



Otras formas de clasificación

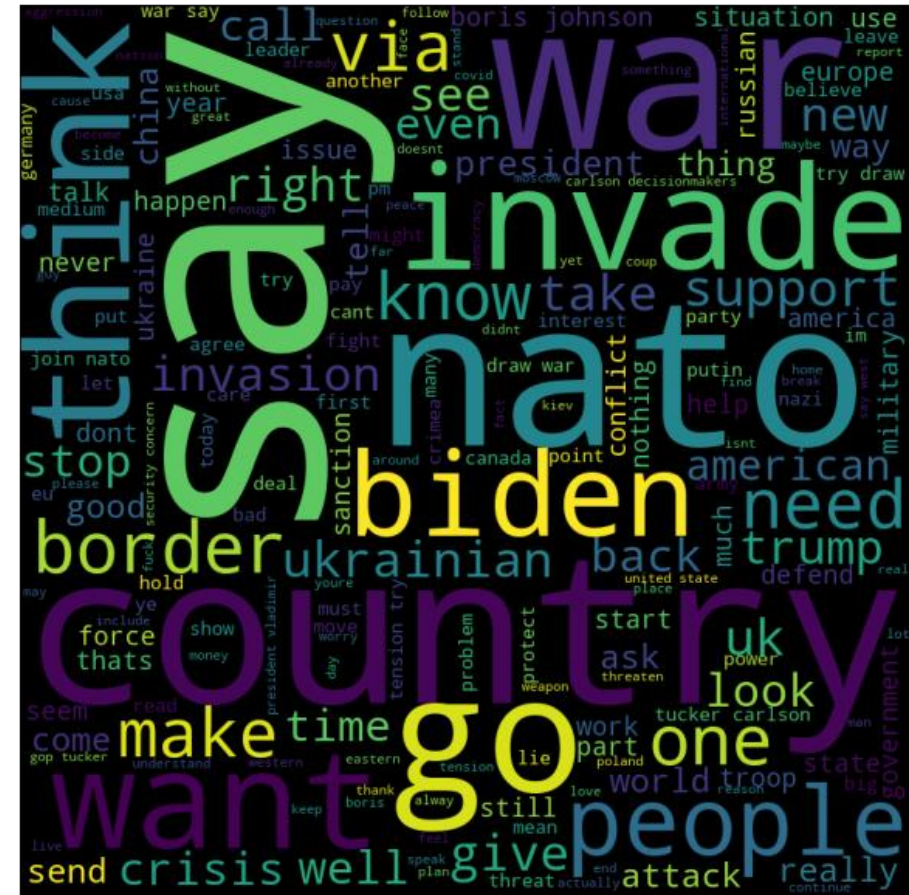
M. en C. Luis Zúñiga



Agenda

- Tipos de Aprendizaje
- Aprendizaje no Supervisado en PLN

- Estudiar información de interés para su extracción.
- Análisis exploratorio: nubes de palabras.
- Análisis exploratorio: modelado de temas.
- Clasificación de polaridad: Machine Learning.
- Exploración de la información clasificado según su polaridad.





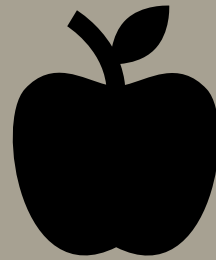
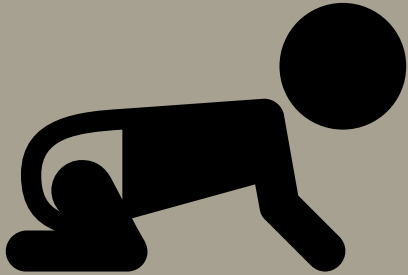
Tipos de Aprendizaje

Recordatorio





Aprendizaje
Supervisado



Aprendizaje
por Refuerzo

supervised learning

Input data



Annotations

These are
apples



Model

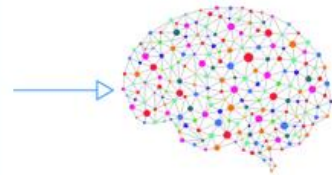


Prediction

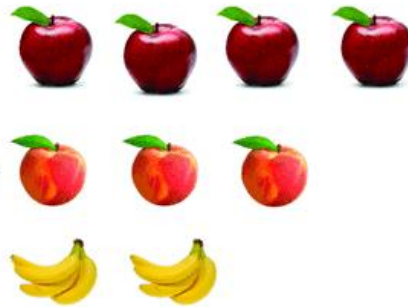
Its an
apple!

unsupervised learning

Input data



Model



Aprendizaje Supervisado



Aprendizaje no Supervisaado

Particularidades del PLN



Aprendizaje no Supervisado

¿Qué alternativas existen al aprendizaje supervisado?

En particular, en el campo del Procesamiento de Lenguaje Natural (PLN), existe una forma de clasificar del tipo no supervisado:

Basado en Reglas

Aprendizaje no Supervisado

- Clasificar requiere aprender, en este caso de un ser humano.
- ¿De qué otra manera se puede transferir este conocimiento?
- ¡Con **lexicones**!

Aprendizaje no Supervisado

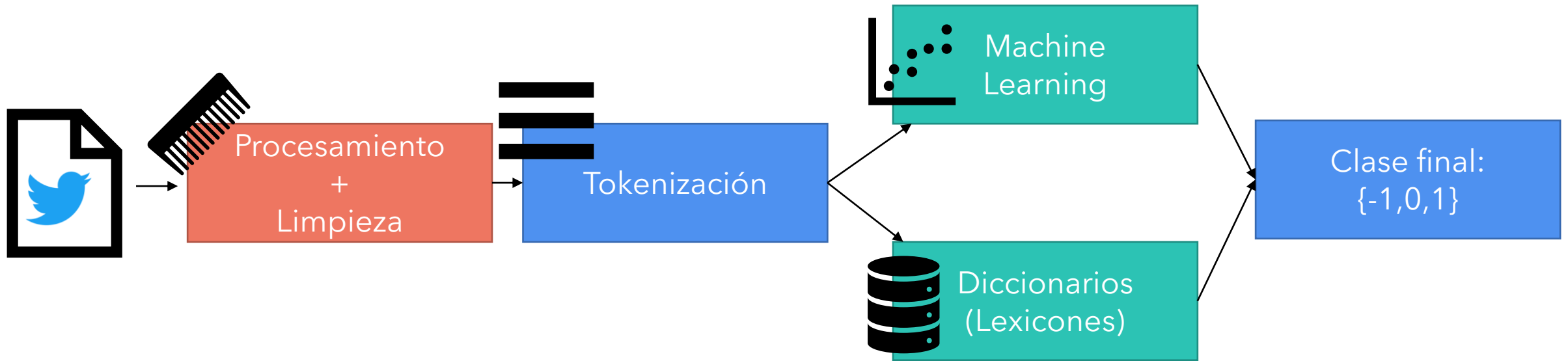
Lexicón¹:

1. m. diccionario (ll repertorio).
2. m. Ling. Conocimiento léxico que un hablante posee sobre una lengua.

Es un repositorio de información léxica que captura relaciones semánticas, categorías y variabilidad.

1. <https://dle.rae.es/lexic%C3%B3n>

Clasificación mediante Lexicones



Clasificación mediante Lexicones

Palabra	Sentimiento
Explosivo	Negativo
Ok	Positivo
:(\	Negativo
Perfecto	Postivo

Clasificación mediante Lexicones

MPQA Opinion Corpus, University of Pittsburgh

- El corpus de opinión de MPQA contiene artículos de noticias de una amplia variedad de fuentes anotadas manualmente para opiniones y otros estados privados (es decir, creencias, emociones, sentimientos, especulaciones, etc.).

Theresa Wilson, Janyce Wiebe, Paul Hoffmann. "Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis".
Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing. 2005

Clasificación mediante Lexicones

SentiWordNet

- Recurso léxico para la minería de opinión el cual se basa (la versión actual) en WordNet 3.0.
- A cada synset de Wordnet se le asignan tres puntuaciones de sentimiento: positividad, negatividad y objetividad.
- Incluido en NLTK.

Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, Manfred Stede. "Lexicon-Based Methods for Sentiment Analysis". *Computational Linguistics* 2011; 37 (2): 267-307.

Clasificación mediante Lexicones

SenticNet

- Iniciativa del MIT Media Laboratory en 2009.
- Diseñado para desarrollar aplicaciones inteligentes que necesiten procesar, identificar y tratar emociones.

Tareas

Revisar y realizar el siguiente tutoriales en Python o Jupyter (Google Colab):

<https://www.analyticsvidhya.com/blog/2021/06/rule-based-sentiment-analysis-in-python/>



Gracias

Luis Zúñiga

lzun.morales@gmail.com