

Lab2 - Expansão de Consultas

[Enviar Tarefa](#)

Vencimento terça-feira por 0:59 **Pontos** 40
Enviando uma caixa de entrada de texto ou uma URL de site
Disponível 5 jun em 1:00 - 14 jun em 0:59 9 dias

Nesta atividade você vai exercitar a noção de expansão de consultas. Considerando a coleção de notícias do lab passado, você primeiro precisa executar os seguintes passos:

1. Escreva uma função que receba uma coleção de documentos e retorne uma matrix de termos-termos contendo as frequências de co-ocorrência de duas palavras consecutivas no texto (bigramas).
2. Escreva uma função que receba um certo termo de consulta e a matriz construída no passo 1 acima e retorneas top-3 palavras em ordem decrescente de frequencia.
3. Expanda a consulta original com os termos retornados no passo 2 acima.
4. Faça uma busca disjuntiva (OR) considerando a nova consulta.

Por exemplo, se o termo de consulta é *petrobrás* e as palavras que mais co-ocorrem com *petrobrás* são: *empresa*, *gasolina* e *estatal*, a consulta agora ficaria: *petrobrás OR empresa OR estatal*.

Escolha livremente três termos de consulta e responda o seguinte:

- Quais os termos retornados para a expansão de cada consulta?
- Você acha que esses termos são de fato relacionados com a consulta original? Justifique.
- Compare os documentos retornados para a consulta original com a consulta expandida. Quais resultados você acha que melhor capturam a necessidade de informação do usuário? Por que?
- A expansão de consultas é mais adequada para melhorar o recall ou o precision? Por que?