**Pràctica**

Presentació

Aquesta Pràctica cobreix tots els mòduls del programa de l'assignatura.

Competències

Les competències que es treballen en aquesta prova són:

-Ús i aplicació de les TIC en l'àmbit acadèmic i professional.

-Capacitat per innovar i generar noves idees.

-Capacitat per avaluar solucions tecnològiques i elaborar propostes de projectes tenint en compte els recursos, les alternatives disponibles i les condicions de mercat.

-Conèixer les tecnologies de comunicacions actuals i emergents i saber-les aplicar convenientment per a dissenyar i desenvolupar solucions basades en sistemes i tecnologies de la informació.

-Aplicació de les tècniques específiques d'enginyeria del programari a les diferents etapes del cicle de vida d'un projecte.

-Capacitat per aplicar les tècniques específiques de tractament, emmagatzematge i administració de dades.

-Capacitat per proposar i avaluar diferents alternatives tecnològiques per a resoldre un problema concret.

Objectius

La correcta assimilació de tots els aspectes treballats al llarg del semestre.

En aquesta Pràctica abordem un problema real de mineria de dades on hem de posar en joc tots els conceptes treballats.

Cal treballar tot el cicle de vida del projecte. Des de l'objectiu del projecte fins la implementació del coneixement trobat passant per la preparació, neteja de les dades, coneixement de les dades, generació del model, interpretació i avaluació.

Descripció de la Pràctica a realitzar

La prova està estructurada en un total d'un exercici pràctic.

Recursos

**Bàsics**

Material docent proporcionat por la UOC.

**Complementaris**

* Els descrits en les anteriors PAC.
* Fitxer gourmetdb.zip

Criteris de valoració

**Exercicis pràctics**

És necessari documentar en cada apartat de l'exercici pràctic què s'ha fet i com s'ha fet, quin era l’objectiu i com s’ha desenvolupat.

Format i data d’entrega

El format de lliurament és: usernameestudiant-practica.zip.

Estructura del contingut del fitxer zip

+ Detall composició del paquet..............  
--Index\_documents\_practica.txt  
+ Respostes................................................  
--- username\_estudiant\_MD\_Practica.pdf/odt/docx/html/rmd

--- username\_estudiant\_MD\_Presentacio.pdf/odt/docx/html/rmd  
+ Fitxer de dades generats a partir dels originals amb transcendència.

--- md\_prac-n.+1.\*

Data de lliurament: 09/01/2019

Cal lliurar la pràctica la bústia de lliuraments de l'aula.

Enunciat

**Introducció:**

Volem plantejar com podria ser un hipotètic cas real de projecte de mineria de dades. Tot parteix d'una presentació del client que exposa un problema de negoci difús i molt genèric que haureu de reconduir com a projecte de mineria de dades. Heu treballat un cicle de vida de projecte que ara té que ser el pal de paller de les respostes.

Transformeu els objectius de negoci en objectius de mineria de dades.

Prepareu i seguiu una estratègia.

Analitzeu bé les dades, enteneu el model que representen les dades.

Preparació de dades, neteja, derivació de dades noves, enriquiment, extracció visual de coneixement, creació del model, avaluació de la qualitat, interpretació i proposta de què es pot fer amb el coneixement extret.

Orienteu la pràctica com podria ser la presentació de resultats que li faríeu al client. Per tant llenguatge clar i ordenat. Podeu fer dos documents, una presentació breu amb les conclusions orientada al client i un altre document recollint tota la feina feta.

Les dades. Estan en fitxers csv. Hi ha un script de SQL que us pot ajudar a entendre la tipologia de les dades i les seves relacions. Si no queda clar pregunteu al Fòrum.

Per preparar les dades creativitat i llibertat absoluta: R, Excel, Acces, Mysql... El que vulgueu. Treballeu en les eines que sigueu més eficients i productius. Si us perdeu o quedeu aturats en algun punt no dubteu en contactar amb mi.

**Cal aplicar els tipus d'algorismes de MD vistos al semestre. Per tant, com a mínim tindreu tres objectius de MD. Atenció: les dades mostren resultats rellevants per models d´agregació i regles d´associació. Els resultats en algorismes de classificació depenen molt del plantejament que es faci i com es treballi. Per tant en aquest cas no és tant important el resultat i la qualitat del model trobat sinó la feina ben plantejada, raonada i executada.**

**Presentació del client:**

*Hola, em dic Joan Schlemmer. Sóc CEO a la cadena de botigues presencials Gourmet. Ens dediquem al negoci de l'alimentació i beguda de qualitat. Tenim presència en botigues a Europa i Estats Units.*

*El negoci no ens va malament però m'agradaria poder augmentar la facturació de l'empresa. No tinc clar si augmentant les vendes o reduint despeses... L'altre dia jugant al golf amb un company que es dedica al negoci de les comunicacions en va comentar que havia engegat un projecte de... Big Data o Mineria de Dades o no se que del coneixement... No ho tinc clar... Em va dir que li varen aconsellar accions concretes a partir de dades recollides del seu negoci i que està content... La qüestió és que us enviaran unes dades extretes del nostre ERP perquè em feu això de la mineria. Si us plau si teniu dubtes amb les dades pregunteu als nostres tècnics. Quan podré tenir un primer informe? D'aquí un mes? Perfecte...*

|  |
| --- |
| **Nota:** Propietat intel·lectual  Sovint és inevitable, al produir una obra multimèdia, fer ús de recursos creats per terceres persones. És per tant comprensible fer-lo en el marc d'una pràctica dels Estudis, sempre que això es documenti clarament i no suposi plagi en la pràctica.  Per tant, al presentar una pràctica que faci ús de recursos aliens, s'ha de presentar juntament amb ella un document en quin es detallin tots ells, especificant el nom de cada recurs, el seu autor, el lloc on es va obtenir i el seu estatus legal: si l'obra està protegida pel copyright o s'acull a alguna altra llicència d'ús (Creative Commons, llicència GNU, GPL ...).  L'estudiant haurà d'assegurar-se que la llicència no impedeix específicament el seu ús en el marc de la pràctica. En cas de no trobar la informació corresponent haurà d'assumir que l'obra està protegida per copyright.  Hauríeu de, a més, adjuntar els fitxers originals quan les obres utilitzades siguin digitals, i el seu codi font si correspon. |