

**Intel·ligència artificial**: Projecte2

Albert Soto Serrano (1361153)

Sergio Paulo Rodríguez (1306688)

Levani Makharashvili (1362905)

**ImageSoft Corporation**

Etiquetatge dels colors d’imatges

**Índex**

**1. Introducció i presentació del problema p. 2**

1.1 Objectius p. 2

1.3. Objectius personals p. 3

1.4. Material utilitzat p. 3

**2. Metodologia de resolució p. 4**

2.1. Algorisme K-means p. 4

2.2. Test de Fisher p. 6

2.3. Altres algorismes p. 8

**3. Resultats p. 9**

3.1. RGB p. 9

3.2. RGB + Potentials p. 10

3.3. HSV + Potentials p. 10

**4. Conclusions p. 11**

**5. Bibliografia p. 12**

**1. Introducció**

L’empresa ImageSoft, els clients de la qual són empreses que venen imatges o que s’encarreguen de gestionar grans bases d’imatges, s’està plantejant un canvi d’estratègia i pretén començar a oferir-los software més ‘intel·ligent’ que permeti incrementar la productivitat d’alguns dels seus processos.

Una de les necessitats que han detectat en aquestes empreses és l’anotació automàtica de les imatges que tenen a la base de dades.

L’objectiu de l’empresa es poder disposar d’un nou software que donada una imatge sigui capaç d’indicar els colors que hi predominen per tal d’agilitzar les funcions de cerca, identificació i gestió.

**Com podem realitzar aquest programari? Quins algorismes caldrà aplicar per obtenir les etiquetes desitjades?**

|  |  |
| --- | --- |
| **Funcionalitat del software** | **Qualitat del software** |
| * Dissenyar un programa capaç d’etiquetar un conjunt d’imatge amb els colors que més hi predominen. * El programa ha de poder treballar amb diferents espais de colors, incloent RGB, Potentials, etc. per tal d’obtenir el millor resultat possible. | * El programa ha de ser lleuger, ràpid i eficaç. * El marge d’error a l’hora d’etiquetar els colors d’una imatge ha de ser el menor possible. * Ha d’implementar correctament algorismes de clústering o d’agrupació de conjunts de dades. |

**1.2. Objectius de la pràctica**

**1.3. Objectius a nivell personal**

Els objectius a nivell personal d’aquesta pràctica es basen en l’aprenentatge d’algorismes de *clústering* o *data mining*. On s’han de gestionar grans conjunts de dades per extreure’n la informació que considerem rellevant.

En el nostre cas concret, treballarem l’algorisme K-means, que és una tècnica d’aprenentatge no supervisat sobre un conjunt de dades.

En el nostre cas aquest conjunt de dades es tracta de punts d’un conjunt d’imatges.

A més, utilitzarem estadístics com el discriminant de Fisher per a determinar el nombre de clústers adients per a cada imatge.

**1.4. Material utilitzat**

Per al desenvolupament de la pràctica hem fet servir el següent programari:

|  |  |
| --- | --- |
| **Nom del programari** | **Ús** |
| **Ubuntu 15.04** | Sistema Operatiu sobre el que hem treballat |
| **Python 2.7** | Llenguatge de programació en què està programada la pràctica.  Es tracta d’un llenguatge interpretat de gran potencial. |
| **Geany, gedit, etc** | Diferents editors de text que hem utilitzat per a desenvolupar el nostre codi. |
| **Codi esquelet proporcionat pel professorat** | Conté l’esquelet de la pràctica i on es produeix la major part del tractament de la imatge en el procés de transformar-la en una estructura de dades fàcilment manejable. |
| **Imatges de prova** | Conjunt d’imatges sobre les que hem realitzat les nostres proves. |
| **Microsoft Office 2013** | Realització de la memòria |

**2. Metodologia de la resolució**

Per a resoldre el problema descrit, farem servir bàsicament dos algorismes: K-means i el de càlcul del discriminant de Fisher.

La planificació del treball ha comportat les següents fases:

1. Comprensió del problema que se’ns planteja.
2. Comprensió dels algorismes que ens permetran resoldre el problema: K-means i test de Fisher.
3. Primera implementació de K-means amb petit conjunt de dades de dues dimensions.
4. Implementació de l’algorisme K-means ja a l’esquelet de la pràctica proporcionada.
5. Implementació del càlcul del discriminant de Fisher.
6. Implementació d’altres espais de colors.
7. Ajustar paràmetres per a obtenir un millor rendiment.

**2.1. Algorisme K-Means**

L’algorisme K-means es un algorisme utilitzat com a tècnica d’aprenentatge no supervisat, on no existeix un conjunt de mostres d’aprenentatge de les que es coneix a quina classe pertany cada una.

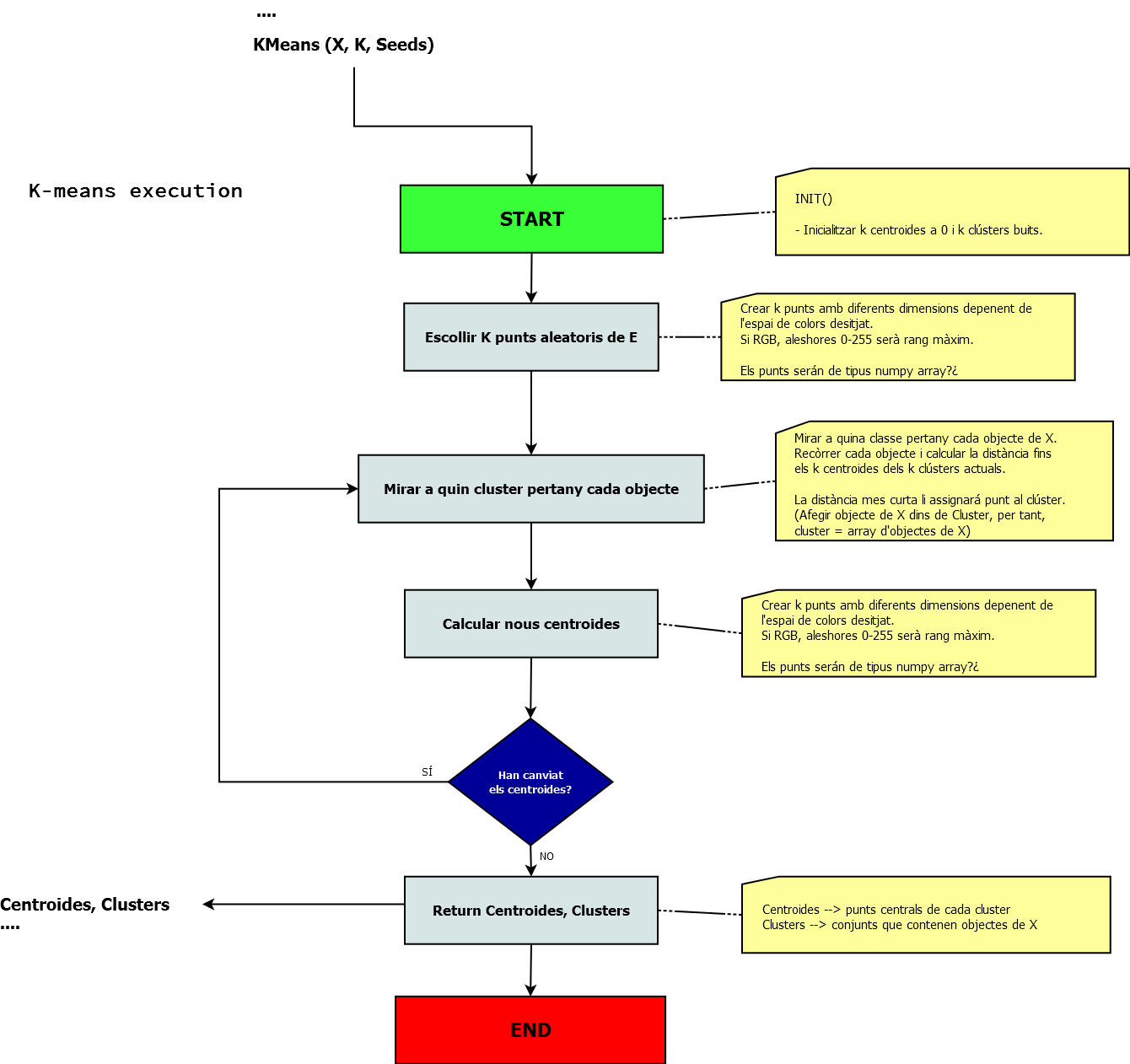
Per tant K-means planteja una solució al buscar automàticament l’estructura de les dades per agrupaments (*Clustering*)

Per a fer-ho, s’estableixen un centroides, que seran els punts centrals dels nostres clústers (agrupacions).

L’algorisme anirà assignant a cada dada un clúster basant-se en la proximitat a aquest i anirà recalculant nous centroides a partir dels nous punts dins de cada clúster.

L’algorisme finalitzarà quan els centroides no variïn d’una a una altra iteració.

A la pàgina següent es pot veure un diagrama de flux on s’explica el funcionament de l’algorisme i de la seva implementació a la pràctica:



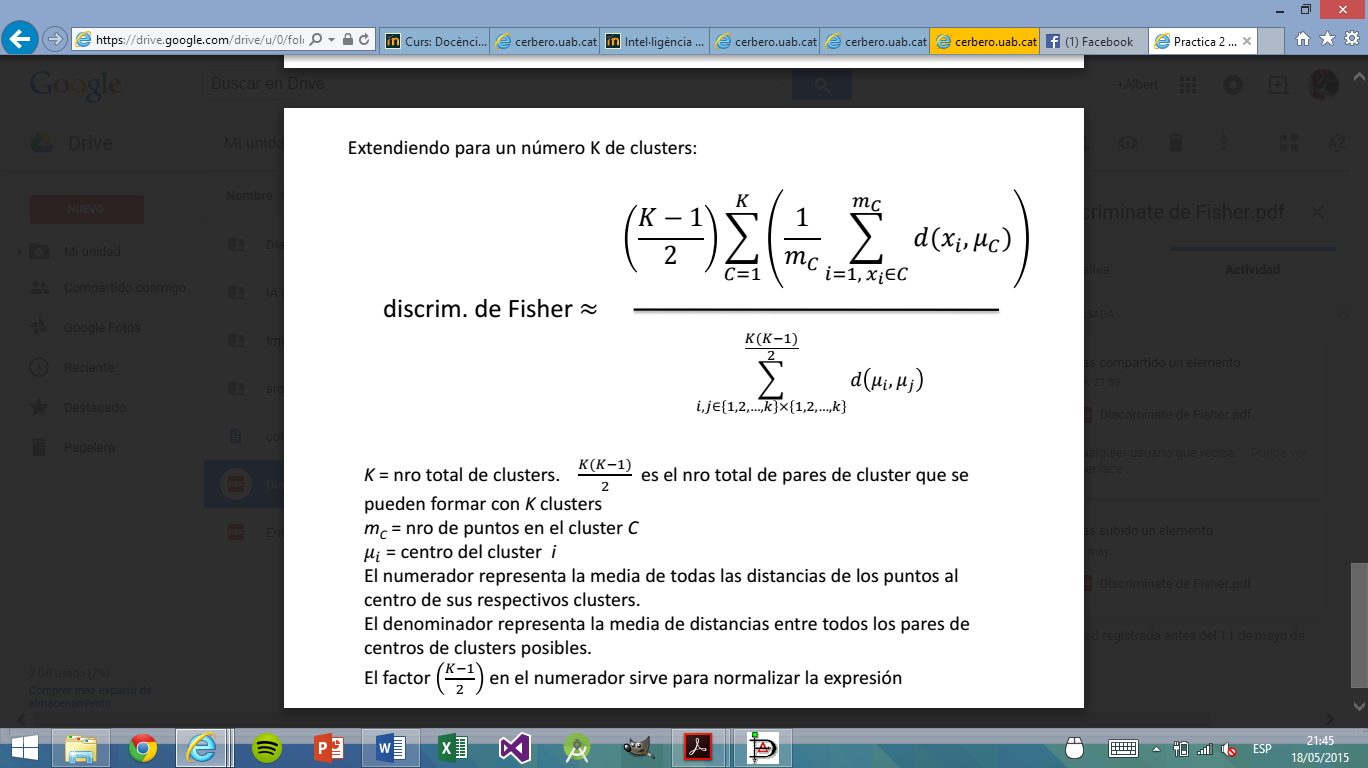
*\*Diagrama de flux de l’algorisme K-means*

|  |  |
| --- | --- |
| **K-Means** | |
| **Fortaleses:** | **Debilitats:** |
| – Relativament eficient: on N és el número d’objectes, K és el número de clústers, i T el número d’iteracions.  Normalment,  K, T << N.  – El procediment sempre acaba satisfactòriament | – No necessàriament troba la configuració més òptima.  – Significativament sensible a la selecció inicial aleatòria dels centres dels clústers.  – Aplicable només quan la mitjana està definida.  – Cal especificar K, el número de clústers, per avançat. |

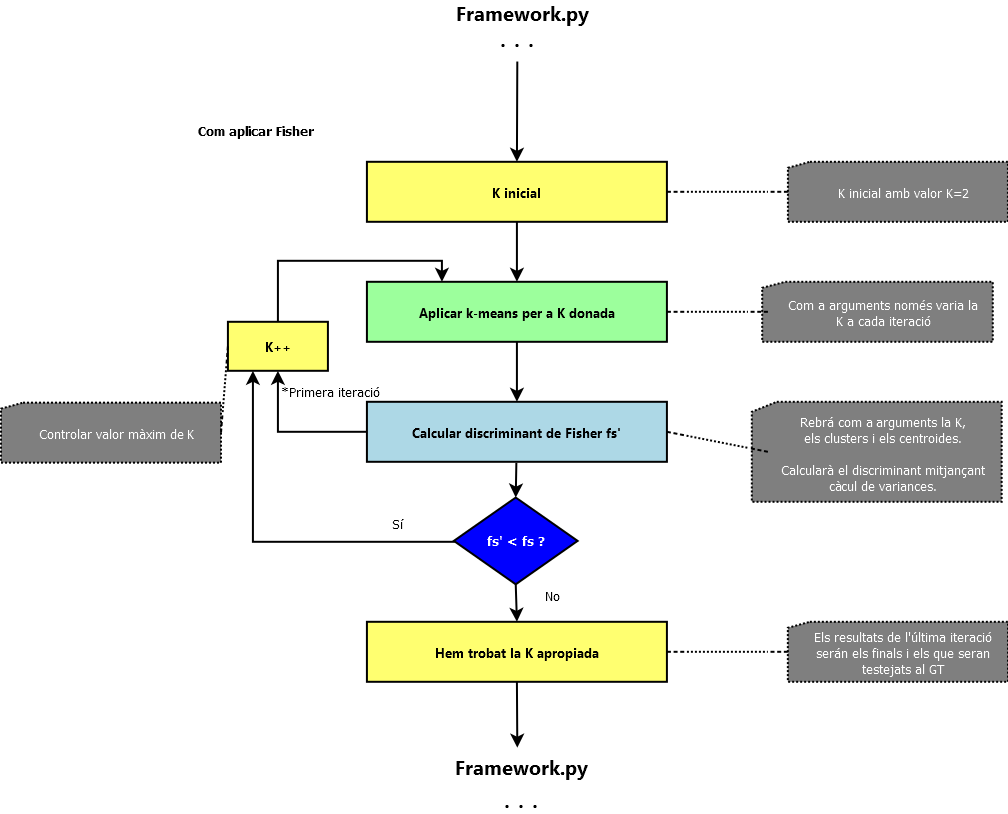
**2.2. Test de Fisher**

A la pràctica implementem el càlcul del discriminant de Fisher que ens permetrà determinar el nombre òptim de clústers/etiquetes a la nostra imatge.

El càlcul del discriminant de Fisher és el següent (\*):



*(\*)Extret del material proporcionat pel professorat de pràctiques*

S’ha realitzat la seva implementació per separat de l’arxiu principal de la pràctica. El seu encaix a l’estructura principal del programa queda detallat al següent diagrama de flux:

Com es pot observar en el diagrama, l’aplicació de l’algorisme de Fisher consistirà en calcular el discriminant de Fisher per a una K donada, mitjançant el càlcul de variàncies a partir dels resultats obtinguts de l’algorisme K-Means.

La tendència del discriminant de Fisher a mesura que vagi augmentant K serà de disminuir, quan aquest pugi, voldrà dir que no té més sentit seguir afegint clústers, ja que no seria el més apropiat donada la nostra distribució de les dades.

**2.3. Altres algorismes**

**Algorisme de selecció de les llavors (Seeds o Centroides inicials)**

Segons el valor que atorguem al paràmetre “seedSelection”, l’algorisme K-Means inicial rebrà uns centroides inicials diferents:

* seedSelection = ‘fixed’ 🡪 l’algorisme rebrà fixats els punts inicials a colors bàsics com el blanc, negre, vermell, gris, etc. A partir de K=5 s’assignaran punts aleatòriament.

Amb aquest mètode hem obtingut millors resultats a l’hora de detectar colors bàsics com el blanc o el negre, però no han representat una millora respecte l’alternativa aleatòria.

* seedSelection = ‘random’ 🡪 l’algorisme rebrà K centroides aleatoris extrets a partir de la pròpia mostra. El fet d’extreure’ls de la pròpia mostra va millorar lleugerament el rendiment de l’algorisme K-means però aquesta millora no va suposar més d’un 5% respecte la versió anterior, que es limitava a, segons l’espai de color, escollir punts aleatoris dins dels marges de l’espai.

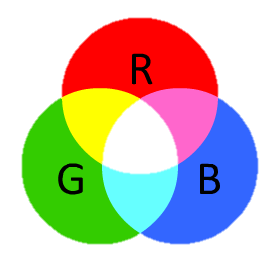
**Mode RGB/HSV + Potentials**

Seguint les indicacions del professor de pràctiques, vam emprar l’espai d’onze dimensions per tal de determinar el color d’un clúster.

Per a fer-ho els passos que fem servir son els següents:

1. S’aplica K-Means i Fisher per a un espai de colors entre a escollir (RGB, HSV).
2. Un cop obtinguts els clústers i els centroides en 3 dimensions, es busca el punt més proper dins de cada clúster al seu centroide.
3. Un cop s’han trobat aquests punts, es busca la seva coordenada a la imatge original, és a dir, el número de píxel.
4. Un cop obtingut el número de píxel, obtenim la imatge en onze dimensions i procedim a crear els nous centroides d’onze dimensions que es troben a aquella “coordenada”.
5. Enviem els nous centroides a l’avaluador, havent canviat els paràmetre d’aquest anteriorment.

**4. Resultats**

**Espai de colors RGB**

Es tracta de l’espai de colors amb que millors resultats hem obtingut. És un espai de 3 dimensions on cada dimensió representa la quantitat de cada color específic: la primera coordenada pel color vermell (Red, en anglès), la segona coordenada pel color verd (Green) i la tercera i última coordenada pel color blau (Blue).

És l’espai de colors més popular. Cada coordenada pot prendre un valor de 0 fins a 255.

Resultats per a 10 execucions:

|  |  |  |  |
| --- | --- | --- | --- |
| **#Execució** | **lablThrs** | **ScaleFactor** | **%Accuracy** |
| 1 | pi/32 | 0,3 | 39,20% |
| 2 | pi/32 | 0,3 | 38,98% |
| 3 | pi/32 | 0,3 | 41,34% |
| 4 | pi/32 | 0,3 | 38,76% |
| 5 | pi/32 | 0,3 | 39,20% |
| 6 | pi/32 | 0,3 | 29,32% |
| 7 | pi/32 | 0,3 | 40,53% |
| 8 | pi/32 | 0,3 | 42,23% |
| 9 | pi/32 | 0,3 | 38,98% |
| 10 | pi/32 | 0,3 | 42,12% |

Com podem observar l’algorisme té variacions a cada execució però obtenim aproximadament un 40% d’encert.

**RGB/HSV + POTENTIALS**

Potentials es un espai de 11 dimensions on cada dimensió representa la probabilitat d’un color determinat. El procediment fet servir per a aplicar aquesta combinació d’espais de colors és el següent:

1. S’aplica K-Means i Fisher per a un espai de colors entre a escollir (RGB, HSV).
2. Un cop obtinguts els clústers i els centroides en 3 dimensions, es busca el punt més proper dins de cada clúster al seu centroide.
3. Un cop s’han trobat aquests punts, es busca la seva coordenada a la imatge original, és a dir, el número de píxel.
4. Un cop obtingut el número de píxel, obtenim la imatge en onze dimensions i procedim a crear els nous centroides d’onze dimensions que es troben a aquella “coordenada”.
5. Enviem els nous centroides a l’avaluador, havent canviat els paràmetre d’aquest anteriorment.

Aquest procediment es troba dins l’arxiu framework.py i per a fer funcionar aquest mode només s’ha de posar la variable “combined\_Potentials” a *True*.

Resultats per a 10 execucions:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **#Execució** | **Espai de color** | **lablThrs** | **ScaleFactor** | **%Accuracy** |
| 1 | **RGB + POTENTIALS** | pi/32 | 0,15 | 31,95% |
| 2 | **RGB + POTENTIALS** | pi/32 | 0,15 | 31,39% |
| 3 | **RGB + POTENTIALS** | pi/32 | 0,15 | 30,17% |
| 4 | **RGB + POTENTIALS** | pi/32 | 0,15 | 29,06% |
| 5 | **RGB + POTENTIALS** | pi/32 | 0,15 | 32,95% |
| 6 | **HSV + POTENTIALS** | pi/32 | 0,15 | 31,28% |
| 7 | **HSV + POTENTIALS** | pi/32 | 0,15 | 28,84% |
| 8 | **HSV + POTENTIALS** | pi/32 | 0,15 | 29,39% |
| 9 | **HSV + POTENTIALS** | pi/32 | 0,15 | 29,17% |
| 10 | **HSV + POTENTIALS** | pi/32 | 0,15 | 29,84% |

**4. Conculsions**

En aquesta pràctica hem aprofundit el nostre coneixement sobre l’àmbit de l’aprenentatge computacional no supervisat.

Concretament hem aprofundit en l’algorisme K-Means i l’hem reforçat amb estadístics com el test de Fisher per a determinar el nombre adient de clústers a partir dels càlculs de les variàncies dels resultats obtinguts al aplicar K-Means.

Ens hagués agradat poder obtenir millors resultats, però ens ha resultat impossible de moment. La col·locació dels centroides inicials juga un gran paper en l’algorisme K-Means i s’hauria de mirar d’optimitzar la col·locació d’aquests punts centrals per tal d’obtenir millores realment significatives.

El rendiment de l’espai RGB ha sigut el millor amb diferència. Tot i que la combinació dels espais RGB/HSV i Potentials no està tant lluny d’obtenir bons resultats i es podrien buscar alternatives per a millorar-lo.

Havent provat algorismes ja implementats de K-Means, ISOSDATA (per determinar la K), etc. no hem obtingut molt millors resultats, però no dir pitjors en alguns casos.

Si bé es veritat que potser no hem configurat correctament els algorismes, però al intentar-ho tampoc no hem aconseguit obtenir-ne gran millora.

**5. Bibliografia**

**Apunts, recomanacions, problemes resolts a classe.**

**Reconeixement de patrons**

http://decsai.ugr.es/~castro/Neuro/B-IntroalRecdePatrones/ppframe.htm

**Algorisme ISODATA**

http://web.pdx.edu/~jduh/courses/Archive/geog481w07/Students/Vassilaros\_ISODATA.pdf

**KMeans de la llibreria sklearn**

http://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html

**Determinar el nombre de clusters en un conjunt de dades**

http://en.wikipedia.org/wiki/Determining\_the\_number\_of\_clusters\_in\_a\_data\_set

**KMeans de la llibreria Scipy**

http://docs.scipy.org/doc/scipy/reference/generated/scipy.cluster.vq.kmeans.html#scipy.cluster.vq.kmeans