

# Applied Deep Learning for NLP

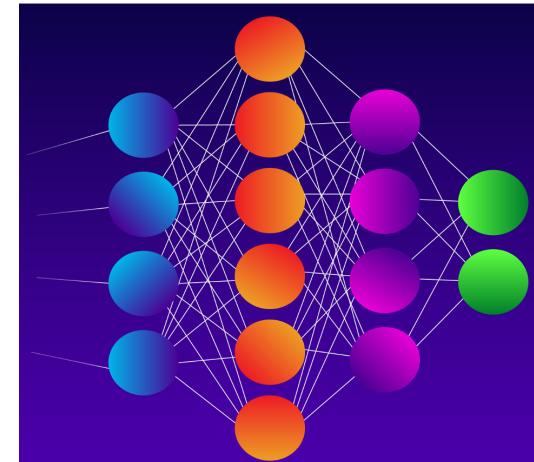
## Week 12 - NLP Outro

Juan Carlos Medina Serrano

Technische Universität München  
Hochschule für Politik  
Political Data Science

Munich, 21. January 2021

political  
data  
science



## Multimodal Tasks

Tasks that include more than only text! Mainly images and videos.

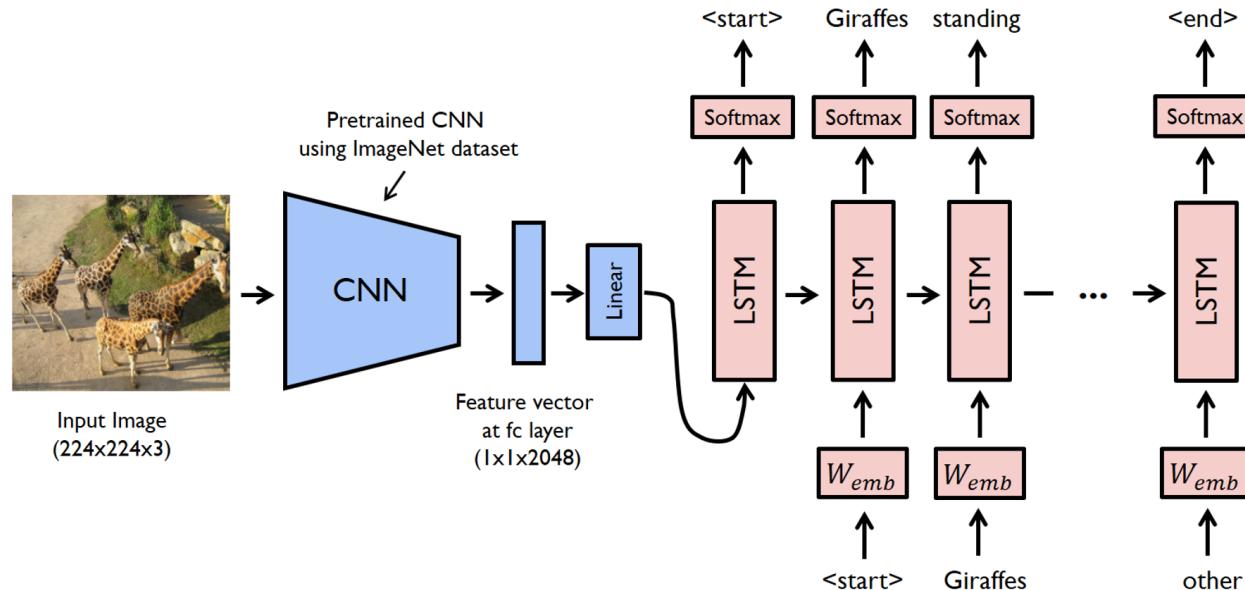
## Image/Video Captioning

Find captions for an image/video

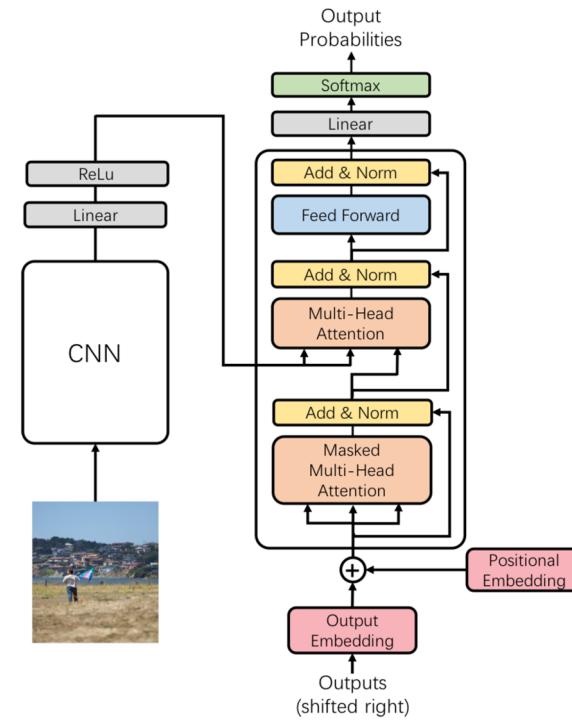


Prediction Caption: the person is riding a surfboard in the ocean <end>

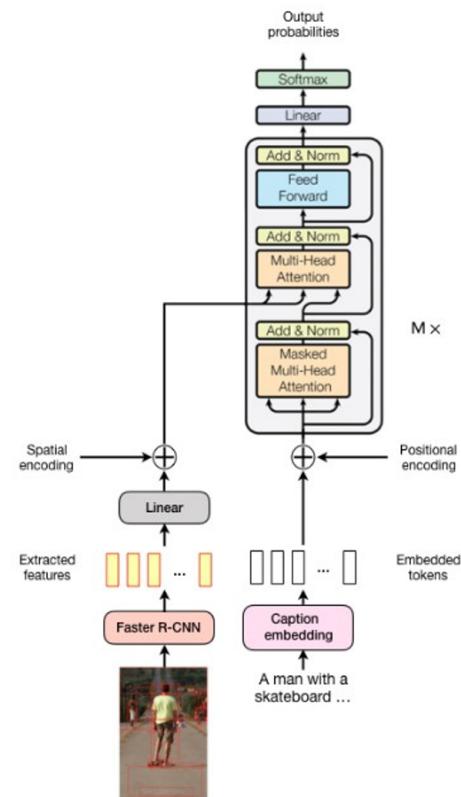
## Image/Video Captioning



## Image/Video Captioning

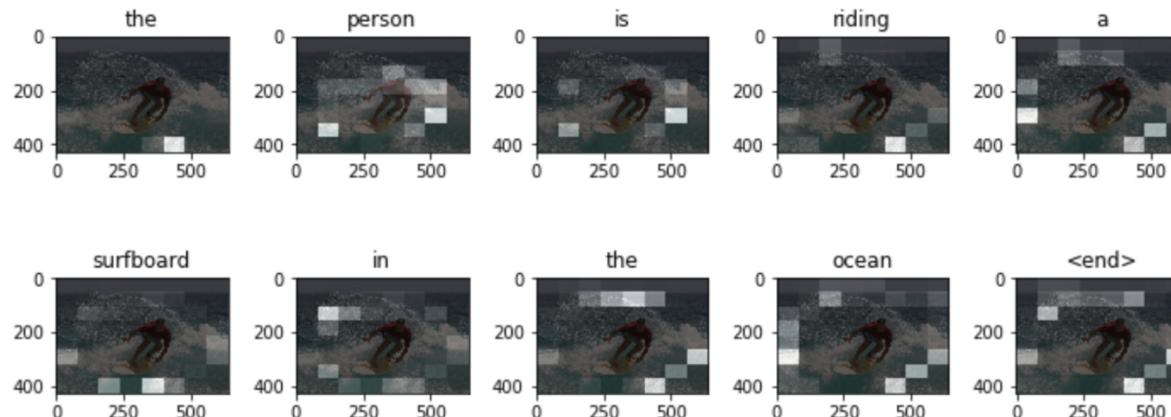


## Image/Video Captioning



## Image/Video Captioning

Prediction Caption: the person is riding a surfboard in the ocean <end>



[https://www.tensorflow.org/tutorials/text/image\\_captioning](https://www.tensorflow.org/tutorials/text/image_captioning)

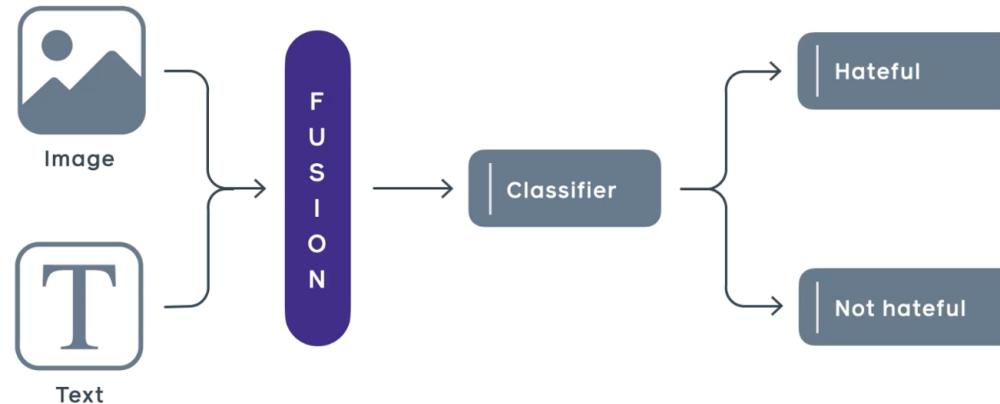
## Multimodal Classification

Facebook's Hateful Meme Dataset



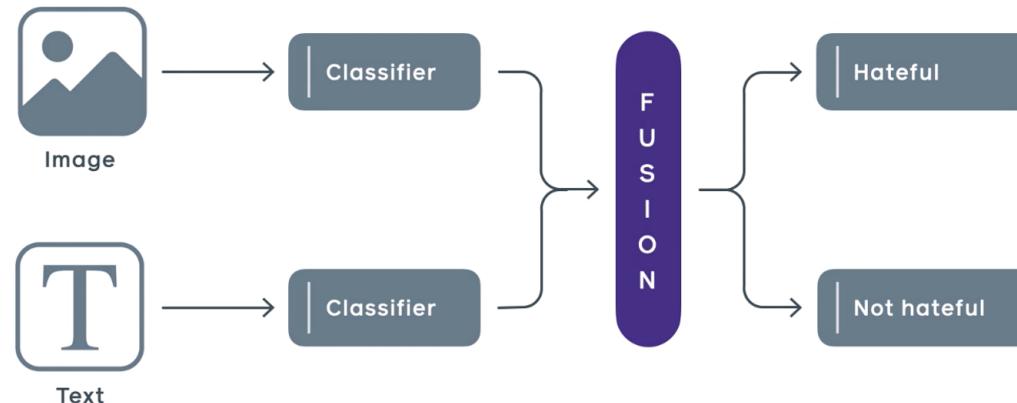
## Multimodal Classification

### Early fusion



## Multimodal Classification

### Late fusion



## Visual Question Answering

### Visual Question Answering

Visual Question Answering (VQA) is the task of generating a answer in response to a natural language question about the contents of an image. VQA models are typically trained and evaluated on datasets such as VQA2.0, GQA, Visual7W and VizWiz.

This page demonstrates ViLBERT for VQA. ViLBERT (short for Vision-and-Language BERT), is a model for learning task-agnostic joint representations of image content and natural language.

<https://demo.allennlp.org/visual-question-answering>

## Multitask Learning

Different to multimodal. It aims to learn multiple different tasks simultaneously while maximizing performance on one or all of the tasks.

T5 is a good example of a multitask system.

## Storytelling

Use a prompt to generate a story. Input can be the beginning of a story (story continuation) or even an image.

\*Generating Stories about Images (using skip-thought vectors)

<https://medium.com/@samim/generating-stories-about-images-d163ba41e4ed>

\*Event Representations for Automated Story Generation with Deep Neural Nets,

<https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17046/15769>

\* Generate recipes or poetry

## Topic Modeling

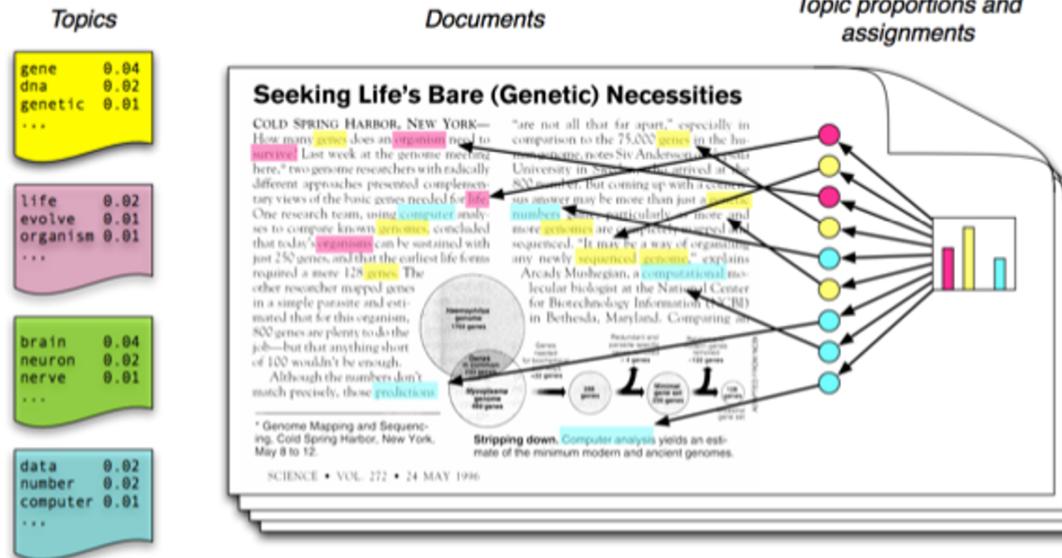


Figure source: Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.

## Topic Modeling

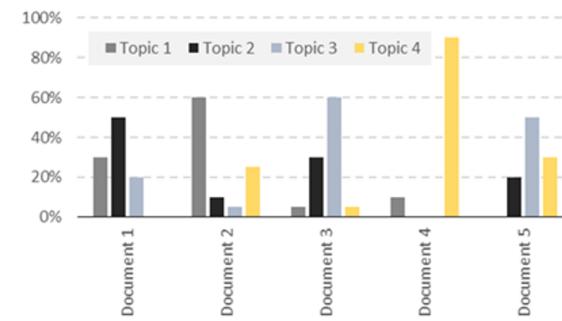


LDA

Creation of topics

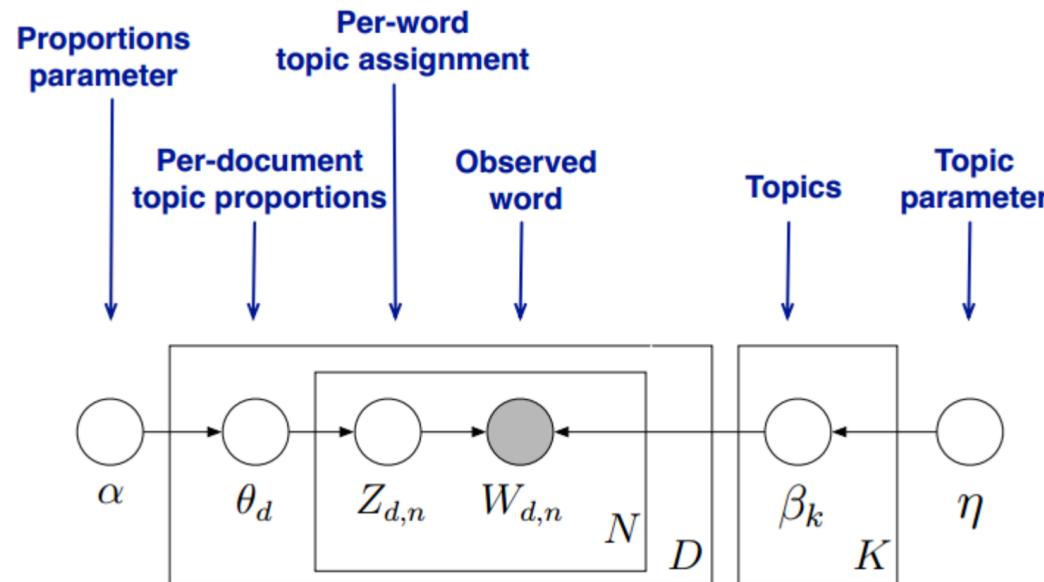
	weight	words
Topic 1	3%	flower
	2%	rose
	1%	plant
...		
Topic 2	2%	company
	1%	wage
	1%	employee

Topics allocation to documents



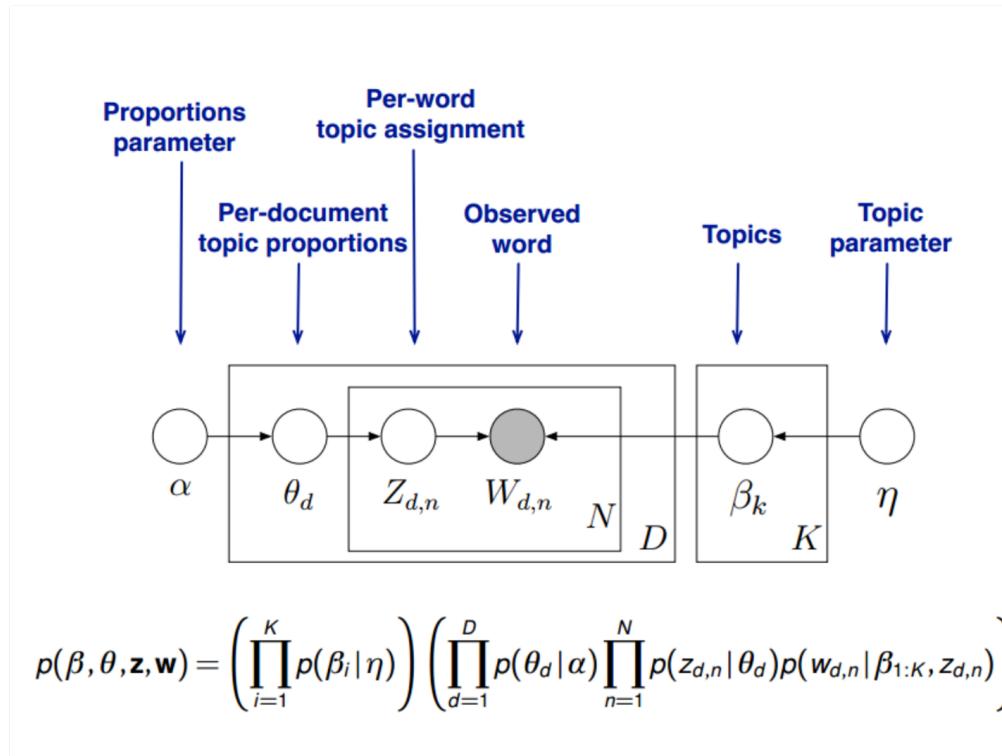
## LDA

Latent Dirichlet Allocation



## LDA

Latent Dirichlet Allocation



## Neural Topic Modeling?

Not there yet. But, research going in that direction

Idea: Use a smart clustering of embeddings to obtain topic representations

<https://arxiv.org/pdf/2004.14914.pdf>

## Other Linguistic Tasks

- ▶ **Natural Language Inference:** Determining whether a *hypothesis* is true (entailment), false (contradiction), or undetermined (neutral) given a *premise*

Premise	Label	Hypothesis
A man inspects the uniform of a figure in some East Asian country.	contradiction	The man is sleeping.
An older and younger man smiling.	neutral	Two men are smiling and laughing at the cats playing on the floor.
A soccer game with multiple males playing.	entailment	Some men are playing a sport.

- ▶ **Semantic Similarity** Find similar texts/questions/captions
- ▶ **Stance Detection** Extraction of a subject's reaction to a claim made by a primary actor (agree/disagree)

## Linguistic Evaluation

**GLUE** and **SuperGLUE** are sets of tasks to evaluate new NLP models

Leaderboards:

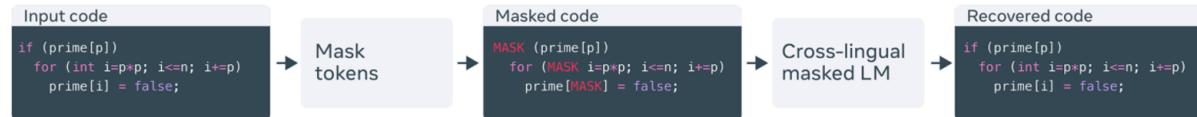
Rank	Name	Model	Score
1	HFL iFLYTEK	MacALBERT + DKM	90.7
2	Alibaba DAMO NLP	StructBERT + TAPT	90.6
3	PING-AN Omni-Sinitic	ALBERT + DAAF + NAS	90.6
4	ERNIE Team - Baidu	ERNIE	90.4
5	T5 Team - Google	T5	90.3

Name	Model	Score
SuperGLUE Human Baselines	SuperGLUE Human Baselines	89.8
T5 Team - Google	T5	89.3
Huawei Noah's Ark Lab	NEZHA-Plus	86.7
Alibaba PAI&ICBU	PAI Albert	86.1
Tencent Jarvis Lab	RoBERTa (ensemble)	85.9

## Neural Translation for Other Tasks

Example: Transcoder

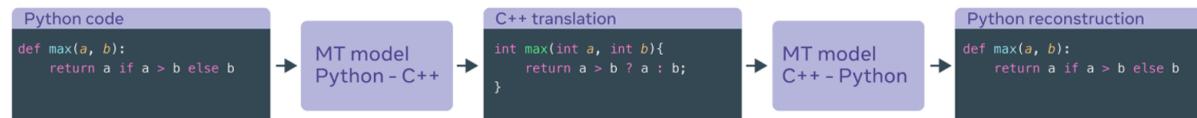
### Cross-lingual masked model pretraining



### Denoising auto-encoding

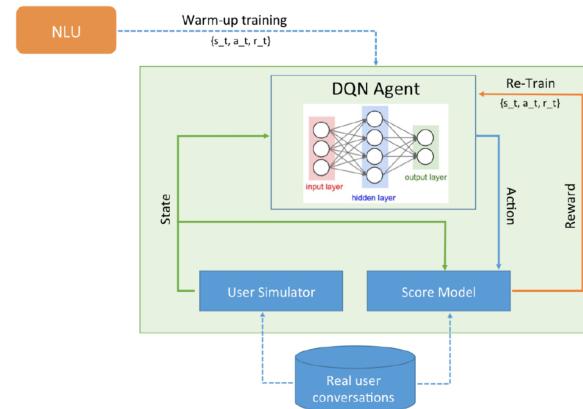


### Back - translation



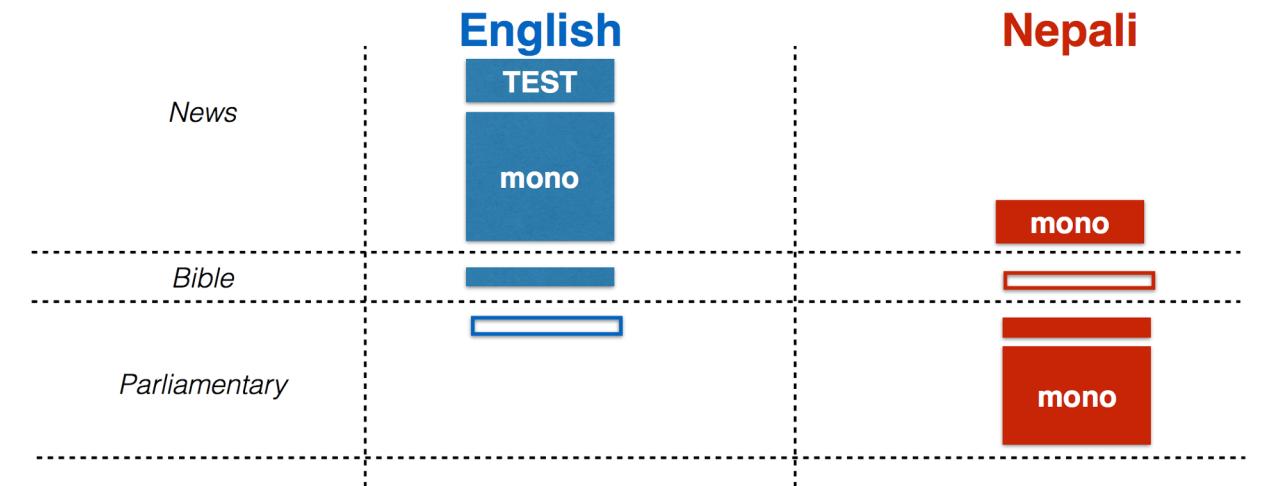
## Deep Reinforcement Learning in NLP

- ▶ Needs specific scores to be used as rewards for the RL agent (Hard to define)
- ▶ Policy learning can be implemented using a Deep Q-Network (DQN)
- ▶ Can be used in all NLG tasks (summarization, translation, dialogue)
- ▶ For a Chatbot:
  - ▶ State: The previous two dialogue turns.
  - ▶ Action: Indefinite
  - ▶ Reward:
    - ▶ Ease of answering: Avoid utterance with a dull response
    - ▶ Information Flow: Penalize semantic similarity between consecutive turns
    - ▶ Semantic Coherence: Penalize ungrammatical or not coherent phrases



## Low Resource NLP

Translation: Number of parallel sentences is in the order of 10,000 or less



## Bias in NLP

He is...



She is...



## Bias in NLP

```
text_to_sentiment("My name is Emily")
```

2.2286179364745311

```
text_to_sentiment("My name is Heather")
```

1.3976291151079159

```
text_to_sentiment("My name is Yvette")
```

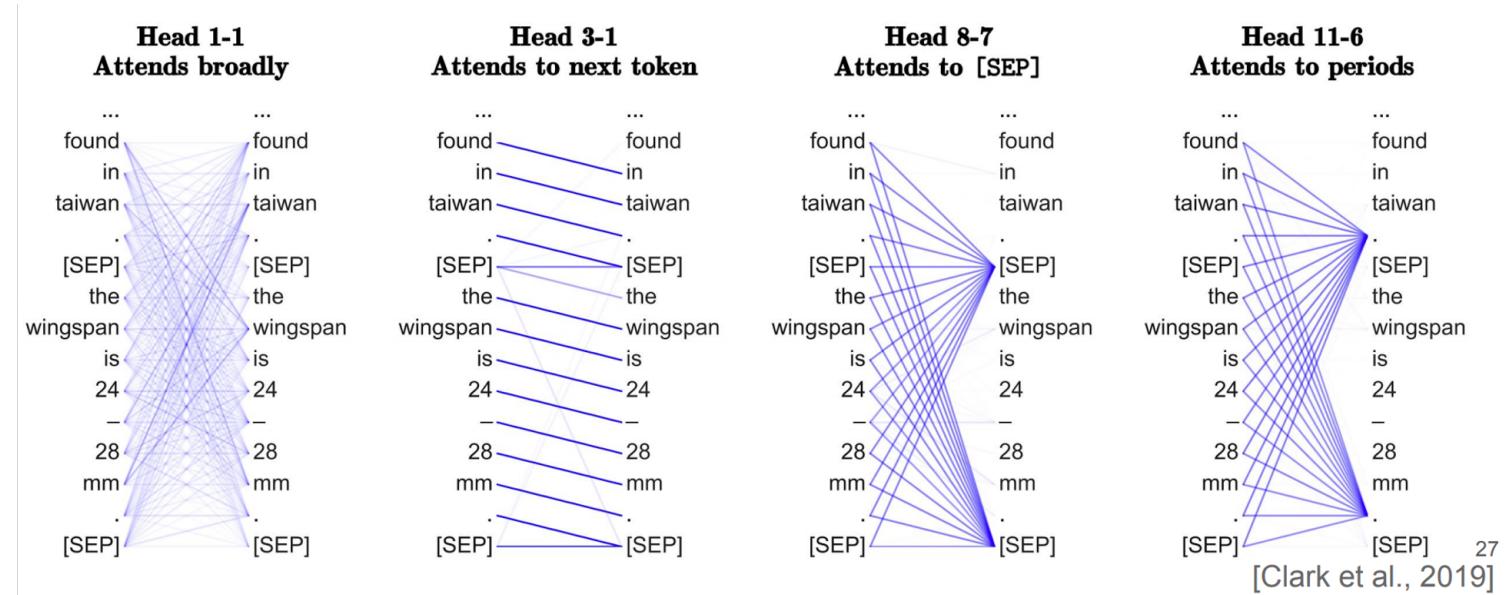
0.98463802132985556

```
text_to_sentiment("My name is Shaniqua")
```

-0.47048131775890656

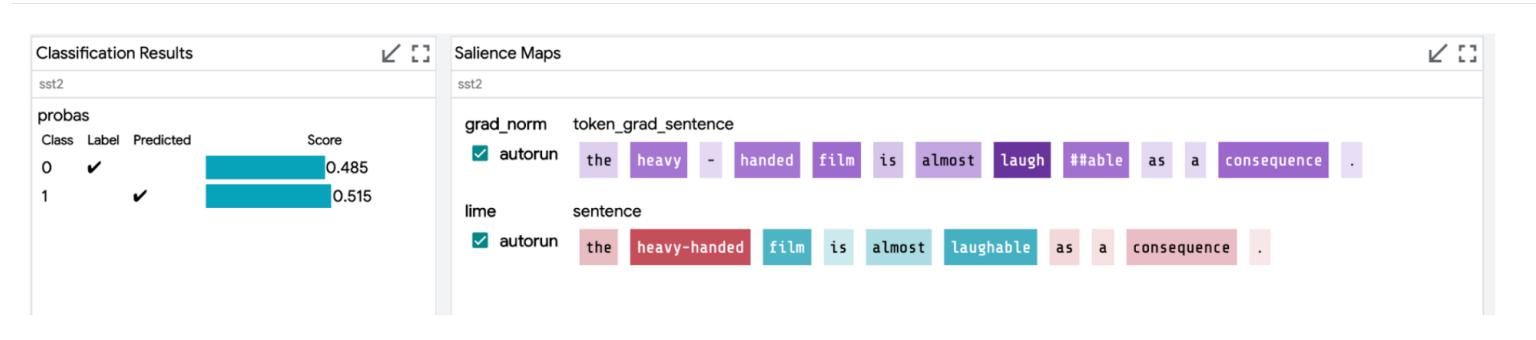
## Interpretability of NLP

Buzzword: Explainable AI



## Interpretability of NLP

### Google's Language Interpretability Tool (LIT)



- ▶ Examine local explanations using **saliency maps**
- ▶ Test model behavior with perturbed (counterfactual) examples using techniques such as back-translation, word replacement, or adversarial attacks

<https://ai.googleblog.com/2020/11/the-language-interpretability-tool-lit.html?m=1>

This is the end

