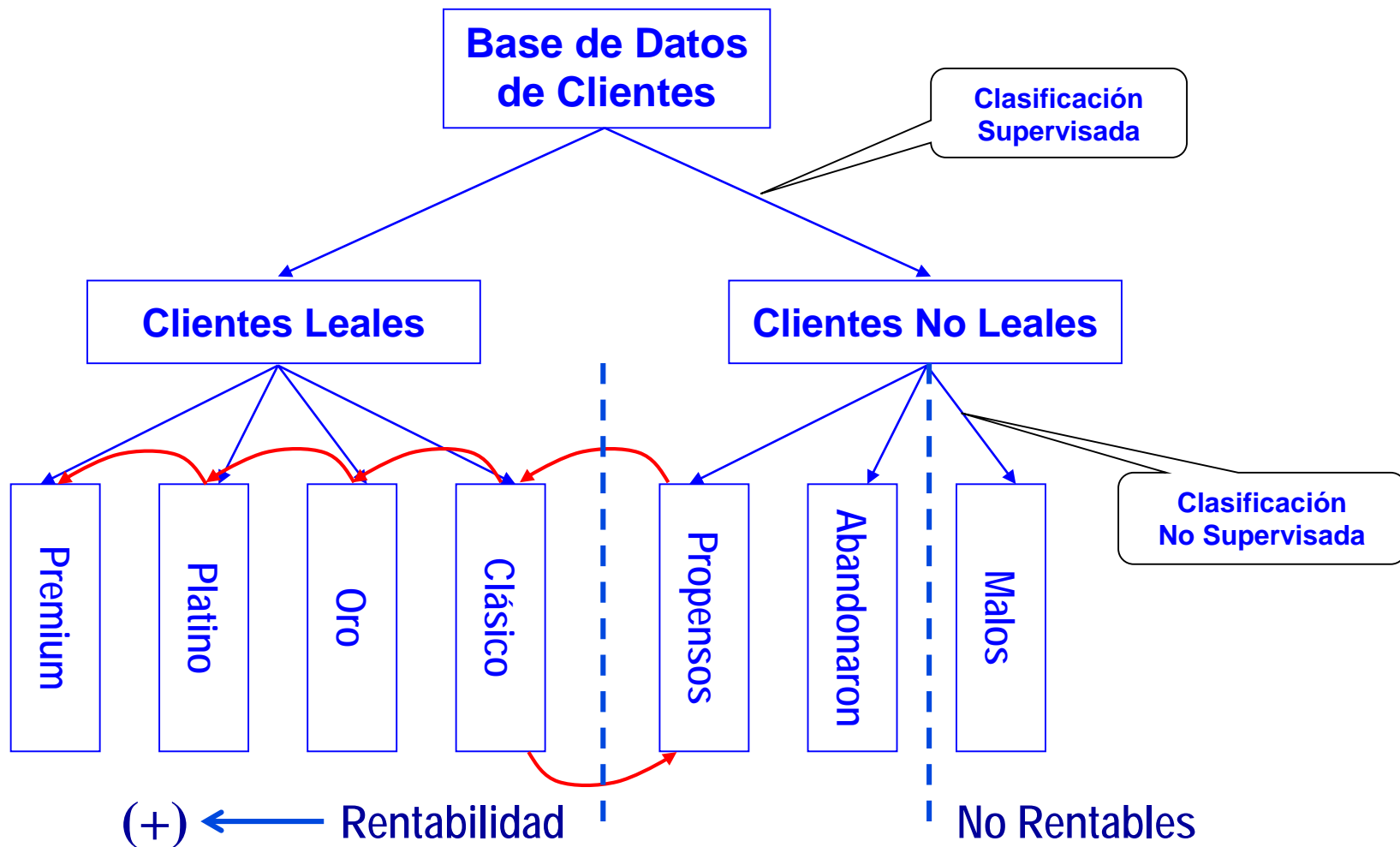
# Proyectos para el análisis de la lealtad de los clientes

- Problema de negocio.
  - ¿Quiénes y por qué abandonaron el producto o servicio?
  - ¿Cuáles clientes, que aún están en la empresa, son más propensos a abandonar el producto o servicio?
  - De los clientes que tienen propensión al abandono ¿Cuáles son los más rentables?
  
- Datos.
  - Variables de clasificación.
    - Edad, sexo, ocupación, ingresos, NSE, etc.
  - Variables de negocios.
    - Facturación, deudas, intereses.
    - Productos, servicios.
    - Antigüedad, rentabilidad, LTV, etc.
  - Construcción de la variable objetivo (criterio de lealtad) con reglas de negocios.
    - Rentabilidad, valor del cliente, otros.

# Proyectos para el análisis de la lealtad de los clientes

- Técnicas analíticas.
  - Clasificación supervisada (con técnicas de modelación predictiva).
    - Árboles de decisión, regresión logística y redes neuronales.
  - Clasificación no supervisada.
    - Análisis de conglomerados y técnicas de Kohonen.
- Resultados.
  - Determinación de clientes leales.
    - Muy rentables, rentables y no rentables.
  - Determinación probabilística de la propensión al desgaste (attriters).
  - Determinación probabilística de la propensión al abandono (churners).

# Representación esquemática del modelo de lealtad de clientes



# Proyectos para el análisis de ventas cruzadas de productos o servicios

## Problema de negocio.

- ¿Qué productos o servicios los clientes compran juntos?
- ¿En qué secuencia los clientes escogen sus productos o servicios?
- ¿En qué secuencia los clientes compran los productos o servicios escogidos?
- ¿Qué clientes tienen mayor propensión a comprar determinados tipos de productos o servicios?

## ■ Datos.

- Historial de compras de los clientes.
- Variable de identificación de cliente.
  - ID del cliente, la cuenta, el contrato, etc.
- Variable objetivo.
  - Productos, categorías de productos, servicios, etc.
- Variable de secuencia.
  - Fechas en que cada cliente ha ido adquiriendo los distintos productos o servicios.

# Proyectos para el análisis de ventas cruzadas de productos o servicios

- Técnicas analíticas.
  - Análisis estadístico de asociaciones.
  - Análisis estadístico de secuencias.
  - Análisis Bayesiano.
    - Determinación de probabilidades condicionales posteriores, a partir de información inicial limitada.
- Resultados.
  - Determinación probabilística de la composición de las canastas de productos o servicios, que los clientes adquieren juntos.
  - Determinación probabilística de la propensión de los clientes a comprar determinados producto o servicios una vez que han adquirido otros inicialmente.
  - Determinación probabilística de la propensión a cierta secuencia de compras de productos o servicios.

# Varios casos exitosos de minería de datos en México

- Análisis de ventas cruzadas y secuenciación de ventas en bebidas y licores.
- Modelo de clasificación de clientes en banca electrónica (variante de modelo de lealtad).
- Análisis de lealtad en una empresa líder en telecomunicaciones (modelo genérico).
- Modelación y optimización del sistema de otorgamiento de crédito al consumidor (credit scoring) en una institución financiera.

# Varios casos exitosos de minería de datos en México

- Modelación del comportamiento (riesgo) de cuentas en un banco (behavioral scoring).
- Modelo de clasificación de clientes de banca por teléfono (variante del modelo de lealtad).
- Segmentación del mercado de autos compactos para determinada marca.
- Modelo predictivo de ingresos de los clientes en un banco.

# Nuevos proyectos de minería de datos en México (propuestas)

- Modelo de evaluación y control de fraudes en línea con el enfoque estadístico de  $6\sigma$  (seis sigmas) en tarjetas en bancos.
- Modelos de activación de tarjetas de crédito en bancos.
- Modelo de análisis de campañas de mercadeos en bancos.
- Modelo de comportamiento de eventos delictivos para el departamento de seguridad pública en una gran ciudad.
- Modelo de segmentación de micro-mercados en una gran ciudad.

# Conclusiones

- La minería de datos se presenta como una vía para la modelación matemática de problemas muy complejos.
- Existen en la actualidad, en el mercado, potentes herramientas de minería (mineros), que permiten, que sea prácticamente posible, modelar cualquier escenario de investigación.
- La minería de datos tiene un profundo enfoque económico, desde la definición inicial del problema de negocios, hasta sus resultados finales (ROI).

# Conclusiones

- La minería de datos está relacionada con la investigación de mercados en bases de datos.
- La minería de datos no suplanta la investigación de mercados tradicional, ya que su enfoque, técnicas analíticas, herramientas de trabajo y escenario de investigación, son diferentes.
- Finalmente, la minería de datos posibilita la obtención de modelos de soporte de gran precisión y confiabilidad, en escenarios de investigación muy grandes y complejos.



# Datos Personales

- Viterbo H. Berberena González.
- Doctor en Ciencias Técnicas.
- Director de Minería de Datos de Pearson S.A. de C.V.
- Homero 223, PH, Colonia Polanco.
- CP 11560, México D.F.
- Teléfonos:(52)(55) 5531-5324, 5531-5560 ext. 145
- Fax: (52)(55) 5203-8230
- mailto: [vberberena@pearson-research.com](mailto:vberberena@pearson-research.com)
- USA phone: (305) 390 8242 Miami.
- 1-800 711 7709 –TOLL FREE.