



Text Summarization of Meeting Dialogues

Erion Çano

UFAL - MFF - UK



Definition and Applications of SA

Sentiment Analysis

Computational examination of sentiments, opinions, and attitudes expressed in text from an opinion holder towards an entity.

Sentiment Classification

Determining the polarity of an opinion in a text unit about an entity. It can be document-level, sentence-level or aspect-level.

Applications

Market surveys and predicitions, brand/product popularity analysis, client/product profiling, political surveys, counter-terrorism, etc.

Word Representations: Bag of Words

I love this movie! It's sweet, but with satirical humor. The dialogue is great and the adventure scenes are fun... It manages to be whimsical and romantic while laughing at the conventions of the fairy tale genre. I would recommend it to just about anyone. I've seen it several times, and I'm always happy to see it again whenever I have a friend who hasn't seen it yet!



it	6
I	5
the	4
to	3
and	3
seen	2
yet	1
would	1
whimsical	1
times	1
sweet	1
satirical	1
adventure	1
genre	1
fairy	1
humor	1
have	1
great	1
...	...

Severe sparsity and dimensionality issues on large V ... 😞

Word Representations: Word Embeddings

“You shall know a word by the company it keeps.”

– J. R. Firth, 1957


Word Embeddings

- 😊 Dense and low-dimensional
- 😊 Complexity scales linearly w.t.r V
- 😊 Preserve word order in phrases
- 😊 Capture semantic and syntactic similarities
- 😞 Require big text corpora to train
- 😞 Computationally expensive to train

Representation of Words

Matrix representaton of “your shirt looks nice”:

$d = 5$



your	0.23	0.18	0.34	0.76	0.62
shirt	0.64	0.23	0.21	0.03	0.83
looks	0.98	0.59	0.76	0.65	0.45
nice	0.11	0.43	0.30	0.22	0.92

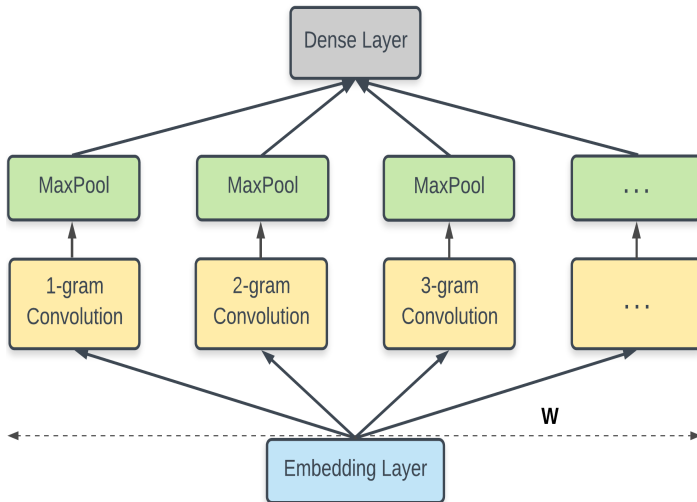
Embeddings sourced from pretrained GoogleNews collection.

Experimental Datasets

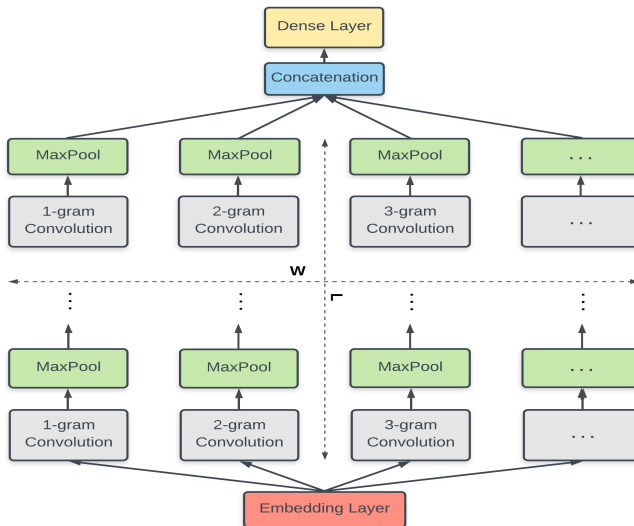
Dataset	Docs	MinL	AvgL	MaxL	UsedL
Mlpn (song lyrics)	5K	23	227	2733	450
Sent (Sentences)	10K	1	17	46	30
Imdb (movie reviews)	50K	5	204	2174	400
Phon (phone reviews)	232K	3	47	4607	100
Yelp (yelp reviews)	598K	1	122	963	270

- Different domain tasks and data types
- Both small (Mlpn) and big (Yelp) datasets
- Both long (Imdb) and short (Phon) documents

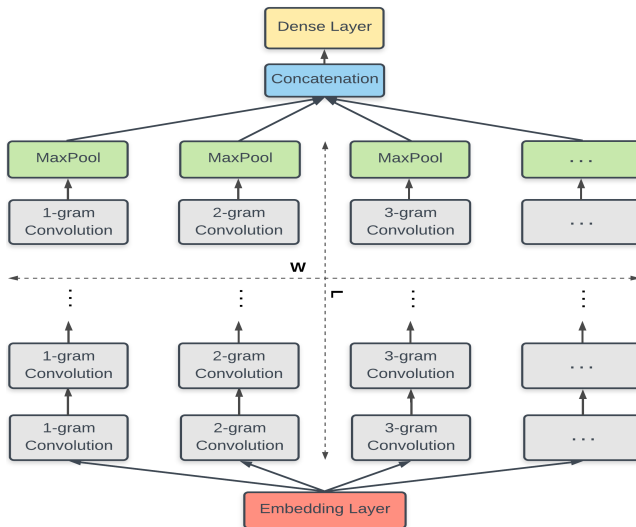
Multi-Channel Network Structures



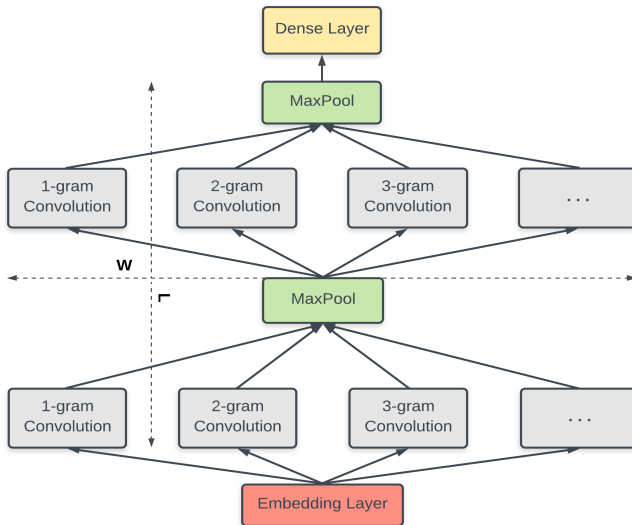
NgramCNN Basic Architecture



NgramCNN Pyramid Architecture



NgramCNN Fluctuating Architecture



Baseline Models

- Single LSTM
- Single Convolution-Pooling
- Bidirectional LSTM with max-pooling
- Bidirectional LSTM with Convolution-Pooling
- Logistic Regression with tf-idf
- Support Vector Machine with tf-idf

Comparative Accuracy Scores

Network	Sent	Imdb	Phon	Yelp
NgCNN Basic	79.87	90.77	<u>95.92</u>	<u>94.88</u>
NgCNN Pyramid	79.52	<u>91.21</u>	95.70	94.83
NgCNN Fluctuate	77.41	89.32	93.45	92.27
Optimized LR	81.63	89.48	92.46	91.75
Optimized SVM	82.06	88.53	92.67	92.36
SingleCNN	81.79	89.84	94.25	93.86
SingleLSTM	80.33	84.93	93.71	90.22
BLSTM-POOL	80.96	85.54	94.33	91.19
BLSTM-2DCNN	<u>82.32</u>	85.70	95.52	91.48

Definition of Text Summarization

Text Summarization (TS)

Distilling the most important information in a text to produce an abridged version.

Types of TS

- Single-document vs Multi-document
- Extractive vs Abstractive
- Generic vs Query-driven
- Informative vs Indicative

Why to Summarize...?

- Simplify and abbreviate text (Abstracts)
- Summary of email threads (Subjects)
- Action or decisions from a meeting (Discussions)
- Generating news about an event (Stories)
- General opinions about an item (Reviews)
- Answering user questions (Queries)

Extractive vs Abstractive

Extractive TS

The generated summary is a selection of relevant sentences from the source text in a copy-paste fashion.

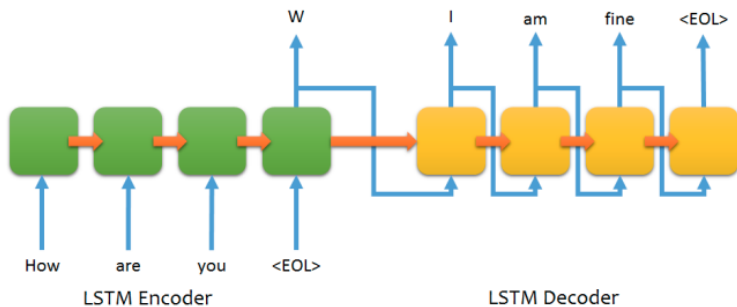
- Simpler and highly explored
- Statistical, Feature-based, Machine Learning, Graph-based

Abstractive TS

The generated summary is a new cohesive text not necessarily present in the original source.

- Hard and challenging
- TS as a neural MT problem; encoder-decoder paradigm

Abstractive TS Problems



Problems

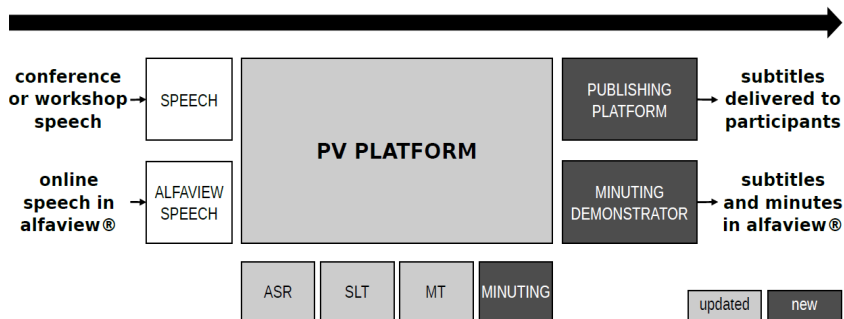
- Generated summary not always meaningful
- Hard to distinguish rare and unknown words
- Grammar errors in the generated summaries

Remote conferencing/meeting



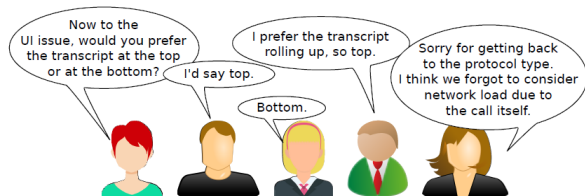
- Important for reducing business expenses
- Market growth: from 5 % (2015) to 60 % (2021)
- Leaders: Cisco **WebEx**, Microsoft **Skype** and **Zoom**

Proposed Workflow of ELITER Project



- Automatic Speech Recognition
- Spoken Language Translation
- Machine Translation

Meeting Dialogues



Original agenda as prepared by the organizer beforehand:

- Protocol type: push or pull?
- Layout of the user interface:
 - Transcript grows at the top or bottom of the document?
 - Or in a side pane?

Shared document, everyone allowed to edit.

Starts with the agenda and gets populated by Automatic Minuting (AM)

- Protocol type: push or pull?
 - (AM) > Pull easier to implement.
 - (AM) > Updates can get lost with push *in case the user*
 - (AM) > Consider network load.
- Layout of the user interface:
 - Transcript grows at the top or bottom of the document?
 - (AM) > Top (AM) > Bottom (AM) > Top, transcript rolling up.
 - Or in a side pane?

Transcript, optionally editable to correct ASR errors:

- 11:03 Sorry for getting back to the protocol type. I think we forgot ...
- 11:02 I prefer the transcript rolling up, so top.
- 11:02 Bottom
- ...

S1: Segmenting Dialogues

- Dialog segments instead of text sentences
- Breaking transcript into segments
- Order and classify each segment
- Using neural networks on labeled texts
- Exploring unsupervised or semi-supervised methods

S2: Summarizing Sentences/Segments

- Speech has disfluencies and redundancy
- Remove redundant words keeping important content
- Extractive supervised: Deep RNNs
- Extractive unsupervised: TextRank
- Abstractive Supervised: Bidirectional LSTMs

S3: Summarizing Documents/Discussions

- Rank segments based on position and topic
- Sentence embedding for merging redundant segments
- Trying Encoder-decoder or seq2seq networks
- Identifying speaker's attitude on the topic
- Infusing external knowledge to eliminate grammatical errors

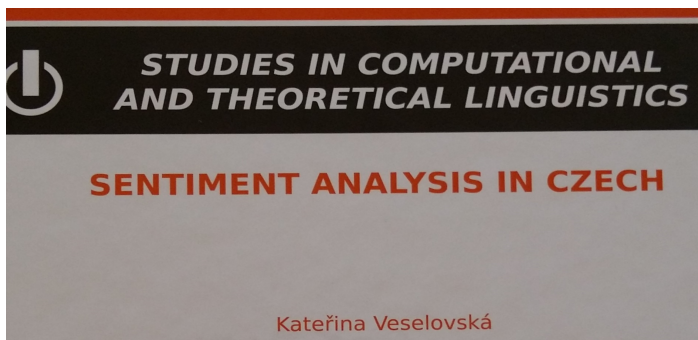
S4: Fitting Segments into Agenda

- Matching segments with agenda points
- Static or dynamic agenda...?
- Computing similarity of segments with agenda points
- Using embeddings of words and segments

Available Datasets

- ICSI meeting corpus (Janin *et al.*, 2003)
- AMI meeting corpus (Carletta, J., 2006)
- Gigaword dataset (Graff *et al.*, 2003)
- Google sentence compression (Filippova and Altun, 2013)
- CNN Daily Mail corpus (Hermann *et al.*, 2015)

Sentiment Analysis: A new context



- 😊 Rigorous dataset creation and assessment
- 😊 Careful linguistic feature processing
- 😞 Old and maybe not optimal classifiers

Sentiment Analysis: Ideas to explore

- Comparing many existing off-the-shelf classifiers
- Trying ensemble learners
- Word embeddings on bigrams
- Neural networks with convolutions

Questions or Suggestions...?



Thank You...😊