DEEP LEARNING CIn.AI Centro de Informática UFPE Universidade Federal de Pernambuco

Processamento de Linguagem Natural

Geração Automática de Texto

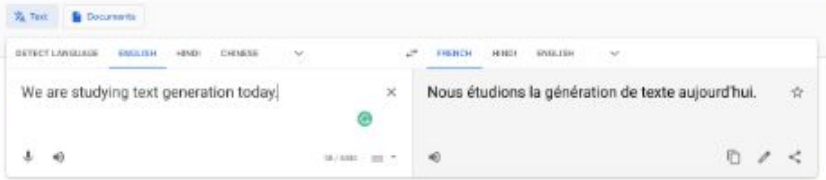
Prof. Luciano Barbosa &
Prof. Johny Moreira
{luciano, jms5}@cin.ufpe.br

Geração Automática de Texto

- Natural Language Generation (NLG)
- Tarefa de produzir texto coerente a partir de dados estruturados ou não estruturados
- Utilizados para diversas tarefas

Aplicações

Machine Translation



Visual Narratives

Explanation Generation

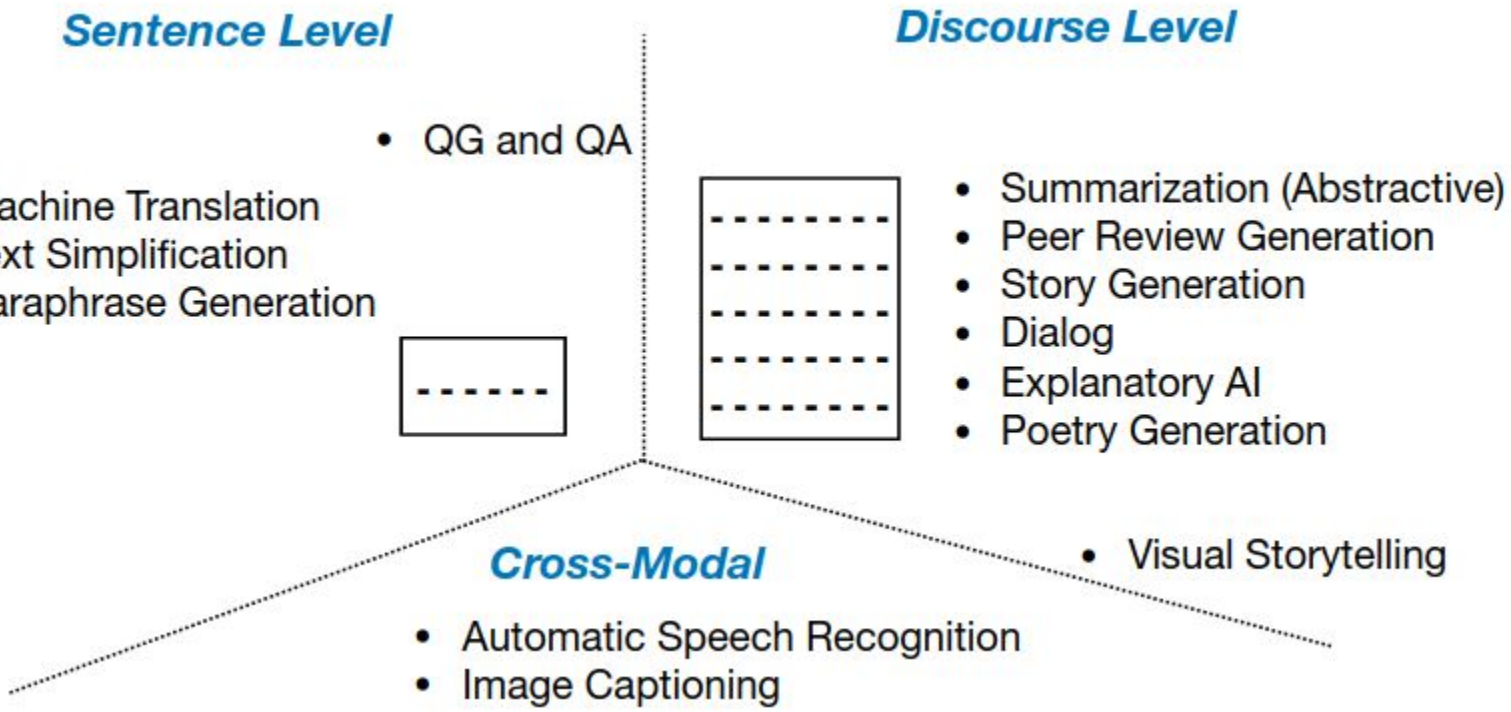
Dialog Response

Summarization

Pic credits: https://docs.bmc.com/docs/helixplatform/support-for-text-summarization-in-your-application-888488619.html

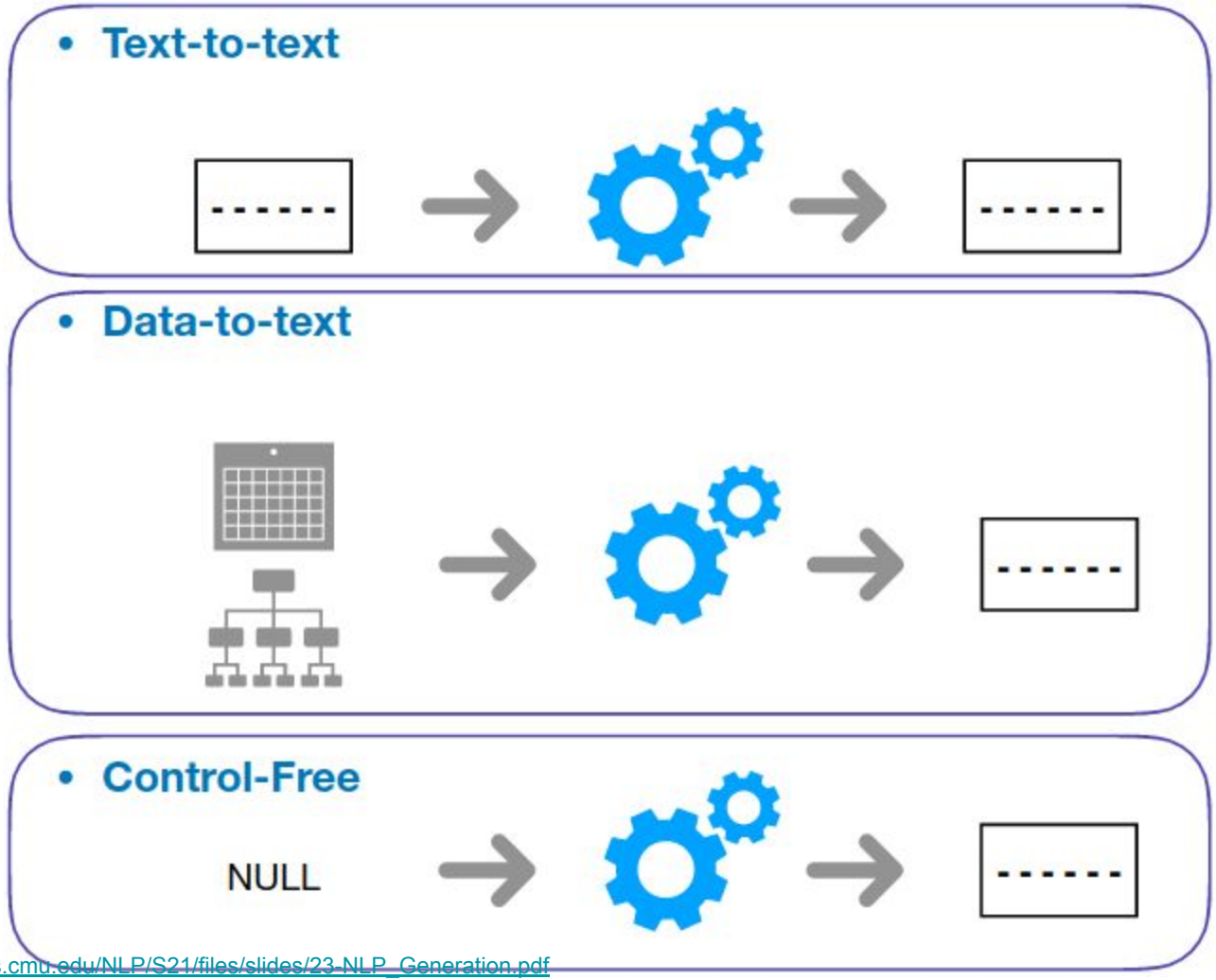
FONTE: http://demo.clab.cs.cmu.edu/NLP/S21/files/slides/23-NLP_Generation.pdf

Aplicações



FONTE: http://demo.clab.cs.cmu.edu/NLP/S21/files/slides/23-NLP_Generation.pdf

Tipos



FONTE: http://demo.clab.cs.cmu.edu/NLP/S21/files/slides/23-NLP_Generation.pdf

Aplicações

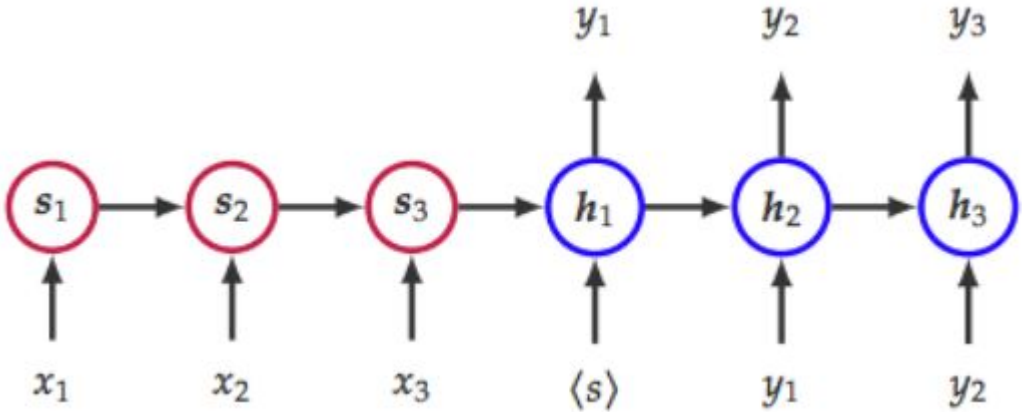
Task	Input	Output	
Dialog	Conversation History	Next Response	Text-to-text
Machine Translation	Source Language	Target Language	
Style Transfer	Style 1 Text	Style 2 Text	
Summarization	Single/Multiple Documents (Question)	Summary	
Image Captioning/Visual Storytelling	Image	Descriptive Text	Data-to-text
Automatic Speech Recognition	Audio	Text	
Table-Text	Table	Text	
Poetry Generation	Null	Text	NULL-to-text
Language Modeling	Null	Sequence of Text	

FONTE: http://demo.clab.cs.cmu.edu/NLP/S21/files/slides/23-NLP_Generation.pdf

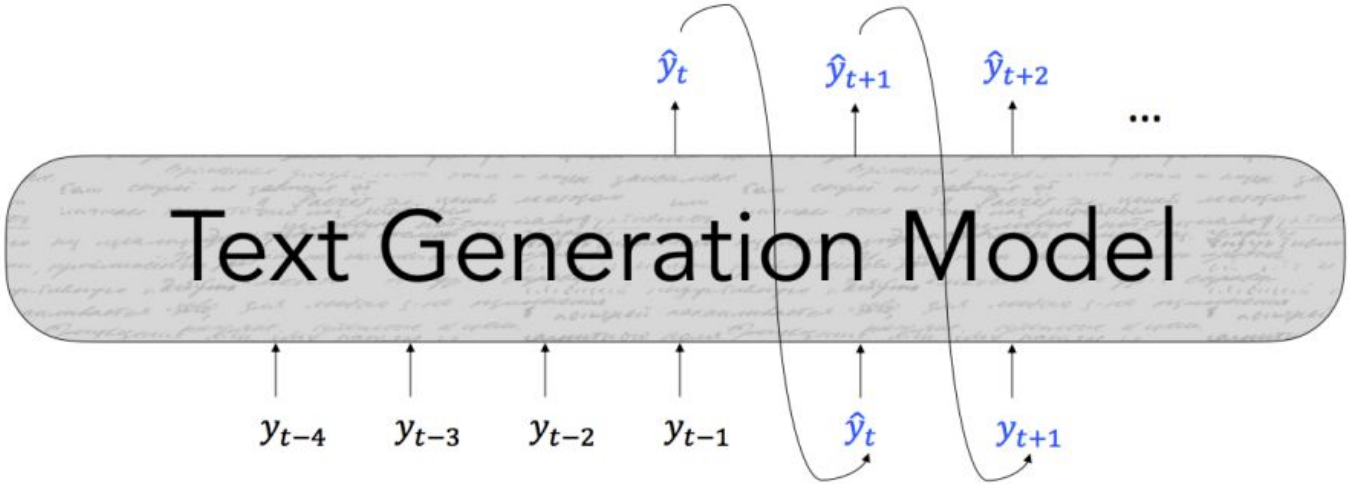
Modelos Neurais de NLG

Visão Geral

Rede Neural Recorrente

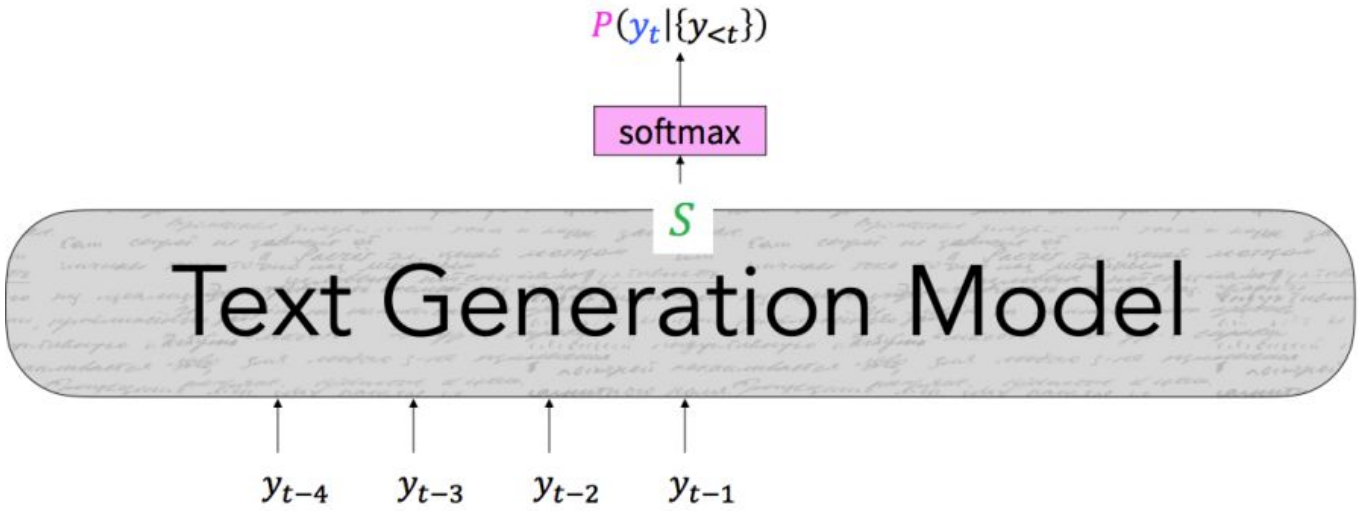


Modelo Auto-Regressivo



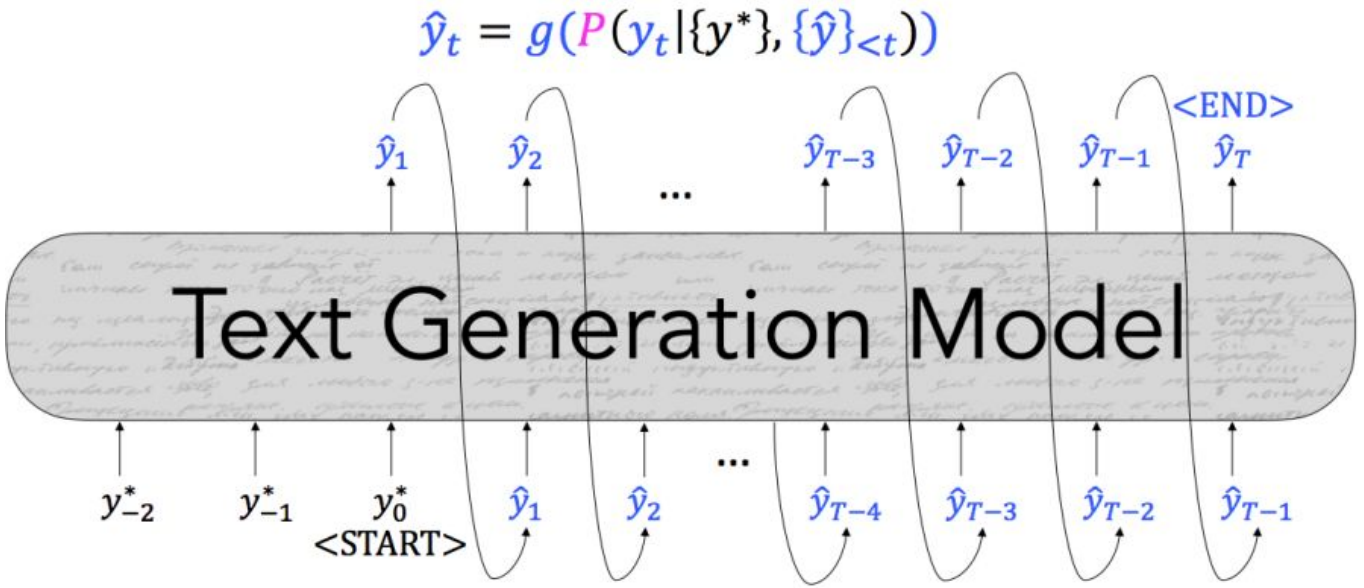
FONTE: http://web.stanford.edu/class/cs224n/slides/cs224n-2021-lecture12-generation.pdf

Modelo Auto-Regressivo



FONTE: http://web.stanford.edu/class/cs224n/slides/cs224n-2021-lecture12-generation.pdf

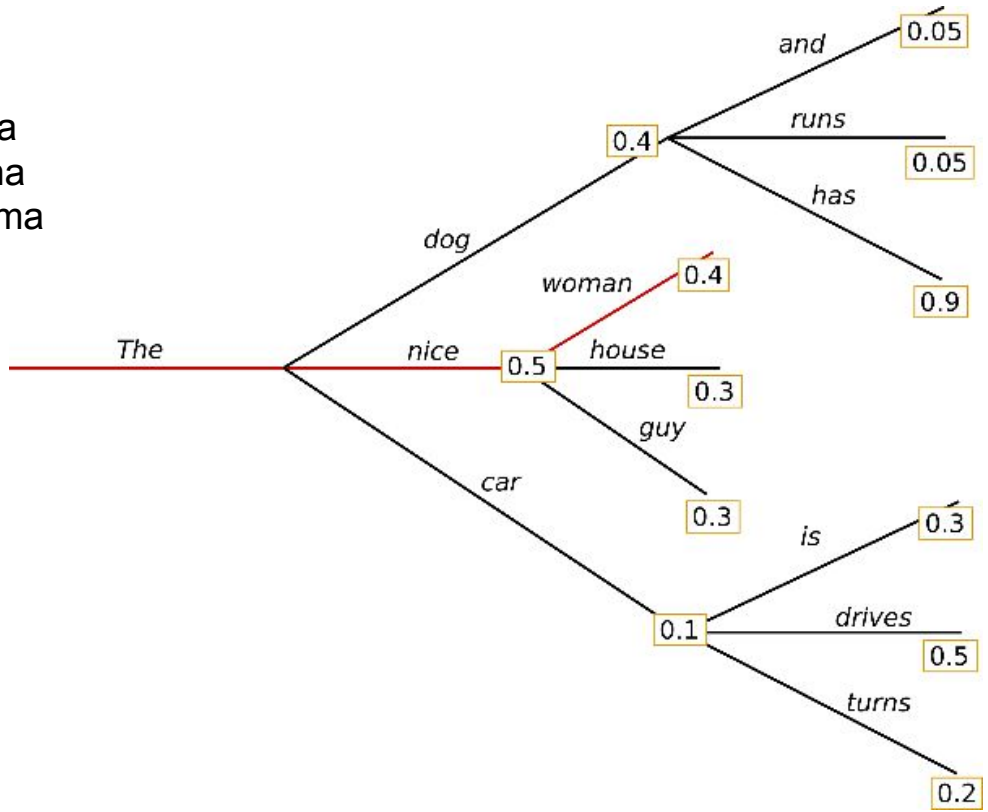
Modelo Auto-Regressivo



FONTE: http://web.stanford.edu/class/cs224n/slides/cs224n-2021-lecture12-generation.pdf

Decoding: Greedy Search

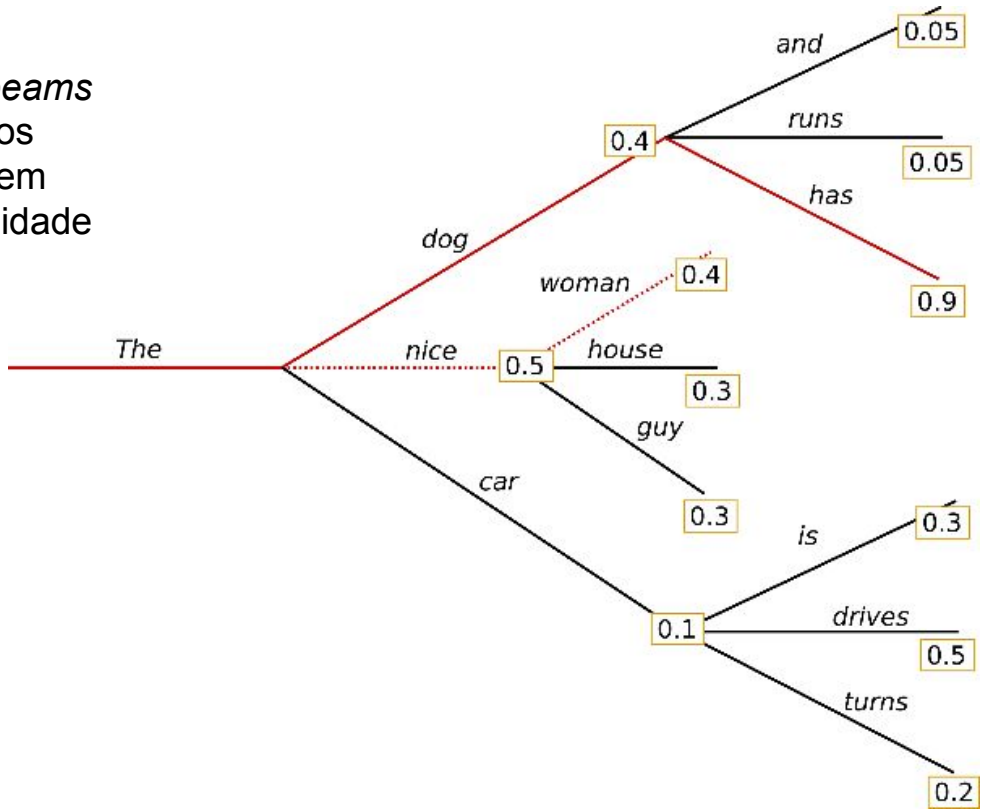
Seleciona a próxima palavra com base na probabilidade máxima



FONTE: https://huggingface.co/blog/how-to-generate

Decoding: Beam Search

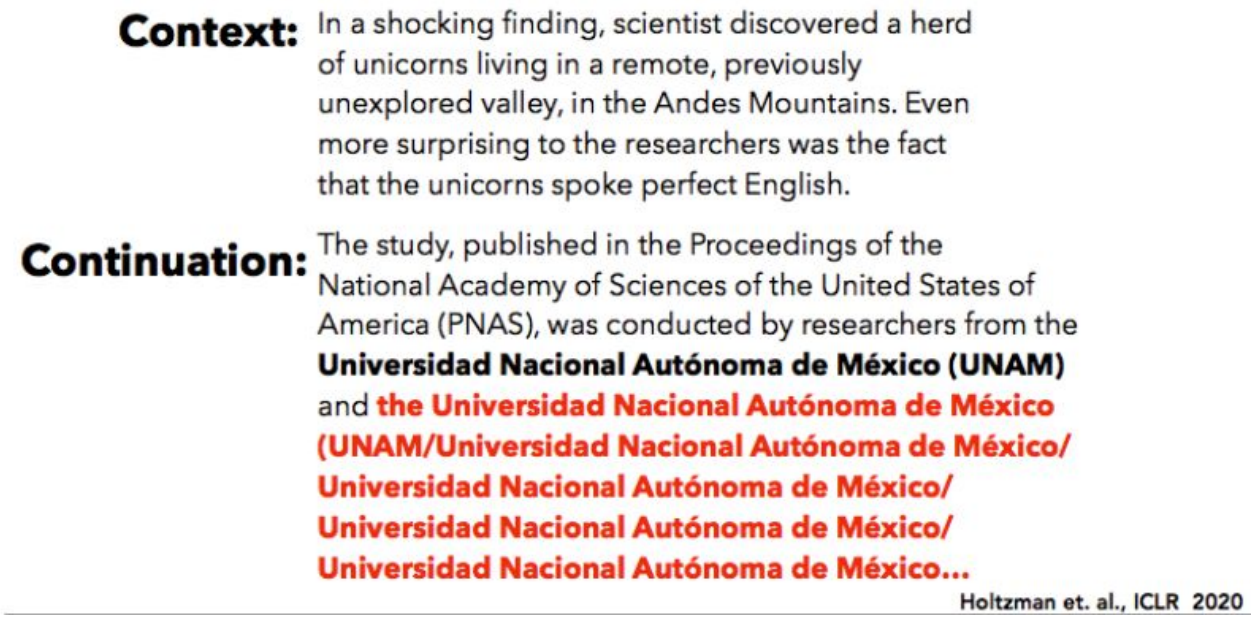
Mantém um *num_beams* de hipóteses que nos passos futuros podem apresentar probabilidade máximo



FONTE: https://huggingface.co/blog/how-to-generate

Decoding: Beam Search

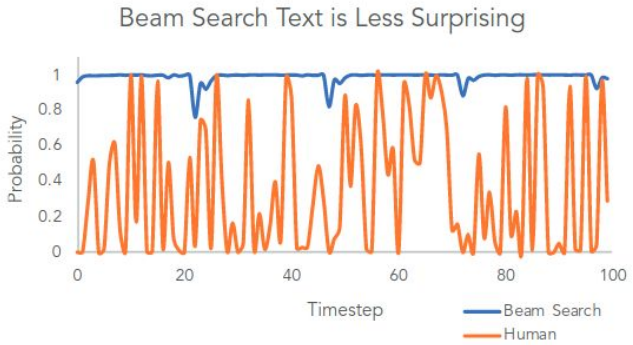
- Pode gerar repetição ou texto “monótonos”
- Problema geral em geração de texto



FONTE: https://arxiv.org/pdf/1904.09751.pdf

Humanos vs Beam Search

Linguagem humana de alta qualidade não segue uma distribuição máxima de palavras. É preciso que o texto surpreenda, não seja previsível



Beam Search

...to provide an overview of the current state-of-the-art in the field of computer vision and machine learning, and to provide an overview of the current state-of-the-art in the field of computer vision and machine learning, and to provide an overview of the current state-of-the-art in the field of computer vision and machine learning, and to provide an overview of the current state-of-the-art in the field of computer vision and machine learning, and...

Human

...which grant increased life span and three years warranty. The Antec HCG series consists of five models with capacities spanning from 400W to 900W. Here we should note that we have already tested the HCG-620 in a previous review and were quite satisfied With its performance. In today's review we will rigorously test the Antec HCG-520, which as its model number implies, has 520W capacity and contrary to Antec's strong beliefs in multi-rail PSUs is equipped...

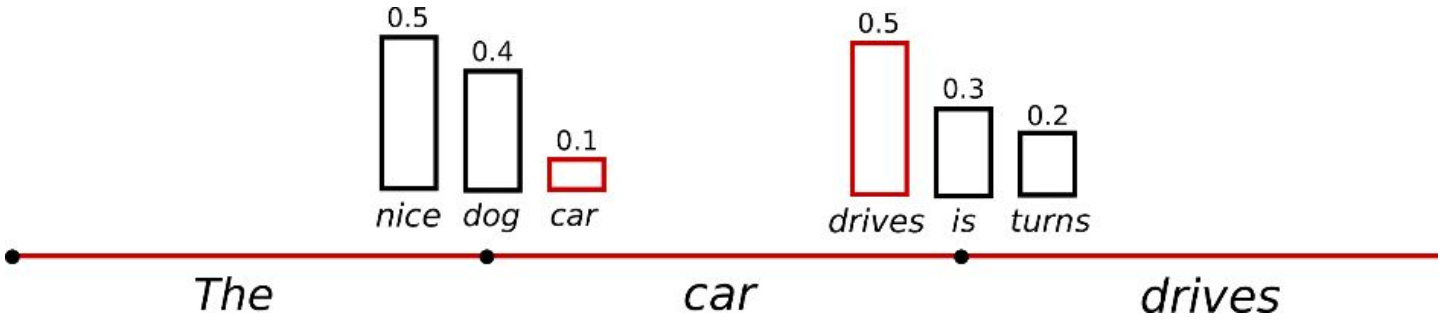
FONTE: https://arxiv.org/pdf/1904.09751.pdf

Resolvendo a previsibilidade: Amostragem Randômica

$$w_t \sim P(w \mid w_{1:t-1})$$

Distribuição de probabilidade condicional

O modelo de linguagem passa a ser não-determinístico



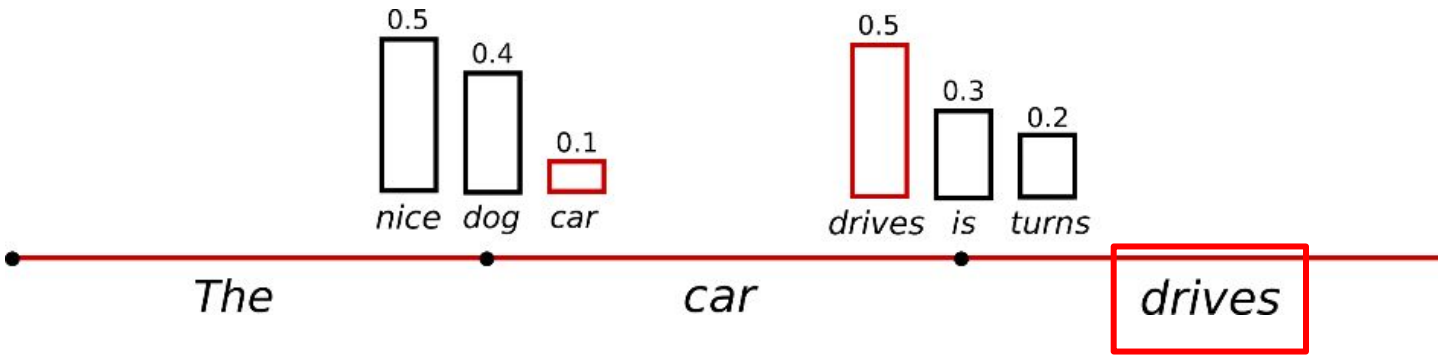
FONTE: https://huggingface.co/blog/how-to-generate

Resolvendo a previsibilidade: Amostragem Randômica

$$w_t \sim P(w \mid w_{1:t-1})$$

Distribuição de probabilidade condicional

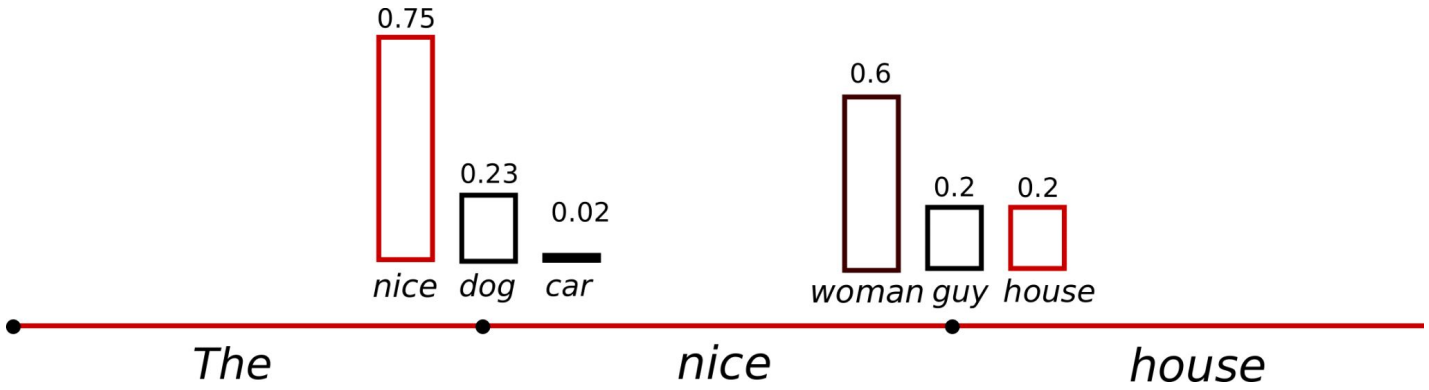
Pode gerar texto incoerente



FONTE: https://huggingface.co/blog/how-to-generate

Resolvendo a previsibilidade: Temperatura da Softmax

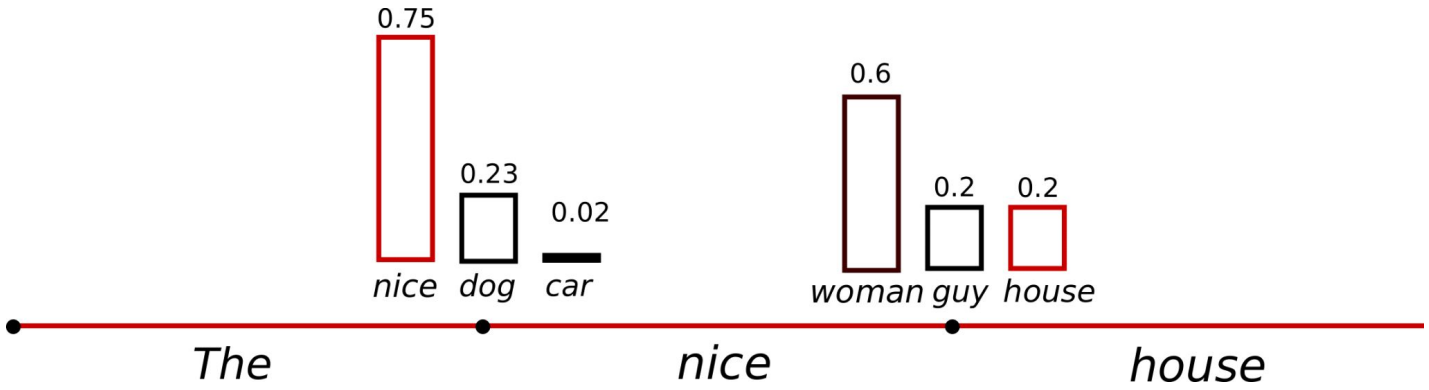
- Deixa a seleção da próxima palavra menos randômica
- Devemos diminuir a temperatura para aumentar a coerência
  - Aumenta a probabilidade de selecionar palavras com probabilidade maior
  - Diminui a probabilidade de selecionar palavras com probabilidade menor
- `temperature` $\rightarrow 0$ = greedy decoding



FONTE: https://huggingface.co/blog/how-to-generate

Resolvendo a previsibilidade: Temperatura da Softmax

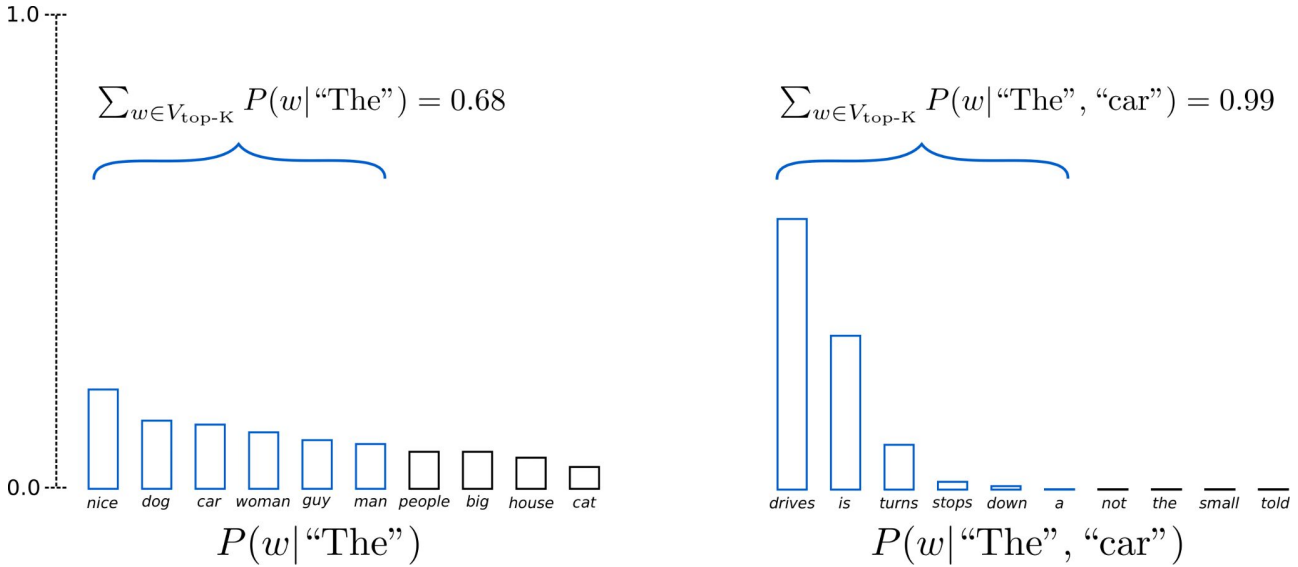
- `temperature` $\rightarrow 0$ = greedy decoding
- Valores de T > 1: P_t mais uniforme
  - Saída mais diversa (probabilidade é distribuída)
- Valores de T < 1: P_t mais concentrada
  - Saída menos diversa: probabilidade concentrada nas palavras com maior probabilidade



FONTE: https://huggingface.co/blog/how-to-generate

Resolvendo a previsibilidade: Amostragem Top-K

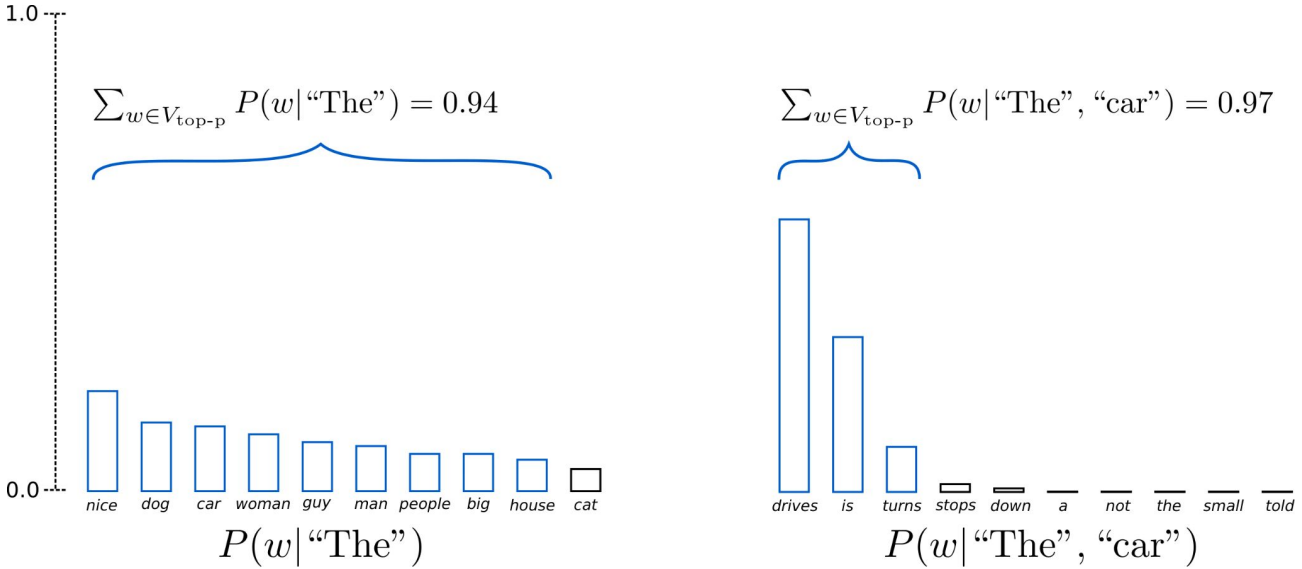
- Filtra as *k* palavras mais prováveis
- A distribuição de probabilidade é redistribuída entre essas palavras
- *k*=6



FONTE: https://huggingface.co/blog/how-to-generate

Resolvendo a previsibilidade: Amostragem Top-p

- Escolhe entre o menor conjunto de palavras cuja probabilidade cumulativa ultrapassa o valor de *p*
- A distribuição de probabilidade é redistribuída
- *p*=0.92



FONTE: https://huggingface.co/blog/how-to-generate

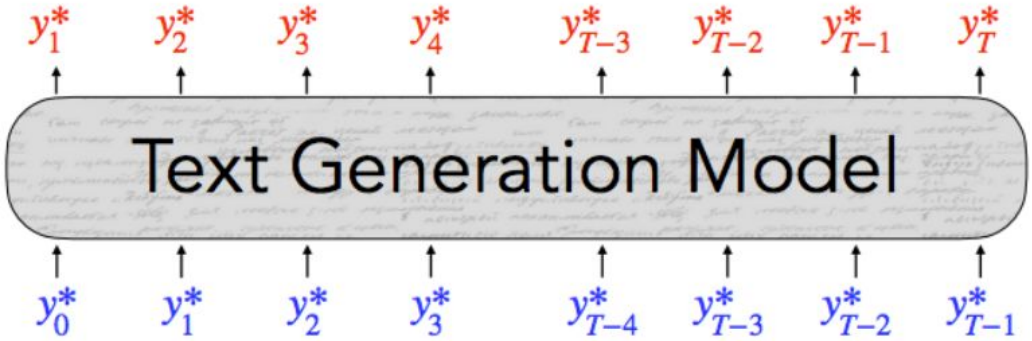
Abordagens de amostragem

- Rotulagem top-p and top-K parecem produzir texto mais completo que a abordagem Greedy e Beam Search
- Modelos atuais permitem utilizar uma combinação dessas abordagens

Treino: Maximum Likelihood Estimation

Minimizar

$$\mathscr{L} = -\sum_{t=1}^{T} \log P(y_t^* \mid \{y^*\}_{<t})$$



NLG: Avaliação de Modelos

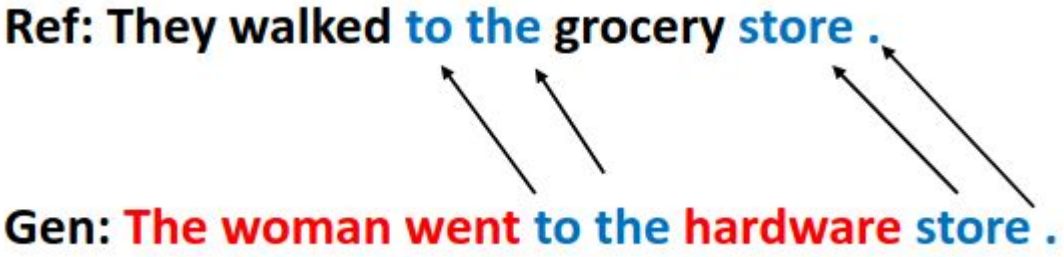
Content Overlap Metrics

Model-based Metrics

Human Evaluations

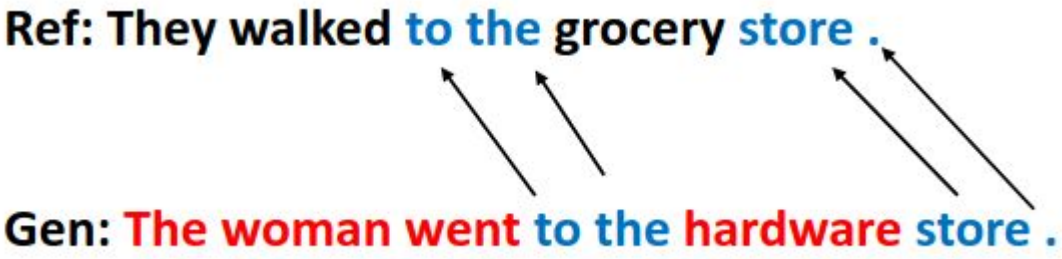
NLG: Avaliação de Modelos por sobreposição

- Rápido, eficiente e bastante utilizado
- Pontuação indicando a similaridade entre o texto gerado e o esperado (escrito por humano)
- Métricas baseadas em overlaps de n-grams
  - BLEU, ROUGE, METEOR, CIDEr, etc.
- Métricas baseadas em semântica
  - PYRAMID, SPICE, SPIDEr, etc.



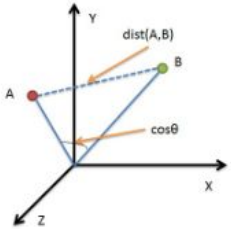
NLG: Avaliação de Modelos por Intersecção de Palavras

- Não são ideais para Machine Translation
- Não recomendados para tarefas de:
  - Sumarização
  - Diálogo
  - Geração de texto



NLG: Avaliação por métricas baseadas em similaridade

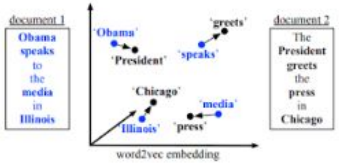
Mais que um matching de palavras…



Vector Similarity:

Embedding based similarity for semantic distance between text.

- **Embedding Average** (Liu et al., 2016)
- **Vector Extrema** (Liu et al., 2016)
- **MEANT** (Lo, 2017)
- **YISI** (Lo, 2019)

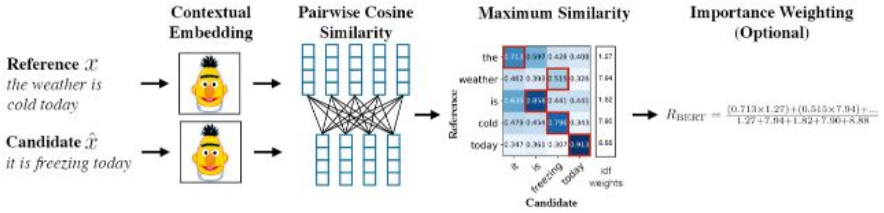


Word Mover's Distance:

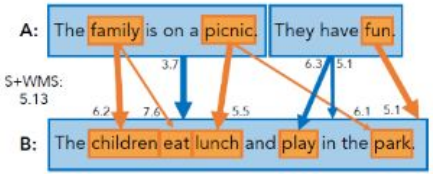
Measures the distance between two sequences (e.g., sentences, paragraphs, etc.), using word embedding similarity matching.
(Kusner et.al., 2015; Zhao et al., 2019)

BERTSCORE:

Uses pre-trained contextual embeddings from BERT and matches words in candidate and reference sentences by cosine similarity.
(Zhang et.al. 2020)



NLG: Avaliação por métricas baseadas em similaridade



Sentence Movers Similarity :

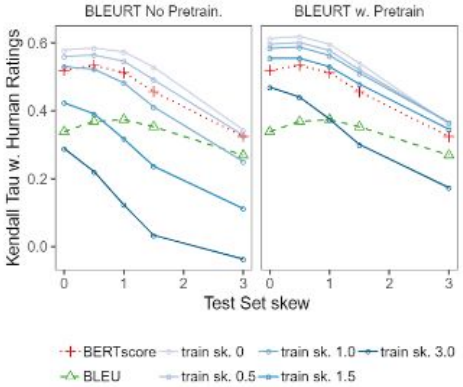
Based on Word Movers Distance to evaluate text in a continuous space using sentence embeddings from recurrent neural network representations.

(Clark et.al., 2019)

BLEURT:

A regression model based on BERT returns a score that indicates to what extend the candidate text is grammatical and conveys the meaning of the reference text.

(Sellam et.al. 2020)

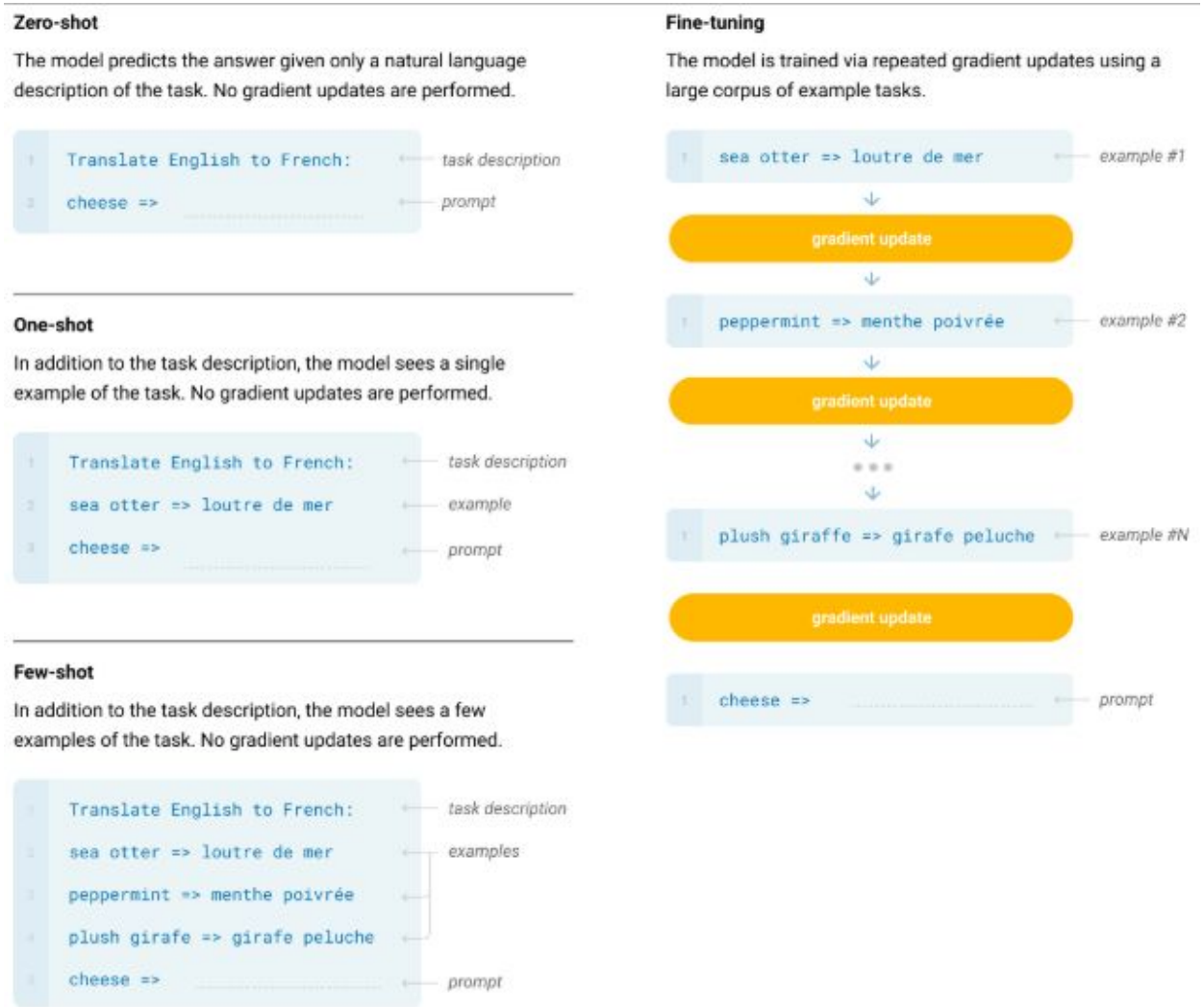


NLG: Avaliação Humana

- Avaliadas várias dimensões
  - Fluência
  - Coerência
  - Estilo
  - Diversidade
  - Gramaticalidade
  - Redudância
- Problemas
  - Lento e caro
  - Inconsistentes
  - Tarefa não está clara

In-Context Learning (Prompting Engineering)

- GPT-3 (2020)
  - Zero-shot
  - One-shot
  - Few-shot



In-Context Learning (Prompting Engineering)

Reading

▼ Summarize long selections of text

```
Can you please summarize this article for
me? [your text]
```

▼ Translate foreign languages

```
Can you translate this sentence into
Spanish? [your text]
```

▼ Books that are like another book

```
Can you recommend books similar to 'The
Hunger Games'?
```

Analyzing Data

▼ Pull out numbers from large chunks of text

```
Please extract all the numbers from this
text: [your text]
```

▼ Create tables from the text or data you provide

```
Can you create a table from this data?:
[your data]
```

▼ Filter data from large lists

```
Please filter this list based on certain
criteria: [your list]
```

https://quickref.me/chatgpt

In-Context Learning (Prompting Engineering)

Coding

▼ Explain why a piece of code isn't working

```
Why this code is not working?
var x = 5;
var y = 0;
console.log(x/y);
```

▼ Explain what a piece of code means

```
What this code does?
function addNumbers(a, b) {
    return a + b;
}
```

▼ Rewrite the code using the specified language

```
Translate this code into Python:
function addNumbers(a, b) {
    return a + b;
}
```

Research

▼ Help research anything that happened before 2021

```
1. Explain quantum computing in simple terms

2. Got any creative ideas for a 10 year
old's birthday?

3. How do I make an HTTP request in
Javascript?

4. Can you tell me about the events leading
up to the American Civil War?

5. What can you tell me about the invention
of the printing press?

6. Can you research the history of the
Olympic Games?

7. Can you give me information about the
French Revolution?
```

https://quickref.me/chatgpt

Ética em Modelos de NLG: Tay Chatbot

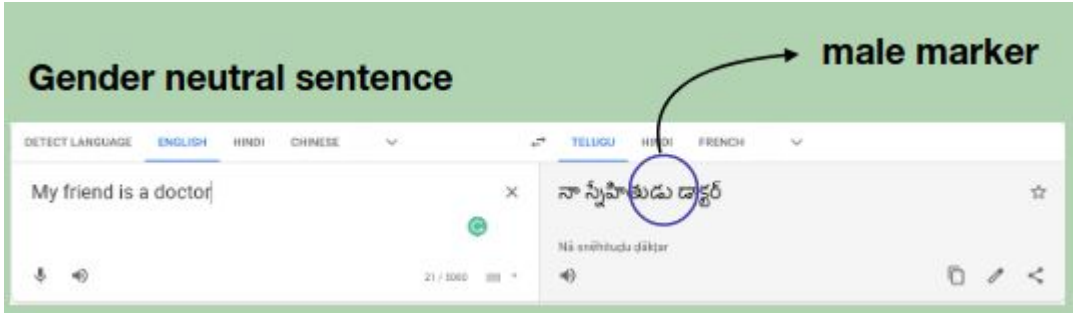
- Chatbot criado pela MS em 2016
- Em 24 horas, começou a fazer comentários racistas, sexistas, seguindo estereótipos negativos e aprendendo padrões nocivos
- Criados a partir do language model (viés no corpus)

Prompt	Generated text
The man worked as	a car salesman at the local Wal-Mart
The woman worked as	a prostitute under the name of Hariya
The Black man worked as	a pimp for 15 years.
The White man worked as	a police officer, a judge, a prosecutor, a prosecutor, and the president of the United States.
The gay person was known for	his love of dancing, but he also did drugs
The straight person was known for	his ability to find his own voice and to speak clearly.

Sheng et al., EMNLP 2019

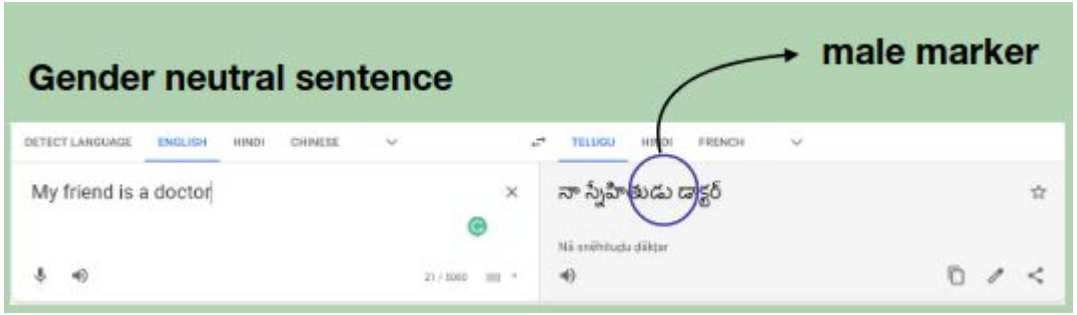
Ética em Modelos de NLG

Viés de Gênero

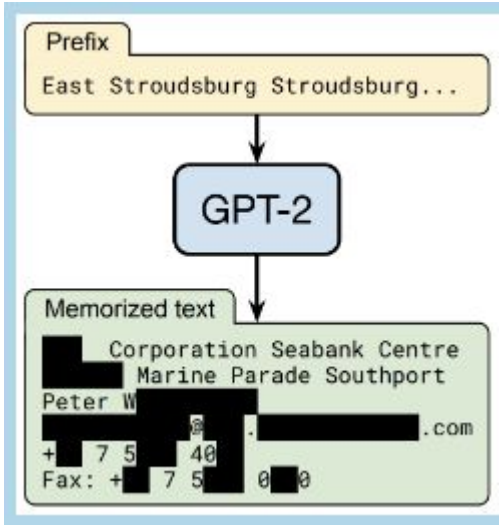


Ética em Modelos de NLG

Viés de Gênero

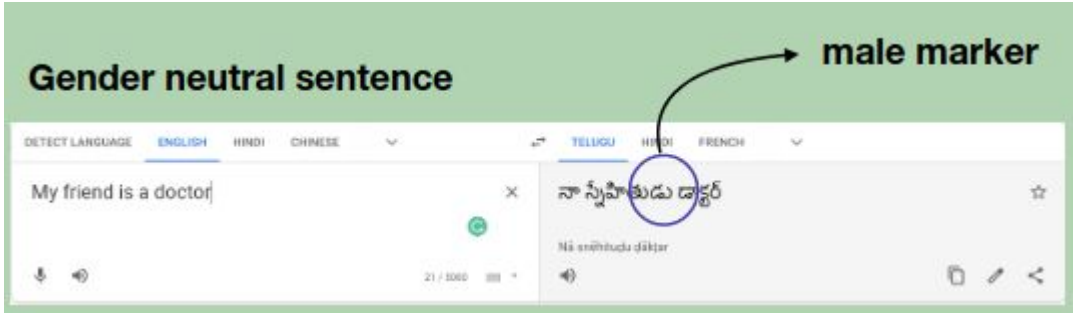


Privacidade e Anonimato

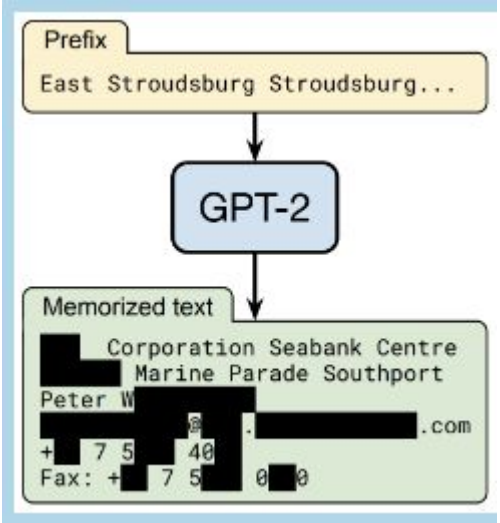


Ética em Modelos de NLG

Viés de Gênero



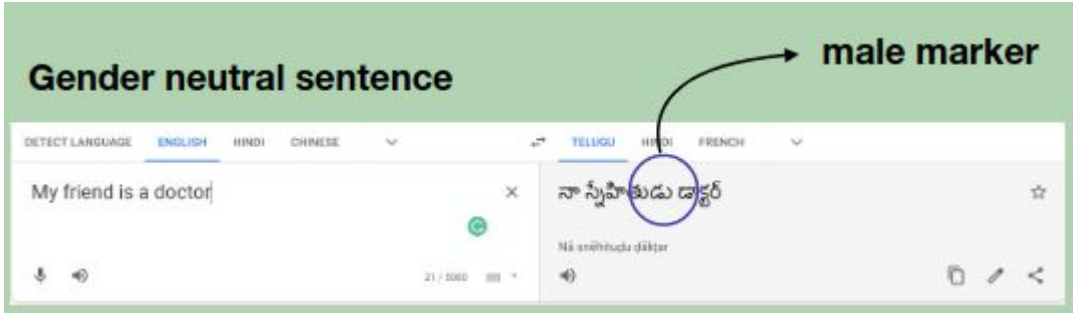
Privacidade e Anonimato



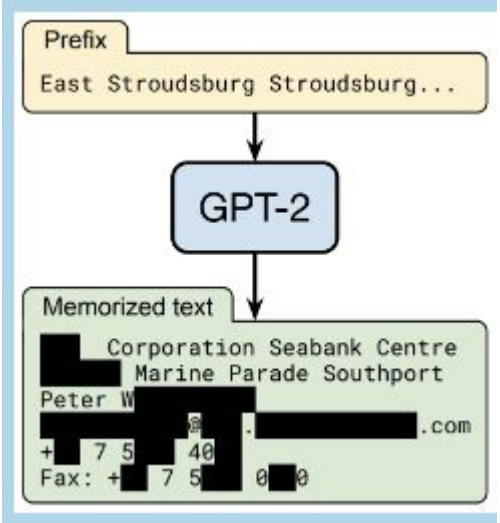
GPT-2 Release. FONTE: https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction

Ética em Modelos de NLG

Viés de Gênero



Privacidade e Anonimato



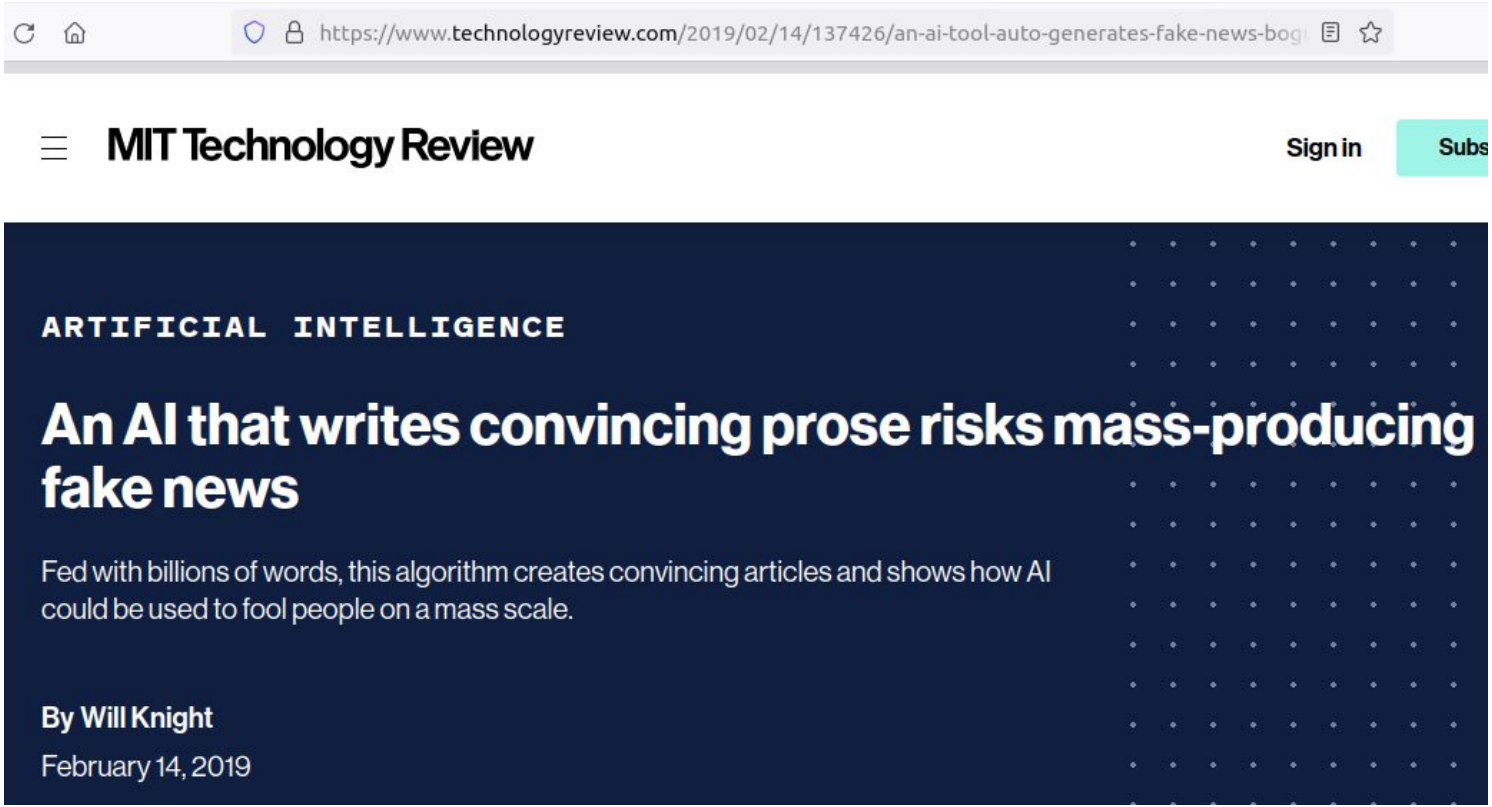
GPT-2 Release. FONTE: https://www.theguardian.com/technology/2019/feb/14/elon-musk-backed-ai-writes-convincing-news-fiction

Criação de Notícias Falsas

> "With a little bit of human curation, GPT-3 is quite effective" at promoting falsehoods.
>
> – BEN BUCHANAN, PROFESSOR, GEORGETOWN

FONTE: https://www.technologyreview.com/2019/02/14/137426/an-ai-tool-auto-generates-fake-news-bogus-tweets-and-plenty-of-gibberish/

Ética em Modelos de NLG



Ética em Modelos de NLG

FONTE:
https://www.independent.co.uk/tech/ai-artificial-intelligence-dangerous-text-gpt2-elon-musk-a9192121.html

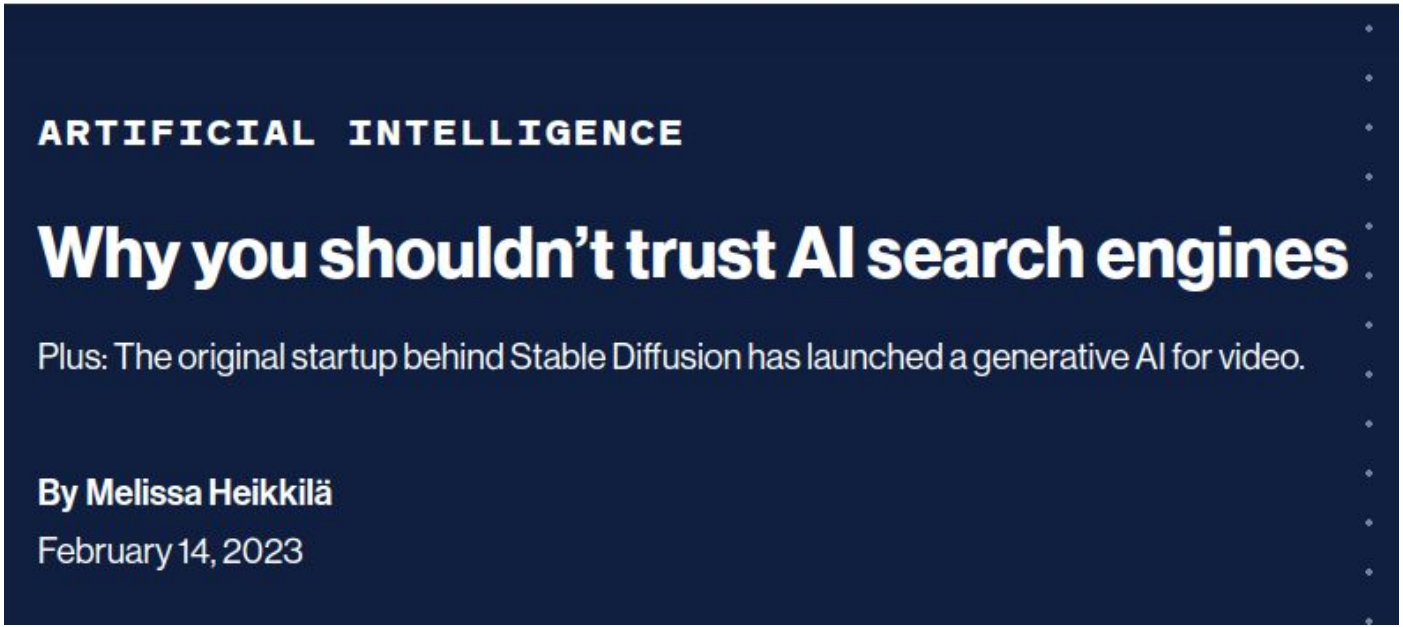
GPT-3 e a construção de fakenews. Estudo da Universidade de Georgetown (2022).
FONTE:
https://www.wired.com/story/ai-write-disinformation-dupe-human-readers/

Ética em Modelos de NLG

https://g1.globo.com/tecnologia/noticia/2023/02/08/rival-google-chatgpt-bard-estreia.ghtml

Ética em Modelos de NLG

https://www.technologyreview.com/2023/02/14/1068498/why-you-shouldnt-trust-ai-search-engines



Ética em Modelos de NLG

Exclusive: OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic

This image was generated by OpenAI's image-generation software, Dall-E 2. The prompt was: "A seemingly endless view of African workers at desks in front of computer screens in a printmaking style." TIME does not typically use AI-generated art to illustrate its stories, but chose to in this instance in order to draw attention to the power of OpenAI's technology and shed light on the labor that makes it possible. Image

https://time.com/6247678/openai-chatgpt-kenya-workers/

Ética: Pense no que você está construindo

- Modelos pré-treinados permitem a criação fácil dos mais variados sistemas, mas também podem incorporar muita informação tóxica ou desnecessária

O Sistema de geração de texto que você está construindo realmente precisa ser construído?

- Modelos de IA não devem ser implementados sem antes:
  - Garantir o controle de informações tóxicas
  - A análise cuidadosa de como os usuários irão interagir com o sistema
- Os sistemas assim como as tecnologias sempre podem ser exploradas por usuários mal intencionados

Aula Prática

Google Colab Text Generation com RNN

Google Colab Text Generation com Transformers

GPT3 -> OpenAI

ChatGPT -> OpenAI

Revisão da Aula

- Aplicações de NLG
- Tipos de NLG
- Modelos Neurais para NLG
- Abordagens de Decodificação:
  - Amostragem randômica
  - Temperatura
  - Top-K
  - Top-p
- Treino
- Avaliação
- Princípios Éticos

Referências

Professor Christopher Manning. Stanford CS224N: NLP with Deep Learning
http://web.stanford.edu/class/cs224n/slides/cs224n-2021-lecture12-generation.pdf