

EXERCICI D'ENTREGA OBLIGATÒRIA

Objectiu: aplicar les diferents tècniques vistes a classe a un problema real. Neteja de dades, analitzar les característiques del problema, estudiar y ajustar els paràmetres del algorismes d'aprenentatge automàtic.

Materials: La pràctica es realitzarà en Python emprant la llibreria scikit-learn[1] (es recomana l'entorn Anaconda[2]) y la entrega serà un PDF resultat d'un Python Notebooks, a més aquest PDF ha de contenir un enllaç a un repositori públic de codi amb tots els fitxers necessaris per la seva execució.

Puntuació: el problema es puntuarà seguint la següent rúbrica

APARTAT	PUNTS	PUNTS TOTALS	PES
General			
L'alumne entén el problema i es capaç de proveir una solució valida al problema	50	100	0,3
La solució presentada és executable sense errors	50		
Data Wrangling			
L'alumne es capaç de carregar les dades i entendre la seva estructura general.	20	100	0,3
L'alumne es capaç de realitzar un anàlisi exploratori correcte i detectar outliers i col·linealitat	20		
L'alumne és capaç de tractar els nuls correctament.	20		
L'alumne avalua la possibilitat si una reducció de la dimensionalitat pot ser beneficiosa .	20		
L'alumne empra visualitzacions amb l'objectiu d'entendre les dades.	20		
Modelització			
L'alumne es capaç d'escollir la mètrica més adient per la selecció i/o avaluació del model.	10	100	0,4
<i>Training + Test + validation approach / cross-validation</i>	20		
El model final es capaç de captar senyal i no renou i es justifica l'elecció d'un model o un altre	10		
L'alumne realitza un model de base simple amb el qual comparar el diferents models candidats.	20		
L'alumne és capaç d'entendre el model. Quin són els seus <i>main drivers</i> ?	10		
L'alumne es capaç d'entregar una solució estructurada i neta	30		
Total			
Total		300	100



Pràctica

Airbnb és una plataforma que permet publicar, descobrir i reservar diferents allotjaments. Disposa de propietats en més de 192 països repartides en 33.000 ciutats d'arreu del món. Es va crear a finals de l'any 2008 i avui dia és un dels recursos més buscats sobretot per viatjants. Únicament a Mallorca disposa de 14900 allotjaments de diferent tipus.

Un dels problemes que es troben els *hosts* quan volen afegir un nou allotjament a la plataforma és determinar el preu de lloguer. En aquest problema la pregunta que haureu de resoldre és: quin preu de lloguer posar a una propietat sabent les seves característiques?

Per respondre a aquesta qüestió disposeu del conjunt de dades real d'*Airbnb* per l'illa de Mallorca (el trobareu a l'aula digital). Finalment us proposem dues opcions:

1. Estimar el preu directament a partir de les característiques dels allotjament.
2. Convertir el preu en rangs i classificar-les en un d'aquests grups.

Data d'entrega: El darrer dia per entregar la pràctica serà el dia 17 de febrer, NO S'ADMETRAN entregues posteriors.

Referències:

[1] <http://scikit-learn.org/stable/>

[2] <https://www.anaconda.com/>