1. Definir o escopo do problema:

| **SUGESTÃO DE TEMA** | |
| --- | --- |
| **Título do trabalho** | ***Dê um título ao seu trabalho, pense que pode ser o título do artigo.*** |
| **Contexto básico do assunto** | ***Aqui você descreve o contexto no qual está fazendo o trabalho.***   * Diferenciar documentos diversos de aleatórios   + Aleatório: Qualquer documento (dentro do mesmo corpus)   + Diverso: Diferente do normal (dentro do mesmo corpus) * Todo aleatório é diverso mas nem todo diverso é aleatório * Explorar documentos diversos mas não aleatórios (objetivo) |
| **Problema identificado** | ***Qual é o problema que você está tentando resolver com seu trabalho? É importante que haja um problema bem definido!***   * Melhorar acurácia de RAG * Hipótese: Documentos relevantes diversos melhoram acurácia |
| **Intervenção proposta** | ***O que será feito para resolver o problema?***  \*Igual ao artigo The Power of Noise   * Usar dataset NQ (\*) * Dada uma pergunta, buscar documentos com Contriever/BM25 (\*) * Rerankear documentos com clusterização   + **1° experimento:** Clusteriza rank do Contriever/BM25 (Usando K-means)   + **2° experimento:** Usa RagFusion e clusteriza   + **3° experimento:** Usa RagFusion mas para no meio, usando as categorias intermediárias   + Injetar documentos aleatórios em cada experimento (\*), totalizando 6 experimentos (3 originais, com e sem a injeção de documentos aleatórios) * Evitar documentos similares para aumentar diversidade |
| **Produção esperada** | ***O que você pretende entregar na conclusão do trabalho?***   * Código referente aos experimentos * Análise das acurácias dos experimentos |
| **Contribuições esperadas** | ***Como o trabalho contribui para a solução do problema definido?***   * Novas metodologias para aplicação RAG |