

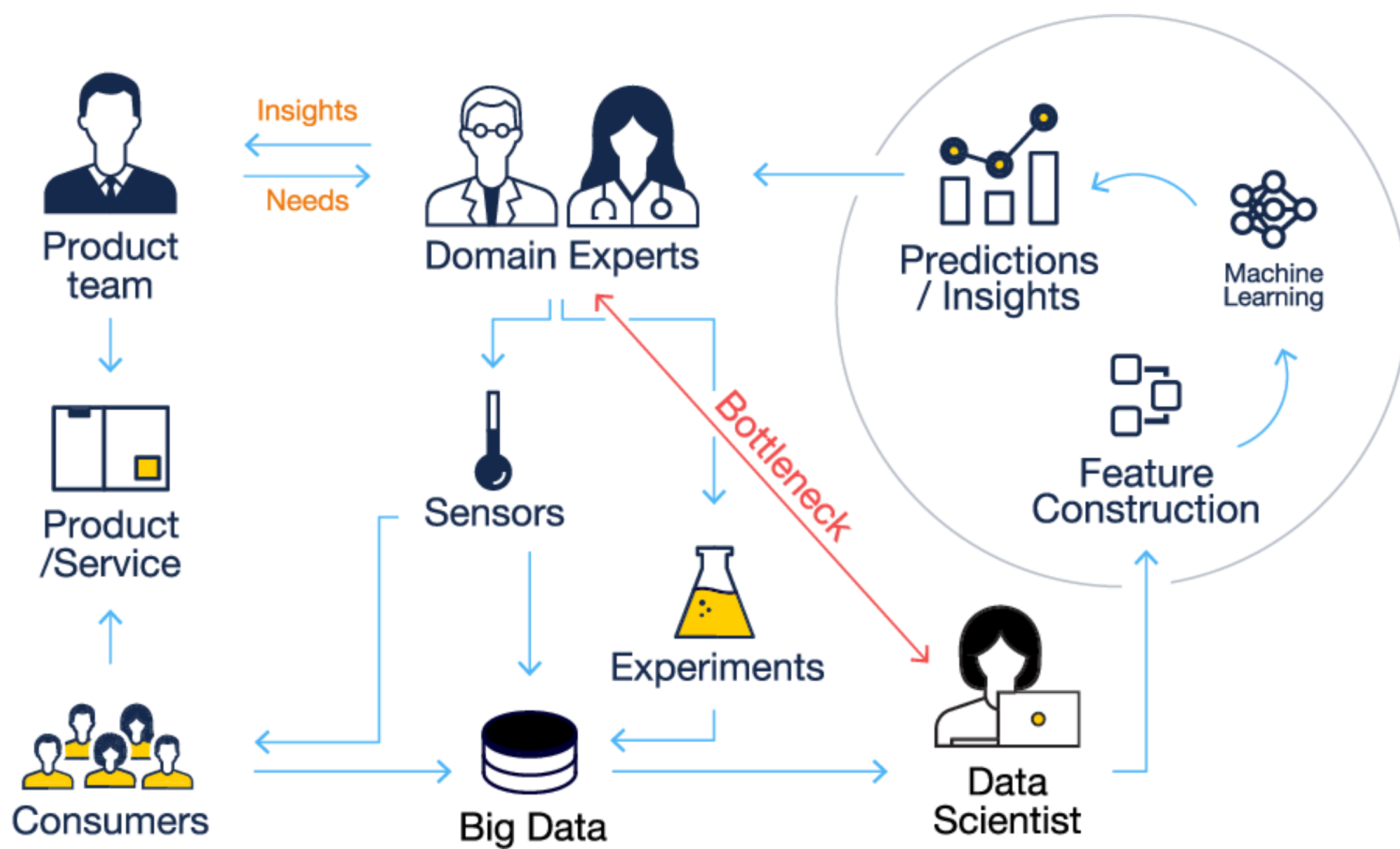


Large Language Model as Personal Data Scientist

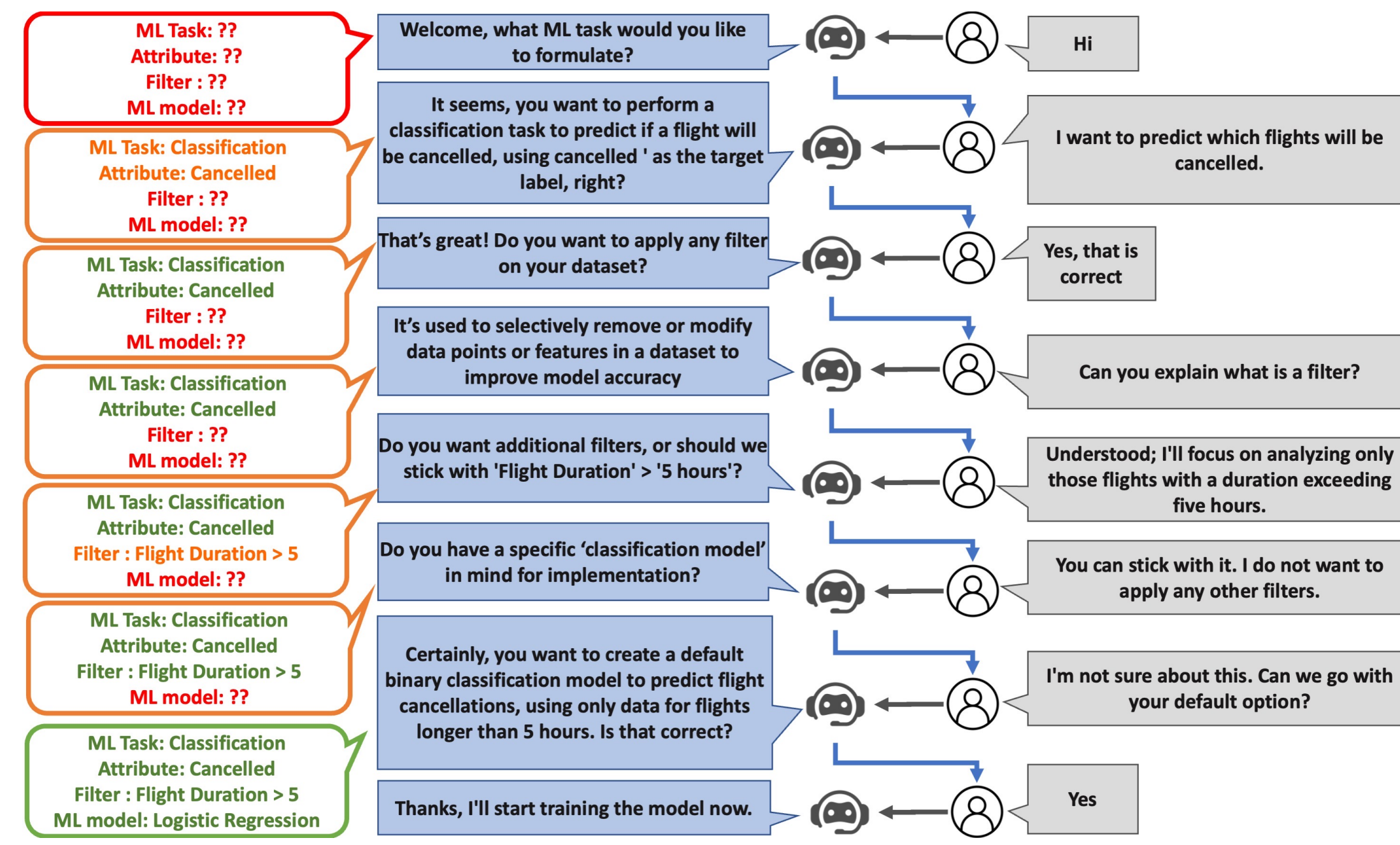
Md Mahadi Hassan, Alex Knipper, Shubhra Kanti Karmaker
Department of Computer Science and Software Engineering

Motivation

- ❖ Small businesses rarely contain a highly competent data science team.
- ❖ Hiring external consultants is often expensive and data privacy often becomes a bottleneck when involving a third party.
- ❖ Reluctance of the ML community to automate task formulation.



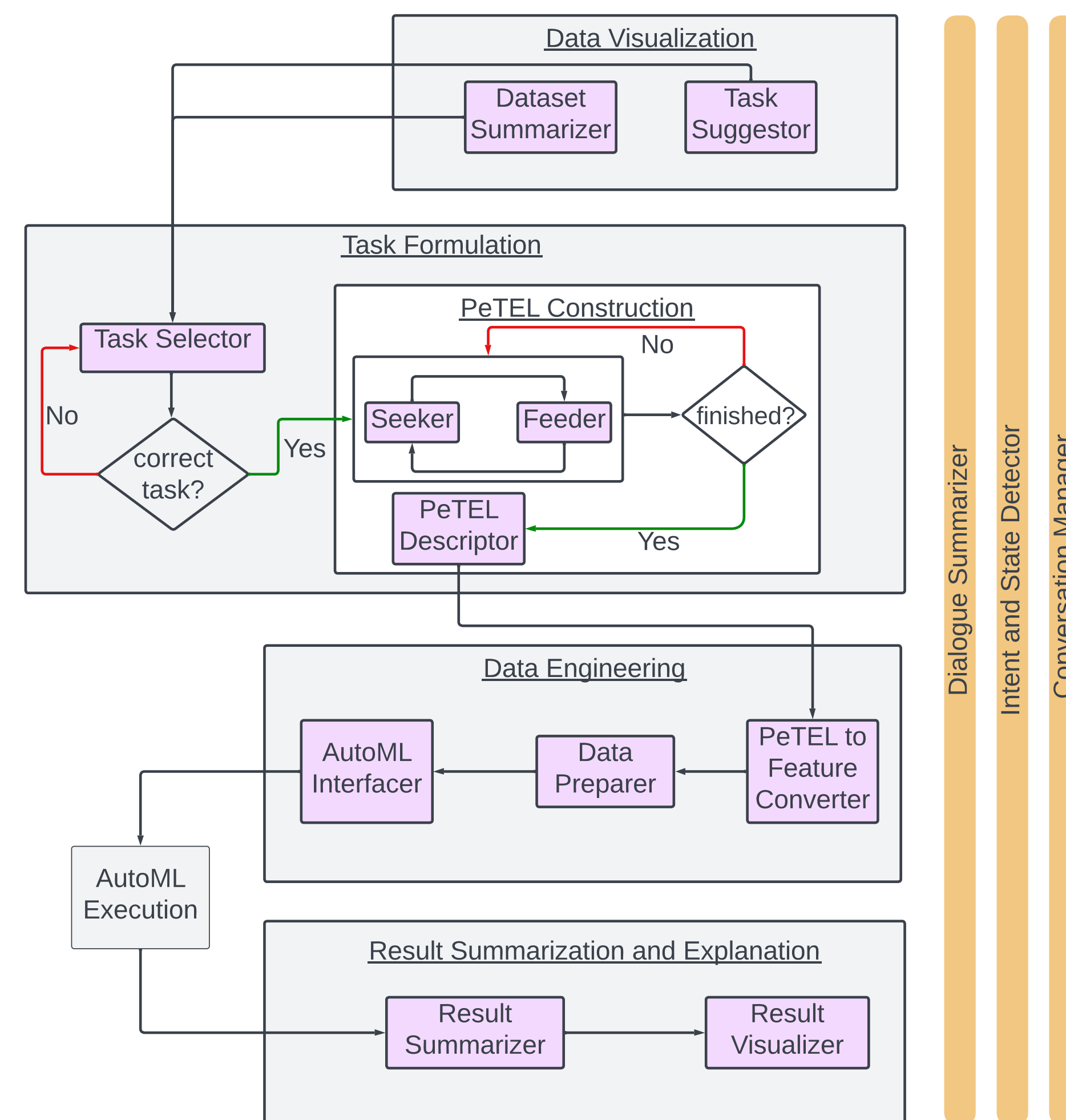
Conversational Data Science



Research Questions

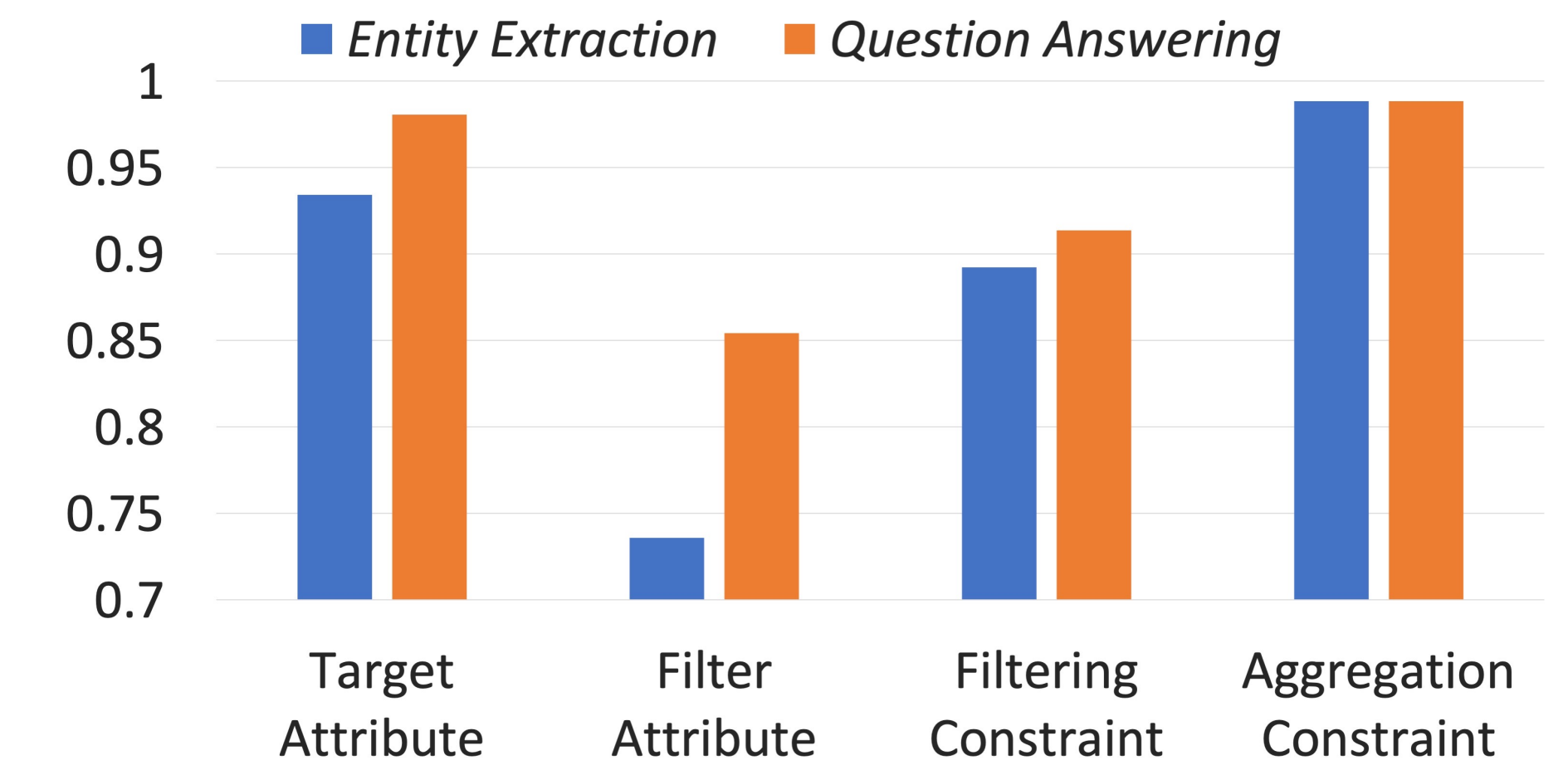
- ❖ Is conversational AI a feasible way to perform data science tasks?
- ❖ How accurate are LLMs in framing and solving ill-defined complex data science tasks?
- ❖ What are the common challenges involved with systems like 'Virtual Interactive Data Scientist'?

Model Architecture

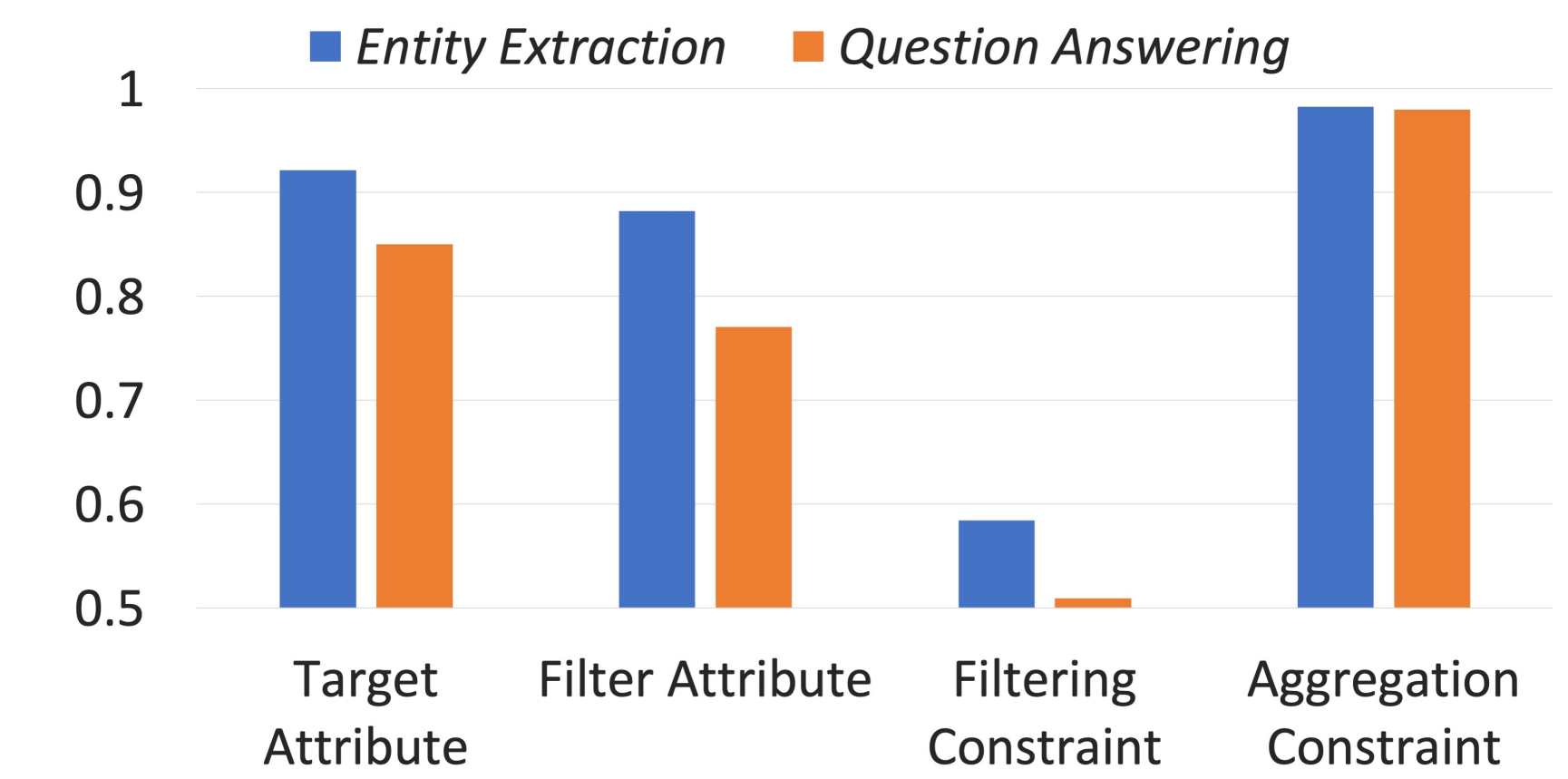


Quantitative Results

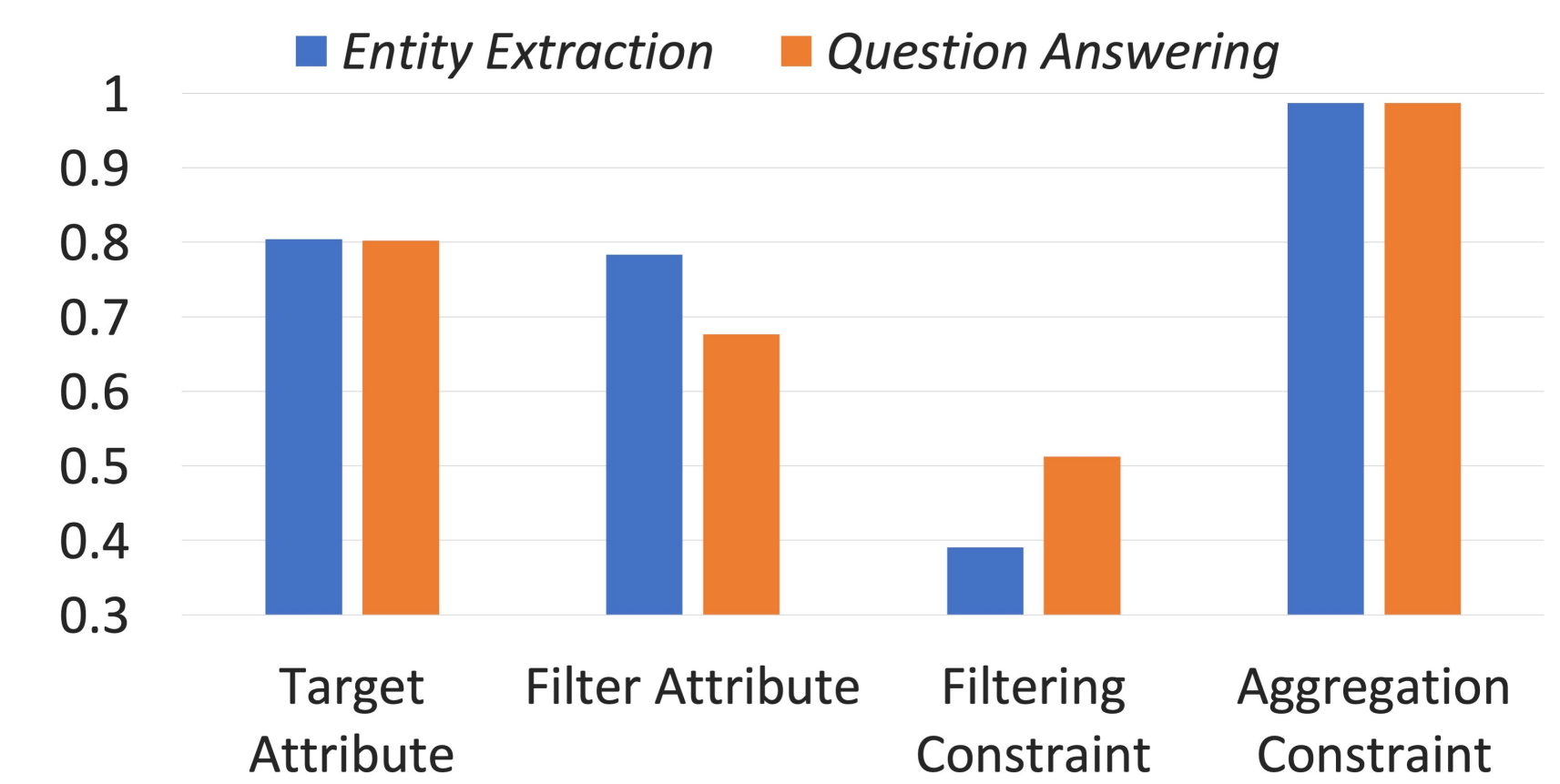
Our findings highlight that, even without LLMs, small models can achieve significant improvements in performance when fine-tuned on synthetic data. This underscores the viability and potential of 'Conversational Data Science'. We experimented with three public dataset.



MRR for Flight Delay dataset



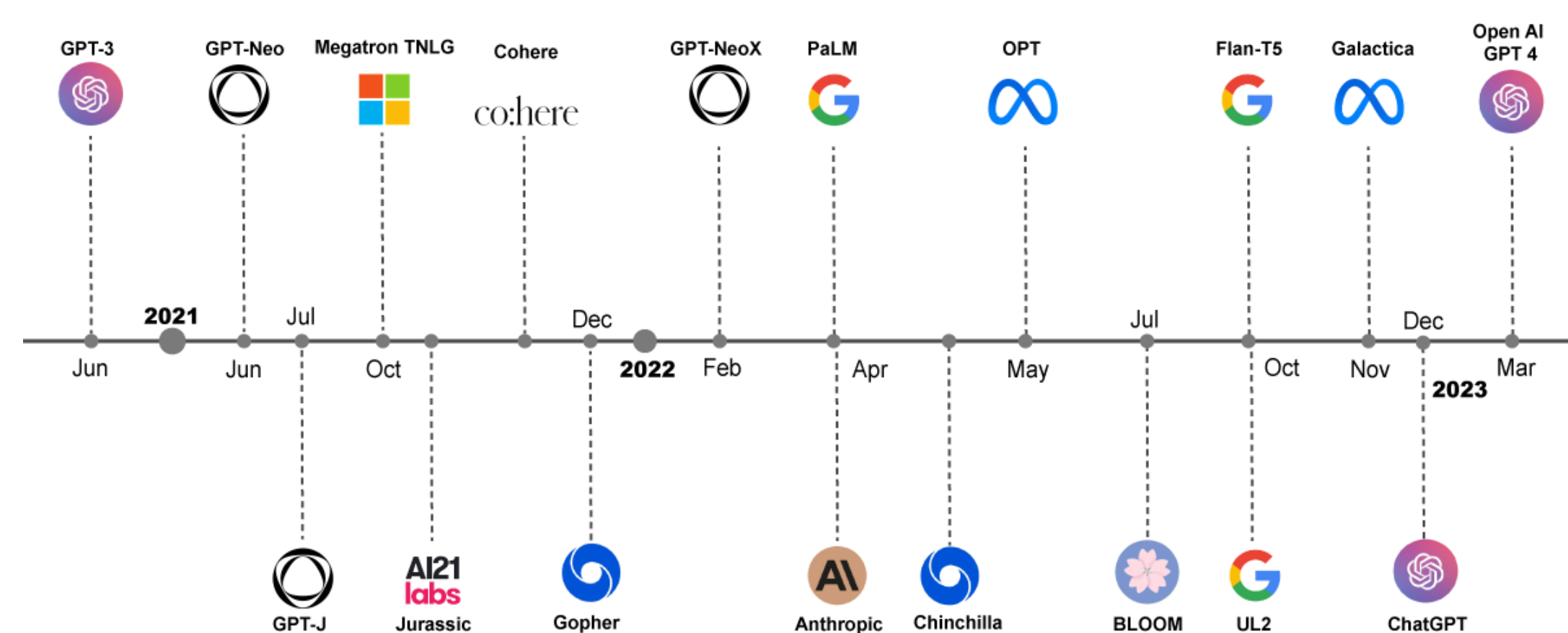
MRR for Online Delivery dataset



MRR for Student Performance dataset.

LLM's Upper Bound

Evolution of LLM



- ❖ LLMs show Massive improvement for language understanding and generation tasks.
- ❖ LLMs are still Unexplored in performing "ill-defined" tasks.
- ❖ Can LLMs help users of Data Science through conversations?

Future Works

- ❖ Begin with AB testing using ChatGPT to assess interaction quality, planning future exploration into the effects of diverse prompting strategies on user engagement.
- ❖ Transitioning to Open Source LLMs to enhance data privacy and control.
- ❖ Advancing conversational UIs with LLM on background to provide fluid result summaries and enable users to explore alternative options, enriching decision-making and user engagement.
- ❖ Continuous learning of user behavior and personalize the responses.