

Resumo dos Artigos

Mineração de Texto 2019.2

Mateus Meneses Fittipaldi

Artigo 1 - Text Mining: Open Source Tokenization Tools - An Analysis

O artigo em questão faz uma análise de sete ferramentas de tokenização que são de código aberto. Na análise, a pesquisadora leva em consideração a limitação e a saída das ferramentas. Dentre todas as ferramentas a que foi considerada a “melhor” foi a Nlpdotnet Tokenizer, pois aceita vários arquivos de texto como entrada e é fácil de trabalhar com o resultado. Com essa análise podemos optar pela melhor ferramenta de tokenização para trabalhar em nosso próprio projeto.

Artigo 2 - Translation of a Selected Text Fragment of a Screen

Esse artigo sugere uma aplicação que permite o usuário selecionar um bloco da tela do dispositivo e fazer a tradução do texto presente neste bloco. O método de tradução é muito bem explicado o que pode ser um ótimo guia para usarmos como base para nosso projeto.