



PREDICIENDO EL FUTURO CON AZURE MACHINE LEARNING + APLICACIÓN EN XAMARIN

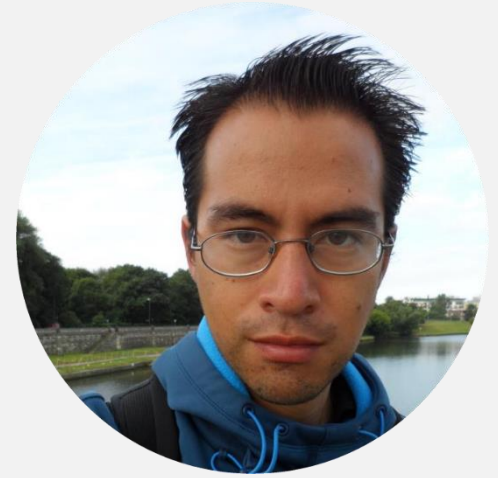
#SGVirtual

Presenta:



@darkicebeam
ISC Luis Beltrán

- Soy un Estudiante de Doctorado en Ingeniería Informática en la Universidad Tomás Bata en Zlín, República Checa.
- También soy docente en el Instituto Tecnológico de Celaya, Guanajuato, México.
- Pertenezco a la red de Microsoft Student Partners en la República Checa y Xamarin Student Partners a nivel mundial.
- Me interesa:
 - Desarrollo de aplicaciones móviles.
 - Inteligencia Artificial.
 - Cómputo en la nube.
 - Compartir mi experiencia y conocimientos



Agenda

1. Introducción
2. Machine Learning
3. Azure Machine Learning
4. Demo

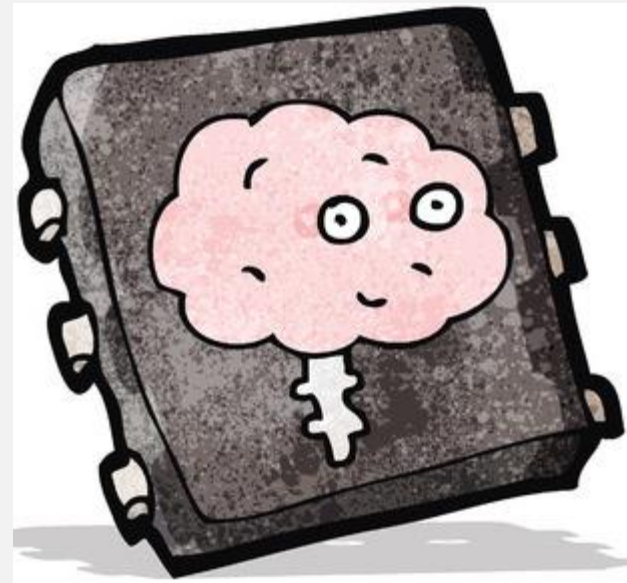
Conclusiones



1. Introducción

¿Qué es la Inteligencia Artificial? (IA)

- Es un campo de estudio que busca explicar y emular el comportamiento inteligente en términos de procesos computacionales.
- Toma influencia de campos tan diversos como las Matemáticas, la Filosofía, la Psicología, la Biología, la Neurociencia y otros más.



LA INTELIGENCIA ARTIFICIAL



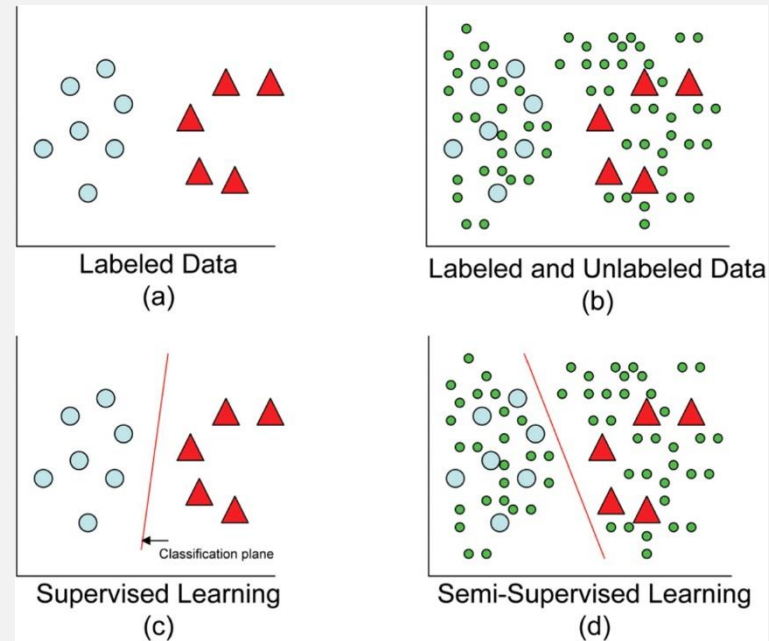
Ahora son esenciales en nuestras vidas, no solo sirven para contar sino para un sinfín de cosas

- La IA implica áreas genéricas:

- Aprendizaje
- Percepción
- Solución de problemas

- Y áreas específicas:

- Jugar una partida de ajedrez
- Diagnosticar enfermedades
- Conducir un automóvil de manera autónoma



¿Por qué es importante la IA hoy en día?

- Es una ventaja competitiva en las empresas.
- Los datos son el **arma secreta** de toda compañía.
- Los datos alimentan el aprendizaje de algoritmos inteligentes para toma de decisiones y predicción del futuro (tendencias, patrones).
- No pretende reemplazar a las personas, sino colaborar con ellas.



Fuente: <http://elliottvelmont.blogspot.mx/>

HOY EN DÍA ES MUY COMÚN VER
COMPUTADORAS EN NUESTRAS
ACTIVIDADES



LA TECNOLOGÍA ESTÁ
REBASANDO FRONTERAS



EN UN PRINCIPIO SE USO
PARA ACTIVIDADES DE
RIESGO



AHORA NOS FACILITAN LAS
ACTIVIDADES DIARIAS



HOY EN DÍA ES MUY COMÚN VER
COMPUTADORAS EN NUESTRAS
ACTIVIDADES



LA TECNOLOGÍA ESTÁ
REBASANDO FRONTERAS



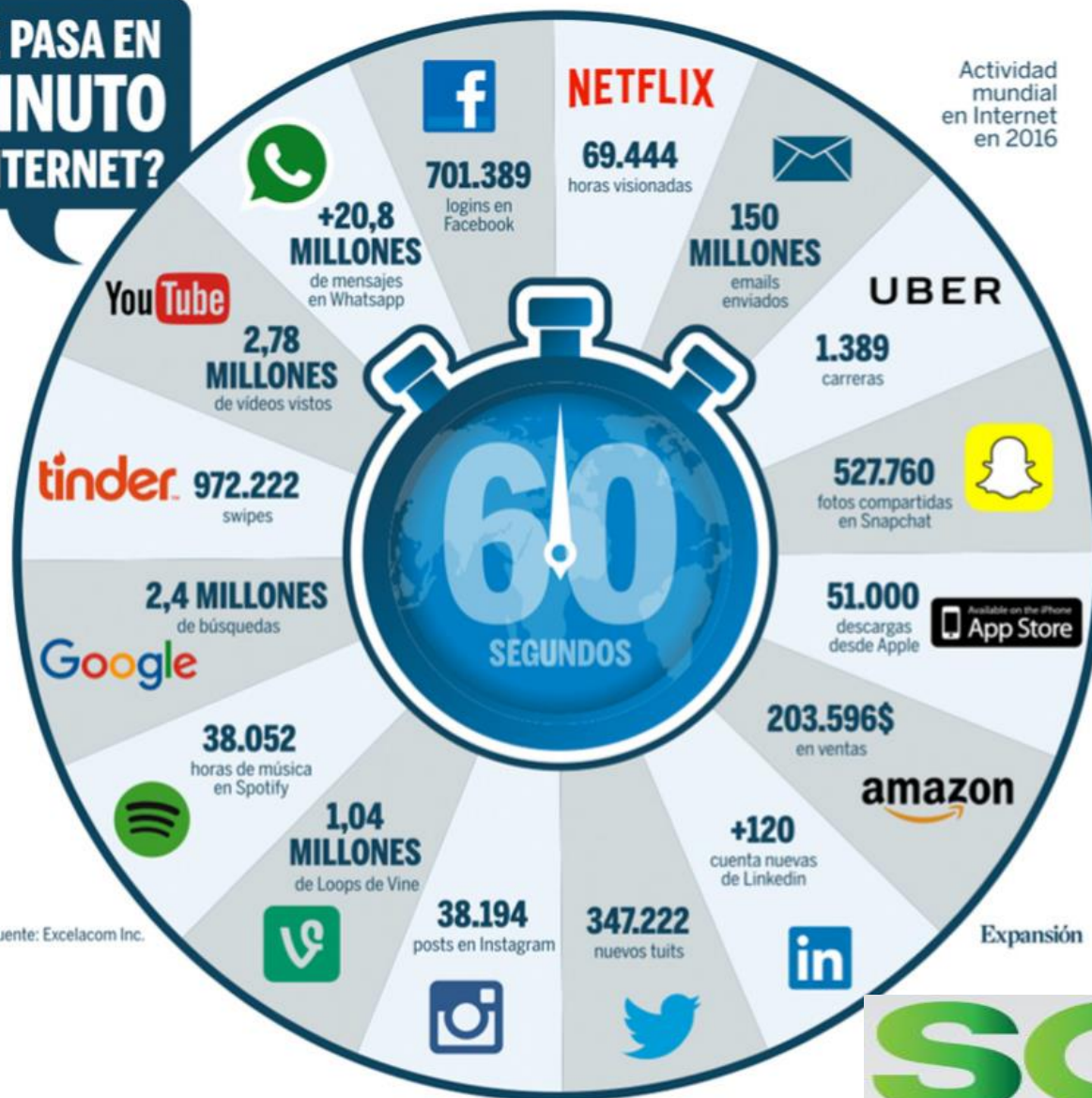
EN UN PRINCIPIO SE USO
PARA ACTIVIDADES DE
RIESGO



AHORA NOS FACILITAN LAS
ACTIVIDADES DIARIAS



¿QUÉ PASA EN 1 MINUTO EN INTERNET?



<http://www.strategiaonline.es/que-pasa-en-un-minuto-en-internet-en-2016/>

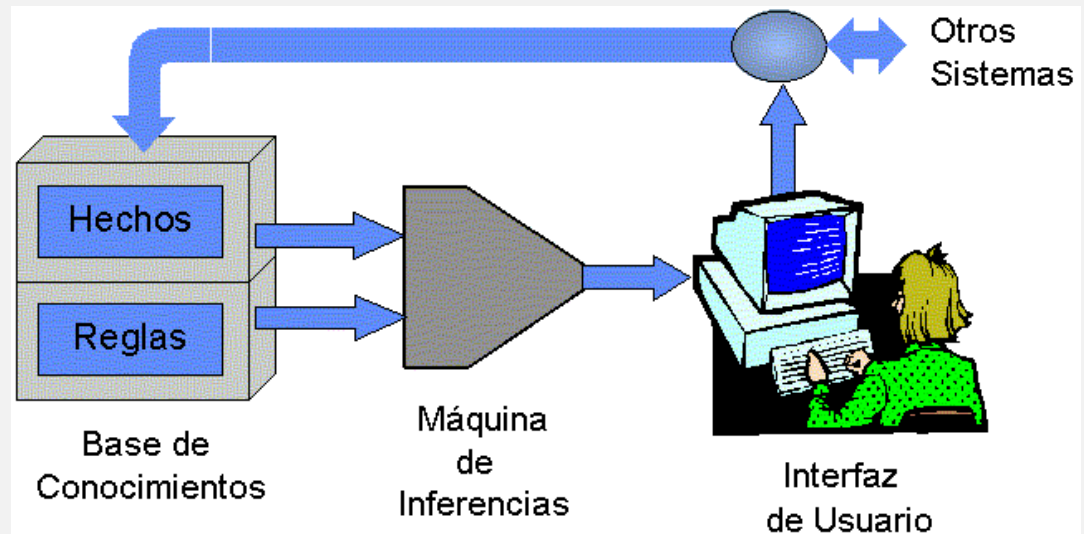
Áreas de la Inteligencia Artificial

- Sistemas Expertos
- Robótica
- Procesamiento de Lenguaje Natural
- Visión de computadora
- Redes Neuronales
- Lógica Difusa
- Algoritmos Genéticos
- Realidad Virtual
- **Aprendizaje Automático**



Áreas de la Inteligencia Artificial

- Sistemas Expertos



- Robótica



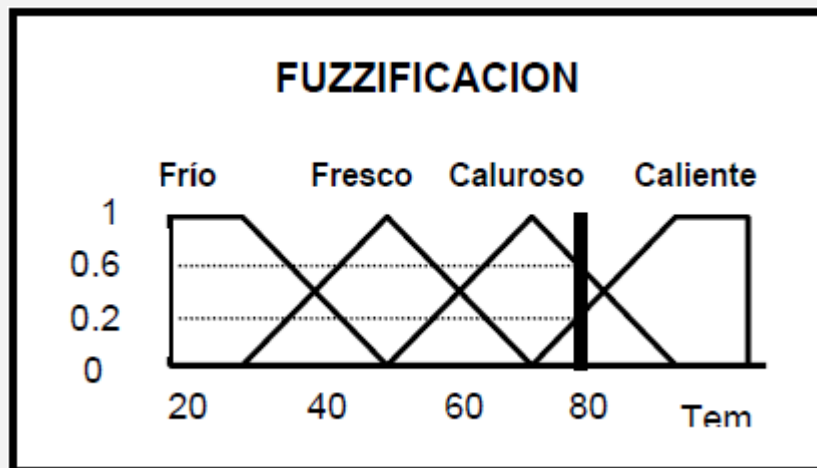
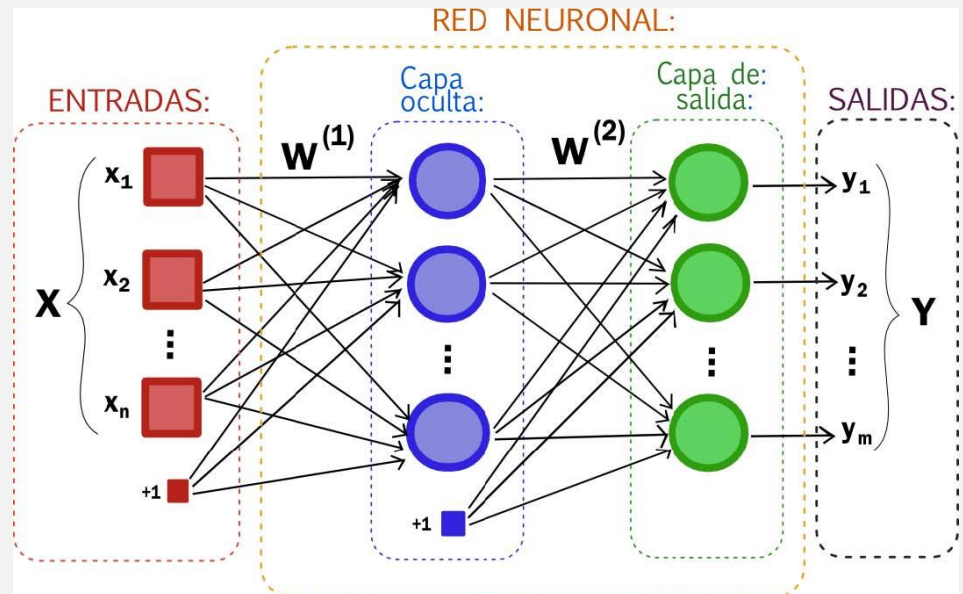
Áreas de la Inteligencia Artificial

- Procesamiento de Lenguaje Natural
- Visión de Computadora



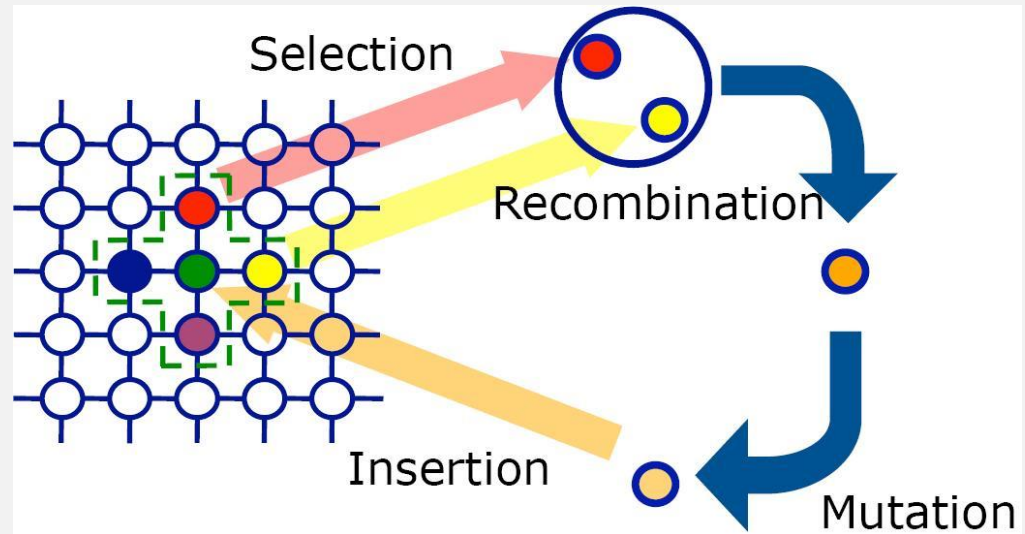
Áreas de la Inteligencia Artificial

- Redes Neuronales
- Lógica Difusa



Áreas de la Inteligencia Artificial

- Algoritmos Genéticos

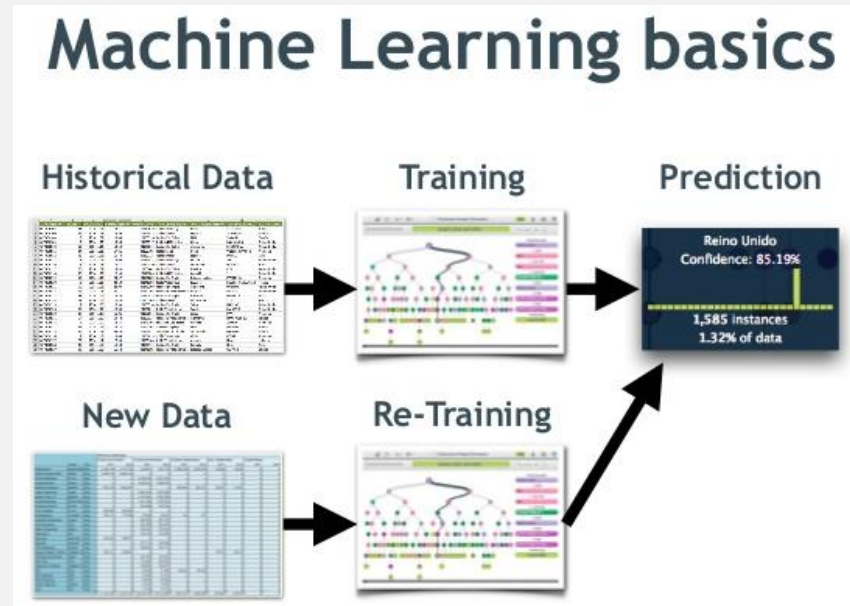


- Realidad Virtual

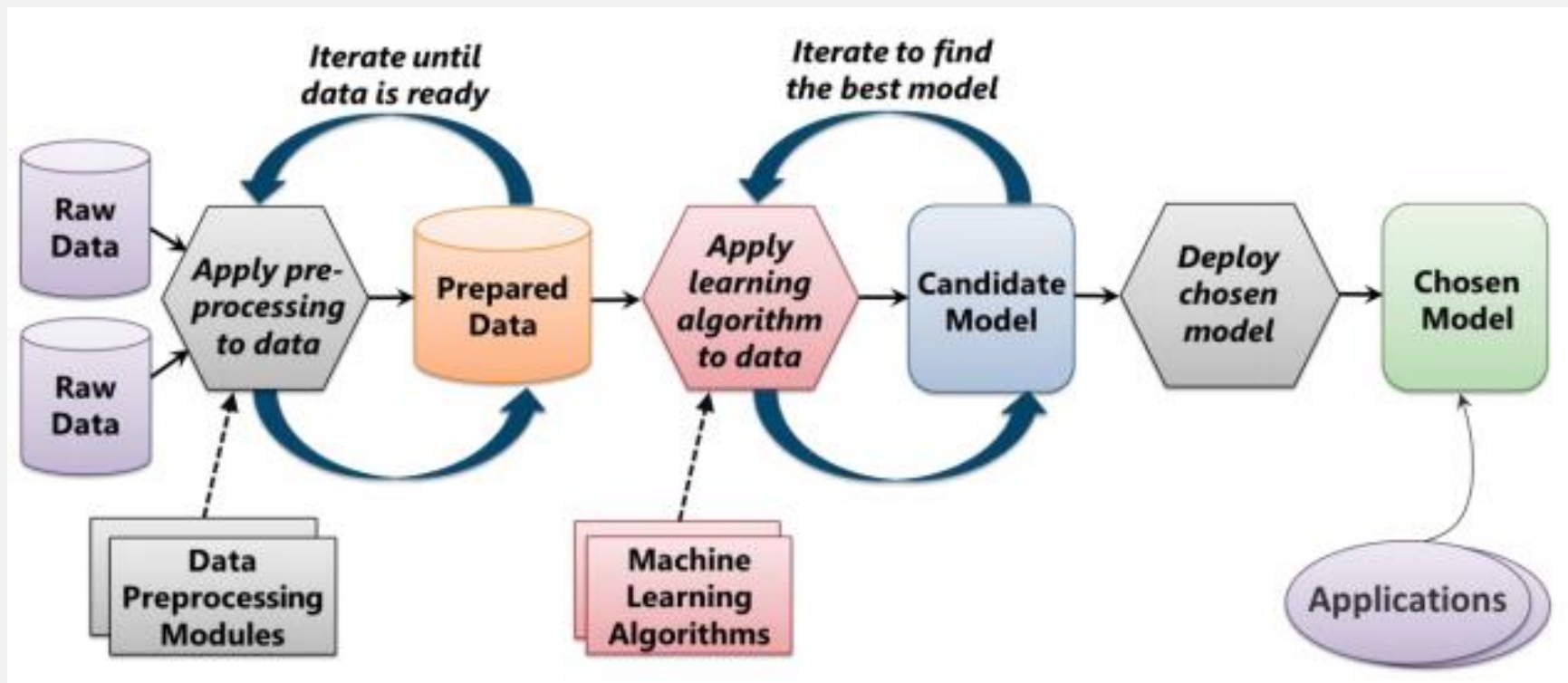


2. Machine Learning

- El Aprendizaje Automático (ML) es un tipo de Inteligencia Artificial capaz de aprender sin ser programado de forma explícita
- Se enfoca en la construcción de programas que pueden adaptarse por sí mismos cuando se presentan nuevos datos.



- Los datos pueden contener información valiosa que no se percibe a primera vista. Si se tiene una cantidad enorme de datos, se puede hacer un análisis inteligente para encontrar patrones que pueden decirnos cómo resolver un problema.
- ML obtiene un modelo que reconoce esos patrones en datos nuevos, diferentes a los utilizados en la fase de construcción del modelo (entrenamiento).



- Entre más información se tenga al inicio, más confiable será el modelo construido.
- Las características de los datos a analizar tienen que ser elegidos de manera cuidadosa.
- Normalmente los datos se pre-procesan a fin de tener datos preparados (limpios).
- Los algoritmos de ML aplican análisis estadístico sobre los datos a fin de generar un modelo que los describa.
- El modelo es utilizado en alguna aplicación a fin de resolver un problema u analizar datos nuevos.
- El modelo generado tienen un nivel de confiabilidad que depende de varios factores.

<i>Name</i>	<i>Amount</i>	<i>Where Issued</i>	<i>Where Used</i>	<i>Age of Cardholder</i>	<i>Fraudulent</i>
Smith	\$2,600.45	USA	USA	22	No
Potter	\$2,294.58	USA	RUS	29	Yes
Peters	\$1,003.30	USA	RUS	25	Yes
Adams	\$8,488.32	FRA	USA	64	No
Pali	\$200.12	AUS	JAP	58	No
Jones	\$3,250.11	USA	RUS	43	No
Hanford	\$8,156.20	USA	RUS	27	Yes
Marx	\$7,475.11	UK	GER	32	No
Norse	\$540.00	USA	RUS	27	No
Edson	\$7,475.11	USA	RUS	20	Yes

The truth is that the pattern the data supports is this: A transaction is fraudulent if the cardholder is in his 20s, the card was issued in the USA and used in Russia, and the amount is more than \$1,000. With some time, you probably would have figured this out, since the data you have to work with isn't very large.

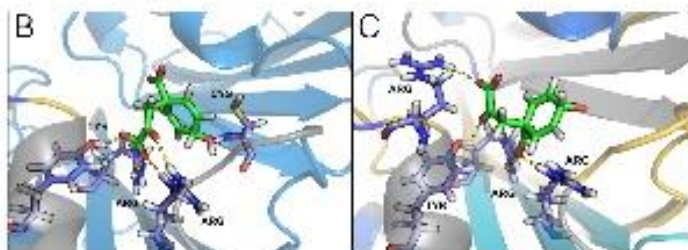
Examples of Machine Learning

Text Recognition



http://commons.wikimedia.org/wiki/File:American_book_company_1916_letter_grydcoe-2.JPG#filelinks
[public domain]

Biology



<https://doi.org/10.1371/journal.pone.0241666.g001>

Spam Filtering

Machine Learning:

It is used in more ways than you think!



Other Movies You Might Enjoy



ML wins Jeopardy!



Consideraciones

- ML puede utilizar millones de registros, cada uno con un número elevado de propiedades que, combinadas, permiten entrenar un modelo a partir de un algoritmo basado en técnicas estadísticas o heurísticas para la generación de conocimiento a partir de nuevos datos.

3. Azure Machine Learning



Azure
Machine Learning

<https://studio.azureml.net/>



- Azure Machine Learning es un servicio completamente administrado desde la nube que facilita la construcción e implementación de soluciones analíticas de predicción, clasificación, categorización y detección de fraudes.
- El proceso de aprendizaje automático no es sencillo. Para facilitar a los científicos de datos o programadores esta tarea, Azure Machine Learning proporciona diferentes componentes.

Microsoft Azure Machine Learning | Home Studio Gallery

Binary Classification: Direct marketing

In draft

Properties

Two-Class Boosted Decision Tree

Create trainer mode
Single Parameter

Maximum number of leav...
20

Minimum number of sam...
10

Learning rate
0.2

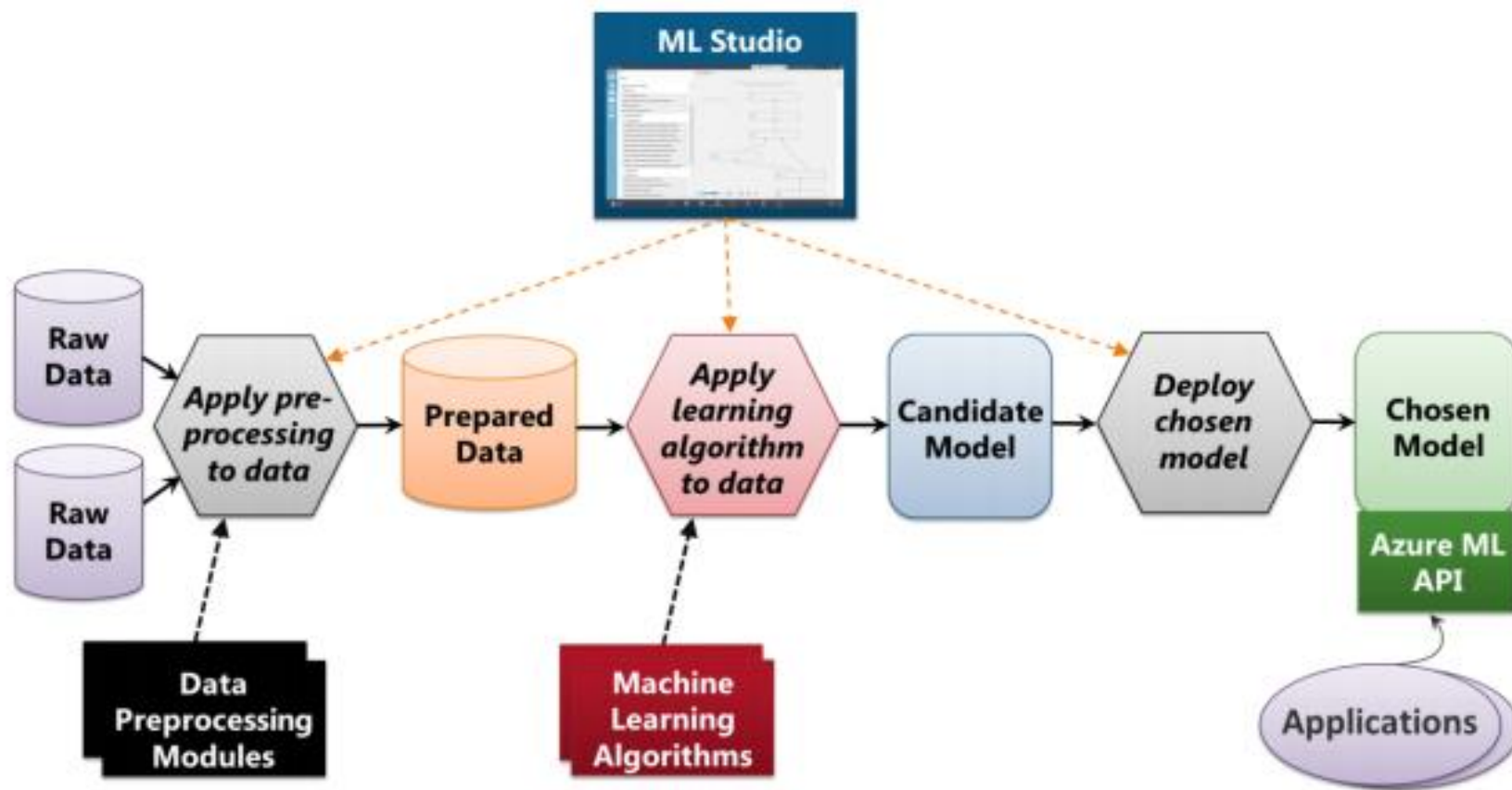
Number of trees construct...
100

Random number seed
0

☒ Allow unknown categ...

Quick Help

Creates a binary classifier using a boosted decision tree algorithm
(more help...)

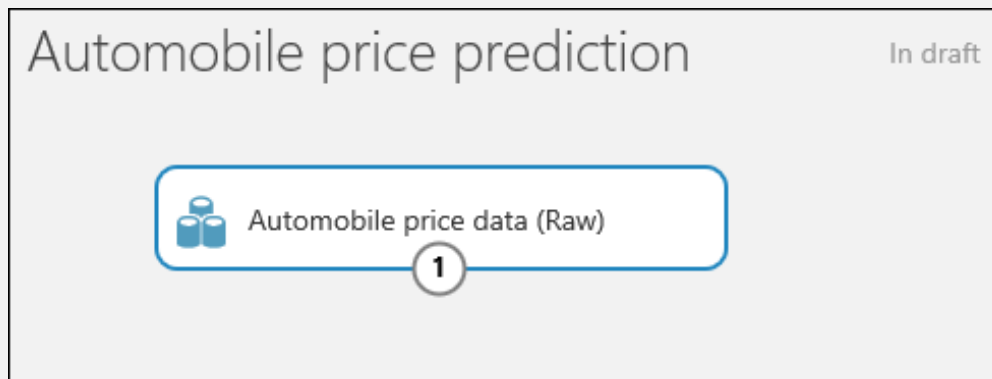
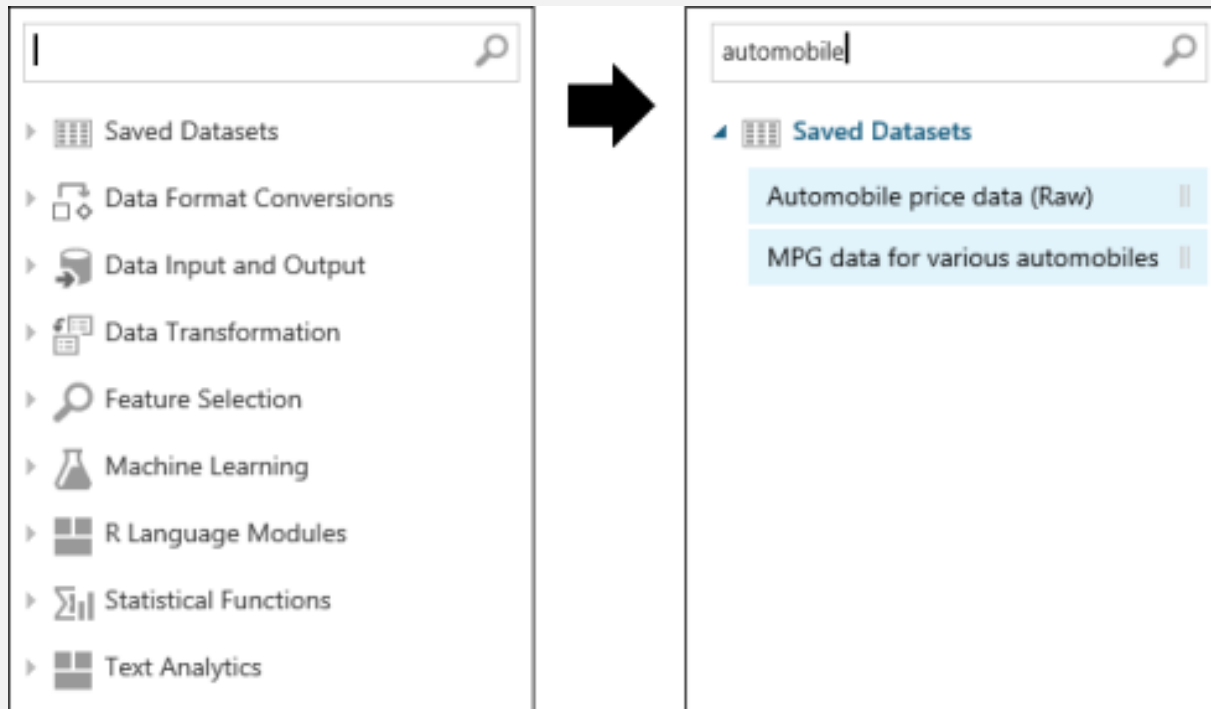


- ML Studio: Una interfaz gráfica utilizada para controlar el proceso desde el inicio hasta el fin. Con esta herramienta se pueden pre-procesar los datos utilizar algoritmos, probar el modelo resultante e incluso implementar el modelo en la nube de Azure
- Un set de módulos de pre-procesamiento de datos.
- Un set de algoritmos de aprendizaje automático.
- Un API para que las aplicaciones accedan al modelo generado una vez desplegado en la nube de Azure.

- Existen 5 pasos fundamentales para crear un experimento en Azure ML:
- **Creación del modelo**
 - Paso 1: Obtener los datos
 - Paso 2: Pre-procesar los datos
 - Paso 3: Definir las características
- **Entrenar el modelo**
 - Paso 4: Elegir y aplicar un algoritmo de aprendizaje
- **Probar el modelo**
 - Paso 5: Predecir, categorizar, clasificar o detectar fraudes a partir de nuevos datos.



- Paso 1. Obtener los datos

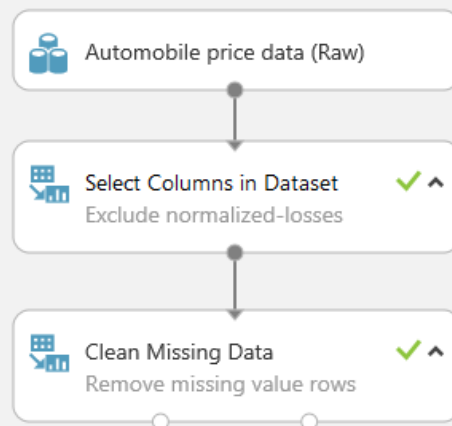


Automobile price prediction > Automobile price data (Raw) > dataset

rows: 205, columns: 26

	symboling	normalized-losses	make	fuel-type	aspiration	num-of-doors	body-style	engine	peak-rpm	city-mpg	highway-mpg	price
view as												
3			alfa-romero	gas	std	two	convertible	5000	21	27		13495
3			alfa-romero	gas	std	two	convertible	5000	21	27		16500
1			alfa-romero	gas	std	two	hatchback	54	5000	19	26	16500
2		164	audi	gas	std	four	sedan	102	5500	24	30	13950
2		164	audi	gas	std	four	sedan	115	5500	18	22	17450
2			audi	gas	std	two	sedan	110	5500	19	25	15250
1		158	audi	gas	std	four	sedan	110	5500	19	25	17710
1			audi	gas	std	four	sedan	110	5500	19	25	18920
1		158	audi	gas	turbo	four	sedan	140	5500	17	20	23875
0			audi	gas	turbo	two	sedan	160	5500	16	22	
2		192	bmw	gas	std	two	sedan	101	5800	23	29	16430
0		192	bmw	gas	std	four	sedan	101	5800	23	29	16925
0		188	bmw	gas	std	two	sedan	121	4250	21	28	20970
0		188	bmw	gas	std	four	sedan	121	4250	21	28	21105
1			bmw	gas	std	four	sedan	121	4250	20	25	24565

• Paso 2. Pre-procesar los datos →



Properties

Clean Missing Data

Columns to be cleaned

Selected columns:
All columns

Launch column selector

Minimum missing value ratio

0

Maximum missing value ratio

1

Cleaning mode

Remove entire row

- Paso 3. Definir las características

Select columns ×

BY NAME
WITH RULES

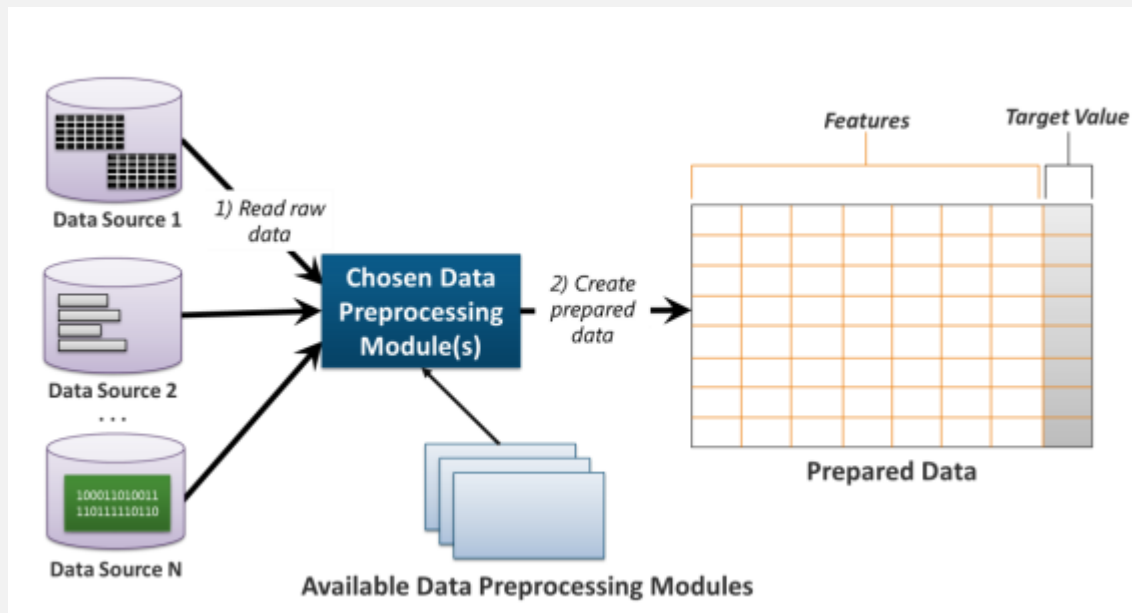
☐ Allow duplicates and preserve column order in selection

Begin With
ALL COLUMNS NO COLUMNS

Include column names

make × body-style × wheel-base × engine-size ×
horsepower × peak-rpm × highway-mpg ×
price ×

+ -



- Paso 4. Elegir y aplicar un algoritmo de aprendizaje

Select a single column

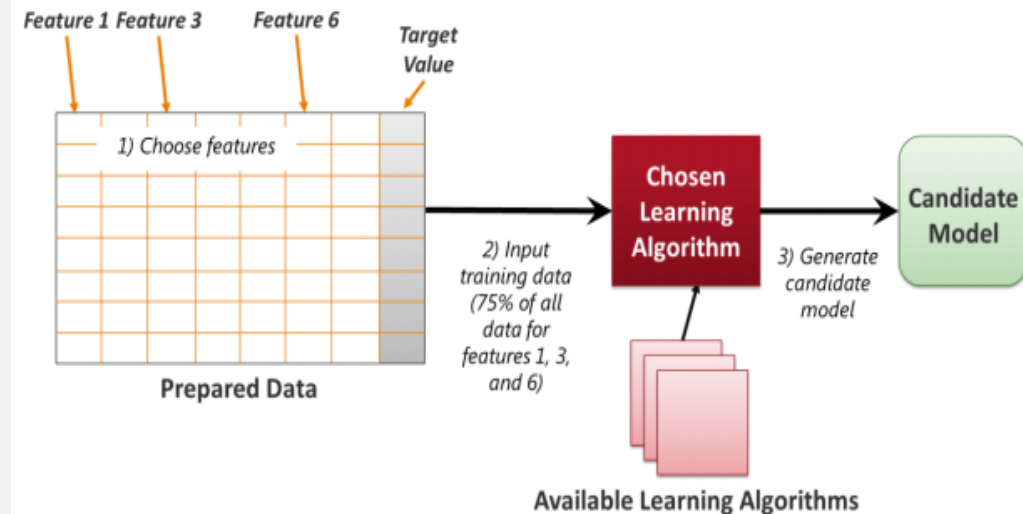
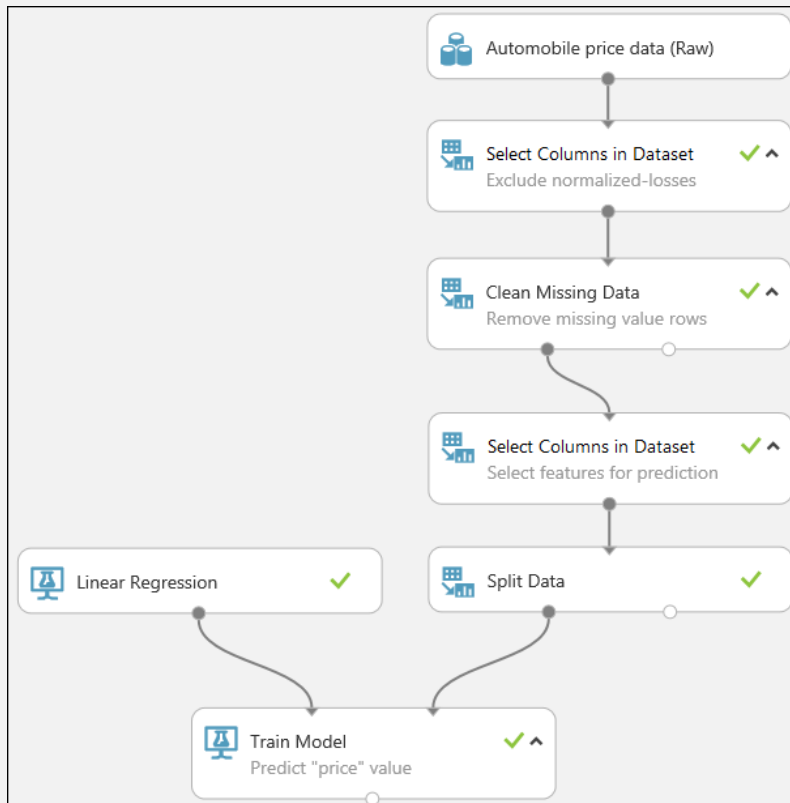
BY NAME

WITH RULES

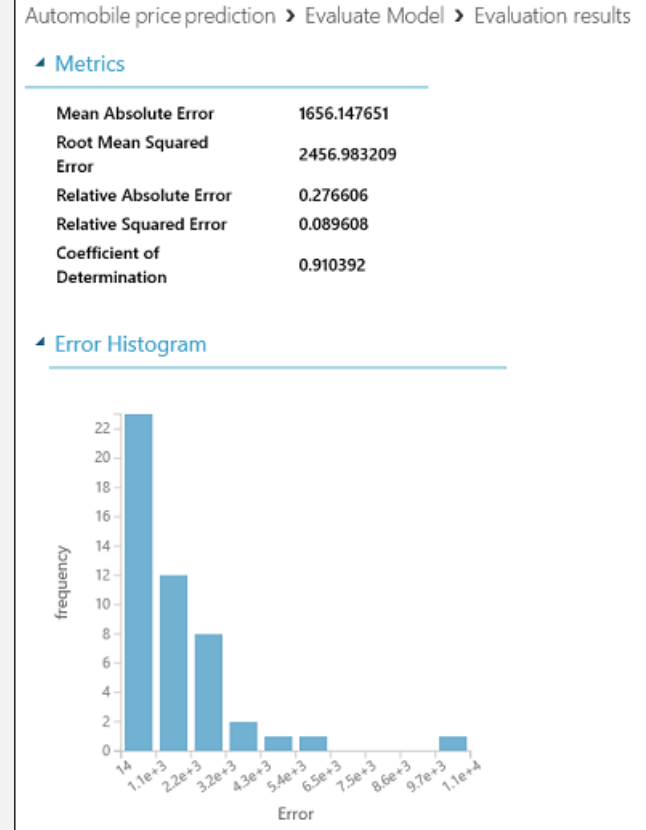
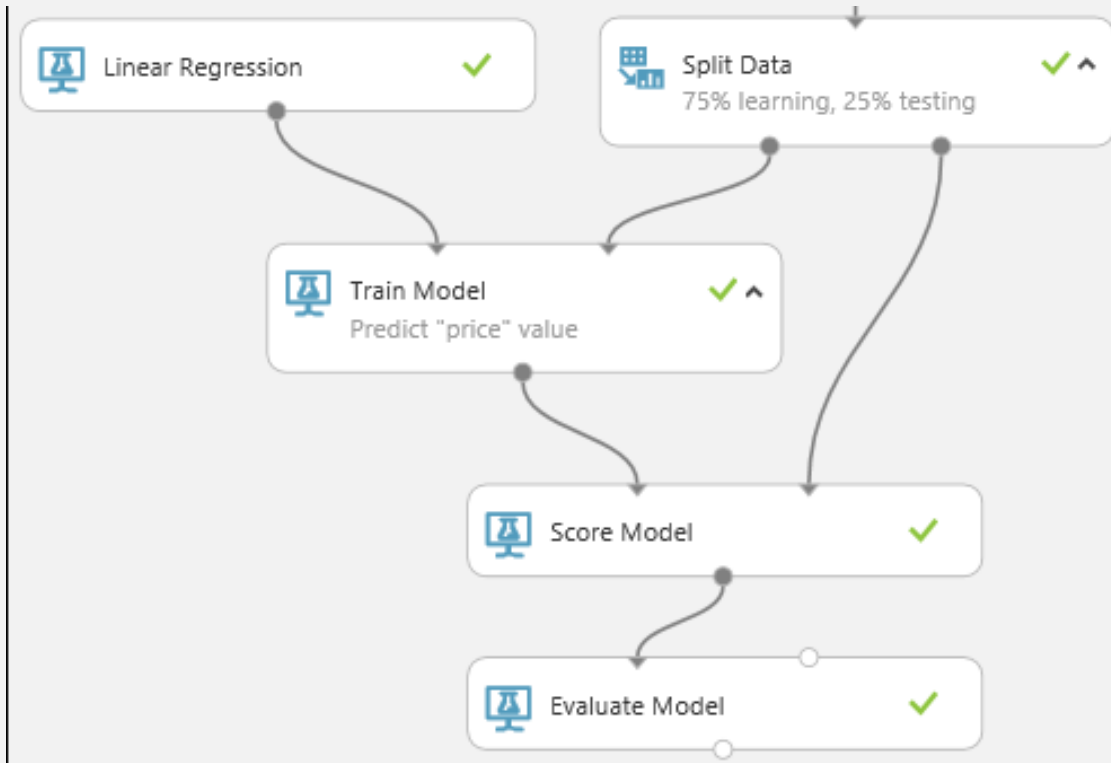
Include ▾

column names ▾

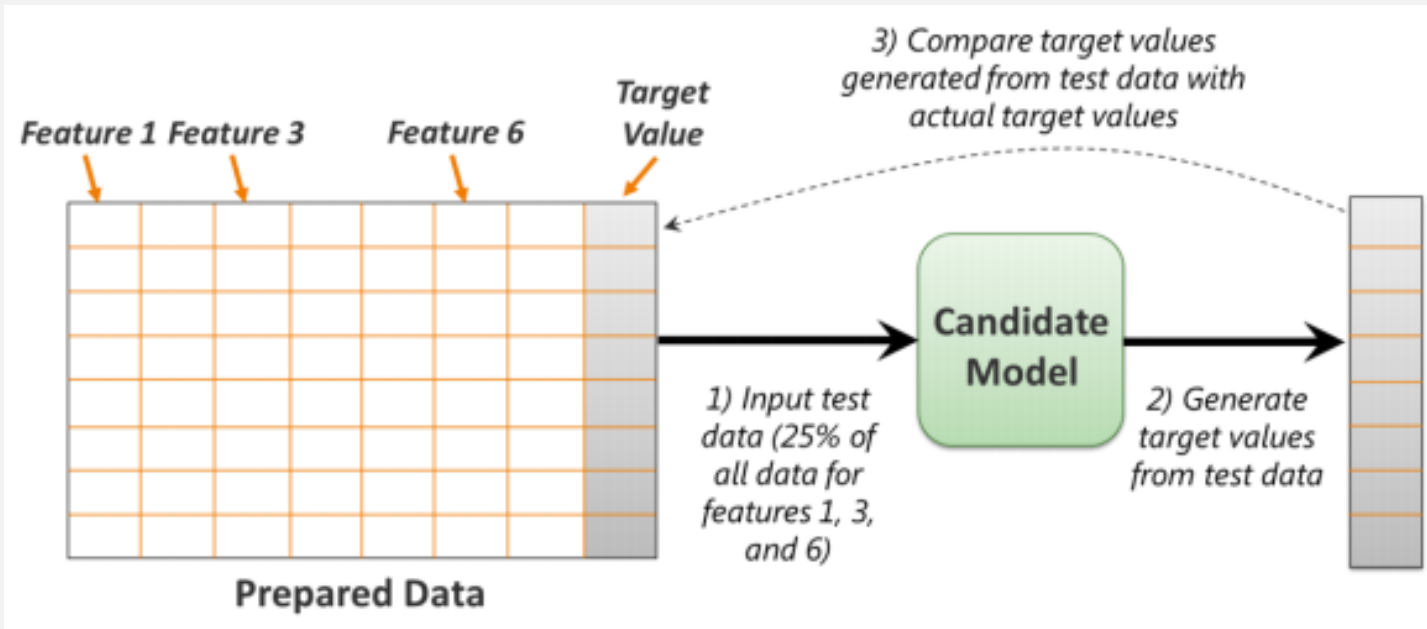
price ✕



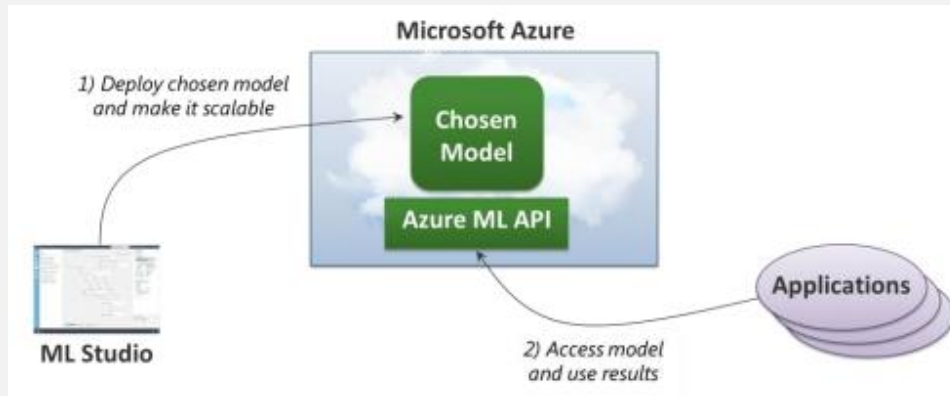
- Paso 5. Predecir, categorizar, clasificar o detectar fraudes a partir de nuevos datos.



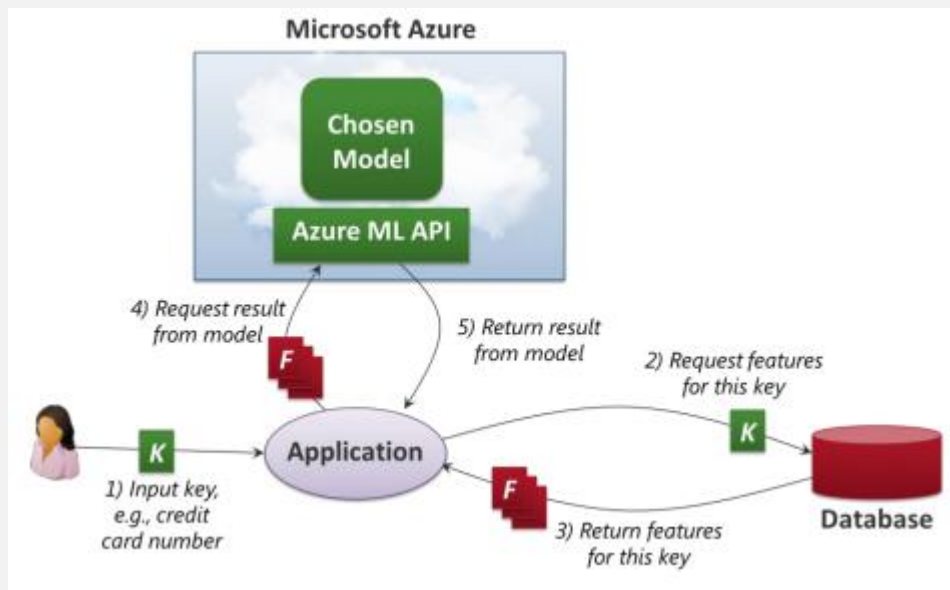
- Paso 5. Predecir, categorizar, clasificar o detectar fraudes a partir de nuevos datos.



- Implementación



- Acceso



Otras características:

- Python y R, dos de los lenguajes utilizados por los científicos de datos, son soportados por Azure ML.
- OpenCV, una librería open-source para procesamiento y reconocimiento de imágenes, también se puede integrar en Azure ML.
- Los experimentos creados pueden ser publicados, compartidos y monetizados en el Azure Marketplace.
- El servicio está orientado para ser utilizado tanto por programadores como por científicos de datos.



Detalles de precios

Precios de Estudio de aprendizaje automático

Estudio de aprendizaje automático se ofrece en dos niveles: Gratis y Estándar.

En la tabla siguiente se comparan las características, por nivel:

	GRATIS	ESTÁNDAR
Precio	Gratis	MXN\$176,82 por puesto al mes MXN\$17,70 por hora de experimentación de Estudio de aprendizaje automático
Suscripción a Azure	No requerido	Requerido
Número máximo de módulos por experimento	100	Ilimitado
Duración máxima de experimentación	1 hora por experimento	Hasta 7 días por experimento con un máximo de 24 horas por módulo
Espacio máximo de almacenamiento	10 GB	Ilimitado (BYO)
Leer datos de SQL local <small>Vista previa</small>	No	Sí
Ejecución/rendimiento	Un nodo	Varios nodos
API web de producción	No	Sí
Contrato de nivel de servicio	No	Sí

Los gastos por hora se aplican únicamente al uso activo del servicio. Si existen varios medidores, estos se aplican simultánea



Precios de API web de producción

Aprendizaje automático de Azure permite implementar soluciones de análisis predictivo como servicios web.

Los servicios web implementados ([nueva versión](#)) están sujetos a los siguientes planes

	DESARROLLO Y PRUEBAS*	ESTÁNDAR S1	ESTÁNDAR S2	ESTÁNDAR S3
Precio de nivel al mes	MXN\$0	MXN\$1.770	MXN\$17.700	MXN\$177.000
Características				
Transacciones incluidas (al mes)	1.000	100.000	2.000.000	50.000.000
Horas de proceso incluidas (al mes)	2	25	500	12.500
Número total de servicios web ¹	2	10	100	500
Tarifas de uso por encima del límite	N/D	MXN\$8,85 por 1000 transacciones MXN\$35,40 por hora de proceso de API	MXN\$4,425 por 1000 transacciones MXN\$26,55 por hora de proceso de API	MXN\$1,77 por 1000 transacciones MXN\$17,70 por hora de proceso de API



4. Demo:

<https://github.com/icebeam7/azureml-xamarin>

- Creando un experimento de predicción en Azure Machine Learning
- Desarrollando un web service
- Consumiendo el web service en una app multiplataforma construida con Xamarin.



Conclusiones

- Machine Learning es un área de gran importancia y utilidad en un mundo donde a cada instante se generan millones de datos. Es la era del Big Data.
- El análisis de Big Data permite crear soluciones a problemas y construir mejores aplicaciones.
- Azure Machine Learning afronta el reto de simplificar el proceso de aprendizaje automático.

- Los científicos de datos siempre serán importantes en el análisis de datos.
- Este servicio de la nube puede ayudar a que personas menos especializadas en las ciencias de los datos jueguen un rol importante en la aplicación de aprendizaje automático en problemas de la vida real.



¡Muchas gracias por su atención! 😊



SG VIRTUAL CONFERENCE

11 a. edición



LUIS BELTRÁN



@darkicebeam



<http://icebeamwp.blogspot.mx>



<fb.me/darkicebeam>



darkicebeam



<linkedin.com/in/darkicebeam>



luis.beltran@itcelaya.edu.mx

#SGVirtual