Getting the Doc's Answers

Steve Hall
Angelina Wedemeyer

Executive Summary

Question-answering is **challenging in highly technical, specific domains** like healthcare

One solution is to **retrieve similar questions** that have been answered correctly in the past, and use those answers as "context" for new questions

Dr. Asma Abacha, in 2019, demonstrated that **question entailment could improve consumer health question answering**

We **extend her research** by applying the latest in NLP architecture (T5) and leveraging the large medical question-answering dataset she developed to approach this problem

Our modeling results show that **question entailment modestly improves** over a baseline pre-trained T5 model, but predicted answers are often inaccurate

Motivation

Online healthcare utilization has grown 38X since before the pandemic[1] borne out of necessity and now driven by growing consumer preference.

Since 2020, **investment in virtual care and digital health has grown 3X**, according to the same McKinsey study..

Accurate and fast healthcare question-answering systems will be necessary to meet the demands of consumers, but the **challenge is monumental**:

- Limited data due to privacy reasons
- Specific terminology (e.g. glaucoma)
- Costly and potentially harmful inaccurate responses

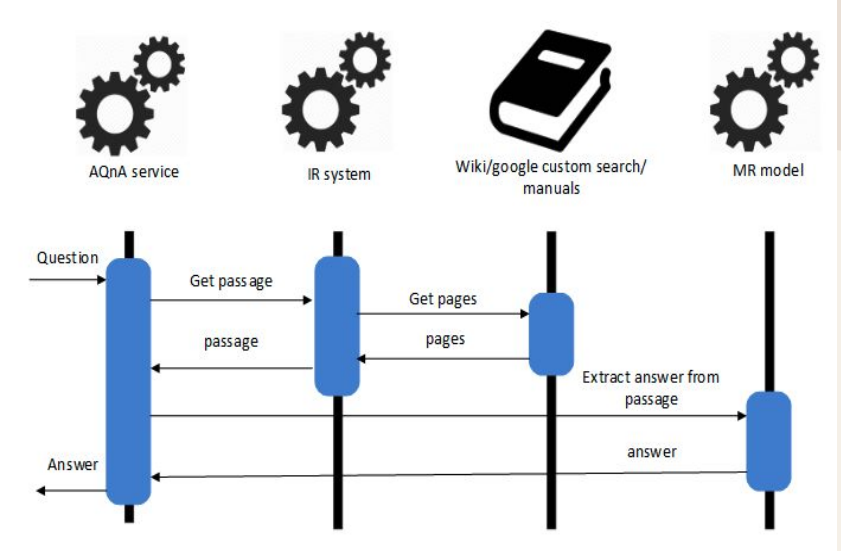
Sources: https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/telehealth-a-quarter-trillion-dollar-post-covid-19-reality

Background

Automated question-answering systems typically involve information retrieval and NLP machine reading comprehension models (MR) to extract the correct answer from given context.

This is adequate for generic, short-form question answering. But what if this is the question?

> My name is [NAME] and I am 76 years old. I have had jock itch (just itch no rash) for 7 years. When I try to stop the itch comes back with a vengeance. I am currently using Ketoconazole 2% cream to control the itch. It appears that I am doomed to no cure but I hope I may be wrong.



Dr. Abacha developed a large medical question-answering dataset and question-entailment approach to tackle this problem. In her conclusion, she recommended to investigate transfer learning. We continue her work by using her dataset and a T5 model to answer consumer health questions.

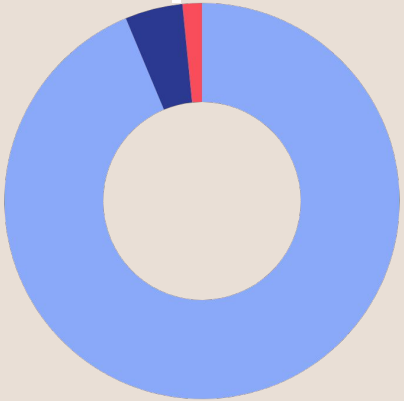
Sources: https://towardsdatascience.com/automatic-question-answering-ac7593432842

Data

Training: MedQuAD

The large dataset Dr. Abacha developed to entail questions consists of 47,457 question-answer pairs extracted from 12 trusted medical sources (e.g. cancer.gov, niddk.nih.gov, MedlinePlus Health Topics, etc.)

- Our training dataset was **11,481** questions
- Our validation dataset was **4,921** questions



Training: TREC-2017 LiveQA

A collection of questions received by the U.S. National Library of Medicine

- Our training dataset was **204** questions
- Our validation dataset was **102** questions

Evaluation: TREC-2017 LiveQA

- Our test dataset was **103** questions

Methodology

We explored the following three approaches to consumer health question-answering:

01 An evaluation of the latest **transformer encoder-decoder architecture** (T5) question-answering system on TREC 2017 LiveQA medical questions [4], our baseline model

02 A deeper study of **transfer learning by fine-tuning** our baseline T5 model on MedQUAD before answering the TREC 2017 Live QA medical questions

03 An extension of Dr. Abacha's **question-entailment approach by clustering** MedQUAD questions, finding the most similar cluster and questions within that cluster that entail the newly posed question, and then using the answer(s) of those entailed questions as context for the question-answering system

Results

The model with additional context performed better on ROUGE-1 and ROUGE-L metrics

	ROUGE-1	ROUGE-2	ROUGE-L
T5 (baseline)	9.40	**1.87**	7.58
T5 + MedQuAD	10.84	1.52	8.54
T5 + MedQuAD + Context	**15.17**	1.21	**9.29**

Example Of Answer Prediction

Patient Question:
- I need to lose fat

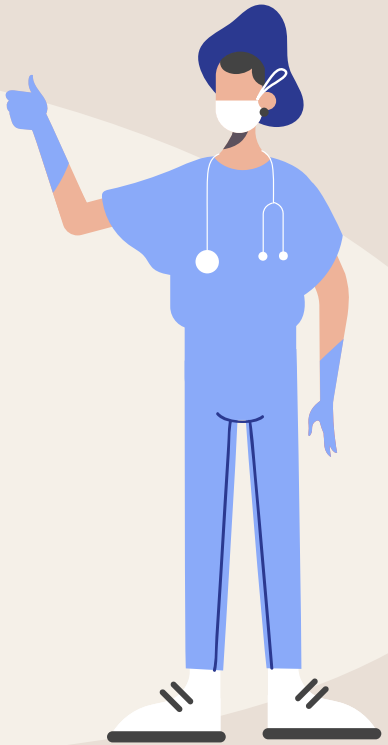
T5 (baseline) Predicted Answer:
- Diet and exercise can help you lose weight.

T5 + MedQuAD Predicted Answer:
- - Lose weight. - Maintain a healthy weight. - Maintain a healthy weight....

T5 + MedQuAD + Context Predicted Answer:
- the most common way to find out whether youre overweight or obese is to figure out your body mass index bmi bmi is an estimate of body fat and its a good gauge of your risk for diseases that occur with more body fat...

Conclusion and Next Steps

- Our results demonstrate that **question similarity can be a powerful approach** to automating question-answering in a highly specific domain, like healthcare.
- To meet the demands of changing consumer healthcare preferences, **greater investment and resources will need to be devoted** to improve the speed and accuracy consumer health question answering.
- In the future, we would **propose exploring question summarization** to reduce the complexity and excessively descriptive information that can be distracting for a NLP model.



Thanks!