

WEEK-3

Date:12-06-2021

OBJECTIVE

Write a program to extract and store key phrases along with associated metrics

ACTIVITY

- ***Task worked on and completed:***

- Identified Data Sources that is 100% related to the project , by manually checking one-by-one.
- Built a program to extract and parse the necessary information from the feeds
- Built a program to execute Key-Phrasing Extraction with the help of YAKE open source library under evaluation metric as TF*IDF (I.e. Term frequency x Inverse Document frequency).

- ***New Challenges:***

- Parsed extraction of feed items. Where extraction of context/summary/description was very difficult and output contained non-english characters.
- Decided to do Key-Phrasing on topic as the dataset was large enough. But still working on summary.
- Built a custom Key-Phrasing which worked only for one word(i.e if “machine learning” was given as input it would consider “machine” as different and “learning” as different).
- Finally ,decided use YAKE.

- ***Planning for next week:***

- Using YAKE for yearly and monthly topics.
- Evaluating different metric to give more acute value.
- Visualisation of the results.

Please go through the code (ipynb) file upload and provide FeedBack.