

# Detección de robots en Twitter

## Aplicaciones de la Lingüística Computacional

Jesús Alonso<sup>1</sup>

<sup>1</sup>Máster IARFID  
Universitat Politècnica de València

3 de julio de 2014

- Twitter: red social y herramienta de microblogging lanzada en 2006
- Robots en Twitter:
  - Benignos: publican, por ejemplo, noticias y actualizaciones de blogs
  - Malignos: difunden spam
- Necesarias técnicas que distingan robots y humanos

- Estudio basado en 3 tipos de características:
  - Entropía: detecta regularidad en la frecuencia de tuiteo
  - Detección de spam: la mayoría de tuits que contienen spam son publicados por robots
  - Propiedades de la cuenta
    - 1 Ratio de URLs
    - 2 Método de tuiteo (web, API, aplicaciones móviles,...)
    - 3 Ratio seguidores / amigos
    - 4 Seguridad de los enlaces
    - 5 Verificación de la cuenta
    - 6 Fecha de registro
    - 7 Ratio de hashtags
    - 8 Ratio de menciones
- Clasificación mediante bosque automático
- Resultados:  $> 90\%$  de clasificación correcta

- Proporcionados por Autoritas Consulting
- Composición:

Proyecto	Tuits	Usuarios	Robots	Porcentaje robots
gasofa	339734	135167	50	0.04 %
castellon	1824	1178	86	7.30 %
senado	64014	29018	622	2.14 %
comite	323692	122296	393	0.32 %
rato	323692	122296	108	0.09 %

- Bolsas de palabras del contenido de los tuits. Utilizando las 400 palabras (sin stopwords) más comunes en los tuits de cada proyecto
- Bolsas de palabras de las biografías. Utilizando las 20 palabras (sin stopwords) más comunes en las biografías de cada proyecto
- Características adicionales:
  - 1 N° de tuits
  - 2 Ratio de URLs
  - 3 Ratio de hashtags
  - 4 Ratio de menciones
  - 5 Ratio amigos / seguidores
  - 6 N° de días desde el registro de la cuenta

- Grupos de características
  - Bolsas de palabras de los contenidos
  - Bolsas de palabras de los contenidos + Bolsas de palabras de las biografías
  - Bolsas de palabras de los contenidos + Características adicionales
  - Bolsas de palabras de los contenidos + Bolsas de palabras de las biografías + Características adicionales
  - Sólo características adicionales
- Método de clasificación: Support Vector Machines

# Entrenamiento y test

- Se han eliminado cuentas de humanos de los conjuntos de entrenamiento debido al reducido número de robots.

Proyecto	Entrenamiento			Test			Total
	Total	Humanos	Robots	Total	Humanos	Robots	
gasofa	1040	1000	40	35127	35117	10	135167
castellon	858	792	66	320	300	20	1178
senado	12472	12000	472	8546	8396	150	29018
comite	10293	10000	293	32003	31903	100	122296
rato	2588	2500	88	32208	32188	20	122296

# Validación cruzada

- Consistente en entrenar con los datos de un proyecto y validar con los de los proyectos restantes
- Medidas de evaluación en media
- Entrenamiento:

Proyecto	Entrenamiento		
	Total	Humanos	Robots
gasofa	1300	1250	50
castellon	1178	1092	86
senado	16172	15550	622
comite	10218	9825	393
rato	2808	2700	108

- Test: conjunto total de datos de los 4 proyectos no considerados en el entrenamiento



- Precision:

$$\frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

- Recall:

$$\frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

- F-Measure:

$$2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

# Resultados

Proyecto	Características	Precision	Recall	F-Measure
gasofa	Cont.	100 %	100 %	1
	Cont. + bio.	90 %	90 %	0.9
	Cont. + adic.	100 %	90 %	0.95
	Cont. + bio. + adic.	100 %	90 %	0.95
	Adic.	55.56 %	100 %	0.71
castellon	Cont.	100 %	95 %	0.97
	Cont. + bio.	100 %	100 %	1
	Cont. + adic.	100 %	100 %	1
	Cont. + bio. + adic.	100 %	100 %	1
	Adic.	62.5 %	25 %	0.36
senado	Cont.	79.17 %	63.33 %	0.7
	Cont. + bio.	76.03 %	61.33 %	0.68
	Cont. + adic.	68.94 %	60.67 %	0.65
	Cont. + bio. + adic.	80.33 %	65.33 %	0.72
	Adic.	68.92 %	34 %	0.46
comite	Cont.	64.03 %	89 %	0.74
	Cont. + bio.	67.36 %	97 %	0.8
	Cont. + adic.	59.21 %	90 %	0.71
	Cont. + bio. + adic.	63.95 %	94 %	0.76
	Adic.	36.14 %	73 %	0.48
rato	Cont.	24.14 %	70 %	0.36
	Cont. + bio.	15.60 %	85 %	0.26
	Cont. + adic.	12.21 %	80 %	0.21
	Cont. + bio. + adic.	15.46 %	75 %	0.26
	Adic.	5.67 %	55 %	0.1

- Validación cruzada

Características	Precision	Recall	F-Measure
Cont.	11.45 %	33.88 %	0.13
Cont. + bio.	14.9 %	27.94 %	0.12
Cont. + adic.	16.14 %	26.03 %	0.12
Cont. + bio. + adic.	17.67 %	28.72 %	0.15
Adic.	10.66 %	27.61 %	0.14

- Los datos de los proyectos gasofa y castellon son anómalos por su elevada tasa de acierto
- Las características adicionales por sí solas aportan los resultados más bajos
- Los resultados de la validación cruzada (más cercana a un entorno real) son significativamente inferiores a los del estudio hecho para cada proyecto
- Trabajo futuro:
  - Nuevas características: frecuencias de actualización, información del grafo social,...
  - Métodos alternativos de clasificación
  - Réplica del análisis mencionado como trabajo previo