

Tecnologías de Inteligencia Artificial de Microsoft Azure

Taller Práctico

Septiembre 2018

Agenda

- Plataforma Azure (45min)
 - Bots
 - Machine Learning
 - Data Science Virtual Machine
- State of the art algoritmos (45min)
 - Sistemas de recomendación
- Laboratorio Azure ML (90 min)
 - Ingesta de datos
 - Configuración de algoritmos
 - Validación
 - Publicación web service
- Bots (90 min)
 - Teoría
 - Práctica
 - Ejemplos



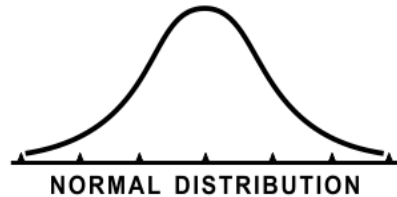
Aprendizaje de máquina

Aprendizaje de máquina (Machine Learning)

- Es una rama de la **inteligencia artificial** cuyo objetivo es desarrollar técnicas que permitan a las computadoras aprender algoritmos capaces de generalizar comportamientos a partir de información no estructurada suministrada en forma de ejemplos.
- El aprendizaje se cataloga como **supervisado, por refuerzo o no supervisado** dependiendo de si el algoritmo debe contar o no con información específica de datos satisfactorios para el objetivo del aprendizaje

Aprendizaje de máquina

- Es la ciencia que estudia como aprender a partir de datos.
- ¿Qué eso no es lo que hace la estadística?



Aprendizaje de máquina

- Empata en muchos aspectos con la estadística, pero el enfoque es distinto:
 - La estadística busca modelos simples que expliquen el porqué de los fenómenos
 - El aprendizaje de máquina busca que las predicciones sean lo más certeras posible
 - El aprendizaje de máquina se enfoca más en el aspecto computacional dada la complejidad de los algoritmos

Estadística tradicional vs. Aprendizaje de máquina

Estadística tradicional

Aprendizaje de máquina

Más interpretables

Mejores predicciones

RISE OF THE MACHINES
Larry Wasserman



Estadística Tradicional versus Aprendizaje de Máquina

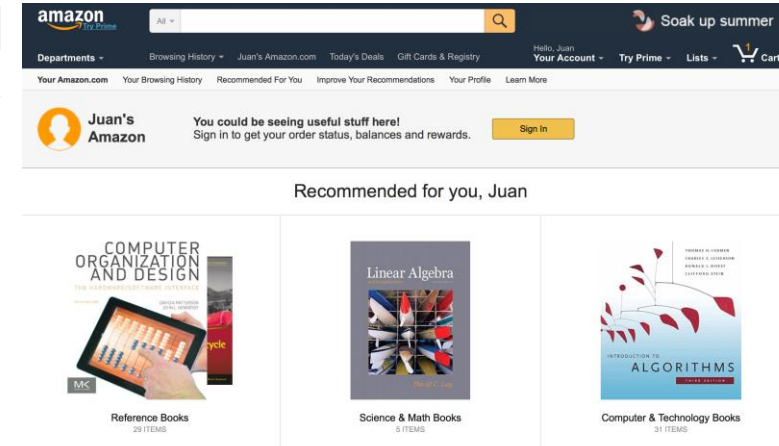
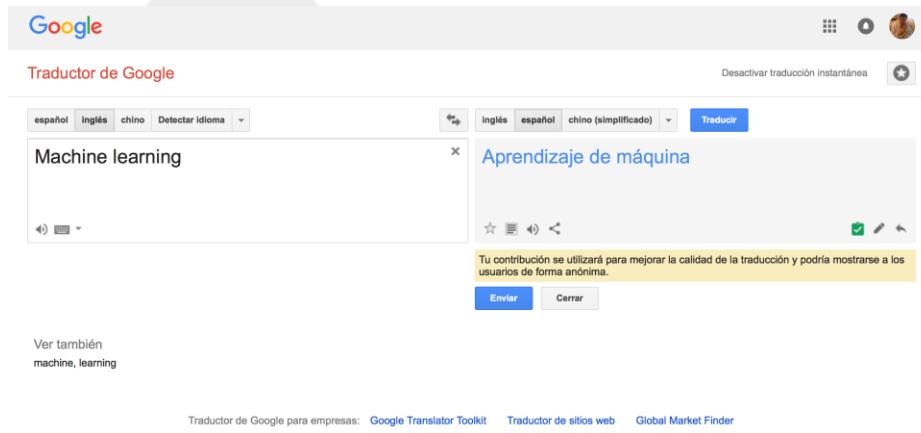
Estadística Tradicional

- Hacen hincapié en la **inferencia** de superpoblación
- Modelos más simples se prefieren a los complejos (parsimonia), aunque los modelos más complejos lo representen mejor
- Énfasis en la capacidad de **interpretar los parámetros**
- Modelado estadístico y los supuestos de muestreo conectan los datos a una población de interés
- Preocupación por los supuestos

Aprendizaje de Máquina

- Hace hincapié en las **predicciones**
- Preocupación por el sobreentrenamiento pero no por la complejidad del modelo per sé
- Énfasis en el **rendimiento**
- La generalización se obtiene a través de la aplicación sobre nuevos conjuntos de datos. Por lo general no hay un modelo de superpoblación específica
- Preocupación por el rendimiento

Ejemplos de programas que utilizan aprendizaje de máquina



El presente y futuro cercano



What if technology could adapt to human emotion?

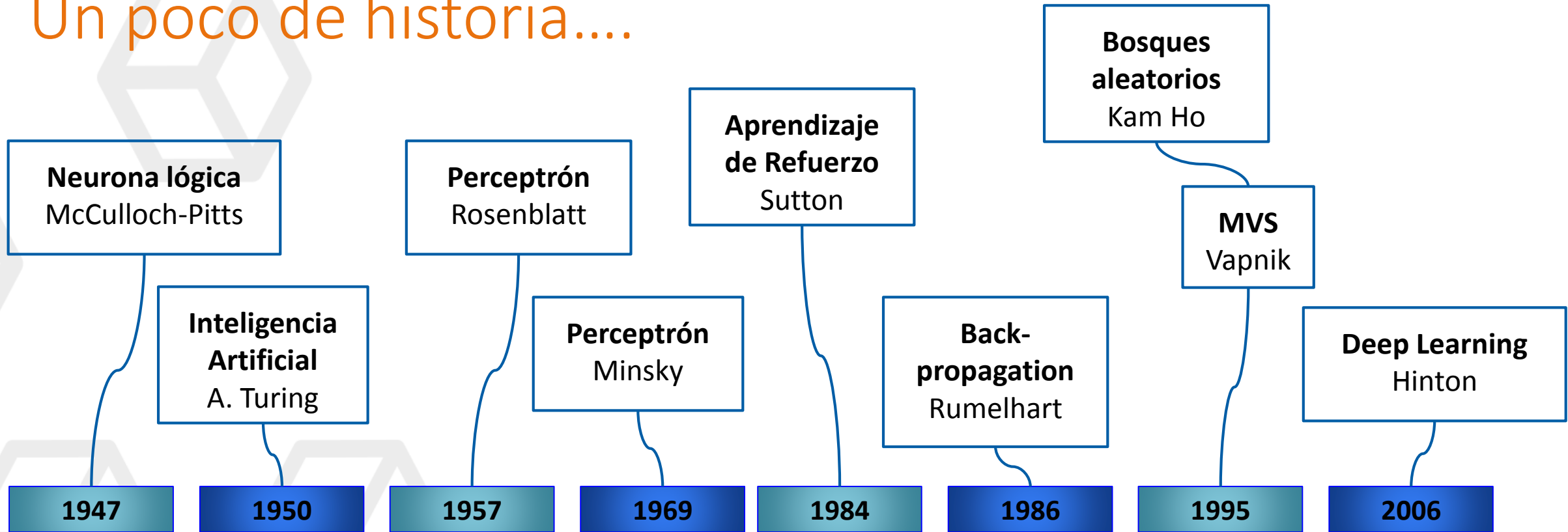
Emotion AI

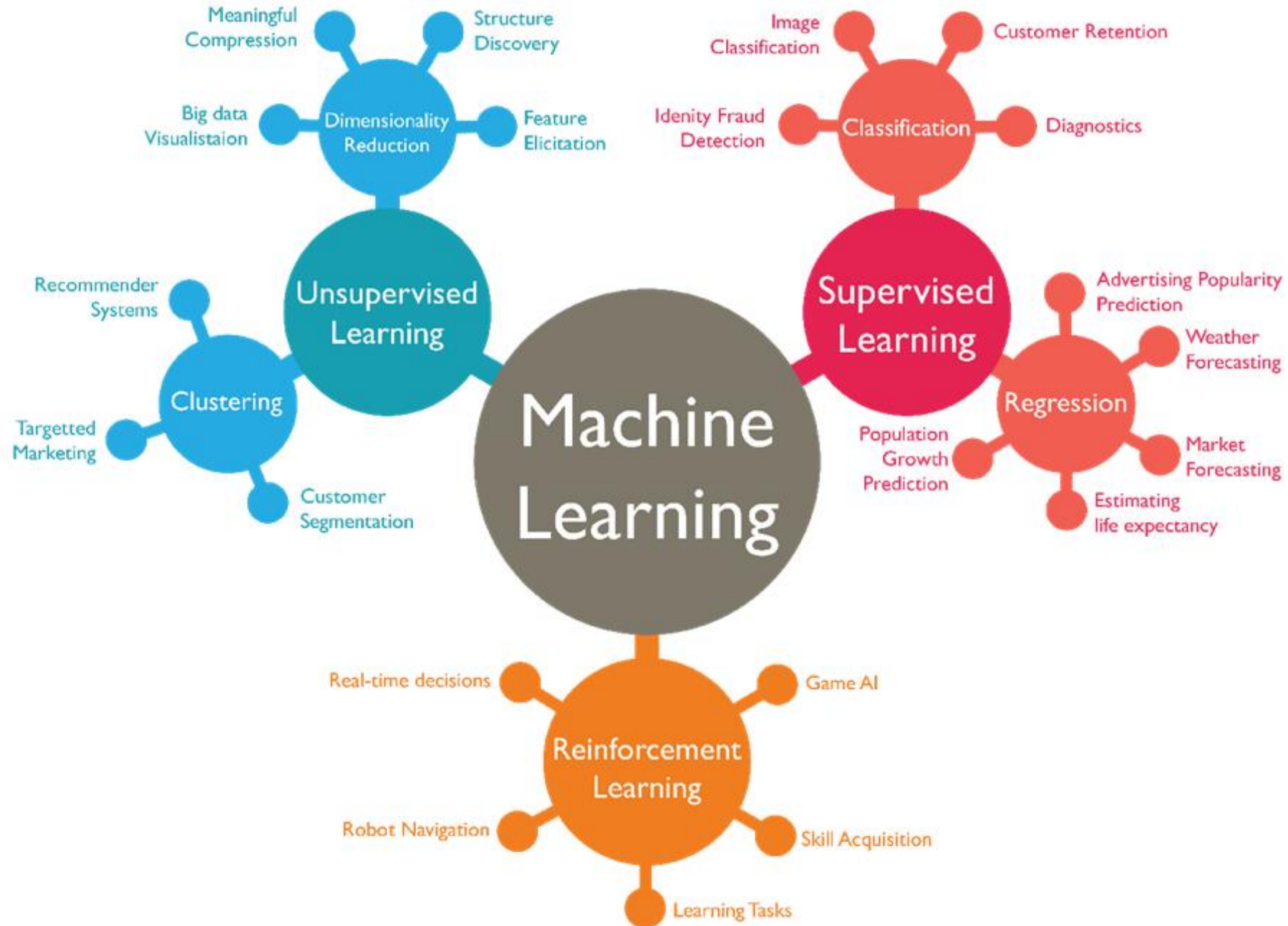
What would you build?

TRY OUR AFFDEX DEMO >

TRY OUR EMOTION RECOGNITION SDK NOW >

Un poco de historia....





Qué significa “aprender” para una máquina?

Extraer patrones reproducibles ocultos de un conjunto de datos

Supervisado:

aprender a partir de datos **etiquetados**
a lo largo del tiempo

ejemplo: spam



No Supervisado:

aprender de datos que **no tienen etiquetas**

ejemplo: clusters

RESEARCH
PLAN

HOMEWORK



BRING
SNACKS

BEST
TRAILS

HIKING
BOOT!

CLICK
TO WIN

MEETING
TODAY

CLASS
CANCELED

Aprendizaje supervisado

- Infiere una función a partir de datos de entrenamiento etiquetados.
- Cada dato de entrenamiento es un par que consta de un objeto de entrada (típicamente un vector) y un valor de salida deseado (también llamada la señal de supervisión).
- Dos tipos de salida
 - Numérico continuo: Regresión
 - Valores discretos (clases): Clasificación

Aprendizaje supervisado

- Se utilizan ampliamente en analítica predictiva
- Algoritmos
 - Regresiones lineales
 - Regresiones logísticas
 - Redes neuronales
 - Máquinas de vectores de soporte
 - Árboles de decisión
 - Bosques Aleatorios
 - Boosting
 - K-vecinos cercanos
 - Filtrado colaborativo

Aprendizaje no supervisado

- Infiere una función que describe la estructura de datos no etiquetados
- No hay señal de error ni de recompensa para evaluar una solución potencial
- Busca resumir y explicar la principales características de los datos

Aprendizaje no supervisado

- Encargado de detectar patrones o asociaciones, no fácilmente observables dentro de los datos.
- Se utiliza principalmente en minería de datos.
- Algoritmos
 - Agrupamiento (clustering)
 - Componentes principales
 - Modelo de mezclas (mixture models)
 - Apriori

Sistemas de recomendación

Películas, libros, comidas, trabajos, parejas... Buenos consejos son necesarios en todos los contextos. Los sistemas de recomendación se encargan de brindar sugerencias de usuarios con intereses similares a los tuyos.



- Cada minuto se agregan 300 horas de video en Youtube. Los sistemas de recomendación son útiles para discriminar y elegir la opción más ad hoc a nuestras preferencias.
- Los sistemas de recomendación son el motivo principal del éxito de Netflix como caso de negocio.
- Puedes no saber exactamente lo que estás buscando.
- El resultado depende del contexto y pretende ser personalizado.

Sistemas de recomendación



- Son útiles para entender mejor la demanda y planificar adecuadamente la oferta a fin de optimizar el proceso de producción.
- Muchos sectores económicos (Transporte, retail, entretenimiento, finanzas) dependen de las preferencias de las personas involucradas en las transacciones.



El caso Netflix

The Netflix Prize



- Inicio del concurso: Octubre 2006.
- Training set: 100 millones de registros incluyendo: usuario, película y calificación
- En promedio, cada usuario calificó 200 películas y cada película contaba en promedio con 5,000 calificaciones.
- Netflix ofreció \$1,000,000 USD al equipo que logrará mejorar en al menos 10% el error cuadrático medio (MSE) del algoritmo de clasificación

NETFLIX

Netflix Prize

COMPLETED

Home Rules Leaderboard Update

Leaderboard

Showing Test Score. [Click here to show quiz score](#)

Display top 20 leaders.

Rank	Team Name	Best Test Score	% Improvement	Best Submit Time
Grand Prize - RMSE = 0.8567 - Winning Team: BellKor's Pragmatic Chaos				
1	BellKor's Pragmatic Chaos	0.8567	10.06	2009-07-26 18:18:28
2	The Ensemble	0.8567	10.06	2009-07-26 18:38:22
3	Grand Prize Team	0.8582	9.90	2009-07-10 21:24:40
4	Opera Solutions and Vandelay United	0.8588	9.84	2009-07-10 01:12:31
5	Vandelay Industries !	0.8591	9.81	2009-07-10 00:32:20
6	PragmaticTheory	0.8594	9.77	2009-06-24 12:06:56
7	BellKor in BigChaos	0.8601	9.70	2009-05-13 08:14:09
8	Dace	0.8612	9.59	2009-07-24 17:18:43
9	Feeds2	0.8622	9.48	2009-07-12 13:11:51
10	BigChaos	0.8623	9.47	2009-04-07 12:33:59
11	Opera Solutions	0.8623	9.47	2009-07-24 00:34:07
12	BellKor	0.8624	9.46	2009-07-26 17:19:11

La competencia convocó a miles de equipos y contribuyó al desarrollo de investigación en el área de aprendizaje de máquina. Tres años después, el equipo BellKor's Pragmatic Chaos finalmente logró mejorar el algoritmo en un 10.06%.

Recommendation System

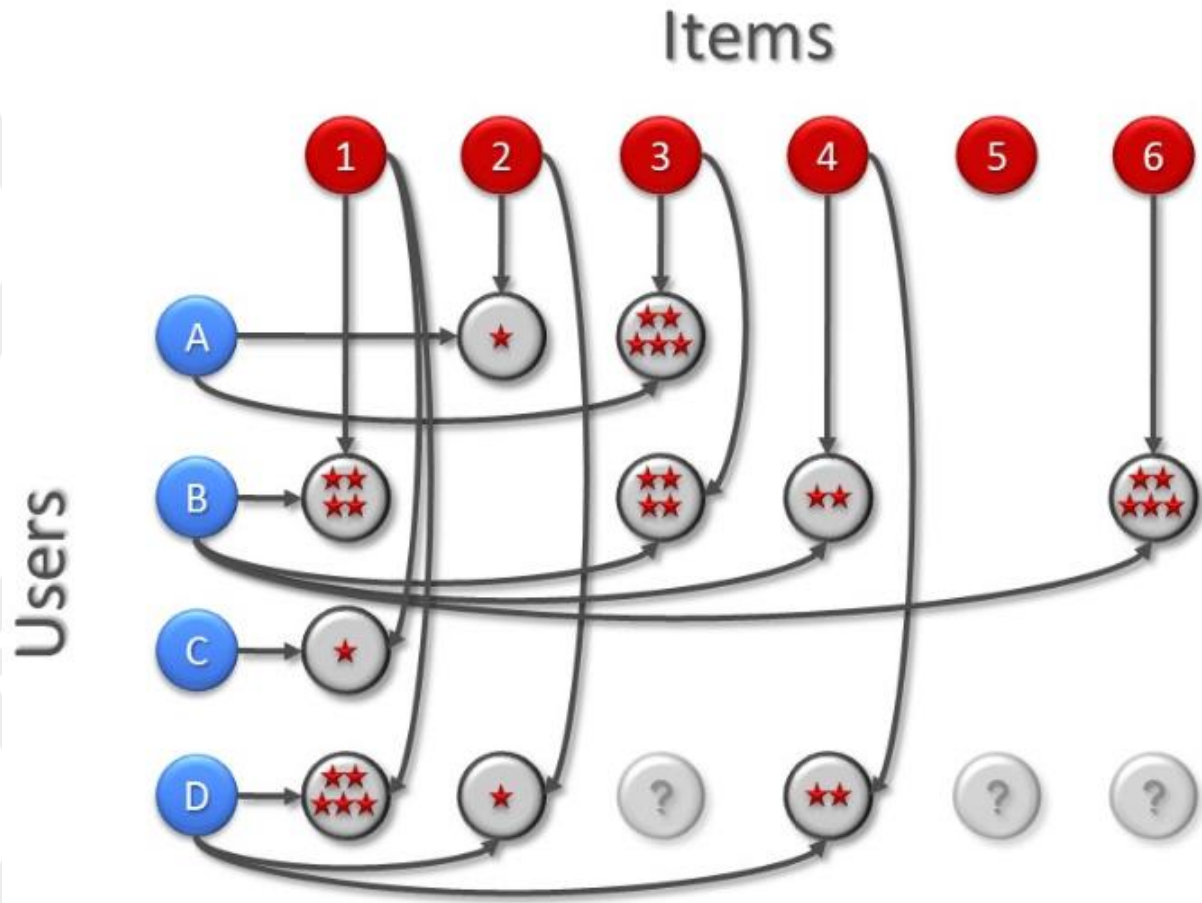


Collaborative filtering

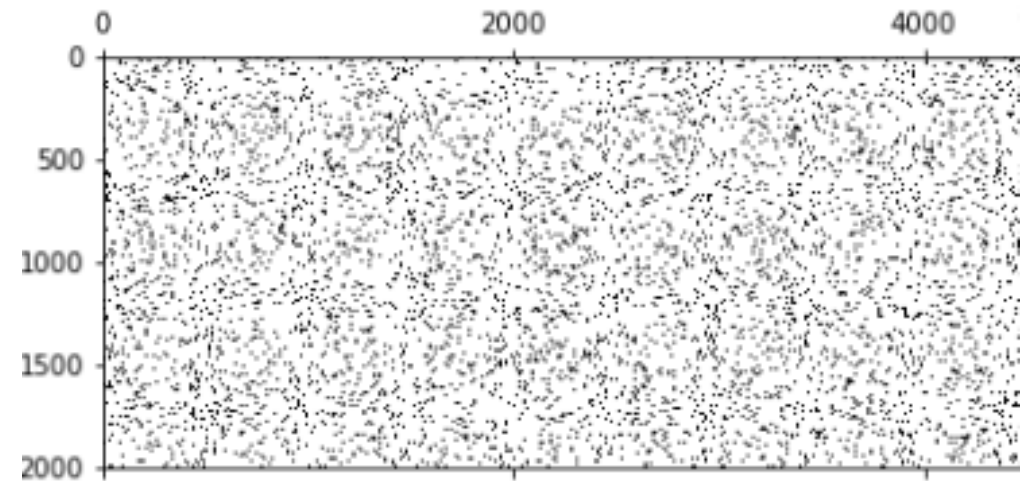


- El objetivo es encontrar otras personas con gustos, opiniones o intereses similares y aplicar sus preferencias para predecir situaciones que podrían gustarme o disgustarme.
- Se filtran o predicen las preferencias usando el conocimiento de otros usuarios

¿Cómo predecir ratings?



- Es necesario construir una matriz con los usuarios como renglones, los productos como columnas y las calificaciones en las entradas.
- Hay un gran porcentaje de datos faltantes porque típicamente los usuarios sólo califican un pequeño subconjunto de productos.



¿Cómo predecir ratings?

	Juan Paco Pedro Mar			
La sociedad de los poetas muertos	5	5	0	0
Cinema Paradiso	5	?	?	0
La lista de Schindler	?	4	0	?
Star Wars	0	0	5	4
Star Trek	0	0	5	?

¿Cómo predecir ratings?

	Juan	Paco	Pedro	Mar	Drama	Acción
La sociedad de los poetas muertos	5	5	0	0	90%	0%
Cinema Paradiso	5	?	?	0	100%	1%
La lista de Schindler	?	4	0	?	99%	0%
Star Wars	0	0	5	4	10%	100%
Star Trek	0	0	5	?	0%	90%

¿Cómo predecir ratings?

	Juan	Paco	Pedro	Mar	Drama	Acción
La sociedad de los poetas muertos	5	5	0	0	90%	0%
Cinema Paradiso	5	?	?	0	100%	1%
La lista de Schindler	?	4	0	?	99%	0%
Star Wars	0	0	5	4	10%	100%
Star Trek	0	0	5	?	0%	90%

¿Cómo predecir ratings?

	Juan	Paco	Pedro	Mar	Drama	Acción
La sociedad de los poetas muertos	5	5	0	0	90%	0%
Cinema Paradiso	5	?	?	0	100%	1%
La lista de Schindler	?	4	0	?	99%	0%
Star Wars	0	0	5	4	10%	100%
Star Trek	0	0	5	?	0%	90%

¿Cómo predecir ratings?

	Juan	Paco	Pedro	Mar	Drama	Acción
La sociedad de los poetas muertos	5	5	0	0	90%	0%
Cinema Paradiso	5	?	?	0	100%	1%
La lista de Schindler	?	4	0	?	99%	0%
Star Wars	0	0	5	4	10%	100%
Star Trek	0	0	5	?	0%	90%

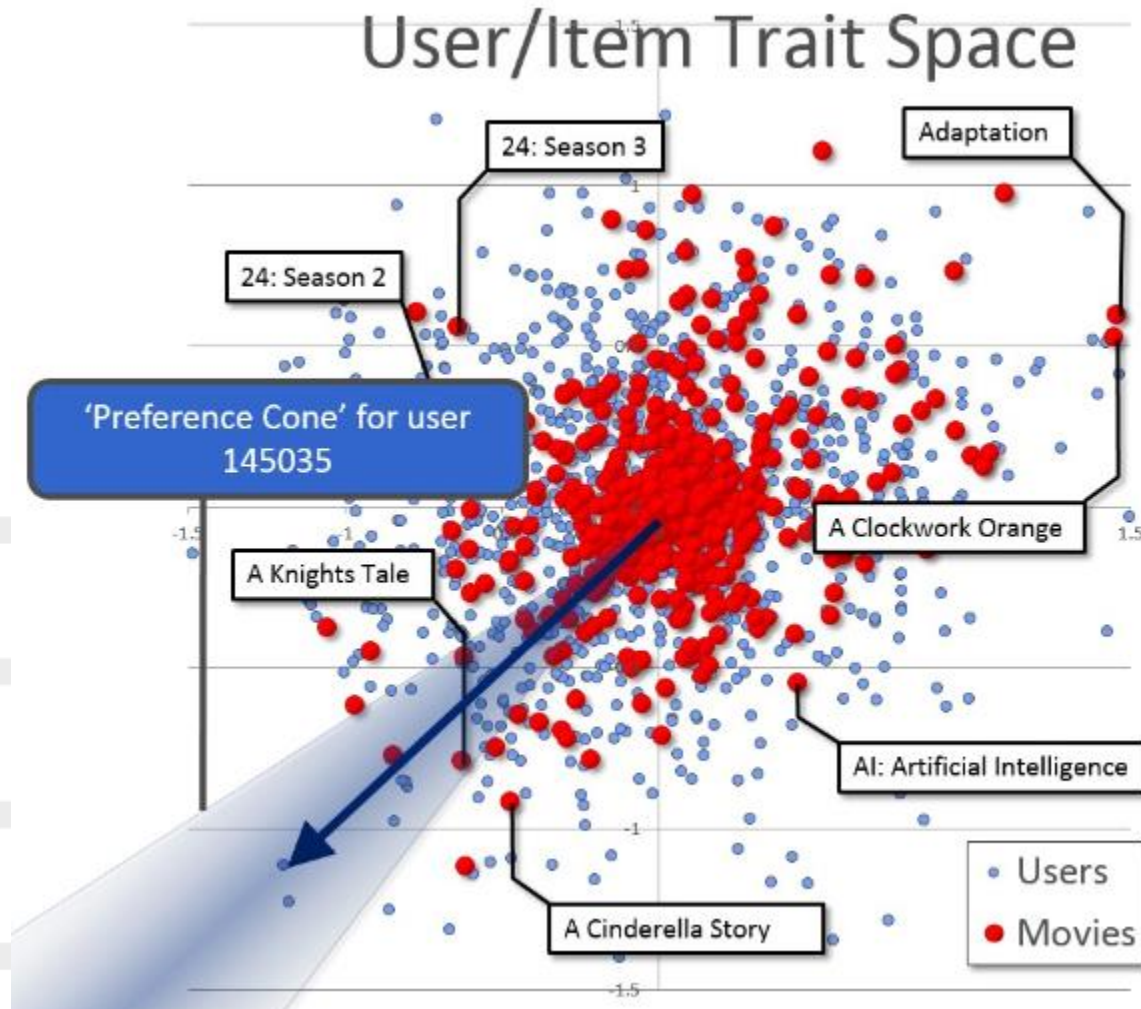
¿Cómo predecir agrupaciones (traits)?

	Juan	Paco	Pedro	Mar	Drama	Acción
La sociedad de los poetas muertos	5	5	0	0	?	?
Cinema Paradiso	5	?	?	0	?	?
La lista de Schindler	?	4	0	?	?	?
Star Wars	0	0	5	4	?	?
Star Trek	0	0	5	?	?	?

¿Cómo predecir agrupaciones?

	Juan Paco Pedro Mar				Drama	Acción
La sociedad de los poetas muertos	5	5	0	0	90%	0%
Cinema Paradiso	5	?	?	0	100%	1%
La lista de Schindler	?	4	0	?	99%	0%
Star Wars	0	0	5	4	10%	100%
Star Trek	0	0	5	?	0%	90%

Bayesian Recommendation System Azure ML



- El modelo representa la información de usuarios y películas en un hiperespacio (latent trait space).
- Un usuario (punto azul) califica una película (punto rojo) positivamente si su vector está alineado con el vector de la película. La preferencia es negativa si están en direcciones opuestas.
- Usuarios y productos similares estarán cerca en el espacio.
- Algunas veces, es posible asignar interpretación a los ejes ("traits"). Por ejemplo, la dirección norte-sur podría ser "adolescentes vs infantiles" y el eje x "mainstream vs culto".



<https://github.com/bdatae/hack-ai-ml-2018>

Lab

The background is a solid light purple color. It features a subtle pattern of white hexagons and molecular-like structures. Some hexagons are solid, while others are outlines. Molecular structures consist of small circles connected by lines, resembling a network or a chemical structure.

Tecnologías de Inteligencia Artificial de Microsoft Azure

Gracias!

Septiembre 2018