

README

NLP Based Analysis of SweCris to understand the dynamically changing trends in Research Fund distribution.

IP Summer Sem 2023

Guide: Dr. N Arul Murugan

Students: Arha Samanta & Abhishek Acharya

IP- N Arul Murugan Sir.ipynb

1. This Jupyter Notebook contains our code that interacts with SweCris API and get a response. The response is stored as response.csv.
2. After that, we extracted the Title and Abstract and found the top words in trend related to diseases. We analysed the whole dataset and also in various year ranges, plotted the change in trends and a final graph showing all the words and their trends. We even did a separate analysis of diseases by analysing trends of diseases over the years, and extracted words before and after common diseases.
3. To run the Notebook in Google Colab, we need to connect the computer to internet and simply Run the Notebook, to get the response by interacting with the SweCris API. As an alternative approach, the response.csv file can be read directly and the rest code remains the same.
4. All the outputs and graphs will be generated one by one. As there are a lot of operations, it takes a few minutes to get all the outputs.