

Title: Intelligent Exam Generator Using Deep Learning



Robust on QA Faculty

like Levenshtein

dataset: excels in tasks

Student Name: Mostafa Abdelaziz Abdullah

Supervisor: Mostafa Salama

METEOR: 0.8540. EM:

49.34, F1: 60.20

QA (LLaMA-2-7B with

LoRA)

Abstract

- Developed an AI tool to automate exam creation, saving educators time.
- Uses NLP and deep learning to generate diverse, personalized questions.
- Enhances scalability and adaptive learning in computer science education.

System Design/Interface

Employs SciFive for question generation, LLaMA-2-7B with LoRA for question answering, and Flan-T5 for distractor creation, ensuring diverse, high-quality exam content.

DistilBERT-based difficulty adjuster dynamically tailors question complexity to student proficiency levels using synthetic profiles.

Features a modular architecture with seamless integration of NLP models, processing BUE ICS datasets for computer science topics.

Programmatically processes input materials to generate customized exams with varied question types efficiently.

. 133 5.113			
Model	Metrics Achieved	Benchmark Comparison	Significance
Question Generation (SciFive)	METEOR: 0.9054, BLEU: 0.1503, ROUGE-L: 0.4175	Surpasses T5 (METEOR: 0.83), GPT-3.5 (0.85), BERT (0.70)	High semantic accuracy on Faculty dataset; training loss reduced from 3.2148 to 0.0599.

Outperforms BERT

(0.70). LLaMA-2-13B

(0.75); rivals GPT-3.5

Results

(0.85)

distance calculation etc.

Competitive with GPT3.5 (BLEU: 0.30,
ROUGE-1: 0.53, METEOR: 0.39

Distractor (Flan-T5)

BLEU: 0.29, ROUGE-1: 0.50);
Surpasses BERT (BLEU: 0.39)

distance calculation etc.

Competitive with GPT3.5 (BLEU: 0.30,
ROUGE-L: 0.50);
Surpasses BERT (BLEU: 0.50);
Surpasses BERT (BLEU: 0.3951.

Conclusions

- Pioneered custom datasets and distractor system for Aldriven exams.
- Enables personalized, scalable assessments in education.
- Future work: Expand question types and optimize computational efficiency.