



Del Big Data al Smart Data

Cómo hacer nuestros negocios más inteligentes y productivos.

Luis Neira, Head of Latam
Hocelot

“El Big Data está de moda”

2018-2024

200 B\$ Tamaño del Mercado Global del Big Data

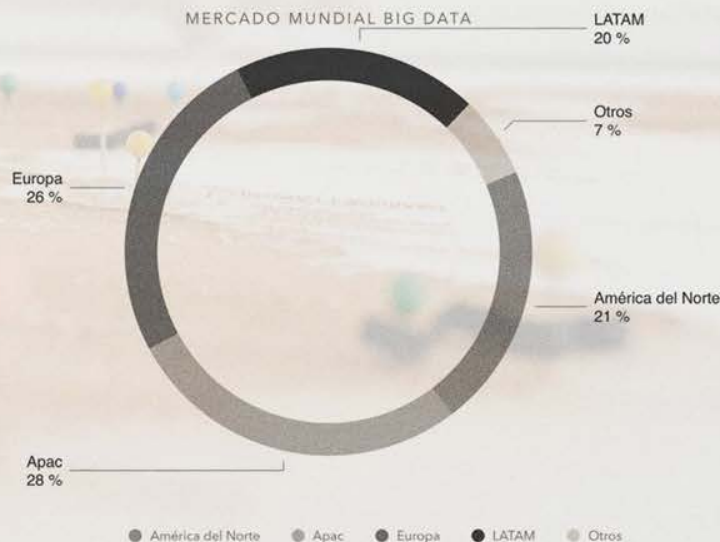
24,4%

Tasa de crecimiento interanual

Fuente: Gold Stein Research

HOCELOT™

Tamaño por Mercados - Continente bbs (2024):



**(bill. USD, crec. interanual)*

México en 2024

Liderará la carrera junto con Brasil.

26,7%

de total de las ventas en Latam .



85%

de los datos en
México son Dirty
Data

57%

es el incremento del
dirty data en México
entre 2017-2019

\$MEX 30MM

Pérdidas por Dirty Data en
2019



Smart Data:

Analiza la calidad del dato, su mejora, estandarización y enriquecimiento del mismo con hasta 1.500 nuevas variables.



Smart Analytics:

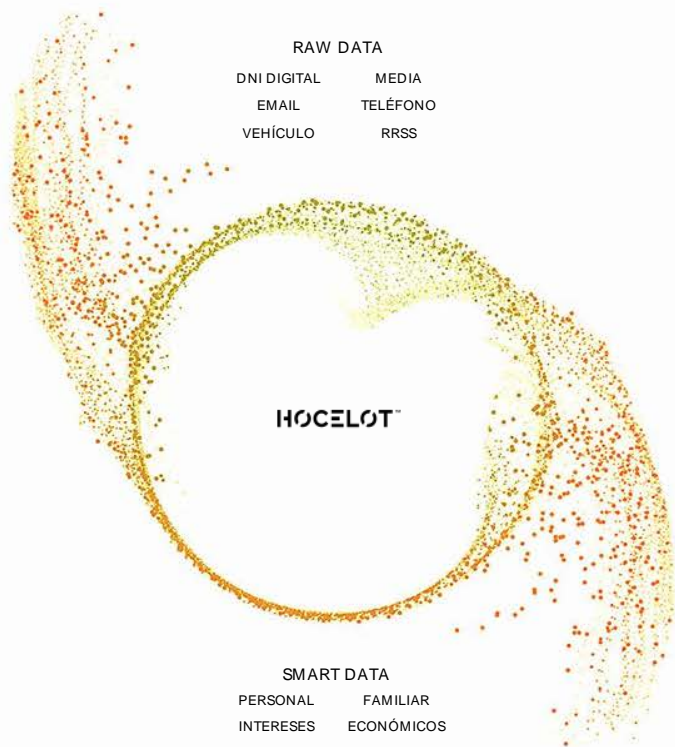
Mediante el uso de algoritmos extraen todo el potencial de las bases de datos y mejoran la toma de decisiones o Reduce costos y riesgos, vender más y solucionar problemas concretos



Las soluciones de analítica avanzada
incrementaron un 37% las ganancias de
empresas mexicanas en 2019



- 01 Se recibe la aplicación
- 02 Proceso de analítica
- 03 Toma de decisión



- 01 Se recibe la aplicación
- 02 Perfilado dinámico del cliente en tiempo real
- 03 Analítica avanzada (ML, DL, IA)
- 04 Toma de decisión



Alcance

Crear un modelo de aprendizaje supervisado o no supervisado que permita obtener segmentos de clientes dirigidos a la propensión de compra de productos.

58.988

Muestra inicial
Clientes

≈17%

Consentimiento

Visión Actual

- 1) Proceso de captación basado en reglas de negocio
- 2) Modelos estadísticos basados en transacciones
- 3) Toma de decisiones en base a la experiencia y criterio del experto

Nuevo Enfoque utilizado en el Caso de Uso

Fuentes de datos

Data lab

Analítica Avanzada

Cliente

Nombre
Fecha de nacimiento
Dirección
Email
Teléfono
Consentimiento
Entidad

Hoce lot

Id Check
Geo Check
Personal Enrichement
Economic Enrichement
Home Enrichement
Family Enrichement
Phone and email Verifcation

cliente_20190416.csv

Hoce lot's Real Time API

PoC basada en Machine Learnig

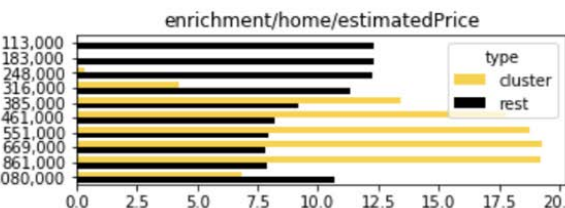
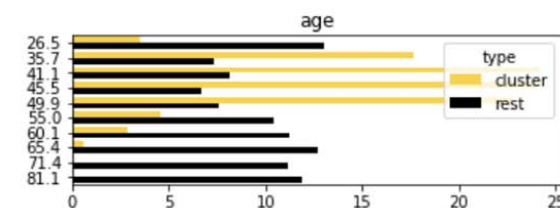
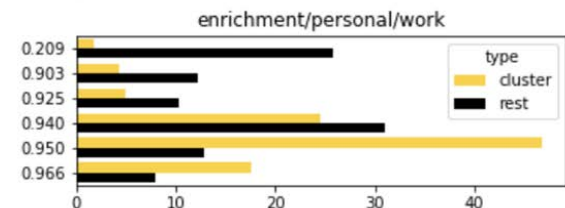
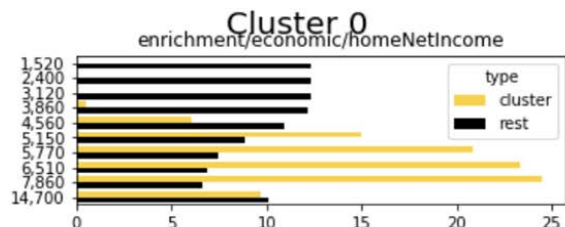
Posibilidad de incluir un gran número de variables, fuentes de datos y horizontes temporales

Posibilidad de automatizar el proceso de segmentación de clientes en base al aprendizaje realizado.

Toma de decisiones a futuro con mayor consistencia, en tiempo real y con fundamento matemático



Resultados: Profesionales liberales



Características

- Alto índice de empleabilidad
- Entre 35 y 50 años
- Ingresos familiares superiores a la media
- Alta capacidad de ahorro
- Invierten mucho en educación

Interés por:

- Productos de inversión. Este cluster puede tener un perfil avanzado en productos financieros, y estarán potencialmente interesados en fondos de inversión y asumir cierto riesgo.
- Productos de ahorro: por ejemplo, planes de pensiones.
- Préstamo para estudios. Con buenas condiciones por su alta solvencia.

Orden Contactabilidad

- Email
- Carta
- SMS
- Llamada



Alcance

Elaboración de modelos de aprendizaje automático para la detección temprana de mora, sobre la muestra de 6 meses un proceso de captación bajo un escenario de datos actualizados.

170.322

Muestra inicial
Clientes

- Definición de la variación de la muestra para cada versión: i.e. 125K -> 170K.
- 85 variables

≈ 50%

Clientes
activados

Visión Actual

- Proceso de captación basado en reglas de negocio
- Modelos estadísticos basados en transacciones
- Tom a de decisiones en base a la experiencia y criterio del experto

Nuevo Enfoque utilizado en el Caso de Uso

Fuentes de datos

Data lab

Analítica Avanzada

Cliente

Nombre
Fecha de nacimiento
Dirección
Email
Teléfono
Importes Impagados
Facturas devueltas
Facturaciones medias

Hoce lot

Id Check
Geo Check
Personal Enrichement
Economic Enrichement
Home Enrichement
Family Enrichement
Phone and email Verification

DATOS_HOCELOT_TARGET_TOTAL.csv

Hoce lot's Real Time API

PoC basada en Machine Learnig

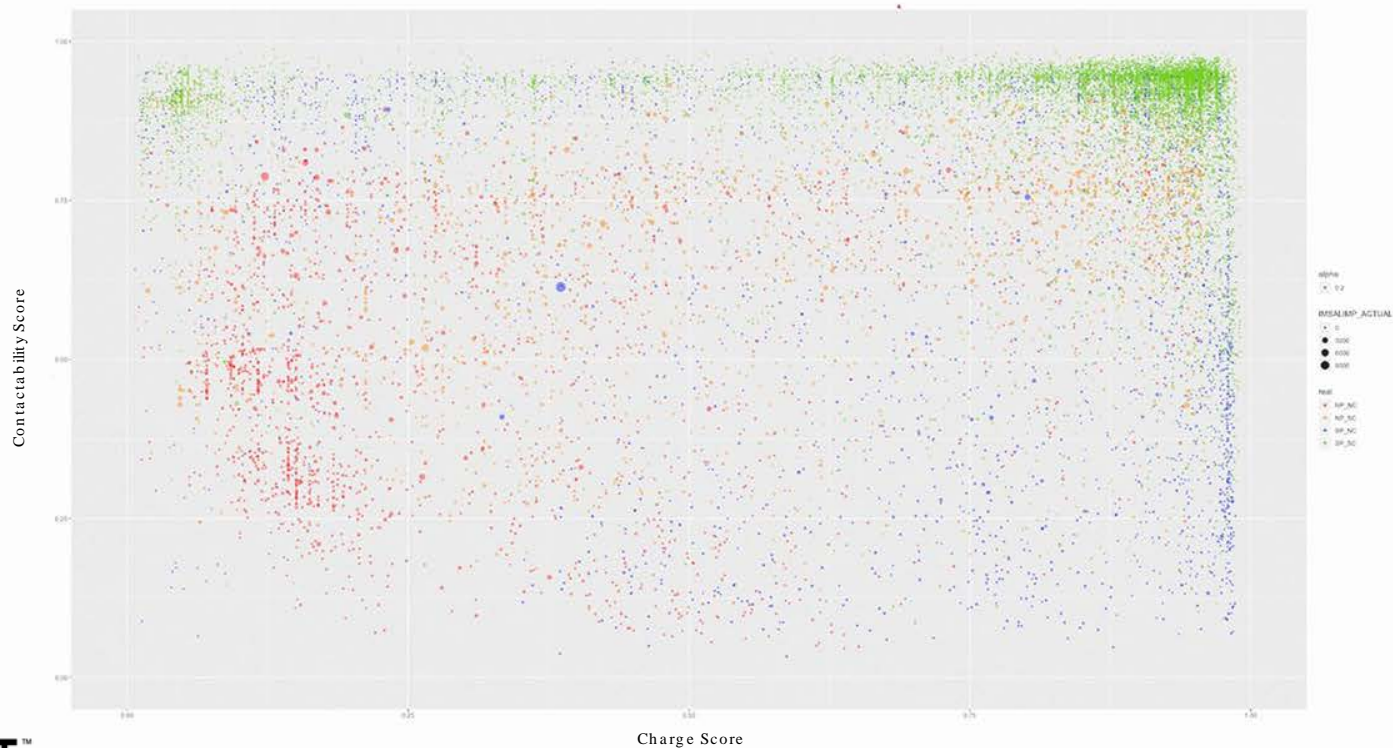
Posibilidad de incluir un gran número de variables, fuentes de datos y horizontes temporales

Posibilidad de automatizar el proceso de captación de clientes en base a la im probabilidad de im pago

Tom a de decisiones a futuro con mayor consistencia, en tiempo real y con fundamento matemático



Tarjetas de Crédito 0- 30 días





Decide better