



UNIVERSIDAD DE CHILE

Deep Learning

Deeper, Better, _____ , Stronger than Machine Learning

Valentin Barriere

Universidad de Chile – DCC

CC6204, Primavera 2024

Proposición de temas de clase

El programa

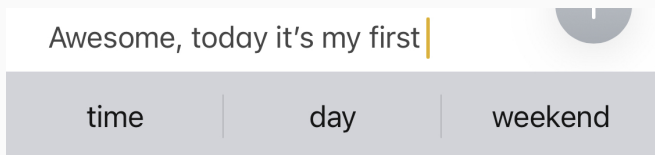
Todo no esta definido! Que quieren ver después las bases?

- **Classical language modeling:** from Shannon to LLM-based Chatbot
- **Audio models:** audio speech recognition, source separation, audio scene classification, general models
- **Multimodal models:** classical ones, CLIP, Diffusion, LLM-based, generative ones for vision or audio
- **Large text generative models (LLM):** Instructions, RLHF, Agent using tools, Retrieved Augmented Generation
- **Efficiency:** how to train/deploy an LLM on a small budget (parameter efficient fine tuning, quantization, optimized specific libraries like vLLM)

Les voy a mandar un encuesta!

Modelos de Lenguaje: what is it?

The goal of a language model is to predict the following word knowing the past ones in a sentence, or longer context (book, conversation, etc...)



$$P(w_{t+1} | w_t, \dots, w_1)$$

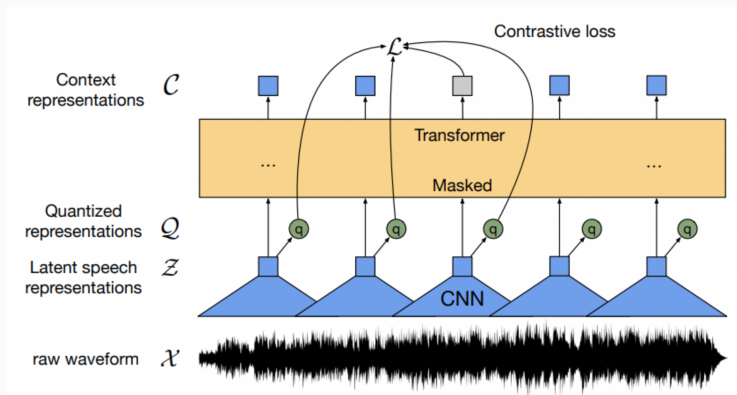
This is the base of all the Large **Language Models!!**

Modelos de Lenguaje: how to do it?

- Naive n-gram model (based on probabilities of transition between words)
- Neural n-gram
- Classical RNN, and RNN-LSTM
- Transformers
- Emerging abilities

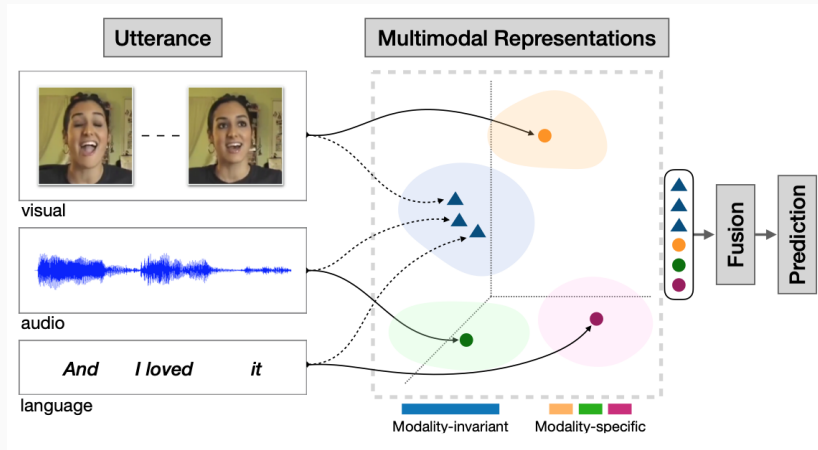
Modelos Audio

- waveNet (CNN): Voice generation
- Jukebox (Transformer): music generation
- Wav2Vec2 (Transformer): Audio Speech Recognition
- Audio Scene Classification



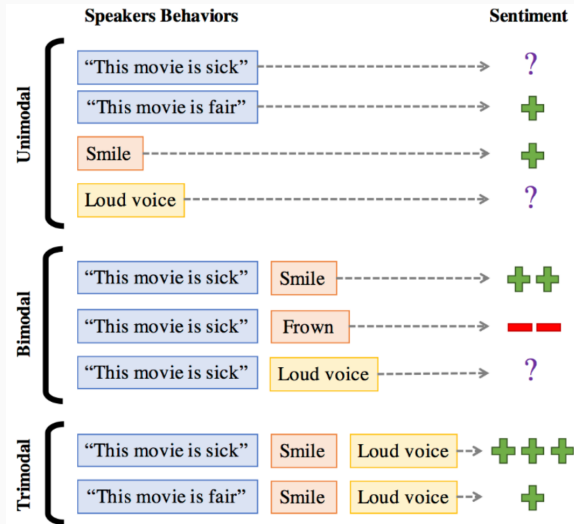
Modelos Multimodales

Modelos que entienden varios tipos de datos: textos, audio, imágenes, vídeos, etc...



Modelos Multimodales

Modelos que entienden varios tipos de datos: textos, audio, imágenes, vídeos, etc...



Modelos Multimodales: from LLM to LMM

ViLBERT

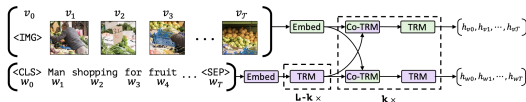


Figure 1: Our ViLBERT model consists of two parallel streams for visual (green) and linguistic (purple) processing that interact through novel co-attentional transformer layers. This structure allows for variable depths for each modality and enables sparse interaction through co-attention. Dashed boxes with multiplier subscripts denote repeated blocks of layers.

Modelos Multimodales: from LLM to LMM

ViLBERT

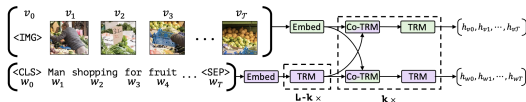
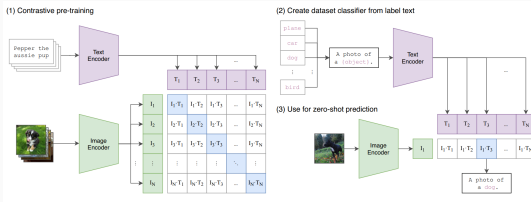


Figure 1: Our ViLBERT model consists of two parallel streams for visual (green) and linguistic (purple) processing that interact through novel co-attentional transformer layers. This structure allows for variable depths for each modality and enables sparse interaction through co-attention. Dashed boxes with multiplier subscripts denote repeated blocks of layers.

Contrastive Language-Image Pre-Training



Modelos Multimodales: from LLM to LMM

VILBERT

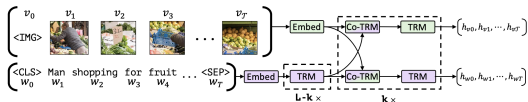
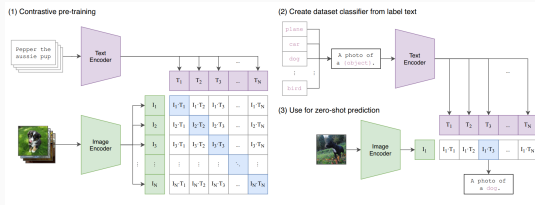
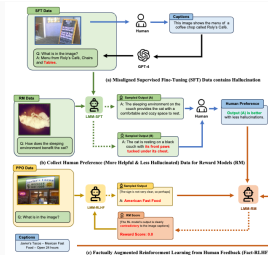


Figure 1: Our ViLBERT model consists of two parallel streams for visual (green) and linguistic (purple) processing that interact through novel co-attentional transformer layers. This structure allows for variable depths for each modality and enables sparse interaction through co-attention. Dashed boxes with multiplier subscripts denote repeated blocks of layers.

Contrastive Language-Image Pre-Training



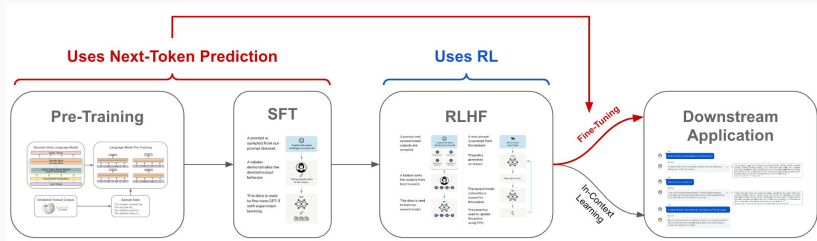
Llava (LLM-based)



Modelos de Lenguaje Largos

Como ir de un LLM básico a un modelo que puede hablar con los usuarios:

- Instrucciones
- Reinforcement Learning on Human Preferences
- LLM using tools
- Retrieved Augmented Generation



Modelos de Lenguaje Largos: RLHF

Step 1

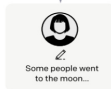
Collect demonstration data, and train a supervised policy.

SFT

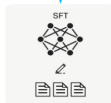
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

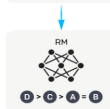
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

RLHF

A new prompt is sampled from the dataset.



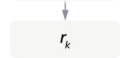
The policy generates an output.



The reward model calculates a reward for the output.

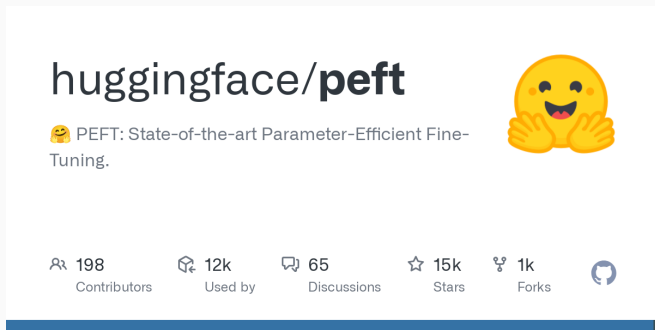


The reward is used to update the policy using PPO.



How to train/deploy an LLM on a small budget?

- Parameter Efficient Fine Tuning: Huggingface's [PEFT](#)
- Quantization
- Adapter
- Soft-prompting
- vLLM (faster inference)



Questions?

