

Processamento de Linguagem Natural

Professor Vitor Kessler

Introdução

- Fazer computadores entenderem, interpretar e manipular a linguagem humana.
- Uso de *machine learning* para processar e interpretar texto e dados.
- Fazendo sucesso agora com a chegada de *big data*.



Introdução

- Um conjunto de teorias de técnicas computacionais para:
 - Analisar e representar naturalmente textos em um ou mais níveis da análise linguística.
 - Realizar o processamento da linguagem humana para um conjunto de tarefas e aplicações.

Introdução

- **Problemas:**

- **Ambiguidade (polissemia):**

- A manga ficou amarela.

- **Linguagens diferentes:**

- Inglês: The white cat.
 - Português: O gato branco.
 - Árabe القط البيض .

Introdução

- A pesquisa em PLN envolve o estudo de:
 - **Som:**
 - Fonologia (como os sons formam sílabas e palavras).
 - **Estrutura:**
 - Análises morfológica (formação das palavras) e sintática (formação das frases).
 - **Significado:**
 - Análises semântica (significado das palavras) e pragmática (significado das palavras dentro do contexto).

Introdução

- **Utilidade:**



Introdução

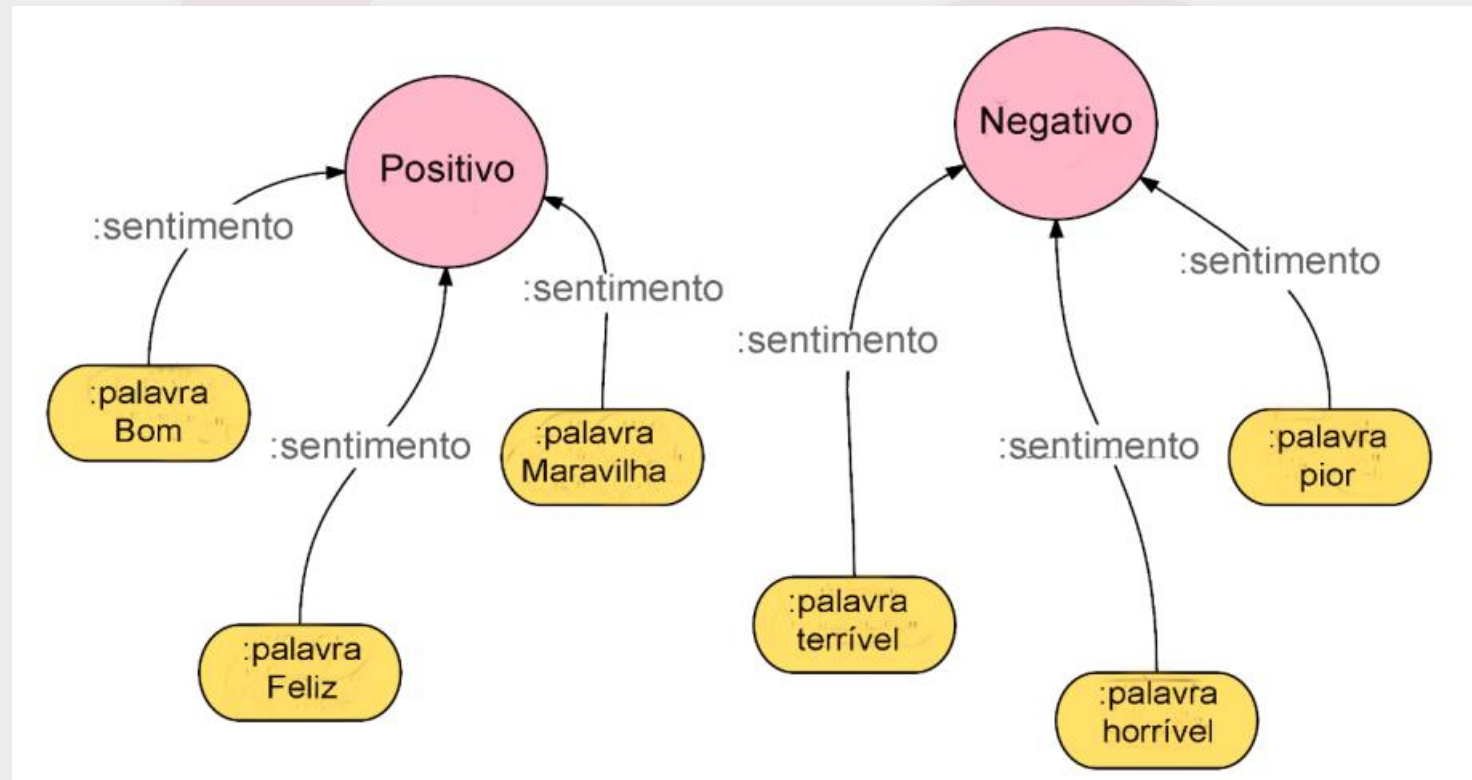
- **Utilidade:**
 - **Categorização de conteúdo:**
 - Um resumo do documento baseado em linguística, que inclui pesquisa e indexação, alertas de conteúdo e detecção de duplicações.
 - **Descoberta e modelagem de tópicos:**
 - Captura com precisão o significado e os temas em coleções de texto, e aplica *advanced analytics* como otimização e *forecasting*.
 - **Extração contextual:**
 - Extrai automaticamente informações estruturadas de fontes textuais.

Introdução

- Utilidade:
 - **Análise de sentimento:**
 - Identifica o estado de espírito ou opiniões subjetivas em grandes quantidades de texto, incluindo o sentimento médio e a mineração de opinião.
 - **Conversão fala-texto e texto-fala:**
 - Transforma comandos de voz em texto escrito e vice-versa.
 - **Sumarização:**
 - Gera sinopses de grandes corpos de texto automaticamente.
 - **Tradução de máquina:**
 - Traduz texto ou fala de um idioma para outro, automaticamente.

Introdução

- **Utilidade:**
 - **Análise de sentimento:**



Introdução

- **Classificação do PLN:**
 - **Interpretação de linguagem natural:**
 - Compreensão de frases e textos.
 - Exemplo: Conversão de linguagem natural para SQL.
 - **Geração de linguagem natural:**
 - Conversão de uma representação interna de um conteúdo semântico pré-definido para linguagem natural.
 - Exemplo: Geração automática de resumos textuais de tabelas de dados.

Introdução

Árvore sintática

