



UNIVERSIDAD
DE GRANADA

Procesamiento del Lenguaje Natural y Análisis de Sentimientos

Carlos Gómez-Rodríguez

carlos.gomez@ucdc.es

Eugenio Martínez Cámará

emcamara@decsai.ugr.es



Índice

- Introducción.
- Representaciones vectoriales del lenguaje.
- Etiquetación (part-of-speech tagging).
- Análisis sintáctico.
- Análisis de sentimientos.

Índice

- Introducción.
- Motivación.
- ¿Qué es el Análisis de Opiniones?
- Análisis de Opiniones en Twitter.
- Ejemplos en la literatura.
- Código.



**¿Te gustaría conocer el nivel de
satisfacción de tus clientes?**

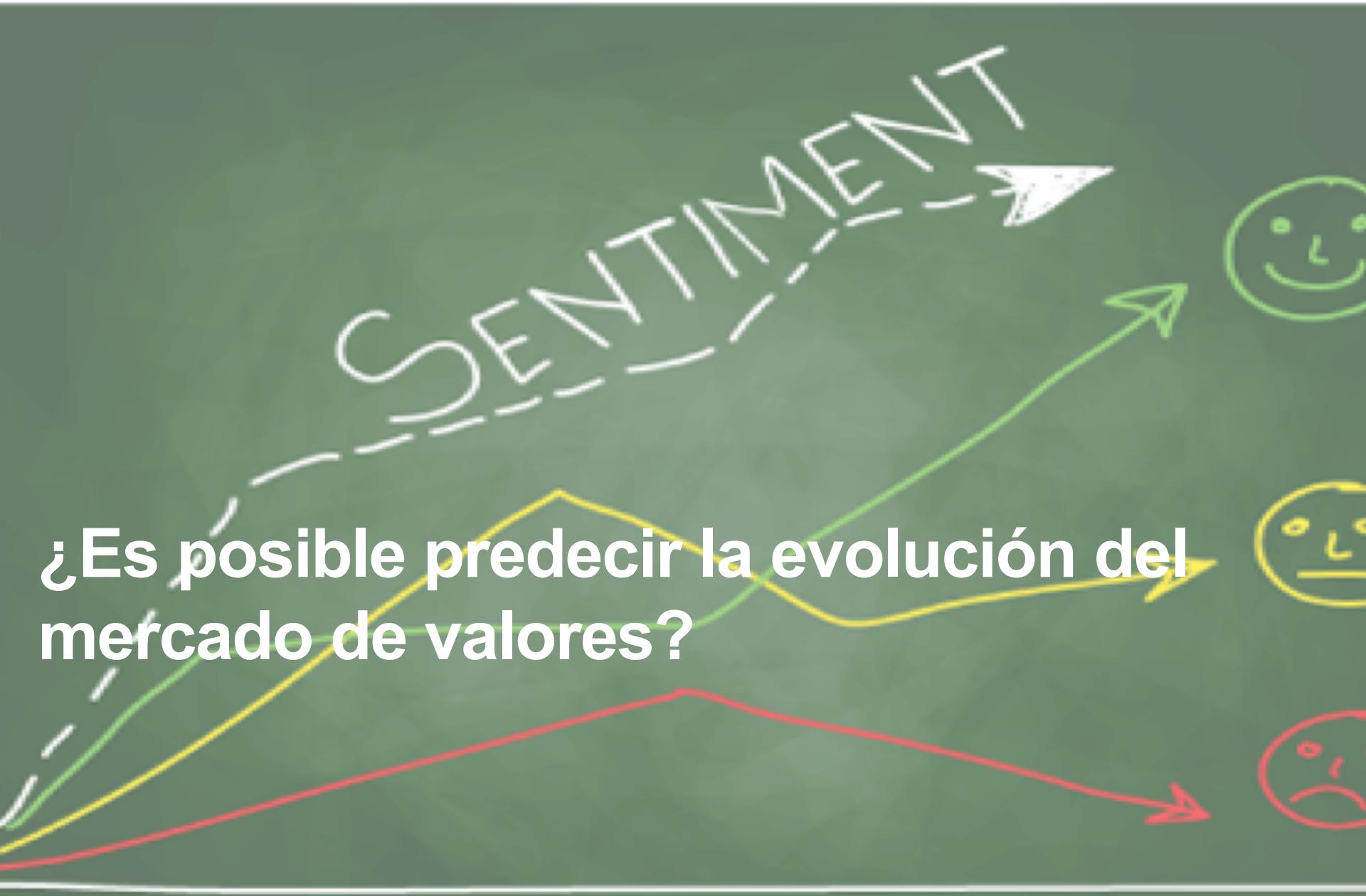


**¿Te gustaría conocer la opinión de tus
votantes?**

A network graph illustrating connections between various 1970s rock bands. The nodes are represented by colored circles, and the connections by lines of different colors. The bands include America, Eagles, Elton John, Dire Straits, 10cc, Phil Collins, Roger Hodgson, Styx, Electric Light Orchestra, Supertramp, Barclay James Harvest, Manfred Mann's Earth Band, Saga, Peter Gabriel, Genesis, and The Moody Blues. The network shows dense clusters of connections between artists like The Moody Blues, Dire Straits, and Genesis.

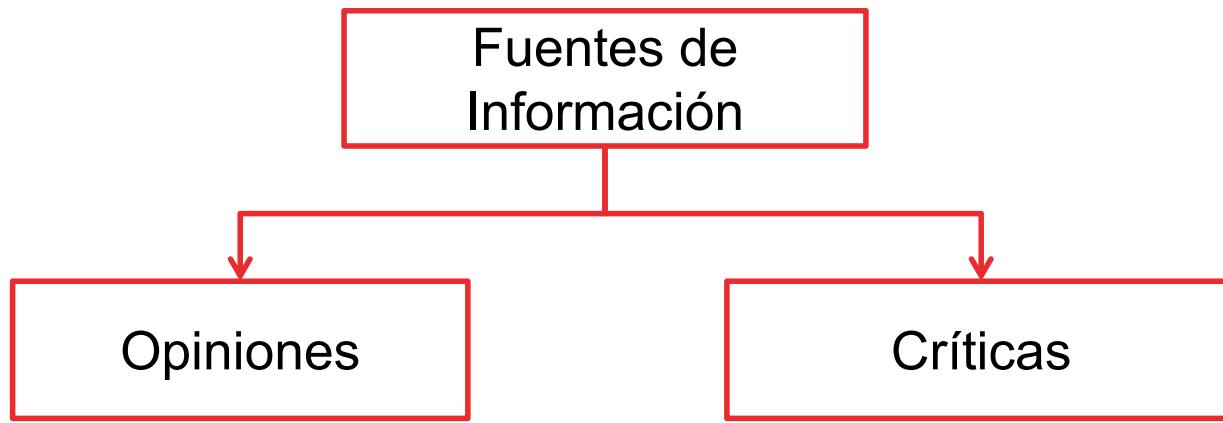
¿Te gustaría recomendar música en función de las opiniones de las canciones?

**¿Es posible predecir la evolución del
mercado de valores?**



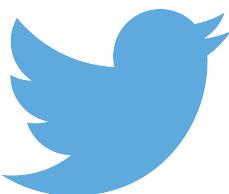
Motivación

- Diariamente tomamos decisiones.
- Tomar decisiones no es sencillo.
- Necesitamos información de experiencias similares.



Motivación

- Al principio, las opiniones se transmitían de boca a boca.
- Pero hoy, la Web 2.0 ha proporcionado nuevos canales de comunicación donde las opiniones se distribuyen a gran velocidad.



Motivación

- Se requiere de sistemas de recuperación, extracción, clasificación y resumen de opiniones sobre entidades, que pueden ser un hotel, un monumento, un restaurante, un producto electrónico...
- La industria, los gobiernos, los partidos políticos, los hospitales quieren conocer la opinión de sus clientes lo más pronto posible, con el fin de utilizarla como realimentación para mejorar sus servicios.

The background features three overlapping circles. The largest circle is teal and contains two white circles representing eyes. A white curved line below them forms a smile. The middle circle is a lighter shade of teal and overlaps the bottom of the first circle. The smallest circle is dark teal and overlaps the top of the first circle, containing two white circles and a white curved line forming a more pronounced smile.

¿Qué es el análisis de opiniones?

¿Qué es el análisis de opiniones?

- El análisis de opiniones (AO) es una tarea del Procesamiento del Lenguaje Natural (PLN).
- PLN: Conjunto de técnicas computacionales destinadas al desarrollo de software capaz de entender y/o generar lenguaje natural.



¿Qué es el análisis de opiniones?

Aplicaciones PLN

Clásicas

"Nuevas"

Traducción
automática

Desambi-
guación

Recuperación
de
información

Resolución
de
preguntas

Extracción
de
información

Generación
de LN

Análisis
de
Opiniones

Generación
de
resúmenes

Simplifi-
cación
de texto

¿Qué es el análisis de opiniones?

Aplicaciones PLN

Clásicas

"Nuevas"

Traducción
automática

Desambi-
guación

Recuperación
de
información

Resolución
de
preguntas

Extracción
de
información

Generación
de LN

Análisis
de
Opiniones

Generación
de
resúmenes

Simplifi-
cación
de texto

¿Qué es el análisis de opiniones?

Tratamiento computacional de la opinión, el sentimiento y la subjetividad en un texto.

(Pang & Lee, 2008)

Conjunto de técnicas computacionales para la extracción, clasificación, entendimiento y evaluación de opiniones expresadas en fuentes de noticias, comentarios en redes sociales y otros contenidos generados por usuarios.

(Cambria & Hussain, 2012)

¿Qué es el análisis de opiniones?

- Objeto de estudio: **la opinión**.
- ¿Qué es una opinión? La evidencia o expresión de un estado de ánimo (Quirk et al. 1985).
- Las opiniones, las emociones y las evaluaciones son expresiones de estados de ánimo.

¿Qué es el análisis de opiniones?

Autor: Pedro Reyes

Fecha: 20/10/2018

La semana anterior pasé unos días en Jaén con mi esposa. Elegimos una habitación bonita en el centro de la ciudad. Nos ha quedado un buen recuerdo del personal del hotel. Debo resaltar que la cama era muy confortable. Por el contrario, mi mujer no pudo dormir porque, según ella, la almohada era muy mala. No nos gustó el baño porque encontramos algunos desagradables pelos cuando nosotros entramos en la habitación. Para terminar, debo decir que lo mejor del hotel fue el desayuno, un día más y reviento.

¿Qué es el análisis de opiniones?

Autor: Pedro Reyes

Fecha: 20/10/2018

La semana anterior pasé unos días en Jaén con mi esposa.

Elegimos una habitación bonita en el centro de la ciudad. Nos ha quedado un buen recuerdo del personal del hotel. Debo resaltar que la cama era muy confortable. Por el contrario, mi mujer no pudo dormir porque, según ella, la almohada era muy mala. No nos gustó el baño porque encontramos algunos desagradables pelos cuando nosotros entramos en la habitación. Para terminar, debo decir que lo mejor del hotel fue el desayuno, un día más y reviento.

¿Qué es el análisis de opiniones?

Autor: Pedro Reyes

Fecha: 20/10/2018

La semana anterior pasé unos días en Jaén con mi esposa.

Elegimos una habitación bonita en el centro de la ciudad. Nos ha quedado un buen recuerdo del personal del hotel. Debo resaltar que la cama era muy confortable. Por el contrario, mi mujer no pudo dormir porque, según ella, la almohada era muy mala. No nos gustó el baño porque encontramos algunos desagradables pelos cuando nosotros entramos en la habitación. Para terminar, debo decir que lo mejor del hotel fue el desayuno, un día más y reviento

¿Qué es el análisis de opiniones?

Autor: Pedro Reyes

Fecha: 20/10/2018

La semana anterior pasé unos días en Jaén con mi esposa.

Elegimos una habitación bonita en el centro de la ciudad. Nos ha quedado un buen recuerdo del personal del hotel. Debo resaltar que la cama era muy confortable. Por el contrario, mi mujer no pudo dormir porque, según ella, la almohada era muy mala. No nos gustó el baño porque encontramos algunos desagradables pelos cuando nosotros entramos en la habitación. Para terminar, debo decir que lo mejor del hotel fue el desayuno, un día más y reviento

¿Qué es el análisis de opiniones?

Autor: Pedro Reyes

Fecha: 20/10/2018

La semana anterior pasé unos días en Jaén con mi esposa.

Elegimos una habitación bonita en el centro de la ciudad. Nos ha quedado un buen recuerdo del personal del hotel. Debo resaltar que la cama era muy confortable. Por el contrario, mi mujer no pudo dormir porque, según ella, la almohada era muy mala. No nos gustó el baño porque encontramos algunos desagradables pelos cuando nosotros entramos en la habitación. Para terminar, debo decir que lo mejor del hotel fue el desayuno, un día más y reviento

¿Qué es el análisis de opiniones?

Autor: Pedro Reyes

Fecha: 20/10/2018

La semana anterior pasé unos días en Jaén con mi esposa.

Elegimos una habitación bonita en el centro de la ciudad. Nos ha quedado un buen recuerdo del personal del hotel. Debo resaltar que la cama era muy confortable. Por el contrario, mi mujer no pudo dormir porque, según ella, la almohada era muy mala. No nos gustó el baño porque encontramos algunos desagradables pelos cuando nosotros entramos en la habitación. Para terminar, debo decir que lo mejor del hotel fue el desayuno, un día más y reviento

$$(e_i, a_{ij}, p_{ijhl}, h_h, t_l)$$

(Liu, 2012)

¿Qué es el análisis de opiniones?

Tareas propias del AO:

1. Extracción y categorización de entidades.
2. Extracción y categorización de aspectos.
3. Extracción del opinador.
4. Extracción del momento en el que tiene lugar la opinión.
5. Clasificación de la polaridad de la opinión: Opiniones explícitas e implícitas.
6. Generación de la quíntupla de la opinión.

¿Qué es el análisis de opiniones?

(hotel, estilo, positivo, Pedro Reyes y esposa, 20/10/2018)

(hotel, personal, positivo, Pedro Reyes y esposa, 20/10/2018)

(hotel, cama, positivo, Pedro Reyes, 20/10/2018)

(hotel, almohada, negativo, esposa de Pedro Reyes,
20/10/2018)

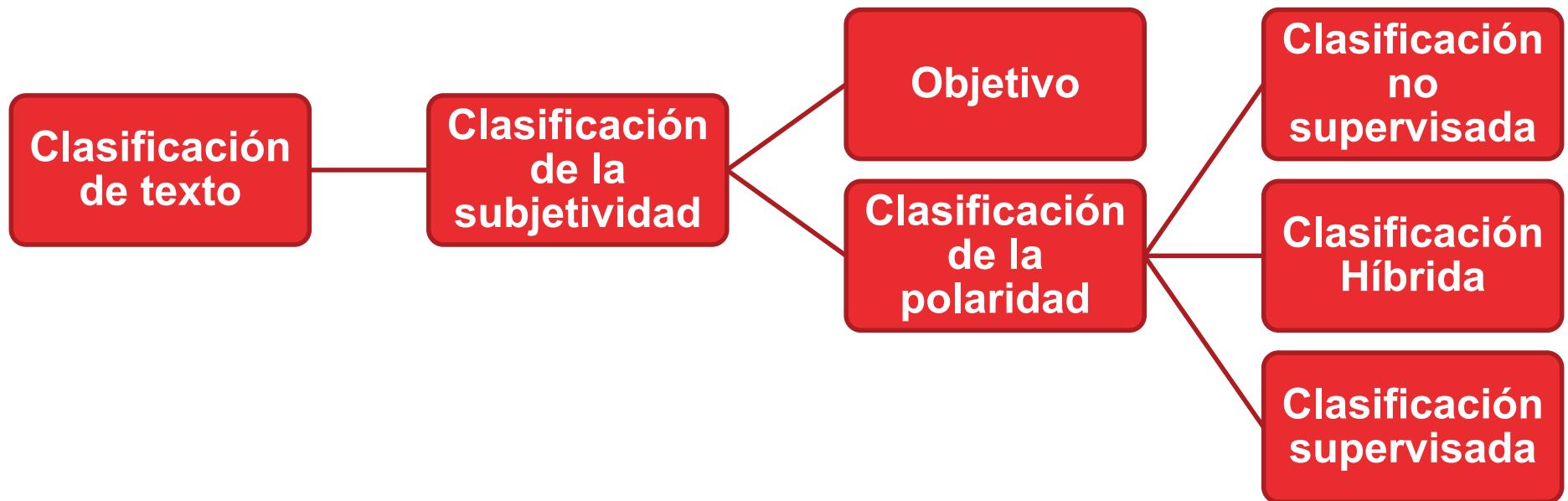
(hotel, desayuno, positivo, Pedro Reyes, 20/10/2018)

¿Qué es el análisis de opiniones?

Tareas propias del AO

1. Extracción y categorización de entidades.
2. Extracción y categorización de aspectos.
3. Extracción del opinador.
4. Extracción del momento en el que tiene lugar la opinión.
5. **Clasificación de la polaridad de la opinión: Opiniones explícitas e implícitas.**
6. Generación de la quíntupla de la opinión.

¿Qué es el análisis de opiniones?



¿Qué es el análisis de opiniones?

Clasificación

- Supervisada:

- Requiere de un corpus/conjunto de datos anotado
- Principalmente basado en el uso de algoritmos de aprendizaje automático: SVM, KNN, Naïve Bayes, Random Forest, Neural Networks, Logistic Regression...
- Tendencia actual: Redes neuronales (Deep Learning)

- No supervisada:

- No requiere de un corpus/conjunto de datos anotado
- Métodos: Lista de palabras de opinión, patrones sintácticos, clustering

- Híbrida:

- Combinación de los métodos anteriores. Por ejemplo:
 - Sistema no supervisado para anotar un pequeño conjunto de elementos con el que construir un clasificador supervisado
 - Multi-clasificador de sistema de clasificación supervisado uy no supervisado

¿Qué es el análisis de opiniones?

Niveles de análisis

Documento

Oración

Entidad/Aspecto

¿Qué es el análisis de opiniones?

• A nivel de documento:

- Objetivo: La clasificación de la orientación de la opinión general expresada en un documento.
- Entrada: Un documento (largo / corto (tuit)).
- Asunción: El documento está formado por opiniones explícitas o implícitas.
- Salida: La opinión del documento.

• A nivel de oración:

- Objetivo: La clasificación de la opinión expresada en una oración.
- Entrada: Un documento / una oración.
- Asunción: El documento está formado por opiniones explícitas o implícitas.
- Salida:
 - Documento: Combinación de la orientación de las opiniones de las oraciones que componen el documento.
 - Oración: La orientación de la opinión expresada en la oración.

¿Qué es el análisis de opiniones?

- A nivel de aspecto:
 - Objetivo: La clasificación de la opinión sobre una entidad o aspecto.
 - Entrada: Una oración/un conjunto de oraciones/un documento.
 - Asunción: Las entidades/aspectos están dados.
 - Salida:
 - El conjunto de entidades y aspectos de las entidades.
 - La polaridad de la opinión sobre cada entidad y aspecto.



¿Qué es el análisis de opiniones? Corpora

Textos largos

- Customer Review Dataset (2004)
 - Documentos: 500.
 - Anotación: A nivel de documento.
 - Dominio: Productos electrónicos (5).
- SFU Review Corpus (2004)
 - Documentos : 400.
 - Anotación: A nivel de documento.
 - Dominio: Opiniones de productos (8).
 - Versión en español: 2009.
- MPQA (2005):
 - Documentos: 535.
 - Anotación: A nivel de oración.
 - Dominio: Noticias.
 - Última versión: 2015.

Textos cortos (Twitter)

- Stanford Twitter Corpus (2009)
 - Documentos:
 - Entrenamiento: 1.600.000 (balanceado). Supervisión distante (*distant supervision*).
 - Test: 182 positivas 177 negativas. Anotado manualmente.
 - Anotación: A nivel de documento.
 - Dominio: General.
 - Idioma: Inglés.
- SemEval Twitter corpus:
 - Desde 2013, la organización publica un nuevo corpus.
 - Idioma: Inglés.
- General Corpus of TASS
 - Documentos:
 - Entrenamiento : 7.219.
 - Test: 60.798.
 - Idioma: Español.
 - Anotación: A nivel de documento. 6 clases.
- InterTASS (2017; 2018)
 - Documentos:
 - Entrenamiento: 1.008.
 - Desarrollo: 506.
 - Test: 1.899.
 - Idioma: Español.
 - Anotación: A nivel de documento. 4 clases.

¿Qué es el análisis de opiniones? Recursos

- **Listas de palabras de opinión:** Listas de palabras positivas y negativas
 - Lista de Bing Liu:
 - Idioma: Inglés
 - Palabras: sobre 6.800 palabras
 - iSOL:
 - Idioma: Español
 - Palabras: 2.509 positivas; 5.626 negativas
 - GermanPolarityWords
 - Idioma: Alemán
 - Palabras: 10.000

¿Qué es el análisis de opiniones? Recursos

- Lexicones: Conjunto de palabras relacionadas por propiedades léxicas
 - SentiWordNet
 - Basado en WordNet.
 - Cada significado/sentido (*synet*) tiene asociado tres valores de polaridad: positivo, negativo y neutro.
 - La suma de los tres valores es igual a 1.
 - Q-WordNet
 - Basado en WordNet.
 - Cada significado/sentido (*synet*) tiene asociado dosvalores de polaridad: positivo y negativo.

¿Qué es el análisis de opiniones?

Retos relacionados con el AO

- Extracción de información
- Extracción de relaciones semánticas
- Resolución de la coreferencia y anáfora
- Desambiguación
- Tratamiento de la negación
- Conocimiento implícito
- Adaptación al dominio
- Tratamiento de ironía y sarcasmo

Mucho trabajo aún por hacer

Análisis de Opiniones en Twitter



AO en Twitter

- La investigación en AO en Twitter avanzó en paralelo al aumento del uso de *microblogs*.
- Los primeros estudios estuvieron más relacionados con la sociología que con la informática
- Boca a boca → Boca a boca electrónico o en línea (EWOM / OWOM)
- Martínez-Cámara, E., Martín-Valdivia, M.T., Ureña-López, L. A. & Montejo-Ráez A. (2014) Sentiment Analysis in Twitter. *Journal of Natural Language Engineering* 20(1):1-28.

AO en Twitter

Textos Largos

- La longitud no está limitada.
- Compuestos por varias oraciones.
- Pueden expresar varios mensajes o ideas.
- El estilo formal es el más común.
- Suele haber un uso correcto de la gramática.
- Hay contexto suficiente para determinar la información del discurso.

Textos cortos

- La longitud es limitada (tuits: 280 caracteres).
- Normalmente compuesto por un máximo de 3 oraciones.
- Usualmente sólo se expresa una oración.
- Prevalece el estilo informal.
- Uso pobre de la gramática.
- Dispersión de los datos. Hay que buscar redundancia con técnicas lingüísticas (implicación textual).
- No suele haber contexto.

Ejemplos en la Literatura

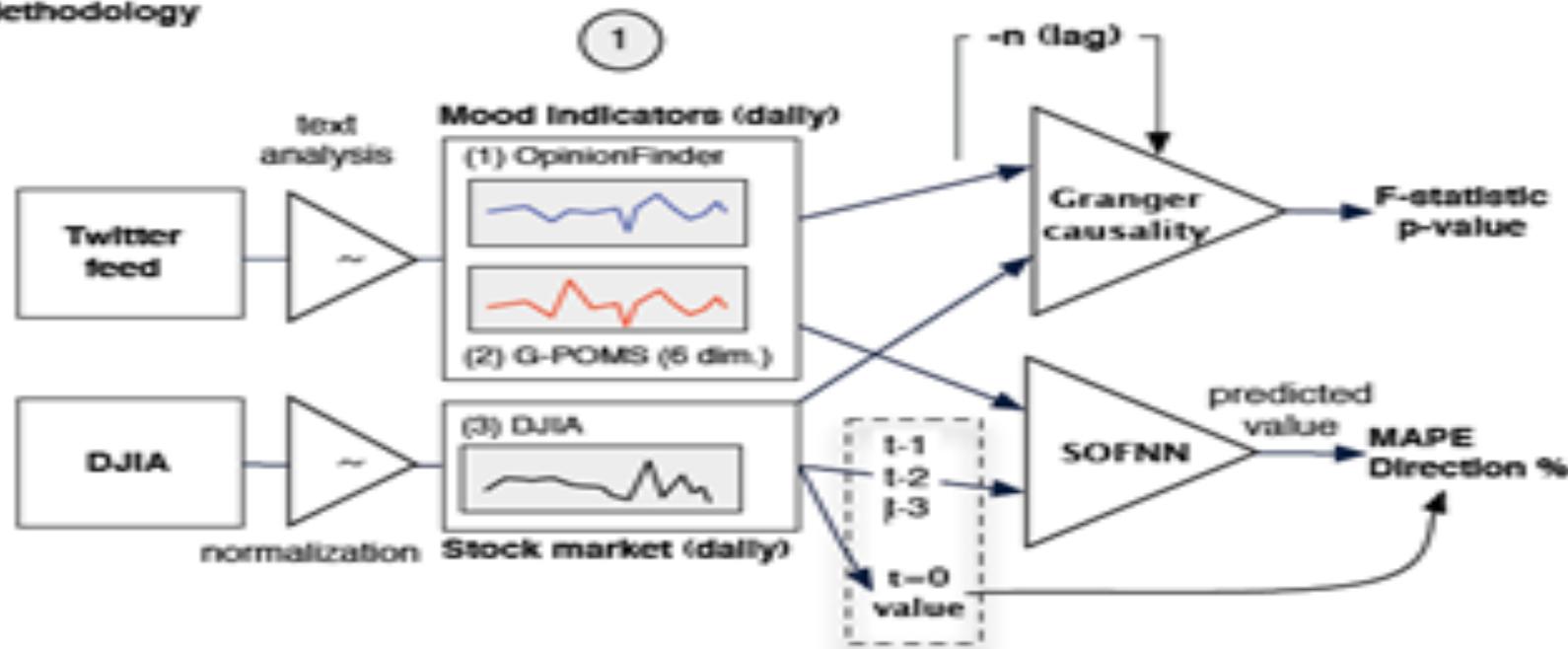


Predicción de mercados de valores

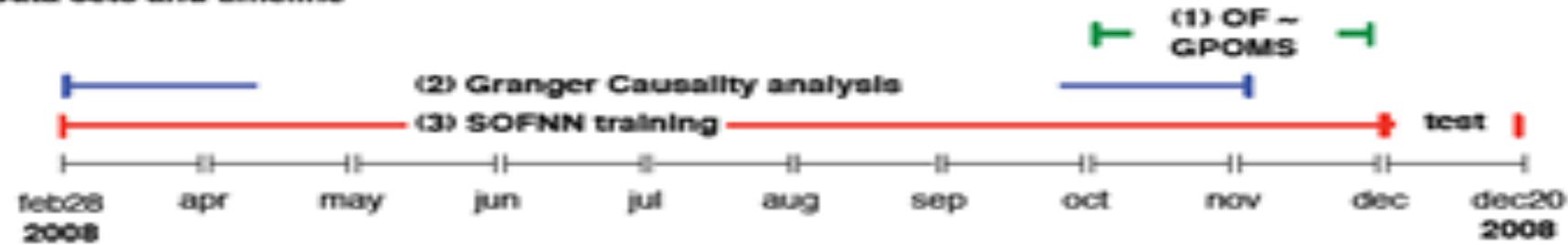
- Se sabe que el estado de ánimo influye en la capacidad de toma de decisiones.
- ¿Puede el estado emocional de una sociedad influir en sus decisiones colectivas?
- ¿Es factible considerar el estado emocional o el estado de opinión de una sociedad como un indicador económico?
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8.

Predicción de mercados de valores

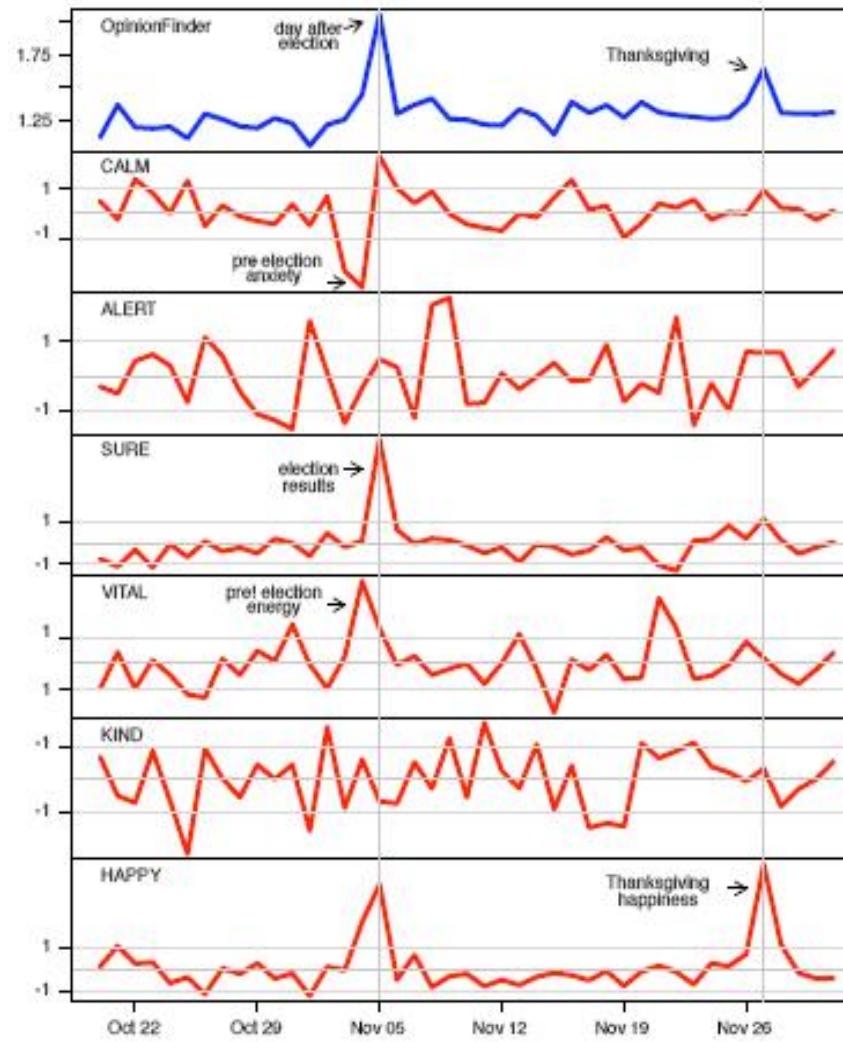
Methodology



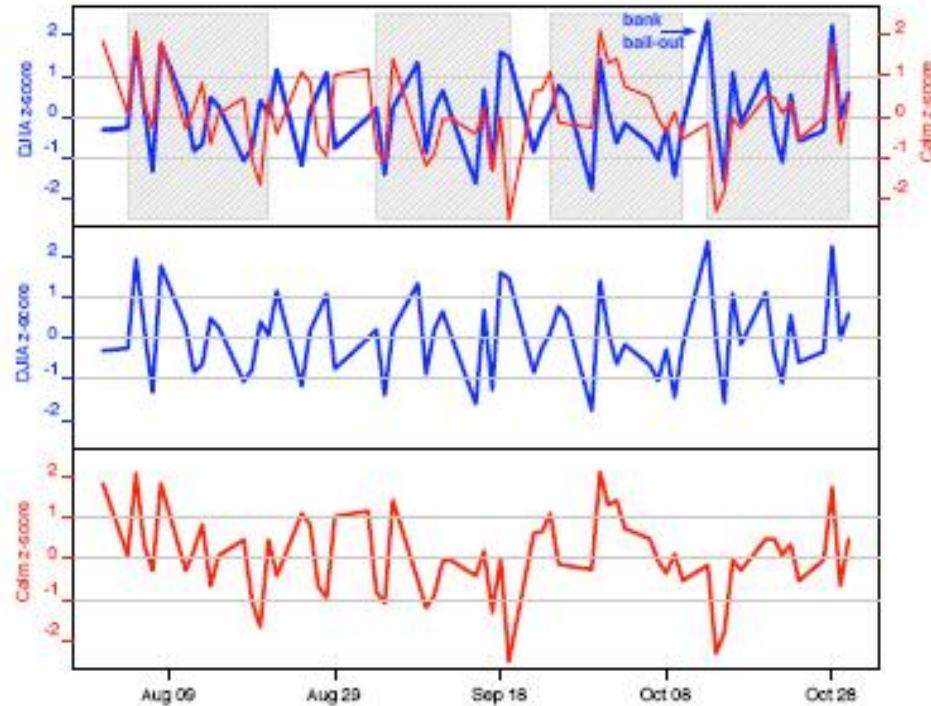
Data sets and timeline



Predicción de mercados de valores



Predicción de mercados de valores



Clasificación de Opinión en Twitter

- Primer trabajo: Go, A., Bhayani, R and Huang, Lei (2009). Twitter Sentiment Classification using Distant Supervision. CS224N Project Report, Stanford, 1, 12
- Corpus de tuits escritos en inglés:
 - Entrenamiento: 1.600.000 (800.000 positivos y 800.000 negativos).
 - Supervisión imprecisa: :) (positivo) :((negativo) de acuerdo a (Read, 2005).
 - Test: 182 positivos y 177 negativos. Etiquetado manual.
 - Algoritmo: SVM, Naïve Bayes, Máxima entropía.
 - Características: Unigramas, bigramas, unigramas+bigramas, pos-tags (categoría morfológica)
 - Mejor configuración:
 - Algoritmo: Máxima entropía.
 - Características: unigrams-bigrams.

Cocacola

Spanish

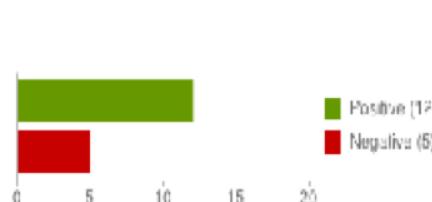
Search

Sentiment analysis for Cocacola

Sentiment by Percent



Sentiment by Count



Tweets about: Cocacola

[MOMtumblr](#): Photoset: Cuando tu madre pide una cocacola. <http://t.co/5IXQgHO8r2>

Posted: 1 minute ago

[MasterSMMAST](#): RT @ChusoSoria: @CocaCola_es crea nuevas formas de atención al cliente, Consumer Care. conversar con cocacola 2.0 sea una sensación social ?

Posted: 3 minutes ago

[RosFernandez7](#): Mi papí es malo. Siempre me deja con intriga, y aún por encima hoy, que le compré una lata de CocaCola con su nombre

Posted: 5 minutes ago

[Beerto10](#): RT @raullja: Comida favorita y bebida favorita ? Patatas fritas con filete y Cocacola <http://t.co/5iu3Df15PW>

Posted: 5 minutes ago

[ijuliansuaza](#): Le metian perico hasta a la cocacola (8)

Posted: 5 minutes ago

The results for this query are: [Accurate](#) [Inaccurate](#)

Negación en AO

- ¿Afecta la negación a la clasificación de la opinión?
- Jiménez-Zafra, S., Martín-Valdivia, M. T., Martínez-Cámara, E., Ureña-López, L. A. (2018). Studying the Scope of Negation for Spanish Sentiment Analysis on Twitter. *IEEE Transactions on Affective Computing*.
- Objetivo: Determinar si la identificación del ámbito de la negación es beneficioso para la clasificación de la opinión en español.

Negación en AO

Metodología:

- Corpus: General Corpus de TASS
 - Entrenamiento: 7.219.
 - Test: 60.798.
 - Clases: P+, P, NEU, N, N+, NONE.
 - Uso:
 - Sólo se usa el conjunto de test; se descarta la clase NONE; se agrupan las clases P+ y P, y N y N+.

P	22.233
NEU	1.305
N	15.844

Negación en AO

Metodología:

- Sistema: Clasificación no supervisada basada en el uso de:
 - La lista de palabras de opinión iSOL.
 - Lista de *hashtags* anotados a nivel de opinión.
 - Anotación de la opinión/emoción de emoticonos.
 - Identificación del ámbito de partículas negativas.
 - Clasificación basada en reglas

$$\text{opinión}(tuit) = \begin{cases} P & \text{si } pv > nv \\ NEU & \text{si } pv = nv \\ N & \text{si } pv < nv \end{cases}$$

Negación en AO

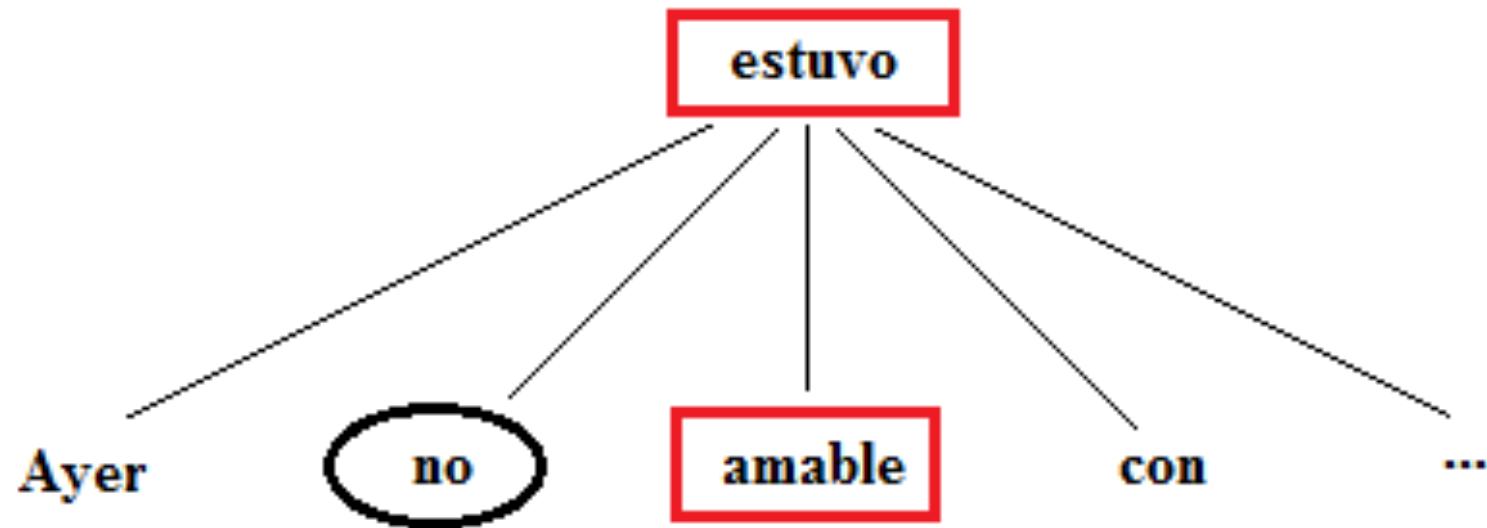
Identificación de la opinión:

- Partículas negativas usadas en español: no, tampoco, nadie, jamás, ni, sin, nada, nunca y ninguno.
- Tres grupos de partículas negativas según su construcción sintáctica.
- Cálculo árbol de dependencias sintácticas. Uso de Freeling.

Partículas Negativas	Regla
No, tampoco, nadie, jamás, ninguno	Nodo padre y todo el árbol formado por el hermano derecho
Ni, sin	Nodos hijos y todos sus descendientes hasta el nodo hoja
Nada, nunca	Nodo padre

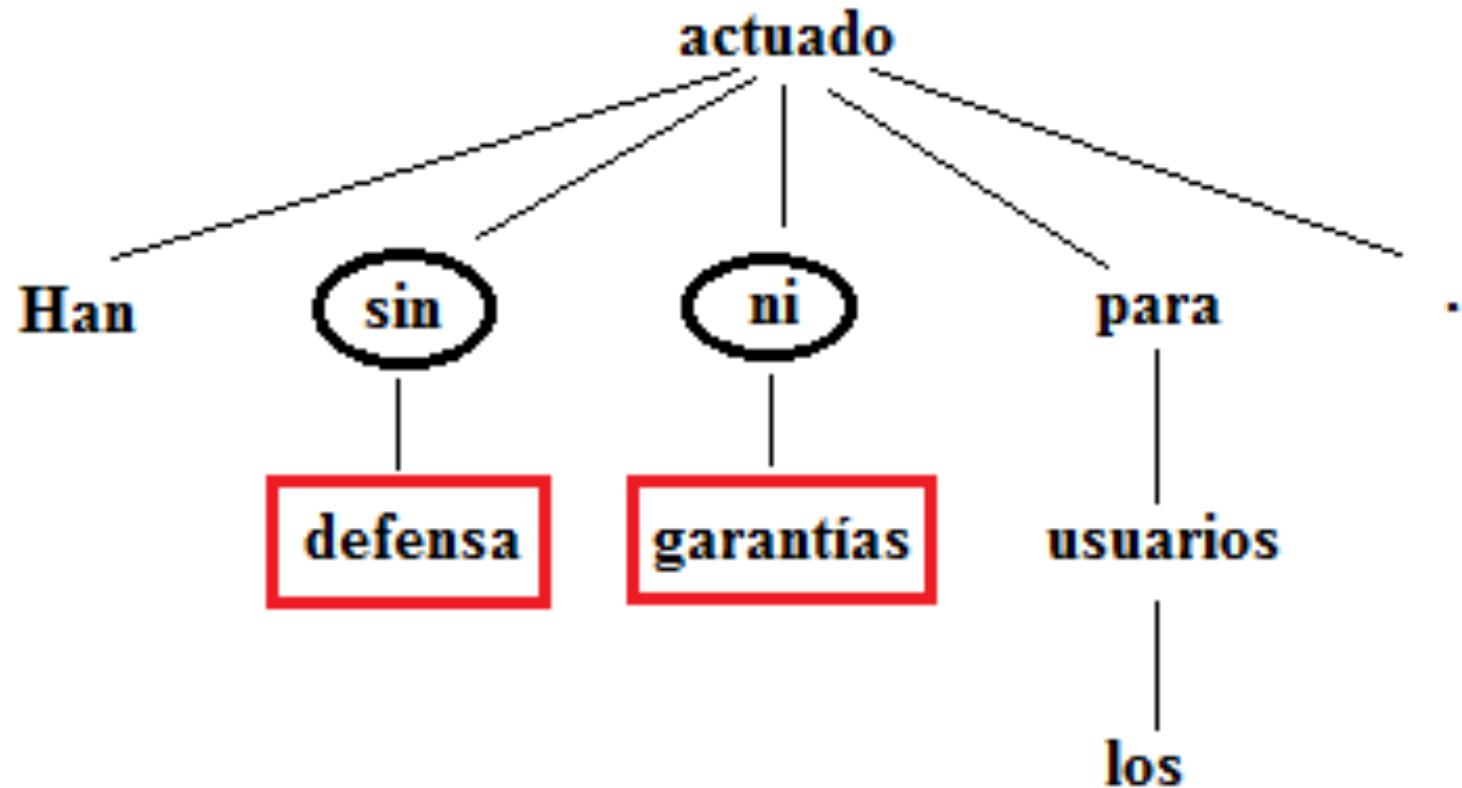
Negación en AO

Grupo partícula NO.



Negación en AO

Grupo partícula NI.



Negación en AO

Grupo partícula NADA.



Negación en AO

Todo arranca de un tweet nada amaaable. #maldad =(

Tokenización

[Todo] [arranca] [de] [un] [tweet] [nada] [amaaable][.] [#maldad]
[=()]

Normalización

[Todo] [arranca] [de] [un] [tweet] [nada] [amable][.] [#maldad]
[=()]

Part-of-Speech Tagging y Lematización

[Todo] [arrancar] [de] [un] [tweet] [nada] [amable][.] [#maldad]
[=()]

Negación en AO

Todo arranca de un tweet nada amaaable. #maldad =(
Negación

[Todo] [arrancar] [de] [un] [tweet] **[nada]** **[amable]**[.] [#maldad]
[=()]

Clasificación de la opinión

[Todo] [arrancar](-1) [de] [un] [tweet] {[**nada**] **[amable]** (+2)}(
-2)[.] [#maldad] (-2)[=() (-2)

Nv = 7

Pv = 0

Opinión = N

Negación en AO

Casos base:

- No considerar la negación (BS)
- Considerar como ámbito de la negación todas las palabras desde la partícula hasta un signo de puntuación (BSN)

Evaluaciones:

- Conjunto completo de datos (Total)
- Tuits con partículas negativas (NegCue)
- Tuits con partículas negativas y términos de opinión (RuleAffect)

Negación en AO

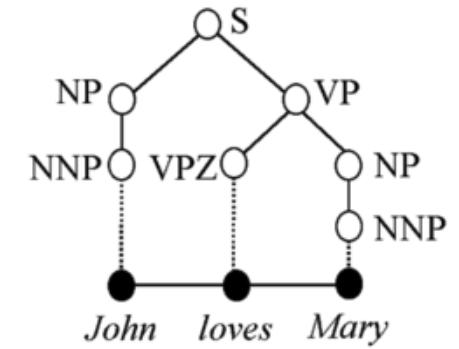
Corpus	Sistema	Macro-P	Macro-R	Macro-F1	Accuracy
Total	BS	0,5764	0,5235	0,5486	0,6258
	BSN	0,5705	0,5190	0,5435	0,6205
	Propuesta	0,5810	0,5296	0,5541	0,6308

Corpus	Sistema	Macro-P	Macro-R	Macro-F1	Accuracy
NegCue	BS	0,4861	0,4702	0,4780	0,4866
	BSN	0,4621	0,4514	0,4567	0,4622
	Propuesta	0,5060	0,4936	0,4997	0,5092

Corpus	Sistema	Macro-P	Macro-R	Macro-F1	Accuracy
RuleAffect	BS	0,3971	0,3949	0,3960	0,4463
	BSN	0,4431	0,4545	0,4487	0,5026
	Propuesta	0,4660	0,4792	0,4725	0,5292

Clasificación Opinión

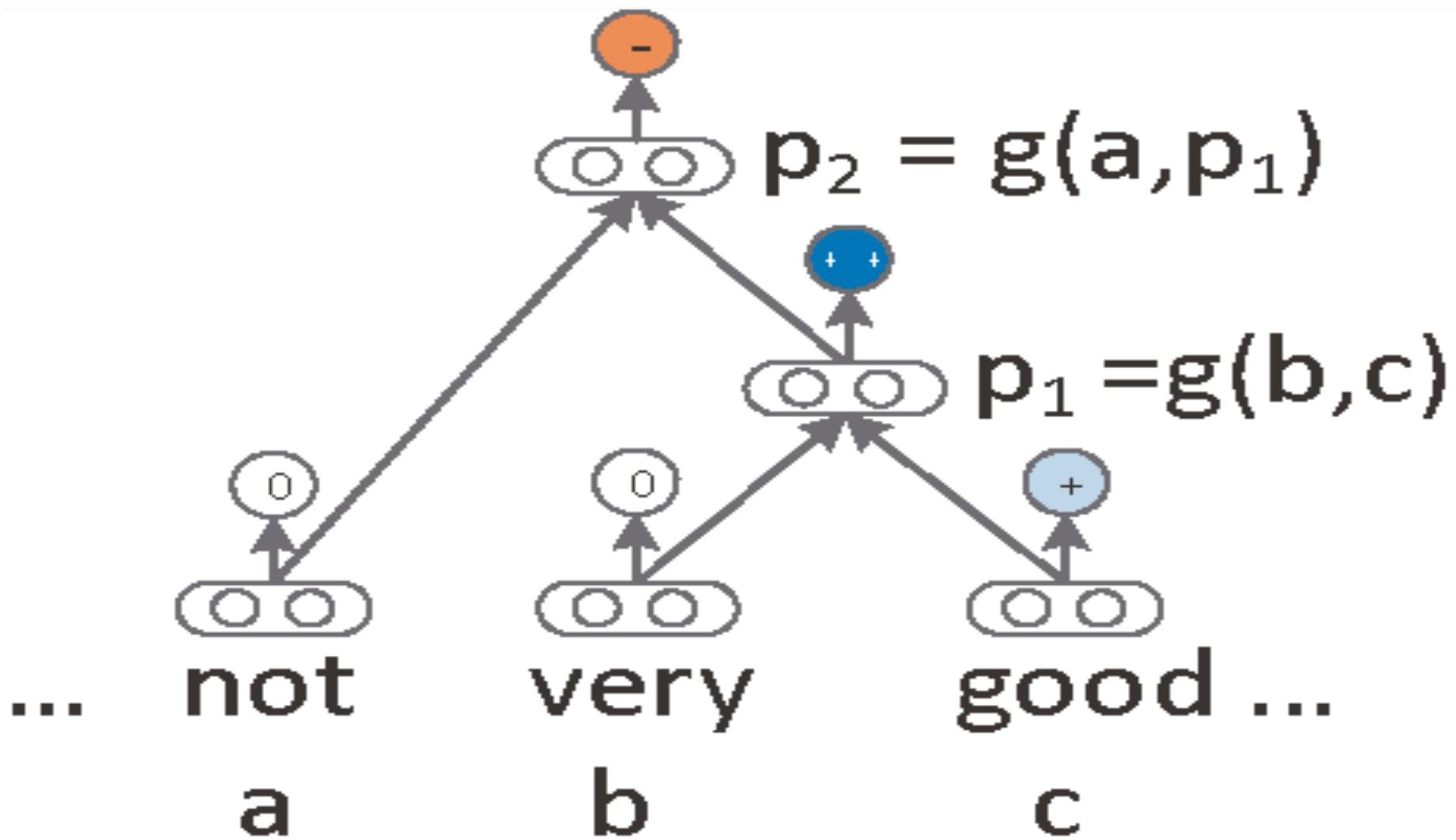
- Las redes neuronales (RN) representan ya el estado del arte en AO.
- El artículo de Socher et al. 2013 presenta una de las primeras aplicaciones con éxito de las RN en AO.
- Los autores argumentan que la opinión de un documento no sólo depende de la polaridad individual de las palabras, sino del significado global de todas ellas → Carácter composicional del lenguaje.
- Generación del primer Treebank de opiniones: Sentiment Treebank.
- Un Treebank es un corpus en el que se encuentra anotada la estructura sintáctica de las oraciones y la categoría morfológica de las palabras.
- Un Treebank de opiniones es un Treebank donde además están anotadas la polaridad de los sintagmas.



Clasificación Opinión

- Se propone una RN que clasifica la opinión de cada palabra.
- La RN es recursiva y recorre el árbol sintáctico de una oración.
- En cada nodo hoja, la RN calcula la polaridad del nodo del árbol teniendo en cuenta las palabras que dependen del nodo y del resultado de la anterior iteración.

Clasificación Opinión

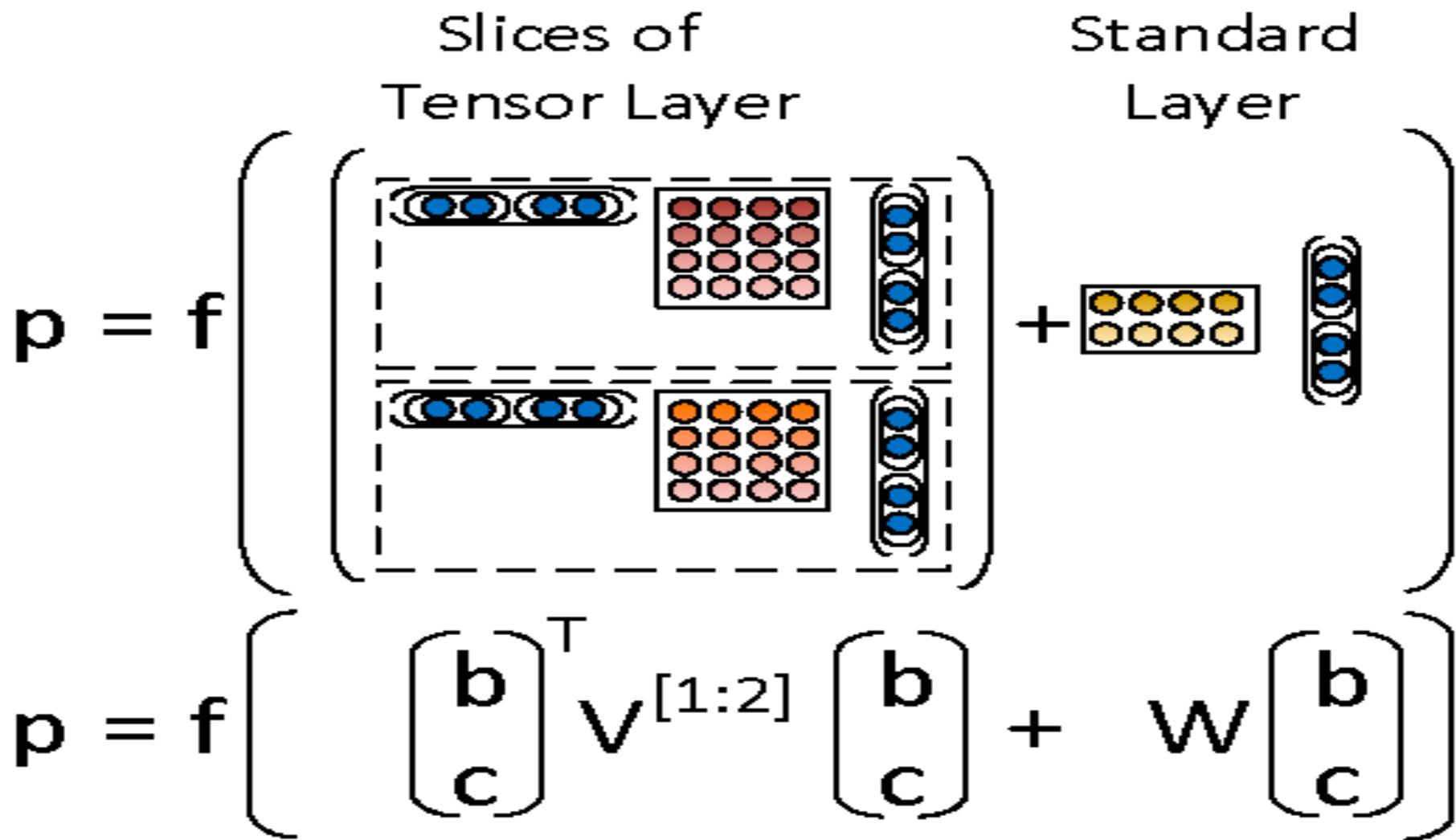


Clasificación Opinión

- Los autores proponen el modelo Tensor Neuronal Recurrente (*Recurrent Neural Tensor*).
- La idea principal es el aprendizaje automático de la función de composición del significado de varias palabras.

$$h = \begin{bmatrix} b \\ c \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} b \\ c \end{bmatrix}$$
$$h_i = \begin{bmatrix} b \\ c \end{bmatrix}^T V^{[i]} \begin{bmatrix} b \\ c \end{bmatrix}$$
$$V^{[1:d]} \in \mathbb{R}^{2d \times 2d \times d}$$

Clasificación Opinión



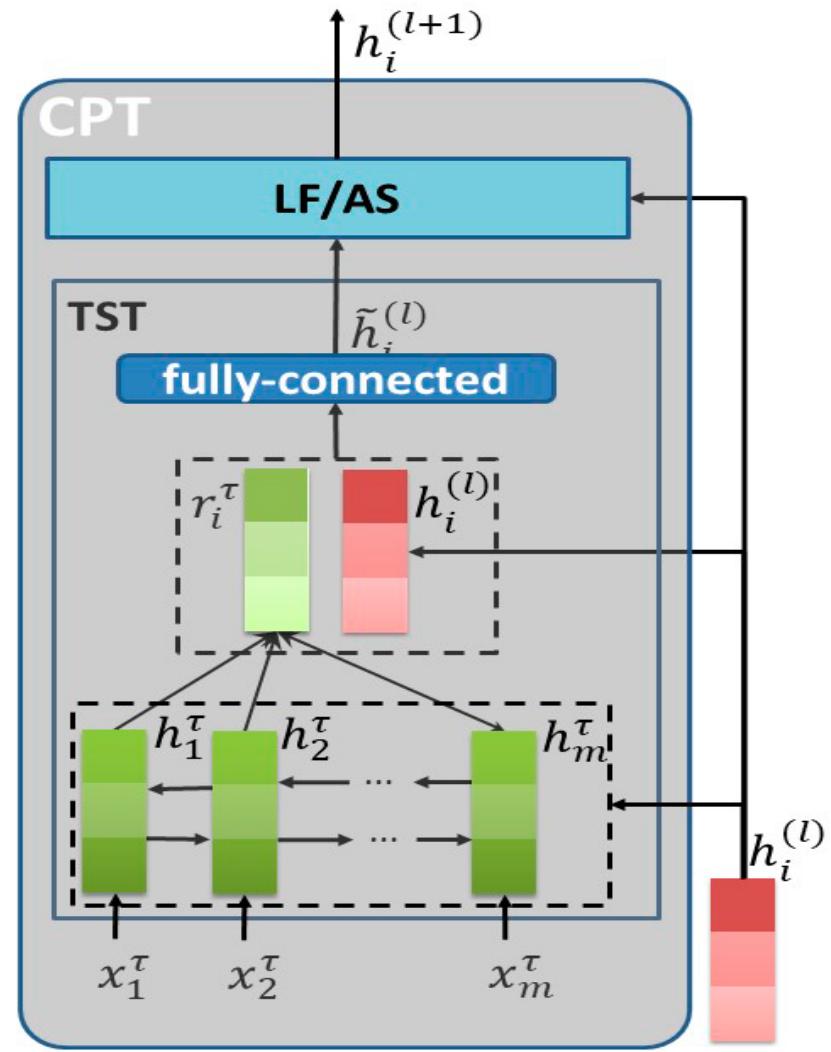
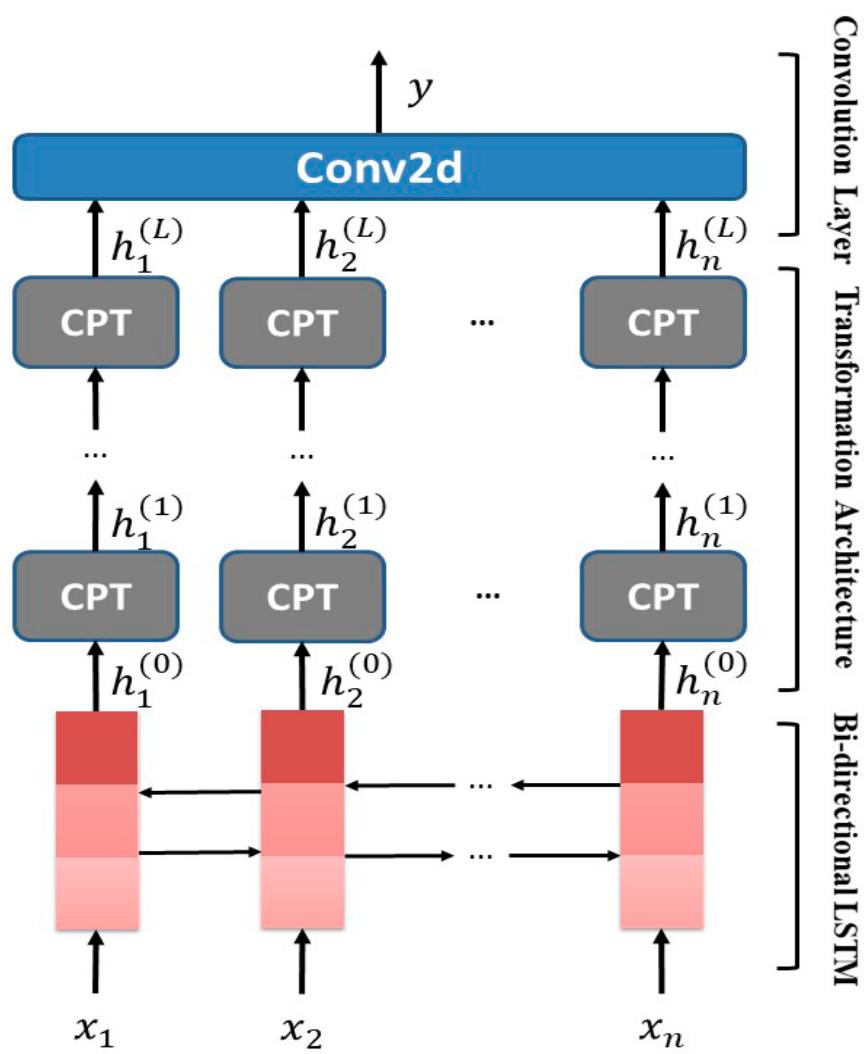
Clasificación Opinión

- En el corpus Movie Review (Pang & Lee 2005) realizan varios experimentos:
 - Clasificación 5 clases: 80,7% Accuracy (No exp. Anterior).
 - Clasificación binaria: 85,4% Accuary. Mejoran Est. Arte \approx 6,75%.
 - Opinión en oraciones positivas negadas: 71,4% de Accuary.
 - Opinión en oraciones negativas negadas: 81,8% de Accuary.
- Ref: Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 conference on empirical methods in natural language processing (pp. 1631-1642).
- Más información en:
<https://nlp.stanford.edu/sentiment/>

Clasificación Opinión a nivel de Aspecto

- La tendencia actual es la de aumentar la codificación de la información del contexto.
- Muy aconsejable para el AO a nivel de aspecto.
- Ej: Li, X., Bing, L., Lam, W., & Shi, B. (2018). Transformation Networks for Target-Oriented Sentiment Classification. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 946-956)
- Corpus: SemEval 2014
- Dominios:
 - Restaurantes: *All the **appetizers** and **salads** were **fabulous**, the **steak** was mouth watering and the **pasta** was delicious!!!*
 - Portátiles: *From the **build quality** to the **performance**, everything about it has been sub-par from what I would have expected from Apple.*

Clasificación Opinión a nivel de Aspecto



Clasificación Opinión a nivel de Aspecto

Sentence	BILSTM-ATT-G	RAM	TNet-LF	TNet-AS
1. Air has higher [resolution] _P but the [fonts] _N are small .	(N ^X , N)	(N ^X , N)	(P, N)	(P, N)
2. Great [food] _P but the [service] _N is dreadful .	(P, N)	(P, N)	(P, N)	(P, N)
3. Sure it ' s not light and slim but the [features] _P make up for it 100% .	N ^X	N ^X	P	P
4. Not only did they have amazing , [sandwiches] _P , [soup] _P , [pizza] _P etc , but their [homemade sorbets] _P are out of this world !	(P, O ^X , O ^X , P)	(P, P, O ^X , P)	(P, P, P, P)	(P, P, P, P)
5. [startup times] _N are incredibly long : over two minutes .	P ^X	P ^X	N	N
6. I am pleased with the fast [log on] _P , speedy [wifi connection] _P and the long [battery life] _P (> 6 hrs) .	(P, P, P)	(P, P, P)	(P, P, P)	(P, P, P)
7. The [staff] _N should be a bit more friendly .	P ^X	P ^X	P ^X	P ^X



Ejemplo

Ejemplo: TASS

- TASS: Taller de Análisis Semántico en la SEPLN.
- Hasta la fecha, principal taller y evaluación competitiva de AO en español.
- Desde 2012 se celebra como actividad del congreso de la SEPLN.



Ejemplo: TASS

- Principales contribuciones:

- Corpus General de TASS.
 - Clases: P+, P, NEU, N, N+, NONE.
 - Entrenamiento/Test: 60.798 tuits y 7.219 tuits.
- InterTASS: Corpus representativo del español de España y América (Perú y Costa Rica)
- Dos corpus en el dominio de la política:
 - Elecciones Generales de 2011
 - STOMPOL: Elecciones regionales y municipales de España de 2015
- Dominio Televisión Social: Corpus Social TV, Final Copa del Rey entre R. Madrid vs F.C. Barcelona de 2014



Ejemplo: Clasificadores

- Corpus: Conjunto de entrenamiento de Corpus General TASS
- Lenguaje: Python3
- Librerías: Scikit-learn, Numpy, Keras, Tensorflow
- Clasificadores:
 - Lineal: SVM. Características: Unigramas ponderados con TF-IDF.
 - No lineal: RNN LSTM. Características: Unigramas ponderados con TF-IDF.
 - No lineal: RNN LSTM. Características: Unigramas con embeddings aleatorios
 - No lineal: RNN LSTM. Características: Unigramas ponderados con embeddings pre-entrenados (FastText).

Ejemplo: Clasificadores

- Ahora nos vamos a Jupyter.

Muchas gracias por su atención

¿PREGUNTAS?