Neural Networks for Sentiment Analysis in Czech

Ladislav Lenc[1,2] & Tomáš Hercig[1,2]

[1]Dept. of Computer Science & Engineering
University of West Bohemia
Plzeň, Czech Republic

[2]NTIS - New Technologies for the Information Society
University of West Bohemia
Plzeň, Czech Republic

llenc,tigi@kiv.zcu.cz

September 18, 2016

Table of Contents

Introduction

- Sentiment analysis - determining sentence polarity
- Aspect-based sentiment analysis (ABSA)
  - Identify aspects of a given target entity
  - Determine sentiment polarity for each aspect
- Focus on polarity detection on various levels - texts, sentences, aspects
- First attempts on Czech using neural networks
- Comparison of results on English

Examples

- **Aspect Term Extraction (TE)** – identify aspect terms.

  `Our` **server** `checked on us maybe twice during the entire` **meal**`.` → {`server, meal`}

- **Aspect Term Polarity (TP)** – determine the polarity of each aspect term.

  `Our server checked on us maybe twice during the entire meal.` → {`server: negative, meal: neutral`}

- **Aspect Category Extraction (CE)** – identify (predefined) aspect categories.

  `Our server checked on us maybe twice during the entire meal.` → {`service`}

- **Aspect Category Polarity (CP)** – determine the polarity of each (pre-identified) aspect category.

  `Our server checked on us maybe twice during the entire meal.` → {`service: negative`}

Examples

The later SemEval's ABSA tasks (2015 and 2016) further distinguish between more detailed aspect categories and associate aspect terms (targets) with aspect categories.

- **1) Aspect Category Detection** – identify (predefined) aspect category – entity and attribute (E#A) pair.

  `The pizza is yummy and I like the atmoshpere.`
  `→ {FOOD#QUALITY, AMBIENCE#GENERAL}`
- **2) Opinion Target Expression (OTE)** – extract the OTE referring to the reviewed entity (aspect category).

  The **pizza** is yummy and I like the **atmoshpere**.
  → {pizza, atmoshpere}
- **3) Sentiment Polarity** – assign polarity (positive, negative, and neutral) to each identified E#A, OTE tuple.

  The pizza is yummy and I like the atmoshpere.
  → {FOOD#QUALITY - pizza: positive,
  AMBIENCE#GENERAL - atmoshpere: positive}

Data Preprocessing and Representation

Preprocessing

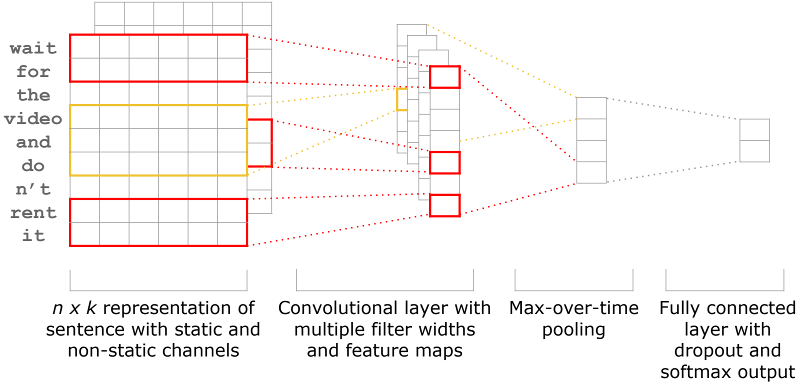
- Noisy data from user reviews - need of preprocessing
- Removing accents
- Converting to lower case
- Replacing numbers with one token
- Stemming (tested with and without stemming)

Data Representation

- One-hot encoding
- Sentence representation - sequence of indexes from dictionary
- Fixed length of sentences - cutting / padding longer / shorter ones
- 50 words for document level, 11 for aspect level
- Dictionary - 20,000 words

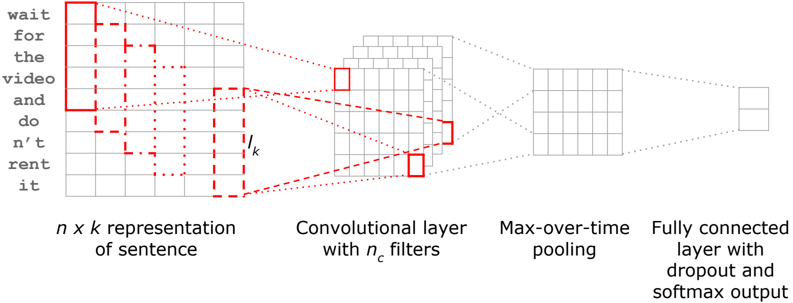
Convolutional Network 1

- Inspired by Kim [1]
- Three filter widths in the convolutional layer



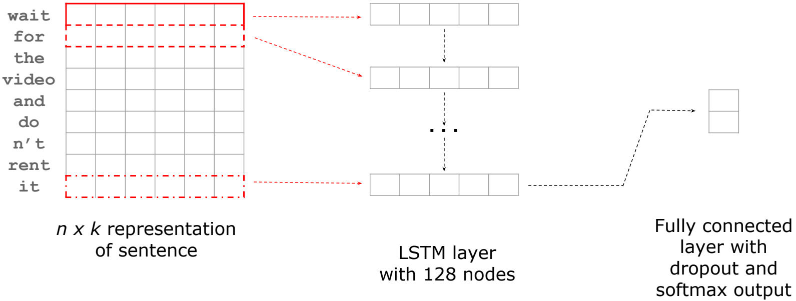
Convolutional Network 2

- Inspired by architecture used for document classification [2]



LSTM

- Basic LSTM architecture



Datasets

Table 1 : Properties of the aspect-level and document-level corpora in terms of the number of *sentences*, *average length* of sentences (number of words), and numbers of *positive*, *negative*, *neutral* and *bipolar* labels.

Aspect-level Sentiment Dataset	Sentences	Avg	Positive	Negative	Neutral	
English 2016 Laptops train + test	3.3k	14	2.1k	1.4k	0.2k	
English 2016 Restaurants train + test	2.7k	13	2.3k	1k	0.1k	
English 2015 Restaurants train + test	2k	13	1.7k	0.7k	0.1k	
Czech Restaurant reviews	2.15k	14	2.6k	2.5k	1.2k	
Czech IT product reviews short	2k	6	1k	1k	–	
Czech IT product reviews long	0.2k	144	0.1k	0.1k	–	
Document-level Sentiment Dataset	**Sentences**	**Avg**	**Positive**	**Negative**	**Neutral**	**Bipolar**
English RT Movie reviews	10.7k	21	5.3k	5.3k	–	–
Czech CSFD Movie reviews	91.4k	51	30.9k	29.7k	30.8k	–
Czech MALL Product reviews	145.3k	19	103k	10.4k	31.9k	–
Czech Facebook posts	10k	11	2.6k	2k	5.2k	0.2k

Model Verification

Table 2 : Accuracy on the English RT movie reviews dataset in %.

Description	Results
Kim [1] randomly initialized	76.1
Kim [1] best result	**81.5**
CNN1	77.1
CNN2	76.2
LSTM	61.7
Confidence Interval	± 0.8

Table 3 : Accuracy on the English SemEval 2016 ABSA datasets in %.

Description	Restaurants	Laptops
SemEval 2016 best result	**88**	**82**
SemEval 2016 best constrained	**88**	75
CNN1	78	68
CNN2	78	71
LSTM	72	68
Confidence Interval	± 3	± 3

Results

Table 4 : F-measure on the Czech document-level datasets in %.

Description	CSFD Movies	MALL Products	Facebook Posts
Supervised Machine Learning [3]	78.5	75.3	**69.4**
Semantic Spaces [4]	**80**	**78**	-
Global Target Context [5]	**81.5**	-	-
CNN1 stemmed	70.8	74.4	**68.9**
CNN2 stemmed	71.0	75.5	**69.4**
LSTM stemmed	70.2	73.5	67.6
Confidence Interval	± 0.3	± 0.2	± 1.0

Results

Table 5 : Accuracy on the Czech aspect-level restaurant reviews dataset in %. *W* denotes words, *S* stemms and *W+S* the combination of these inputs.

	Term Polarity			Class Polarity		
Description \ Features	W	S	W+S	W	S	W+S
CNN1	65	66	**67**	65	66	68
CNN2	64	65	66	67	68	**69**
LSTM	61	62	62	65	65	64
Confidence Interval	±2	±2	±2	±2	±2	±2

State-of-the-art results 72.5% TP and 75.2% CP [6].

Conclusions / Future Work

- Experiments
  - Two English corpora to confirm comparability with existing work
  - Three Czech corpora for document-level SA
  - One Czech corpus for ABSA
  - → First attempt with basic features, not fine-tuned
  - → The tested networks don't achieve as good results as the state-of-the-art approaches.
  - → The most promising results were obtained when using the CNN2 architecture
  - → Czech is much more complicated than English in terms of SA (e.g. double negative, sentence length, comparative and superlative adjectives, or free word order)
- Future work
  - Error analysis, word embeddings layer initialization, experiment with automatic translation of Czech into English, explore aspect term extraction and aspect category extraction, and new neural network architectures for sentiment analysis

Yoon Kim,
"Convolutional neural networks for sentence classification,"
arXiv preprint arXiv:1408.5882, 2014.

L. Lenc and P. Král,
"Deep neural networks for Czech multi-label document classification,"
in *International Conference on Intelligent Text Processing and Computational Linguistics*, Konya, Turkey, April 3 - 9 2016.

Ivan Habernal, Tomáš Ptáček, and Josef Steinberger,
"Sentiment analysis in Czech social media using supervised machine learning,"
in *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Atlanta, GA, USA, June 2013, pp. 65–74, Association for Computational Linguistics.

Ivan Habernal and Tomáš Brychcín,
"Semantic spaces for sentiment analysis,"
in *Text, Speech and Dialogue*, Berlin, 2013, vol. 8082 of *Lecture Notes in Computer Science*, pp. 482–489, Springer-Verlag.

Tomáš Brychcín and Ivan Habernal,
"Unsupervised improving of sentiment analysis using global target context,"
in *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, Shoumen, Bulgaria, September 2013, pp. 122–128, INCOMA Ltd.

Tomáš Hercig, Tomáš Brychcín, Lukáš Svoboda, Michal Konkol, and Josef Steinberger,
"Unsupervised methods to improve aspect-based sentiment analysis in Czech,"
Computación y Sistemas, in press.