

PROJECTE 2: **Etiquetador**

Intel·ligència Artificial

Universitat Autònoma de Barcelona

Objectiu: Fer un agent que faci **etiquetatge automàtic d'imatges** que permeti fer cerques intel·ligents en llenguatge natural en una botiga on-line que actualitza el llistat de productes constantment.

El vostre sistema hauria d'assignar dos tipus d'etiquetes: color i forma del producte, per a que els usuaris puguin fer cerques en un llenguatge molt directe, com ara: "red shirt" o "Black sandals"



Com que pot ser molt complex!!! \rightarrow Farem algunes simplificacions

Simplificacions:

- Les etiquetes seran en <u>anglès</u>
- Etiquetarem només 8 tipus de roba:
 - ✓ Dresses
 - ✓ Flip Flops
 - **√** Jeans

 - **√** Sandals

- ✓ Shirts
- ✓ Shorts
- ✓ Socks
- **√** Handbags



• Etiquetarem els colors predominants de cada peça de roba només amb els 11 color bàsics universals:

✓ Red

- **√** Green
- **✓** Black

- **√** Orange
- **√** Blue
- **√** Grev

- √ Brown
- ✓ Purple
- √ White

- ✓ Yellow
- ✓ Pink

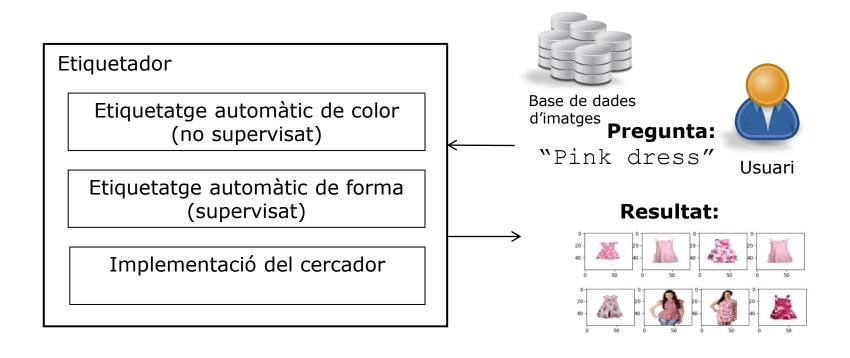


• Per reduir el temps d'execució de tots els algorismes treballarem amb imatges de baixa resolució (60x80 píxels). Farem servir la base de dades: Fashion Product Images Dataset del Kaggle

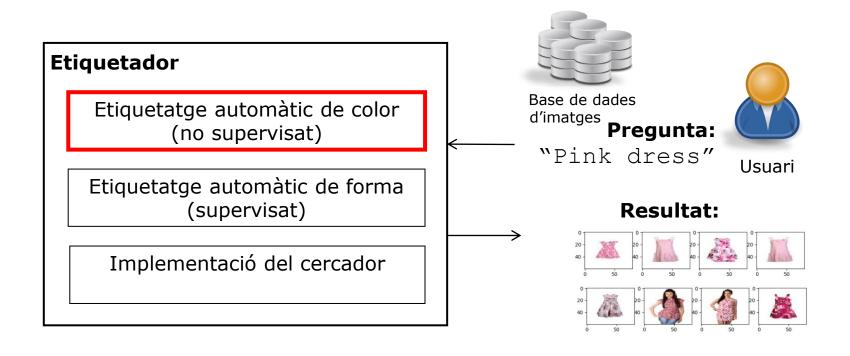
https://www.kaggle.com/paramaggarwal/fashion-product-images-dataset

Kaggle és una repositori de bases de dades que es comparteixen per recerca en Data-Science

Problemes a resoldre per fer aquest etiquetador:



Problemes a resoldre per fer aquest Etiquetador:



Etiquetador Etiquetatge automàtic de color (no supervisat) Etiquetatge automàtic de forma (supervisat) Etiquetatge automàtic de forma (supervisat) Implementació del cercador

Com podem resoldre el problema d'etiquetar automàticament el color?



Etiquetes de colors predominants:

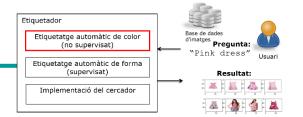
Yellow, Orange, Blue, Black, Green, White

3 Qüestions:

Com es representa el color?



- Com podem trobar els colors predominants d'una imatge?
- Com podem assignar un nom als colors predominants?

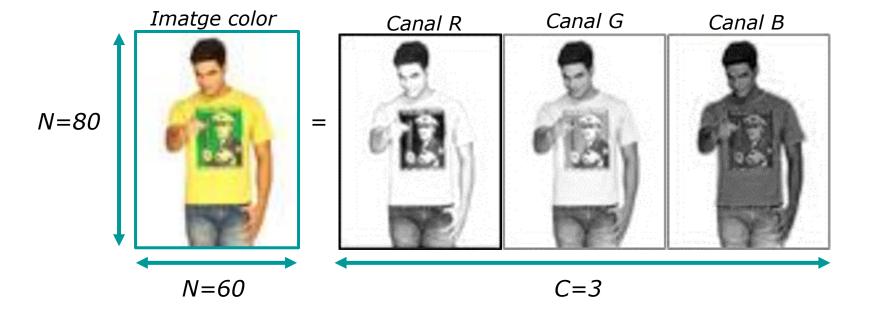


Com es representa el color?

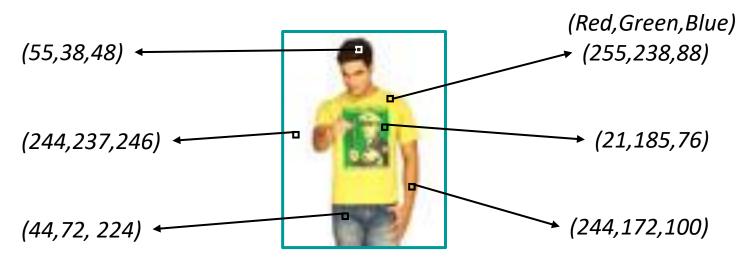
La resposta està vinculada a com es representa una imatge?

Una imatge en color és una matriu de dimensions: N x M x C

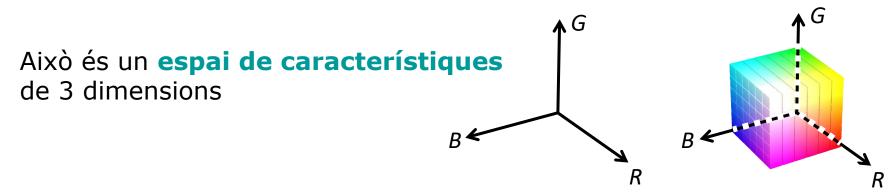
Exemple: Imatge en color 80x60x3 (files x columnes x canals) Imatge en nivell de gris 80x60x1



Mirem-ho a nivell de cada píxel:



Tornem a la pregunta inicial: Com es representa el color?



Etiquetador Etiquetatge automàtic de color (no supervisat) Etiquetatge automàtic de forma (supervisat) Etiquetatge automàtic de forma (supervisat) Implementació del cercador

Com podem resoldre el problema d'etiquetar automàticament el color?



Etiquetes de colors predominants:

Yellow, Orange, Blue, Black, Green, White

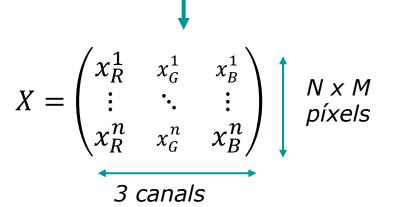
3 Qüestions:

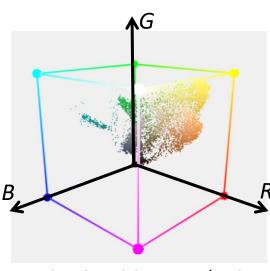
- Com es representa el color? ${f V}$
- Com podem trobar els colors predominants d'una imatge?
- Com podem assignar un nom als colors predominants?

Com podem trobar els colors predominants d'una imatge?



Imatge en color $(N \times M \times 3)$ Núm. de píxels = files x columnes



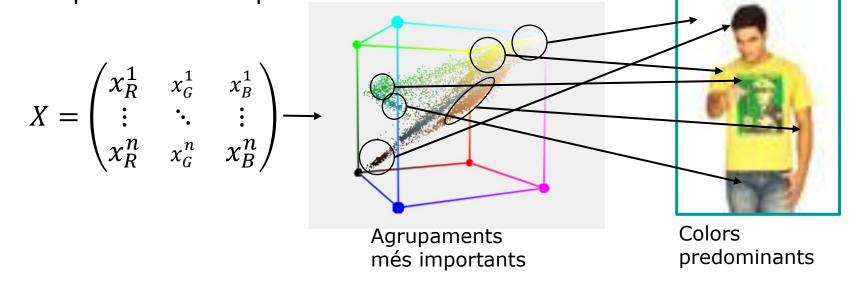


Els colors dels punts són els colors de l'RGB de cada píxel



Com podem trobar els colors predominants?

Objectiu: Tenim una mostra de punts en un espai de tres dimensions i <u>necessitem trobar els agrupaments més importants</u> d'aquest núvol de punts



Solució: Agrupament no supervisat de punts

Com ho fem? Algorisme k-means

En aquest projecte l'algorisme K-means el treballarem a

Fitxer: Kmeans.py

Classe: Kmeans

Entrades de la classe Kmeans

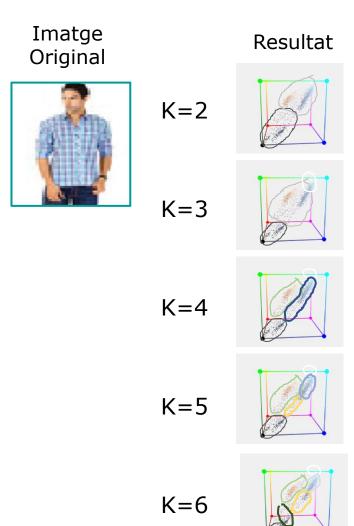
- X: Imatge que volem analitzar.
- K: Numero de clústers que utilitzarem
- options: Opcions addicionals (mètode d'inicialització de centroides, màxim numero d'iteracions,...)

Primer s'inicialitzaran totes les variables necessàries quan es cridi: Kmeans (X, K=3, options=None)

I finalment s'aplicarà l'algorisme que iterarà fins a convergir Kmeans.fit()

es guardaran els centroides obtinguts a la variable centroids

Exemple: aplicació del K-means per diferents valors de K





Nota: Aquesta visualització es fa amb la funció **visualize k means()**

Problema del k-means: Quina k és la millor?

A Teoria hem vist com estimar la millor k automàticament:

Es pot estimar una **Mesura sobre la qualitat de la classificació**, i estudiar com varia per a diferents números de classes (k=2, 3, 4, ...).

(Normalment aquest estudi es basa en un anàlisi de la variància de les classes)

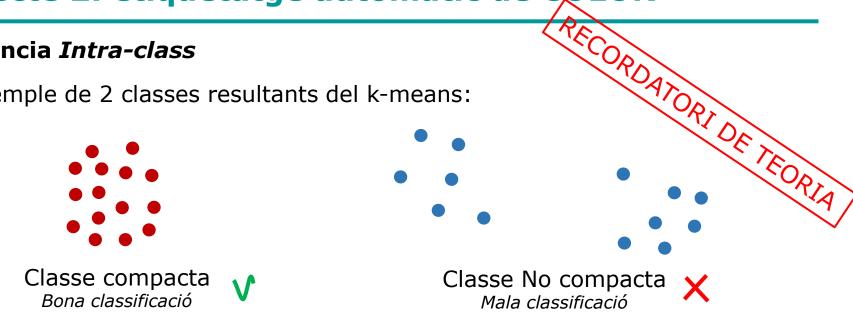
farem servir aquesta!!

Alguns estadístics interessants:

- Distància Intra-class
- Distància Inter-class
- Discriminant de Fisher

Distància Intra-class

Exemple de 2 classes resultants del k-means:

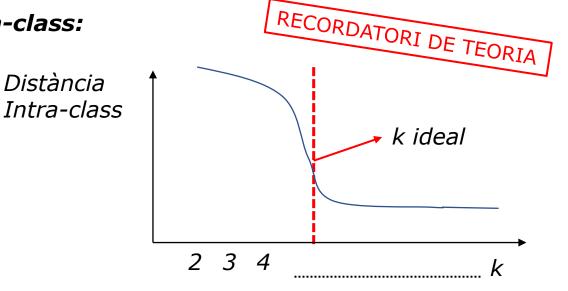


Estimació: Suma per a totes les classes de la mitjana de les distàncies entre tots els parells de punts d'una classe

$$D(C) = \frac{2}{m(m-1)} \sum_{j=1}^{m} \sum_{i=j+1}^{m} d(\vec{x}^i, \vec{x}^j) : \vec{x}^i, \vec{x}^j \in C, i, j: 1...m$$

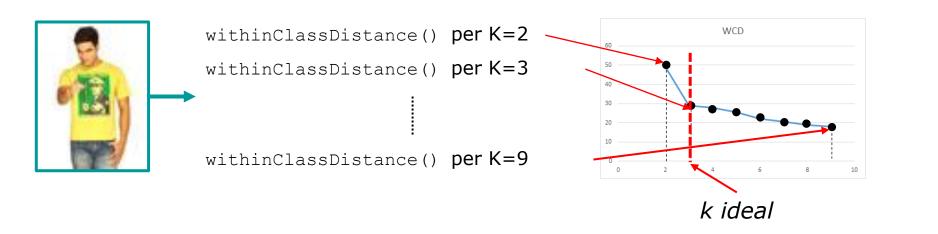
$$\sum_{i=1}^{k} D(C_i) \longrightarrow \text{ és bo que sigui petita !!}$$

Estudi de la Distància Intra-class:



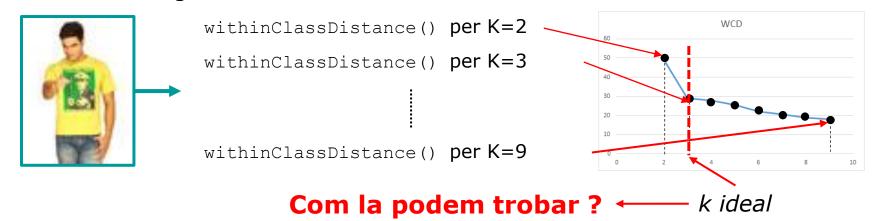
En el nostre cas:

Donada una imatge



Estudi de la Distància Intra-class:

Donada una imatge



Podem calcular el % de Decrement:

$$\%DEC = 100 \frac{WCD_k}{WCD_{k-1}}$$

Un llindar possible és agafar la k a a partir de la qual

$$100 - \%DEC < 20\%$$
 (exemple)

	K	WCD	%DEC	10	00-%DE	С	
ļ	2	49.09					
ı	3	20 11	50 20		40.71		→ k ideal
L	5	29.11	39.29		40.71		7. 7.4.04.7
Ī	4	27.95	96.03		3.97		
	5	25.68	91.86		8.14		
	6	22.00	85.70		14.30		200/
	7	20.61	93.65		6.35		→ < 20%
	8	18.82	91.31		8.69		
	9	18.09	96.15		3.85		

Problema: Quina k és la millor?

Per a calcular distancia intra-class programareu la funció:

whitinClassDistance()

Entrada: self

Retorna: valor wcD

Per a seleccionar la millor k programareu la funció:

find_bestK()

Entrada: self, max K

Actualitza: K

Etiquetador Etiquetatge automàtic de color (no supervisat) "Pink dress" Etiquetatge automàtic de forma Implementació del cercador

Com podem resoldre el problema d'etiquetar automàticament el color?



Etiquetes de colors predominants:

Yellow, Orange, Blue, Black, Green, White

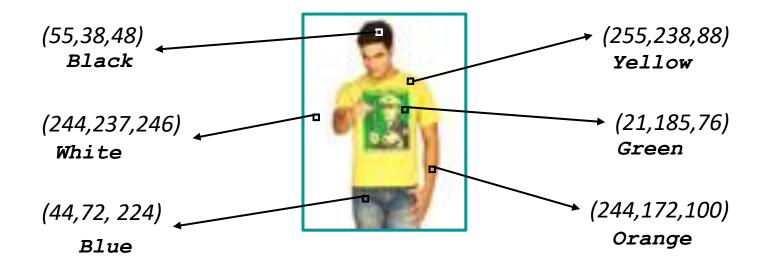
3 Qüestions:

- Com es representa el color? V
- Com podem trobar els colors predominants d'una imatge?
- Com podem assignar un nom als colors predominants?





Com podem assignar un nom als colors predominants?



Aquest problema requereix simular com els humans perceben el color !!!

Aquest problema ja ha estat resolt de manera multidisciplinar:

Experiments en **Antropologia**

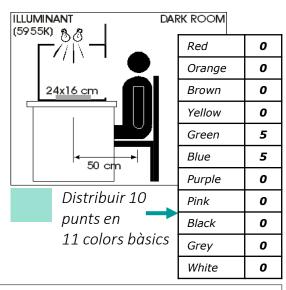
+

Experiments en **Psicologia Experimental**

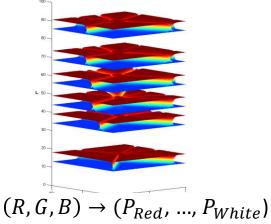
+

Models Matemàtics en **Visió Artificial**

Estudis sobre 78 llengües ha demostrat que existeixen 11 noms de color bàsics universals compartits per les llengües més evolucionades



R. Benavente, M. Vanrell, R. Baldrich (2006) A dataset for fuzzy color naming, Color Research and Applications



Codi disponible a:

http://www.cvc.uab.cat/colour_naming

R. Benavente, M. Vanrell, R. Baldrich (2008)

Parametric fuzzy sets for automatic color naming, Journal of the OSA.

Berlin, B., & Kay, P. (1991) Basic color terms: Their universality and evolution. Univ of California Press.

Nosaltres usarem aquest codi!!!

Com podem assignar un nom als colors predominants?

Fent servir els resultats dels treballs anteriors, passarem de l'espai RGB a l'espai dels 11 noms de colors:

$$(R,G,B) \rightarrow (P_{Red},P_{Orange},P_{Brown},P_{Yellow},P_{Green},P_{Blue},P_{Purple},P_{Pink},P_{Black},P_{Grey},P_{White})$$

per a cada RGB retorna un vector d'11 probabilitats de que un humà assigni cada un dels noms de color.

La funció que fa aquesta conversió us la donem programada:

Funció: get_color_prob()

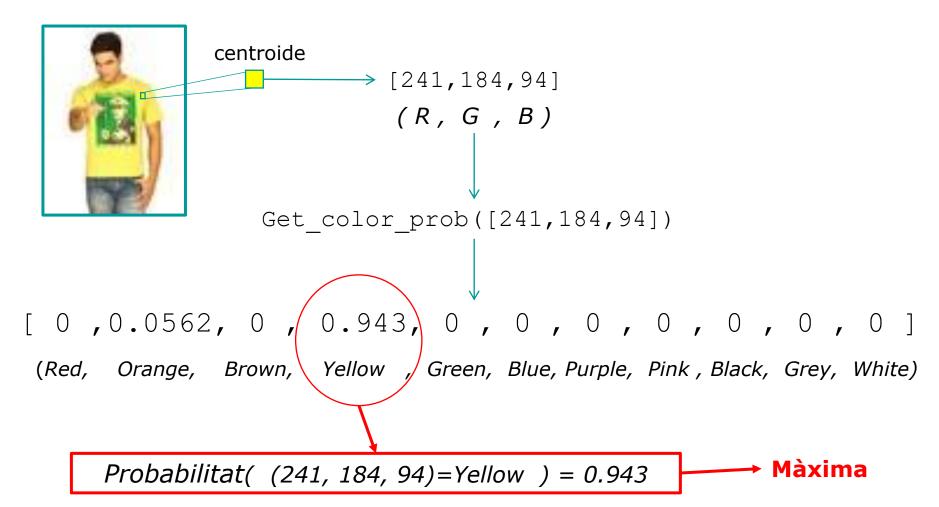
Fitxer: utils.py

Per assignar etiquetes a tots els colors predominants programareu:

Funció: get_color()

Fitxer: kmeans.py

Exemple:



Exemple: aplicació de les etiquetes per diferents resultats

Imatge Original



Resultat



K=3

K=2

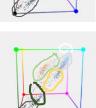


K=4



K=6

K=5



Píxels assignats a centroides









Etiquetes dels centroides

[Gris, Gris]

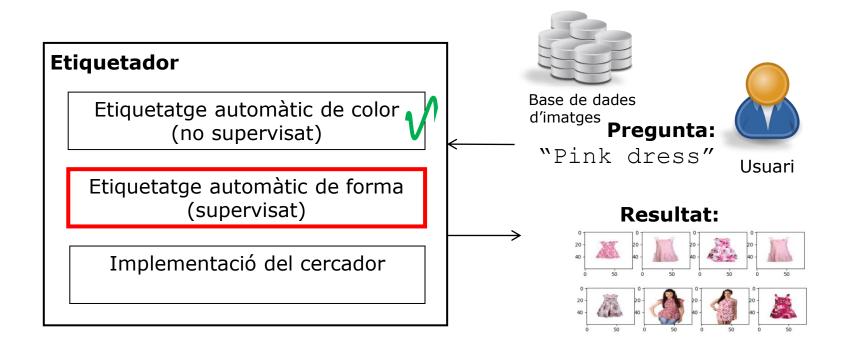
[Blanc, Gris, Negre]

[Blanc, Blau, Taronja, Negre]

[Blanc, Blau, Lila Taronja, Negre]

[Blanc, Blau, Lila, Marro, Taronja, Negre]

Problemes a resoldre per fer aquest Etiquetador:



Com podem resoldre el problema d'etiquetar automàticament la roba?



Etiqueta de forma:

Shirt



3 Qüestions:

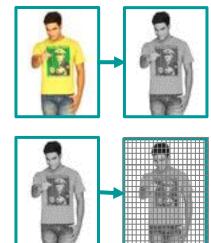
- Com podem representar la forma de la roba?
- Com podem aprendre a classificar la roba?
- Com assignem l'etiqueta de tipus de roba a una nova imatge?

Com podem representar la forma de la roba?

Quin **espai de característiques** podríem fer servir per representar la forma de la roba?

Aquest és un **problema de visió per computador** que com que encara no en sabem prou ho resoldrem de manera molt simple de la següent manera:

- 1) Eliminarem el COLOR, ja que no el necessitem per representar la forma
- 2) Agafarem directament els píxels de la imatges com la característica de cada posició de la imatge.



Com podem representar la forma de la roba?

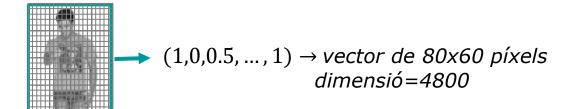
Extracció de les característiques d'una imatge

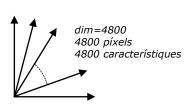
1) Eliminarem el COLOR, ja que no el necessitem per representar la forma

$$(R,G,B) \rightarrow \left(\frac{R+G+B}{3}, \frac{R+G+B}{3}, \frac{R+G+B}{3}\right)$$



2) Agafarem directament els píxels de la imatges com la característica de cada posició de la imatge.





Com podem resoldre el problema d'etiquetar automàticament la roba?



Etiqueta de forma:

Shirt



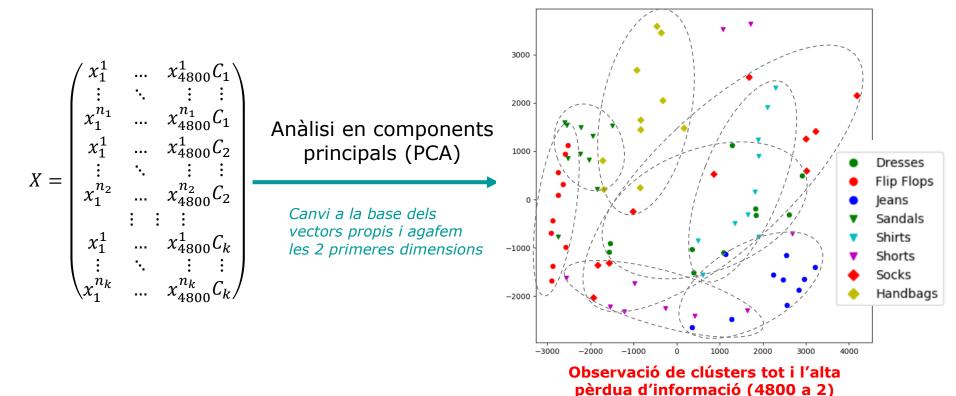
3 Qüestions:

- Com podem representar la forma de la roba? \(
 \bar{\text{l}}\)
- Com podem aprendre a classificar la roba?
- Com assignem l'etiqueta de tipus de roba a una nova imatge?



Com podem aprendre a classificar la roba?

Donada la mostra que farem servir com a conjunt d'aprenentatge podem intentar visualitzar aquest espai de 4800 dimensions a un espai observable de 2 dimensions:



Com podem aprendre a classificar la roba?

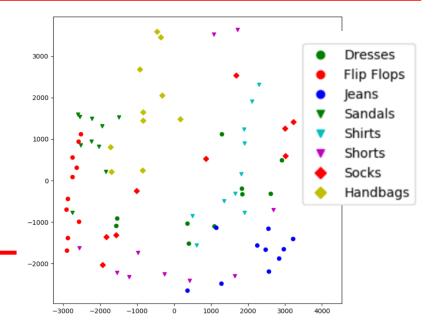
A Teoria hem vist diferents dues famílies de classificadors:

- Classificador lineal
- Classificador no lineal
- Classificador probabilístic

Quan les dades presenten un model clar (lineal, no-lineal, probabilístic, ...)

 Classificador del k-veí més proper (KNN)

Quan no existeix un model clar



Com podem resoldre el problema d'etiquetar automàticament la roba?



Etiqueta de forma:

Shirt



3 Qüestions:

Com podem representar la forma de la roba?



Com podem aprendre a classificar la roba?



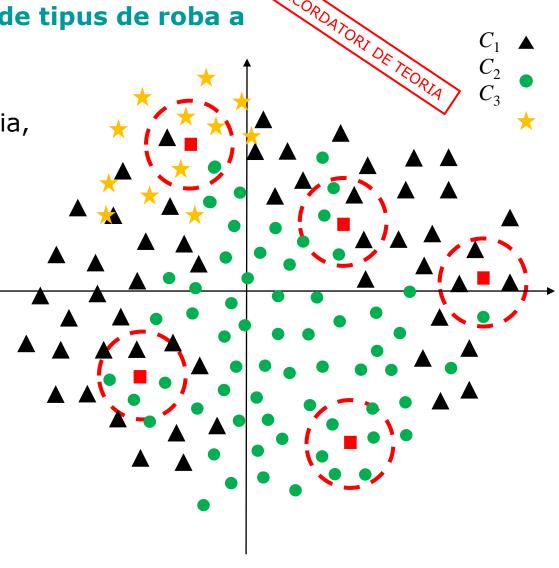
Com assignem l'etiqueta de tipus de roba a 🖊 una nova imatge?



Com assignem l'etiqueta de tipus de roba a una nova imatge?

Algorisme K-NN vist a teoria,

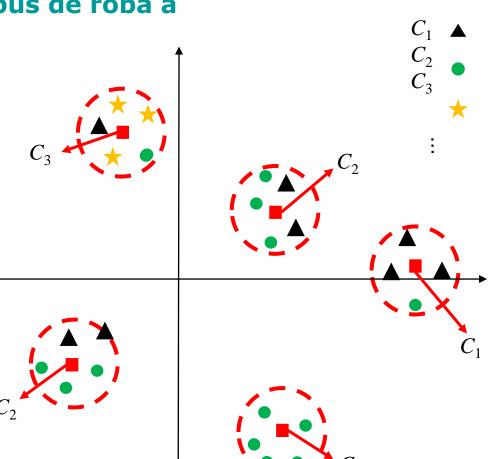
Idea: Es basa la decisió en els veïns més propers, considerant a quina classe pertanyen els N-veïns més propers.



Com assignem l'etiqueta de tipus de roba a una nova imatge?

Algorisme K-NN vist a teoria,

Idea: Es basa la decisió en els veïns més propers, considerant a quina classe pertanyen els N-veïns més propers.



Com assignem l'etiqueta de tipus de roba a una nova imatge?

Algorisme K-NN

Funció de decisió

(per classificar \overrightarrow{y})

$$Per\left(\vec{x}^{j} \in X\right) fer$$

$$Llista = inserir(\left[d(\vec{y}, \vec{x}^{j}), C_{j}\right], Llista)$$

$$fPer$$

$$Veins = Primers_k(ordenar_d(Llista))$$

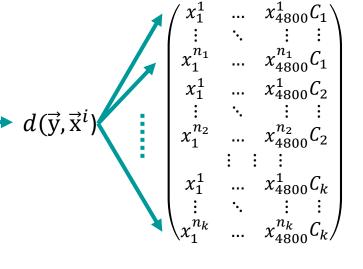
$$Si\left(comptar(Veins, C_{1}) > comptar(Veins, C_{2})\right)$$

$$\vec{y} \in C_{1}$$

$$Sin\acute{o}$$

$$\vec{y} \in C_{2}$$

$$fSi$$



Exemple:

$$\vec{y} = [1,1,1,0.2,0.5,0,0,1,...,1]$$



$$d=23$$

$$d=59 \rightarrow [0,0,0,0.7,0.5,1,0,1,...,1]$$

[1,1,1,0.1,0.4,0.2,1,1,...

$$d=103$$
 [0,0,0,0.2,1,1,0,0.4,... ,0]





Com podem resoldre el problema d'etiquetar automàticament la roba?



Etiqueta de forma:

Shirt



3 Qüestions:

- Com podem representar la forma de la roba?
- Com podem aprendre a classificar la roba?
- Com assignem l'etiqueta de tipus de roba a una nova imatge?







La resposta a les 3 questions en el codi:

Com podem representar la forma de la roba?

Funció: read_dataset()

Directori: images/train

Fitxer: utils data.py

Com podem aprendre a classificar la roba?

Funció: KNN. init ()

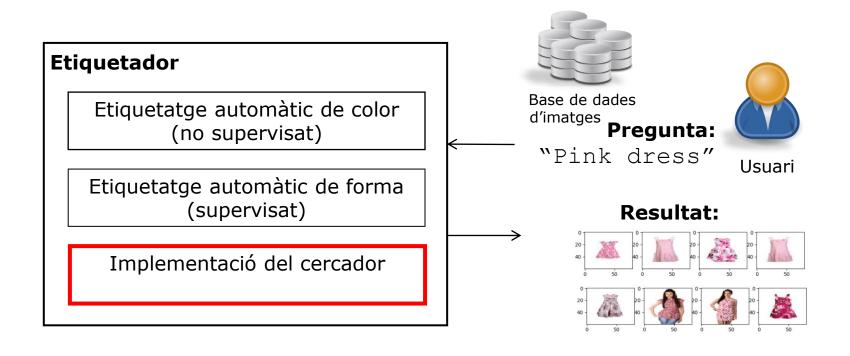
Fitxer: KNN.py

 Com assignem l'etiqueta de tipus de roba a una nova imatge?

Funció: KNN.predict()

Fitxer: KNN.py

Problemes a resoldre per fer aquest Etiquetador:



Projecte 2: Cercador d'imatges

Com implementem un cercador basat en etiquetes de color i forma?

Ja tenim imatges etiquetades amb COLOR i FORMA,

Per fer les cerques amb etiquetes programareu les funcions:

```
Retrieve_img_by_color()
Retrieve_img_by_class()
Retrieve_combine()
```

Fitxer: my_labeling.py

Planificació

Part 1: PROGRAMACIÓ Kmeans i color

Sessió Virtual de seguiment: Setmana del 27 de Març

Entrega: Què? Exercicis indicats al guió de la Part 1 (GuiaP2 Part1.pdf)

Quan? Abans del Dilluns 10 d'Abril a les 23:55h.

Part 2: PROGRAMACIÓ kNN i forma

Sessió Virtual de seguiment: Setmana del 17 d'Abril

Entrega: Què? Exercicis indicats al guió de la Part 2 (GuiaP2 Part2.pdf)

Quan? Abans del **Dimarts 2 de Maig** a les 23:55h.

Part 3: Anàlisi del Rendiment

Sessió Virtual de seguiment: Setmana del 8 de Maig

Entrega: Què? Informe indicat al guió de la Part 3 (GuiaP2_Part3.pdf)

Quan? Abans del Dijous 18 de Maig a les 23:55h.

Presentació ORAL, explicació de tot el projecte

Sessió Virtual de seguiment: Setmana del 22 de Maig

Entrega: Què? Diapositives de la presentació

Quan? abans del **Dijous 1 de Juny** a les 23:55h.

Consells pràctics per a la primera entrega:

- Trobareu els exercicis al fitxer <GuiaP2_Part1.pdf> al cv.uab.cat Secció>Pràctiques>Projecte2. Aquest document us va guiant en tot el que heu d'anar programant.
- Guardareu totes les funcions al fitxer <Kmeans.py>
- Programeu les funcions tal i com s'especifiquen quant als paràmetres d'entrada i el que retorna cada funció.
- L'entrega es farà via CV, entregareu un fitxer que conté totes les funcions que es treballen a la Part 1.
- Aconsellem assistir a la sessió virtual de pràctiques amb els exercicis pràcticament resolts per poder aprofitar millor la sessió i poder comentar dubtes amb el professor de pràctiques o per poder avançar amb la pràctica.

Consells pràctics per a la segona entrega:

- Trobareu els exercicis al fitxer <GuiaP2_Part2.pdf> al cv.uab.cat
 Secció>Pràctiques>Projecte2. Aquest document us va guiant en tot el que heu d'anar programant.
- Guardareu totes les funcions al fitxer <knn.py>
- L'entrega es farà via CV, entregareu un fitxer que conté totes les funcions que es treballen a la Part 2.
- Aconsellem assistir a la sessió virtual de pràctiques amb els exercicis pràcticament resolts per poder aprofitar millor la sessió i poder comentar dubtes amb el professor de pràctiques o per poder avançar amb la pràctica.

Consells pràctics per a la 3a entrega:

- Trobareu els exercicis al fitxer <GuiaP2_Part3.pdf> al cv.uab.cat Secció>Pràctiques>Projecte2. Aquest document us va guiant ent tot el que heu d'anar programant.
- Guardareu totes les funcions al fitxer <my_labeling.py>
- Entregareu al cv.uab.cat, els fitxers Kmeans.py, KNN.py, i
 my_labeling.py amb totes les funcions de tot el projecte, i <u>l'informe</u>
 on explicareu els anàlisis que heu fet i tots els resultats
- Aconsellem assistir a la sessió virtual de pràctiques amb els exercicis pràcticament resolts per poder aprofitar millor la sessió i poder comentar dubtes amb el professor de pràctiques o per poder avançar amb la pràctica.

Consells pràctics per preparar l'Informe i la Presentació Oral:

Ambdós s'haurien d'organitzar de la següent manera:

- Introducció (llista de continguts i resum del que s'ha fet diferent)
- 3 anàlisis com a mínim d'alguna de les parts opcionals que es mencionen a la guia 3. Un anàlisi hauria de contenir el següent:
 - Breu introducció sobre quin paràmetre fem l'anàlisi
 - Comparació del resultat original amb els nous resultats
 - Explicació dels resultats (Per que funciona millor? per que funciona pitjor? es més eficient? Podríeu trobar casos on un mètode funcioni millor que l'altre?

Conclusió

Principals problemes trobats, què heu après?, què milloraríeu?

Possibles paràmetres d'anàlisi:

- Mètode d'inicialització de centroides
- Utilitzar un espai de colors diferent a l'RGB
- Utilitzar diferents valors de K
- Utilitzar diferents metodes per trobar la millor K (Interclass variance, Fisher,...)
- Utilitzar altres mètodes d'etiquetatge de color (Múltiples etiquetes, introduir nous colors,...)
- Utilitzar diferents caràcteristiques per al KNN → Diferents tamanys d'imatge
- Utilitzar diferents caràcteristiques per al KNN → Caracteristiques calculades per separat (Valor mitja de tots els pixels, pixels dreta vs esquerra, superiors vs inferiors,...)

- ...

Avaluació:

Nota Projecte 2 = 0.75 * Nota Grup + 0.25 * Nota Individual

- ❖ Nota Grup = 0.6 * Nota Codi + 0.3 * Informe + 0.1 * Presentació Grup
- ❖ Nota Individual = 0.5 * Presentació Individual + 0.5 * Participació Grupal