

UNIDAD TEMÁTICA: Introducción a ML

EJERCICIO 1

¿Qué es “Machine Learning”?

hacer una búsqueda rápida y listar al menos 3 definiciones de “Machine Learning” y responder:

- ¿qué tiene en común y en qué se diferencia de “Inteligencia Artificial”?
- ¿qué tiene en común y en qué se diferencia de “Análisis Estadístico”?
- ¿Cómo se diferencia con Data Mining?
- ¿en qué se aplica?

Genera un breve documento de resumen. Podrás incluirlo en tu portafolio de ML como introducción o carátula!.

EJERCICIO 2

Breve investigación de herramientas y plataformas para Machine Learning.

- Listar los nombres, url donde se la describe y características más importantes:
- Funcionalidades disponibles (bloques / scripts / librerías) para
 - Tratamiento de datos
 - Algoritmos
 - Prueba y validación
- Capacidades de integración con aplicaciones (embebidas, cloud/web-services)

Desarrolla un breve documento de resumen, que podrás adaptar e incluir en el portafolios. Asegúrate de incluir RapidMiner y Weka.

EJERCICIO 3

Procesos de “Data Science”

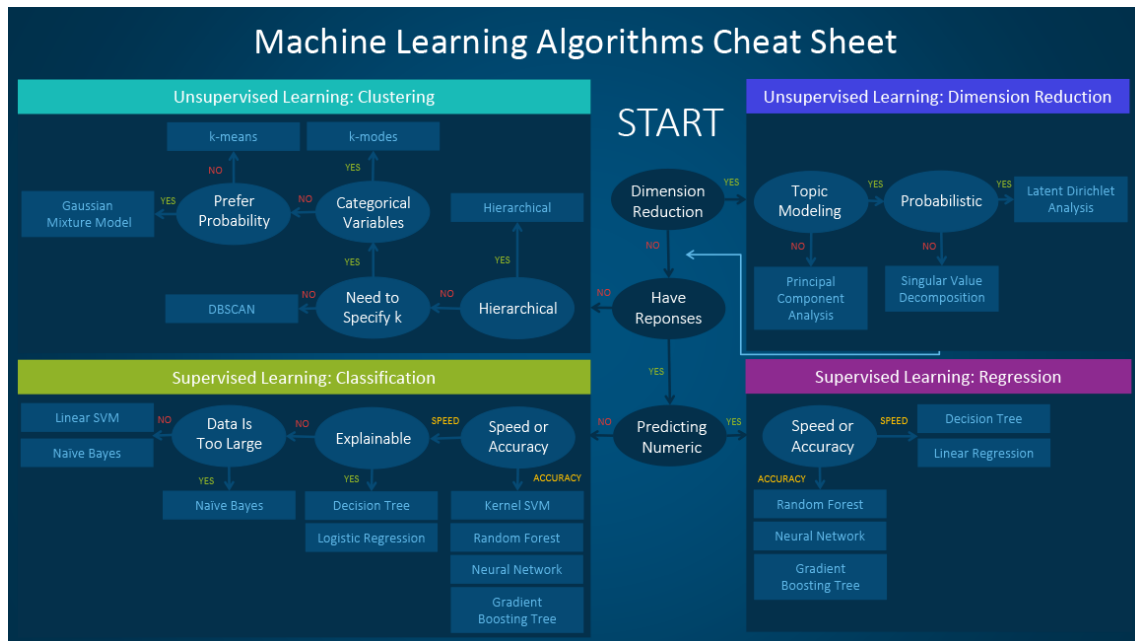
Busca información detallada del proceso CRISP-DM y realiza un breve resumen.

¿Qué otros procesos similares existen?

¿En qué se asemejan a CRISP-DM y en qué se diferencian?

EJERCICIO 4

La siguiente figura muestra una clasificación típica de algoritmos y aplicaciones de los mismos.



A partir de ésta, analiza algunas de las herramientas / plataformas que identificaste en el Ejercicio 2, observando cuáles algoritmos son soportados, y haciendo una revisión rápida de los parámetros que aceptan. Haz una tabla comparativa para al menos 2 herramientas. Asegúrate de incluir RapidMiner.

EJERCICIO 5

Utilizando los datasets del repositorio UCI: <http://archive.ics.uci.edu/ml/datasets.html>:

- Selecciona un dataset interesante, complejo, y que preferentemente haya sido objeto de muchas investigaciones
- Identifica el problema que aborda
- Estudia los atributos : tipos de datos, significado, rangos, cómo afectan a la(s) variables de predicción
- ¿qué tipos de algoritmos de Machine Learning pueden aplicarse para resolverlo? (revisa rápidamente algunas publicaciones referenciadas)