

# LINGÜÍSTICA COMPUTACIONAL

## **Parte 03: Processamento de textos usando Redes neurais**

Marcelo Finger and Felipe Salvatore

Departamento de Ciência da Computação  
Instituto de Matemática e Estatística  
Universidade de São Paulo

2º Semestre 2019

# DESAFIOS EM PROCESSAMENTO

# CRESCENTE COMPLEXIDADE: CONTEXTOS

FÁCIL Pouco ou nenhum contexto

DOMINANDO Sensibilidade média ao contexto

DIFÍCIL Dependência de contexto ilimitado

DIFICÍLIMO Parte do contexto fora do texto

# CRESCENTE COMPLEXIDADE: CONTEXTOS

**FÁCIL** Pouco ou nenhum contexto

Tokenização; Consultas em BDs

**DOMINANDO** Sensibilidade média ao contexto

**DIFÍCIL** Dependência de contexto ilimitado

**DIFICÍLIMO** Parte do contexto fora do texto

# CRESCENTE COMPLEXIDADE: CONTEXTOS

**FÁCIL** Pouco ou nenhum contexto

Tokenização; Consultas em BDs

**DOMINANDO** Sensibilidade média ao contexto

Identificação de ponto final; Etiquetagem morfofossintática

**DIFÍCIL** Dependência de contexto ilimitado

**DIFICÍLIMO** Parte do contexto fora do texto

# CRESCENTE COMPLEXIDADE: CONTEXTOS

**FÁCIL** Pouco ou nenhum contexto

Tokenização; Consultas em BDs

**DOMINANDO** Sensibilidade média ao contexto

Identificação de ponto final; Etiquetagem morfosintática

**Difícil** Dependência de contexto ilimitado

Tradução; Parseamento; Desambiguação

**DIFICÍLIMO** Parte do contexto fora do texto

# CRESCENTE COMPLEXIDADE: CONTEXTOS

**FÁCIL** Pouco ou nenhum contexto

Tokenização; Consultas em BDs

**DOMINANDO** Sensibilidade média ao contexto

Identificação de ponto final; Etiquetagem morfossintática

**Difícil** Dependência de contexto ilimitado

Tradução; Parseamento; Desambiguação

**DIFICÍLIMO** Parte do contexto fora do texto

Ironia; Anáfora temporal e sintagmática; Interpretação textual

# AMBIGUIDADE COM POUCO CONTEXTO

POR EXEMPLO, A PALAVRA **como** [KEPLER 2005]

- Sabei de certo que já **como** a muito boas horas ;
- (...) além de nos explicar **como** se ensinava a Gramática Latina , (...)
- Mas , seja **como** for , estas não são razões (...)
- (...) e o **como** se pode ordenar uma Gramática (...)



# AMBIGUIDADE COM POUCO CONTEXTO

## POR EXEMPLO, A PALAVRA **como** [KEPLER 2005]

- Sabei/VB-I de/P certo/ADJ que/C já/ADV **como** a muito boas horas ;
- (...) além/ADV de/P nos/CL explicar/VB **como** se ensinava a Gramática Latina , (...)
- Mas/CONJ ,/, seja/SR-SP **como** for , estas não são razões (...)
- (...) e/CONJ o/D **como** se pode ordenar uma Gramática (...)

# AMBIGUIDADE COM POUCO CONTEXTO

## POR EXEMPLO, A PALAVRA **como** [KEPLER 2005]

- Sabei/VB-I de/P certo/ADJ que/C já/ADV **como**/VB-P a muito boas horas ;
- (...) além/ADV de/P nos/CL explicar/VB **como**/WADV se ensinava a Gramática Latina , (...)
- Mas/CONJ ,/, seja/SR-SP **como**/CONJS for , estas não são razões (...)
- (...) e/CONJ o/D **como**/ADJ se pode ordenar uma Gramática (...)

## AMBIGUIDADE COM POUCO CONTEXTO

### POR EXEMPLO, A PALAVRA **como** [KEPLER 2005]

- Sabei/VB-I de/P certo/ADJ que/C já/ADV **como**/VB-P a/P muito/Q boas/ADJ-F-P horas/N-P ;/.
- (...) além/ADV de/P nos/CL explicar/VB **como**/WADV se/SE ensinava/VB-D a/D-F Gramática/NPR Latina/ADJ-F ,/, (...)
- Mas/CONJ ,/, seja/SR-SP **como**/CONJS for/SR-SR ,/, estas/D-F-P não/NEG são/SR-P razões/N-P (...)
- (...) e/CONJ o/D **como**/ADJ se/SE pode/VB-P ordenar/VB uma/D-UM-F Gramática/NPR (...)

# AMBIGUIDADE COM CONTEXTO ILIMITADO

- Palavra **mais** como advérbio de intensidade:

*Faça **mais**/ADV-I força*

- Palavra **mais** como comparador:

*Eu sei **mais**/ADV-R que/C você*

# DEPENDÊNCIAS DE DISTÂNCIA ILIMITADA

POR EXEMPLO, O PAR **mais/ADV-R** **que/C**

- Eu sei **mais/ADV-R** **que/C** você.

# DEPENDÊNCIAS DE DISTÂNCIA ILIMITADA

POR EXEMPLO, O PAR **mais/ADV-R** **que/C**

- Eu sei **mais/ADV-R** **que/C** você.
- Ele é **mais/ADV-R** bonito **que/C** você.

# DEPENDÊNCIAS DE DISTÂNCIA ILIMITADA

POR EXEMPLO, O PAR **mais/ADV-R** **que/C**

- Eu sei **mais/ADV-R** **que/C** você.
- Ele é **mais/ADV-R** bonito **que/C** você.
- Ele é **mais/ADV-R** pão duro **que/C** você.

# DEPENDÊNCIAS DE DISTÂNCIA ILIMITADA

POR EXEMPLO, O PAR **mais/ADV-R** **que/C**

- Eu sei **mais/ADV-R** **que/C** você.
- Ele é **mais/ADV-R** bonito **que/C** você.
- Ele é **mais/ADV-R** pão duro **que/C** você.
- Dedique-se **mais/ADV-R** aos seus estudos **que/C** à balada.



# DEPENDÊNCIAS DE DISTÂNCIA ILIMITADA

POR EXEMPLO, O PAR **mais/ADV-R** **que/C**

- Eu sei **mais/ADV-R** **que/C** você.
- Ele é **mais/ADV-R** bonito **que/C** você.
- Ele é **mais/ADV-R** pão duro **que/C** você.
- Dedique-se **mais/ADV-R** aos seus estudos **que/C** à balada.
- **Mais/ADV-R** vale um asno que me carregue **que/C** um cavalo que me derrube.

# DEPENDÊNCIAS DE DISTÂNCIA ILIMITADA

POR EXEMPLO, O PAR **mais/ADV-R** **que/C**

- Eu sei **mais/ADV-R** **que/C** você.
- Ele é **mais/ADV-R** bonito **que/C** você.
- Ele é **mais/ADV-R** pão duro **que/C** você.
- Dedique-se **mais/ADV-R** aos seus estudos **que/C** à balada.
- **Mais/ADV-R** vale um asno **que/WPRO** me carregue **que/C** um cavalo que me derrube.

# REDES NEURAIS

# VANTAGENS OFERECIDAS PELAS REDES NEURAIS

- capazes de capturar contextos cada vez maiores

# VANTAGENS OFERECIDAS PELAS REDES NEURAIS

- capazes de capturar contextos cada vez maiores
- capazes de fazer processamento de texto em vez de processamento de palavras

# VANTAGENS OFERECIDAS PELAS REDES NEURAIS

- capazes de capturar contextos cada vez maiores
- capazes de fazer processamento de texto em vez de processamento de palavras
- ao contrário dos métodos probabilísticos clássicos, não requerem dados de entrada independentes

# VANTAGENS OFERECIDAS PELAS REDES NEURAIAS

- capazes de capturar contextos cada vez maiores
- capazes de fazer processamento de texto em vez de processamento de palavras
- ao contrário dos métodos probabilísticos clássicos, não requerem dados de entrada independentes
- implementações disponíveis livremente e com comunidades ativas que as mantém

# ARQUITETURA BÁSICA PARA PROC DE TEXTO

**sentença:**  $w_1 \quad \dots \quad w_n \quad \langle pad \rangle_{n+1} \dots \langle pad \rangle_N$

**embedding:**  $\begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{d1} \end{bmatrix} \quad \dots \quad \begin{bmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{dn} \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \dots \quad \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

**codificação:**  $\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}$



# POSSÍVEIS SAÍDAS

- Classificação:

Exemplos: análise de sentimento, detecção de spam

- Geração de texto:

Exemplos: tradução, sequência-para-sequência (seq 2seq)

# ARQUITETURAS NEURAIS PARA TEXTOS

- redes multicamadas
- redes convolucionais
- redes recorrentes
- redes LSTM (long short term memory) e GRU
- Atenção neural
- Transformer
- BERT

# ARQUITETURAS NEURAIIS PARA TEXTOS

- redes multicamadas
- redes convolucionais
- redes recorrentes
- **redes LSTM (long short term memory) e GRU**
- Atenção neural
- Transformer
- BERT