Tópicos Especiais em Sistemas de Informação

Aula 02

Processamento de Linguagem Natural

Prof. Msc. Francisco de Assis de Souza Rodrigues

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Visão Geral

- Introdução
- Fundamentos do PLN
- Aplicações Práticas
- Desafio



Objetivo

Introduzir conceitos fundamentais do processamento de linguagem natural e aplicar técnicas básicas usando **Python** e a biblioteca **NLTK**, **,** entre outras **bibliotecas** e **frameworks.**

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Introdução - O que é PLN?

PLN (Processamento de Linguagem Natural) é uma área da inteligência artificial (IA) que busca permitir que máquinas compreendam, interpretem e gerem a linguagem humana de maneira significativa. Ele combina linguística computacional e aprendizado de máquina para processar textos e falas de forma automatizada.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Introdução - Importância do PLN

PLN é fundamental na interação entre humanos e máquinas, pois permite que computadores analisem grandes volumes de texto, extraiam informações relevantes e respondam de maneira coerente. Sua importância cresce com a necessidade de automação de tarefas que envolvem linguagem natural, como suporte ao cliente, análise de sentimentos e tradução de idiomas.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Introdução - Exemplos Práticos de PLN

Assistentes Virtuais (ex.: Siri, Alexa, Google Assistant)

- Usam PLN para interpretar comandos de voz e responder a perguntas.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Introdução - Exemplos Práticos de PLN

Tradução Automática (ex.: Google Tradutor, DeepL)

- Analisa textos em um idioma e os traduz para outro de forma automatizada.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Introdução - Exemplos Práticos de PLN

Chatbots (ex.: atendimento automatizado em sites)

- Responde a perguntas frequentes e simula conversas humanas para suporte ao cliente.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Introdução - Exemplos Práticos de PLN

Análise de Sentimentos (ex.: monitoramento de redes sociais)

- Identifica emoções e opiniões em textos, auxiliando empresas a entenderem o feedback dos clientes.

PLN está presente em diversas aplicações do dia a dia e continua evoluindo com avanços em IA e aprendizado de máquina.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Introdução
Desafios do Processamento de Linguagem Natural (PLN)

Apesar dos avanços no PLN, existem diversos desafios que dificultam a compreensão e geração da linguagem natural pelas máquinas. Entre os principais desafios estão:

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Ambiguidade Lexical

As palavras podem ter múltiplos significados dependendo do contexto. O PLN precisa determinar qual sentido é mais apropriado.

Exemplo:

- "Banco" pode significar uma instituição financeira ou um assento.
- "Tirar" pode significar "remover" (tirar o casaco) ou "fotografar" (tirar uma foto).

Polissemia e Sinonímia

- **Polissemia** ocorre quando uma mesma palavra tem vários significados relacionados.
  - Exemplo: "Cabeça" pode se referir à parte do corpo ou a um líder de equipe.
- **Sinonímia** refere-se a diferentes palavras que possuem significados semelhantes, dificultando a padronização do entendimento.
  - Exemplo: "Bonito" e "lindo" podem ser usados de forma intercambiável, mas com nuances diferentes dependendo do contexto.

Gírias, Ironia e Contexto

- **Gírias** variam de acordo com a região e o público, dificultando a interpretação correta.
  - Exemplo: "Essa roupa tá muito top!" – onde "top" significa algo positivo.
- **Ironia** é especialmente difícil para sistemas de PLN, pois muitas vezes o significado é o oposto do que está sendo dito.
  - Exemplo: "Nossa, que ótimo serviço!" pode ser um elogio ou uma crítica sarcástica.
- **Contexto** influencia muito a interpretação do significado. O mesmo texto pode ser compreendido de maneiras diferentes dependendo do assunto da conversa e do conhecimento prévio do usuário.
  - Exemplo: "O jogo foi incrível!" pode se referir a futebol, videogame ou um tabuleiro, dependendo do contexto.

Introdução
Desafios do Processamento de Linguagem Natural (PLN)

Os desafios do PLN mostram a complexidade da linguagem humana e a necessidade de algoritmos mais sofisticados para lidar com ambiguidades, nuances e variações no uso das palavras. O uso de aprendizado profundo (Deep Learning) e modelos de linguagem avançados, como o GPT, tem ajudado a melhorar a interpretação, mas ainda há muito a ser desenvolvido para tornar as máquinas capazes de compreender a linguagem de forma tão natural quanto os humanos.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Fundamentos do PLN

Principais **Tarefas** do Processamento de Linguagem Natural (PLN)

O PLN envolve diversas tarefas essenciais para permitir que máquinas compreendam e processem a linguagem humana.

Algumas das principais são:

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Tokenização – Dividir frases em palavras

A tokenização é o processo de dividir um texto em unidades menores chamadas *tokens*, que podem ser palavras, frases ou até caracteres, dependendo do nível de segmentação.

Exemplo:
Entrada: *"O cachorro late à noite."*
Saída: ["O", "cachorro", "late", "à", "noite", "."]

Lematização e Stemming – Reduzir palavras à sua raiz

- **Stemming:** Reduz palavras às suas raízes, sem levar em conta regras gramaticais.
- **Lematização:** Reduz palavras à sua forma base (lematizada), respeitando a gramática.

Exemplo:

- **Stemming**: "correr", "correndo" → "corr"
- **Lematização**: "correr", "correndo" → "correr"

A lematização é mais precisa, mas o stemming é mais rápido.

POS Tagging (Marcadores Gramaticais) – Identificar a função das palavras

O *Part-of-Speech Tagging* (POS Tagging) identifica a classe gramatical das palavras, como substantivos, verbos e adjetivos, ajudando na análise sintática do texto.

Exemplo:

Entrada: *"O gato dorme feliz."*

Saída: [("O", DET), ("gato", NOUN), ("dorme", VERB), ("feliz", ADJ)]

Reconhecimento de Entidades Nomeadas (NER) – Extrair nomes próprios, locais e organizações

O *Named Entity Recognition* (NER) identifica e classifica entidades como nomes de pessoas, locais, empresas, datas e valores monetários.

Exemplo:
Entrada: *"A Apple lançou um novo iPhone nos EUA em 2024."*
Saída: [("Apple", EMPRESA), ("iPhone", PRODUTO), ("EUA", LOCAL), ("2024", DATA)]

Fundamentos do PLN
Principais Tarefas do Processamento de Linguagem Natural (PLN)

Análise de Sentimentos – Determinar polaridade textual

A análise de sentimentos avalia a opinião expressa em um texto, classificando-o como positivo, negativo ou neutro.

Exemplo:
Entrada: *"O atendimento foi horrível, nunca mais volto!"*
Saída: Sentimento: Negativo

Essa técnica é amplamente usada em redes sociais, análise de feedback de clientes e marketing digital.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Fundamentos do PLN
Principais **Tarefas** do Processamento de Linguagem Natural (PLN)

As **tarefas** de **PLN** são fundamentais para diversas aplicações, como **chatbots**, **assistentes virtuais**, **busca inteligente** e **análise de dados textuais**. Modelos avançados, como Transformers e Redes Neurais Profundas, continuam aprimorando a precisão dessas tarefas, tornando a interação entre humanos e máquinas cada vez mais natural.

Experimento:
https://colab.research.google.com/drive/1vfVr8vuDSJrBwSpbsQNPfamLSIILra2a?usp=sharing

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Sugestão para Leitura

Artigo: https://aclanthology.org/2020.lrec-1.895.pdf

Dissertação:
https://repositorio.ufpe.br/bitstream/123456789/37648/1/DISSERTA%c3%87%c3%83O%20Francisco%20de%20Assis%20de%20Souza%20Rodrigues.pdf

Apresentão:
https://docs.google.com/presentation/d/1ZikI7zQzA91jUQTqKrywmjeUY48DkhMLzhbox4bgIDk/edit#slide=id.p1

Website: https://deepnlpf.github.io/site/

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Referências

OpenAI. ChatGPT (modelo GPT-4). Auxílio na elaboração de plano de aula sobre Processamento de Linguagem Natural. 2025.

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO

Perguntas?!



UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO



Obrigado!

UniFAP
CENTRO UNIVERSITÁRIO PARAÍSO