

SOCIAL EVENT DETECTION

GRUP 3.2

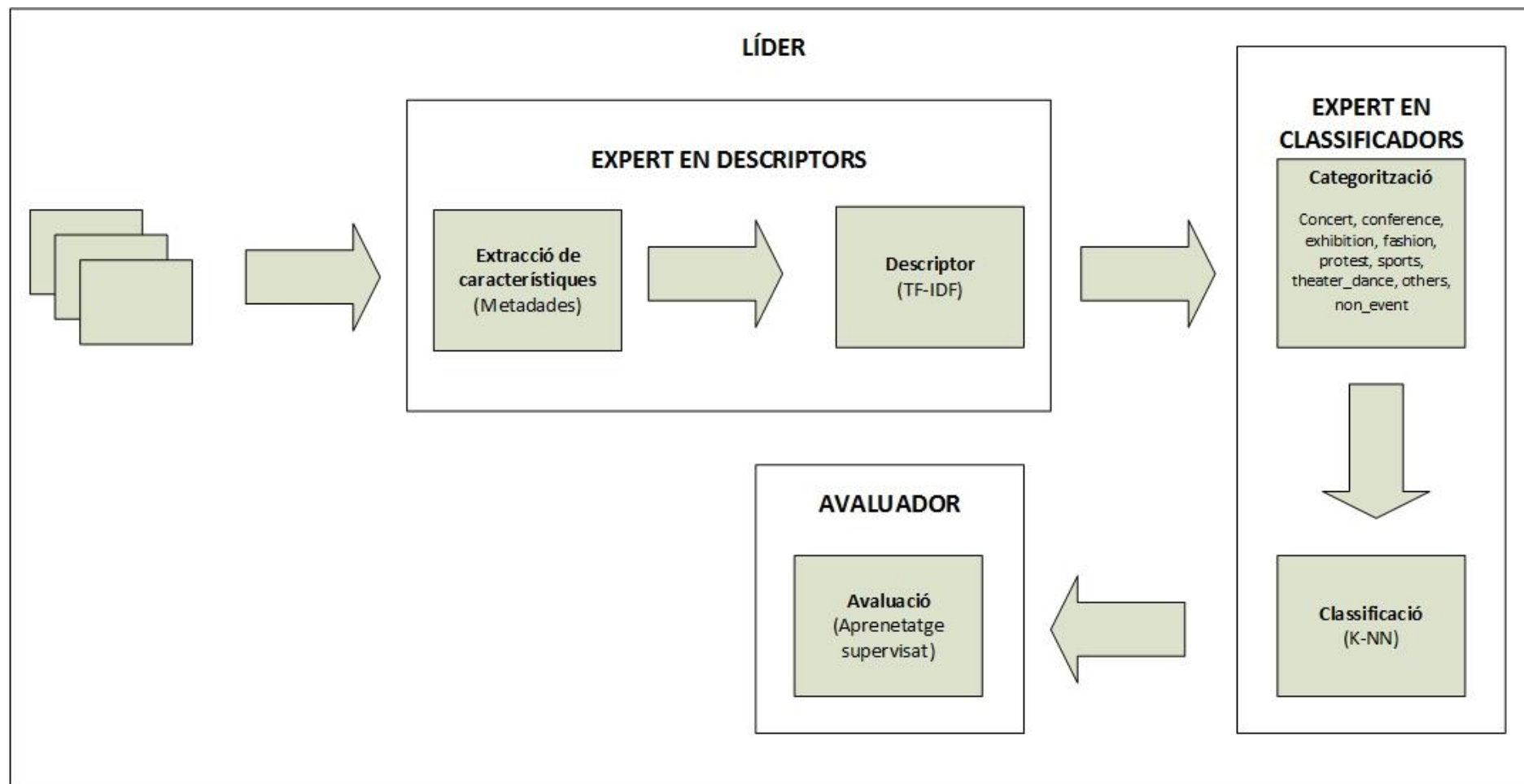
Antonio J. Urbano Torres

Sergio Tomás Fernández

Enric Carbonell Fages

Xavier Serrat Bordas

Diagrama de blocs



Reflexió de les decisions presses

- **Descriptor de text: TF-IDF**

Dóna importància a les paraules amb poques repeticions → Paraules clau → Millor identificació dels títols i de continguts

- **Classificador: Classificació per màxims**

Algoritme simple. Es busca el màxim dels hashtags i una vegada trobat, s'assigna la classe.

- **Repositori de codi i dades compartides: GitHub**

Pràctic a l'hora de fer-lo funcionar i facilitat de compartir el codi amb versions.

Llenguatges de programació: MATLAB

1. Conegut per tots els membres del grup.
2. Compatible amb les biblioteques de OpenCV i Weka.
3. Molta informació de les funcions
4. Llenguatge àgil d'utilitzar

Descripció

1. Carregar les metadates de les fotografies (2815): ID, títol, classificació.
2. Escollir el format de SQL per extreure metadates.
3. Classificar en categories els títols i aplicar TF-IDF per a cada categoria.

Classificació

1. Busquem el màxim hashtag → És el que té més pes dins cada categoria.
2. Una vegada trobat el màxim, es classifica la foto en aquella categoria.
3. Trobem 9 categories: concert, conference, exhibition, fashion, non-event, other, protest, sports, theatre dance

Avaluació

1. Importar arxiu .csv que conté la veritat terreny (27000 imatges).
2. Importar l'arxiu .txt dels resultats del classificador.
3. A una funció externa que es diu `getcm` li passem aquests resultats i ens diu la matriu de confusió, el número de correctes, precisió, record i F-score.

Llicències de programari lliure

GNU General Public License → utilitzar, estudiar, compartir i modificar software.

Referències

1. www.github.com
2. www.mathworks.es/discovery/matlab-opencv.html
3. ww.mathworks.com/matlabcentral/fileexchange/21204-matlab-weka-interface
4. www.mathworks.es
5. www.multimediaeval.org/mediaeval2013/sed2013