

# EARLY FUSION EN CARACTERÍSTICAS TRADICIONALES Y PROFUNDAS PARA LA DETECCIÓN DE IRONÍA EN TWITTER



CIMAT

Hairo Ulises Miranda Belmonte

Adrián Pastor López-Monroy

Centro de Investigación en Matemáticas A.C.

## Resumen

El presente trabajo es un resumen realizado para otro foro con el fin de abordar la tarea de detección de ironía mediante características tradicionales y profundas. Se utiliza la corpora del concurso IroSvA en detección de ironía de variantes al español (México, España y Cuba). Para la detección de la ironía se concatenan tres vectores distintos de características y se introducen a un clasificador (i.e., *early fusion*), capturando información de: n-gramas, *embeddings* y unidades ocultas de una red recurrente. Se contrastan los resultados respecto al desempeño individual de cada espacio de características. Como resultado se supera el *baseline* del concurso. Los resultados presentan fuerte evidencia de la utilidad en la representación de los textos para la identificación de la ironía, independientemente de la variante del idioma.

## Bag of Words

**Bag of Words (BoW)** : extrae características tradicionales basadas en contar palabras. Se utiliza el conocido esquema de ponderación de Frecuencia TF-IDF:

$$\text{TF-IDF} = TF * IDF$$

$$IDF(t) = \log_e \frac{\# \text{ total de documentos}}{\# \text{ de documentos con en el trmino } t}$$

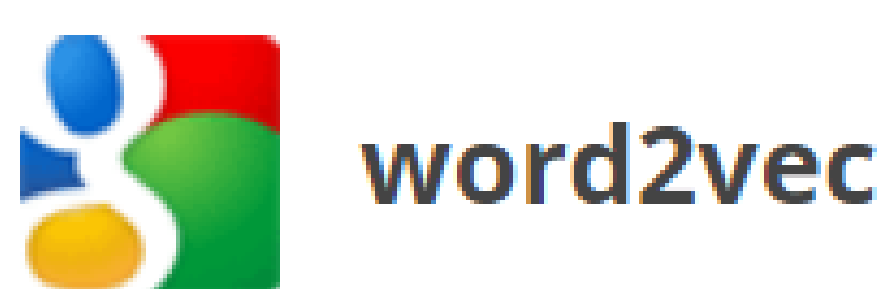
$$TF(t) = \frac{\# \text{ de veces que el trmino } t \text{ aparece en un documento}}{\# \text{ total de trminos en el documento}}$$

En este trabajo se le refiere como Bolsa de Términos (BoT) a la generalización de características, i.e., n-gramas a nivel de caracteres y palabras.

## Word Embeddings

**Embeddings**: vectores de palabras aprendidas usando Word2Vec. Se evalúa *embeddings* de modelos pre-entrenados:

i) Embedding de 300 dimensiones de Google [2]



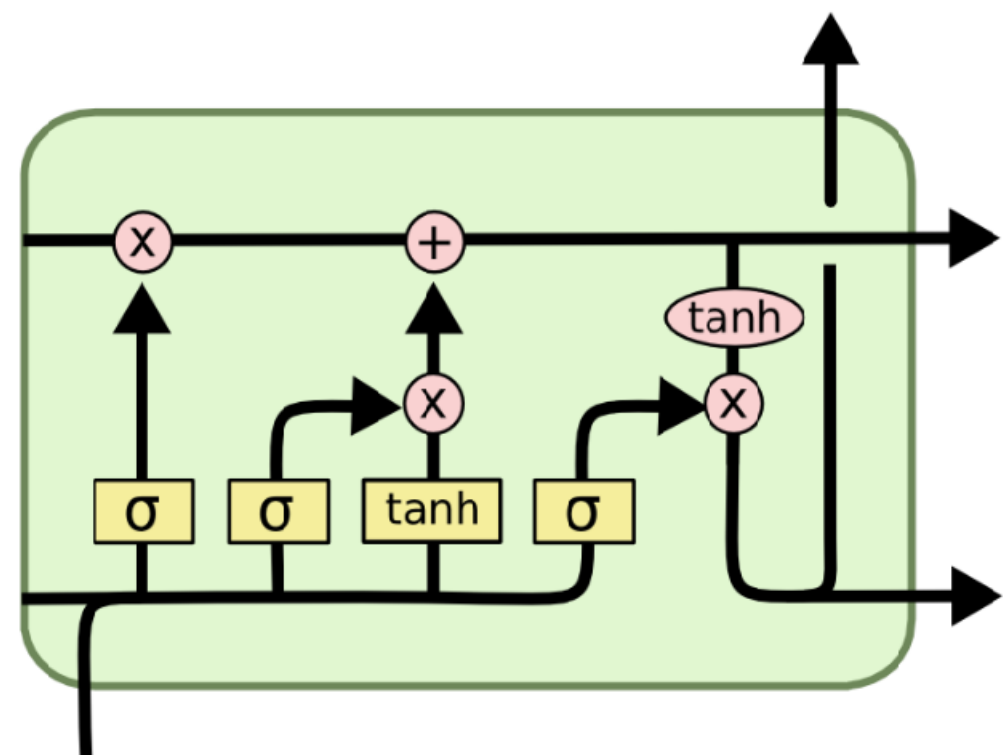
ii) Embedding de 400 dimensiones de Twitter [1].



La idea es representar a los textos promediando los vectores de características para cada palabra.

## Long Short Term Memory

**Long Short Term Memory Neural Networks (LSTMs)**: captura dependencias a largo plazo en los textos.



Se evalua LSTM de dos maneras:

i) se codifica el tweet con LSTM y se utiliza la capa oculta en el último paso para alimentar una máquina de vectores de soporte (SVM).

ii) se hace uso de la función *sigmoide* para clasificar con el LSTM en la capa de salida.

## Datos: Irosva (*Irony Detection in Spanish Variants*)

- La corpora consiste en 9000 mensaje cortos sobre diferentes tópicos escritos en Español -3000 de Cuba, 3000 de México y 3000 de España- con aproximadamente el 80% (2400 twitters), de cada corpus, utilizados como datos de entrenamiento y el 20% (600 twitters) como conjunto de prueba.



## Hiperparámetros

- SVM**: kernel lineal, párametro de penalización optimizado.
- Red Neuronal**: 1 capa oculta con 512 nodos, Dropout de 0.25, batch size de 50 y optimizador Adam
- LSTM**: 256 unidades ocultas, batch size de 50 y optimizador Adam
- BoT (n-gramas)**: con 1 a 5 gramas a nivel palabra y carácter. Se selecciona empíricamente los mejores *n-gramas* para cada tamaño de *n* con la métrica de  $\chi^2$ . La configuración particular que se tiene es: i) 5k unigramas, 3k bigramas y 1k trigramas a nivel palabra, y ii) 5k 3-gramas, 5k 4-gramas y 5k 5-gramas a nivel carácter.

## Resultados

**Tabla 1. Evaluación detección de Ironía. Resultados muestran métrica F-1 *average* de distintas representaciones que son ajustadas por un SVM**

Representation	México	España	Cuba
BoT	<b>64.27</b>	<b>68.50</b>	60.10
W2V-Google	62.53	66.00	60.64
W2V-Twitter	62.60	65.25	60.65
LSTM	63.30	66.75	<b>60.85</b>

**Tabla 2. Evaluación detección de Ironía. Resultados muestran métrica F-1 *average* de distintas representaciones que son ajustadas por una Red Neuronal**

Representation	México	España	Cuba
BoT	<b>60.47</b>	<b>67.31</b>	<b>57.90</b>
W2V-Google	59.75	61.35	55.57
W2V-Twitter	55.65	62.36	52.82
LSTM	60.21	63.25	56.41

**Tabla 3. Evaluación detección de Ironía. Resultados muestran métrica F-1 *average* de distintas representaciones combinadas que son ajustadas por un SVM.**

Representation	México	España	Cuba
BoT+W2V-Google	65.40	<b>69.77</b>	61.88
BoT+W2V-Twitter	63.58	67.68	62.15
BoT+LSTM	64.57	68.66	61.81
BoT+W2V-Google+LSTM	66.12	69.04	62.33
BoT	64.27	68.50	60.10
BoT(n-grams)	65.01	68.79	62.05
BoT(n-grams)+W2V-Google+LSTM	<b>67.09</b>	64.49	<b>65.96</b>

## Conclusión

- Los patrones locales de los n-gramas, *embeddings* y dependencias largas entre palabras con el LSTM, ayudan a mejorar el rendimiento de clasificación.
- Usando estos atributos, el clasificador mantiene buenas tasas de clasificación debido a la relación entre diferentes tipos de características.
- Se muestran mejores resultados experimentales que el estándar BoT, Word2Vec y LSTM, que han demostrado ser útiles pero mejores cuando se usan conjuntamente.

## Referencias

- Godin, F., Vandersmissen, B., De Neve, W., Van de Walle, R.: Multimedia lab @ acl wnwt ner shared task: Named entity recognition for twitter microposts using distributed word representations. pp. 146–153 (2015)
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. pp. 3111–3119 (2013)