

Processamento de Linguagem Natural

Introdução ao Curso

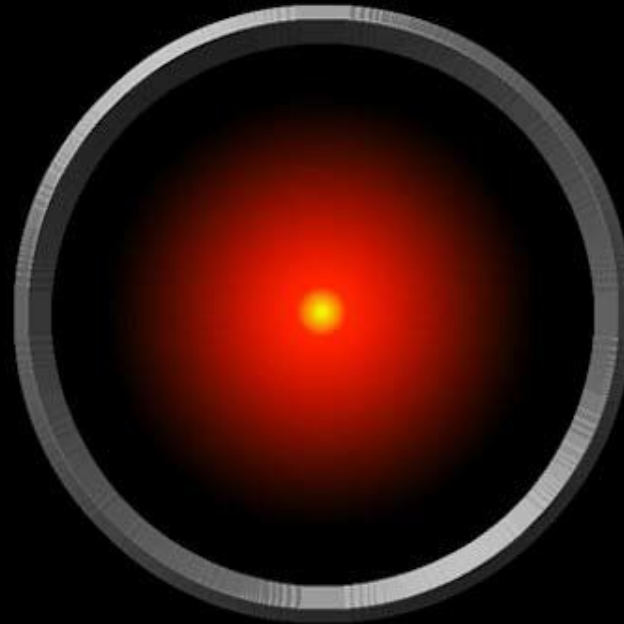
Prof. Luciano Barbosa &
Prof. Johny Moreira
{luciano, jms5}@cin.ufpe.br

O que é PLN

"Natural language processing (NLP) is a field of computer science, artificial intelligence (also called machine learning), and linguistics concerned with the interactions between computers and human (natural) languages. Specifically, the process of a computer extracting meaningful information from natural language input and/or producing natural language output "

Wikipedia

I'm sorry Dave,
I'm afraid I can't do that.



ChatGPT



Examples

"Explain quantum computing in simple terms"

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

"How do I make an HTTP request in Javascript?"



Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021



Free Research Preview: ChatGPT is optimized for dialogue. Our goal is to make AI systems more natural to interact with, and your feedback will help us improve our systems and make them

Motivação para o Uso de PLN

- Linguagem é complexa e envolve várias atividades humanas
 - Leitura, escrita, fala, audição
- Desejo de extrair conhecimento de textos
 - Artigos científicos, notícias etc

Aplicações




Aplicações: Tradução

Portuguese ▼

↔

English ▼

Enter text



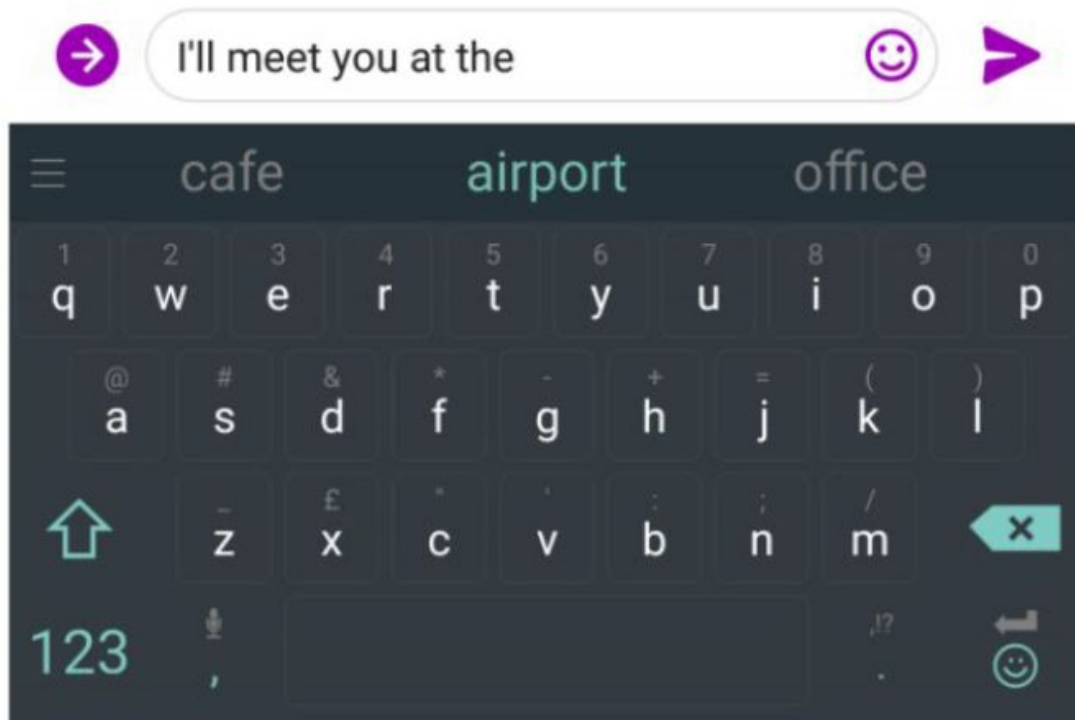
Translation

[Open in Google Translate](#) • [Feedback](#)

Aplicações: Reconhecimento de Fala



Aplicações: Predição de Próxima Palavra



Desafio: Ambiguidade

- Ex: I made her duck
 - Significados:

I cooked waterfowl for her.

I cooked waterfowl belong to her.

I created the (plaster?) duck she owns.

I caused her to quickly lower her head or body.

I waved my magic wand and turned her into undifferentiated waterfowl.

Desafio: Ambiguidade

I saw a girl with a telescope




A Disciplina

- Objetivo:
 - Aprender principais técnicas, tarefas e aplicações de Processamento de Linguagem Natural
- Avaliação: projeto
- Página: https://profluciano.github.io/pln_dl/

Projeto: Implementar Análise de Sentimentos em Avaliações de Produtos

- Coletar avaliações de um produto selecionado: texto e nota
- Treinar classificador: SVM + bow, SVM + embeddings, BERT
- Apresentação reportando resultados (F1 e acurácia) e análises (slides + vídeo de máximo 15 minutos)

 Luana Gimenes
★★★★★ **Atende as expectativas**
Avaliado no Brasil 🇧🇷 em 16 de fevereiro de 2023
Compra verificada

Velo tudo certinho, conforme anúncio, aqui em casa compramos a linha Redmi Note há anos, então sabemos bem o que esperar, estou muito satisfeita, pra mim não tem celular intermediário melhor nessa faixa de preço. Velei antes do prazo, não veio nota fiscal mas eu pedi e o vendedor enviou no dia seguinte. Uma semana de uso e nada a reclamar.

 Emily Azevedo
★★★★★ **Dispensa comentários!!**
Avaliado no Brasil 🇧🇷 em 13 de fevereiro de 2023
Compra verificada

Ele é muito bom, a câmera dispensa comentários. Meu marido está amando, a tela é bem grande, a bateria dura bastante tbm e veio tudo certinho, com carregador e bem embalado, tudo lacrado. Só não gostei pq o envio foi feito pelos correios que é uma bosta, eu tive que ir fazer a retirada.



Avaliações

Bibliografia

- Speech and Language Processing (3rd ed. draft). Dan Jurafsky and James H. Martin
- Foundations of Statistical Natural Language Processing. Christopher D. Manning and Hinrich Schütze
- Natural Language Processing. Jacob Eisenstein
- Lin, Tianyang, Yuxin Wang, Xiangyang Liu, and Xipeng Qiu. "A survey of transformers." AI Open (2022).

Tópicos

- Processamento de texto
- Representação de palavras
- Extração de informação
- Chatbots
- Modelos de linguagem
- Geração automática de texto

Processamento de Texto: Segmentação de Palavras

- Primeira tarefa em um sistema de PLN

毎年うちの研究室の学生が1-2名国語研でアルバイトさせてもらっているので、今日は新しくアルバイトする B4 学生の紹介である。

- Estado da arte utiliza machine learning
- Obtêm alta acurácia

Classificação de Texto: Análise de Sentimentos

"I am happy with this water bottle."



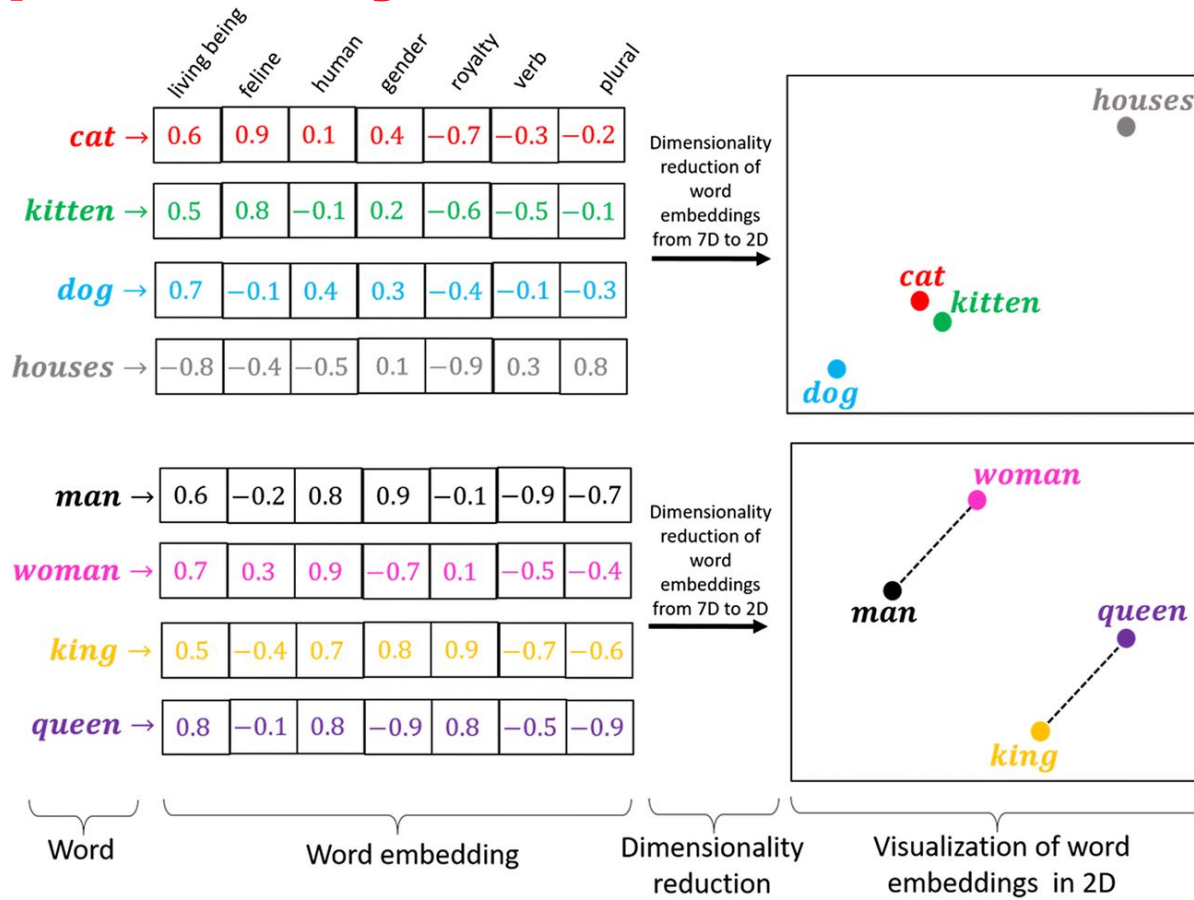
"This is a bad investment."



"I am going to walk today."



Representação de Palavras: Embeddings



Information Extraction

We have decided to meet tomorrow at 10:00am in the lab.

To do : meeting
Time : 10:00 am, 22/3/2012
Venue : Lab

Chatbot: ELIZA

User1: You are like my father in some ways.

ELIZA1: WHAT RESEMBLANCE DO YOU SEE

User2: You are not very aggressive but I think you don't want me to notice that.

ELIZA2: WHAT MAKES YOU THINK I AM NOT AGGRESSIVE

User3: You don't argue with me.

ELIZA3: WHY DO YOU THINK I DON'T ARGUE WITH YOU

User4: You are afraid of me.

ELIZA4: DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU

Modelo de Linguagem



natural language processing

natural language processing **algorithms**

natural language processing **tutorial**

natural language processing **python**

natural language processing **machine learning**

natural language processing **course**

natural language processing **examples**

natural language processing **application**

natural language processing **pdf**

Machine Translation

मेरा नाम रजत है | => My name is Rajat.

Grass is greener on the other side. => दूर के ढोल सुहावने |

Google's Translation : घास दूसरी तरफ हरियाली है |

Q&A

Quora



who was the highest IQ

Who has the highest recorded IQ of all time?



Answer



Follow · 311



Request



3



Ad by Forge of Empires



Through the ages!

Develop a civilization from the stone age to the future in this award-winning game.



Learn More

81 Answers



Alan Muller, Self-taught

Answered December 16, 2016



William James Sidis is another individual whose exact IQ score is unknown. Experts have pegged his IQ somewhere in the range of 250 to 300, which would give him the highest IQ score in history.

William James Sidis was an american child prodigy with exceptional mathematical abilities and a claimed mastery of many languages. After his death, his sister made the unverifiable claim that his IQ was "the very highest that had ever been obtained", but any records of any IQ testing that Sidis actually took have been lost to history.

He entered Harvard at age 11 and, as an adult, was claimed to be conversant in over 40 languages and dialects. It was later acknowledged, however, that some of the claims made were exaggerations, with a researcher stating "I have been researching the veracity of primary sources of various subjects for about twenty-eight years, and never before have I found a topic so satiated with lies, myths, half-truths, exaggerations, and other forms of misinformation as is in the history behind William Sidis"

Geração de Texto

Instruction: Given an input question, respond with syntactically correct PostgreSQL. Be creative but the SQL must be correct.

Input: how many users signed up in the past month?

GPT-3 Response: SELECT COUNT(*) FROM users
WHERE signup_time > now() - interval '1 month'