



Título: Compressão de Sentenças sobre Domínios com Disponibilidade Limitada de Dados Rotulados

Horário: 09:00h

Data: 27/11/2019

Local: Bloco 952 (Sala de Seminários)

Resumo:

Dia após dia, volumes gigantescos de dados são produzidos na Web. São grandes quantidades de vídeos, imagens e textos que armazenam informação de maneira não estruturada. Sistemas de sumarização de textos foram criados com o intuito facilitar a apresentação de grandes quantidades de dados textuais para usuários assim como para facilitar a recuperação de informações nesses tipos de dados. A tarefa de compressão de

sentenças surgiu para melhorar a qualidade dos sumários gerados por esses sistemas. Este trabalho propõe um modelo de compressão de sentenças capaz de atingir resultados competitivos com uma quantidade de dados menor que outros modelos baseados em redes neurais, aliado a uma estratégia de redução de palavras raras que busca maximizar a ocorrência de sentenças com palavras comuns. Em domínios com pouca disponibilidade de dados de compressão de sentenças rotulados, modelos baseados em redes neurais tem grandes dificuldades de extrair todas as informações que eles precisam. Para melhorar a performance destes modelos neste cenário, é necessário que parte da informação seja extraída e adaptada antes de serem utilizadas para o treinamento. Este trabalho compara o modelo proposto com outros modelos baseados em redes neurais e verifica que este se sai melhor em reproduzir as compressões de sentenças dos dados de avaliação, com quantidades menores de dados de treinamento.

Banca:

- Prof. Dr. José Antônio Fernandes de Macêdo(MDCC/UFC - Orientador)
- Prof. Dr. Ticiania Linhares Coelho da Silva (UFC)
- Prof. Dr. João Paulo Pordeus Gomes(MDCC/ UFC)
- Prof. Dr. Vlândia Célia Monteiro Pinheiro(Unifor)