

Modelos lineales y modelos lineales generalizados

Rolando Gonzales Martinez, PhD

Fellow postdoctoral Marie
Skłodowska-Curie

Universidad de Groningen
(Países Bajos)

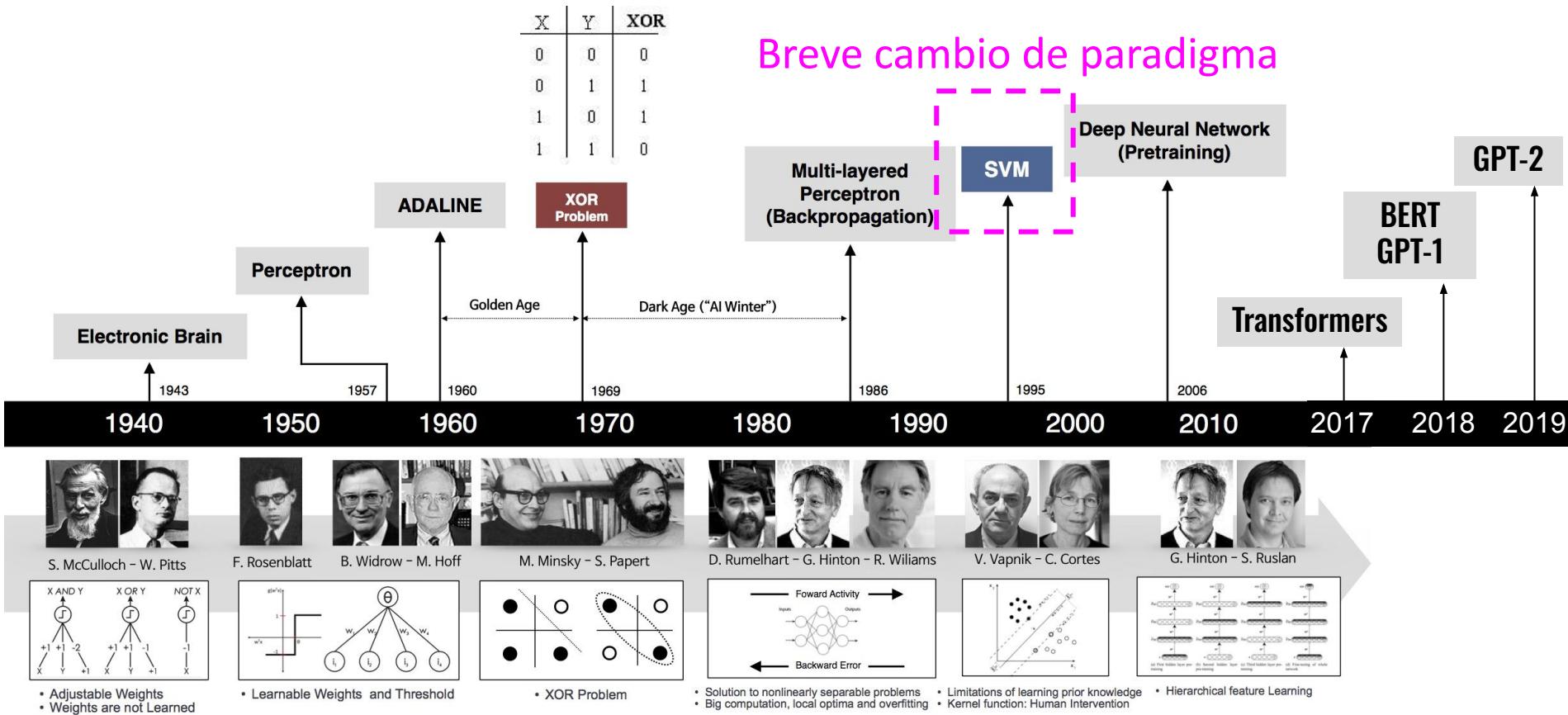
Investigador (researcher)
Iniciativa de Pobreza y Desarrollo
Humano de la Universidad de
Oxford (UK)

Contenido del curso

(6) Modelos de Lenguaje de Gran Escala (LLMs)

- Fundamentos y arquitecturas principales de LLMs: Transformers y GPT.
- Aplicaciones en traducción automática, generación de texto, y comprensión de lenguaje.
- Laboratorio: Aplicación de LLMs en tareas de NLP.
- Aplicaciones Prácticas y Casos de Estudio

Línea de tiempo del aprendizaje profundo y LLMs



Large Language Models (LLMs)

- LLM: Siglas de "Large Language Models": "Modelos de Lenguaje Grandes".
- Un modelo de lenguaje grande (LLM) es un modelo computacional capaz de generar lenguaje u otras tareas de procesamiento del lenguaje natural (NLP).
- Actualmente, los modelos LLM de mayor capacidad están basados en neuronales artificiales construidas con una arquitectura basada en transformadores (transformers).

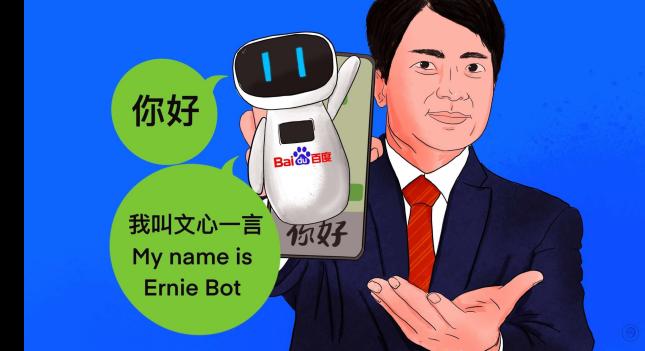
Large Language Models (LLMs)

- Los modelos LLM se entrenan con grandes cantidades de texto.
- Gran número de parámetros (millones), lo que les permite capturar patrones intrincados en el lenguaje y generar texto coherente y contextualmente relevante.
- Una vez pre-entrenados, estos modelos pueden afinarse en tareas específicas utilizando cantidades más pequeñas de datos etiquetados.

LLMs

Ejemplos de LLMs:

- LLaMA (Large Language Model Meta AI)
- Bard/Gemini (basado en LaMDA de Google).
- Enhanced Representation through Knowledge Integration (Ernie Bot)
- ChatGPT
- Claude
- ...



ChatGPT

<https://chat.openai.com/>

- Chat Generative Pre-Trained Transformer.
- 2022: GPT-3.5, 2023: GPT-4. 2024: GPT-4o
- Entrenado en Azure (30 mil NVIDIA GPUs) con información hasta 2021.
- Basado en aprendizaje reforzado con feedback humano, e.g. trabajadores humanos de Kenia que categorizaban contenido ofensivo por \$2 la hora



Bard

- LLM de Google entrenado con gran cantidad de datos de texto. Aun en desarrollo.
- Capaz de comunicarse y generar texto creativo, como poemas, código, guiones, piezas musicales, correo electrónico, cartas, etc.

The image shows a screenshot of the Bard AI interface. At the top left, it says "Bard Experimental". On the right are icons for refresh, help, settings, and a user profile. Below that, there's a "Nuevo chat" button and a "Reciente" section with a green bar containing the message "que es Bard?". The main area shows a conversation where the user asks "que es Bard?" and the AI responds with a text block: "Bard es un gran modelo de lenguaje, también conocido como IA conversacional o chatbot entrenado para ser informativo y completo. Estoy entrenado en una gran cantidad de datos de texto y soy capaz de comunicarme y generar texto similar al humano en respuesta a una". There are buttons for "Ver otras versiones" and a speaker icon. At the bottom, there's an input field "Introduce una petición aquí" with a microphone icon, and a send button with a speech bubble icon.

Gemini Advanced

Gemini Advanced ▾



X

⌚ Gemini se acaba de actualizar. [Ver actualización](#)

exception就是jos except no decuen就是jos except vno就是jos except prevent就是jos
exceptAutowired就是jos exceptjos except就是jos except就是bid就是jos
exceptket就是jos exceptäng就是jos except— !jos except就是thous就是jos
exceptPay就是jos except !njos就是josIGHT就是jos except— !jos except就是äng就是
jos except"></mages就是jos thous就是jos exceptäng就是jos except ii就是jos except
decided就是jos except !njos就是josFull就是jos exceptproperty就是jos exceptFull就是
jos excepthest就是jos exceptPay就是jos except就是jos exceptPayosen就是jos
except guy就是jos except ii就是jos excepthest就是jos except Who就是jos except
prevent就是jos except exceptAutowired就是jos Who就是jos except Who就是jos
except Who就是jos except sua就是jos excepttob就是jos except except就是jos

Introduce una petición aquí



Procesamiento del Lenguaje Natural (NLP)

Rama de la inteligencia artificial y la lingüística computacional que se centra en la interacción entre las computadoras y el lenguaje humano en forma de texto o voz.

El objetivo principal del NLP es permitir que las máquinas comprendan, interpreten y generen lenguaje humano de manera similar a como lo hacen los seres humanos.

NLP implica:

- Análisis sintáctico y morfológico: Comprender la estructura gramatical de las oraciones y la relación entre las palabras.
- Análisis semántico: Interpretar el significado de las palabras y las oraciones en función del contexto.
- Análisis pragmático: Comprender las implicaciones y los aspectos contextuales del lenguaje, como el sarcasmo, la ironía y las inferencias.

Procesamiento del Lenguaje Natural (NLP): Aplicaciones

- Generación de texto: Crear texto de manera coherente y contextualmente relevante.
- Extracción de información: Identificar información clave en un texto, como nombres de personas, lugares o fechas.
- Resumen automático: Crear resúmenes concisos de textos largos.
- Clasificación de texto: Categorizar textos en diferentes clases o etiquetas.
- Análisis de sentimiento: Determinar la actitud emocional expresada en un texto, como positiva, negativa o neutral.
- Respuesta a preguntas: Proporcionar respuestas a preguntas formuladas en lenguaje natural.

Codificadores: Encoders y Decoders

Encoders de texto (codificadores de texto):

- En NLP, los codificadores se utilizan para convertir secuencias de palabras o caracteres en representaciones numéricas que pueden ser procesadas por algoritmos de aprendizaje automático.
- "word2vec" por ejemplo representa las palabras como vectores densos en un espacio vectorial continuo.
- Codificadores más avanzados como LSTM, BERT y GPT (Transformador preentrenado generativo) pueden capturar información contextual compleja y relaciones entre palabras

Análisis de sentimiento

- Técnica NLP que tiene como objetivo determinar la polaridad emocional de un texto, ya sea positiva, negativa o neutra.
- Algoritmos y modelos de aprendizaje automático que pueden clasificar automáticamente el tono emocional de un texto dado.
- Análisis de opiniones en redes sociales, comentarios de productos en marketing.

Análisis de sentimiento

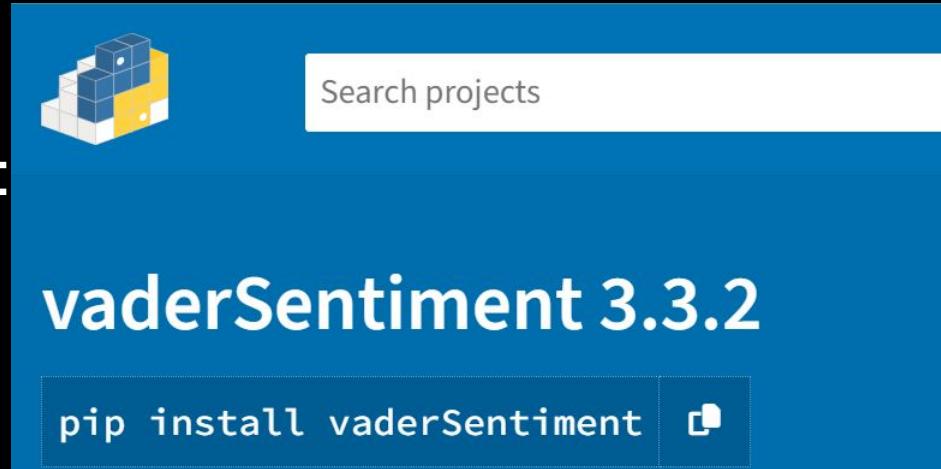
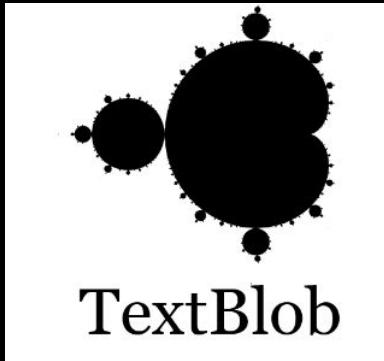
Pasos:

- Preprocesamiento del texto: el texto se limpia y se prepara para el análisis. Esto puede incluir la eliminación de caracteres especiales, tokenización (división del texto en palabras o unidades más pequeñas), y la eliminación de palabras irrelevantes (stop words: y, o).
- Asignación de polaridad: un modelo de aprendizaje automático se aplica para asignar una polaridad emocional a cada fragmento de texto.
- Clasificación: positiva, negativa o neutra, o categorías más finas, como "muy positiva" o "muy negativa". Similar a una escala de Likert

Análisis de sentimiento

Bibliotecas y herramientas:

- VADER
- TextBlob
- NLTK
- spaCy



Tokenizacion

Proceso de dividir un texto en unidades más pequeñas, llamadas "tokens".

En NLP, un token puede ser una palabra individual, una subpalabra, un carácter o incluso una unidad más grande en algunos casos:

- Tokenización de palabras: El texto se divide en palabras individuales.
- Tokenización de subpalabras: "aplicaciones" podría tokenizarse en ["ap", "lic", "acciones"].
- Tokenización en caracteres individuales: por ejemplo, la palabra "hola" se tokenizaría en ["h", "o", "l", "a"].
- Tokenización de unidades más grandes: frases completas pueden ser tratadas como tokens.

Tokenizacion y embedding

"Universidad Mayor de San Andrés"

["Universidad", "Mayor", "de", "San", "Andres"]

"Universidad": [0.12, 0.45, -0.23, ...]

"Mayor": [0.03, 0.67, 0.11, ...]

"de": [-0.34, 0.78, 0.56, ...]

"San": [0.21, -0.89, 0.33, ...]

"Andres": [0.09, 0.23, -0.76, ...]

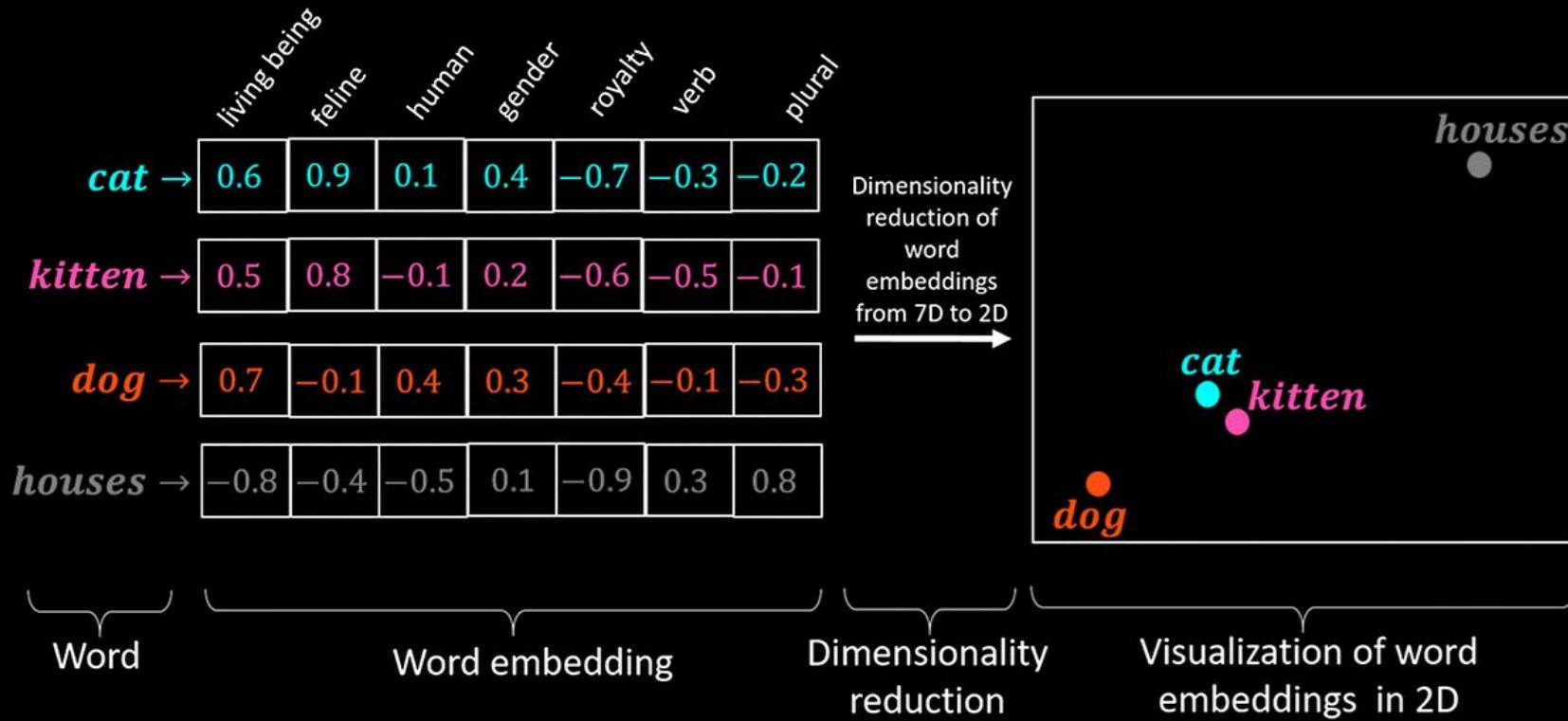
Embedding

- Técnica en el NLP para producir una representación numérica densa y de baja dimensionalidad de un objeto o entidad, como palabras, frases, imágenes u otros elementos, en un espacio vectorial.
- En el contexto del procesamiento de lenguaje natural, los "embeddings" de palabras son representaciones vectoriales de palabras en un espacio continuo y numérico, donde palabras similares en significado están cerca unas de otras en ese espacio.
- Técnica es ampliamente utilizada para capturar la semántica y relaciones entre palabras en una forma que los algoritmos de aprendizaje automático puedan entender y procesar.

Embedding

- Vector numérico denso: no es un vector numérico disperso
- Por ejemplo, en un modelo de lenguaje entrenado, las palabras como "gato" y "perro" podrían tener "embeddings" cercanos en el espacio vectorial, ya que son términos relacionados en el contexto de animales domésticos.
- Del mismo modo, los "embeddings" pueden capturar otras relaciones, como género, pluralidad, y ciertos aspectos de significado.

Tokenización y embedding



Transformers

- Modelos de aprendizaje automático para el procesamiento del lenguaje natural (NLP).
- Introducidos por primera vez en el artículo "Attention is All You Need" publicado por Vaswani et al. en 2017.
- El mecanismo de atención que permite al modelo procesar y entender relaciones de dependencia entre las diferentes partes de una secuencia de entrada.
- A diferencia de los enfoques anteriores que procesan secuencialmente la entrada, los Transformers pueden capturar relaciones de largo alcance entre palabras debido al uso de múltiples capas de redes neuronales

Transformers

Un Transformer consta de dos componentes principales: el codificador y el decodificador:

- El codificador analiza la entrada y genera representaciones contextuales
- El decodificador genera la salida en tareas de generación de lenguaje, como la traducción automática o la generación de texto o imágenes.

Arquitectura de transformers

Transformaciones numéricas
(matriciales)
atención multi-cabeza
con redes neuronales

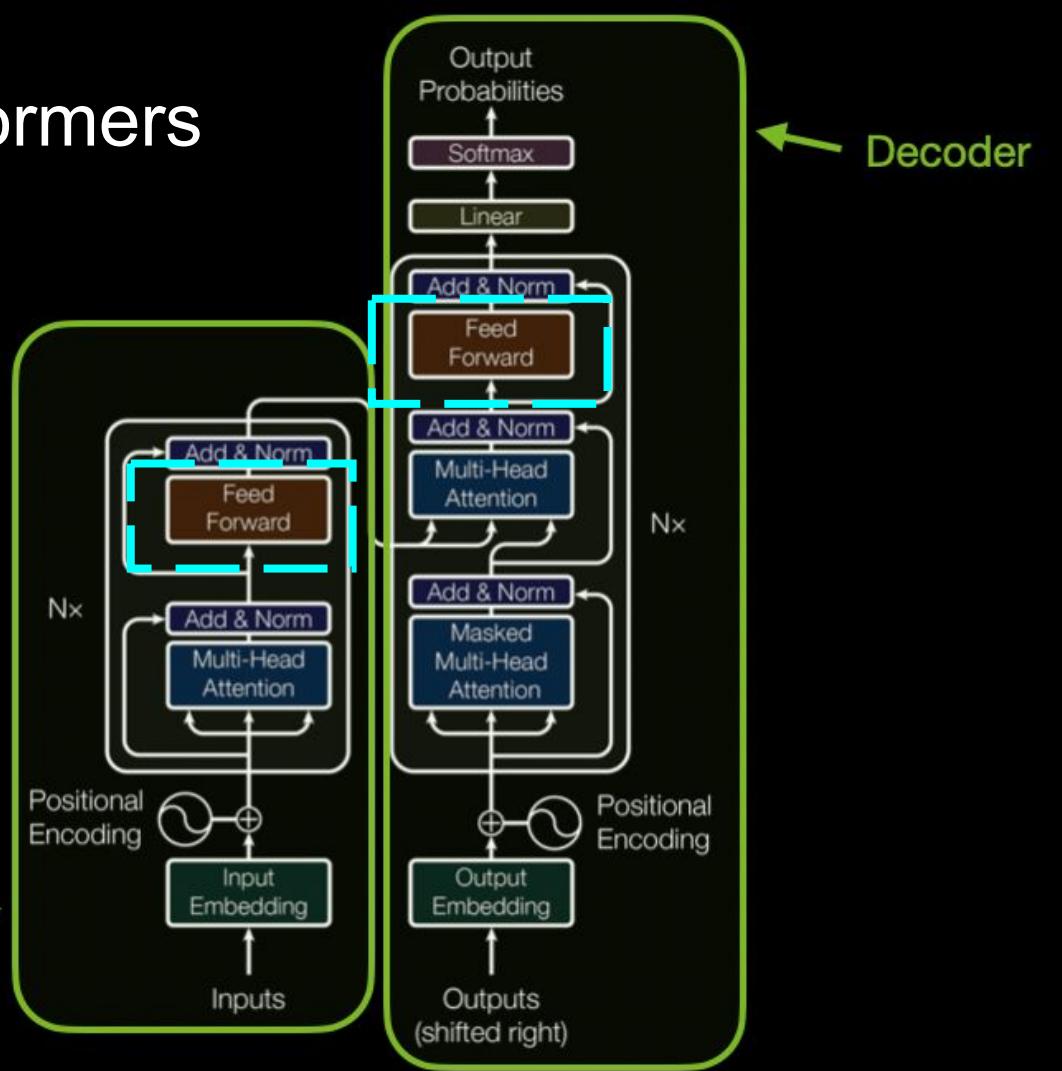
$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)$$

$$\text{where } d_k = \frac{d_{\text{model}}}{h}$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O$$

$$\text{where } \text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$$

Encoder →

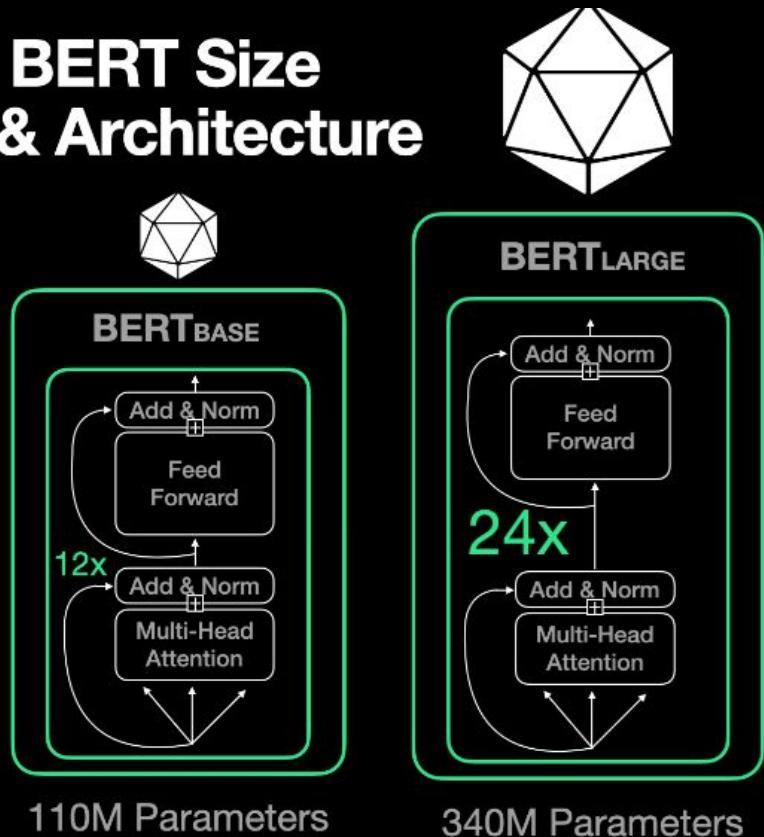


Arquitectura de transformers

El Transformer original dio paso a varias variantes y modelos más avanzados:

- BERT (Bidirectional Encoder Representations from Transformers). Bidireccional: toma en cuenta las palabras anteriores como las posteriores.
- RoBERTa (Robustly Optimized BERT Pretraining Approach).
- BART (Bidirectional and Auto-Regressive Transformers).
- GPT (Generative Pre-trained Transformer) y sus versiones sucesivas como GPT-2 y GPT-3.

BERT Size & Architecture



Aplicaciones en Finanzas

Análisis de sentimiento:

- De acuerdo con la hipótesis del mercado eficiente, toda la información pasada se refleja en los precios de las acciones y la nueva información se absorbe instantáneamente para determinar los precios futuros de las acciones.
- Por tanto, la pronta extracción de sentimientos positivos o sentimientos negativos de las noticias es muy importante para la toma de decisiones de inversión.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 13, 2020, accepted July 1, 2020, date of publication July 16, 2020, date of current version July 29, 2020.
Digital Object Identifier 10.1109/ACCESS.2020.3009626

Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers

**KOSTADIN MISHEV^{D1}, ANA GJORGJEVIKJ^{ID1}, IRENA VODENSKA², LUBOMIR T. CHITKUSHEV²,
AND DIMITAR TRAJANOV^{D1}, (Member, IEEE)**

¹Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, 1000 Skopje, North Macedonia
²Financial Informatics Lab, Metropolitan College, Boston University, Boston, MA 02215, USA
Corresponding author: Kostadin Mishev (kostadin.mishev@finki.ukim.mk)
This work was supported in part by the Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Skopje.

Aplicaciones en Finanzas

Análisis de sentimiento:

Utilizaron encoders especiales,
porque los encoders
predifinidos no distinguijan
entre:

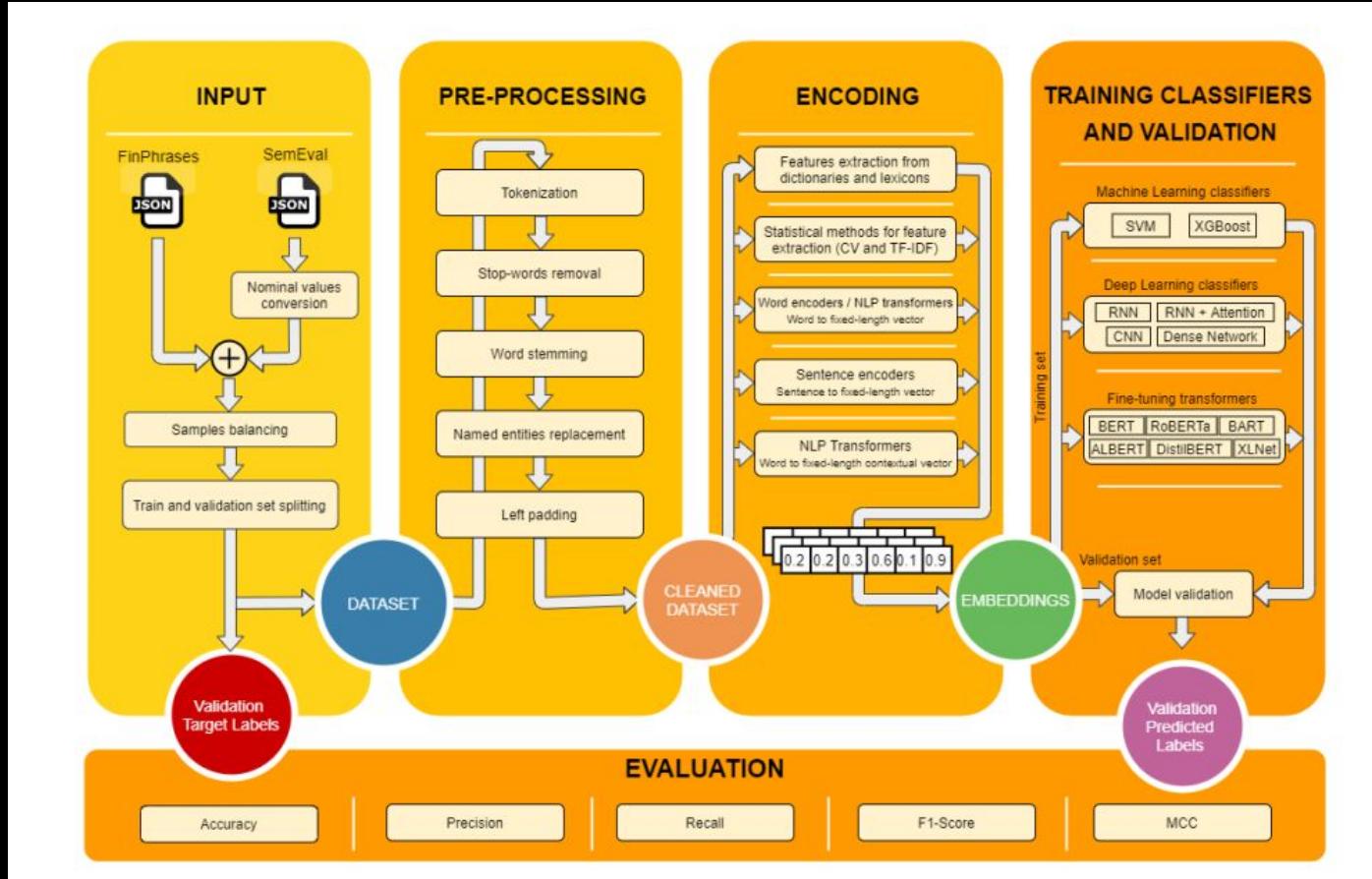
“Apple Inc performed well this year.”

En comparación con:

‘Apple fruits are exported to various countries.’

The image shows the IEEE Access journal cover. At the top, it features the IEEE Access logo with the text "Multidisciplinary | Rapid Review | Open Access Journal". Below the logo, a horizontal line separates it from the publication details. The text "Received June 13, 2020, accepted July 1, 2020, date of publication July 16, 2020, date of current version July 29, 2020." is followed by "Digital Object Identifier 10.1109/ACCESS.2020.3009626". The main title of the article, "Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers", is displayed prominently in large blue text. Below the title, the authors are listed: KOSTADIN MISHEV¹, ANA GJORGJEVIKJ¹, IRENA VODENSKA², LUBOMIR T. CHITKUSHEV², AND DIMITAR TRAJANOV¹, (Member, IEEE). The superscript numbers 1 and 2 indicate the affiliations of the authors. Below the authors' names, there are two small lines of text: "1Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, 1000 Skopje, North Macedonia" and "2Financial Informatics Lab, Metropolitan College, Boston University, Boston, MA 02215, USA". At the bottom of the cover, it says "Corresponding author: Kostadin Mishev (kostadin.mishev@finki.ukim.mk)" and "This work was supported in part by the Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Skopje."

Aplicaciones en Finanzas



Aplicaciones en Finanzas

Análisis de sentimiento:

- Tweets de economistas españoles y sitios de noticias centrados en el sector financiero entre noviembre de 2017 y octubre 2021
- En el proceso de etiquetado se identifica los tweets con los siguientes sentimientos: muy positivo, positivo, neutral, negativo, muy negativo, fuera de dominio y no-sabe-no-responde.

Received 5 January 2023, accepted 7 February 2023, date of publication 10 February 2023, date of current version 15 February 2023.
Digital Object Identifier 10.1109/JACCESS.2023.3244065



Smart Analysis of Economics Sentiment in Spanish Based on Linguistic Features and Transformers

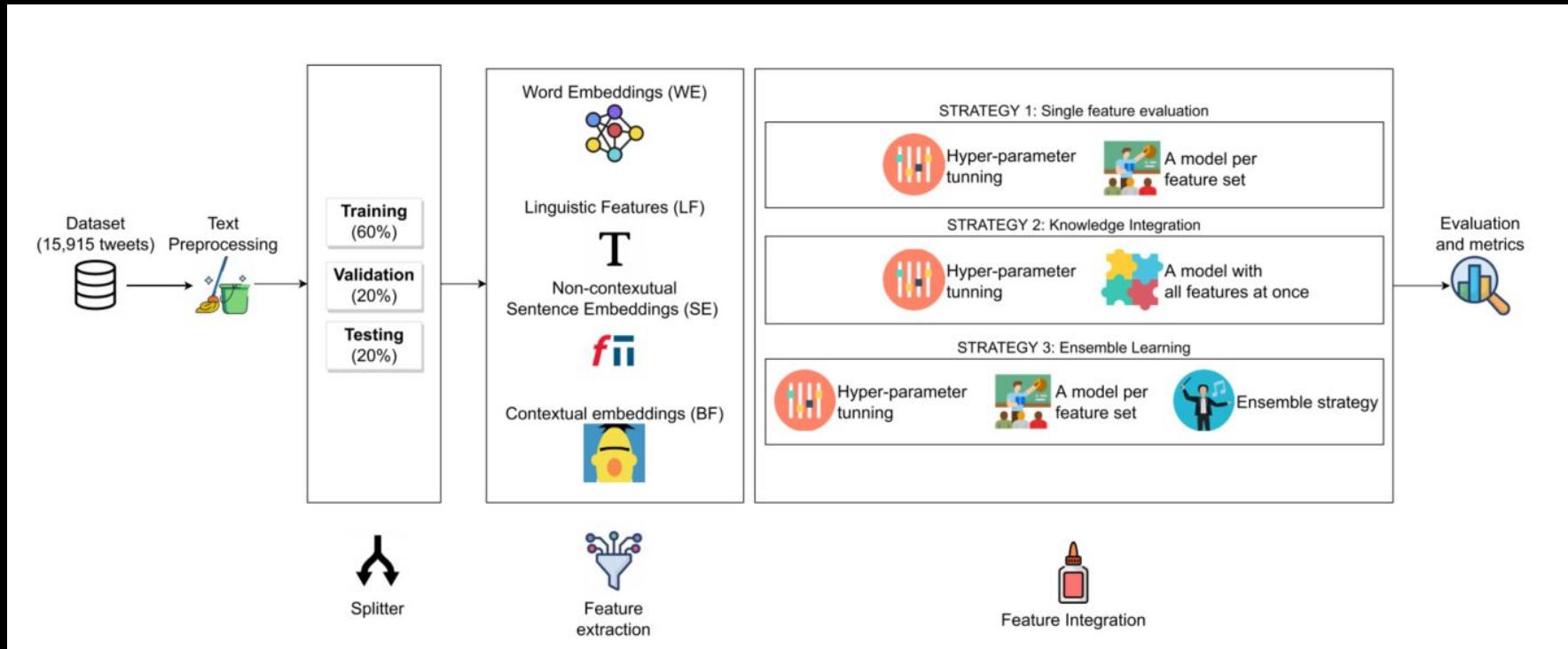
JOSÉ ANTONIO GARCÍA-DÍAZ[✉], FRANCISCO GARCÍA-SÁNCHEZ[✉],
AND RAFAEL VALENCIA-GARCÍA[✉]
Facultad de Informática, Universidad de Murcia, Campus de Espinardo, 30100 Murcia, Spain
Corresponding author: Rafael Valencia-García (valencia@um.es)

This work is part of the research projects AllInFunds (PID2021-121112-I00) and LT-SWM (TED2021-131167B-I00) funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR. This work is also part of the research project LaTe4PSP (PID2019-107652RB-I00/AEI/10.13039/501100011033) funded by MCIN/AEI/10.13039/501100011033. Besides, it was partially supported by the Seneca Foundation-the Regional Agency for Science and Technology of Murcia (Spain)-through project 20963/PI/18. In addition, Jose Antonio García-Díaz is supported by Banco Santander and the University of Murcia through the Doctorado Industrial programme.

ABSTRACT Texts related to economics and finances are characterized by the use of words and expressions whose meaning (and the sentiments they convey) substantially depend on the context. This poses a major challenge to Natural Language Processing tasks in general, and Sentiment Analysis in particular. For low-resource languages such as Spanish, this situation becomes even more acute. Yet, the latest advancements in the field, including word embeddings and transformers, have allowed to boost the performance of Sentiment Analysis solutions. In this work we explore the impact of the combination of different feature sets in the accuracy of Sentiment Analysis in Spanish financial texts. For this, a corpus with 15,915 tweets has been compiled and manually annotated as either positive, negative, or neutral. Then, feature sets based on contextual and non-contextual embeddings along with linguistic features were evaluated both individually and combined. The best results, with a weighted F1-score of 73.15880%, were obtained with a combination of feature sets by means of knowledge integration.

INDEX TERMS Sentiment analysis, financial, transformers, feature engineering, deep learning.

Aplicaciones en Finanzas



Aplicaciones en Negocios

Chapman, C. L., Hillebrand, L., Stenzel, M. R., Deußer, T., Biesner, D., Bauckhage, C., & Sifa, R. (2022, August). Towards generating financial reports from tabular data using transformers.

In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction* (pp. 221-232). Cham: Springer International Publishing.

Insurance and Guarantee Program Liabilities as of September 30, 2020, and 2019

(In billions of dollars)	2020	2019
Insurance and Guarantee Program Liabilities:		
Defined Benefit Pension Plans - Pension Benefit Guaranty Corporation	187.3	181.1
Federal Crop Insurance - Department of Agriculture	7.7	8.9
National Flood Insurance Programs - Department of Homeland Security	2.8	3.4
Ginnie Mae's Mortgage-Backed Securities - Department of Housing and Urban Development	1.3	1.0
Other insurance and guarantee programs	0.2	0.1
Total insurance and guarantee program liabilities	199.3	194.5



[CLS]Insurance and Guarantee Program Liabilities of September 30, 2020 and 2019</section_title> 2020</col_header> 2019</col_header></row> Federal Crop Insurance Department of Agriculture</row_header> 7.7</cell> 8.9</cell></row>[SEP]



Aplicaciones en Negocios

Chapman, C. L., Hillebrand, L., Stenzel, M. R., Deußer, T., Biesner, D., Bauckhage, C., & Sifa, R. (2022, August). Towards generating financial reports from tabular data using transformers. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction* (pp. 221-232). Cham: Springer International Publishing.

[CLS]Insurance and Guarantee Program Liabilities of September 30, 2020 and 2019</section_title> 2020</col_header> 2019</col_header></row> Federal Crop Insurance Department of Agriculture</row_header> 7.7</cell> 8.9</cell></row>[SEP]

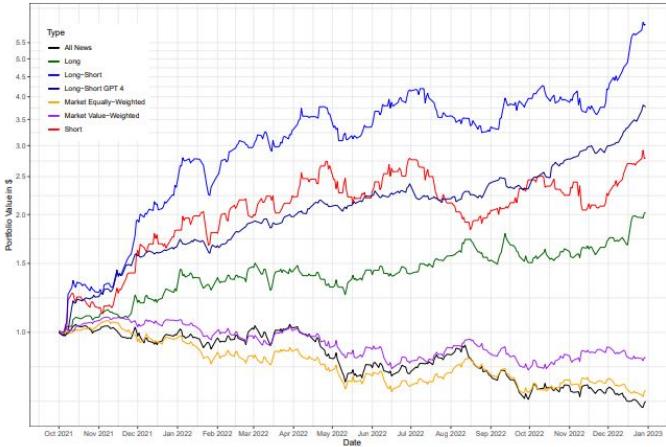
Transformer

"As of September 30, 2020, and 2019, \$7.7 billion and \$8.9 billion, respectively, pertain to USDA's Federal Crop Insurance Program."

ChatGPT

- Pronósticos de precios del mercado de valores

Figure 1: Cumulative Returns of Investing \$1 (Without Transaction Costs)



Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models

29 Pages • Posted: 10 Apr 2023 • Last revised: 12 May 2023

Alejandro Lopez-Lira

University of Florida - Department of Finance, Insurance and Real Estate

Yuehua Tang

University of Florida - Department of Finance

Date Written: April 6, 2023

Abstract

We examine the potential of ChatGPT, and other large language models, in predicting stock market returns using sentiment analysis of news headlines. We use ChatGPT to indicate whether a given headline is good, bad, or irrelevant news for firms' stock prices. We then compute a numerical score and document a positive correlation between these "ChatGPT scores" and subsequent daily stock market returns. Further, ChatGPT outperforms traditional sentiment analysis methods. We find that more basic models such as GPT-1, GPT-2, and BERT cannot accurately forecast returns, indicating return predictability is an emerging capacity of complex models. ChatGPT-4's implied Sharpe ratios are larger than ChatGPT-3's; however, the latter model has larger total returns. Our results suggest that incorporating advanced language models into the

Análisis de sentimiento



Opiniones de cientos de expertos en cambio climático en una encuesta de Damian Carrington

Inteligencia artificial generativa

- Modelos diseñados para crear contenido nuevo y “original”, como imágenes, música, texto o incluso videos, imitando el estilo y la estructura de los datos con los que fueron entrenados.
- Estos modelos funcionan mediante la generación de datos que no han sido vistos previamente en el conjunto de entrenamiento, lo que los hace capaces de producir resultados creativos y sorprendentes.



Midjourney y DALL-E

- Midjourney y DALL-E son programas de inteligencia artificial generativa para crear imágenes a partir de descripciones textuales
- Estas descripciones textuales se llaman *prompts*



Midjourney

Midjourney está basado en LLM y modelos de difusión



DALL-E

<https://labs.openai.com/>

- DALL-E es parte de OpenAI
- DALL-E es una implementación multimodal de GPT-3 con 12 mil millones de parámetros.
- El código fuente no está disponible.



 DALL·E 2

Aplicaciones de imágenes generadas con IA en negocios

Reducción de costos
aumentos de la
productividad en las
empresas en los ámbitos
del:

- Marketing
- Diseño

Particularmente útil para
start-ups, PyMEs



Examen teorico



Examen teorico



Early identification of total loss through image analysis

What is the business problem?

Customer experience is of the utmost importance for LV=, therefore we are constantly seeking new ways to exploit data to improve accuracy and efficiency of all domains across the business. One of such areas is to rapidly identify claims that are total loss. In this project, for the first time we utilise customer images as well as more traditional structured data in order to classify claims as total loss or non total loss.

Creating training data

Before we started collecting training data, we had to make a list of decisions such as:

- What image labels should we ask for?
- How many images do we need?
- Who is labelling the images?
- How can this be scaled?

Building the model

In order to automatically analyse the content of the images, we use Convolutional Neural Networks (CNNs). However, instead of training our network from scratch, we did use transfer learning - we took a pretrained network and transferred the knowledge it has gained by initially being trained on a very large dataset to our custom classification problem by using much smaller training set we have gathered with our .Net labelling app.

Diagram illustrating the image analysis process:

- Input:** Input photo.
- Process:** Convolution + ReLU, Pooling, Convolution + ReLU, Pooling, Flatten, Fully Connected, Connected.
- Output:** Heatmap.

Initial filtering:

- Claims with images → Structured data model, Image models.
- Structured data model, Image models → Combine structured data and image models.
- Combine structured data and image models → Non Total Loss, Total Loss.

When submitting a claim, customers get asked to upload multiple images taken from different angles in order to make an accurate decision on the case.

However, the aim of this project is to categorise a given «claim» as total loss or non total loss. We therefore had to come up with a clever way to combine image model outputs at claim level without losing vital information extracted from individual models we built to analyse images.

Examen teorico

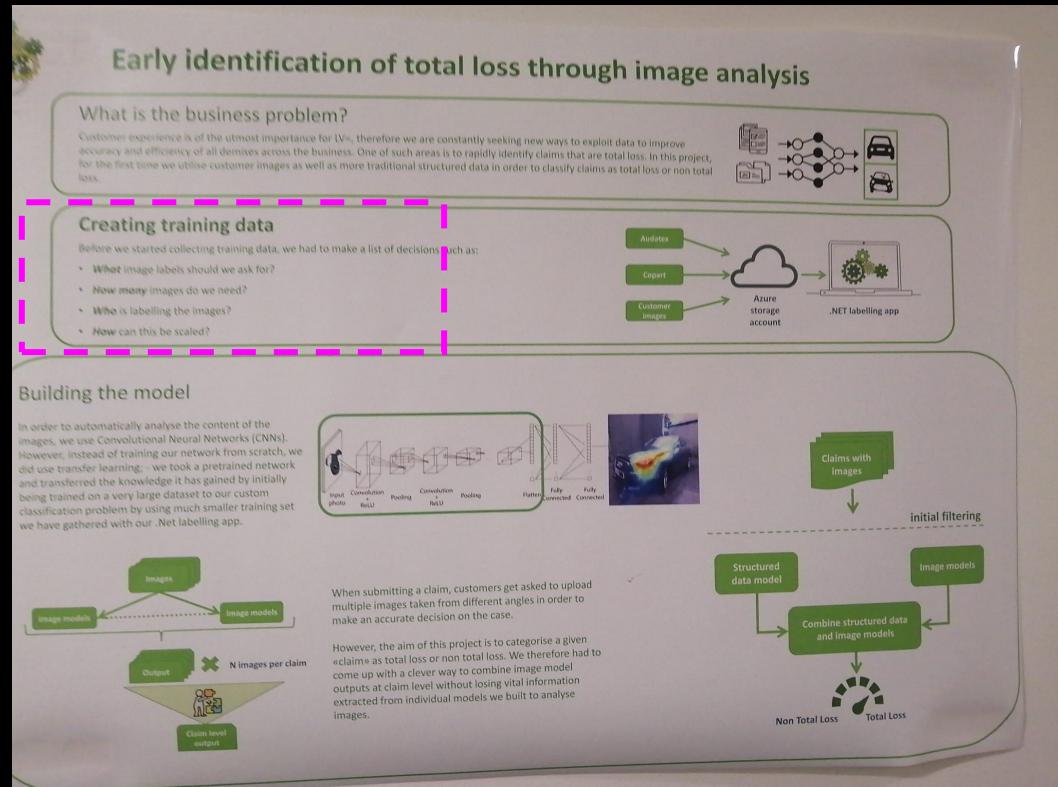
1. Objetivo de tener un conjunto de entrenamiento

¿Cuál es el objetivo principal de usar un conjunto de entrenamiento en Machine Learning?

a) Ajustar el modelo a los datos conocidos para que pueda aprender patrones y relaciones.

b) Evaluar el rendimiento final del modelo antes de su uso en producción.

c) Incrementar la complejidad computacional



Examen teorico

2. Partición de datos

¿Cuál es el objetivo principal de dividir los datos en conjuntos de entrenamiento, validación y prueba?

- a) Minimizar el sobreajuste (overfitting)
- b) Aumentar el rendimiento computacional
- c) Reducir el tiempo de entrenamiento

Examen teorico

3. Métricas de desempeño en Machine Learning

¿Cuál de las siguientes métricas es más apropiada para un problema de clasificación binaria desbalanceada?

- a) Exactitud (Accuracy)
- b) Error cuadrático medio (MSE)
- c) F1-Score

Examen teorico

4. Exactitud en la matriz de confusión

¿Cuál de las siguientes afirmaciones es correcta acerca de la exactitud (accuracy) en un problema de clasificación?

- a) La exactitud mide la proporción de predicciones correctas sobre el total de predicciones.
- b) La exactitud solo mide la tasa de verdaderos positivos
- c) La exactitud solo mide la tasa de verdaderos negativos

Examen teorico

5. Tipos de regularización

¿Qué tipo de regularización agrega una penalización igual a la suma de los valores absolutos de los coeficientes?

- a) L1 (Lasso)
- b) L2 (Ridge)
- c) Dropout

Examen teorico

6. Validación cruzada

En una validación cruzada estratificada, ¿cuál es el objetivo de "estratificar" los datos?

- a) Asegurar que cada fold tenga una distribución similar de clases.
- b) Reducir el tiempo de cómputo.
- c) Hacer que el modelo sea más simple.

Examen teorico

7. Hypertuning (tuneo de hiperparámetros)

¿Qué técnica se utiliza comúnmente para ajustar los hiperparámetros de un modelo de Machine Learning?

- a) Descomposición de valores singulares (SVD)
- b) Grid Search
- c) Algoritmo de retropropagación

Examen teorico

8. Arquitectura de redes neuronales artificiales

¿Qué define el término "capa oculta" en una red neuronal artificial?

- a) La capa que recibe las entradas iniciales.
- b) Cualquier capa que esté entre la capa de entrada y la capa de salida.
- d) La capa encargada de normalizar los datos.

Examen teorico

9. Entrenamiento de redes neuronales

¿Cuál es el propósito de la retropropagación en el entrenamiento de redes neuronales?

- a) Ajustar los pesos mediante la propagación del error hacia atrás a través de las capas de la red.
- b) Aumentar la cantidad de neuronas en las capas ocultas.
- c) Reducir el número de capas ocultas en la red.

Examen teorico

10. Funciones de activación

¿Cuál es una ventaja clave de la función de activación ReLU frente a la función sigmoide?

- a) ReLU no tiene problemas con el "vanishing gradient" en capas profundas.
- b) ReLU es más costosa computacionalmente.
- c) ReLU siempre produce valores negativos.

