

System Setup and Program Running

- Next, we need to input and read the database
 - Make sure that [Python version \(3.6+\)](#) is installed
 - Install necessary libraries:
 - [Pandas](#)
 - [Numpy](#)
 - [Matplotlib](#)
 - [SentenceTransformer](#)
 - [Nltk](#)
 - [Sklearn](#)
 - [Pytorch](#)
 - [String](#)
 - [Cleanlab](#)
 - Import Seg_algorithm.py to use the segmentation algorithm.
- Ensure the program URL variable matches the selected dataset link.
 - Note: Saving dataset to GitHub and using GitHub link worked well for us.
- Run the main.py file using python
- File might take several minutes depending on how big the dataset is, and the system being used