

# SOCIAL EVENT DETECTOR (v2)

---

GDSA14 – 3.1

Eduardo Bernal · Mohamed el Bouchti · Luís Varas · Aitor Niñerola

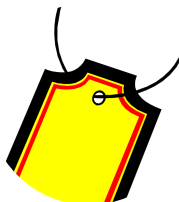


Nova imatge

# ESTRUCTURA



Anàlisi textual



Esdeveniment  
(etiqueta final)



Non Event



Anàlisi Visual



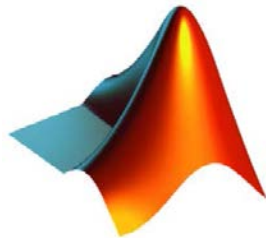
Etiqueta final

# Observacions sobre l'estructura

- Aquesta estructura en cascada es deu a les diferents proves que vam fer amb l'obtenció dels resultats. Amb la classificació només textual obteníem uns resultats i amb la classificació només visual uns altres però si els combinàvem seguint el criteri establert, la puntuació del sistema augmentava lleugerament.
- Val a dir que això depèn dels paràmetres de configuració quan es treballa amb la part visual, ja que una disminució de la informació pot fer reduir el bon funcionament de la classificació.
- **ATENCIÓ:** Amb els paràmetres **finals** establerts respecte als que teníem inicialment (reducció d'informació pel temps de processat i mida de les variables) la classificació visual empitjora lleugerament i, consegüentment, la final combinada.

# LLENGUATGES DE PROGRAMACIÓ

- PROCESSAT TEXTUAL: **C++ i MySQL**
- PROCESSAT VISUAL: **Matlab** juntament amb la llibreria VLFeat
- AVALUACIÓ: **C++** amb consulta a la base de dades MySQL.



# PROCESSAMENT TEXTUAL

- Basat en les cerques en taules amb índex FULLTEXT de MySQL.
- Aquestes cerques ponderen els termes que conté cada document segons la TF-IDF i retorna el document més rellevant de la col·lecció segons la consulta executada.
- S'assigna la classe del PRIMER (KNN=1) document retornat pel motor de la base de dades, és a dir, el que té major puntuació.

**Més informació sobre el funcionament:**

<http://dev.mysql.com/doc/refman/5.7/en/fulltext-search.html>

<http://mysqlserverteam.com/rankings-with-innodb-full-text-search/>

# PROCESSAMENT VISUAL

- **ENTRENAMENT**

- 1) Extracció de característiques vl\_phow (funcions VLFEAT) (*dense sift*)
- 2) Selecció aleatòriament de 4.000 característiques
- 3) Agrupament k-means en 600 clústers
- 4) Creació de KDTrees per a cada classe

- **CLASSIFICACIÓ**

- 1) Extracció de característiques vl\_phow (funcions VLFEAT) (*dense sift*)
- 2) Selecció aleatòriament de 4.000 característiques
- 3) Agrupament k-means en 600 clústers
- 4) Cerca de les distàncies als diferents KDTrees
- 5) Selecció de la distància mínima
- 6) Assignació de la classe corresponent a la distància mínima

# RESULTATS

---

ITEMPS DE PROCESSAT

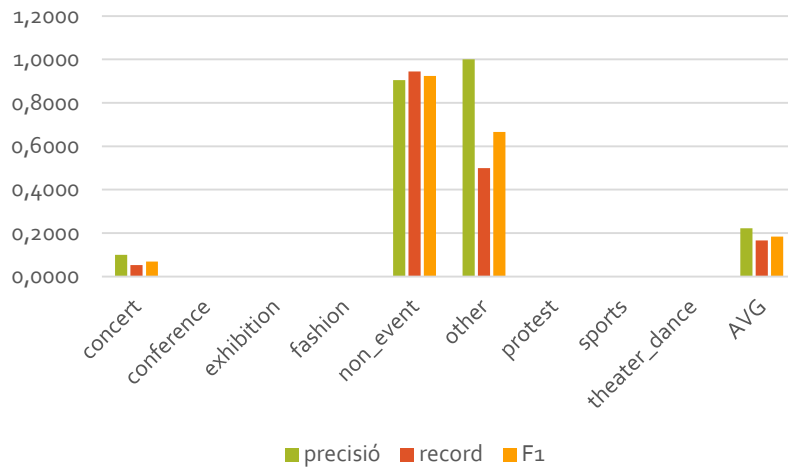
# RESULTATS NUMÈRICS

	SESSIÓ 5 (TEST 33-TRAIN 1)			TEST 2 - TRAIN 1			SESSIÓ 6 (TEST 1 - TRAIN 2)		
	<i>precisió</i>	<i>record</i>	<i>F1</i>	<i>precisió</i>	<i>record</i>	<i>F1</i>	<i>precisió</i>	<i>record</i>	<i>F1</i>
concert	0,1000	0,0526	0,0690	0,2126	0,2784	0,2411	0,0822	0,0261	0,0396
conference	0,0000	0,0000	0,0000	0,0500	0,1250	0,0714	0,0000	0,0000	0,0000
exhibition	0,0000	0,0000	0,0000	0,0638	0,1304	0,0857	0,0270	0,0357	0,0308
fashion	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
non_event	0,9042	0,9440	0,9237	0,8977	0,8536	0,8751	0,8680	0,9468	0,9057
other	1,0000	0,5000	0,6667	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
protest	0,0000	0,0000	0,0000	0,2000	0,1500	0,1714	0,0000	0,0000	0,0000
sports	0,0000	0,0000	0,0000	0,1538	0,1333	0,1429	0,0000	0,0000	0,0000
theater_dance	0,0000	0,0000	0,0000	0,0909	0,1250	0,1053	0,0000	0,0000	0,0000
AVG	0,2227	0,1663	0,1844	0,1854	0,1995	0,1881	0,1086	0,1121	0,1085

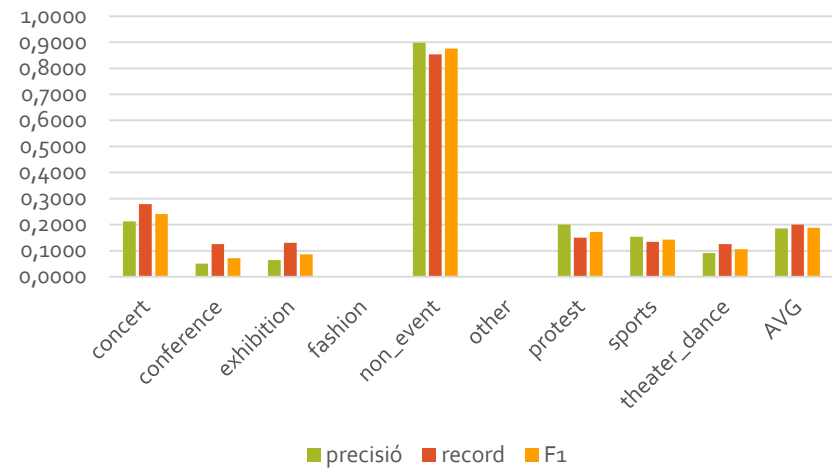


# GRÀFICS

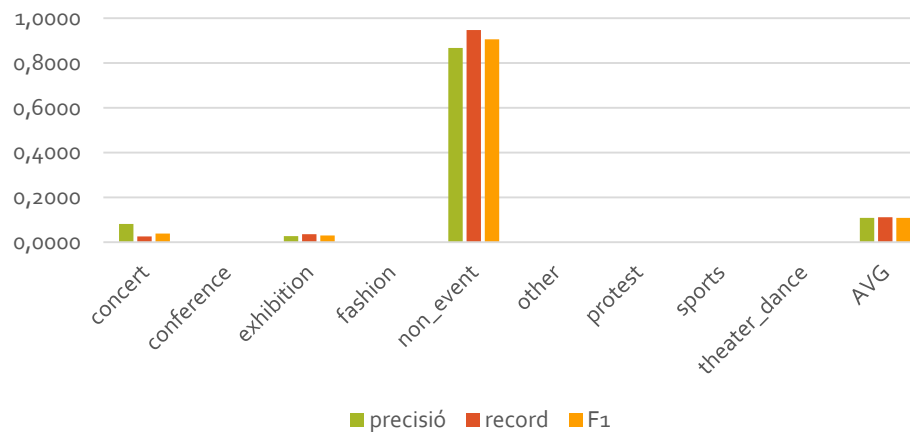
SESSIÓ 5 (TEST 33 - TRAIN 1)



TEST 2 - TRAIN 1



SESSIÓ 6 (TEST 1 - TRAIN 2)



# TEMPS D'EXTRACCIÓ DE CARACTERÍSTIQUES

- Obtaining visual descriptors:0 - Elapsed time is 996.314991 seconds.
- Obtaining visual descriptors:1 - Elapsed time is 799.979601 seconds.
- Obtaining visual descriptors:2 - Elapsed time is 871.000275 seconds.
- Obtaining visual descriptors:3 - Elapsed time is 915.283556 seconds.
- Obtaining visual descriptors:4 - Elapsed time is 1020.583028 seconds.
- Obtaining visual descriptors:5 - Elapsed time is 915.925338 seconds.
- Obtaining visual descriptors:6 - Elapsed time is 956.349341 seconds.
- Obtaining visual descriptors:7 - Elapsed time is 985.286600 seconds.
- Obtaining visual descriptors:8 - Elapsed time is 1008.922571 seconds.
- Obtaining visual descriptors:9 - Elapsed time is 841.008085 seconds.

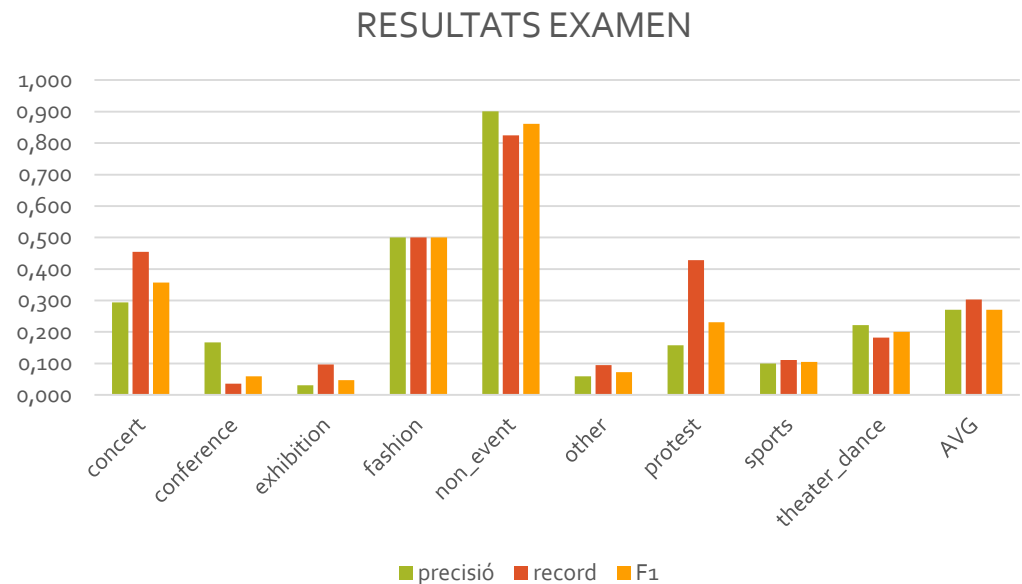
# TEMPS CLASSIFICACIÓ 2.800 fotos (*aprox*)

- Classificació textual (TOTAL): 67.457 s.
- Entrenament visual (2.500 fotos aprox): 3425.760631 seconds.
- Classificació visual : Elapsed time is 3036.190012 seconds.

# RESULTATS EXAMEN

- Entrenament: totes les fotografies acabades en 1-2-3-4-5
- Prova: subconjunt d'entre 2.000 i 3.000 fotografies acabades en 0-6-7-8-9 (*conjunt utilitzat per a l'examen de descriptors*)

	precisió	record	F1
concert	0,294	0,455	0,357
conference	0,167	0,036	0,059
exhibition	0,031	0,097	0,047
fashion	0,500	0,500	0,500
non_event	0,901	0,825	0,861
other	0,059	0,095	0,073
protest	0,158	0,429	0,231
sports	0,100	0,111	0,105
theater_dance	0,222	0,182	0,200
AVG	0,270	0,303	0,270



Classificació textual: Temps total emprat: 85.33 s.

Entrenament visual: Elapsed time is 4122.697810 seconds.

Classificació final (visual condicionada): Elapsed time is 836.518112 seconds.

# Possibles millores

- Les cerques amb índex FULLTEXT de MySQL tenen moltes opcions i diversos paràmetres modificadors que permeten ponderar o forçar paraules. Amb un ús més extensiu d'aquests modificadors i un filtratge dels termes utilitzats es podrien millorar els resultats del classificador textual.
- La classificació visual és molt millorable utilitzant sistemes com els SVM enlloc de la classificació manual que utilitzem nosaltres. També es pot optimitzar la metodologia del processat visual i la quantitat de característiques amb què treballar per fer el procés més eficient.

# GESTIÓ I DISTRIBUCIÓ DE SENYALS AUDIOVISUALS

---

ESCOLA D'ENGINYERIA DE TERRASSA – CURS 2014/25

Grup 3.1

Eduardo Bernal · Mohamed el Bouchti · Luís Varas · Aitor Niñerola