

Universidade Federal de Campina Grande
Unidade Acadêmica de Sistemas e Computação
Curso de Bacharelado em Ciência da Computação

Projeto de Inteligência Artificial I

Predição de Texto sob Contexto usando Cadeias de Markov

Rafae N Cipriano
rafael.cipriano@ccc.ufcg.edu.br



Definição do problema

Em certas interfaces de digitação é conveniente a presença de um sistema de predição de texto acurado que acelere a síntese da mensagem

Existem diversas formas de se relacionar dados de entrada de modo que se maximize esta acurácia, envolvendo conceitos de dedução e aprendizado não-supervisionado



Conceitos de IA Envolvidos

- aprendizado não supervisionado
- dedução



Solução Implementada

A solução proposta utiliza Cadeias de Markov para prever as palavras candidatas. Cadeias de Markov são uma sequência de eventos estocásticos baseados em probabilidade onde o estado atual de uma variável independe dos estados passados, exceto do atual, ou seja, não há contexto



Solução Implementada

Deste modo, para considerar o contexto na predição, serão utilizadas duas Cadeias de Markov: uma para mapear as transições entre estados que representam partes de palavras e palavras inteiras (possíveis complementos) e outra para mapear as transições entre estados que representam palavras (possíveis contextos)



Solução Implementada

Exemplos de processamento de palavras:

'só' \rightarrow {'s', 'só'}

'sair' \rightarrow {'s', 'sair'} {'sa', 'sair'} {'sai', 'sair'}

'vou' \rightarrow {'v', 'vou'} {'vo', 'vou'}

'eu' \rightarrow {'e', 'eu'}



Solução Implementada

Exemplos de processamento de contexto:

'eu vou sair' \rightarrow {'eu', 'vou'} {'vou', 'sair'}

'eu vou só' \rightarrow {'eu', 'vou'} {'vou', 'só'}

Solução Implementada

Cadeias de Markov:

Palavras:

{'e' : {'eu' : 2}} {'v' : {'vou' : 2}} {'vo' : {'vou' : 2}} {'s' : {'só' : 1, 'sair' : 1}} {'sa' : {'sair' : 1}} {'sai' : {'sair' : 1}}

Contexto:

{'eu' : {'vou' : 1}} {'vou' : {'sair' : 1, 'só' : 1}}



Solução Implementada

Exemplos de saída:

'eu v' → contexto identificado: 'eu';
candidatos: 'vou'; selecionado: 'vou'
(repesagem)

'eu vou s' → contexto identificado: 'vou';
candidatos: {'sair', 'só'}; selecionado: 'sair'
(repesagem)



Resultados Obtidos

A metodologia aplicada ao problema apresenta soluções satisfatórias de modo que, a partir do momento em que uma palavra é repetida uma quantidade suficiente de vezes, o sistema reposiciona essa palavra na lista de palavras candidatas e, inclusive, de acordo com o contexto



Resultados Obtidos

O sistema foi implementado com sucesso, ordenando os candidatos do modo esperado, sendo possível inclusive treinar o sistema em tempo de execução, o que facilita testes e demonstração

Foi possível concretizar a proposta de considerar o contexto nos cálculos da predição, utilizando Cadeias de Markov em conjunto para tal



Demonstração Prática

- Favor executar o sistema!



Obrigado!