

## IR Mini Project - Understanding Text in Health Domain - (Find Doctor)

### ➔ Team Details

Patel Kartik Niranjan – S20210010171

Virendra Yadav – S20210010241

### ➔ Patel Kartik Niranjan – S20210010171

- Task Done
  1. Web Crawling (Practo.com)
    - a. Crawling
  2. Build Web Application
  3. Inverted Indexing (posting list for all terms)
  4. Finding TF-IDF value of data
    - a. Calculation of IDF
    - b. Calculation of TF using freq. method
    - c. Calculation of TF-IDF
  5. User Feedbacks
  6. Plot P-R curve using matplotlib based on user feedback

### ➔ Virendra Yadav – S20210010241

- Task Done
  1. Web Crawling
    - Data Processing
  2. Query processing
    - Finding Tf-Idf for query
  3. Cosine Similarity – using TF-IDF score
  4. Retrieval of top 20 Documents
  5. Evaluation : Calculate P-R values based on feedbacks
    - Calculated Precision values
    - Calculated Recall Values