Very good notebook to predict delay a investigar

<https://medium.com/analytics-vidhya/using-machine-learning-to-predict-flight-delays-e8a50b0bb64c>

<https://github.com/JaHerbas/Predicting_Flight_Delays>

Other very good notebook to predict delay a investigar

<https://developers.amadeus.com/blog/flight-delay-prediction-machine-learning>

Exercici 1

Ejemplo Web....notebook en GitHub

<https://medium.com/analytics-vidhya/using-machine-learning-to-predict-flight-delays-e8a50b0bb64c>

<https://medium.com/@banikanusheela/flight-departure-delay-prediction-417240a72ea4>

**DECISION TREE MODEL**

<https://www.datacamp.com/community/tutorials/decision-tree-classification-python>

**LOGISTIC REGRESSION**

<https://www.datacamp.com/community/tutorials/understanding-logistic-regression-python>?

**K-nearest neighbors (KNN)**

<https://stackabuse.com/k-nearest-neighbors-algorithm-in-python-and-scikit-learn/>

*Finally, the KNN algorithm doesn't work well with categorical features since it is difficult to find the distance between dimensions with categorical features.*

**D. SVM**

<https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python>

**E. XGBoost**

<https://www.datacamp.com/community/tutorials/xgboost-in-python>

**G. DEEP NEURAL NETWORK**

<https://medium.com/analytics-vidhya/using-machine-learning-to-predict-flight-delays-e8a50b0bb64c>

Exercici 3

#### **Ver document del curso Laia** [¶](http://localhost:8888/notebooks/Documents/IT%20ACADEMY/13.%20Aprenentatge%20Supervisat%20-%20Classificaci%C3%B3/S13%20T01_Aprenentage%20Supervisat.ipynb#Ver-document-de-Laia-) -> MachineLearning\_2021\_05\_18

**Decision Tree Performance Optimization**

<https://www.datacamp.com/community/tutorials/decision-tree-classification-python>

Exercici 4 TBC

Exercici 2: Compara els models de classificació utilitzant la precisió (accuracy), una matriu de confiança i d’altres mètriques més avançades.

Exercici 4: Compara el seu rendiment utilitzant l’aproximació traint/test o utilitzant totes les dades (validació interna)

Me parece entender que en el 2 te pide calcular las métricas con y\_train y predicciones x\_train y el 4 sería calcular las métricas con y\_test y predicción x\_test (o con cross-validation), no?

si

<https://www.analyticsvidhya.com/blog/2018/05/improve-model-performance-cross-validation-in-python-r/>

entenc que validació interna es refereix a cross-validation

<https://www.analyticsvidhya.com/blog/2018/05/improve-model-performance-cross-validation-in-python-r/>

Exercici 6

A ver ideas en <https://medium.com/analytics-vidhya/using-machine-learning-to-predict-flight-delays-e8a50b0bb64c>

TBC webs

<https://medium.com/@banikanusheela/flight-departure-delay-prediction-417240a72ea4>

<https://towardsdatascience.com/how-to-best-evaluate-a-classification-model-2edb12bcc587>

<https://www.ritchieng.com/machine-learning-evaluate-classification-model/>

<https://towardsdatascience.com/hackcvilleds-4636c6c1ba53>