TEMA 0

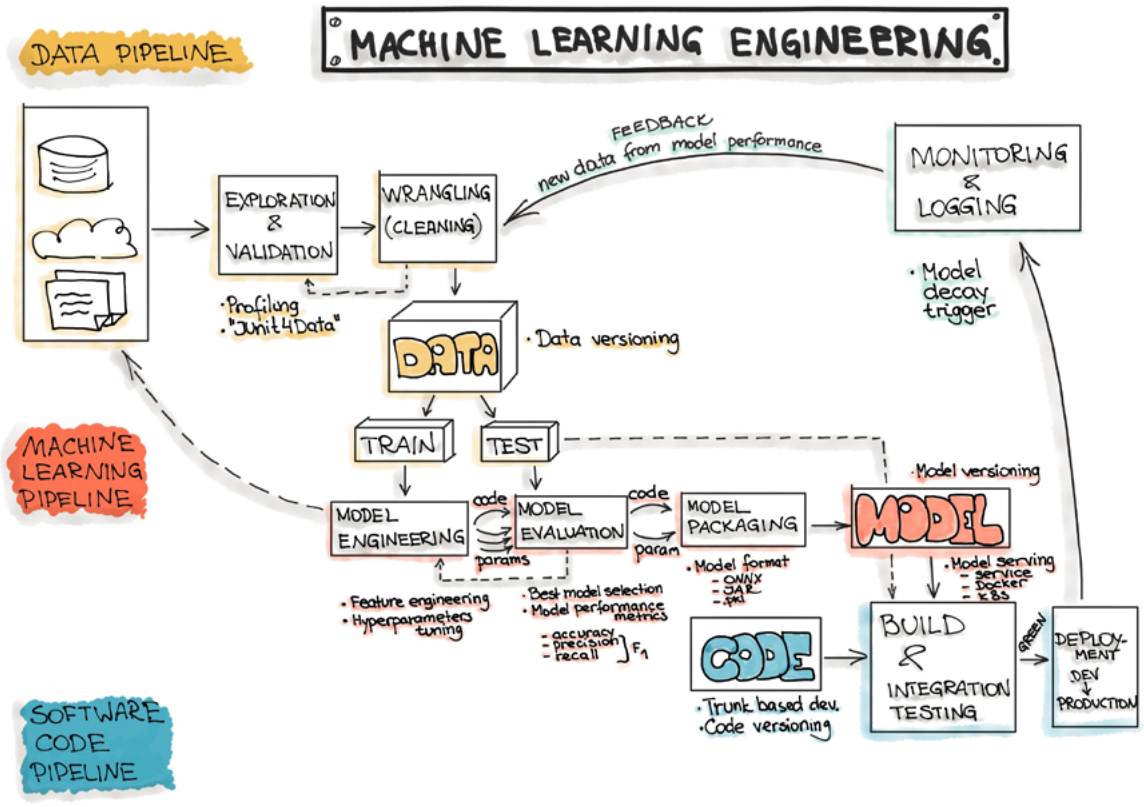
MÓDULO:
PROJECT MANEGEMENT

INTRODUCCIÓN AL BLOQUE 1
MÓDULO 8

Institut de Formació Contínua-IL3
UNIVERSITAT DE BARCELONA

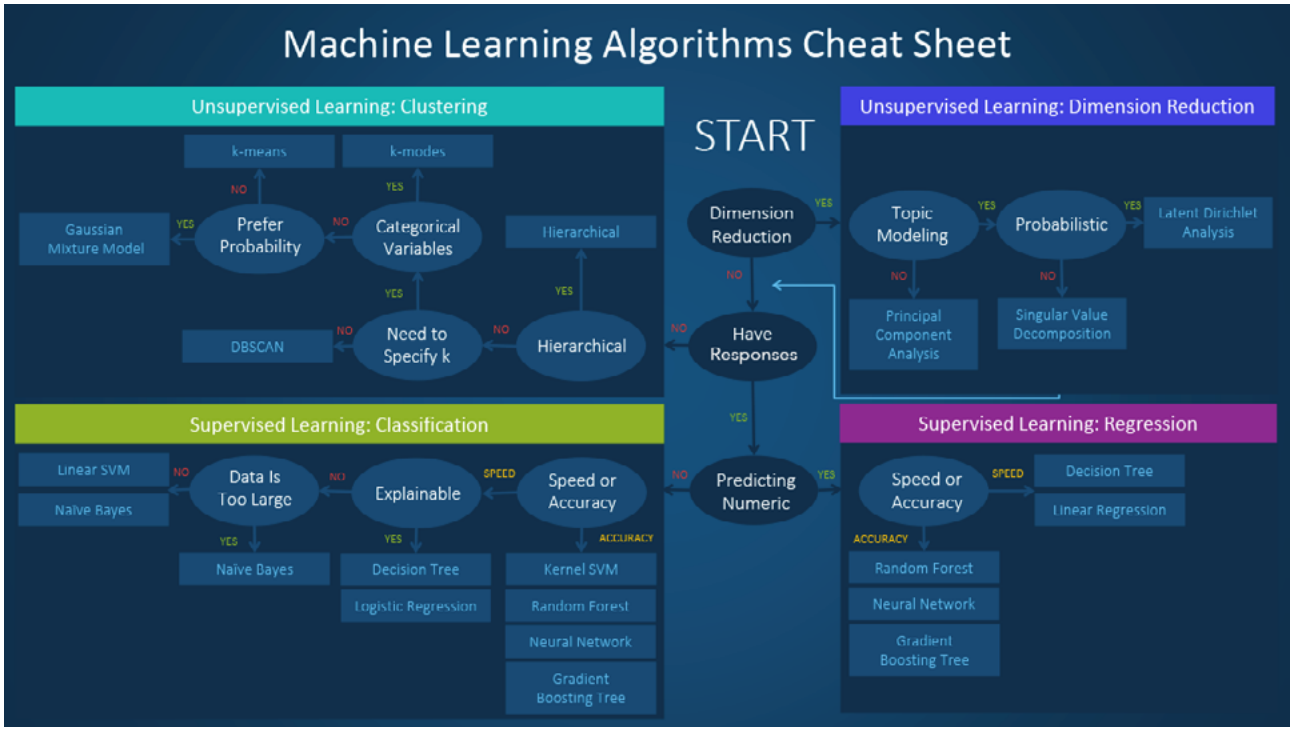
1. MIND MAP

- Tener presente en cada momento el mapa de la ingeniería del aprendizaje automatizado.



Fuente: https://ml-ops.org

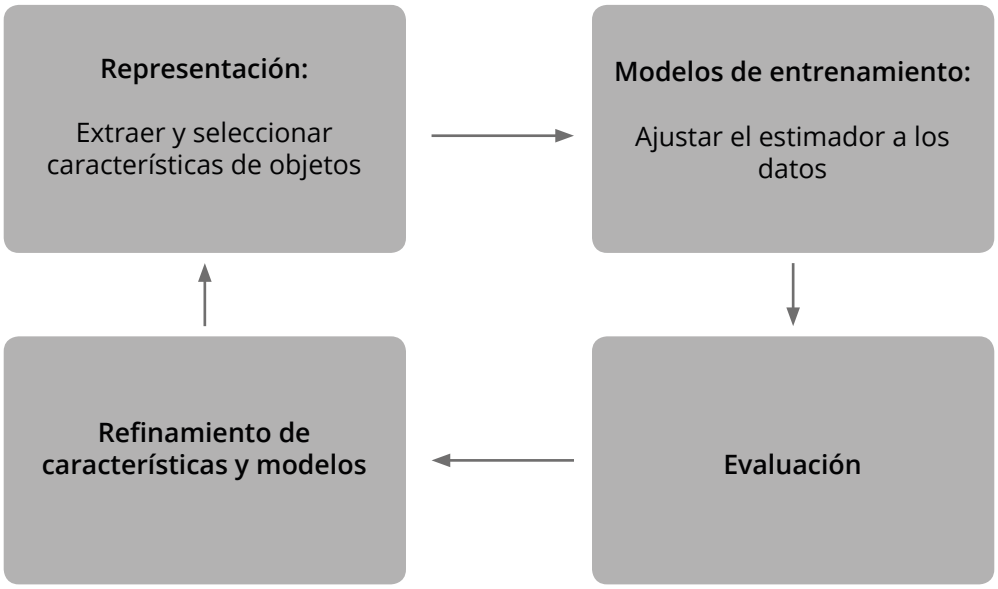
- Recordar siempre la diferencia entre probabilidad y estadística.
- Define el problema.
  - Tener siempre presente la hoja de referencia (Cheat Sheet):
    - ¿Qué problemas quieres solucionar?
  - ¿Cómo empezar? Empieza de forma simple.
  - Familiarízate con los datos y los resultados de referencia.
  - Después, prueba algo más complicado.



Fuente: https://blogs.sas.com/content/subconsciousmusings/2020/12/09/machine-learning-algorithm-use/

2. METODOLOGÍA PARA GARANTIZAR UNA SOLUCIÓN VIABLE

- Representar / capacitar / evaluar / perfeccionar el ciclo:



Fuente propia

3. VALIDACIÓN DE MODELOS

- Regresión lineal:
  - Error absoluto medio (EAM).
  - Error cuadrático medio (ECM).
  - R cuadrado.
  - R cuadrado ajustado.
  - El estudio de los residuos.
  - AIC.
  - BIC.
- Clasificación Binaria
  - Matrices de confusión.
  - Accuracy.
  - Precisión.
  - Recall (Sensibilidad).
  - FPR (Tasa de Falsos Positivos).
  - F1.
  - ROC – AUC.
  - Lift.
- Clasificador de clases múltiples:
  - Extensión del clasificador binario:
    - Matriz de confusión.
    - Informe de clasificación.
- Árboles de decisión:
  - Índice de Gini.
  - Entropía.

OBJETIVOS DEL BLOQUE

- Entender la ingeniería y validación de los modelos.
- Entender las fases de Diseño de aplicaciones impulsada por Machine Learning, su experimentación, su implementación, las operaciones asociadas con un buen despliegue y su posterior mantenimiento.

EVALUACIÓN

Evaluación continua del trabajo realizado en clase mediante la resolución de 3 partes:

1. **Prueba teórica:** al superar los **test** con éxito se alcanzará la posición de ***Initiate Level.***
2. **Prueba individual:** al superar el **trabajo individual** se logrará la posición ***Padawan Level.***
3. **Prueba grupal:** superar el **trabajo colectivo** supondrá conseguir la posición ***Knight Level .***

Criterios Mínimos

El *alumno/a Padawan* para alcanzar el nivel debe superar con éxito los siguientes hitos:

- ***Initiate Level:*** prueba de asentamiento de conceptos teóricos, para superar está parte deberás obtener una calificación superior a 5.

  Nota: Las preguntas que no se contesten de forma correcta restará puntos (indicado en cada actividad).

- ***Padawan Level:*** Realizar, al menos una práctica individual, defendiéndola y justificándola adecuadamente.

- ***Knight Level:*** Realizar al menos una práctica colectiva (participación activa en reuniones y discusiones de grupo, así como en la elaboración de informes, etc.), defendiéndola y justificándola adecuadamente.

Los porcentajes de cada hito estarán reflejados en el plan docente y en cada actividad.

Para aprobar el módulo, la media de todos los hitos debe ser superior al 5.

Recuerda que es evaluación continua por lo que cuantas más prácticas realices más posibilidades tendrás de alcanzar el máximo nivel Padawan.

BIBLIOGRAFÍA

L. Breiman, J. Friedman, R. Olshen y C. Stone, “Árboles de clasificación y regresión”, Wadsworth, Belmont, CA, 1984.

T. Hastie, R. Tibshirani y J. Friedman. “Elementos del aprendizaje estadístico”, Springer, 2009.

L. Breiman y A. Cutler, “Random Forests”,

Christopher Bishop, Pattern Recognition and Machine Learning, ISBN: 978-0-387-31073-2, XX, 738 Springer-Verlag New York, 2006.

Kuhn, Max, and Kjell Johnson. Applied predictive modeling. Vol. 26. New York: Springer, 2013.

Murphy, Kevin P. Machine learning: a probabilistic perspective. MIT press, 2012.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. The elements of statistical learning. Vol. 1. No. 10. New York: Springer series in statistics, 2001.

https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm

Chollet F. Deep Learning with Python. Ed. Manning, 2021.

Goodfellow I, Bengio Y, Courville A. Deep Learning. Ed. MIT Press, 2016.

James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning. Ed. Springer, 2013.

https://github.com/Azarodnyuk/MLatURL2021

https://medium.com/analytics-vidhya/ understanding-the-p-value-in-regression-1fc2cd2568af

https://blog.minitab.com/en/adventures-in-statistics-2/how-to-interpret-regression-analysis-results-p-values-and-coefficients

https://mathvault.ca/hub/higher-math/math-symbols/probability-statistics-symbols/

https://medium.com/usf-msds/choosing-the-right-metric-for-machine-learning-models-part-1-a99d-7d7414e4

https://en.wikipedia.org/wiki/Regression_validation

https://en.wikipedia.org/wiki/Bayesian_information_criterion

https://en.wikipedia.org/wiki/Akaike_information_criterion