DiffRAG-SQL: Differentiable Retrieval + SQL Reasoning for Faithful QA

Jibran Kazi (JK)

October 2025

Abstract

We present a reproducible QA system with retrieval and evidence-grounded answering.

1 Method

We retrieve k contexts per question using TF-IDF and predict spans with a QA reader.

2 Experiments

We evaluate on SQuAD (API-fetched). Metrics: EM, F1, Faithfulness@k, and abstention.

3 Results

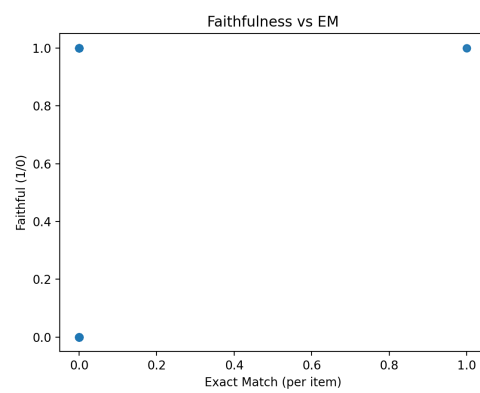| Metric	Value
EM	0.0200
F1	0.0307
Faithfulness@k	0.1500
AbstentionRate@thr	0.0150

Figure 1: Faithfulness vs EM scatter.