

Introducción

Analítica predictiva de datos

Agenda

- ✓ Introducción al Analítica predictiva.
- ✓ Supervised Learning.
- ✓ Unsupervised Learning.
- ✓ Reinforcement Learning.
- ✓ Deep Learning.
- ✓ Aplicación de Analítica predictiva.
- ✓ Consideraciones finales.



Compitiendo con Analítica

Competir con analítica es implementar el uso intensivo de datos, análisis estadísticos con data cuantitativa y cualitativa, modelos explicativos, modelos predictivos y una gestión basada en eventos para impulsar decisiones y acciones.

“Competing on Analytics”

Enterprise Analytics. Thomas Davenport. Pearson Publishing.

Tipos de preguntas analíticas

	Descriptivas	Predictivas	Prescriptivas
Preguntas	<p>¿Qué esta pasando?</p> <p>¿Qué genero esta situación?</p> <p>¿Cuál es el problema exactamente?</p> <p>¿Qué acciones son necesarias?</p>	<p>¿Qué va a pasar si no hacemos nada?</p> <p>¿Qué tan probable es que ocurra?</p> <p>¿Si cambiamos algo cuál será el efecto esperado?</p>	<p>¿Qué deberia hacer?</p> <p>¿Porqué debería hacerlo?</p> <p>¿En que rango obtengo un escenario óptimo?</p> <p>¿Me conviene aplicar este cambio?</p>
Instrumentos	<ul style="list-style-type: none"> • Reportes • Dashboards • Tablas • Alertas 	<ul style="list-style-type: none"> • Data Mining • Text Mining • Web/Media Mining • Forecasting 	<ul style="list-style-type: none"> • Optimización • Simulación • Modelamiento
Resultados	<p>Problemas y oportunidades bien indentificados.</p>	<p>Proyecciones confiables del comportamiento de un proceso.</p>	<p>Proponer el escenario que nos da mayor beneficio.</p>

 **¡En las empresas del grupo Cencosud queremos trabajar con los MEJORES!**

DATA SCIENTIST

Si te apasiona el mundo de los datos, tienes mucha curiosidad, poner tu talento en practica en escenarios reales, te invitamos a postular para el puesto de **Practicante Profesional**.

Requerimos que tengas el conocimiento de : R, Python

Postula y sé parte de

Envía tu CV a: Maryorie.najar@cencosud.com
Asunto: CV Data Scientist

Breca Innovation Hub está buscando un PRACTICANTE de ADVANCED ANALYTICS!

Buscamos personas apasionadas por el mundo de los datos y que quieran aprender a desarrollar y construir soluciones analíticas.

Como integrante del equipo de Data Scientists del Center for Advanced Analytics del Grupo BRECA, tendrás la oportunidad de aprender sobre diversos retos sobre Data & Analytics, aplicados a problemas reales de las diferentes unidades de negocio!

Requisitos:

- * Estudiante entre 7mo y 9no ciclo (Ing Sistemas, Ing Informática, Ing Estadística o afines)
- * Conocimientos en Bases de Datos (SQL)
- * Conocimientos en R y/o Python orientado a análisis de datos
- * ¡Muchas ganas de aprender!

Si te encuentras interesado por favor envía tu cv a anapaula@brein.pe hasta el 18 de Agosto

Únete al Centro de Analítica Avanzada

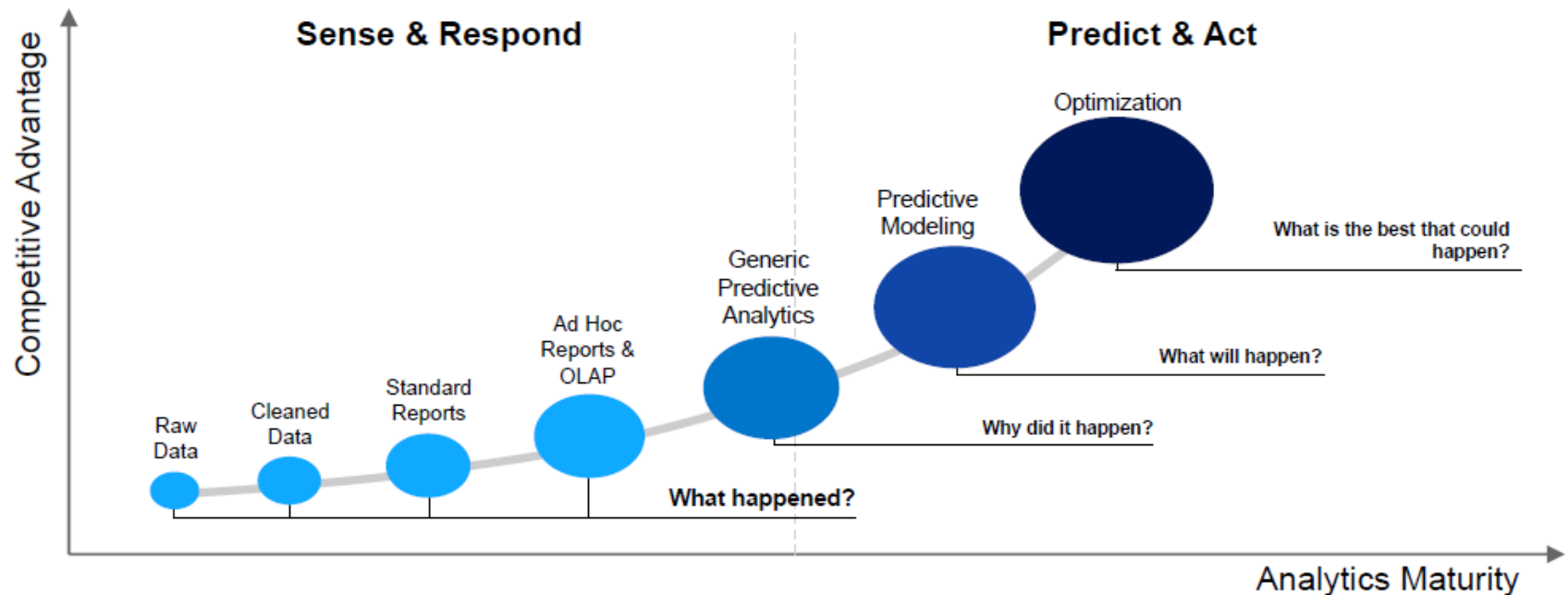
del Hub de Innovación del Grupo Breca



Compitiendo con Analítica

Predictive Analytics

Customer needs and the value propositions



The key is unlocking data to move decision making from sense & respond to predict & act

Fuente: Sap Hana

¿Qué es el Analítica predictiva

Traditional Programming



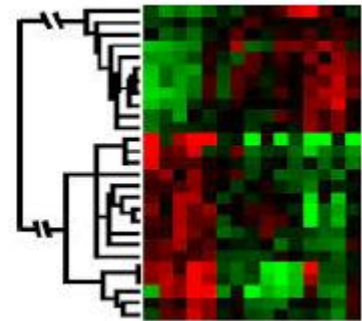
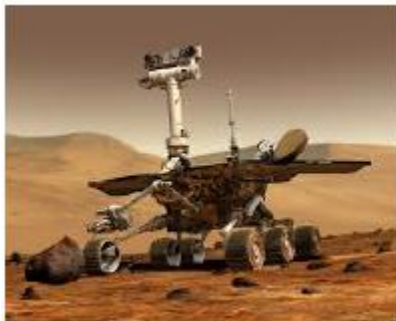
Machine Learning



¿Qué es el Analítica predictiva?

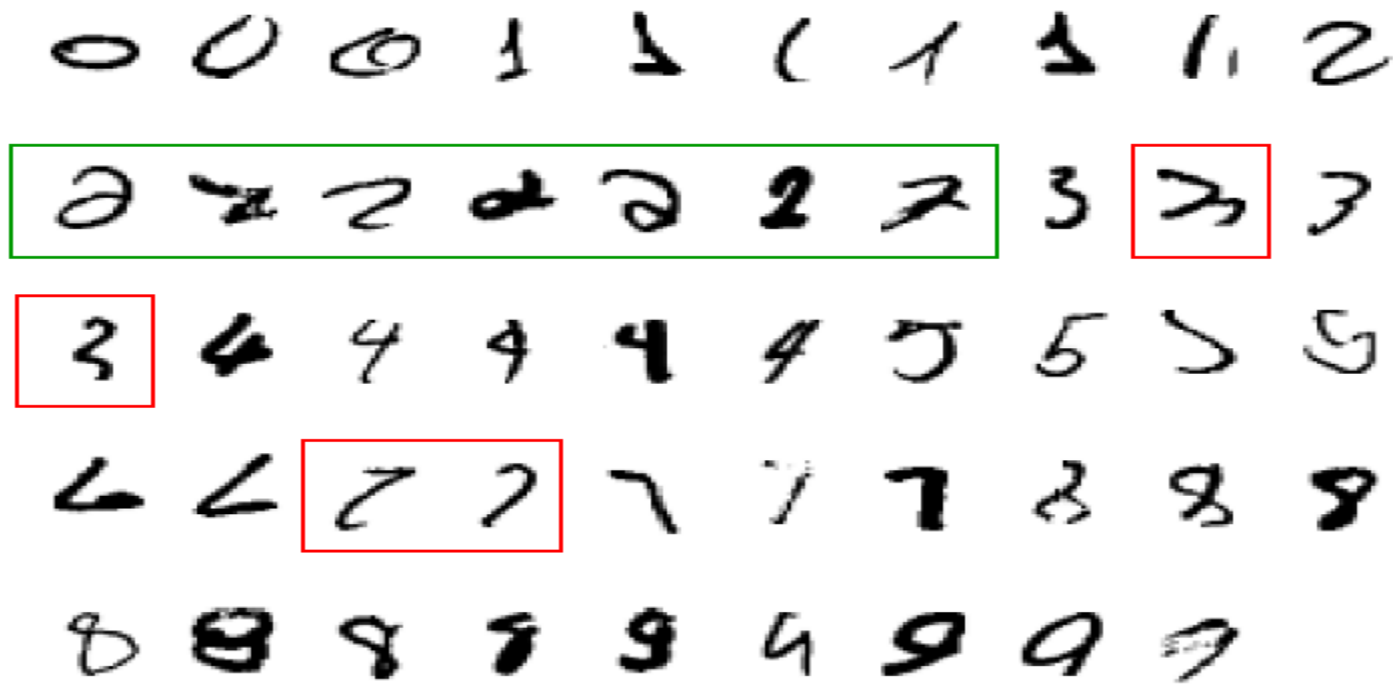
➤ Analítica predictiva se usa cuando:

- La experiencia humana no existe (navegar en Marte)
- Los humanos no pueden explicar su experiencia (reconocimiento de voz)
- Los modelos deben ser personalizados (medicina personalizada)
- Los modelos se basan en grandes cantidades de datos (genómica)



¿Qué es el Analítica predictiva?

- ✓ Un ejemplo clásico de una tarea que requiere aprendizaje automático: Es muy difícil decir qué hace un 2



¿Qué es el Analítica predictiva?

➤ **Multitud de algoritmos** con aplicaciones específicas.

➤ Ramas de Analítica predictiva:

- ✓ Supervised Learning
- ✓ Unsupervised Learning
- ✓ Reinforcement Learning
- ✓ Deep Learning

Una breve historia de Analítica predictiva

- **1950s**

- Samuel's checker player
- Selfridge's Pandemonium

- **1960s:**

- Neural networks: Perceptron
- Pattern recognition
- Learning in the limit theory
- Minsky and Papert prove limitations of Perceptron

- **1970s:**

- Symbolic concept induction
- Winston's arch learner
- Expert systems and the knowledge acquisition bottleneck
- Quinlan's ID3
- Michalski's AQ and soybean diagnosis
- Scientific discovery with BACON
- Mathematical discovery with AM

Una breve historia de Analítica predictiva

- **1980s:**

- Advanced decision tree and rule learning
- Explanation-based Learning (EBL)
- Learning and planning and problem solving
- Utility problem
- Analogy
- Cognitive architectures
- Resurgence of neural networks (connectionism, backpropagation)
- Valiant's PAC Learning Theory
- Focus on experimental methodology

- **1990s**

- Data mining
- Adaptive software agents and web applications
- Text learning
- Reinforcement learning (RL)
- Inductive Logic Programming (ILP)
- Ensembles: Bagging, Boosting, and Stacking
- Bayes Net learning

Una breve historia de Analítica predictiva

- **2000s**

- Support vector machines & kernel methods
- Graphical models
- Statistical relational learning
- Transfer learning
- Sequence labeling
- Collective classification and structured outputs
- Computer Systems Applications (Compilers, Debugging, Graphics, Security)
- E-mail management
- Personalized assistants that learn
- Learning in robotics and vision

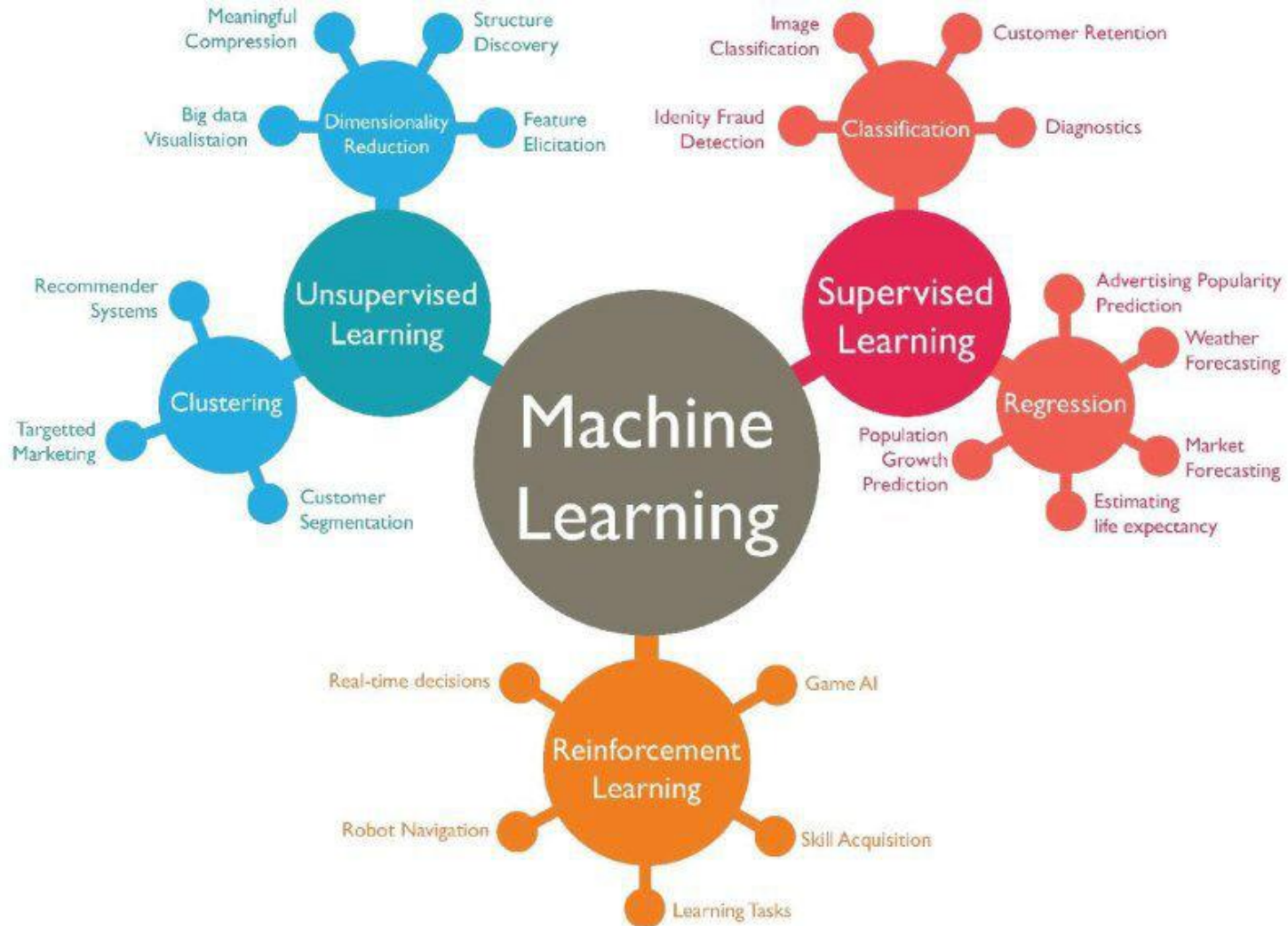
- **2010s**

- Deep learning systems
- Learning for big data
- Bayesian methods
- Multi-task & lifelong learning
- Applications to vision, speech, social networks, learning to read, etc.
- ???

Definiciones Básicas

- **Conjunto de Datos (Data Set):** El total del conjunto de datos sobre los que queremos desarrollar un algoritmo de Analítica predictiva con el fin de obtener un modelo que lo represente lo mejor posible. Contendrá variables independientes y dependientes.
- **Variables Independientes (Features), (VI):** Aquellas columnas del Data Set que serán usadas por el algoritmo para generar un modelo que prediga lo mejor posible las variables dependientes.
- **Variables dependientes (Labels,Target), (VD):** Columna del data set que responde a una correlación de VI y que debe ser predicha por el futuro modelo
- **Conjunto de Datos de Entrenamiento (Training Set):** Subconjunto del Data Set que será utilizado para entrenar el modelo que se pretende generar.
- **Conjunto de Datos de Test (Test Set):** Subconjunto del data set que se le pasará al modelo una vez haya sido entrenado para comprobar, mediante el uso de diferentes métricas, sus indicadores más importantes de calidad.

Introducción



Supervised Learning (Modelos Supervisados)

- Se tiene una **variable objetivo** (Variable de Salida).
- Variables que ayudan a predecir a la variable de salida (Variables de entrada).
- Existe una dependencia de las variables de entrada con las variables de salida.



Supervised Learning

- Género.



- Rangos de Edad.



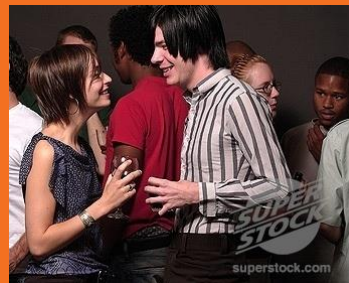
➤ Si Compra

- Ingresos.



➤ No Compra

- Estado Civil.



Unsupervised Learning (Modelos No Supervisados)

- No hay una variable objetivo (Variable de Salida).
- No hay variables que ayudan a predecir a la variable de salida.



- Todas las variables tienen la misma importancia.
- Se busca la interdependencia de las variables.

Modelos No Supervisados



UNIVERSIDAD
DE LIMA

Reinforcement Learning (Aprendizaje por Refuerzo)

- El algoritmo de aprendizaje recibe un tipo de valoración acerca de la idoneidad de la respuesta dada.
- Cuando la decisión es correcta es muy parecido al aprendizaje supervisado, sin embargo difiere mucho cuando la decisión es incorrecta.



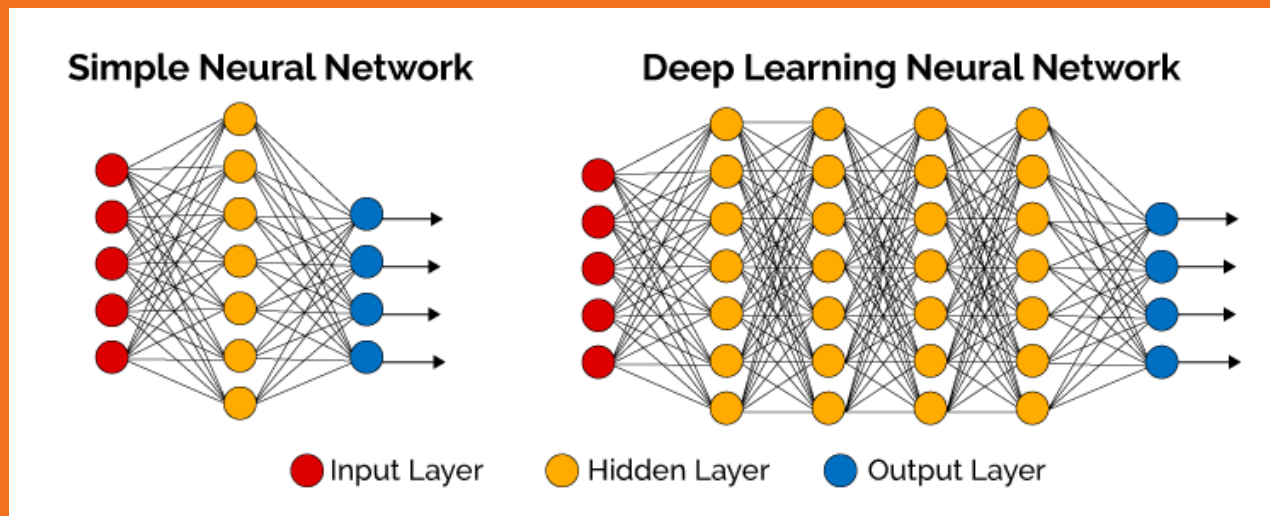
Reinforcement Learning (Aprendizaje por Refuerzo)



<https://www.youtube.com/watch?v=4cgWya-wjgY>

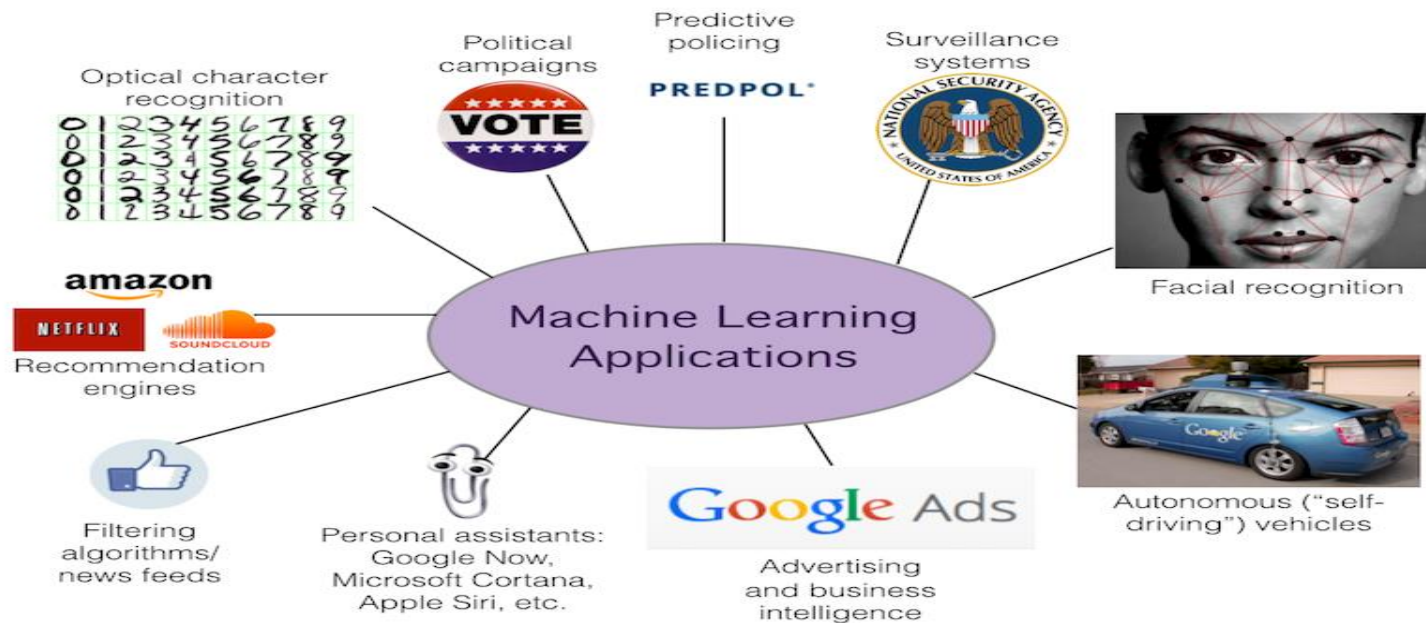
Deep Learning (Aprendizaje Profundo)

- Es un conjunto de algoritmos de Analítica predictiva que intenta modelar abstracciones de alto nivel usando arquitecturas compuestas como redes neuronales profundas, redes neuronales convolucionales y redes de creencia profunda para resolver problemas como visión del computador, reconocimiento automático del habla, reconocimiento del audio y música, etc.

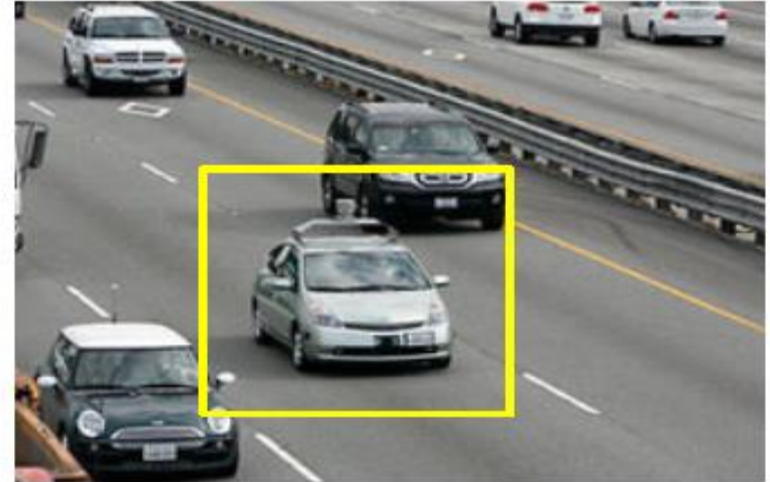


Aplicaciones Analítica predictiva

- Sistemas de Recomendación.
- Detección de Spam.
- Natural Language Processing (NLP).
- Photo OCR (Optical Character Recognition).
- Visión Artificial.
- Diagnósticos médicos.
- Conducción Autónoma.
- AMD Ryzen ANN.



Aplicaciones Analítica predictiva

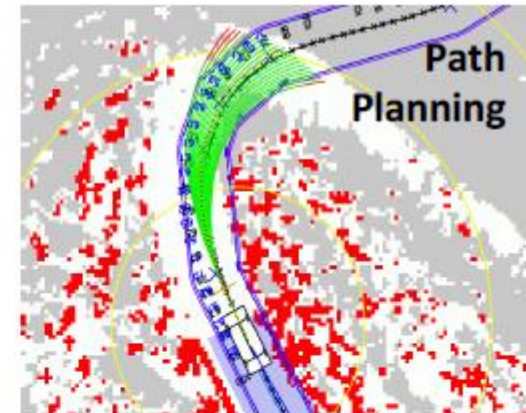


- Nevada lo hizo legal para autos autónomos para conducir carreteras en junio de 2011
- A partir de 2013, cuatro estados (Nevada, Florida, California y Michigan) han legalizado autos autónomos



Autos autónomos

Aplicaciones Analítica predictiva



Images and movies taken from Sebastian Thrun's multimedia website.

Tecnologías autos autónomos

Aplicación del Analítica predictiva

Deserción Académica



Fuga de Clientes



Lavado de Activos



Detección de Fraudes



Generar Perfiles o Grupos



Venta Cruzada

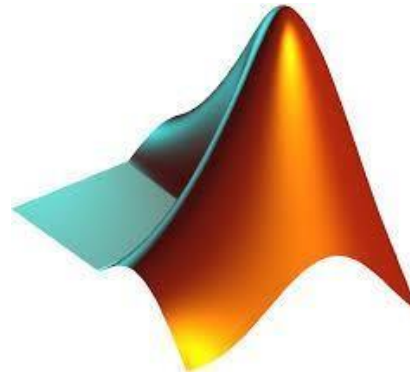
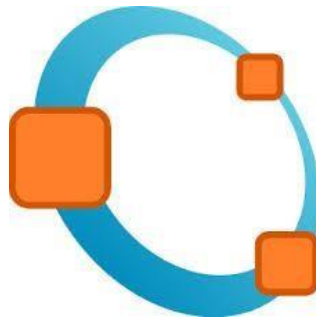


Panorama Tecnológico : Softwares

Analítica predictiva



SPSS Modeler



A word cloud of the word "thank you" in various languages. The words are arranged in a circular shape, with the most prominent words in the center. The colors of the words vary, including shades of blue, green, red, orange, and purple. The words include:

- mercí
- obrigado
- danke
- gracias
- teşekkür ederim
- thank you
- sukriya
- hvala
- vinaka
- shukriya
- matondo
- bayaralaa
- bedankt
- dank je
- tapadh leat
- takk
- manana
- nanni
- agat
- maith
- raibh
- go
- dekuji
- asante
- grazzi
- gratias ago
- leaki
- tau
- welalin
- spas
- rahmat
- ació
- dakujem
- krap
- khun
- kop
- mesí
- grazie
- blagodaram
- nandri
- akun
- djere dieuf
- sobodi
- lack
- rahmat
- mamnun
- dankon
- dhanyavadagalu
- xiexie
- taalefai lava
- gracies
- merce
- kia ora
- murakoze
- sagolun
- taiku
- arigato
- enkosi
- terima kasih
- ngiyabonga
- chnearakaloutioun
- mochhakkeram
- trugarez
- chekrane
- koszulom
- barka
- dhanyavad
- mauruuru
- misaotra
- lanemint
- mahalo