**# Overview**

Built a domain adapted space science information retrieval model that retrieve relevant context over thousands of documents and generate human like crisp and accurate answers

**# Domain adaptation overview**

By domain adaptation we mean the adaptation of an existing sentence transformer to new topics /domains.

Effective domain adaptation is incredibly helpful in taking models pretrained on large, existing datasets, and helping them understand a new domain that lacks any labeled datasets.

For example, any model trained on data from before 2019 will be unaware of COVID-19 and everything that comes with it

GPL hopes to solve this problem by allowing us to take existing models and adapt them to new domains using nothing more than unlabeled data. By using unlabeled data we greatly enhance the ease of finding relevant data, all we need is unstructured text.

**# Pipeline**

![A picture containing text, screenshot, diagram, design

Description automatically generated]()

**# Answer comparison**

Attached excel 'Model\_comparison' consists of examples comparing domain adapated model trained on space science data vs ChatGPT