Automated Video Indexing

Arpan Mukherjee [MT17007]
Kunal Suryavanshi [MT17023]
Shubhi Tiwari [MT17057]

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

OUTLINE

- Problem Definition
- Dataset Collection
- Data Preprocessing
- Model
- Evaluation
- Result

Problem Definition

- Video indexing is the process of providing viewers a way to access and navigate contents easily, similar to book indexing.

- The selection of indexes derived from the content of the video to help organize video data and metadata that represents the original video stream.

Dataset

- Different playlist follows different pattern
- Decided to stick to single genre of the playlist.
- Coursera Course Video
  - Big Data
- YouTube NPTEL Video Playlists for the following courses:
  - Machine Learning
  - Artificial Intelligence
  - Database Management System
- Introduction to Sorting (https://www.youtube.com/watch?v=kMgIlOpq_Dg)
- Video transcript of the playlist was manually scrapped

Data Preprocessing

- Used python module pysrt for processing of subtitles.
- Corrected the timing information of combined videos(5 videos combined in one) so that it can be treated as a single video srt file
- Data lemmatization, stop word removal etc was done
- POS tagging for extracting the noun words
- These nouns later used to identify topics covered in the video titles.

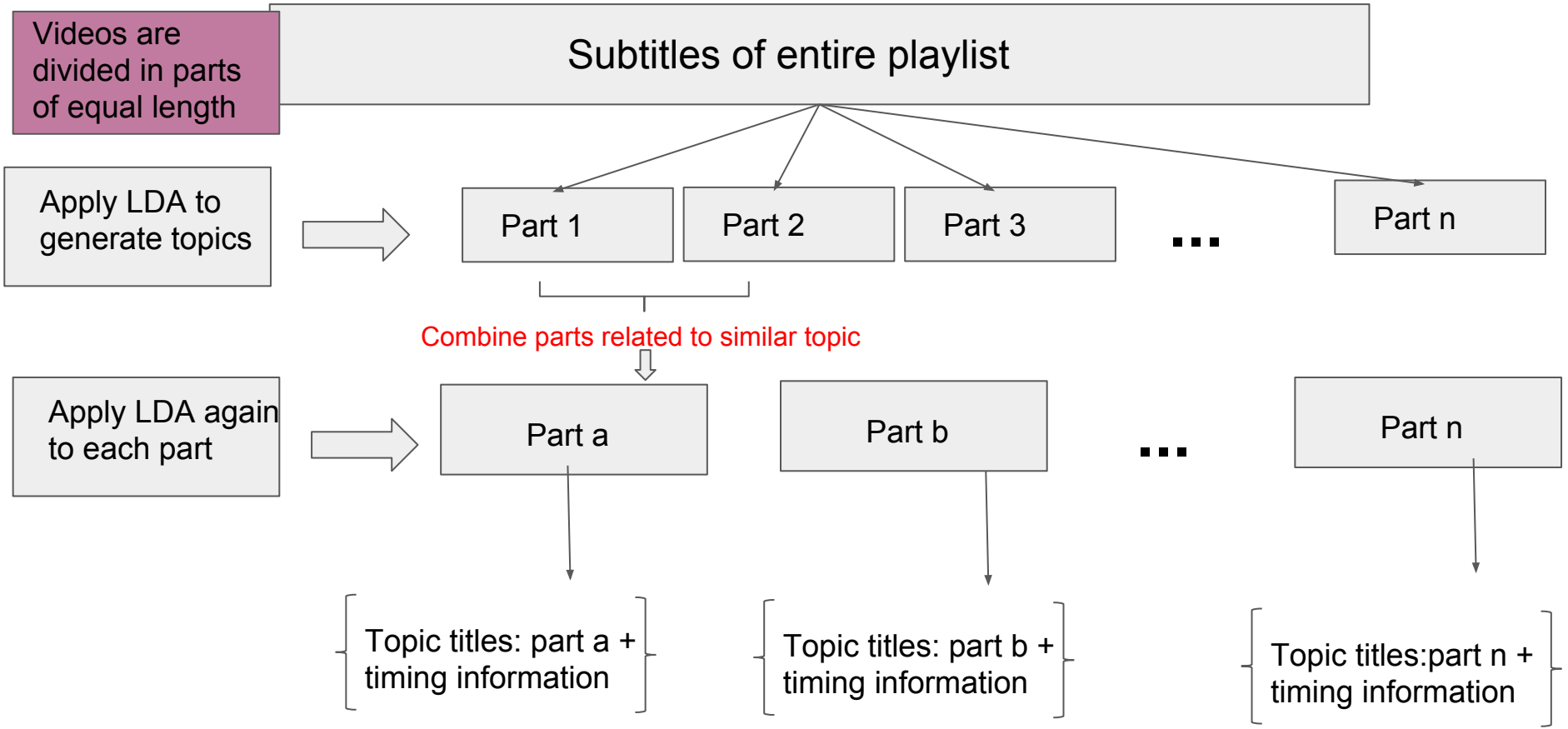
Proposed Methodologies

- Introductory video inference
  - We tried to infer the introductory video of the NPTEL playlist to find the topics which will be covered during the course.
  - Along with keywords, decided to use video name and description as feature
  - Dropped this idea because not all NPTEL playlist contains mandatory introductory and not all video contains description
- Template based Methods
  - Tried to create different templates for single genre of videos
  - Handcrafted rules to identify video initialization, e.g. - “Let us start with...”, “Today we will be talking about...” etc.
  - We could not infer handcrafted rules, as there was no symmetricity between the playlists.

Proposed Methodologies contd.

- Tree-based Topic Modeling Approach(Novel Approach)
  - Divided merged subtitles in the window size of 50 subtitles
  - Extracted keywords using LDA and NMF for the fixed window.
  - In bottom-up manner we merged two windows of topics and their timings, if the jaccard similarity between the two is below a given threshold
  - Another iteration of LDA/NMF was applied to get the final top 10 words describing each merged topic.

Topic extraction from subtitles :Tree based LDA Model

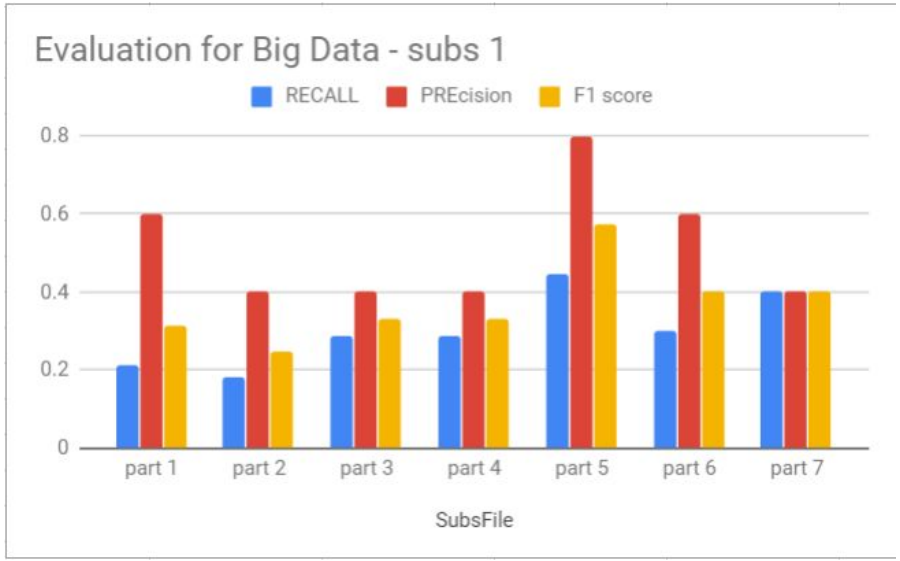


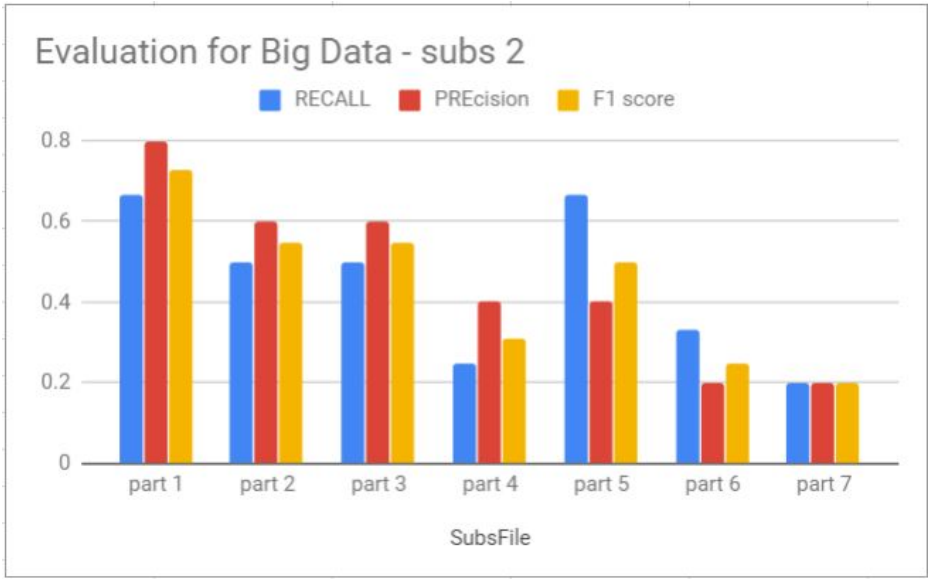
Evaluation:

- For videos on Big Data from coursera.org, video titles serve as ground truth
- For other videos, a human judge indexes the videos and the we compare
- Precision, Recall , F1- score

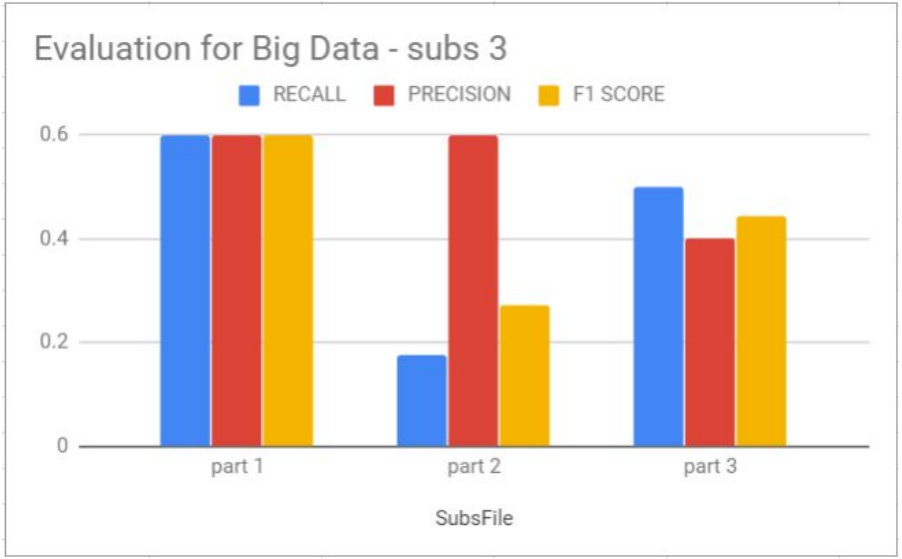
Ground Truth		Predicted (indexer output)	
Timing	**Topics**	**Timing**	**Topics**
0:0:0 - 0:2:21	What, why,sorting, arrangement, algorithm	0:0:0 - 0:2:49	Apply, applications , data, required, sorting
0:2:22 - 0:3:45	Sorting, applications,examples	0:2:49 - 0:5:32	element', 'largest', 'array', 'unsorted', 'sorting
0:3:46 - 0:11:16	Selection, sort, algorithm, array, example,code	0:5:32 - 0:8:32	okay', 'element', 'portion', 'array', 'unsorted'
		0:8:32,0:11:35	element', '17', 'portion', 'array', 'unsorted'
0:11:17 - 0:21:16	Bubble, sort, algorithm, array, example,code	0:11:35,0:14:21	'element', 'portion', 'array', 'unsorted', 'end'
		0:14:23,0:17:32	compare', 'inner', 'greater', '10', '14'
		0:17:32,0:20:20	'element', 'bubble', 'portion', 'array', 'unsorted'}
0:21:17 - 0:33:22	Quick, sort, algorithm, array, example,pseudo , code	0:20:23 -0:23:35	sort', 'quicksort', 'smaller', 'algorithm', 'problem'
		0:23:35 - 0:26:29	left', 'solution', 'solve', 'array', 'problem'
		0:26:29,0:29:50	partition', 'quicksort', 'method', 'portion', 'value'

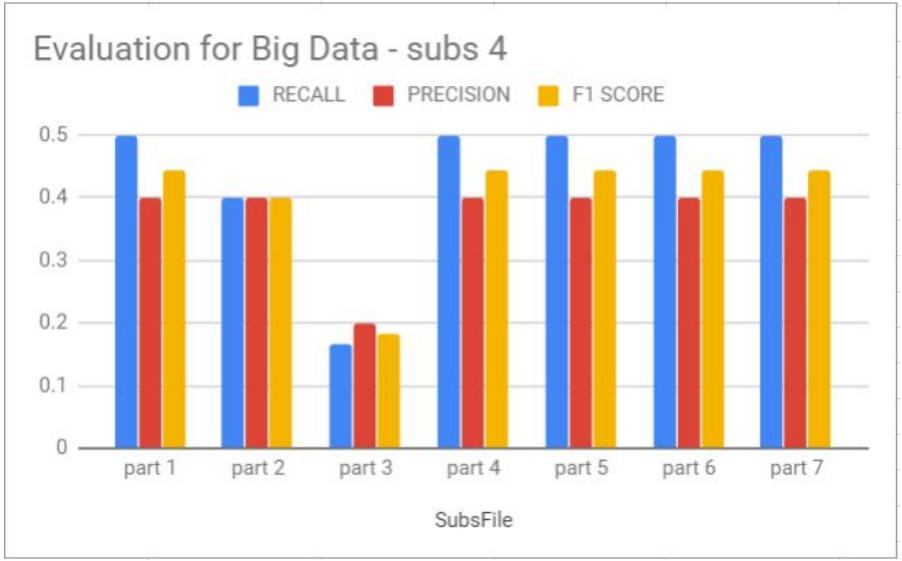
Results:


Evaluation for Big Data - subs 1
RECALL
PREcision
F1 score
0.8
0.6
0.4
0.2
0
part 1
part 2
part 3
part 4
part 5
part 6
part 7
SubsFile


Evaluation for Big Data - subs 2
RECALL
PREcision
F1 score
0.8
0.6
0.4
0.2
0
part 1
part 2
part 3
part 4
part 5
part 6
part 7
SubsFile

Results:


Evaluation for Big Data - subs 3
RECALL
PRECISION
F1 SCORE
0.6
0.4
0.2
0
part 1
part 2
part 3
SubsFile


Evaluation for Big Data - subs 4
RECALL
PRECISION
F1 SCORE
0.5
0.4
0.3
0.2
0.1
0
part 1
part 2
part 3
part 4
part 5
part 6
part 7
SubsFile

Conclusion and Future Work

- We proposed a novel approach for indexing videos, considering no work has been done on this domain.
- Ground Truth based accuracy , precision based on manually short indexed videos
- Human similarity score is suggested, as even though predicted results are somewhat similar, but not exact words are the output of the model.(e.g.- medicine, health, life equivalent to patient)
- Keywords representing each segment can be reconstructed in such a way, that it will summarize the segment better.