

INTRODUCCIÓN A CIENCIA DE DATOS

Dr. Mauricio Rosales Rivera



¿QUÉ ES UN CIENTÍFICO DE DATOS?

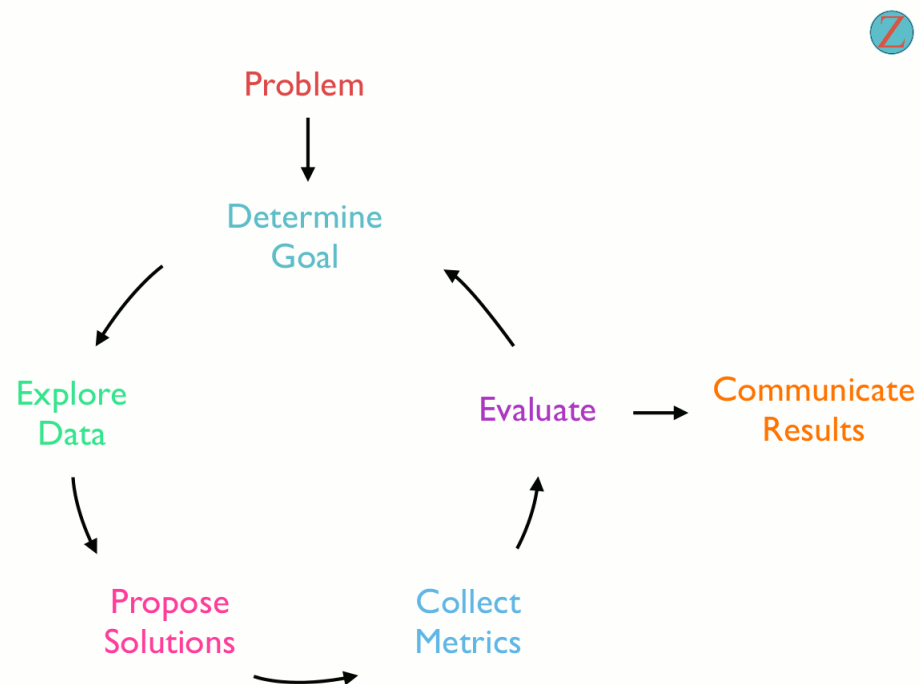
"Científico de datos (n): Persona que sabe más de estadística que cualquier programador y que a la vez sabe más de programación que cualquier estadístico"...

Un poco más en serio, un científico de datos es un profesional dedicado a **analizar e interpretar grandes bases de datos**. O lo que es lo mismo, uno de los profesionales más importantes en cualquier empresa de internet hoy en día.

¿PORQUÉ SE HA PUESTO DE MODA?

La tecnología actual no solo necesita del mejor talento sino de **datos, mucho datos**. Muchos...

Lo que vale para los entornos de *learning*, vale para casi cualquier



¿PORQUÉ SE HA PUESTO DE MODA?

Lo curioso es que este gran valor de los datos contrasta con que precisamente los datos son el recurso más abundante del planeta (se calcula que se crean 2.5 trillones de bytes de información nuevos al día).



¿CÓMO ES POSIBLE QUE ALGO TAN ABUNDANTE SEA TAN VALIOSO?

Acumular datos es algo trivial..... **Lo complejo es procesarlos.**

“La ciencia de datos es la disciplina de hacer que los datos sean útiles.”

Un mapa de Ciencia de Datos



Minería de datos
o Analítica

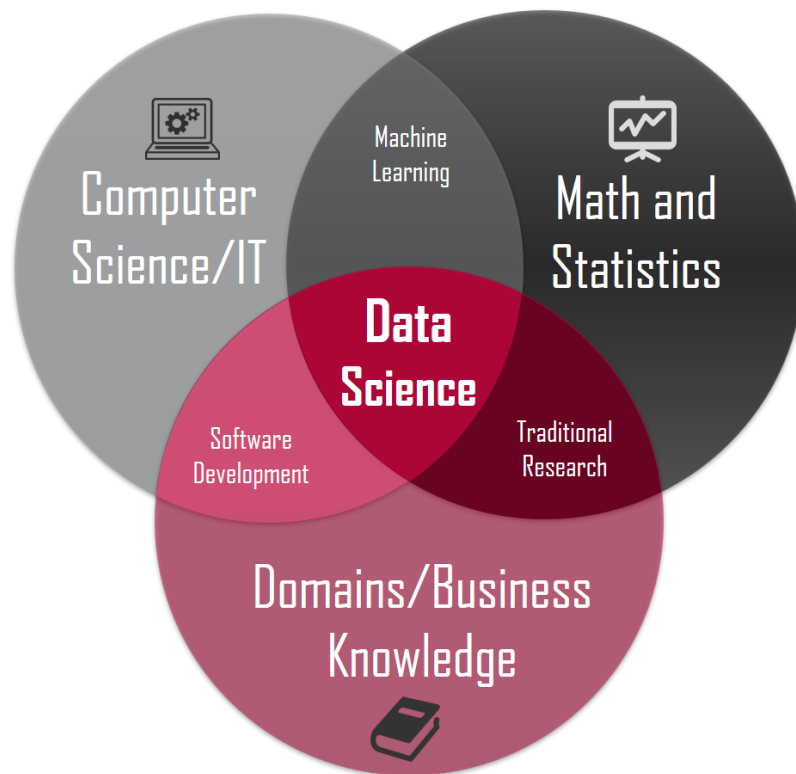


Estadística
inferencial



Machine
Learning

LA CIENCIA DE DATOS ES LA DISCIPLINA DE **HACER QUE LOS DATOS SEAN ÚTILES**



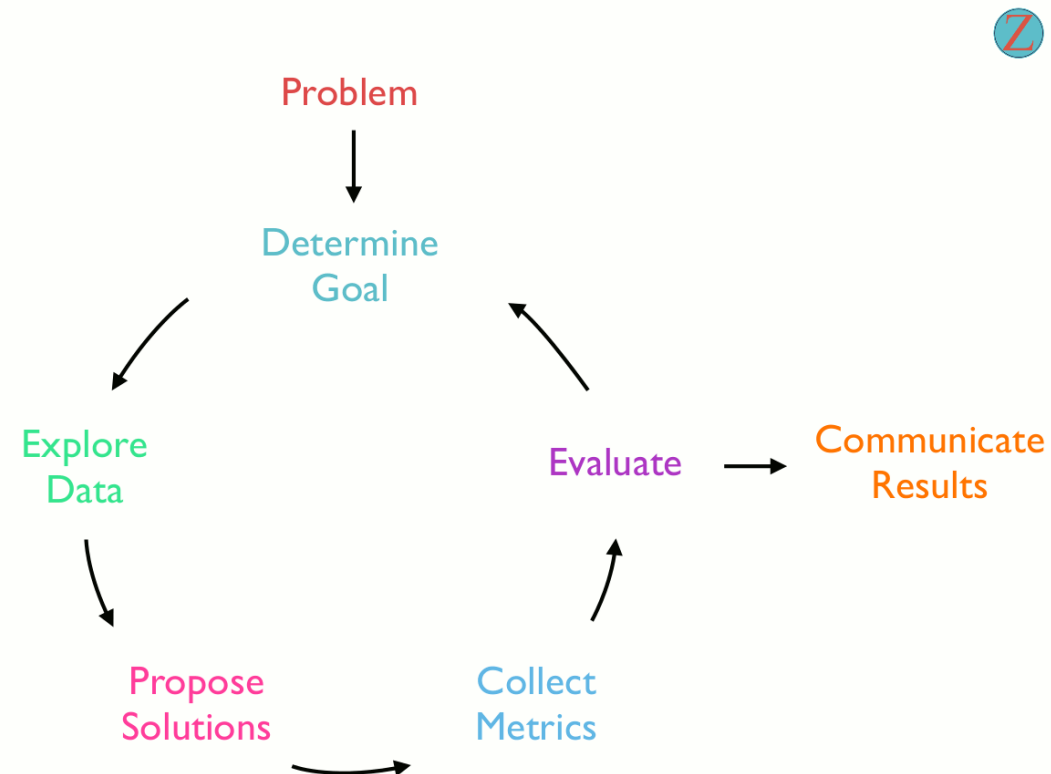
Desde hace unos años nos encontramos ante la primera gran 'democratización' de estas técnicas. Y, con ello, el boom de los científicos de datos: nadie quiere tener una mina de oro sin aprovechar.

Es a través de nuestras acciones—nuestras decisiones—que afectamos el mundo que nos rodea.

LA CIENCIA DE DATOS ES LA DISCIPLINA DE **HACER QUE LOS DATOS SEAN ÚTILES**

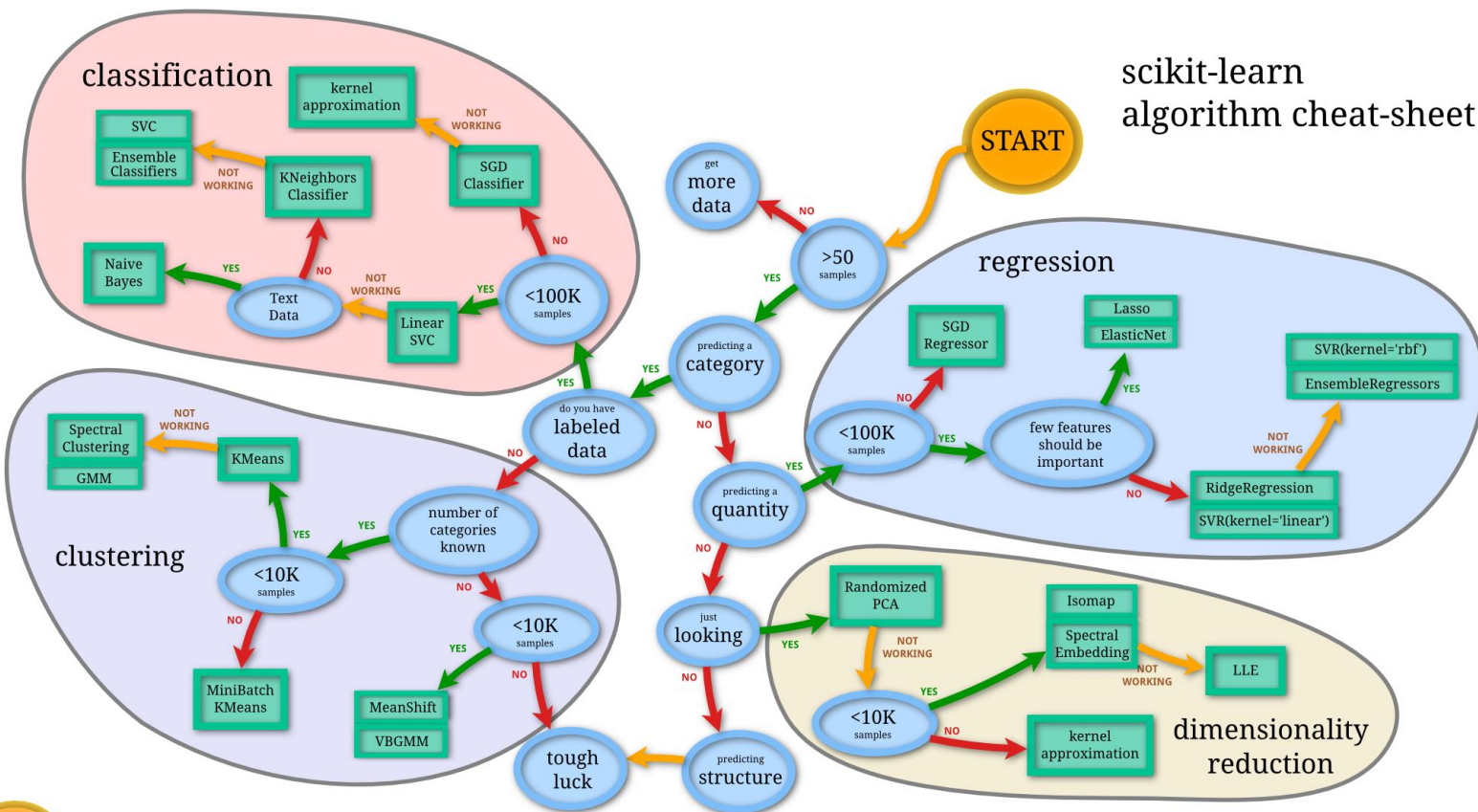


LA CIENCIA DE DATOS ES LA DISCIPLINA DE **HACER QUE LOS DATOS SEAN ÚTILES**



ALGUNAS TÉCNICAS DE *MACHINE LEARNING*





Con el tiempo, se han desarrollado muchos enfoques estadísticos de aprendizaje automático.

Puede usar este [link](#) del equipo de *scikit-learn* como guía para la mayoría métodos populares.