SounderFM

September 27, 2022   12:46 PM

Take home – highlight engaging clips
- Definition of engaging
  - Engaging topics
    - PySS3
  - Text sentiment
  - Interesting entities, trending in current news
  - Topic change detection :
    - Paper : https://arxiv.org/pdf/2106.12978.pdf
    - Link : https://github.com/gdamaskinos/unsupervised_topic_segmentation/blob/cefa8ba65f964220435d35a0cb39722790f9c0b4/core.py
    - TextTiling : https://github.com/gdamaskinos/unsupervised_topic_segmentation/blob/cefa8ba65f964220435d35a0cb39722790f9c0b4/core.py
  - https://www.diva-portal.org/smash/get/diva2:1576969/FULLTEXT01.pdf
  - PQ selection : https://github.com/tannerbohn/AutomaticPullQuoteSelection
  - Awesome sentiment : https://github.com/laugustyniak/awesome-sentiment-analysis/blob/master/README.md
  - READLESS : https://github.com/AdiChat/senpai
  - https://github.com/Ighina/DeepTiling
  - https://github.com/riedlma/topictiling
  - https://github.com/chschock/textsplit
  - Audio sentiment:
    - https://github.com/audeering/w2v2-how-to

POST /analyze
- Accepts transcript or audio file
  Binary or text file
- Triggers analysis
- Responds with ID

Analysis Class
- Interacts with API

Test
- Tests for API level
- Analysis class
  - Test interface
- Text analysis
- Audio analysis

POST /highlights/{id}
- Returns a JSON
  - Begin time
  - Endtime

Analysis Text

- Topic segmentation
  - BERT based or TextTiling based
  - TextTiling looks for paragraph breaks
- Feature extraction
- Text classification
- Sentiment Analysis
- Text emotion recognition
  - https://huggingface.co/kiri-ai/t5-base-qa-sum
  - https://pypi.org/project/text2emotion/
  - VAD analysis of text : https://github.com/bag
    - https://github.com/bagustris/text-vad,

- Return top candidate segments
  - Maintain an ordered candidate list of senten
    - Sentence ID would be needed

Analysis Audio
- Audio feature extraction
- Audio sentiment
- Audio highlights

(Can use to re-rank segments)