Aplicaciones de la Ciencia de Datos

- Creo que una de las nuevas aplicaciones útiles

de la ciencia de datos esta en el campo de la medicina,

al igual que en la administración de medicamentos o el tratamiento del cáncer.

- Creo que es un caso muy interesante

como ahora las empresas pueden usar toda la información

que recopilan de sus clientes

para desarrollar nuevos productos

que responden a las necesidades

de los clientes.

- Una de las nuevas aplicaciones útiles de la ciencia de los datos

fue la alta tendencia de las noticias referentes a Pokémon Go,

ellos usaron Ingress.

Utilizaron los datos de la app Ingress,

la última app de la misma empresa,

y eligieron las localizaciones

para los pokémons y los gimnasios

según los datos de la última app,

así que aprendieron de sus errores.

- El buscador de Google es una aplicación de la ciencia de los datos,

la búsqueda de Google, cuando queremos buscar algo.

Creo que todo se debe a la ciencia de datos.

Lo que llegó a ser Google ahora, todo es consecuencia de la ciencia de datos.

- La realidad aumentada es mi favorita,

una nueva implementación de la ciencia de los datos.

Creo que no puedes mirar una nueva tecnología

y no ver ciencia de datos ahí,

pero la realidad aumentada

es la que más me emociona.

La capacidad de caminar y ver cosas en las paredes

o a nuestro alrededor que no estan ahí realmente.

Pokémons son solo el comienzo.

- Entonces, lo que ha pasado es que ahora las herramientas estan disponibles

y los conjuntos de datos estan disponibles.

Las personas estan usándolas sin tener mucho cuidado

y creo que uno de los casos más extraños

que se informo en los periódicos es sobre la historia

de un padre que fue a una tienda Target en los Estados Unidos

quejándose del hecho

que la tienda Target estaba enviando correos a su hija adolescente

sobre pañales y leche, fórmula para bebés.

El estaba muy molesto por esto,

él dijo: "¿Por qué te gustaría

que mi hija adolescente tenga un bebé?"

y él obviamente estaba perturbado

por estos correos o la campaña publicitaria,

ellos obviamente se disculparon.

Pero luego el padre regresó dos semanas después

y se disculpó con ellos

diciendo que él no sabía que su hija estaba embarazada.

Ahora la pregunta es, ¿cómo sabía Target sobre esto?

antes de que el padre supiera.

Y lo que pasó es que ellos pudieron observar

el comportamiento de compra de los individuos

Así que si estas comprando algún tipo de suplementos o vitaminas

entonces sabes que este es el primer trimestre del embarazo.

Por ende, ellos saben que productos enviarte

asumiendo que las personas

que compraron esos productos estaban embarazadas.

Ahora esta es una gran historia acerca de la ciencia de datos

y como la ciencia de datos puede prever y predecir

estos comportamientos del consumidor

incluso antes de que la familia se enterara.

Y lo encuentro inquietante, extraño y raro

por muchas razones.

En primer lugar, para cada predicción correcta

tienes cientos de predicciones incorrectas

a los que llamamos falsos positivos

y ningún científico de datos en realidad anuncia

sus falsos positivos.

Nosotros solamente anunciamos y fomentamos los que salieron bien,

pero cuando nos equivocamos cientos de veces no lo decimos.

Lo segundo es el abuso de datos,

eso básicamente no te entrega mucha información.

Tu puedes encontrar una correlación,

pero alguien podría estar encontrando lo mismo

para alguien más.

Así que las probabilidades de equivocarse

y las probabilidades de obtener falsos positivos es mucho mayor.

Entonces, me parece extraño y creo que da una falsa sensación

de nuestra capacidad de predecir el futuro.

La realidad sobre la ciencia de datos

y lo más importante

para el científico de datos al inicio es saber

que todas las predicciones son erróneas,

son útiles, pero son erróneas.

Y así uno no debe poner su fe

en el hecho de que ahora podemos realizar análisis predictivo

y que podemos resolver todos los problemas.

Creo que un buen ejemplo es la búsqueda de Google.

Google publicó un artículo diciendo que

pueden predecir las epidemias de gripe

antes del Centro de Control de Enfermedades

y lo que hicieron fue observar

en las personas que buscaban en Google síntomas de la gripe.

Así Google vio los síntomas de gripe buscados

antes que nadie y ellos fueron capaces de predecirlo.

La cosa es que estas búsquedas son buenas

y se correlacionan con algunos resultados

pero no necesariamente todo el tiempo.

Así que en ese momento, cuando Google anunció

fue algo grande y a que todos les gustó mucho

y bueno, eso fue una nueva era del análisis predictivo.

Solo que unos años después se dieron cuenta

que Google empezó a predecir falsos positivos,

predecían cosas que realmente no estaban ahí

o las predicciones no eran tan precisas

por una variedad de razones.

Ellos cambiaron probablemente sus algoritmos

y los conjuntos de datos que no estaban realmente correlacionados

con los resultados.

Entonces, ¿cuál es la lección que se aprende aquí?

Uno tiene que evitar un concepto que llamamos la arrogancia de datos,

que tu no debes creer mucho en tus resultados

porque puede desviarte.

La ciencia de datos tiene tremendo potencial para generar cambios

en partes del mundo, en partes de nuestra sociedad

que han sido marginados por años.

Se ven grandes ejemplos de la ciencia de datos

especialmente en los países en desarrollo

donde estan dirigidos a los esfuerzos de rescate.

Se dirigen a la comida

y otras ayudas particulares,

a lugares que no han sido elegidos anteriormente.

y la razón que esta sucediendo ahora

es la mayor disponibilidad de datos, modelos y análisis

para poder señalar donde están las mayores necesidades.

La habilidad de diseñar y realizar experimentos

para ver si uno dará micro créditos,

pequeños prestamos a hogares muy pobres,

en partes del mundo en desarrollo.

Para ver como afectan

la capacidad individual del hogar para salir de la pobreza,

y también la habilidad de la comunidad

para mejorar colectivamente su bienestar económico

por medio de muy pequeñas infusiones de efectivo o crédito.

Asi que estos experimentos suceden en todo el mundo,

estan permitiendo que sea un resultado directo

de nuestra capacidad de analizar datos

y poder diseñar experimentos

y luego de desplegar esfuerzos gigantescos

en proporcionar alivio, proporcionar crédito,

proporcionar una oportunidad

a aquellos que fueron privados anteriormente,

una oportunidad para unirte al resto del mundo

en la prosperidad,felicidad y salud.