

Spare IT - LLM Report Summary

Friday, September 8, 2023

4:06 PM

Step	Description
Image Segmentation	1. Segment the images into logical groups or themes. This could be done manually or using an image segmentation algorithm.
Image to Text Conversion	2. For each image segment, use a Vision AI model like DALL-E to generate a text description of the image content.
Text Aggregation	3. Aggregate the text descriptions for each segment.
Language Model Fine-tuning	4. Fine-tune a large language model like GPT-3 on a dataset of reports in your domain. The reports should have an abstractive summary reflecting the key points.
Prompt Engineering	5. Construct a prompt for each image segment using the generated text descriptions. Include instructions to generate an abstractive summary reflecting the key points.
Summary Generation	6. Feed the prompts to the fine-tuned language model to generate a summary for each segment.
Report Aggregation	7. Aggregate the summaries for all segments to form the full report.
Refinement	8. Refine and polish the final report using tools like Langchain and having a human editor

	review.
Evaluation and Iteration	9. Evaluate the reports by having target users review and provide feedback. Refine the models based on the feedback.
Interpretability	Deep models can be hard to interpret
Size	Smaller compared to BERT & GPT
Tokenization	Character-based