

Student Name: Mostafa Abdelaziz Abdullah

Supervisor: Mostafa Salama

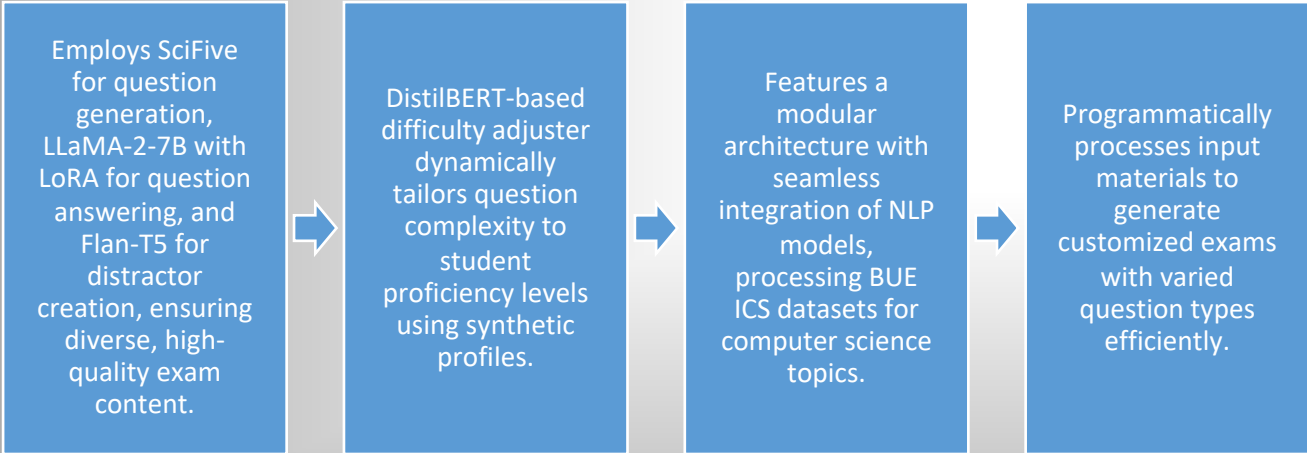
Abstract

- 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.

Results

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.
QA (LLaMA-2-7B with LoRA)	METEOR: 0.8540, EM: 49.34, F1: 60.20	Outperforms BERT (0.70), LLaMA-2-13B (0.75); rivals GPT-3.5 (0.85)	Robust on QA Faculty dataset; excels in tasks like Levenshtein distance calculation etc.
Distractor (Flan-T5)	BLEU: 0.29, ROUGE-1: 0.53, METEOR: 0.39	Competitive with GPT-3.5 (BLEU: 0.30, ROUGE-L: 0.50); surpasses BERT (BLEU: 0.25)	Generates plausible distractors for 626 MCQs; training loss dropped from 11.69 to 0.3951.

System Design/ Interface



Conclusions

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