Text Categorization

A Machine Learning Approach

Sabino Miranda-Jiménez, Mario Graff, Eric S. Téllez

Updated: 2019/11/08

Overview

1. Text Categorization
2. Playground - Twitter
3. Supervised Learning
4. μTC and EvoMSA
5. Our approach: MicroTC, EvoMSA
6. Usage

Text Categorization

Text Categorization

The aim is the classification of documents into a fixed number of predefined categories.

Text Categorization

The aim is the classification of documents into a fixed number of predefined categories.

Tasks

- Sentiment Analysis
  (positive, negative, neutral)
- Emotion Ordinary Classification
  (intensity of emotion: 0, 1, 2, 3)
- Aggressive Detection
  (aggressive, no aggressive)

Text Categorization

The aim is the classification of documents into a fixed number of predefined categories.

Tasks (cont.)

- Toxic Comment Detection
  (toxic, obscene, threat, insult)
- Author profiling
  (male, female, age)
- Language variety identification
  (Spanish: argentina, colombia, mexico, peru, spain)
- ...

Example 1: Multilingual sentiment analysis

Sentiment analysis

Positive or Negative message?

قائد في الحرس يعترف بفقدان السيطرة الأمنية في شرقي وغربي إيران

Sentiment analysis

Positive or Negative message?

قائد في الحرس يعترف بفقدان السيطرة الأمنية في شرقي وغربي إيران

Negative

Sentiment analysis

Positive or Negative message?

قائد في الحرس يعترف بفقدان السيطرة الأمنية في شرقي وغربي إيران

Negative

"A Guard commander admits the loss of security control in eastern and western Iran"

"Un comandante de la guardia admite la pérdida del control de la seguridad en el este y oeste de Irán"

Sentiment analysis

Positive or Negative message?

قائد في الحرس يعترف بفقدان السيطرة الأمنية في شرقي وغربي إيران

Negative

"A Guard commander admits the loss of security control in eastern and western Iran"

"Un comandante de la guardia admite la pérdida del control de la seguridad en el este y oeste de Irán"

Positive or Negative message?

Жизнь грустная, зато зарплата смешная. добрыйдень

Sentiment analysis

Positive or Negative message?

قائد في الحرس يعترف بفقدان السيطرة الأمنية في شرقي وغربي إيران

Negative

"A Guard commander admits the loss of security control in eastern and western Iran"

"Un comandante de la guardia admite la pérdida del control de la seguridad en el este y oeste de Irán"

Positive or Negative message?

Жизнь грустная, зато зарплата смешная. добрыйдень

Negative

Sentiment analysis

Positive or Negative message?

قائد في الحرس يعترف بفقدان السيطرة الأمنية في شرقي وغربي إيران

Negative

"A Guard commander admits the loss of security control in eastern and western Iran"

"Un comandante de la guardia admite la pérdida del control de la seguridad en el este y oeste de Irán"

Positive or Negative message?

Жизнь грустная, зато зарплата смешная. добрыйдень

Negative

"Life is sad, but the salary is ridiculous. good afternoon"

"La vida es triste, pero el salario es ridículo. buenas tardes"

Sentiment analysis

Positive or Negative message?

Niltze nocniuhtzitzinhuan, moztla amo tiquitazqueh ompa amoxnamacoyan

Sentiment analysis

Positive or Negative message?

Niltze nocniuhtzitzinhuan, moztla amo tiquitazqueh ompa amoxnamacoyan

Negative

Sentiment analysis

Positive or Negative message?

Niltze nocniuhtzitzinhuan, moztla amo tiquitazqueh ompa amoxnamacoyan

Negative

"Hi friends, tomorrow I will not go with you to the bookstore"

"Hola amigos, mañana no iré a la biblioteca con ustedes"

Example 2: Multilingual author profiling

Female or Male?

1 Lo que se sufre este partido csm!!!
2 NO PUEDO CREERLO
3 La canción ”soy peor” de Bad Bunny es de la csm!!!!
4 Acabo de ver mi foto cuando cumplí 17 años. ¡Cuanto he cambiado!
5 Habrá algo más rico que la chela?
6 Necesito estar bien en unas horas, no puedo pasar mis 22 años en este estado :(
7 Estas cosas no me pueden suceder
8 Hay un lugar tan especial donde yo contigo quisiera estar ♫♫
9 Tengo sueño y no puedo dormir. Esto se está volviendo cotidiano.
10 @Anapaulavega24 búscala en tu corazón

Female or Male?

1 Lo que se sufre este partido csm!!!
2 NO PUEDO CREERLO
3 La canción ”soy peor” de Bad Bunny es de la csm!!!!
4 Acabo de ver mi foto cuando cumplí 17 años. ¡Cuanto he cambiado!
5 Habrá algo más rico que la chela?
6 Necesito estar bien en unas horas, no puedo pasar mis 22 años en este estado :(
7 Estas cosas no me pueden suceder
8 Hay un lugar tan especial donde yo contigo quisiera estar ♫♫
9 Tengo sueño y no puedo dormir. Esto se está volviendo cotidiano.
10 @Anapaulavega24 búscala en tu corazón

male

Female or Male?

1 Te digo adiós para toda la vida, aunque toda la vida la pase pensando en ti
2 El optimismo es la fe, que conduce al logro
3 Nunca sabemos cuando sera la última vez que tendremos la oportunidad de ver o sentir a ese ser querido
4 definitivamente ningún esfuerzo por el , vale la pena
5 que son los amigos acaso ? - NADA ! TAN SOLO UN ESPEJISMO
6 El optimismo es la fe, que conduce al logro
7 entre mas deseo alejarme,mas presente lo tengo
8 esta realmente loco quererme nada mas como amiga,sabiendo que seria la mejor mujer a su lado
9 Prefiero distancias honestas a cercanías hipócritas
10 y me estoy enamorando mas, de tus ojos de tu boca . . .

Female or Male?

1 Te digo adiós para toda la vida, aunque toda la vida la pase pensando en ti
2 El optimismo es la fe, que conduce al logro
3 Nunca sabemos cuando sera la última vez que tendremos la oportunidad de ver o sentir a ese ser querido
4 definitivamente ningún esfuerzo por el , vale la pena
5 que son los amigos acaso ? - NADA ! TAN SOLO UN ESPEJISMO
6 El optimismo es la fe, que conduce al logro
7 entre mas deseo alejarme,mas presente lo tengo
8 esta realmente loco quererme nada mas como amiga,sabiendo que seria la mejor mujer a su lado
9 Prefiero distancias honestas a cercanías hipócritas
10 y me estoy enamorando mas, de tus ojos de tu boca . . .

female

Female or Male?

```
1 @PoliticsNewz what Trump tries to say is hook not ban
2 @nytimes regret to hear that no President in past 50 years ever ca
lled Media bias
3 @puppymnkey @FoxNews FoxNews is root of fake news
4 @markfollman @Wiininiskwe @MotherJones FoxNews create fake news wh
ile Canadian police arrested French student who killed people
5 @ACLU @DemResistance @politico President Obama you are always Rock
 and in the heart of people in USA and around the world
6 @USATODAY very saddened that Somalian living in a difficult situat
ion
7 @euronews_pe thanks for sharing
8 @dtman199 @nationalpost very true and I do agree with you 1 millio
n times
9 @ananavarro very true
10 @MMFlint you are rock since 2015 in front of Trump tower
```

Female or Male?

```
1 @PoliticsNewz what Trump tries to say is hook not ban
2 @nytimes regret to hear that no President in past 50 years ever ca
lled Media bias
3 @puppymnkey @FoxNews FoxNews is root of fake news
4 @markfollman @Wiininiskwe @MotherJones FoxNews create fake news wh
ile Canadian police arrested French student who killed people
5 @ACLU @DemResistance @politico President Obama you are always Rock
 and in the heart of people in USA and around the world
6 @USATODAY very saddened that Somalian living in a difficult situat
ion
7 @euronews_pe thanks for sharing
8 @dtman199 @nationalpost very true and I do agree with you 1 millio
n times
9 @ananavarro very true
10 @MMFlint you are rock since 2015 in front of Trump tower
```

male

Multilingual Analysis

Female or Male?

1 @ashrafthewand دقتعاال ربوك شودنعام تالمجا فوشناه ديكا لمعيواه يغت
ر ىف تاشتام بختنملا هماقلا ول عفني
2 @dina_fikry ىبيباحاي امه انيز نبساح ان بختنملا بعت لذبو دوهجم
3 عقبال امنبقى من الدعشره الااوائل عالميأ نستاهل واالله https://t.co/7X
DGAVdhhB
4 @MRSherbeny ال واالداره شم هزياع حم است ىرضحلا 😏 ان ول مهم هاوقله
تعاالى احسن لود هاوعيضيعوان ه
5 @baherabbas @joe02543 فسالل مزال لصحت هبيصم هريبك ناشلع كرحتن رايو
ب اميتفتحوش تانى الا بالقانون والترخيص الرسميه او قلتهم احسن
6 ألعف هبيعل ةدهجم أدج وكسبوا بعصا ماتشين ورا ضعب
عقبال الناهئى بأذن االله https://t.co/7EgwPpE8tg
7 ازاى بقى اننا كنا خرجنا من الدور الاول بعد ال٣ماتشات زى الوصولنا للن
هائى ..اكيد شم مستاوين https://t.co/09ipH4hN20
8 الحمدلله
المنتخب الكاميرونى بيتحدونا وعازين يكسبونا وثأيرو امنني علشان خسارت
مه الكتير الساقبه امنن
ربنا يستر والنص… https://t.co/u1dlrv2ZPj
9 @ashrafthewand من الواضح انه التعبو الغتشلو سله بيايخ د من هفورصم ام
متهمه
اناسمتغربه انك متهم بالكلمه الصأ
ربنا يوفقك ويكرمك على قد تعبك وطيبتك🌹
10 ازاى بقى اننا كنا خرجنا من الدور الاول بعد ال٣ماتشات زى الوصولنا لل
نهائى ..اكيد شم مستاوين https://t.co/09ipH4hN20

Female or Male?

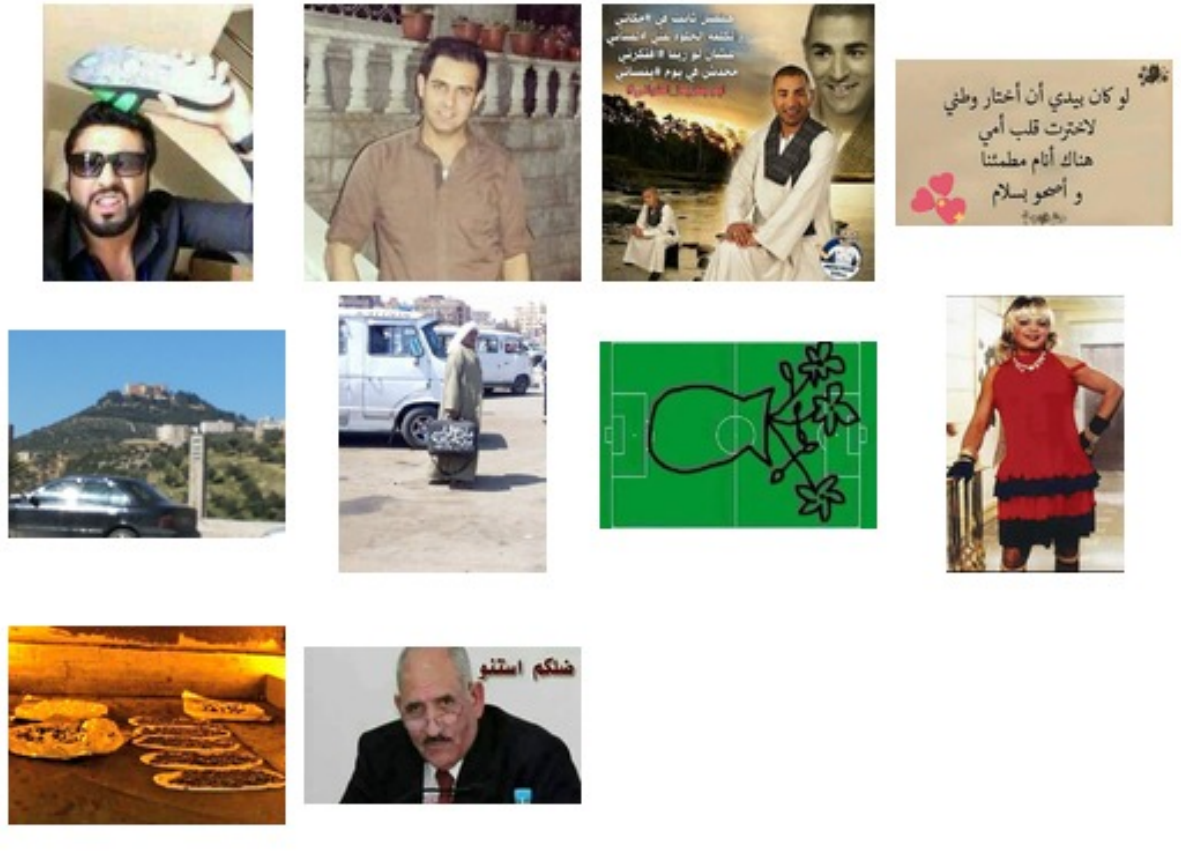
1 @ashrafthewand الاعتقد ربوك ماعندوش مجاملات هاوشوف اكيد هايعمل تغي
ر فى ماتشات المنتخب القادمه لو عفني
2 @dina_fikry ياحبايبى امه انيز حاسبن ان المنتخب بعت وبذل دوهجم
3 عقبال مانبقى من العشره الاوائل عالميأ نستاهل والله https://t.co/7X
DGAVdhhB
4 @MRSherbeny ال والداره مش ءزيزه حماست الحضرى 😏ان لو مهمه هلوقوا
تعاالى احسن دول هاوعيضيه
5 @baherabbas @joe02543 فسالل مزال لصحت هبيصم هريبك علشان كرحتن وياير
نسحا مهتلق وا هيمسرلا صيخارتلاو نونقالاب ال ىنات شوحتفتيام ب
6 ألعف هبيعل هدهجم أدج وكسبوا بعصا ماتشين وار ضعب
عقبال النهائى بأذن الله https://t.co/7EgwPpE8tg
7 ازاى بقى اننا كنا خرجنا من الدور الاول بعد ال٣ماتشات زى الوصول لنن
هائى .. اكيد مش متساوين https://t.co/09ipH4hN20
8 الحمدلله
تراسخ نا علشان اننا اوراثيو انوبسكي نيزياعو انودحتيب ىنوريم الكا المنتخب
مه الكتير السابقه اننا
ربنا يستر واللمص… https://t.co/u1dlrv2ZPj
9 @ashrafthewand من الواضح انه التعب والشغل وسله بيايخد مصروفه من ما
متمه
أصلا مه الكبت متمه كنا هبرغتسم انا
كتبيطو🌹 كبعت قد على كمركيو كقفوي ربنا
10 ازاى بقى اننا كنا خرجنا من الدور الاول بعد ال٣ماتشات زى الوصول لل
نهائى .. اكيد مش متساوين https://t.co/09ipH4hN20

female

Example 3: Beyond text analysis: Multimodal author profiling

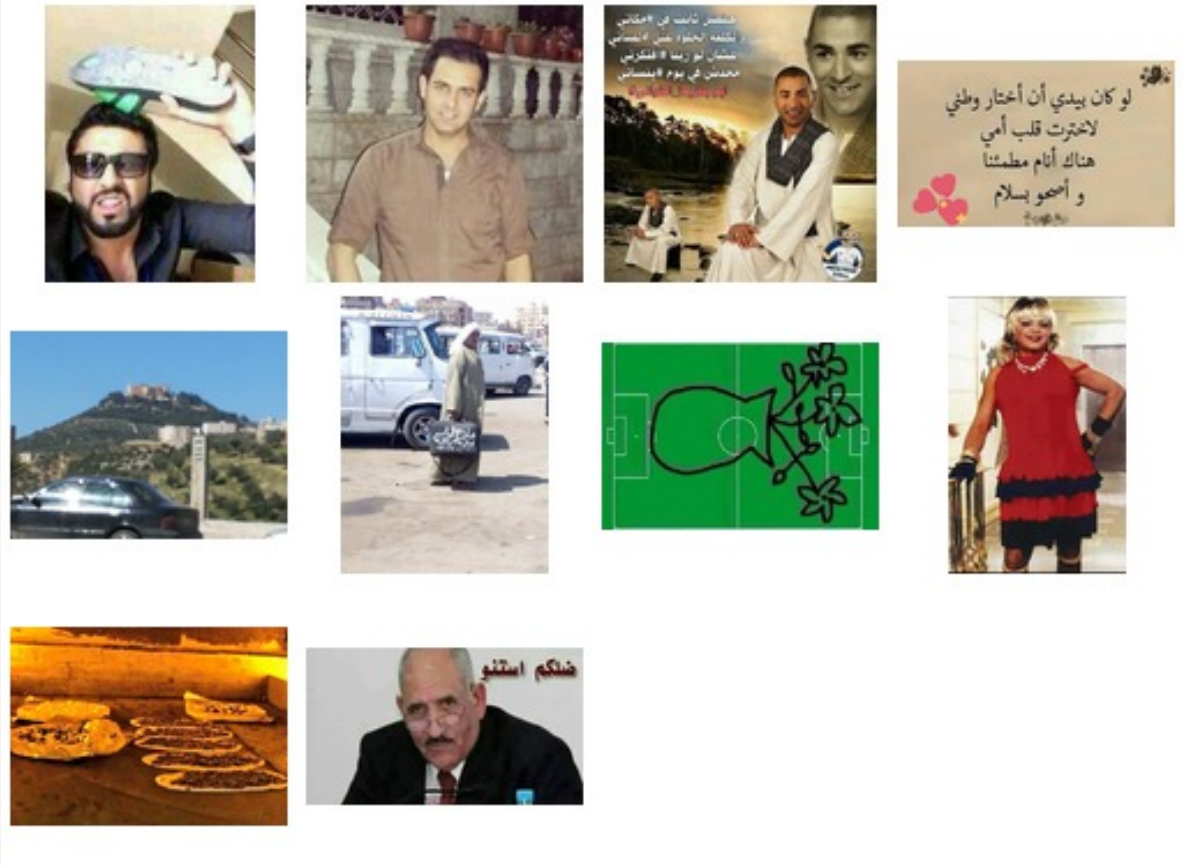
User profiling - Arabic

Male or female?



User profiling - Arabic

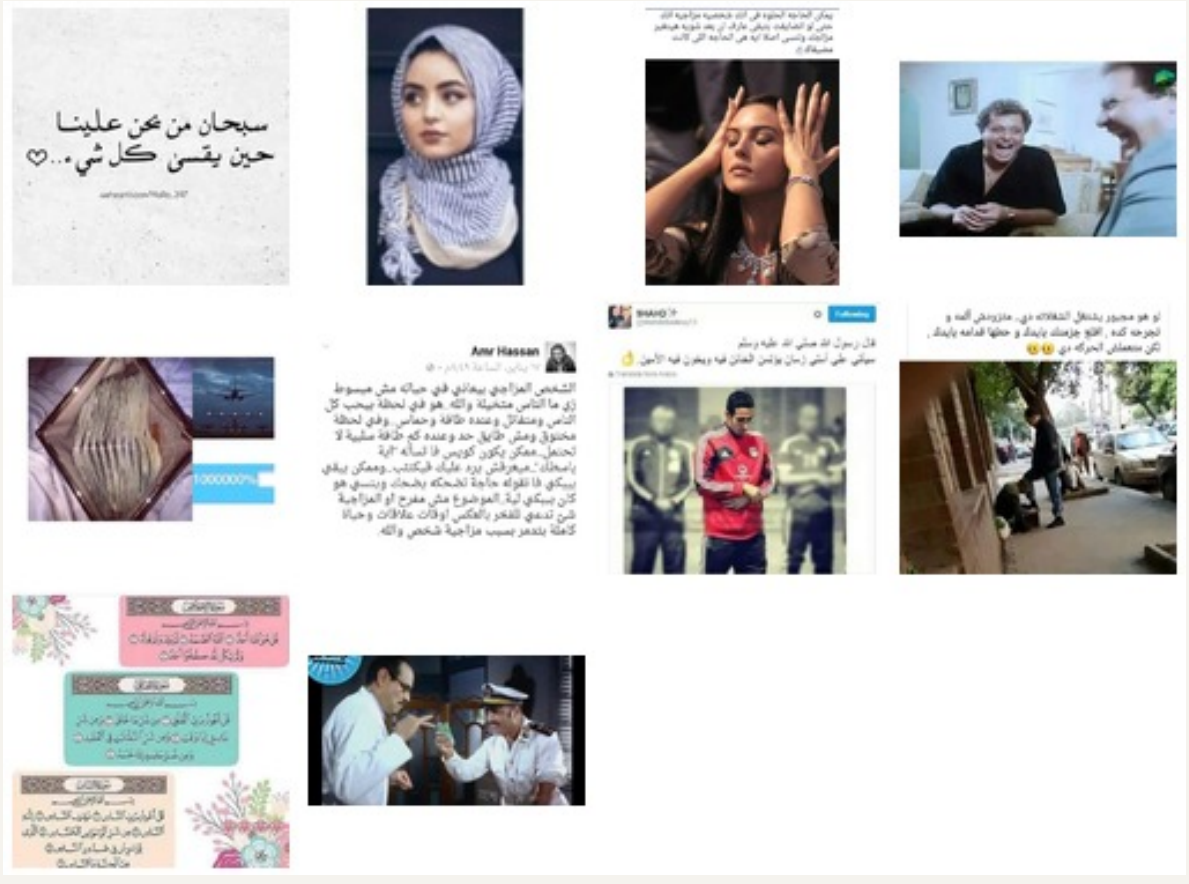
Male or female?



male

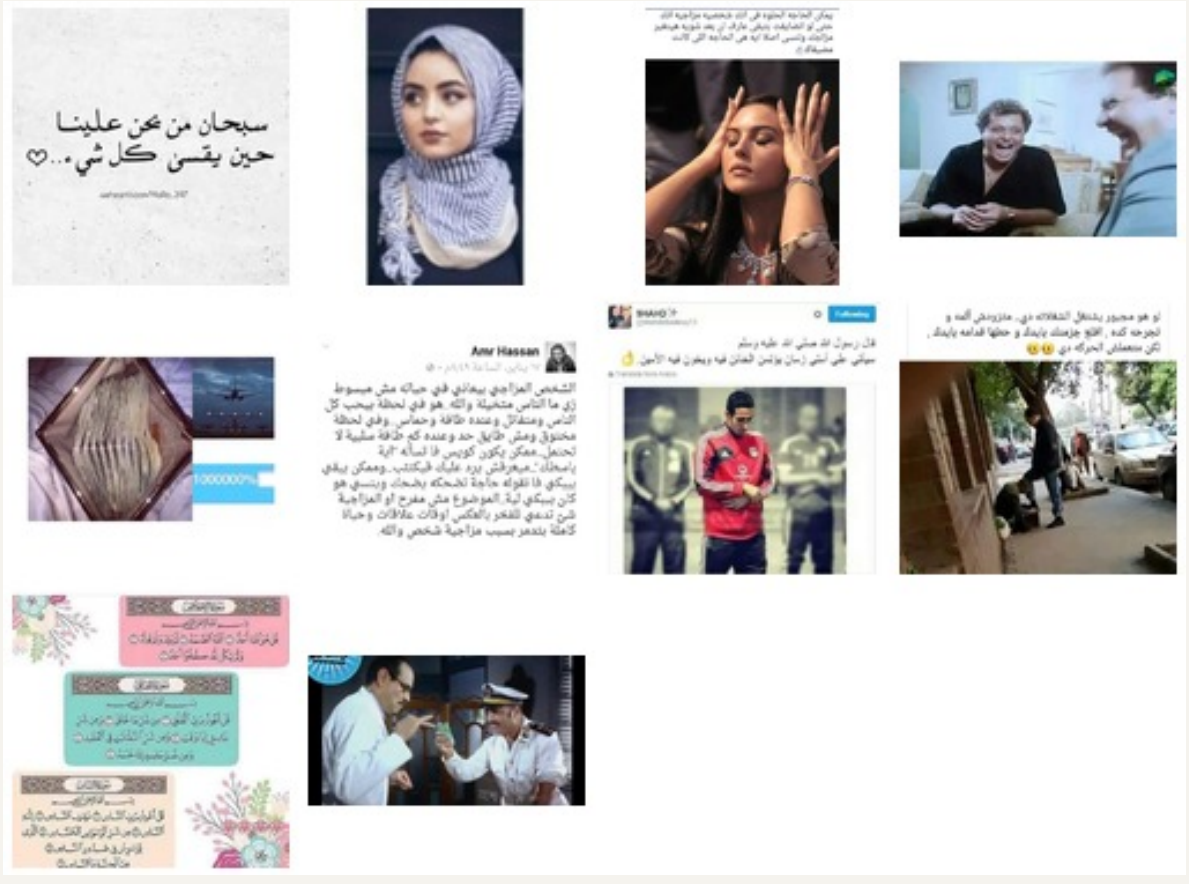
User profiling - Arabic language

Male or female?



User profiling - Arabic language

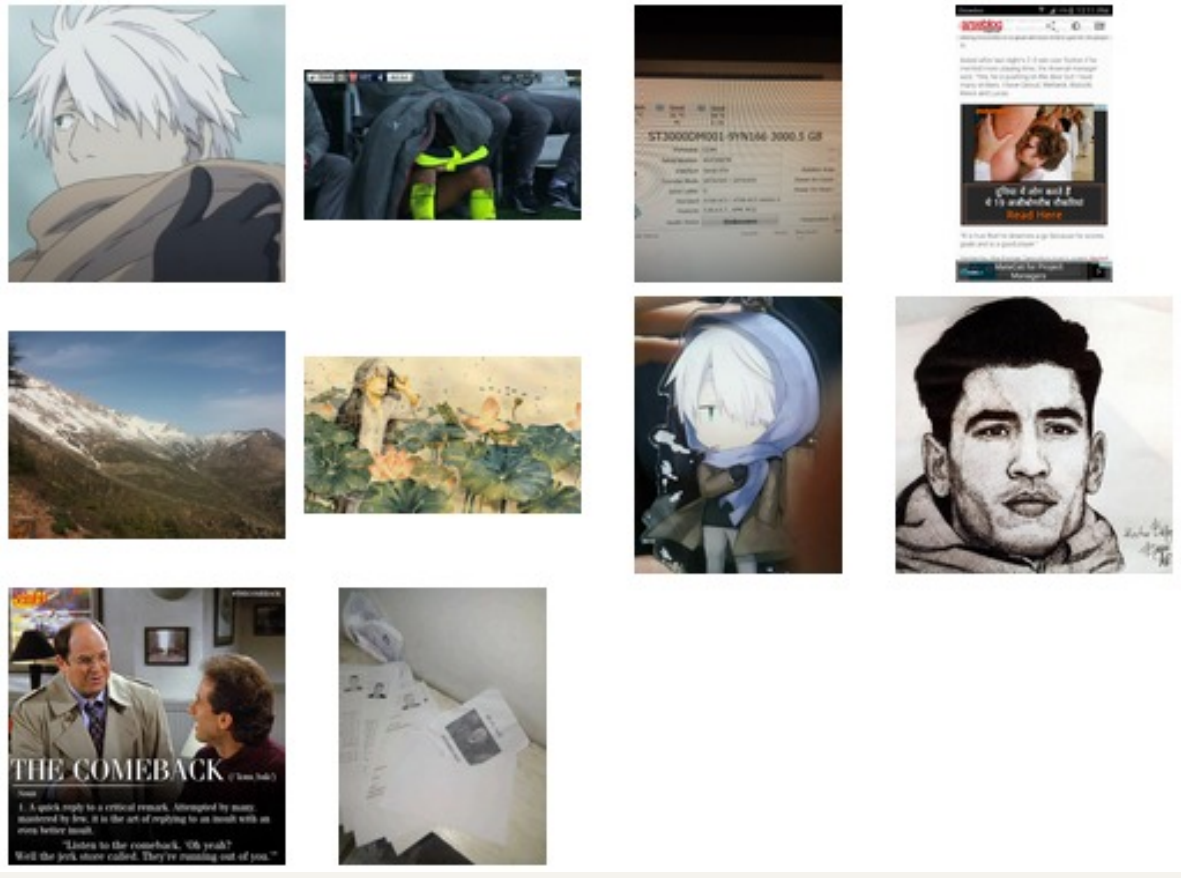
Male or female?



female

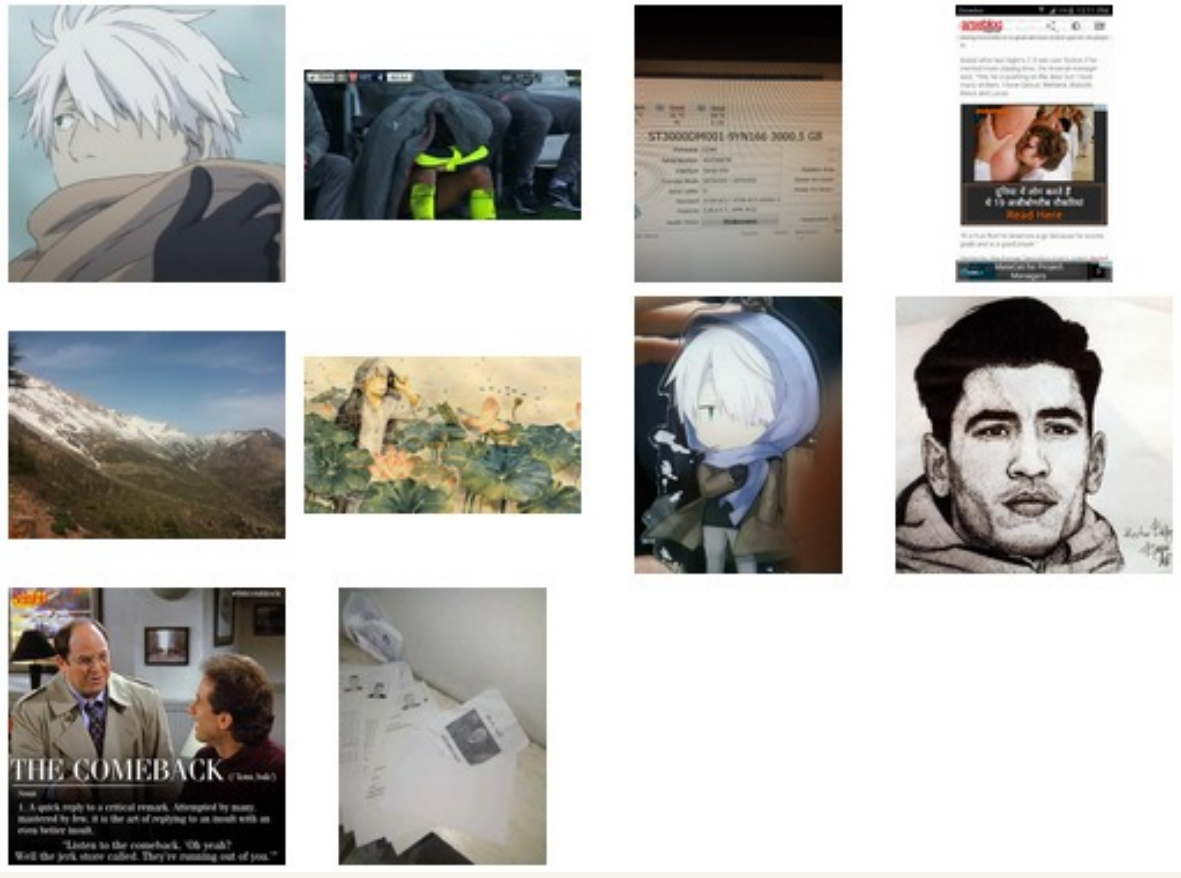
User profiling - Arabic language

Male or female?



User profiling - Arabic language

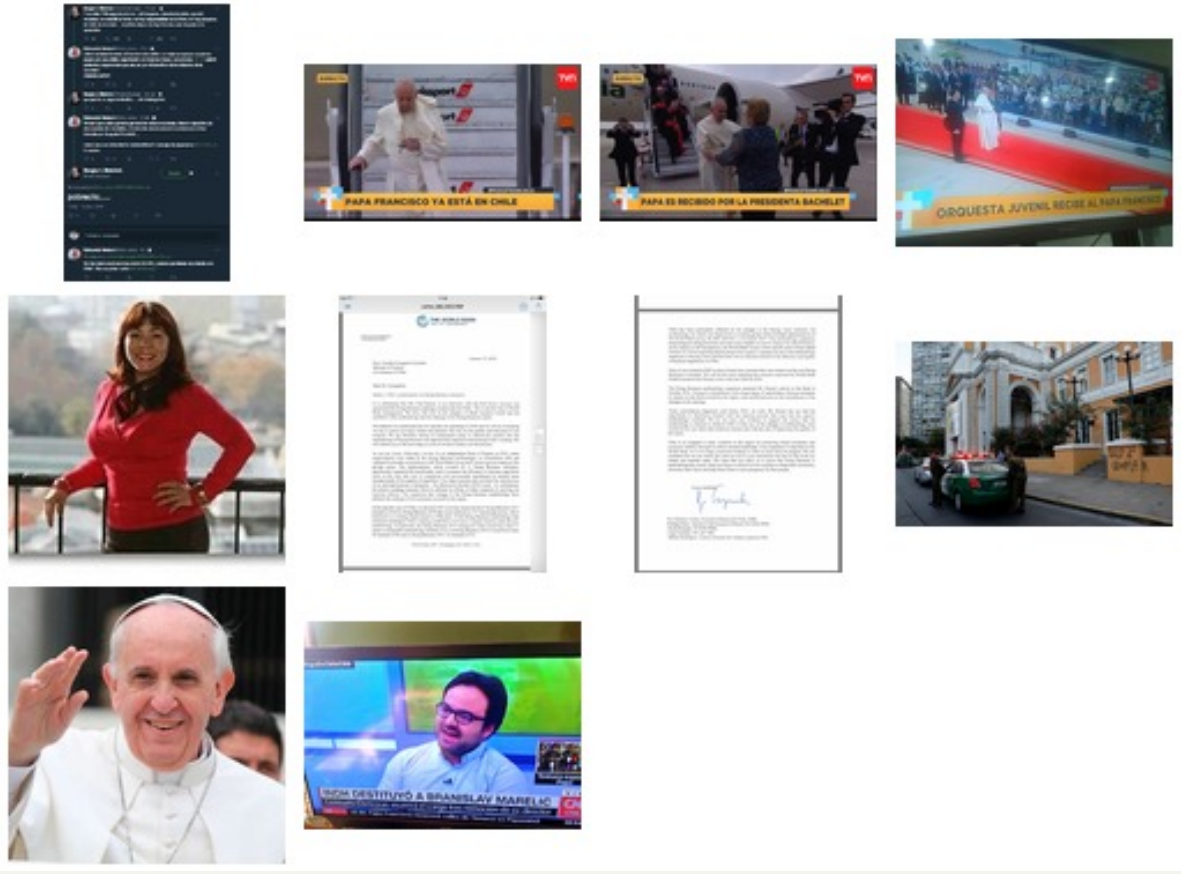
Male or female?



male

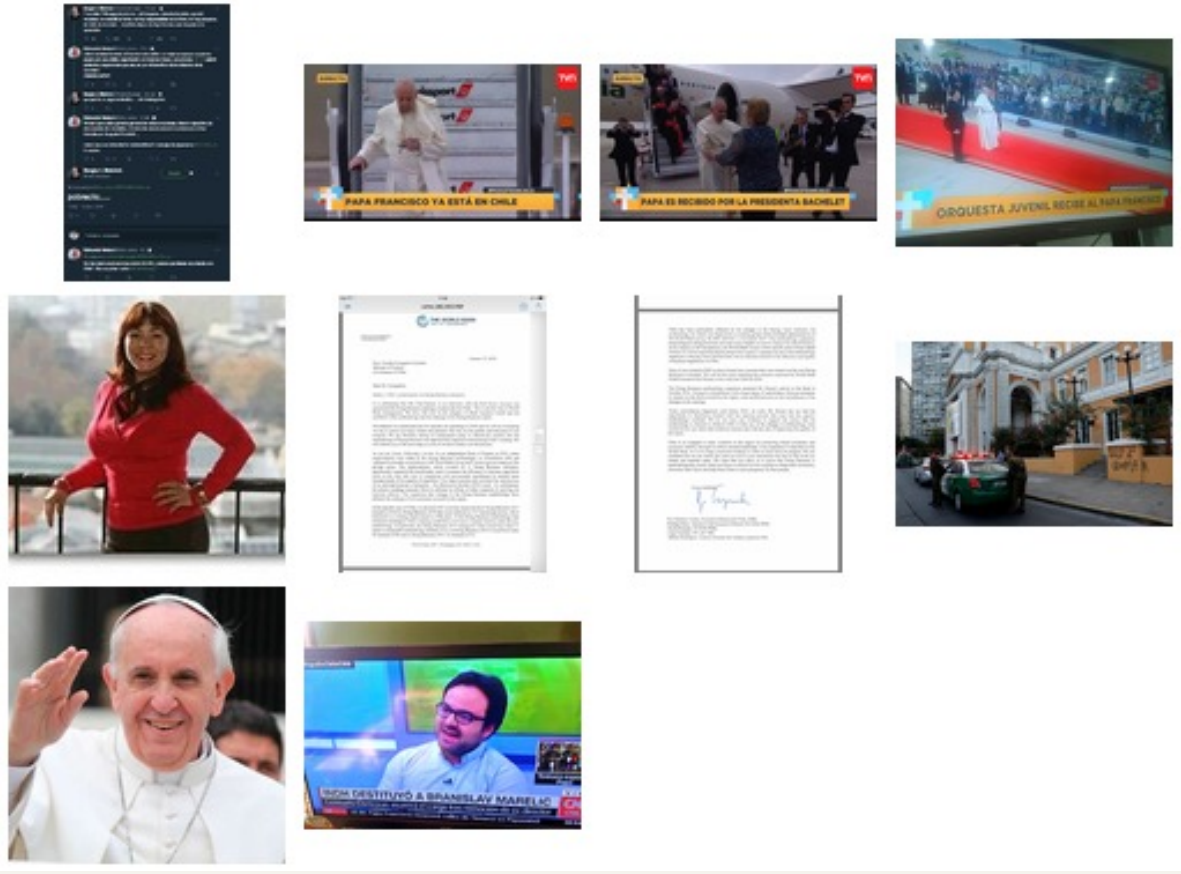
User profiling - Spanish

Male or female?



User profiling - Spanish

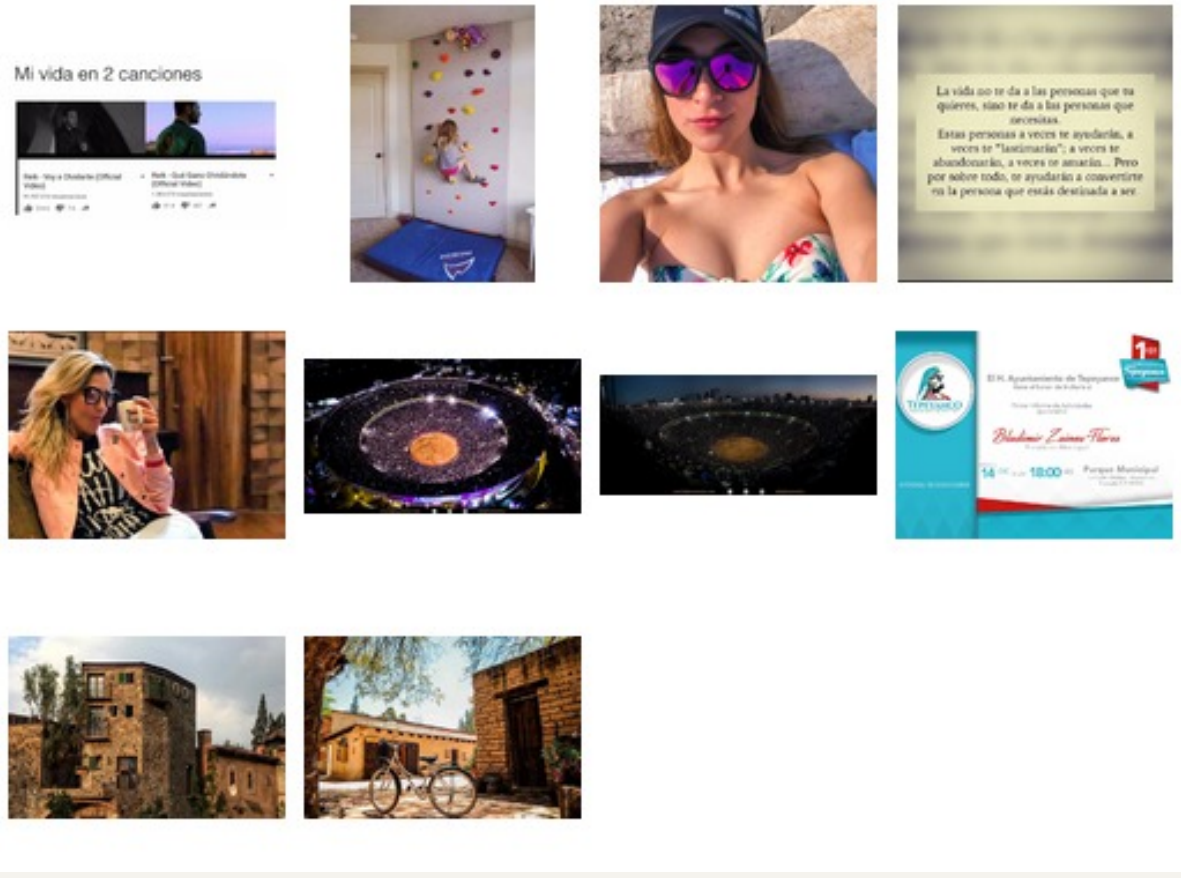
Male or female?



female

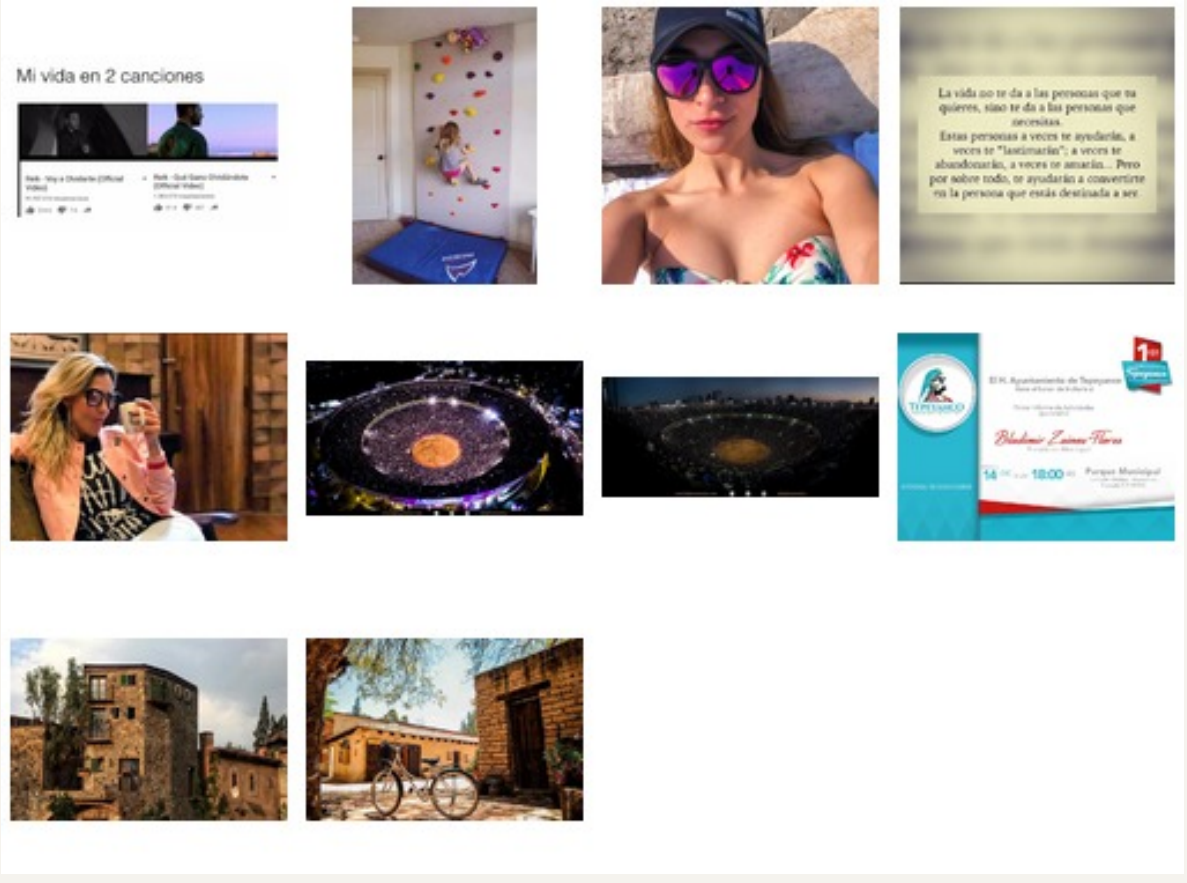
Author profiling - Spanish language

Male or female?



Author profiling - Spanish language

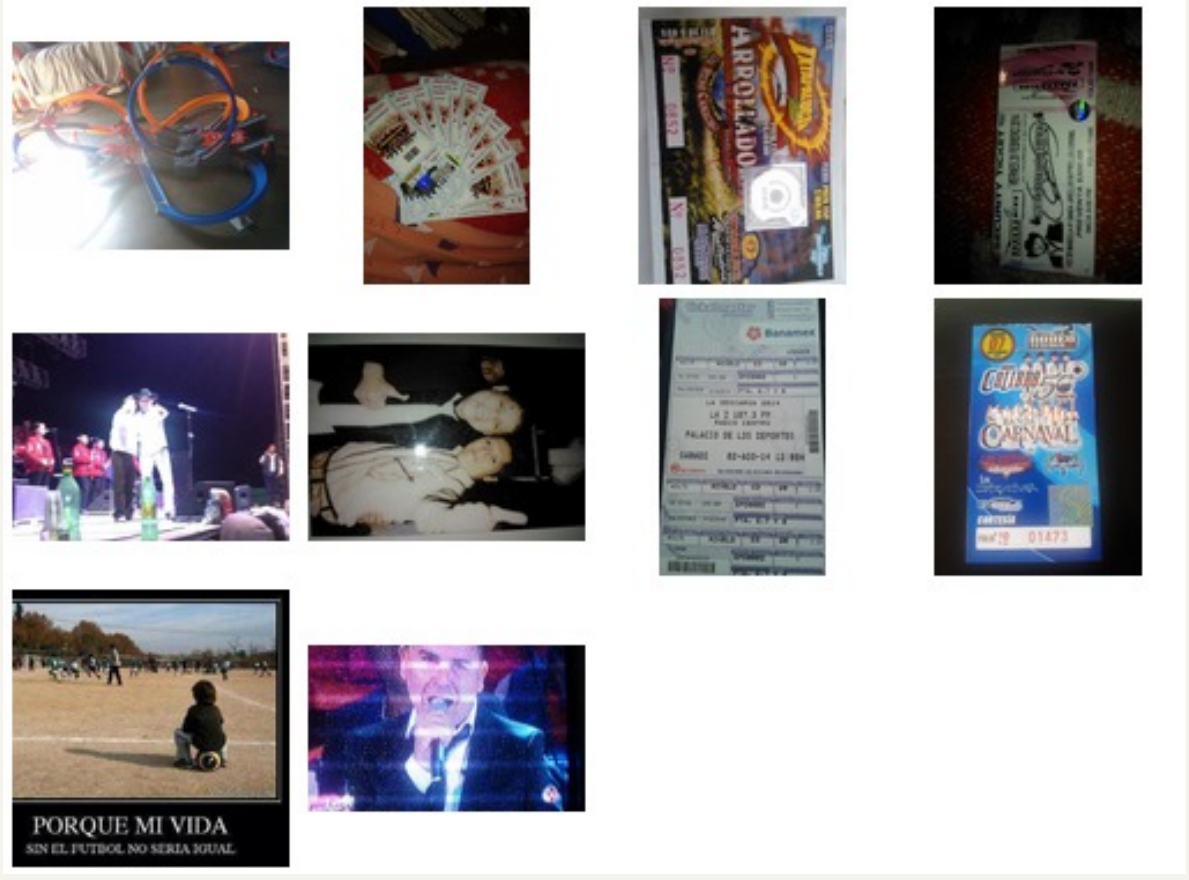
Male or female?



female

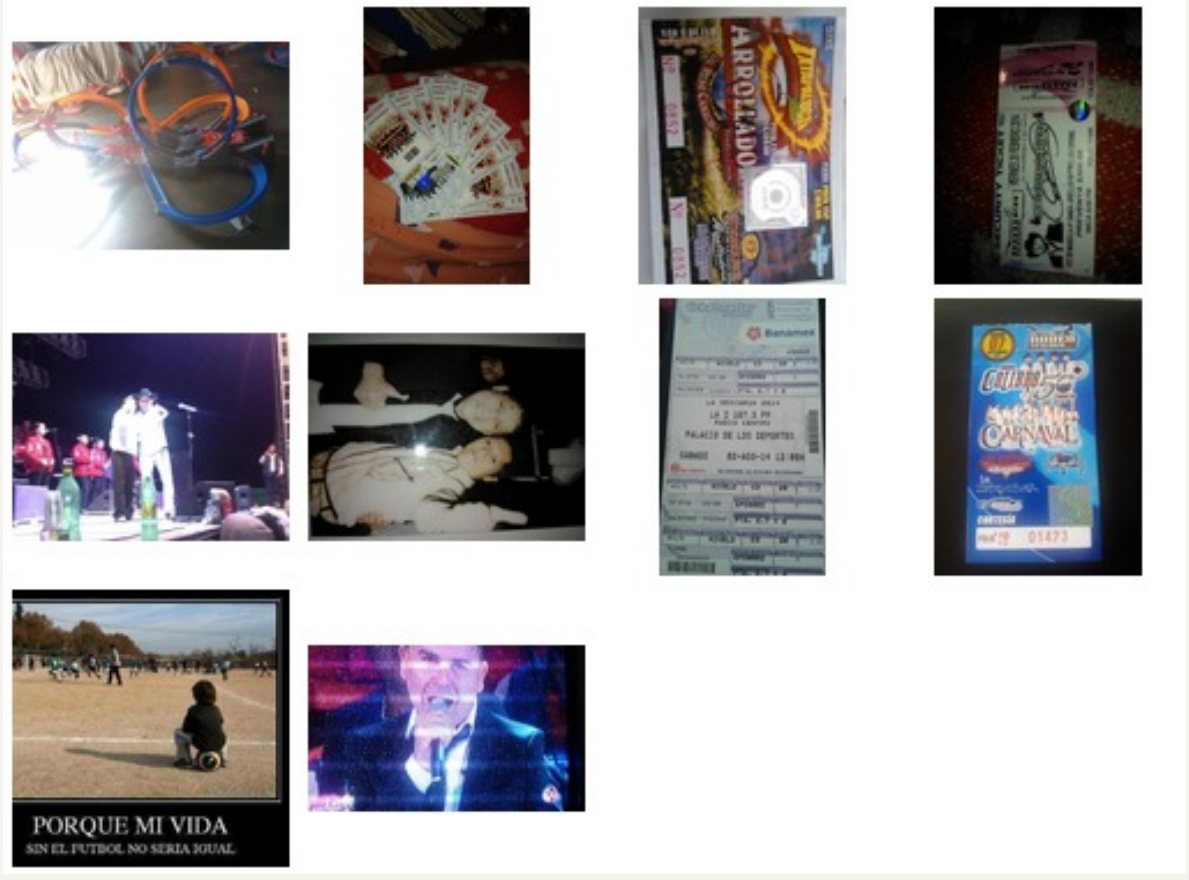
Author profiling - Spanish language

Male or female?



Author profiling - Spanish language

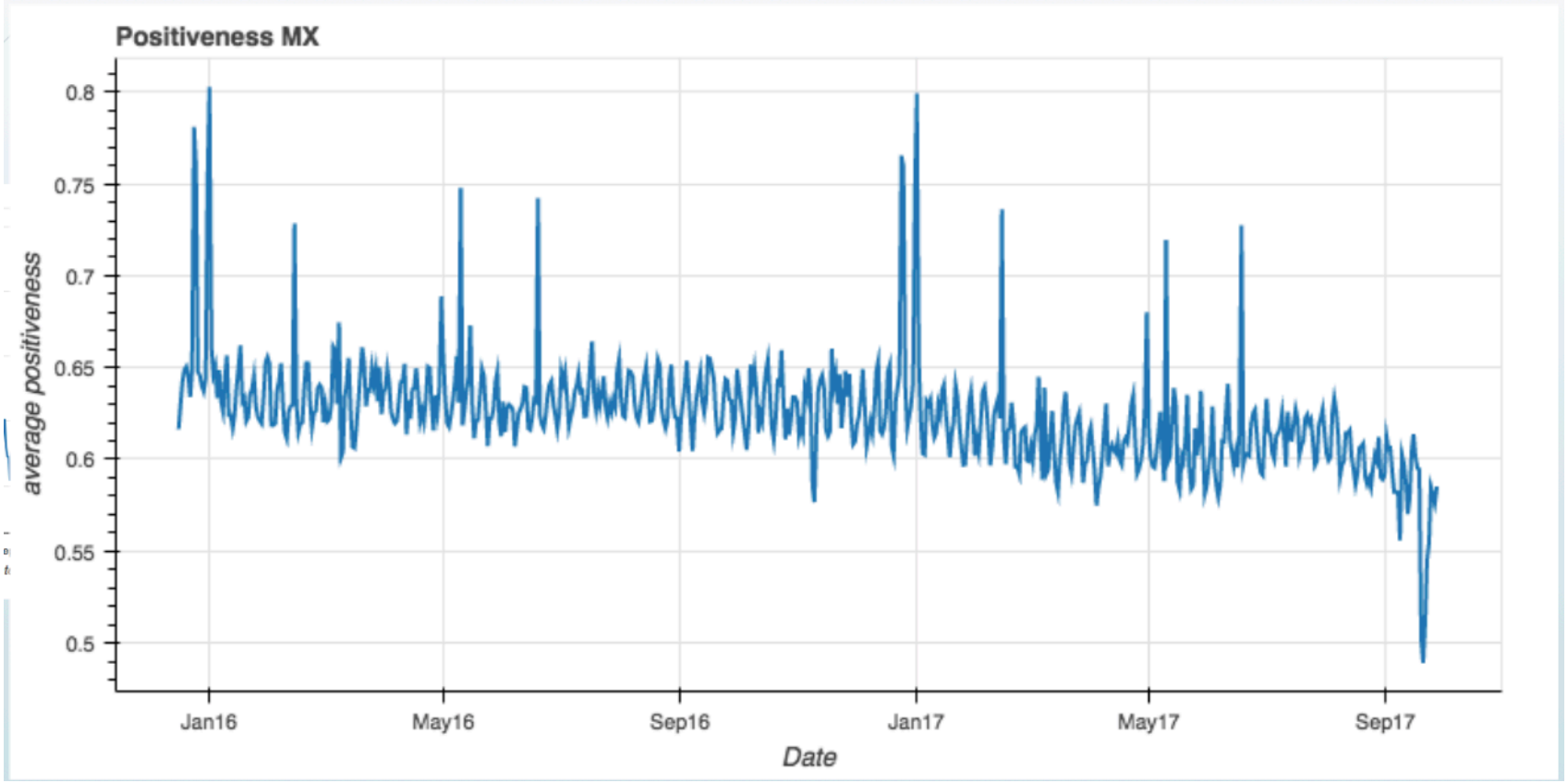
Male or female?



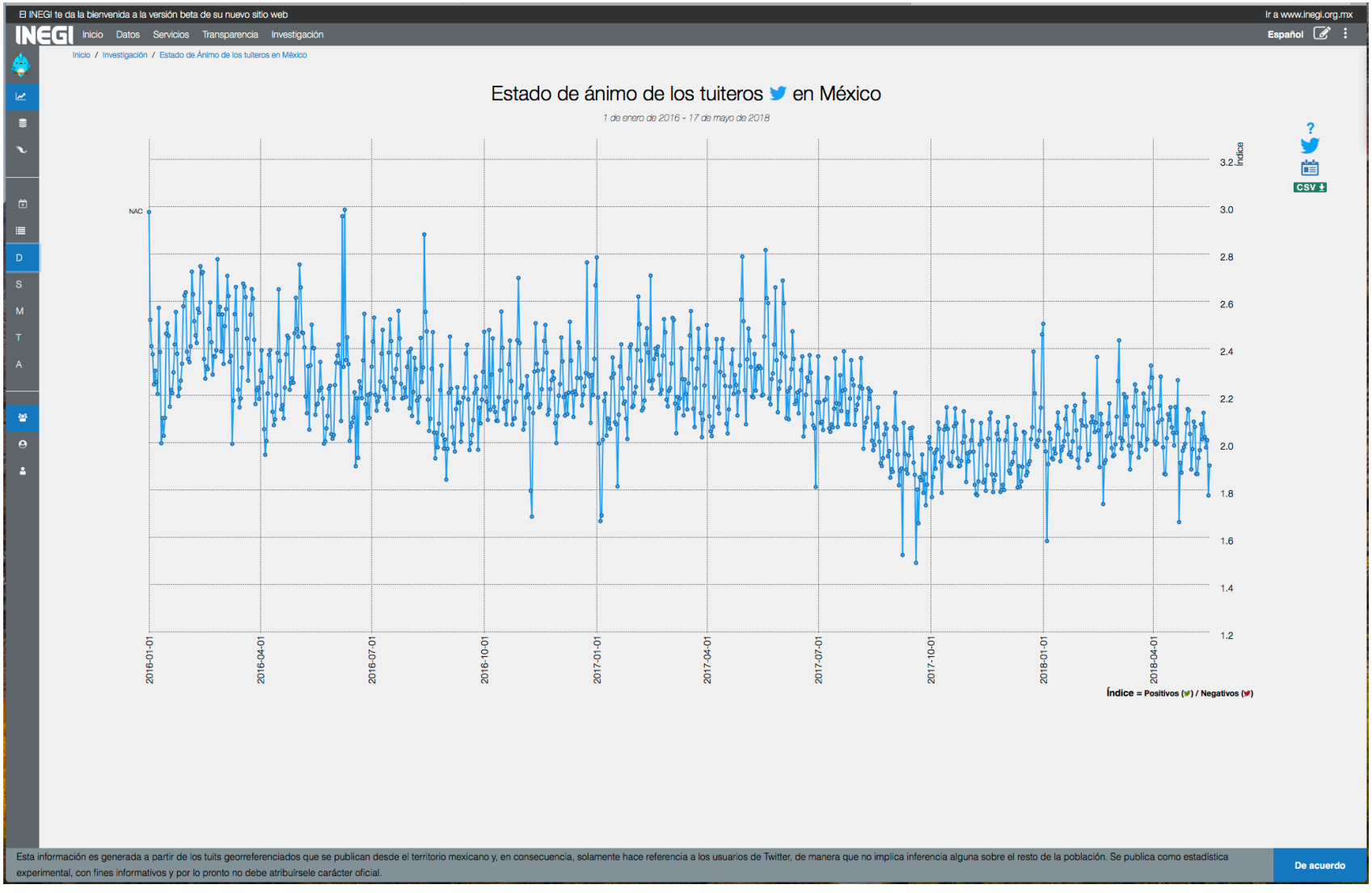
male

Playground - Twitter

Mexico's Positiveness

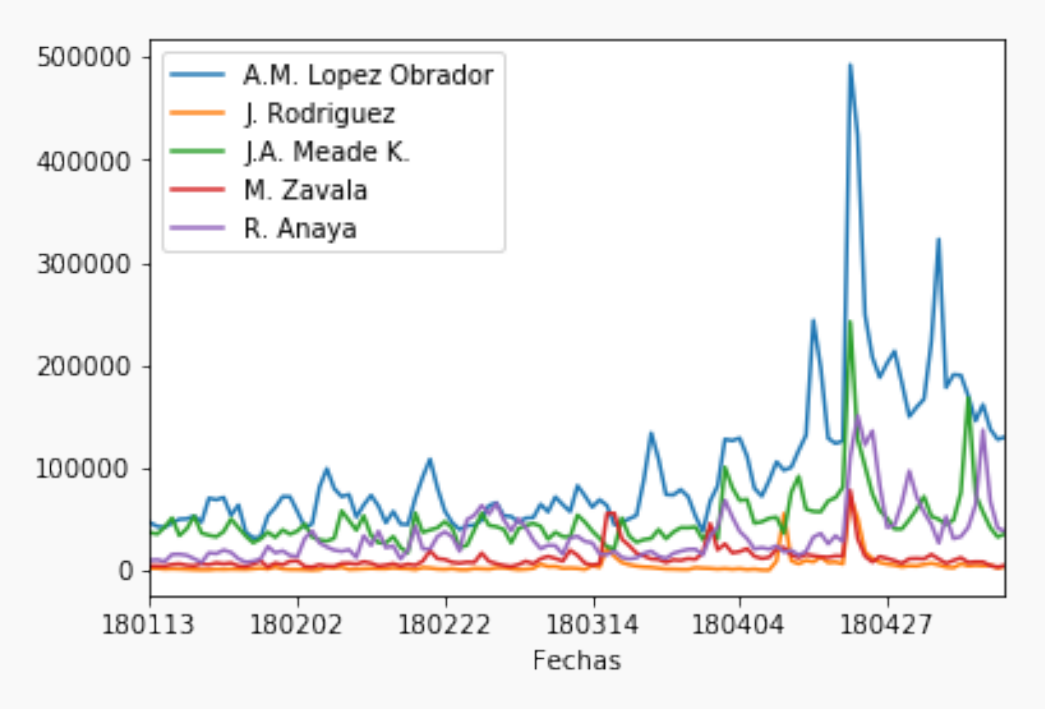


Estado de Ánimo de los tuiteros



Mexico's Earthquake - 19-S

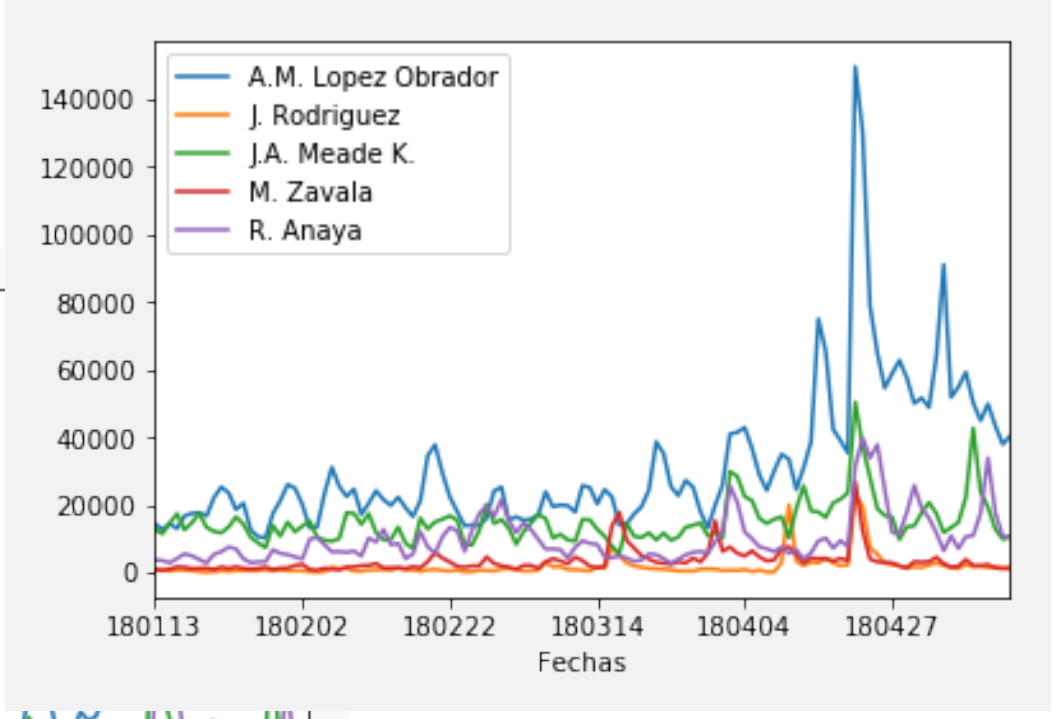
Mexico's 2018 Presidential Campaign in Twitter – mentions

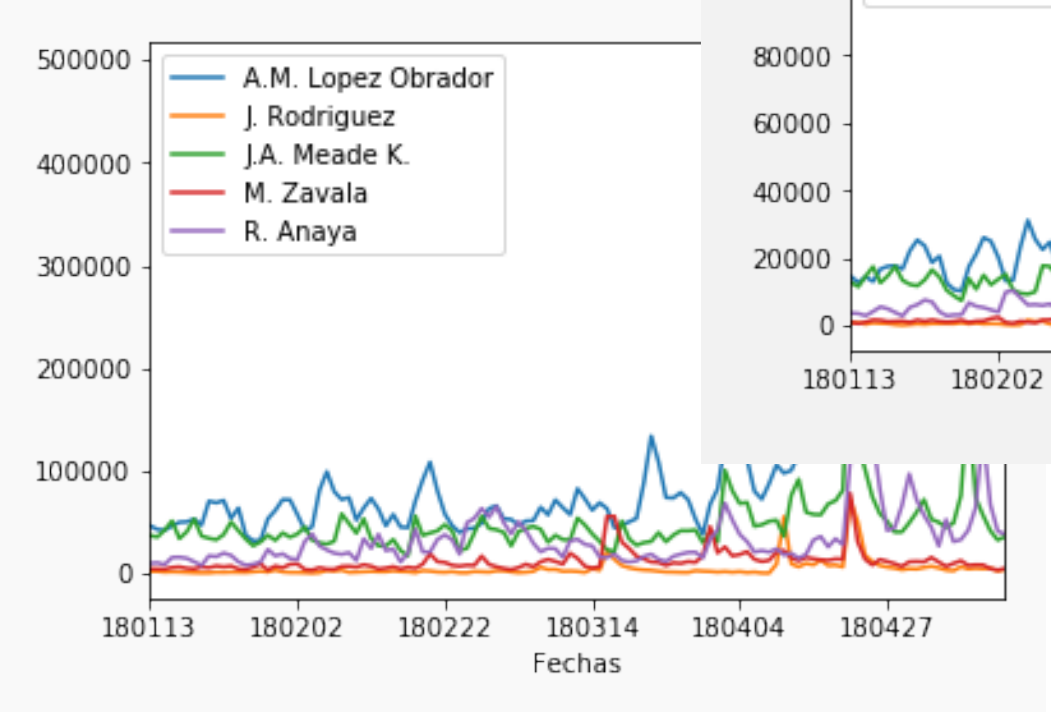


Per tweet

Mexico's 2018 Presidential Campaign in Twitter – mentions

Per user

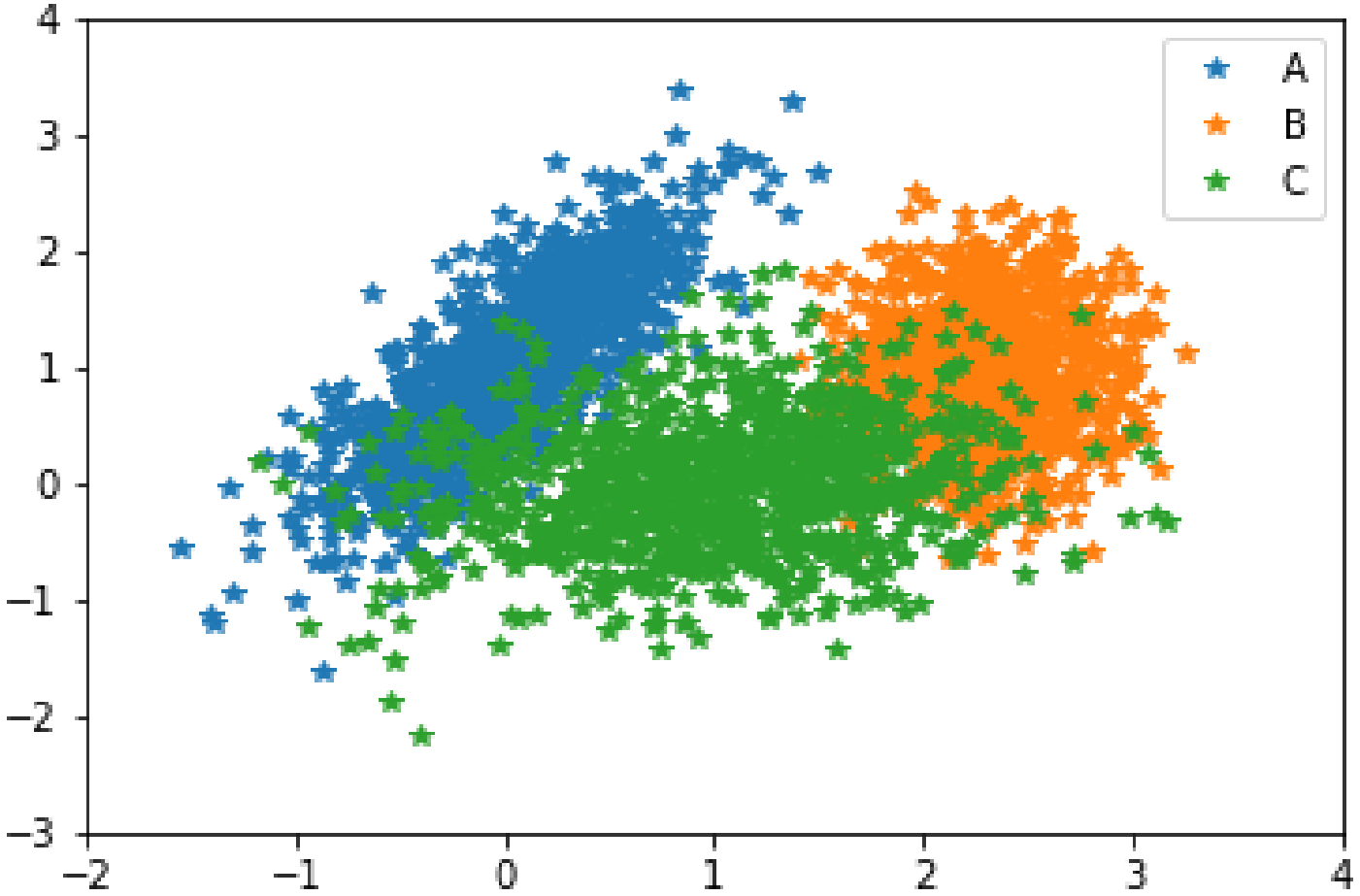




Per tweet

Supervised Learning

Supervised Learning

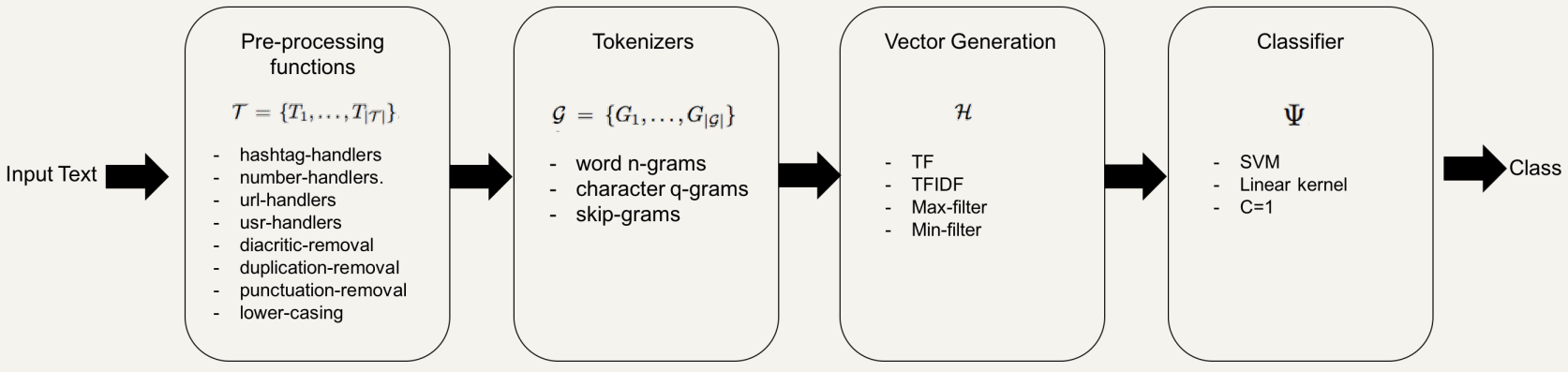


μTC and EvoMSA

Features

- A minimalist combinatorial framework for finding a competitive text classifier
- An error robust representation based on skip n-grams, n-grams (words) and q-grams (characters)
- Language and domain independent
- SVM-based classifier
- Efficient configuration space sampling (Random Search, Hill Climbing)
- Open source and publicly available

Pipeline



- E. S. Tellez, D. Moctezuma, S. Miranda-Jiménez, M. Graff, An automated text categorization framework based on hyperparameter optimization. Knowledge-Based Systems, vol. 149, pp. 110-123, 2018, ISSN 0950-7051.
- https://github.com/INGEOTEC/microtc

Examples of configurations

Text transformation	English	Arabic
remove diacritics	no	yes
remove duplicates	no	yes
remove punctuation	no	yes
emoticons	none	delete
lowercase	yes	yes
numbers	group	group
urls	group	delete
users	delete	none
Term weighting		
TF-IDF	yes	yes
Tokenizers		
n-words	$\{2, 3\}$	$\{2\}$
q-grams	$\{3, 5, 9\}$	$\{3, 5\}$
skip-grams	–	$\{(3, 1)\}$

q-grams

- q-grams capture suffixes (similar a stemming process), borders between words, and tolerate errors inside words

Example:

$T =$ `I_like_vanilla` $\quad T' =$ `I_lik3_vanila`

Extracting 3-grams objects are more similar:

$Q_3^T =$

{`I_l`, `_li`, `lik`, `ike`, `ke_`, `e_v`, `_va`, `van`, `ani`, `nil`, `ill`, `lla`}

$Q_3^{T'} =$

{`I_l`, `_li`, `lik`, `ik3`, `k3_`, `3_v`, `_va`, `van`, `ani`, `nil`, `ila`}

The similarity using Jaccard's coefficient

$$\frac{|Q_3^T \cap Q_3^{T'}|}{|Q_3^T \cup Q_3^{T'}|} = 0.448.$$

q-grams (cont.)

Example:

$T =$ `I_like_vanilla` $\quad T' =$ `I_lik3_vanila`

Using unigrams (words), Jaccard's coefficient:

$$\frac{|\{\texttt{I, like, vanilla}\} \cap \{\texttt{I, lik3, vanila}\}|}{|\{\texttt{I, like, vanilla}\} \cup \{\texttt{I, lik3, vanila}\}|} = 0.2$$

Performance

- Random Search (RS) selects the best performing configuration among the set $\mathscr{C}'$ randomly chosen from $\mathscr{C}$

$$\arg\max_{c \in \mathscr{C}'} \mathsf{score}(c),$$

- Hill Climbing explores greedily the neighborhood $N(c)$, (the best sample in RS) finding the best performing configuration in $N(c)$.
- Score functions could be F1, Accuracy, Precision, or Recall to measure the quality of the text classifier

Datasets

Table: Datasets used in the analysis of n-grams vs q-grams

benchmark		classes				
name	part	positive	neutral	negative	none	total
INEGI	train	2,908	986	1,110	409	5,413
	test	26,911	8,868	9,571	3,361	48,711
						54,124
TASS'15	train	2,884	670	2,182	1,482	7,218
	test	22,233	1,305	15,844	21,416	60,798
						68,016

- A case study of Spanish text transformations for twitter sentiment analysis. ES Tellez, S Miranda-Jiménez, M Graff, D Moctezuma, OS Siordia, EA Villaseñor. Expert Systems with Applications 81, 457-471, 2017.

Metrics

$$\text{accuracy} = \frac{\text{total TP} + \text{total TN}}{\text{total samples}}$$

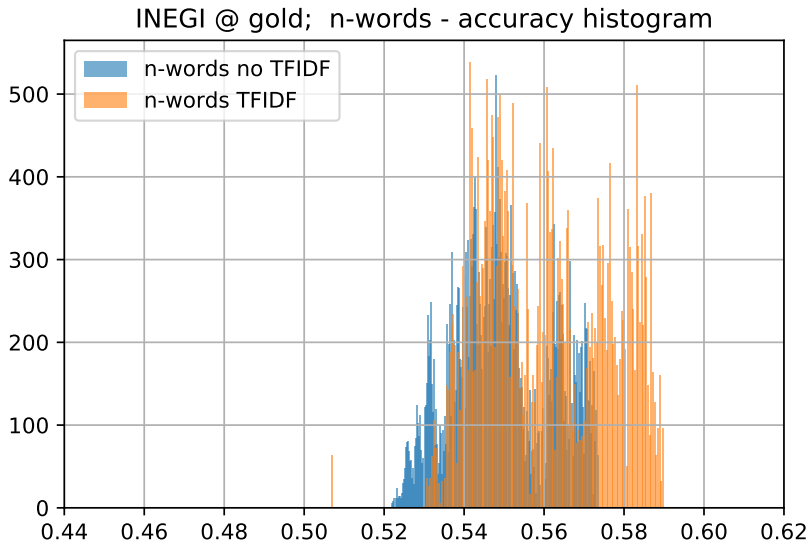
$$\text{precision}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FP}_c}$$

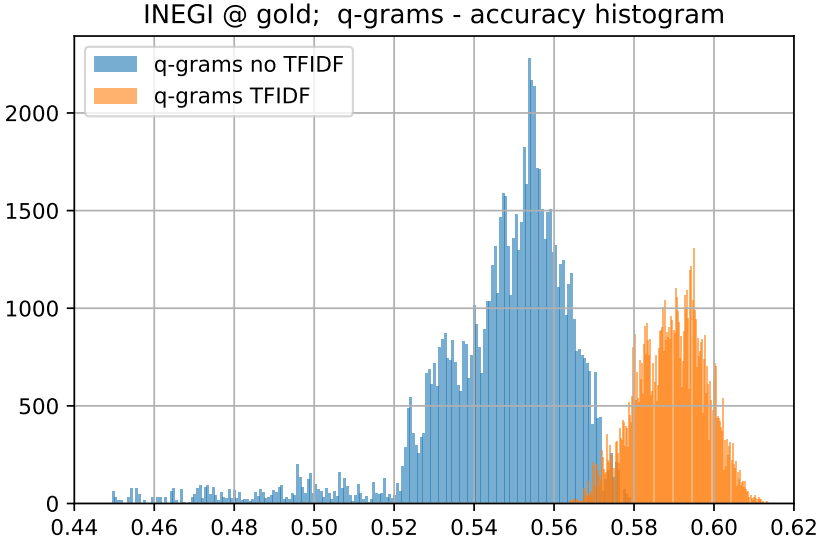
$$\text{recall}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c}$$

$$F_{1,c} = \frac{2 \cdot \text{accuracy}_c \cdot \text{recall}_c}{\text{accuracy}_c + \text{recall}_c}$$

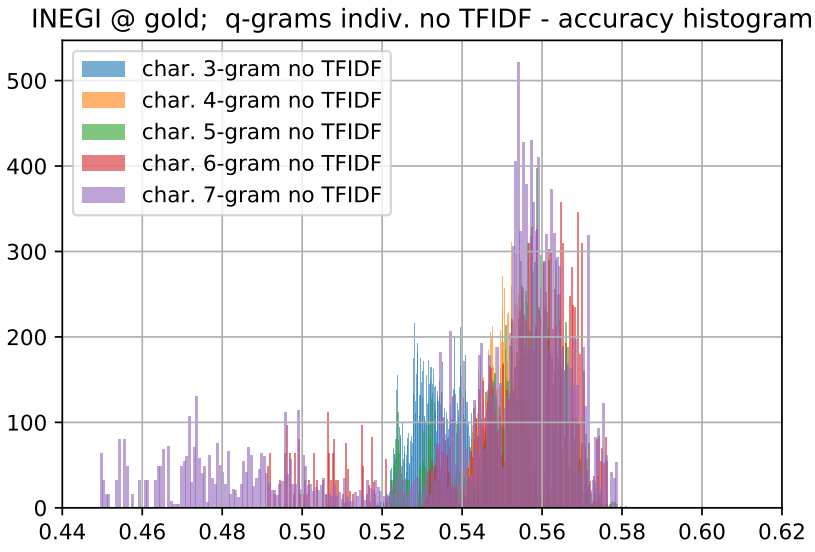
$$\text{macro-}F_1 = \frac{1}{|\mathcal{L}|} \sum_{c \in \mathcal{L}} F_{1,c}$$

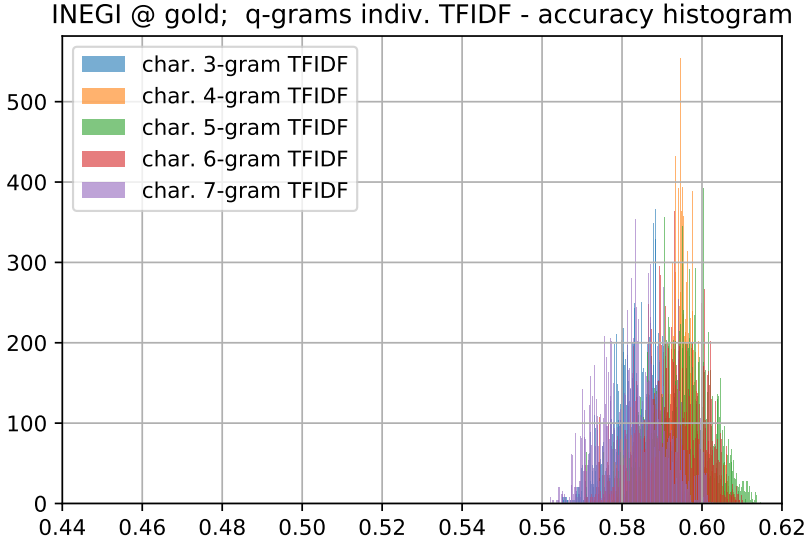
Performance n-words vs q-grams (INEGI-gold)



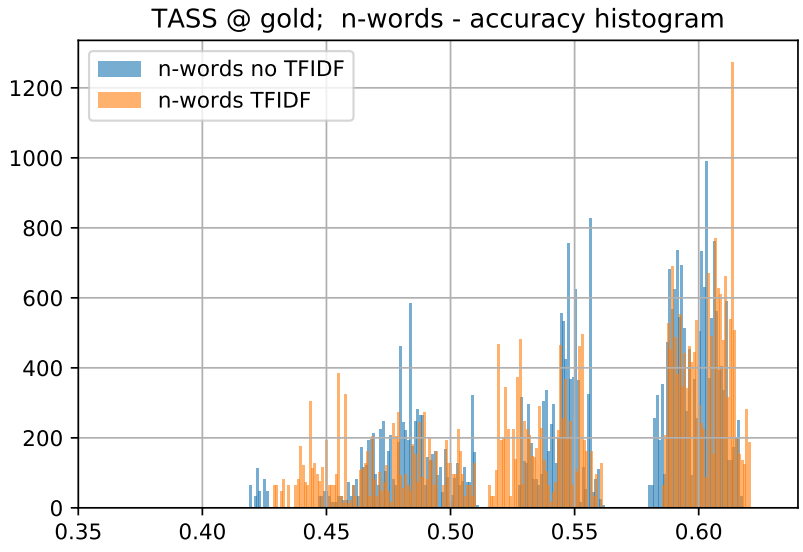


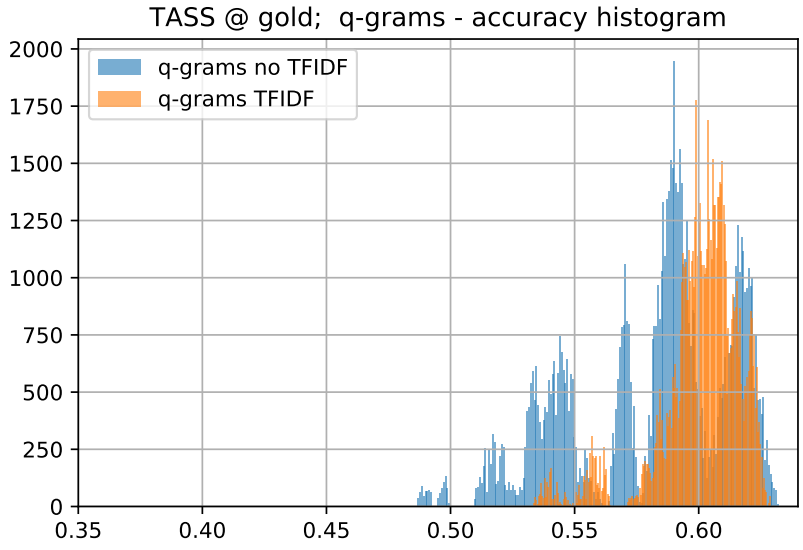
Performance q-grams (INEGI-gold)



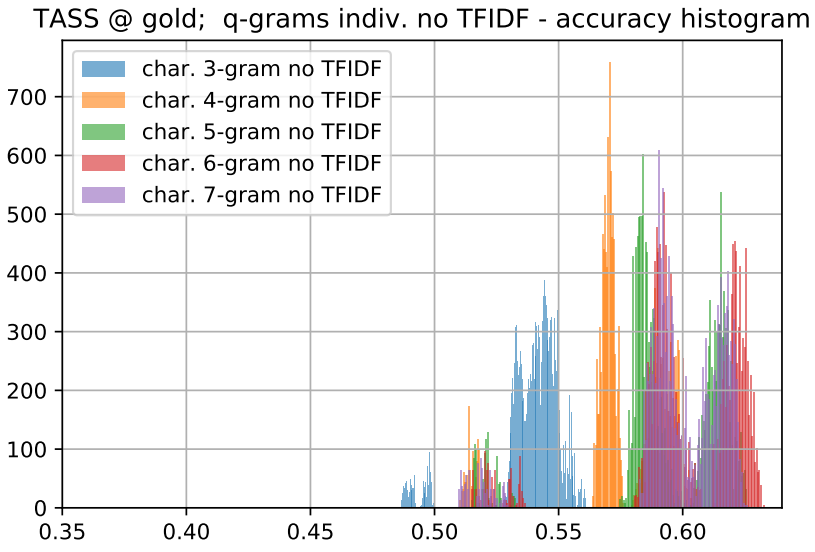


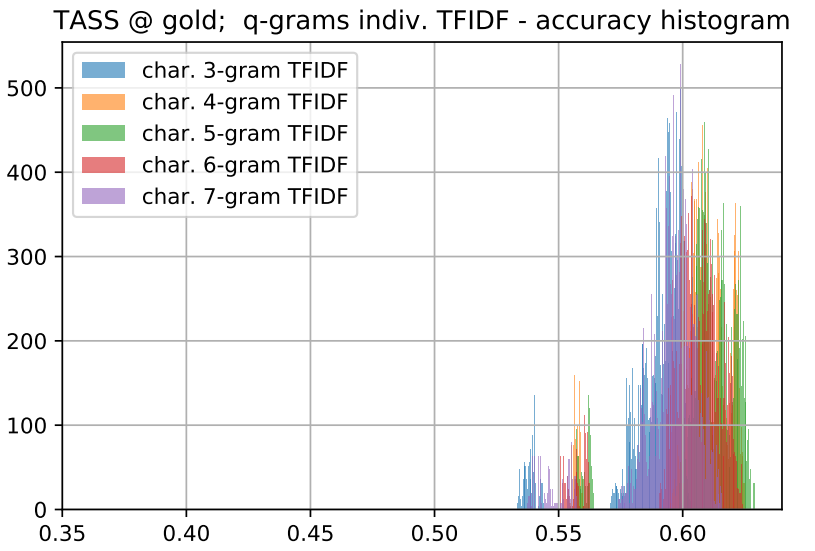
Performance n-words vs q-grams (TASS-gold)





Performance q-grams (TASS-gold)





Our approach: MicroTC, EvoMSA

MicroTC (μTC) is our framework to create text classifiers based on solving the text classification task as a model selection problem.

It selects a competitive configuration from a vast universe of possible ones. Each configuration is composed of a list of text transformations (normalizations and generic transformations), a combination of tokenizers, and a weighting schemes.

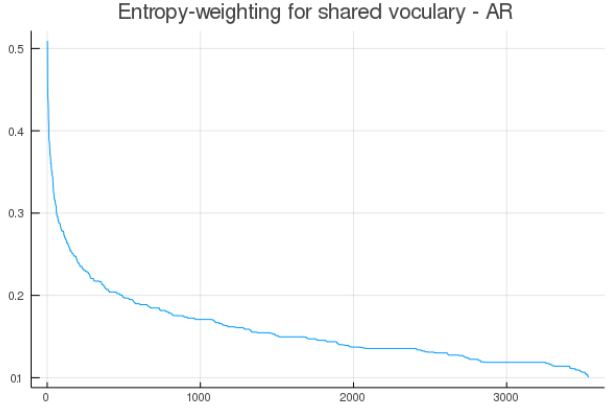
It is free and open source under the Apache 2.0 Licence.

Site: `https://github.com/INGEOTEC/microtc`

Distribution for Arabic language

Arabic vocabularies for both

PAN 17 & PAN 18



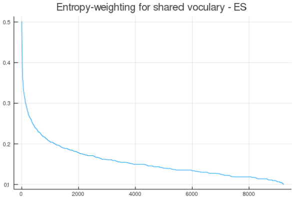
ÁRABE - DETECTADO	INGLÉS	ESPAÑOL	FRANCÉS	ESPAÑOL	INGLÉS	FRANCÉS
الدوري	0.5092			Liga		
حاسة	0.5045			Sentido		
ماما	0.4555			0.4555 mamá		
مش~عارفة~	0.4414			0.4414 está en el buen camino		
ش~قادرة~ا	0.4392			0.4392 u ~ capaz ~ a		
~مش~عارفة	0.4363			0.4363 ~ No ~ sabiendo		
~رونالدو~	0.4320			~ Ronaldo ~		
لدوري~الا	0.4232			0.4232 para League ~ Lala		
فاهمة	0.4114			0.4114 Comprensión		
💖	0.4104			0.4104 💖		
عرص	0.3938			0.3938 libras		
عارفة	0.3884			0.3884 bien informado		
لاعب	0.3874			0.3874 jugadores		
مش~قادرة~	0.3854			0.3854 No puede ~		
رونالدو	0.3848			Ronaldo		
انا~عارفة	0.3803			0.3803 estoy ~ sabiendo		
العرص	0.3780			0.3780 Arp		
~مش~فاهم~	0.3717			Va por un camino distinguido		
~مش~قادرة	0.3717			0.3717 ~ No ~ capaz		
مش~عارفه~	0.3707			0.3707 está en el buen camino		
حبيبتي	0.3684			0.3684 bebé		
💗	0.3656			0.3656 💗		
تشكيلة	0.3602			0.3602 alineación		
مش~عارف~ش	0.3602			0.3602 Malla ~ Aref ~ u		
نا~عارفة~	0.3602			Va por un camino distinguido		

Distribution for English language

English vocabularies for both PAN 17 & PAN 18

Distribution for Spanish language

Spanish vocabularies for both PAN 17 & PAN 18

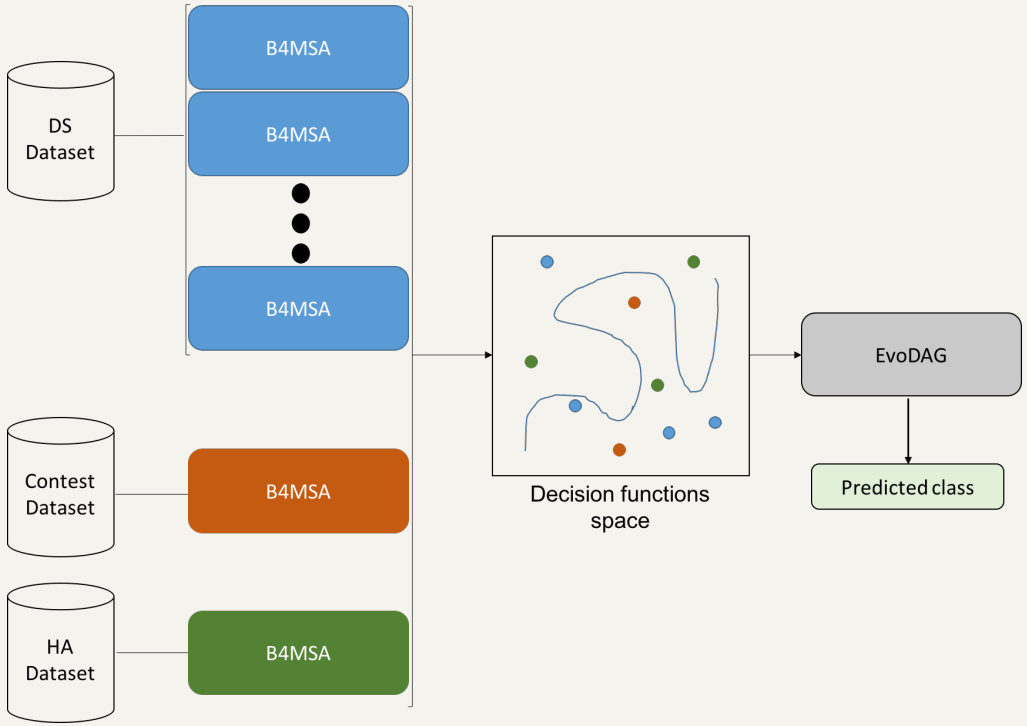


EvoMSA

Features

- Sentiment Analysis System based on B4MSA (μTC) and EvoDAG
- Genetic Programming with semantic operators
- Combines in a GP-based classifier the decision functions come from other classifiers (such as SVMs) trained on different datasets.
- Processes text with language-independent techniques
- Easy to add new knowledge (Distant Supervision approaches)
- Open source and publicly available

Architecture



- A case study of Spanish text transformations for twitter sentiment analysis. ES Tellez, S Miranda-Jiménez, M Graff, D Moctezuma, OS Siordia, EA Villaseñor. Expert Systems with Applications 81, 457-471, 2017.
- A Simple Approach to Multilingual Polarity Classification in Twitter. ES Tellez, S Miranda-Jiménez, M Graff, D Moctezuma, RR Suárez, OS Siordia. Pattern Recognition Letters, 2017.

https://github.com/INGEOTEC/EvoMSA
https://github.com/mgraffg/EvoDAG
https://github.com/INGEOTEC/b4msa

Input datasets

- Distant Supervision
  - Emoticon based and heuristic rules
  - Lexicon based and heuristic rules
  - Hundreds of millions of data (collected)
- Public Human Annotated Datasets
- Contest Datasets

Usage

Install

- Using conda

  $ conda install -c ingeotec microtc

- Using pip - Better to use conda to handle dependencies

  $ pip install microtc

- Installing from github

  $ git clone https://github.com/INGEOTEC/microtc.git

  $ cd microtc

  $ python setup.py install

Step 1: Finding parameters

$ microTC-params -k3 -Smacrorecall -s24 -n24 -o user-profiling.params train.json

- -n Number of cores
- -o Output file
- train.json Training set in json format user-profiling.json is database of exemplars, one json-dictionary per line with text and klass keywords
- -k3 three folds
- -s24 specifies that the parameter space should be sampled in 24 points and then get the best among them

Step 1: Finding parameters (cont.)

$ microTC-params -k3 -Smacrorecall -s24 -n24 -o user-profiling.params train.json

- -n24 let us specify the number of processes to be launch, it is a good idea to set -s as a multiply of -n.
- -o user-profiling.params specifies the file to store the configurations found by the parameter selection process, in best first order
- -S or –score the name of the fitness function (e.g., macrof1, microf1, macrorecall, accuracy, r2, pearsonr, spearmanr)

Step 2: Train

```
$ microtc-train -o user-profiling.model -m
user-profiling.params train.json
```

μTC Usage

Step 2: Train

```
$ microtc-train -o user-profiling.model -m
user-profiling.params train.json
```

Step 3: Predict

```
$ microtc-predict -m user-profiling.model -o
user-profiling-predicted.json test-user-profiling.json
```

μTC Usage

Step 2: Train

```
$ microtc-train -o user-profiling.model -m
user-profiling.params train.json
```

Step 3: Predict

```
$ microtc-predict -m user-profiling.model -o
user-profiling-predicted.json test-user-profiling.json
```

Step 4: Performance

```
$ microtc-perf gold-user.json user-profiling-predicted.json
```

Install

- Using conda

  $ conda install -c ingeotec evomsa

- Using pip - Better to use conda to handle dependencies

  $ pip install evomsa

- Installing from github

  $ git clone https://github.com/INGEOTEC/EvoMSA.git

  $ cd EvoMSA

  $ python setup.py install

EvoMSA Usage

Train

$ EvoMSA-train -n2 -o evomsa.model train.json

- -n Number of cores
- -o Output file
- train.json Training set in json format

EvoMSA Usage

Train

$ EvoMSA-train -n2 -o evomsa.model train.json

- -n Number of cores
- -o Output file
- train.json Training set in json format

train.json

{"klass"= "positive", "text": "good life" }
{"klass"= "neutral", "text": "the computer" }
{"klass"= "negative", "text": "I have a headache" }

EvoMSA Usage

test.json

```
{"text": "Talking about text categorization" }
{"text": "Research public available" }
```

EvoMSA Usage

test.json

{"text": "Talking about text categorization" }
{"text": "Research public available" }

Predict

$ EvoMSA-predict -n2 -o out.json -m evomsa.model test.json

- -m EvoMSA model
- test.json Dataset to be predicted

EvoMSA with Extra-Knowledge

External dataset

```
$ EvoMSA-train -n2 -o evomsa.model train.json external.json
```

Exogenous models - no related to $\mathcal{X}$

```
$ EvoMSA-train -n2 –exogenous-model ex.model -o
evomsa.model train.json
```

Exogenous variables - no related to $\mathcal{X}$

```
$ EvoMSA-train -n2 –exogenous ex.json -o evomsa.model
train.json
```

Conclusions

- Systems tailored to informal text
- A multilingual multi-domain text classifier
  μTC
  - easy to add new text transformations into configuration space
  - easy to add weighting schemes into configuration space
- Easy to add knowledge into EvoMSA Framework

Questions

Questions or comments?

http://ingeotec.mx

https://github.com/INGEOTEC

Thank
You!