



La analítica en el Big Data:

La Ciencia de Datos como factor impulsor de crecimiento empresarial

M.C. Naturaleza Cossío |

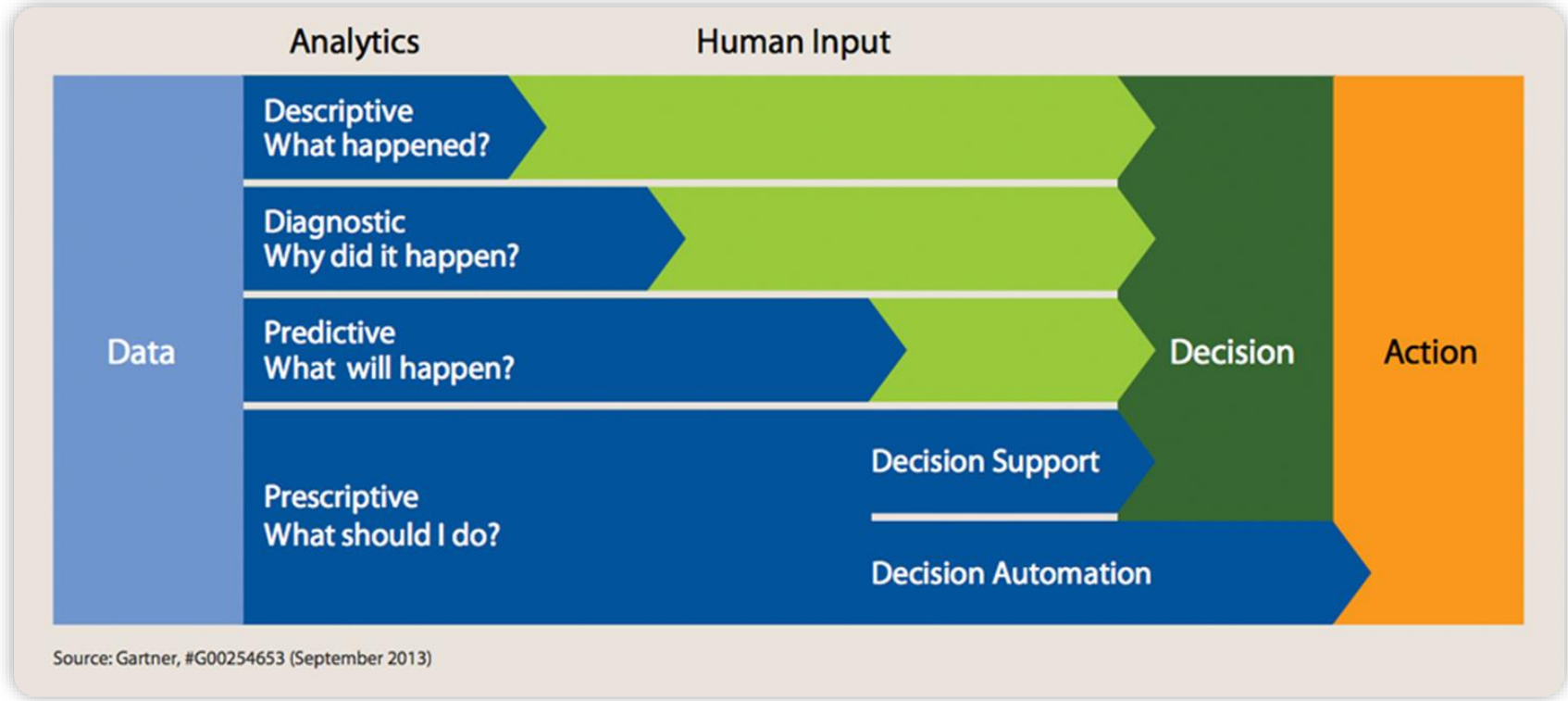
Consultor Senior de Analítica en SAS México y Caribe



Contenido

- ¿Qué es la analítica?
- Aplicaciones de la analítica
- Modelación matemática
- El Big Data
- La ciencia de datos
- Mapa estratégico empresarial
- Recomendaciones

¿Qué es la analítica?



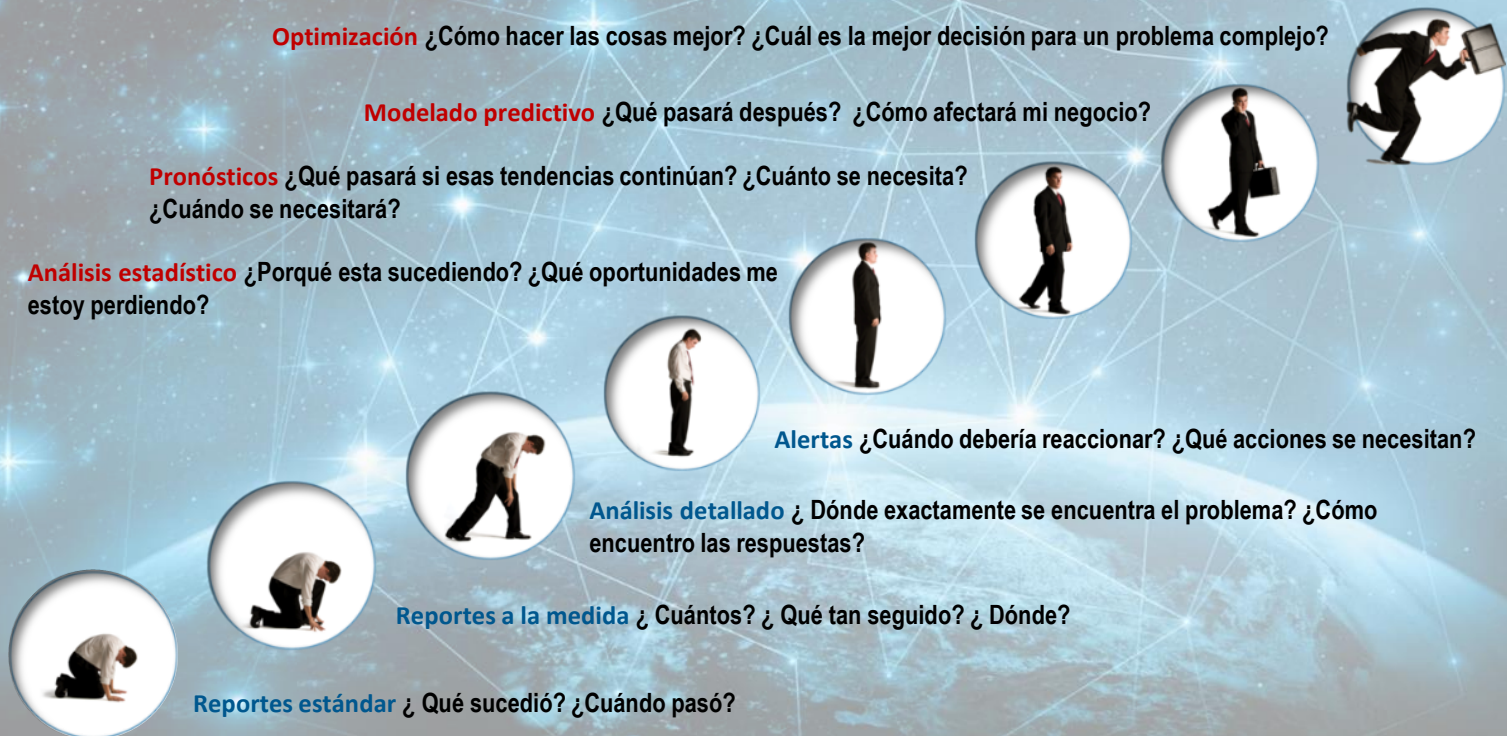
“Extend Your Portfolio of Analytics Capabilities”

¿Hacia dónde ir?

Valor (Valor de la ventaja competitiva)

Business Analytics

Business Intelligence



Tiempo (madurez)

Aplicaciones estratégicas de la analítica

Banco



- Detección de fraude
- Next Best Offers
- Análisis de crédito
- Lealtad del cliente
- Retención del cliente
- El valor del cliente

Gobierno



- Planeación estratégica del gasto público
- Logística de servicios
- Monitoreo de dependencias públicas
- Procesamiento de lenguaje natural

Salud



- Diagnósticos predictivos
- Canalización paciente enfermedad
- Monitorización de salud

Manufactura y energía



- Cadena de suministros
- Optimización de inventarios
- Planeación de la demanda
- Pronósticos

Communications y retail



- Marketing dirigido
- Visualización de oportunidades de negocio
- Hábitos de consumo
- Segmentación
- Ubicación de anuncios en tiempo real

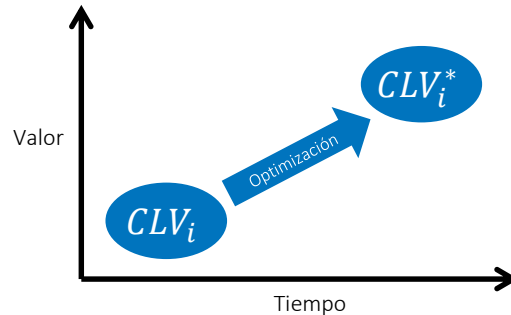
Modelación matemática

Optimización del CLV

Objetivo.- Optimizar el *Customer Equity*¹ (*CE*) de la Aseguradora,

$$CE = \sum_{i=1}^{NTC} CLV_i$$

mediante la mejora del *CLV* de cada cliente, lo cual se logra al explotar el valor potencial positivo de cada uno de sus componentes.



¹El *CE* es el valor de los posibles ingresos futuros de los clientes de la Aseguradora.

Modelo de optimización combinatoria

Dado el modelo del CLV

$$CLV_i = VA_i + \underbrace{\sum_{j \in PiP_i} \sum_{n=1}^2 M_{i,j,n}(1+r)^{-n} + \sum_{k \in PnP_i} \sum_{n=1}^2 M_{i,k,n}(1+r)^{-n}}_{= VF_i}$$

Modelo de optimización combinatoria

Dado el modelo del CLV

$$CLV_i = VA_i + \underbrace{\sum_{j \in PiP_i} \sum_{n=1}^2 M_{i,j,n} (1+r)^{-n}}_{=VF_i} + \underbrace{\sum_{k \in PnP_i} \sum_{n=1}^2 M_{i,k,n} (1+r)^{-n}}_{=VF_i}$$

Se busca

$$CLV_i^* = VA_i + \underbrace{\sum_{j \in PiP_i} \sum_{n=1}^2 M_{i,j,n} I_{i,j,n} (1+r)^{-n} + \sum_{k \in PnP_i} \sum_{n=1}^2 M_{i,k,n} I_{i,k,n} (1+r)^{-n}}_{=VF_i^*}$$

donde

$$I_{i,x,n} = \begin{cases} 0 & \text{si su venta o renovación representa una pérdida} \\ 1 & \text{si su venta o renovación representa un beneficio} \end{cases}$$

para $x = j, k$. Lo cual garantiza que

$$CLV_i \leq CLV_i^*$$

para toda $i = 1:NTC$.

Algoritmo de búsqueda local

Criterio de asignación.

$M_{i,x,2}$ $M_{i,x,1}$	< 0	≥ 0	
< 0	< 0	$ M_{i,x,1} \leq M_{i,x,2}$	≥ 0 4
		$ M_{i,x,1} > M_{i,x,2}$	< 0 5
≥ 0	?	≥ 0	1
	$M_{i,x,1} + M_{i,x,2}$		

Caso	$I_{i,x,1}$	$I_{i,x,2}$
1	1	1
2	1	0
3	0	0
4	1	1
5	0	0

Algoritmo de búsqueda local

Criterio de asignación.

$M_{i,x,2}$ $M_{i,x,1}$	< 0	≥ 0	
< 0	< 0	$ M_{i,x,1} \leq M_{i,x,2}$	≥ 0
		$ M_{i,x,1} > M_{i,x,2}$	< 0
≥ 0	$?$	≥ 0	
	$M_{i,x,1} + M_{i,x,2}$		

Caso 1

Caso	$I_{i,x,1}$	$I_{i,x,2}$
1	1	1
2	1	0
3	0	0
4	1	1
5	0	0

Ejemplo:

Cliente $i \rightarrow$

Producto $x \rightarrow$

Periodo 1: Comportamiento positivo (vender o renovar)

Periodo 2: Comportamiento positivo (renovar)

Algoritmo de búsqueda local

Criterio de asignación.

$M_{i,x,2}$ \ $M_{i,x,1}$	< 0	≥ 0	
< 0	< 0	$ M_{i,x,1} \leq M_{i,x,2}$ ≥ 0	4
		$ M_{i,x,1} > M_{i,x,2}$ < 0	5
≥ 0	?	≥ 0	1
	$M_{i,x,1} + M_{i,x,2}$		

Ejemplo: Caso 2

Cliente $i \rightarrow$

Producto $x \rightarrow$

Periodo 1: Comportamiento positivo (vender o renovar)

Periodo 2: Comportamiento negativo (NO renovar)

Caso	$I_{i,x,1}$	$I_{i,x,2}$
1	1	1
2	1	0
3	0	0
4	1	1
5	0	0

Algoritmo de búsqueda local

Criterio de asignación.

$M_{i,x,2}$ $M_{i,x,1}$	< 0	≥ 0
< 0	< 0 3	$ M_{i,x,1} \leq M_{i,x,2}$ ≥ 0 4 $ M_{i,x,1} > M_{i,x,2}$ < 0 5
≥ 0	? 2	≥ 0 1
	$M_{i,x,1} + M_{i,x,2}$	

Caso 3

Caso	$I_{i,x,1}$	$I_{i,x,2}$
1	1	1
2	1	0
3	0	0
4	1	1
5	0	0

Ejemplo:

Cliente $i \rightarrow$

Producto $x \rightarrow$

Periodo 1: Comportamiento negativo (NO vender o renov.)

Periodo 2: Comportamiento negativo (NO renovar)

Algoritmo de búsqueda local

Criterio de asignación.

$M_{i,x,2}$ $M_{i,x,1}$	< 0	≥ 0
< 0	< 0 3	$ M_{i,x,1} \leq M_{i,x,2}$ ≥ 0 4 $ M_{i,x,1} > M_{i,x,2}$ < 0 5
≥ 0	? 2	≥ 0 1
	$M_{i,x,1} + M_{i,x,2}$	

Caso 4

Caso	$I_{i,x,1}$	$I_{i,x,2}$
1	1	1
2	1	0
3	0	0
4	1	1
5	0	0

Ejemplo (análisis costo-beneficio):

Cliente $i \rightarrow$

Producto $x \rightarrow$

Periodo 1: Comportamiento negativo (vender o renov.)

Periodo 2: Comportamiento positivo mayor al

comportamiento negativo del periodo 1 (renovar)

Algoritmo de búsqueda local

Criterio de asignación.

$M_{i,x,2}$ $M_{i,x,1}$	< 0	≥ 0	
< 0	< 0 3	$ M_{i,x,1} \leq M_{i,x,2}$ ≥ 0 4	$ M_{i,x,1} > M_{i,x,2}$ < 0 5 ← Caso 5
≥ 0	? 2	≥ 0 1	
	$M_{i,x,1} + M_{i,x,2}$		

Ejemplo (análisis costo-beneficio):

Cliente $i \rightarrow$

Producto $x \rightarrow$

Periodo 1: Comportamiento negativo (NO vender o renov.)

Periodo 2: Comportamiento positivo menor al

comportamiento negativo del periodo 1 (NO renovar)

Caso	$I_{i,x,1}$	$I_{i,x,2}$
1	1	1
2	1	0
3	0	0
4	1	1
5	0	0

Algoritmo de búsqueda local

Criterio de asignación.

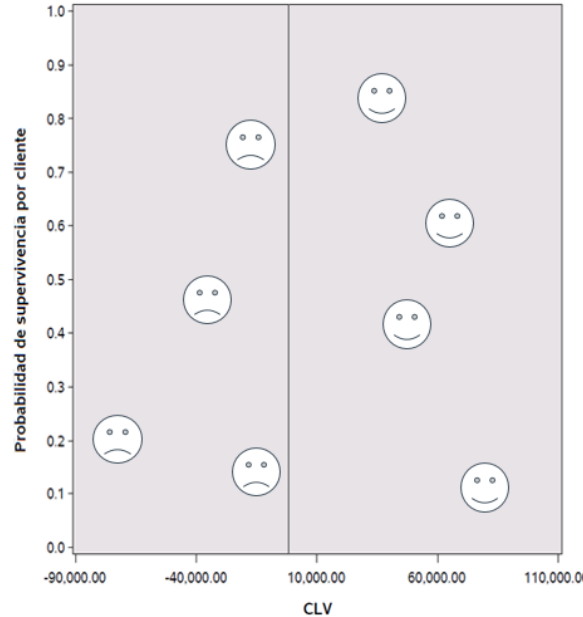
$M_{i,x,2}$ $M_{i,x,1}$	< 0	≥ 0	
< 0	< 0	$ M_{i,x,1} \leq M_{i,x,2}$	≥ 0
		$ M_{i,x,1} > M_{i,x,2}$	< 0
≥ 0	$?$	≥ 0	
	$M_{i,x,1} + M_{i,x,2}$		

$$\text{Sea } E_i = \left\{ \{I_{i,j,n}\}_{j \in P_i P}, \{I_{i,k,n}\}_{k \in P_n P} \right\}_{n=1,2}.$$

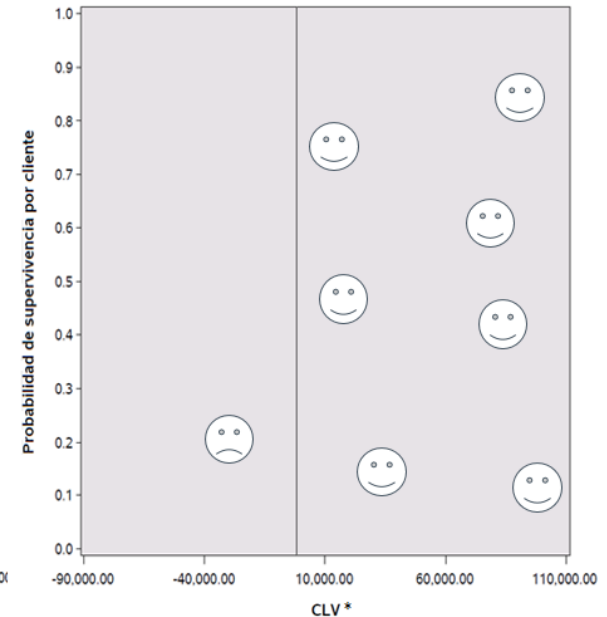
Caso	$I_{i,x,1}$	$I_{i,x,2}$
1	1	1
2	1	0
3	0	0
4	1	1
5	0	0

Beneficios

- Dada la estimación actual del *CLV*, a través del modelo de optimización es posible aumentar su valor mediante la vinculación de acciones de marketing que exploten el valor potencial del cliente en un horizonte de tiempo.
- Permite mejorar el *CE* sin aumentar el número de clientes.
- Este modelo de optimización establece, para cada cliente de la aseguradora, su portafolio de productos óptimo, restringido a los productos probables a adquirir.



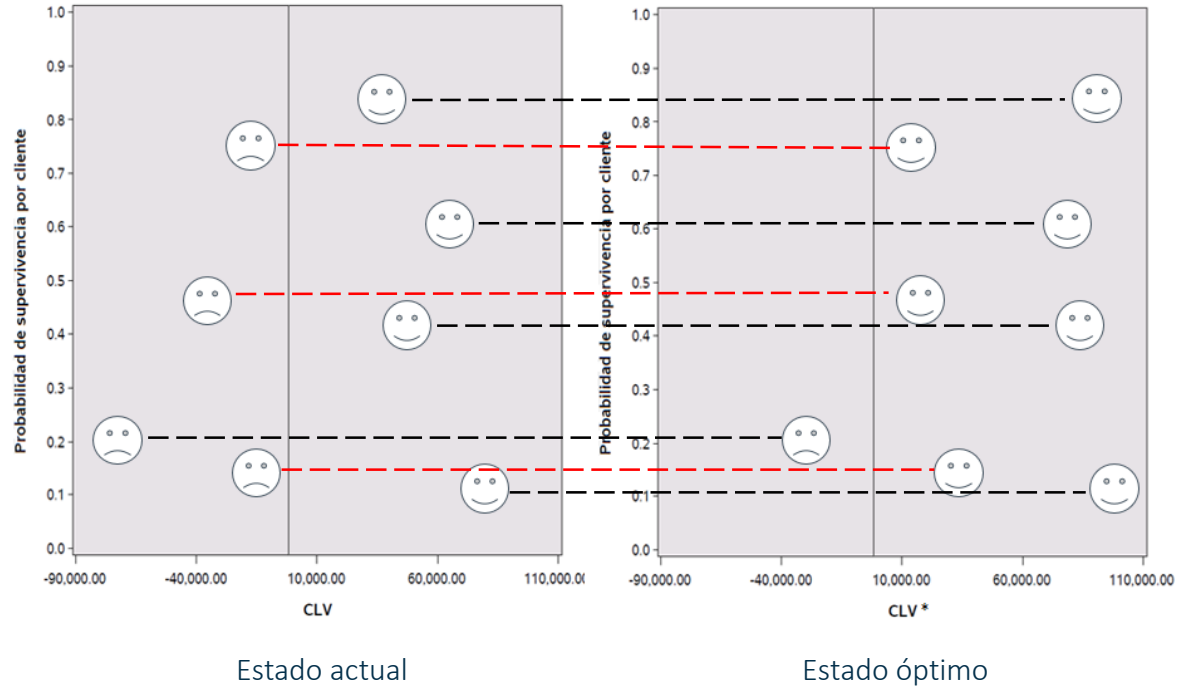
Estado actual



Estado óptimo

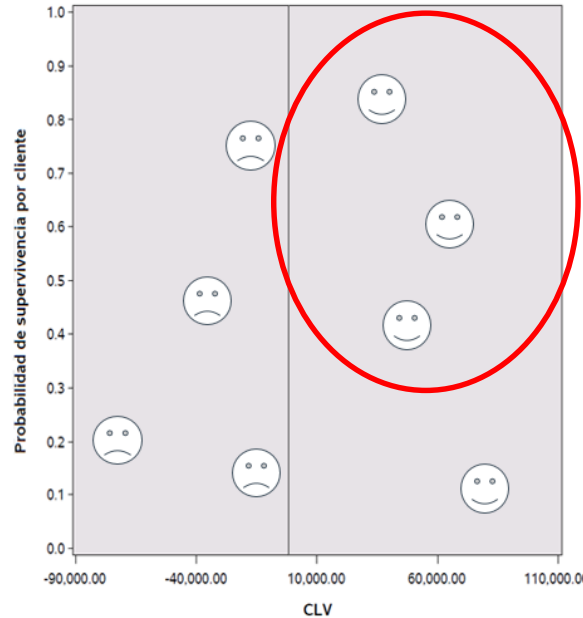
Beneficios

- Dada la estimación actual del *CLV*, a través del modelo de optimización es posible aumentar su valor mediante la vinculación de acciones de marketing que exploten el valor potencial del cliente en un horizonte de tiempo.
- Permite mejorar el *CE* sin aumentar el número de clientes.
- Este modelo de optimización establece, para cada cliente de la aseguradora, su portafolio de productos óptimo, restringido a los productos probables a adquirir.

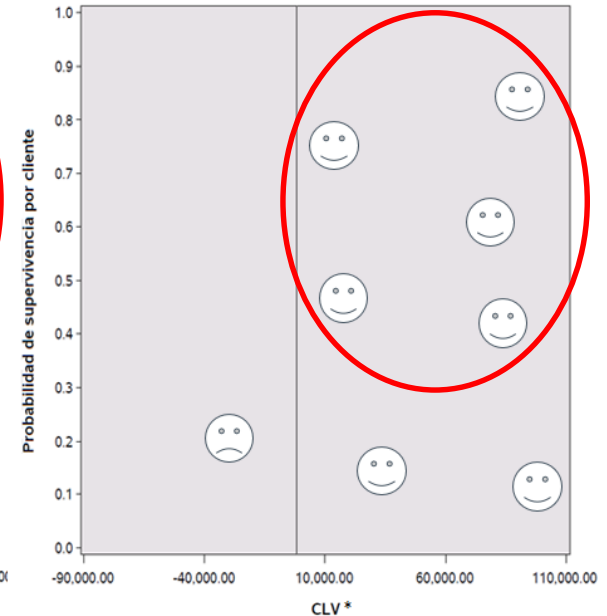


Beneficios

- Dada la estimación actual del *CLV*, a través del modelo de optimización es posible aumentar su valor mediante la vinculación de acciones de marketing que exploten el valor potencial del cliente en un horizonte de tiempo.
- Permite mejorar el *CE* sin aumentar el número de clientes.
- Este modelo de optimización establece, para cada cliente de la aseguradora, su portafolio de productos óptimo, restringido a los productos probables a adquirir.



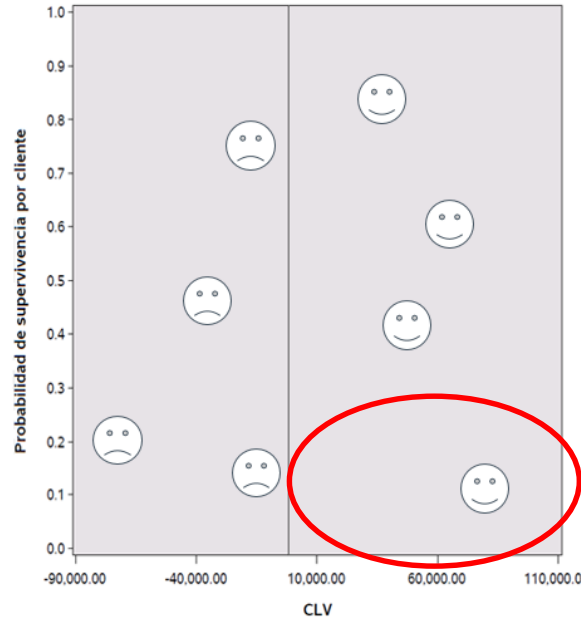
Estado actual



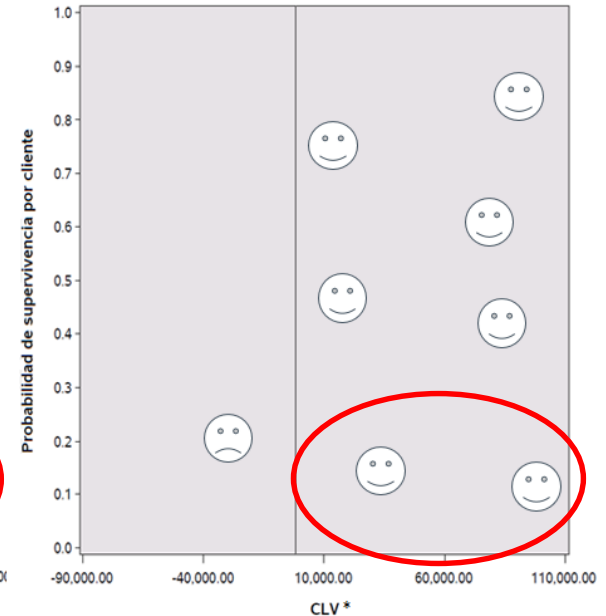
Estado óptimo

Beneficios

- Dada la estimación actual del *CLV*, a través del modelo de optimización es posible aumentar su valor mediante la vinculación de acciones de marketing que exploten el valor potencial del cliente en un horizonte de tiempo.
- Permite mejorar el *CE* sin aumentar el número de clientes.
- Este modelo de optimización establece, para cada cliente de la aseguradora, su portafolio de productos óptimo, restringido a los productos probables a adquirir.



Estado actual

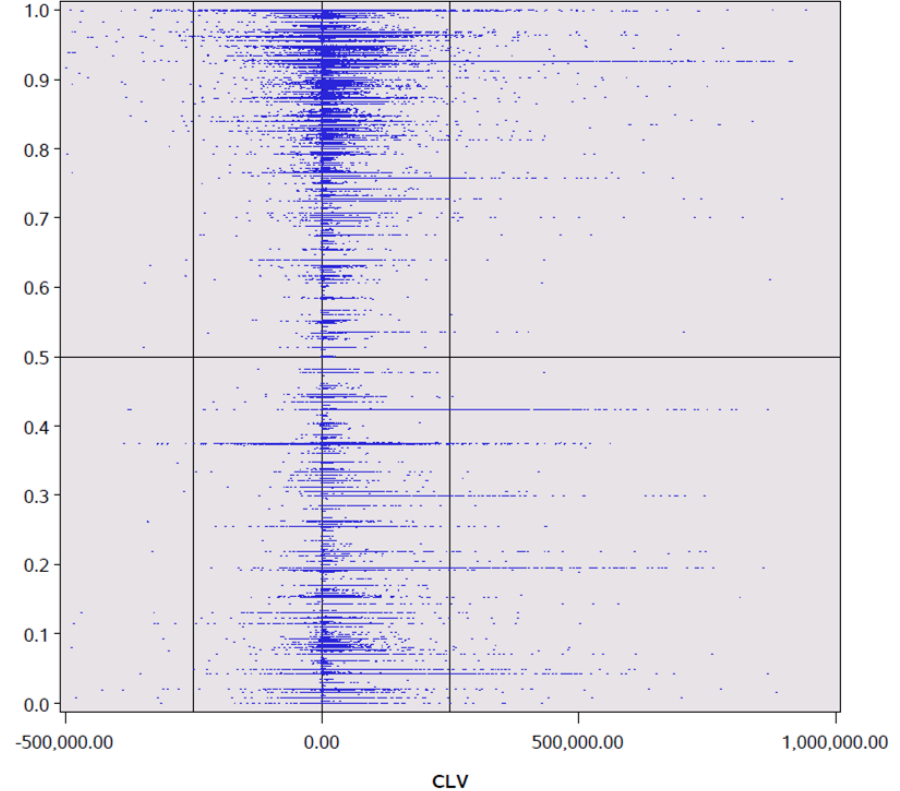


Estado óptimo

Beneficios

- Dada la estimación actual del *CLV*, a través del modelo de optimización es posible aumentar su valor mediante la vinculación de acciones de marketing que exploten el valor potencial del cliente en un horizonte de tiempo.
- Permite mejorar el *CE* sin aumentar el número de clientes.
- Este modelo de optimización establece, para cada cliente de la aseguradora, su portafolio de productos óptimo, restringido a los productos probables a adquirir.

Probabilidad de supervivencia por cliente



Modelo de optimización combinatoria

Dado el modelo del CLV

$$CLV_i = VA_i + \underbrace{\sum_{j \in PiP_i} \sum_{n=1}^2 M_{i,j,n} (1+r)^{-n}}_{=VF_i} + \underbrace{\sum_{k \in PnP_i} \sum_{n=1}^2 M_{i,k,n} (1+r)^{-n}}_{=VF_i}$$

Se busca

$$CLV_i^* = VA_i + \underbrace{\sum_{j \in PiP_i} \sum_{n=1}^2 M_{i,j,n} I_{i,j,n} (1+r)^{-n} + \sum_{k \in PnP_i} \sum_{n=1}^2 M_{i,k,n} I_{i,k,n} (1+r)^{-n}}_{=VF_i^*}$$

donde

$$I_{i,x,n} = \begin{cases} 0 & \text{si su venta o renovación representa una pérdida} \\ 1 & \text{si su venta o renovación representa un beneficio} \end{cases}$$

para $x = j, k$. Lo cual garantiza que

$$CLV_i \leq CLV_i^*$$

para toda $i = 1:NTC$.



Cada día se generan 2.5 exabytes de datos

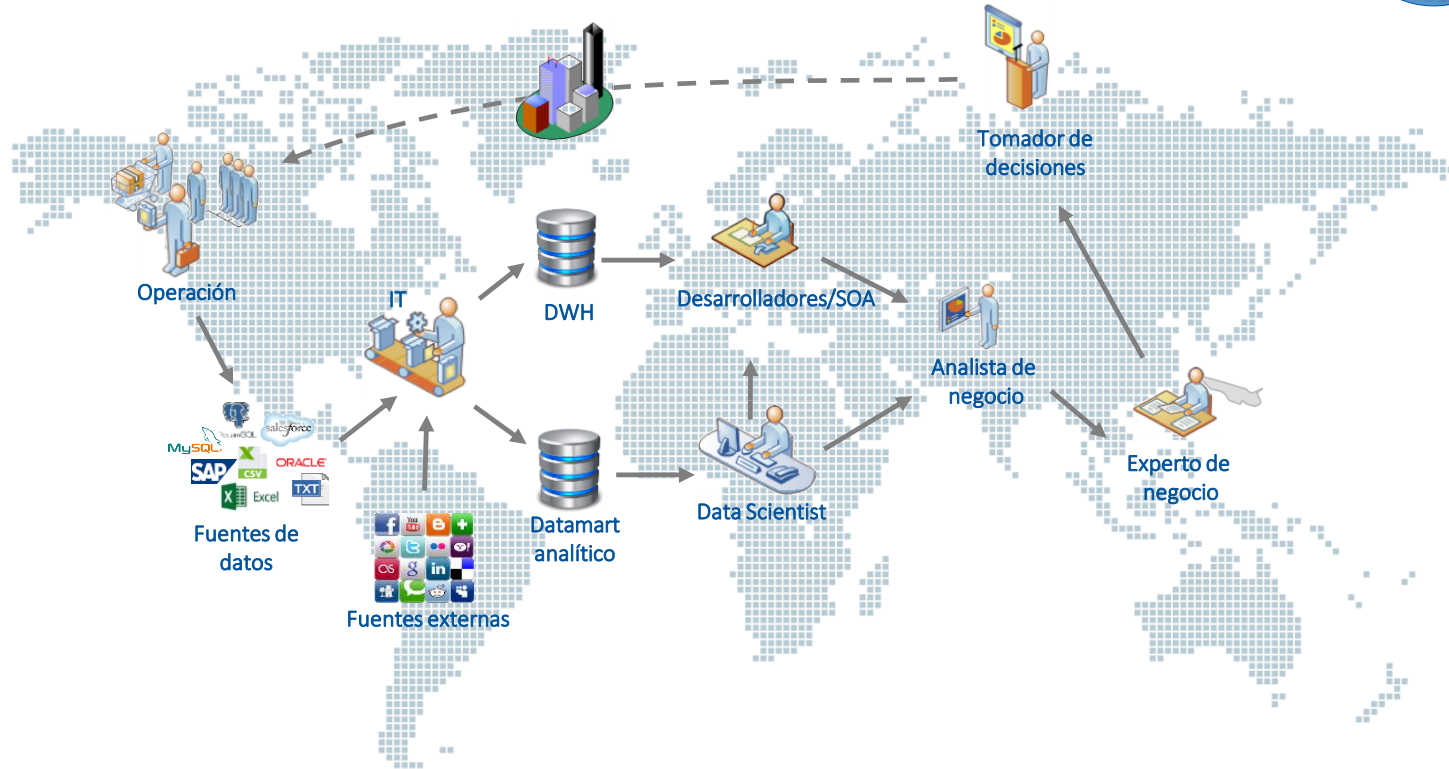
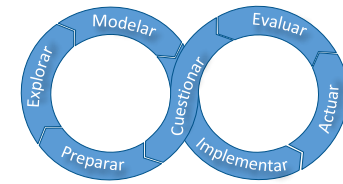
Equivalentes a:

-  150,000,000 iPhone
-  5,000,000 laptops
-  90 años de videos HD

Fuente: Northeastern University's Level Analytics



La ciencia de datos en la organización



Mapa estratégico

Beneficios

Aumentar lealtad del cliente

Aumentar las ventas

Aumentar el Market Share

Mejorar la percepción de la marca

Disminuir costos

Iniciativas

Retención de clientes

Posicionar un producto

Mejorar la experiencia del cliente

Incrementar la cartera de clientes

Conocer la satisfacción del cliente

Incrementar adquisición de productos

Reducción del riesgo crediticio

Reducir desgaste del cliente

Prevención de fraude

Marketing dirigido

Reducir tiempo de atención

Analítica

Credit Scoring

CLV

Modelos probabilistas

Simulación / OR

Minería de texto

Modelo de Churn

Análisis de secuencias Cross Sell / Up sell

Procesos estocásticos

Segmentación / perfilamiento

Datos

Clientes

Ventas

Demografía

Saldos

Sucursales

Canales

Proveedores

Transacciones

Facturación

Redes sociales

Portal web

Socio-demográfica

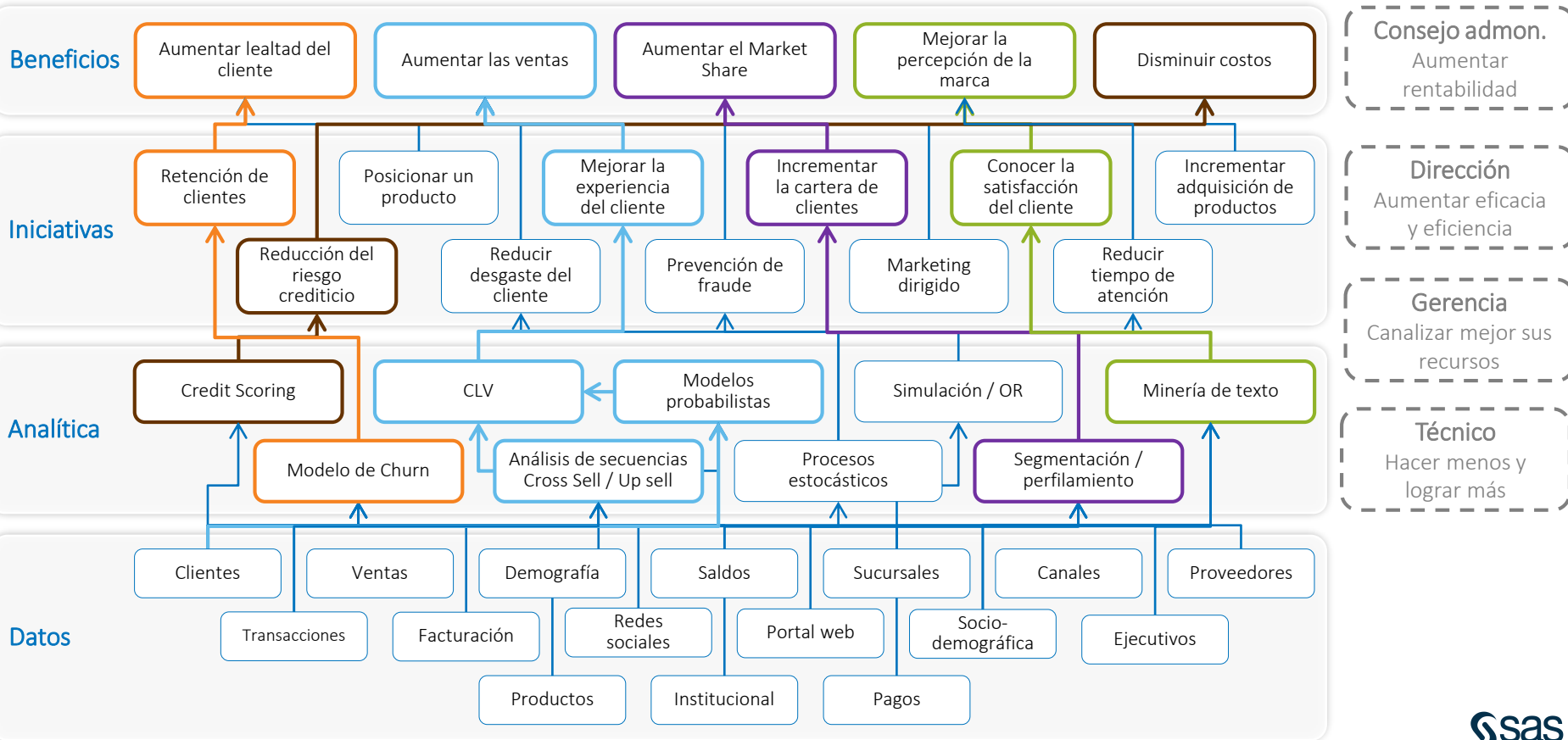
Ejecutivos

Productos

Institucional

Pagos

Mapa estratégico



A word cloud visualization of digital media and technology terms. The background is a light blue gradient with a subtle grid pattern. The words are arranged in a dense, overlapping manner, with some larger than others. Prominent words include "TECH", "DIGITAL", "WEB", "ONLINE", "MEDIA", "CONTENT", "SOCIAL", "NETWORK", "COMMUNICATION", "INFORMATION", "TELEVISION", "RADIO", "NEWSPAPER", "MAGAZINE", "JOURNAL", "BOOK", "FILM", "VIDEO", "AUDIO", "TEXT", "IMAGE", "GRAPHIC", "ANIMATION", "INTERACTIVE", "MULTIMEDIA", "HYPERTEXT", "HYPertext", "LINKED", "DOCUMENT", "PUBLICATION", "DISTRIBUTION", "CIRCULATION", "REACH", "IMPACT", "INFLUENCE", "ENGAGEMENT", "CONNECTION", "RELATIONSHIP", "COMMUNITY", "GROUP", "FORUM", "DISCUSSION", "DEBATE", "OPINION", "COMMENTARY", "CRITIQUE", "REVIEW", "ANALYSIS", "REPORT", "STUDY", "RESEARCH", "FINDING", "CONCLUSION", "SUMMARY", "OVERVIEW", "INTRODUCTION", "BACKGROUND", "CONTEXT", "FRAMEWORK", "METHODS", "RESULTS", "DISCUSSION", "CONCLUSIONS", "REFERENCES", "APPENDICES", "GLOSSARY", "INDEX", "TABLE OF CONTENTS", "EXECUTIVE SUMMARY", "ABSTRACT", "KEYWORDS", "TAGS", "HASHTAGS", "METADATA", "XML", "JSON", "CSV", "PDF", "DOCX", "PPTX", "MP3", "MP4", "AVI", "MOV", "WMV", "FLV", "SWF", "GIF", "PNG", "JPEG", "TIFF", "EPS", "AI", "PSD", "XD", "SKETCH", "FIGURE", "CHART", "GRAPH", "MAP", "TABLE", "FORM", "TEMPLATE", "LAYOUT", "DESIGN", "UI", "UX", "PROTOTYPING", "TESTING", "ITERATION", "REFinement", "RELEASE", "LAUNCH", "MARKETING", "PROMOTION", "ADVERTISING", "SALES", "CUSTOMER SERVICE", "SUPPORT", "TRAINING", "DEVELOPMENT", "MAINTENANCE", "UPDATES", "BUG FIXES", "SECURITY", "COMPLIANCE", "REGULATIONS", "STANDARDS", "BEST PRACTICES", "CASE STUDIES", "EXAMPLES", "TUTORIALS", "HOW-TO", "GUIDES", "CHECKLISTS", "TEMPLATES", "TOOLS", "SOFTWARE", "APPLICATIONS", "SERVICES", "PRODUCTS", "COMPANIES", "INDUSTRIES", "SECTORS", "MARKETS", "ECONOMY", "POLITICS", "CULTURE", "SOCIETY", "ENVIRONMENT", "HEALTH", "WELLNESS", "FOOD", "DRINKS", "TRAVEL", "ENTERTAINMENT", "SPORTS", "LEISURE", "HOBBIES", "INTERESTS", "PASSIONS", "VALUES", "BELIEFS", "OPINIONS", "ATTITUDES", "BEHAVIORS", "CHARACTERISTICS", "PERSONALITIES", "IDIOSYNCRASIES", "QUIRKS", "idiosyncrasies", "quirks". Various social media and technology icons are scattered throughout the word cloud, including the Twitter bird, Facebook 'f' logo, YouTube play button, RSS feed symbol, Apple App Store icon, Google Play Store icon, Amazon logo, Microsoft Windows logo, and various email client icons like Gmail, Outlook, and Yahoo Mail. Some icons are slightly tilted or rotated. The overall composition is dynamic and visually rich, representing the interconnectedness of digital media and technology.

¿Cómo tomar ventaja de la ciencia de datos en las organizaciones?

1. Definir una estrategia
2. Establecer un gobierno de datos
3. Formar un equipo especializado
4. Democratizar la analítica
5. Conectar el mundo offline y online

¿Cómo tomar ventaja de la ciencia de datos en las organizaciones?

1. Definir una estrategia



2. Establecer un gobierno de datos

3. Formar un equipo especializado

4. Democratizar la analítica

5. Conectar el mundo offline y online

¿Cómo tomar ventaja de la ciencia de datos en las organizaciones?

1. Definir una estrategia
2. Establecer un gobierno de datos
3. Formar un equipo especializado
4. Democratizar la analítica
5. Conectar el mundo offline y online



¿Cómo tomar ventaja de la ciencia de datos en las organizaciones?

1. Definir una estrategia
2. Establecer un gobierno de datos
3. Formar un equipo especializado
4. Democratizar la analítica
5. Conectar el mundo offline y online



¿Cómo tomar ventaja de la ciencia de datos en las organizaciones?

1. Definir una estrategia
2. Establecer un gobierno de datos
3. Formar un equipo especializado
4. Democratizar la analítica
5. Conectar el mundo offline y online



¿Cómo tomar ventaja de la ciencia de datos en las organizaciones?

1. Definir una estrategia
2. Establecer un gobierno de datos
3. Formar un equipo especializado
4. Democratizar la analítica
5. Conectar el mundo offline y online



Naturaleza Cossío



Naturaleza.Cossio@sas.com



[/in/naturalezacossio](https://www.linkedin.com/company/naturalezacossio)

sas.com