

El futuro de la informática

Mesa Redonda Novática, 10/05/2016

Introducción

- Llevo trabajando en Big Data aproximadamente un año
- Unido al hecho de que se me da mejor predecir el pasado que el futuro
- No me planteo predecir que va a ser de Big Data en unos años
- Máxime cuando se trata de un campo tecnológicamente convulso
- Por ejemplo una tecnología como Storm, que tiene pocos años, está siendo desplazada ya por Flink
- En lo que sí soy un especialista, que no experto, es en tecnologías de Procesamiento del Lenguaje Natural y Analítica de Texto
- Intentaré hacer una reflexiones que conecten ambos temas

Big Data

- Big Data es una de las palabras clave de moda, junto con Cloud y otras
- En general hace referencia a un conjunto de tecnología para manejar cantidades masivas de datos - del orden de teras diarios
 - Almacenar, organizar, analizar
- Estas tecnologías dan soporte a las grandes aplicaciones de Internet
 - Google, Yahoo, Microsoft, Facebook, Twitter, Youtube, Amazon, Ebay, Salesforce
- Así que son una realidad y no solo una moda
- Centran el foco sobre los datos

Big Data y algoritmos

- Los algoritmos pasan a ser menos importantes
- Hace unos años los ingenieros de Google publicaron un paper justificando el papel de los datos
 - Algoritmos sencillos sobre grandes cantidades de datos vs. algoritmos muy sofisticados
 - Acaban siendo equivalentes
- Sin embargo, si los algoritmos no son tan importantes, ¿por qué el interés actual en tecnologías como el Deep Learning?
 - Por parte de los propios ingenieros de Google, que han hecho pública su tecnología TensorFlow, incluso distribuida a nivel de cluster
- Porque hay problemas que aún son difíciles incluso en presencia de grandes cantidades de datos - análisis del lenguaje natural, visión artificial, genoma

Big Data y algoritmos en medicina

- En el ámbito de la medicina existe una tendencia - medicina basada en la evidencia
 - Que los datos hablen
- Pero el proyecto del genoma humano ha dado muy magros resultados
- No se están alcanzando los santos griales de la medicina
 - Medicina personalizada - adaptada a nuestro propio código genético
 - El diseño de medicamentos basado en el análisis de proteínas
 - E incluso muchos de los estudios científicos están errados - falta de validez estadística
- La comprensión de las leyes de la biología debe estar guiada por teorías - el equivalente a algoritmos
 - Por ejemplo el descubrimiento del Bosón de Higgs requirió del análisis de petabytes de datos, pero en una búsqueda guiada por la teoría

Algoritmos para procesamiento del lenguaje

- El procesamiento del lenguaje humano sigue siendo un problema sin resolver
- Se han realizado avances significativos, muy orientados al usuario
 - Sistemas de respuesta a preguntas - IBM WATSON
 - Asistentes personales, Siri, Ok Google, Cortana
- Pero los usuarios siguen demandando aquello que mejor se adapta a sus usos y costumbres
- Por eso ha surgido una nueva tendencia, los robots conversacionales
- El chat como la interfaz de usuario universal
- Hay bots que simplifican muchas tareas
 - Transferir dinero, reservar vuelos, pagar en un restaurante, pedir una pizza

Chat bots

- Los bots se plantean como la nueva interfaz de usuario de las aplicaciones, desplazando a las Apps en los móviles
- 2016 es el año de los bots
- Evidencias
 - Proveedores de entornos conversacionales abren sus APIs
 - Servicios de Chat como Kik, WeChat, Telegram, Facebook Messenger
 - Plataformas de desarrollo de chat bots como Microsoft
- Algunas tareas son más eficientes con apps - pedir una pizza
- Los bots aplican en tareas en las que el objetivo no está claro, abiertas

Referencias

- Alon Halevy, Peter Norvig, and Fernando Pereira: [The unreasonable effectiveness of data](#).
- Peter Coveney, Edward R Dougherty: [Big data has not revolutionised medicine – we need big theory alongside it](#).
- Will Knight: [Is the Chatbot Trend One Big Misunderstanding?](#)
- Dan Driscoll, Mike Hall: [Building a Conversational Bot: From 0 to 60](#).