

---

# Introducción a la programación y análisis de texto con R

---

Clase 9 – PLN, Análisis de Sentimiento,  
Modelos de Lenguaje

---

# Introducción

---

## ¿Qué es el PLN?

*El **Procesamiento de Lenguaje Natural (PLN)** es una subdisciplina de la Inteligencia Artificial que intenta resolver con **computadoras** tareas vinculadas al **lenguaje humano**, permitiendo la comunicación entre el humano y la computadora a través del lenguaje natural o resolviendo diferentes tareas que implican algún tipo de **procesamiento de texto o habla**. (Jurafsky & Martin, 2008)*

---

# Introducción

---

– Recuperación de información



Busca en Google o escribe una URL 

– Traducción automática

意大利蛋白  
酥皮食譜 → receta de  
merengue  
italiano

– Chatbots



– Análisis de sentimiento



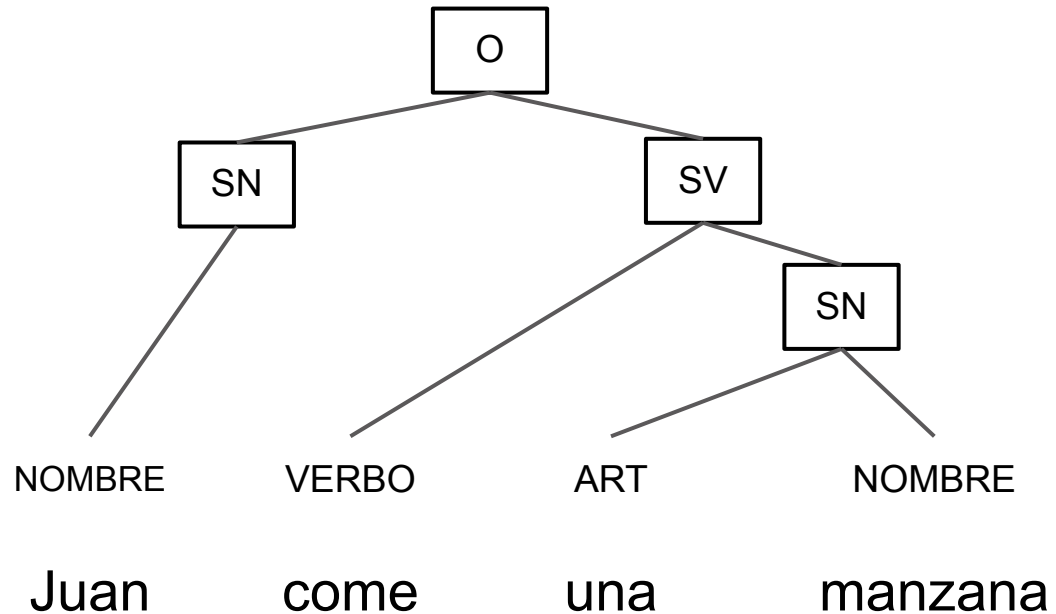
# Introducción

---

– Parsing

– POS Tagging

– mucho más ...



# Análisis de Subjetividad

---

- Análisis de Sentimiento: clasificación de textos según su polaridad, típicamente, positivo/negativo/neutro.
  - Análisis de emociones: se refina el análisis de sentimiento para clasificar según emociones: alegría, tristeza, enojo, ...
  - Análisis de humor: clasificación entre humor/no humor, grado de gracia de un chiste, mecanismo humorístico, tema del chiste, ...
    - Identificación de ironía y sarcasmo.
  - Detección de discurso de odio (racismo, misoginia, ...)
-

# Análisis de Subjetividad: emociones

---

Definición del conjunto de emociones:

- esquema que surge de la combinación de tres dimensiones: valence (polaridad), arousal (grado de activación), dominance (grado de control) (Russell, 1980).
  - esquema de 6 emociones: surprise, happiness, anger, fear, disgust, sadness (Ekman, 1999).
  - esquema de 8 emociones: 4 pares de opuestos: joy–sadness, anger–fear, trust–disgust, and anticipation–surprise (Plutchik, 1980).
    - de este esquema de 8 emociones se derivan otras
-

# Análisis de Subjetividad: emociones

---

Definición del conjunto de emociones:

- esquema basado en tres dimensiones, esquema de 6 emociones, esquema de 8 emociones.

Recursos:

- Léxicos anotados con emociones.
- Conjuntos de textos clasificados según diferentes esquemas de emociones.

Clasificación de tweets en español:

- EmoEvalEs (IberLEF 2021, SEPLN): corpus desbalanceado (esquema de 6 valores), clases difíciles de distinguir (fear/disgust), clase “other” se confunde con las demás.

Trabajo de nuestro grupo: Rosá et al., 2021

---

# Análisis de Sentimiento

---

## Análisis de Sentimiento o Minería de Opiniones

(Turney, 2002; Wiebe et al., 2005; Pang & Lee, 2008; Liu, 2010/ 2020; competencias TASS – IberLEF, SemEval)

### Diferentes tareas:

- Clasificación de textos (tweets, opiniones sobre películas, hoteles, noticias, etc.) según su polaridad (positivos o negativos)
  - Extraer opiniones a partir de textos (por ejemplo, prensa).
  - Para cada opinión:
    - Autor
    - Polaridad (P, N, P+, N+, Neu, etc.)
    - Aspectos: celular -> batería P, pantalla N
    - ...
-



# Análisis de Sentimiento

---

## Usos del análisis de sentimiento:

- Relevancia y clasificación de textos según su polaridad
  - Proceso de toma de decisiones
  - Análisis de la tendencia de opinión
  - Campañas políticas y de mercado
-

# Análisis de Sentimiento

---

Cada vez se habla más del análisis de sentimiento:

- Mayor disponibilidad de datasets para entrenar algoritmos de aprendizaje automático.
- Aumento de la eficiencia de los métodos de aprendizaje automático, procesamiento del lenguaje natural y recuperación información.
- Incremento masivo del uso de las redes sociales por el público en general para expresar opiniones.



# Análisis de Redes Sociales

---

## Desafíos:

### ➤ Lenguaje informal

El lenguaje usado tiene errores ortográficos y no siempre ayudan los correctores

*“kereees decirM ques esssssooo #Tasahi”*



# Análisis de Redes Sociales

---

## Desafíos:

- Formas sutiles de expresar sentimientos:
    - “Mario Benedetti es un excelente escritor” **Palabras con polaridad**
    - “Se trata de la fragancia que quería, por favor, úsela exclusivamente en casa y cierre las ventanas” **Sin palabras negativas**
    - “Si, seguro, cómo no!!” **Ironía**
    - “Vamos a ponerse las pilas” **Expresiones coloquiales**
    - “Si creíste que iba a ser una buena película, no fue una buena elección” **Negación**
-

# Clasificación de tweets

---

Qué bien que  
marcaron a Messi



Gracias a la gente de  
Brookfield por su excelente  
atención. <https://t.co/r38i4q7cIV,P>



pues a mi me ha  
encantado #sherlock



@user top secret: el #psoe de  
andalucía despilfarra otros 90 millones de  
euros en subvenciones. <http://t.co/aiwhnxrj>




# Clasificación de tweets

---

Qué *bien* que marcaron a Messi 

*Gracias* a la gente de Brookfield por su *excelente* atención. <https://t.co/r38i4q7clV,P> 

pues a mi me ha *encantado* #sherlock 

@user top secret: el #psoe de andalucía *despilfarra* otros 90 millones de euros en subvenciones. <http://t.co/aiwhnxrj> 

---

# Clasificación de tweets



---


*Entró la magia al camp nou vamo #Messi* 

*Uruguay sin mucho fútbol pero con abundante marca y actitud* 

*volvió la #bestia... volvió Lionel Andrés #Messi!!! #GIGANTE #FCBDEP* 

*Qué fantasma este pelado! Roja inexistente.* 

*Después de esto como para que no lo echaran a Sanmartino*  

*ajaja muy bueno <https://t.co/i4mfeb1qdY>* 

---

# Clasificación de tweets

---

Tareas previas al análisis del contenido de un texto:

- Obtener los textos de interés: tweets, comentarios en foros o blogs, mails, etc.
  - Limpiarlos: eliminar o sustituir símbolos, enlaces, etiquetas.
  - Guardarlos: definir el modelo de datos y los campos relevantes.
-



# Clasificación de tweets

---







## Primer enfoque: Reglas Manuales

- Escritura de reglas en base al análisis de ejemplos
  - Las reglas incluyen conocimiento lingüístico:
    - Lematización
    - POS tagging (categorías gramaticales de palabras)
    - Parsing (estructura sintáctica de oraciones)
  - También conocimiento específico del dominio o del problema:
    - Léxicos afectivos para análisis de sentimiento
    - Negadores e intensificadores
-

# Clasificación de tweets

---

➤ Léxicos afectivos

<i>excelente</i>		<i>horrible</i>	
<i>útil</i>		<i>despilfarrar</i>	
<i>felicitar</i>		<i>guerra</i>	
...		...	

# Clasificación de tweets

---

- Léxicos afectivos
- Lematización

*sugerencia/sugerencias → sugerencia*  
*despilfarra/despilfarraron/despilfarrando → despilfarrar*  
*bueno/buena/buenos/buenas → bueno*

# Clasificación de tweets

---

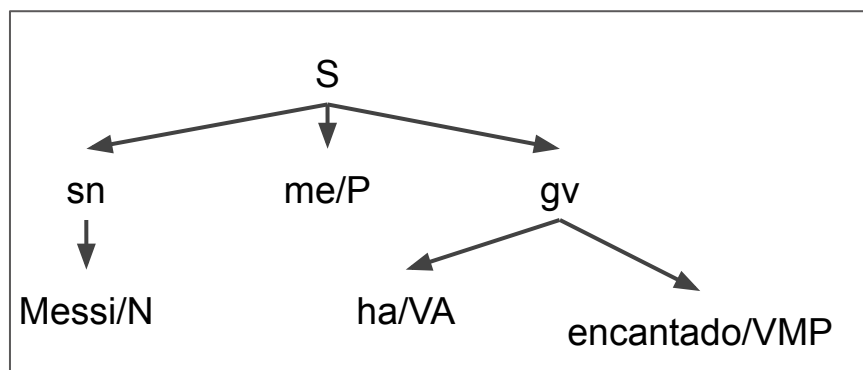
- Léxicos afectivos
- Lematización
- POS-tagging

Messi	→	NP
me		PP
ha		VAIP
encantado		VMP

# Clasificación de tweets

---

- Léxicos afectivos
- Lematización
- POS-tagging
- Parsing



# Clasificación de tweets

---

Ejemplo (1):

*Gracias a la gente de Brookfield por su excelente atención.*

Posible regla:

busco palabras en léxicos afectivos

si `cantidad_palabras_positivas > cantidad_palabras_negativas`  
    el tweet es positivo

si no

    el tweet es negativo

---

# Clasificación de tweets

---

Ejemplo (2):

*Gracias a la gente de Brookfield por sus excelentes sugerencias.*

Posible regla:

```
busco lemas en léxicos afectivos
si cantidad_lemas_positivos > cantidad_lemas_negativos
    el tweet es positivo
si no
    el tweet es negativo
```

# Clasificación de tweets

---

## Ejemplo (3):

*La atención de Brookfield es **excelente**, pero los productos son de **mala** calidad.*

## Posible regla:

busco lemas en léxicos afectivos

si  $\text{cantidad\_lemas\_positivos} > \text{cantidad\_lemas\_negativos}$

el tweet es positivo

si no, si  $\text{cantidad\_lemas\_positivos} < \text{cantidad\_lemas\_negativos}$

el tweet es negativo

si no

el tweet es neutro

---



# Clasificación de tweets

---

## Ejemplo (3):

*La atención de Brookfield es **excelente**, pero los productos son de **mala** calidad.*

busco lemas en léxicos afectivos

```
si cantidad_lemas_positivos = 0 y cantidad_lemas_negativos = 0
  el tweet es neutro (o no_sent)
```

```
si no
```

```
  si cantidad_lemas_positivos > cantidad_lemas_negativos
    el tweet es positivo
```

```
  si no, si cantidad_lemas_positivos < cantidad_lemas_negativos
    el tweet es negativo
```

```
  si no
```

```
    el tweet es mixto (o neutro)
```

---

# Clasificación de tweets

---

## Ejemplo (4):

*La atención de Brookfield es excelente, pero los productos no son buenos.*

## Posible regla:

```
busco lemas en léxicos afectivos
invierto valor afectivo de palabras negadas
si cantidad_lemas_positivos > cantidad_lemas_negativos
    el tweet es positivo
si no, si cantidad_lemas_positivos < cantidad_lemas_negativos
    el tweet es negativo
si no
    el tweet es neutro
```

---

# Clasificación de tweets

---

## Primer enfoque: Reglas Manuales

- Es muy difícil abarcar todos los casos escribiendo reglas manuales.
  - Se amplía la cobertura con grandes cantidades de ejemplos y métodos de aprendizaje automático supervisado, también con métodos híbridos.
  - Esto es posible si se cuenta con conjuntos de datos anotados.
  - Actualmente, los grandes modelos de lenguaje logran buenos resultados sin corpus anotados específicamente para la tarea.
-

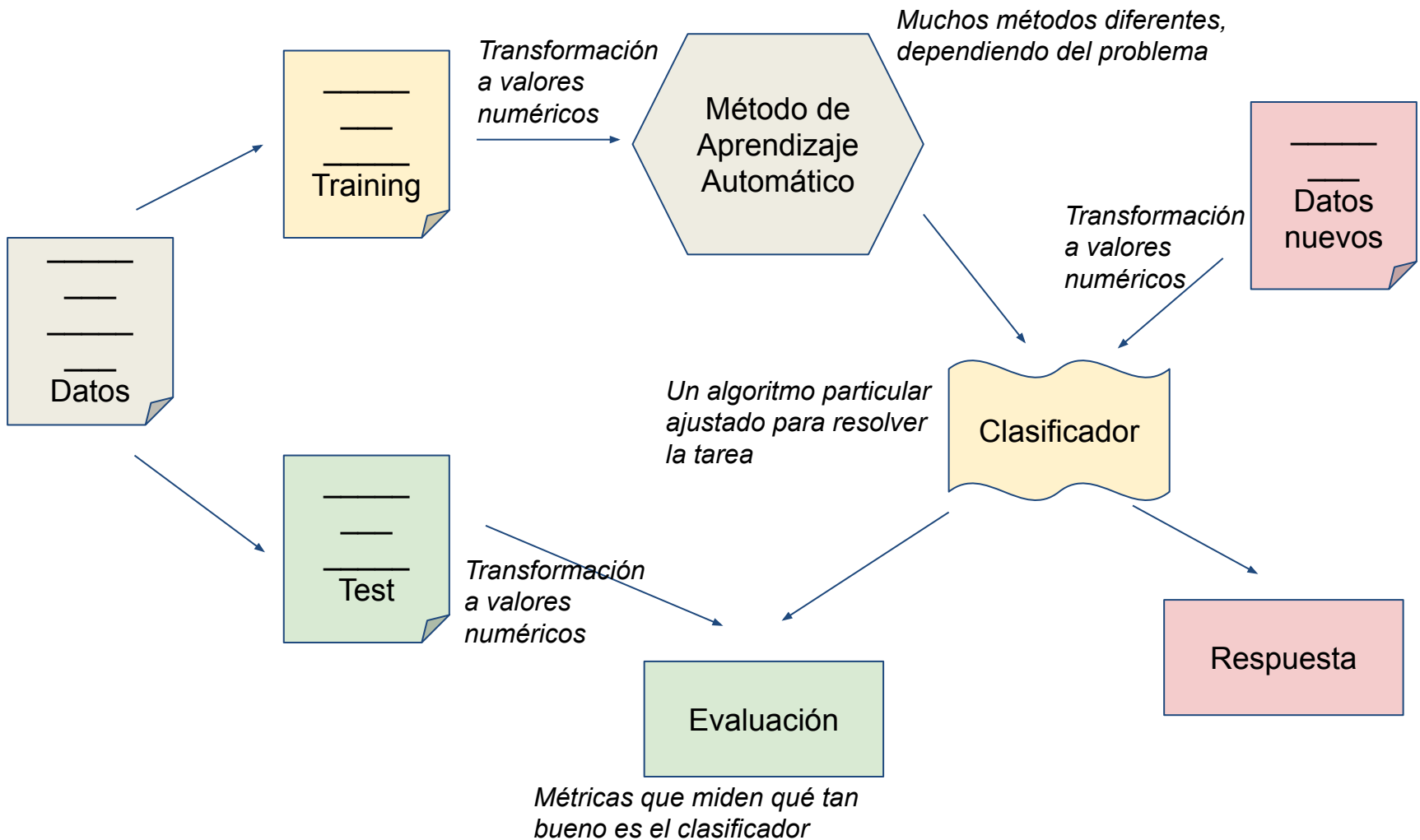
# Clasificación de tweets

---

## Segundo enfoque: Aprendizaje automático basado en *features* manuales

- Corpus de tweets
    - Cada uno con su clasificación (positivo, negativo, etc...)
    - Etiquetados a mano
  - Aprender una función que prediga la clase dado el tweet:
    - Entradas de la función: conjunto de atributos (*features*)
    - Salidas: la clase predicha
  - En estos casos hablamos de aprendizaje automático supervisado
-

# Aprendizaje Automático



# Clasificación de tweets

---

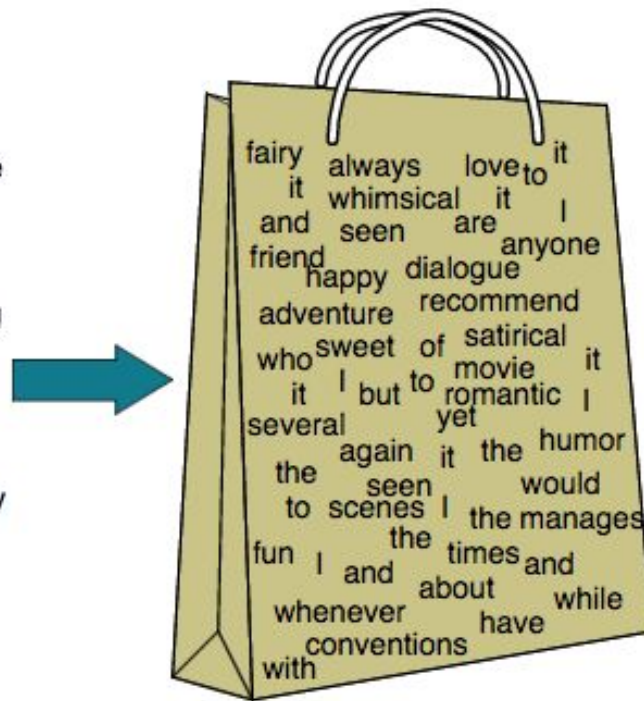
Segundo enfoque: Aprendizaje automático basado en *features* manuales

- Posibles métodos: Naïve Bayes, Árboles de Decisión, SVM...
  - Algunos atributos posibles para análisis de sentimiento:
    - palabras (*bag of words*)
    - lemas
    - categorías gramaticales
    - cantidad de palabras positivas/negativas
    - presencia de negación
    - información sintáctica
    - ...
-

# Clasificación de tweets

Ejemplo: Representación de atributos de un texto (*bag of words*)

I love this movie! It's sweet, but with satirical humor. The dialogue is great and the adventure scenes are fun... It manages to be whimsical and romantic while laughing at the conventions of the fairy tale genre. I would recommend it to just about anyone. I've seen it several times, and I'm always happy to see it again whenever I have a friend who hasn't seen it yet!



it	6
I	5
the	4
to	3
and	3
seen	2
yet	1
would	1
whimsical	1
times	1
sweet	1
satirical	1
adventure	1
genre	1
fairy	1
humor	1
have	1
great	1
...	...

¿Qué problemas puede tener?

# Clasificación de tweets

---

## Segundo enfoque: Aprendizaje automático basado en *features* manuales

- Posibles métodos: Naïve Bayes, Árboles de Decisión, SVM, MLP, ...
  - Algunos atributos posibles para análisis de sentimiento:
    - ~~palabras (*bag of words*)~~ vectores de palabras (*word embeddings*)
    - lemas
    - categorías gramaticales
    - cantidad de palabras positivas/negativas
    - presencia de negación
    - información sintáctica
    - ...
-



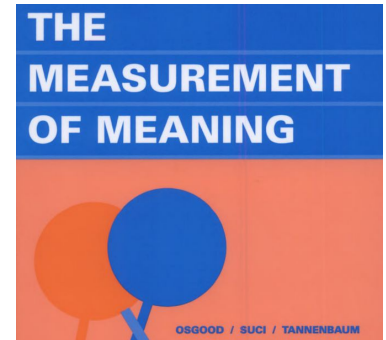
# Vectores de palabras

Antecedente histórico: Osgood et al. (1957) proponen que las palabras varían según tres dimensiones:

*valencia*

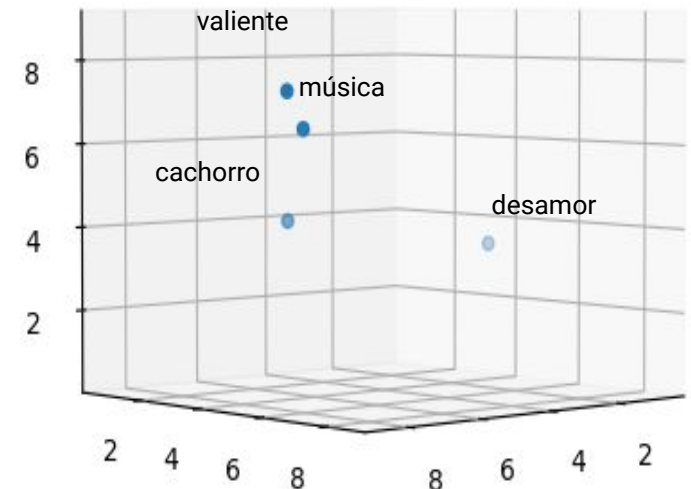
*entusiasmo*

*dominancia*



Con estos valores, cada palabra podría representarse como un vector de tres dimensiones

	<i>valencia</i>	<i>entusiasmo</i>	<i>dominancia</i>
valiente	8,05	5,50	7,38
música	7,67	5,57	6,50
desamor	2,45	5,65	3,58
cachorro	6.71	3.95	4.24



# Vectores de palabras

---

Procesos para construir vectores de palabras (p.e. Mikolov et al., 2013)

A cada palabra le corresponde un vector

Palabras similares tienen vectores cercanos, palabras diferentes tienen vectores lejanos

Grandes cantidades de datos gracias a internet (miles de millones de palabras!)

Mayor poder de cómputo

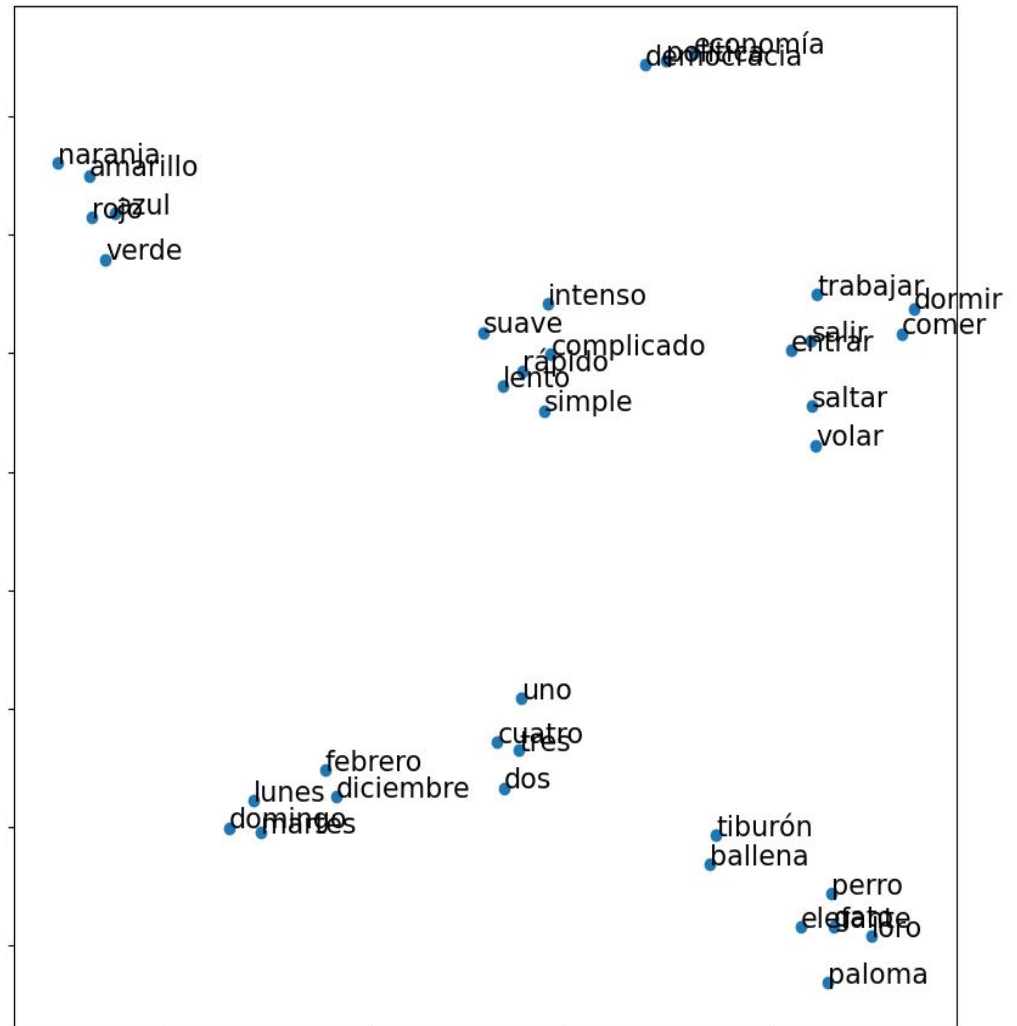
---

# Vectores de palabras

Vectores de 300 dimensiones,  
entrenados sobre 1500 millones  
de palabras en español  
(Cardellino et al., 2016)

Graficamos un conjunto de  
palabras:

animales, días, meses,  
adjetivos, verbos, colores, ...

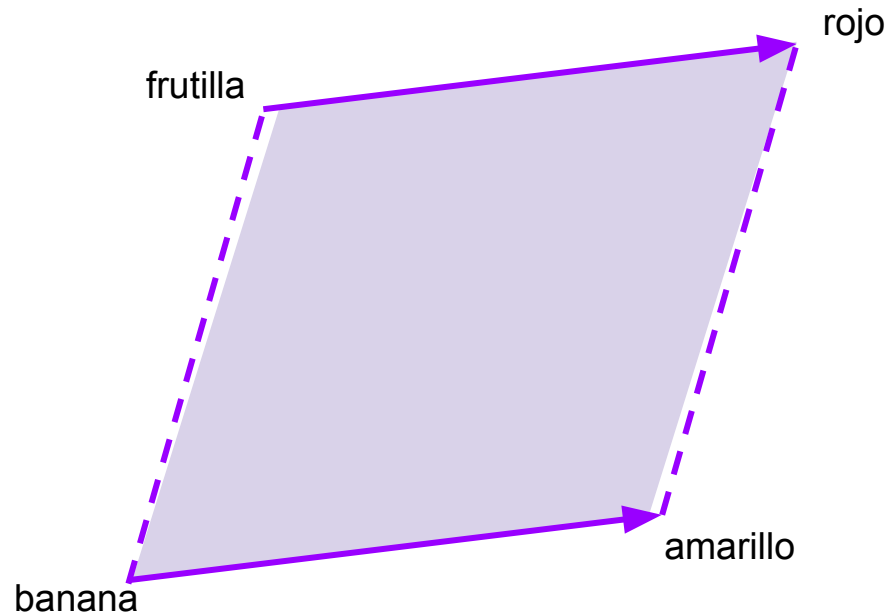


# Vectores de palabras

---

Modelo del paralelogramo para razonamientos sobre analogías (Rumelhart y Abrahamson, 1973)

Resolver: “frutilla es a rojo como banana es a \_\_\_\_\_”



¡Se trasladan muy bien a operaciones con vectores de palabras!

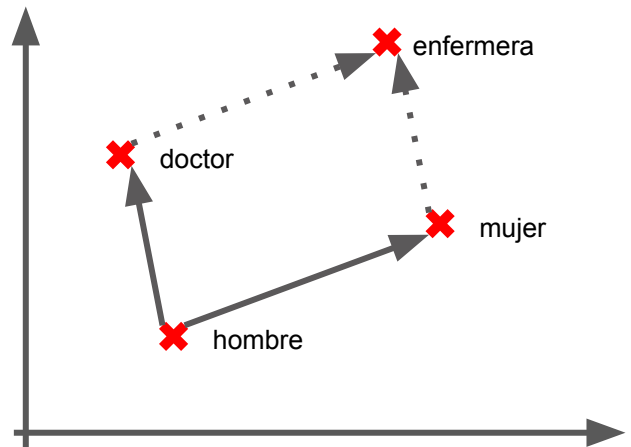
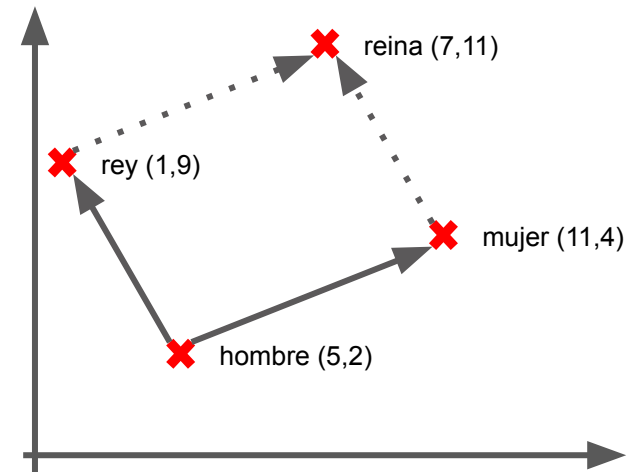
---

# Vectores de palabras

Las operaciones matemáticas sobre vectores pueden descubrir relaciones entre palabras:

rey - hombre + mujer  $\cong$  reina

uruguay - montevideo + francia  $\cong$  parís



¡Cuidado! También puede amplificar sesgos incorrectos encontrados en los datos

# Clasificación de tweets

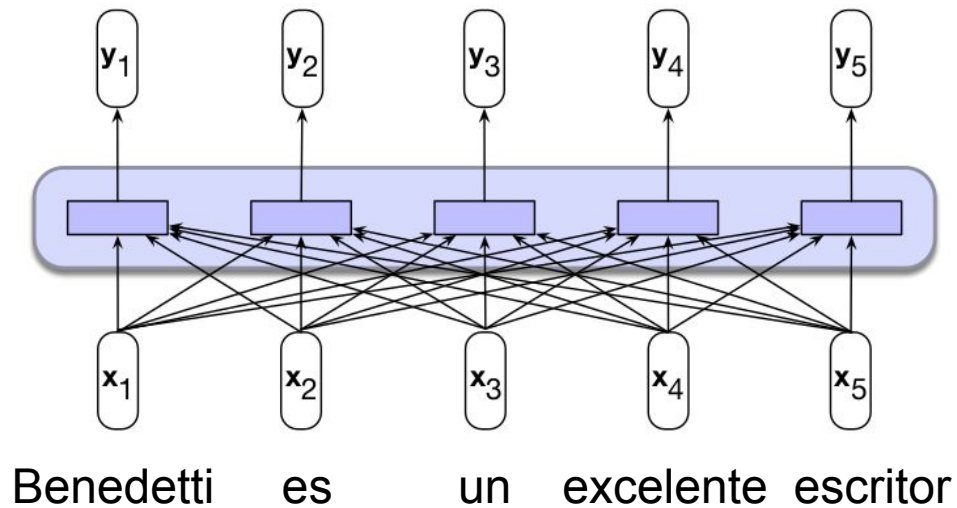
---

## Tercer enfoque: Aprendizaje Profundo (Deep learning)

- métodos de aprendizaje automático basados en redes neuronales de varias capas
  - vectores de palabras (*word embeddings*) como entrada de la red
  - diferentes arquitecturas:
    - convolucional
    - bi-LSTM
    - modelo de lenguaje neuronal tipo BERT (transformers) + fine tuning
-

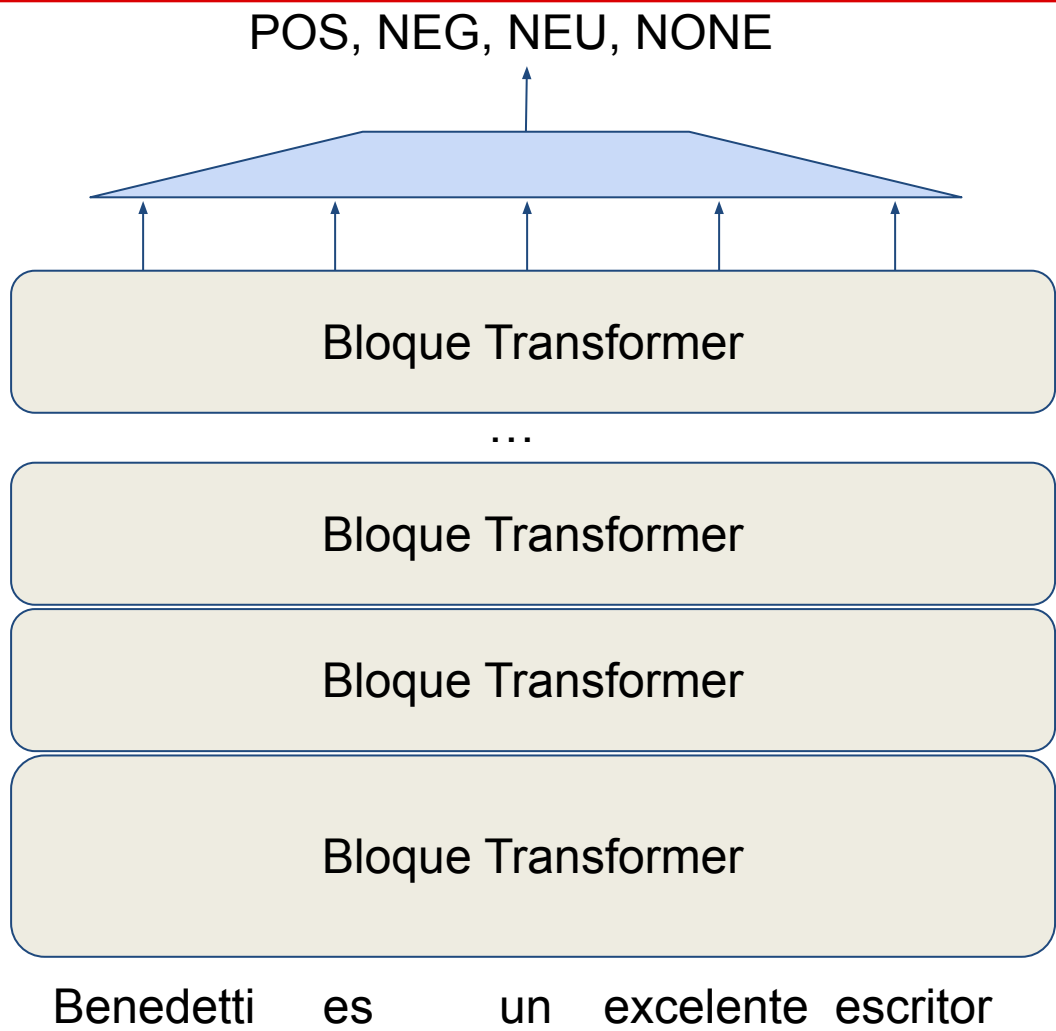
# BERT + fine tuning

---



# BERT + fine tuning

---





# Competencia TASS

---

## Análisis de sentimiento en tweets en español TASS (IberLEF, SEPLN)

- La organización distribuye datos para entrenamiento y validación a los participantes.
  - 1 mes y medio para desarrollo de sistemas.
  - Se publican datos de testeo (sin anotaciones).
  - Los participantes envían los resultados sobre testeo, que son evaluados contra el gold standard y rankeados.
  - Medidas de evaluación: Macro-F y Accuracy.
-

# Competencia TASS

---

Conjunto de etiquetas del TASS (hasta 2019): P, N, NEU, NONE

- *“@user top secret: el #psoe de andalucía despilfarra otros 90 millones de euros en subvenciones. <http://t.co/aiwhnxrj>” → **N***
  - *“pues a mi me ha encantado #sherlock” → **P***
  - *“el principio ha sido súper iconic pero con tanto movimiento iba muerta la pobre #vmas” → **NEU***
  - *“hoy conoceremos datos definitivos de 2011 del padrón municipal. datos ine: en españa hay casi un 1% más de mujeres q de hombres” → **NONE***
-

# Competencia TASS

---

Sistemas presentados por el equipo de PLN-InCo (RETUYT) en TASS 2017, 2018, 2019:

- SVM y MLP con varias features calculadas sobre un léxico subjetivo
- CNN con diferentes arquitecturas
- LSTM optimizado para neutros
- Combinación entre salidas de SVM y CNN
- Adaptación de BERT

En todos los enfoques, incluso SVM y MLP, representamos los tweets en base a los **embeddings** de sus palabras.

Para SVM y MLP calculamos el vector promedio de los embeddings del tweet.

---

# Competencia TASS

---

## Resultados:

- Enviamos resultados de tres sistemas para cada corpus de testeo provisto por la organización.
  - 2017: 3 corpus
  - 2018: 6 corpus - diferentes variantes de español: España, Perú, Costa Rica, combinaciones entre variantes
  - 2019: 10 corpus - España, Perú, Costa Rica, México, **Uruguay** (anotado por nosotros)
  - En varios corpus nuestros sistemas quedaron primeros.
  - Principal problema: tweets neutros. Son pocos y el criterio de anotación no parece muy homogéneo.
  - Resultados generales son bajos: ¡Macro-F aproximadamente 50%!
    - 2020 (tres clases): mejores resultados fueron 67% Macro-F en datos de una sola variante, y 50% en unión de variantes..
-

# Modelo de lenguaje para tweets en español

---

- Entrenamiento de un modelo de lenguaje basado en Roberta (basado a su vez en BERT) con un gran corpus de tweets en español: RoBERTuito (Pérez et al., 2022).
    - 500M tweets
  - Mejoras significativas en una herramienta para análisis de subjetividad, que incluye análisis de sentimiento: pysentimiento (Pérez et al, 2021).
  - La mejor configuración alcanza un 70.7% de Macro F para análisis de sentimiento, sobre el dataset de TASS 2020, usando la unión de todas las variantes.
  - <https://github.com/pysentimiento/pysentimiento>
  - [Notebook con ejemplos](#)
-

# Modelo de lenguaje para prensa uruguaya

---

- ROUBERTa
  - Entrenado con datos de prensa uruguaya
    - El Observador, El País, Montevideo Portal, La Diaria
    - 400M palabras (BERT usó más de 3300M palabras)
  - Evaluado con tareas que involucran texto de prensa uruguayo
    - Análisis de sentimiento
    - Sentimiento económico
    - Respuestas a preguntas sobre COVID-19
    - Predicción de texto
  - <https://huggingface.co/filevich/robertita-cased-finetuned-fact>
-

# Referencias

---

- Jurafsky, Daniel, and James H. Martin. "Speech and Language Processing", 3rd Edition, 2023 draft
  - Turney, P. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. Proceedings of the 40th Annual Meeting of the ACL, Philadelphia.
  - Pang, B., Lee, L. 2008. Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval Vol. 2, No 1–2, 1–135.
  - García Vega, M. et al.. 2020. Overview of TASS 2020: Introducing Emotion Detection . Proceedings of TASS 2020, co-located with 36th SEPLN Conference (SEPLN 2020), Malaga, Spain.
  - Rosá, A., Chiruzzo, L., Etcheverry, M., Castro, S. 2017. RETUYT en TASS 2017: Análisis de Sentimiento de tweets en español utilizando SVM y CNN. Proceedings of TASS 2017, co-located with 33rd SEPLN Conference (SEPLN 2017), Murcia, Spain.
  - Chiruzzo, L., Rosá, A. 2018. RETUYT-InCo at TASS 2018: Sentiment Analysis in Spanish Variants using Neural Networks and SVM. Proceedings of TASS 2018, co-located with 34nd SEPLN Conference (SEPLN 2018), Sevilla, Spain.
  - Pastorini, M, Pereira, M, Zeballos, N, Chiruzzo, L, Rosá, A, Etcheverry, M. RETUYT-InCo at TASS 2019: Sentiment Analysis in Spanish Tweets. Proceedings of TASS 2019, co-located with 35th SEPLN Conference (SEPLN 2019), Bilbao, Spain.
  - Pérez, JM., Furman, D., Alemany, L., Luque, F. RoBERTuito: a pre-trained language model for social media text in Spanish. Proceedings of the Thirteenth Language Resources and Evaluation Conference, 2022.
  - Pérez, JM., Giudici, JC., Luque, F. pysentimiento: A Python Toolkit for Sentiment Analysis and SocialNLP tasks, arXiv cs.CL, 2021. <https://arxiv.org/abs/2106.09462>
  - Rosá, A.; Chiruzzo, L. Emotion Classification in Spanish: Exploring the Hard Classes. Information 2021, 12, 438. <https://doi.org/10.3390/info12110438>.
  - Rosá, A., Chiruzzo, L., Bouza, L., Dragonetti, A., Castro, S., Etcheverry, M., Góngora, S., Goycochea, S., Machado, J., Moncecchi, G. and Prada, J.J., Overview of QuALES at IberLEF 2022: Question Answering Learning from Examples in Spanish, 2022.
-

---

# Gracias!

---

Grupo PLN – Instituto de Computación – Facultad de Ingeniería – Udelar

<https://www.fing.edu.uy/inco/grupos/pln/>

 [pln@fing.edu.uy](mailto:pln@fing.edu.uy)

 @PLN\_UdelaR

---