**Introducción a Minería de Datos**

**David Fernández Barrero**

Minería de datos está estrechamente relacionado con Machine Learning.

La principal idea de Machine Learning es descubrir patrones, relaciones de ciertos grupos de datos. De tal manera que estos patrones nos van a indicar pistas y dar información para resolver problemas reales.

Con machine Learning tratamos de extraer ese conocimiento.

Estos patrones se usarán para tomar decisiones. Decisiones basadas en hechos, no en ideas preconcebidas.

El proceso es (semi) automático.

El Machine Learning intenta predecir el futuro.

**Minería de datos**

Existe una necesidad de sacar valor de los datos.

El objetivo del minero de datos es extraer un modelo matemático.

Data Mining: Statistics + Databases + AI

Machine Learning: Field of AI

Statistics: Field of Mathematics

Big data: A lot of data

ML engineer: Professional role. Construye el modelo matemático.

Data scientist: Professional role. Analiza los datos establece modelos y traduce a un lenguaje que las personas puedan entender.

KDD (Knowledge Discovery in Databases): A process.

Big data necesita de Machine Learning.

La mineria de datos tiene dos enfoques, el modelo explicativo y el modelo predictive.

**KDD**

* Son procesos no triviales
* Son validos
* Se buscan modelos nuevos que no se conozcan
* Debe ser útil
* El modelo debe ser entendible por el humano o por la máquina

**Aplicaciones del Machine Learning**

* Detección de fraudes
* Aprobación de préstamos
* Sistemas de recomendación
* Redes sociales
* Detección de radicalización de células terroristas