

Question Answering Pipeline for Closed Domain Questions

Research Work for Data Science Project
Competition

Presentation overview

Problem Description	Question-Answering system for the NLB Group.
Proposed Solution	Test various implementations of a QA LLM.
Experimentation	Model selection and fine-tuning methods.
Presentation of Results	Metric scores for models and pipelines.
Discussion	Usability and possible improvements.
Questions	For any additional information, the project repository is available here .

Problem Description

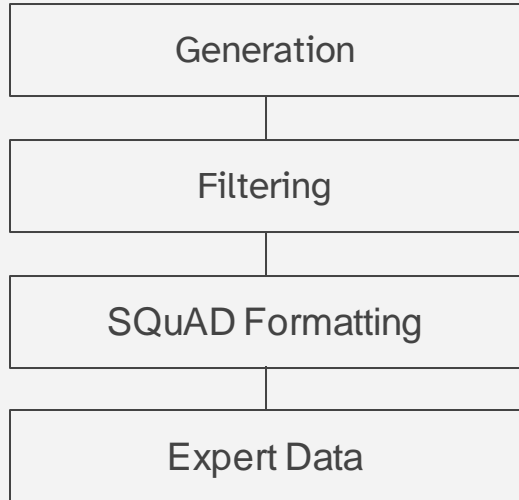
Extract from the Data Science Project Competition Website

In56ight, in collaboration with the NLB banking group, is exploring the idea of developing a chatbot that would be trained on data from annual reports and would be capable of answering different questions regarding the banking group performance.

Year 2023

Proposed Solution

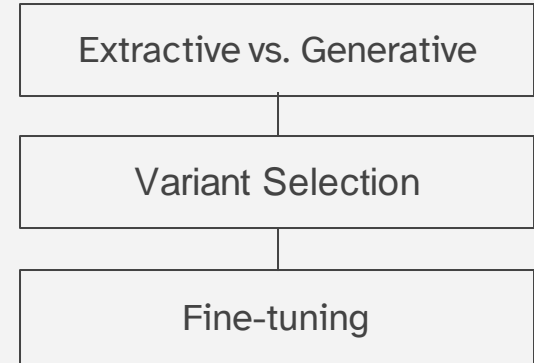
Question Answering Dataset



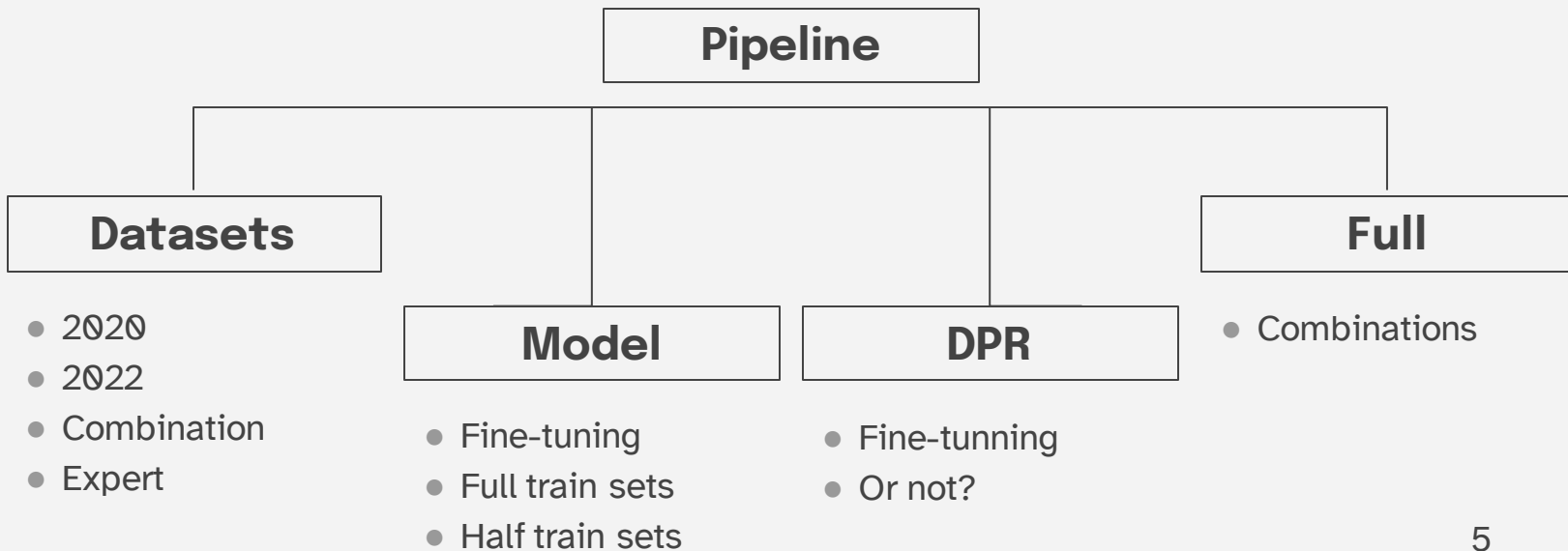
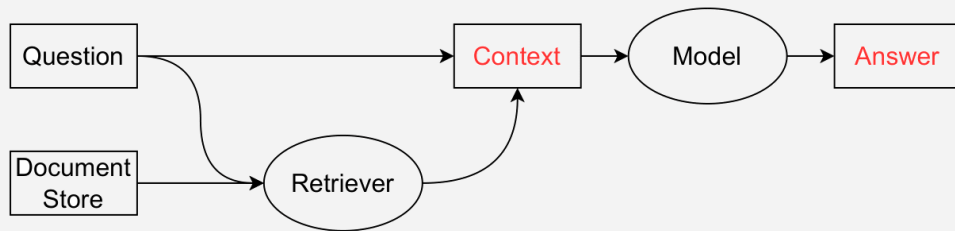
Context Retrieval



Large Language Models



Experimentation



Results: Metrics

		Bert Precision	Bert Recall	Bert F1	Bleu	SQuAD Exact	SQuAD F1
DPR not fine-tuned	Distilbert	86.98	86.94	86.93	0	4.76	11.58
	RoBERTa	88.15	87.33	87.63	4.61	4.76	21.53
	T5-Small	88.78	87.44	88.04	13.48	/	/
	T5-Base	86.11	88.20	87.06	24.60	/	/
DPR fine-tuned	Distilbert	87.41	87.19	87.25	0	4.76	16.87
	RoBERTa	88.78	87.75	88.21	5.1	14.29	27.57
	T5-Small	86.92	85.23	86.03	0.88	/	/
	T5-Base	85.45	87.42	86.38	18.75	/	/

Extract from Table 4 of the report: Pipeline performance using fine-tuned models (expert data only)

Results: Qualitative Analysis

		Correctness				Sensibility				"None"
		1	2	3	Avg[%]	1	2	3	Avg[%]	
2020/22	Distilbert	13	1	7	57	7	5	9	70	0
	RoBERTa	13	1	7	57	7	5	9	70	0
	T5-Small	13	1	7	57	7	6	8	68.3	0
	T5-Base	13	1	7	57	7	5	9	70	0
Expert	Distilbert	17	3	1	41.3	8	6	7	65	0
	RoBERTa	18	3	0	38	11	4	6	58.7	0
	T5-Small	16	5	0	41.3	12	3	6	57	7
	T5-Base	17	3	1	41.3	14	1	6	54	0

Table 5 of the report: Qualitative analysis using defined metrics

Results: Qualitative Analysis

Question: How much did the bank reduce the use of paper in 2022 compared to the previous year?

Ground truth answer: 19%

Retrieved context: In 2022, another step to achieving this goal has been made as paper usage (prints) was reduced by 17% compared to 2021. Thus, we saved 866 trees. Paper usage in NLB was lower by 26%, which is the highest reduction group – wide. All banking subsidiaries experienced reduction as well, ranging from 8% to 19%, respectively.

Our answer(s): 17%

ChatGPT answer: Based on the given context, the bank reduced the use of paper by 17% compared to the previous year (2021). However, the exact amount of paper usage in 2021 or any specific quantity of paper used is not mentioned.

Discussion

Findings about the Data	<ul style="list-style-type: none">• More data would be required.• Fine-tuning on expert data yields the most improvements.• Specific types of questions
Model Usefulness	<ul style="list-style-type: none">• Useful for more straightforward questions.• Highly depended on the DPR model.• The pipeline requires improvements before production.
Possible Improvements	<ul style="list-style-type: none">• More high quality data• DPR improvements• Larger models

Thank you for your attention

All models are available on **HuggingFace.**

Special thanks to

- prof. dr. Marko Robnik Šikonja
- dr. Branislava Šandrih Todorović
- Grega Jerkič and In615ht