

# Baseline Results Documentation

## Introduction:

The purpose of this report is to present the results obtained from evaluating the prototype of our proposed solution for information retrieval. The prototype was designed to demonstrate the effectiveness of our approach in retrieving relevant information from various datasets. This report outlines the metrics used for evaluation and provides an analysis of the obtained results.

## Prototype Description:

- The prototype is a basic implementation of an information retrieval system.
- It retrieves relevant documents based on user queries.
- The prototype uses different datasets, including
  - CourseMaterial.csv
  - PaperChecking.csv
  - SingleQA.csv
  - MCQ.csv, for evaluation.

## Evaluation Metrics:

To evaluate the performance of the prototype, the following metrics were used:

- Precision
- Recall
- F1-score

## Baseline Results:

### CourseMaterial.csv:

```
Metrics for CourseMaterial.csv:

Precision : 0.12698412698412698
Recall    : 0.7272727272727273
F1-score  : 0.2162162162162162
```

### PaperChecking.csv:

```
Metrics for PaperChecking.csv:

Precision : 0.2
Recall    : 0.9090909090909091
F1-score  : 0.32786885245901637
```

#### SingleQA.csv:

##### Metrics for singleQA.csv:

```
Precision : 0.32
Recall    : 0.48484848484848486
F1-score  : 0.38554216867469876
```

#### MCQ.csv:

##### Metrics for MCQ.csv:

```
Precision : 0.45
Recall    : 0.8181818181818182
F1-score  : 0.5806451612903226
```

#### Analysis:

- The prototype exhibits varying performance across different datasets, indicating the influence of dataset characteristics on retrieval effectiveness.
- Higher precision and recall scores are observed for datasets with clearer and more specific queries, such as PaperChecking.csv and MCQ.csv.
- Lower precision and recall scores for CourseMaterial.csv and singleQA.csv suggest the presence of noise or ambiguity in the data.
- Further optimization of the retrieval system, such as query expansion techniques or advanced ranking algorithms, could improve performance across all datasets.
- The evaluation results provide valuable insights for refining the prototype and enhancing its effectiveness in retrieving relevant information from diverse datasets.

#### Conclusion:

The evaluation of the prototype suggests that our proposed solution for information retrieval shows promising results across various datasets. The prototype effectively retrieves relevant information, as evidenced by the calculated precision, recall, and F1-score. Further improvements and optimizations can be made based on these results to enhance the prototype's performance.

This concludes the evaluation report of our IR prototype. Further development and refinement will be guided by the insights gained from this evaluation process.