

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



Conceitos de IA Envolvidos

Este problema pode ser formalizado com base no conceito de Rede Bayesiana de modo que seja possível mapear as sequências de caracteres e palavras em transições entre estados com relações probabilísticas por ordem de ocorrência. Podemos realizar a predição dos candidatos a partir deste modelo.



Conceitos de IA Envolvidos

A representação da rede bayesiana é feita por meio de um grafo direcionado acíclico no qual os nós representam variáveis de um domínio e os arcos representam a dependência condicional ou informativa entre as variáveis. Para representar a força da dependência, são utilizadas probabilidades, associadas a cada grupo de nós pais-filhos na rede (PEARL, 1988)



Conceitos de IA Envolvidos

Cadeias de Markov finitas e discretas podem ser descritas por meio de um grafo dirigido (orientado), onde cada aresta é rotulada com as probabilidades de transição de um estado a outro sendo estes estados representados como os nós conectados pelas arestas



Solução Implementada

A solução proposta faz uso de duas cadeias de Markov com pesos ajustados de acordo com as entradas capturadas: na primeira serão mapeadas as palavras candidatas para uma parte de palavra enquanto na segunda cadeia serão mapeadas as transições entre palavras



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



Demonstração Prática

- Favor executar o sistema!



Obrigado!