



Guia docent

230338 - APATP - Aprenentatge Automàtic: de la Teoria a la Pràctica

Última modificació: 17/06/2024

Unitat responsable: Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona
Unitat que imparteix: 739 - TSC - Departament de Teoria del Senyal i Comunicacions.

Titulació: GRAU EN ENGINYERIA DE TECNOLOGIES I SERVEIS DE TELECOMUNICACIÓ (Pla 2015). (Assignatura optativa).
GRAU EN ENGINYERIA ELECTRÒNICA DE TELECOMUNICACIÓ (Pla 2018). (Assignatura optativa).

Curs: 2024 **Crèdits ECTS:** 2.0 **Idiomes:** Català, Castellà

PROFESSORAT

Professorat responsable: Josep Vidal

Altres: Josep Vidal, Verónica Vilaplana

CAPACITATS PRÈVIES

Estadística

REQUISITS

GRAU EN ENGINYERIA DE TECNOLOGIES I SERVEIS DE TELECOMUNICACIÓ (GRETST)
PROBABILITAT I ESTADÍSTICA - Prerequisit
GRAU EN ENGINYERIA ELECTRÒNICA DE TELECOMUNICACIÓ (GREELEC)
PROBABILITAT I PROCESSOS ESTOCÀSTICS - Prerequisit

METODOLOGIES DOCENTS

El temps es repartirà entre classes de teoria i laboratori en Python.

OBJECTIUS D'APRENTATGE DE L'ASSIGNATURA

Aquesta assignatura fa una revisió de la teoria de la classificació i la regressió basats en dades des d'una perspectiva matemàtica i al mateix temps aplicada a diferents àrees. Es pretén donar una visió general de la teoria de decisió bayesiana, tècniques de classificació paramètrica (regressió logística, xarxes neuronals, arbres de decisió, combinació de classificadors, etc) i aprenentatge no supervisat amb una experimentació simultània en diferents aplicacions d'entre les que destaquem decisió sobre dades biomèdiques, classificació d'imatges, detecció d'esdeveniments, detecció de spam, etc..

Els crèdits de l'assignatura es reparteixen en crèdits teòrics i crèdits pràctics (laboratori en Python) per tal de paral·lelitzar els coneixements teòrics amb el desenvolupament d'aplicacions.

En les darreres setmanes del curs, tots els alumnes participaran en una competició interna de machine learning proposada pels professors.



HORES TOTALES DE DEDICACIÓ DE L'ESTUDIANTAT

Tipus	Hores	Percentatge
Hores grup gran	10,0	20.00
Hores aprenentatge autònom	30,0	60.00
Hores grup petit	10,0	20.00

Dedicació total: 50 h

CONTINGUTS

1. Introducció a l'aprenentatge automàtic

Descripció:

Introducció a l'aprenentatge automàtic. Aprenentatge supervisat i no supervisat. Regressió i classificació

Activitats vinculades:

Laboratori 1. Introducció a Python, Google Colab i Scikit-Learn. Lectura de marcs de dades i anàlisi exploratòria de dades mitjançant el conjunt de dades Iris i el conjunt de dades sobre el càncer de mama. Exemples senzills que mostren com resoldre tasques supervisades (classificació) i no supervisades (reducció de dimensionalitat i agrupació) mitjançant el conjunt de dades Iris

Dedicació: 4h 15m

Grup gran/Teoria: 2h

Activitats dirigides: 2h

Aprenentatge autònom: 0h 15m

2. Regressió

Descripció:

Regressió. Models, funcions de cost i formació. Regressió lineal i polinomial. Entrenament de descens de gradients. Sobreajustament, subadaptació i generalització. Regularització: llaç, cresta i xarxa elàstica. Validació. Mesures de rendiment: MSE, MAP, coeficient R, correlació de Pearson. Reptes en l'aprenentatge associats a la qualitat de les dades.

Activitats vinculades:

Pràctica 2. Regressió lineal i regularització mitjançant el conjunt de dades MTCars. Solució de forma tancada i solució amb descens de gradient. Regressió polinòmica.

Dedicació: 4h 15m

Grup gran/Teoria: 2h

Activitats dirigides: 2h

Aprenentatge autònom: 0h 15m

3. Classificació

Descripció:

Classificació de dues classes, classes múltiples i etiquetes múltiples. Regressió logística. Límits de decisió. Funció de cost i entrenament per gradient. Mesures de rendiment: precisió, precisió, recordatori, puntuació F, matriu de confusió, ROC. Regularització.

Activitats vinculades:

Pràctica 3. Regressió logística utilitzant el conjunt de dades Iris i el conjunt de dades del càncer de mama. Matriu de confusió i mètriques de classificació. Parts de prova de tren. Presentació del concurs d'aprenentatge automàtic.

Dedicació: 4h 15m

Grup gran/Teoria: 2h

Activitats dirigides: 2h

Aprenentatge autònom: 0h 15m

4. Xarxes neuronals

Descripció:

Xarxes neuronals. Estructures. Regressió/classificació. Funcions de costos. Algorisme de retropropagació. Tècniques de descens de gradients. Entrenament per lots. Sobreajustament i generalització. Regularització i evitació del sobreajust. RNN i DNN. Regles d'una competició d'aprenentatge automàtic. Validació del plegat en K.

Activitats vinculades:

Pràctica 4. Classificador de xarxes neuronals multicapa. Conductes Sklearn. Cerca d'hiperparàmetres mitjançant GridSearchCV. Regularització. Aplicació a conjunts de dades Iris i MNIST.

Dedicació: 4h 15m

Grup petit/Laboratori: 2h

Activitats dirigides: 2h

Aprenentatge autònom: 0h 15m

5. Arbres de decisió

Descripció:

Arbres de decisió. Entrenament, sobreajustament i poda. Rendiment en termes de biaix i variància. Bosc aleatori i aprenentatge ensamble.

Activitats vinculades:

Laboratori 5. Arbres de decisió: formació, visualització, regularització. Ensamble classificadors: bosc aleatori, embolcall, boosting. Classificadors de votació (tou i dur). Aplicació a conjunts de dades Iris i MNIST.

Dedicació: 4h 15m

Grup gran/Teoria: 2h

Activitats dirigides: 2h

Aprenentatge autònom: 0h 15m

6. Aprenentatge no supervisat

Descripció:

Aprenentatge automàtic no supervisat. K-means i models de mescleres gaussianes. Algorisme d'expectació-maximització. Selecció de l'ordre del model: mètodes BIC i silueta.

Dedicació: 2h 15m

Grup gran/Teoria: 2h

Aprenentatge autònom: 0h 15m



SISTEMA DE QUALIFICACIÓ

Qüestionaris: 50%

Competició: 50%

NORMES PER A LA REALITZACIÓ DE LES PROVES.

No es pot fer servir calculadora, telèfons mòbils ni apunts.

BIBLIOGRAFIA

Bàsica:

- Aurelien Geron. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. 2. UK: O'Reilly Media, 2022. ISBN 978-1098125974.

Complementària:

- Bishop, C.M. Pattern recognition and machine learning. New York: Springer, 2006. ISBN 0387310738.

- Kuncheva, L.I. Combining pattern classifiers: methods and algorithms [en línia]. 2nd ed. Hoboken (NJ): J. Wiley & Sons, 2014 [Consulta: 21/09/2018]. Disponible a: <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118914564>. ISBN 9781118914564.

- Hastie, T.; Tibshirani, R.; Friedman, J.H. The elements of statistical learning: data mining, inference and prediction [en línia]. 2nd ed. New york [etc.]: Springer, 2009 [Consulta: 18/07/2023]. Disponible a: <https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/978-0-387-84858-7>. ISBN 9780387848570.

RECURSOS

Altres recursos:

Cursos online

<https://es.coursera.org/learn/machine-learning?> />Competicions

<http://www.kdd.org/kdd-cup> /><http://www.kaggle.com/competitions> /><https://datahack.analyticsvidhya.com/contest/all/> />