Sentiment Analysis
State of the Art in Research and Industry

Mark Cieliebak
Zurich University of Applied Sciences

16.9.2016 – SDS 2016 – Zurich

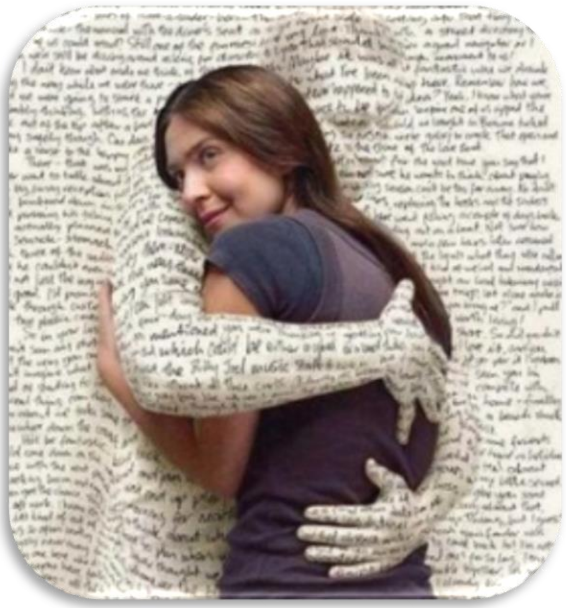
Mark Cieliebak

+ PhD in Theoretical Computer Science
+ IT Consultant in Major Swiss Bank
+ CIO at Netbreeze (bought by Microsoft)
+ >30 Publications

Lecturer	CEO	Conference Chair

Zurich University of Applied Sciences

Sentiment Analysis

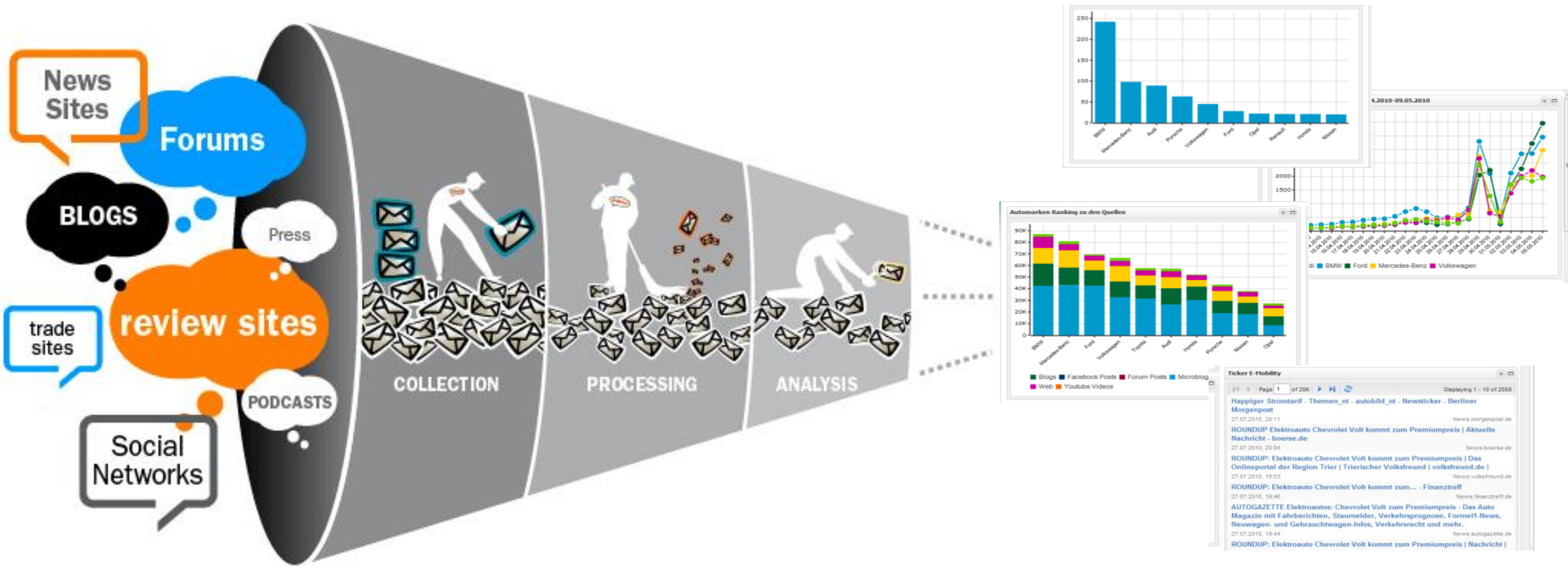


Goal: Decide whether a text expresses positive or negative emotion.

" This is a nice conference! "

Insights for Marketing and Sales

Sentiment Analysis can identify trends in Social Media



Characteristics of Sentiment Analysis

Labels:

- Positive
- Negative
- Neutral

- Mixed
- (unknown)

Tasks:

- Single sentence

- Complete document

- Specific aspect/target

- Quantification

Sentiment-Analysis sounds easy

…but it isn't

@francesco_con40 2nd worst QB. DEFINITELY Tony Romo. The man who likes to share the ball with everyone. Including the other team

@prodnose is this one of your little jokes like Elvis playing at the Marquee next Tuesday?

Tim Tebow may be availible ! Wow Jerry , what the heck you waiting for ! http://t.co/a7z9FBL4

#YouCantDateMe if u still sag ur pants super hard...dat shit is played the fuck out!!!

A Remark about Tool Quality

"They all suck...and we suck, too."

CEO of a sentiment analysis company (2013)

Evaluation of Commercial Sentiment Analysis Tools in 2013

Zurich University of Applied Sciences

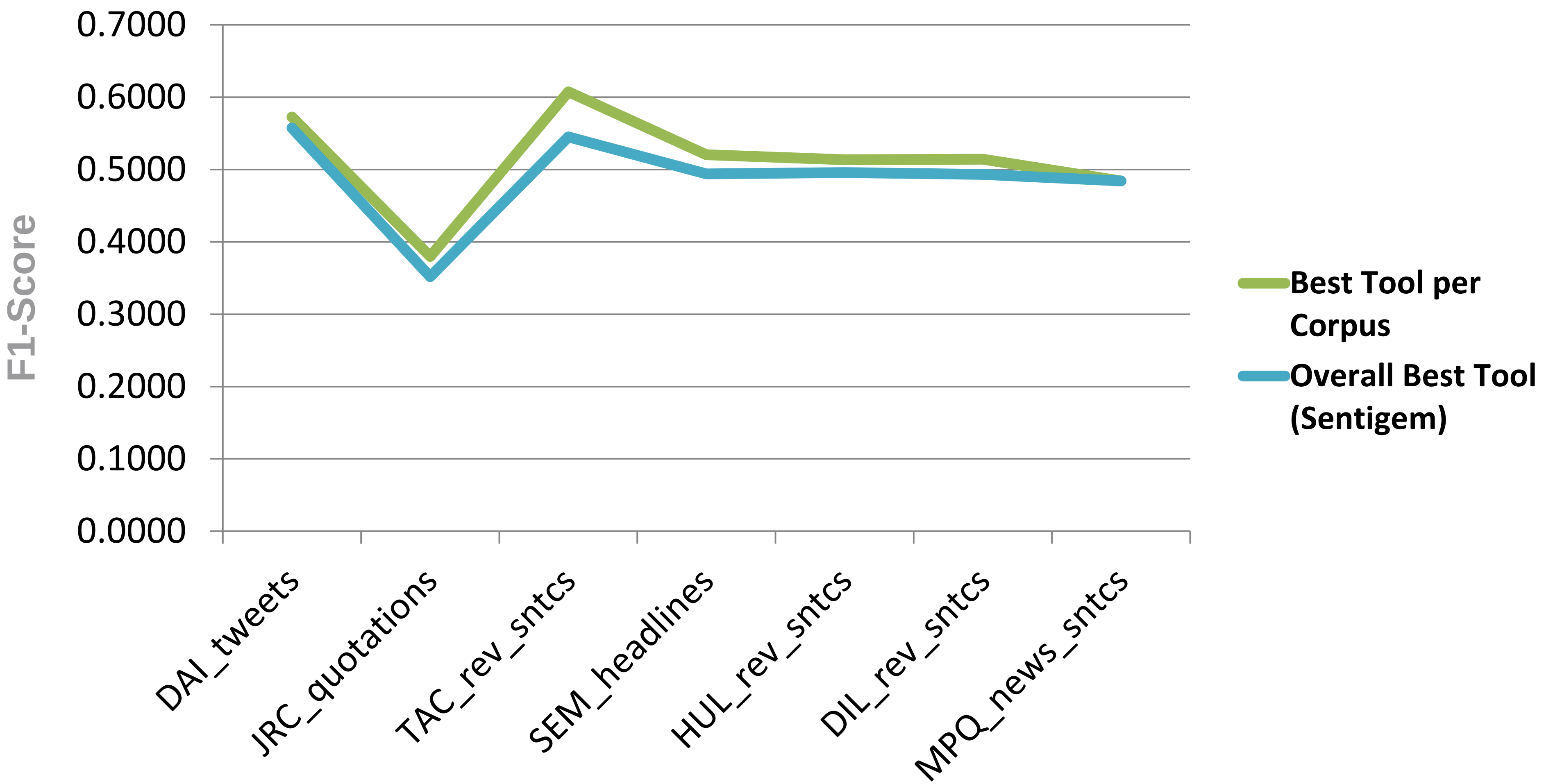
7 Text Corpora

- Single statements
- Various media types (tweet, news, reviews, speech transcripts etc.)
- Total: 28'653 texts

9 Commercial APIs

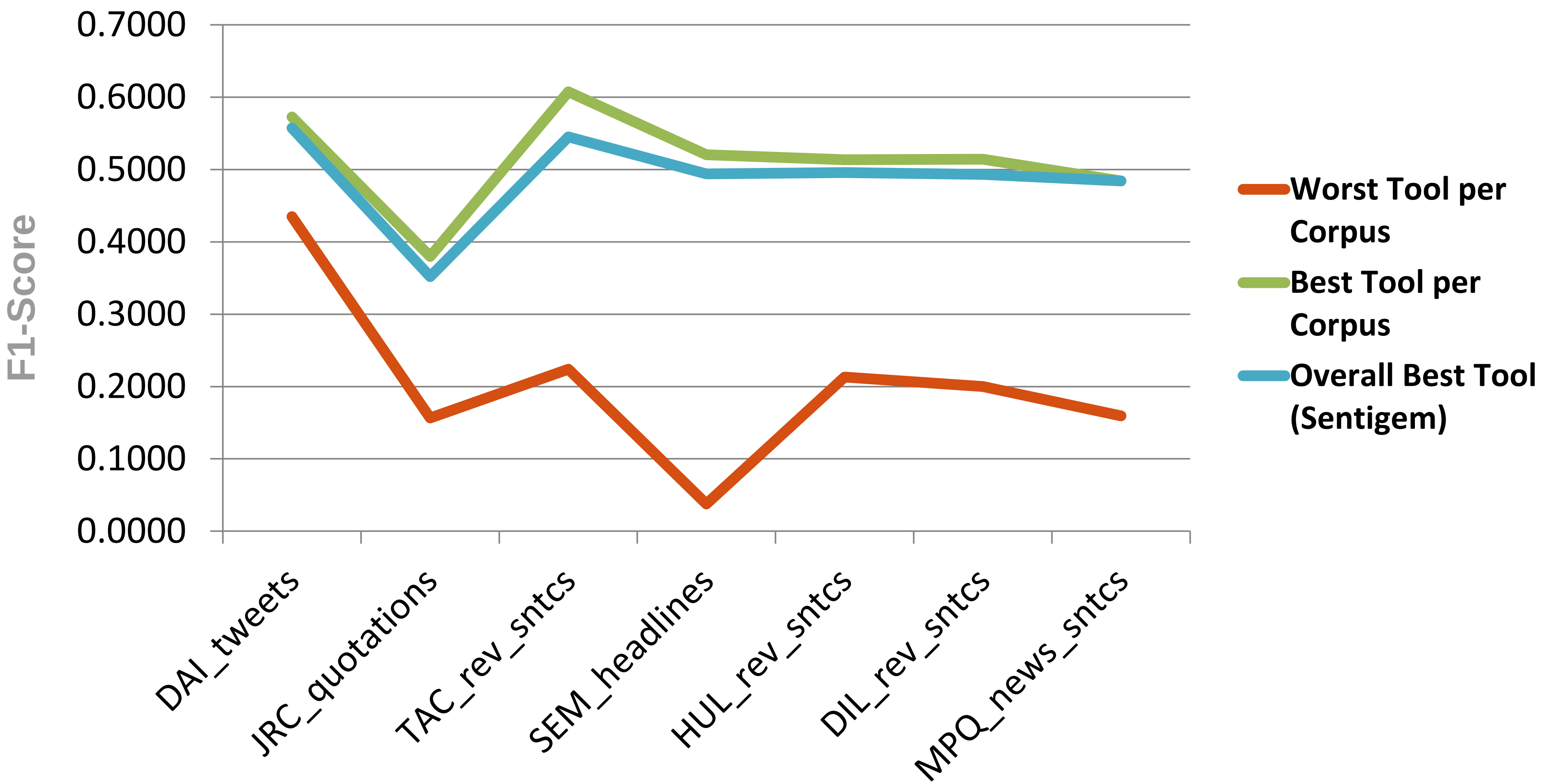
- Stand-alone
- Free for this evaluation
- English text

Quality of Commercial Tools in 2013



Source: M. Cieliebak et al.: Potential and Limitations of Commercial Sentiment Detection Tools, ESSEM 2013.

Quality of Commercial Tools in 2013



Source: M. Cieliebak et al.: Potential and Limitations of Commercial Sentiment Detection Tools, ESSEM 2013.

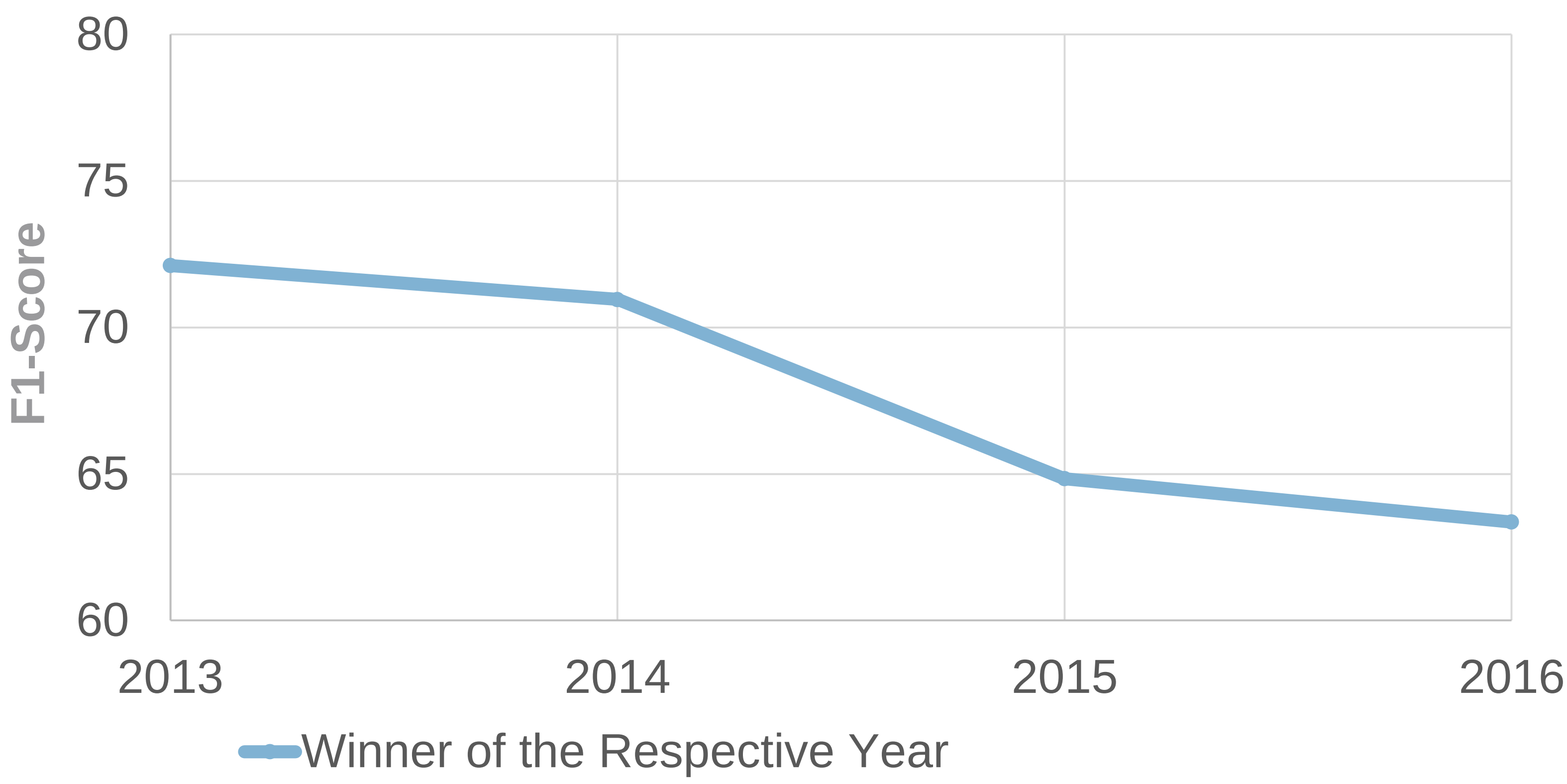
SemEval: International Competition for Sentiment Analysis

Zurich University of Applied Sciences

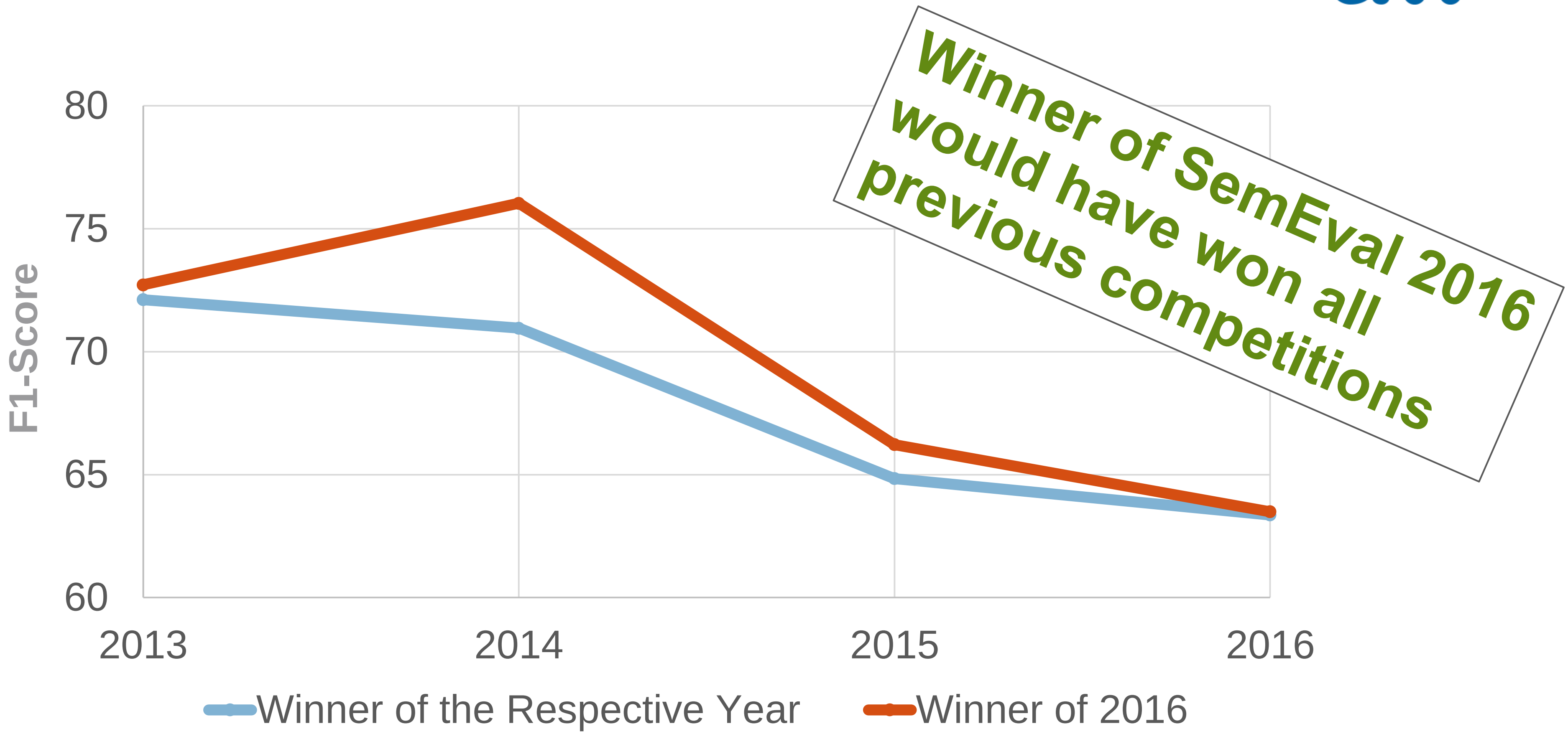
Task: Build a system for sentiment analysis (pos, neg, neutral) on tweets in English

Year	Winning Team	F1-Score	Winning Technology	Remarks
2013	NRC Canada	**69.02**	Features + large dictionaries	First run of the competition
2014	TeamX	**72.12**	Similar approach as in 2013	First two participants using deep learning
2015	Webis	**64.84**	Ensemble of 4 approaches from previous years	
2016	SwissCheese	**63.30**	CNN+Distant Supervision	30'000 new tweets Dominance of deep learning among submissions

Did Sentiment Technology Improve?



Did Sentiment Technology Improve?

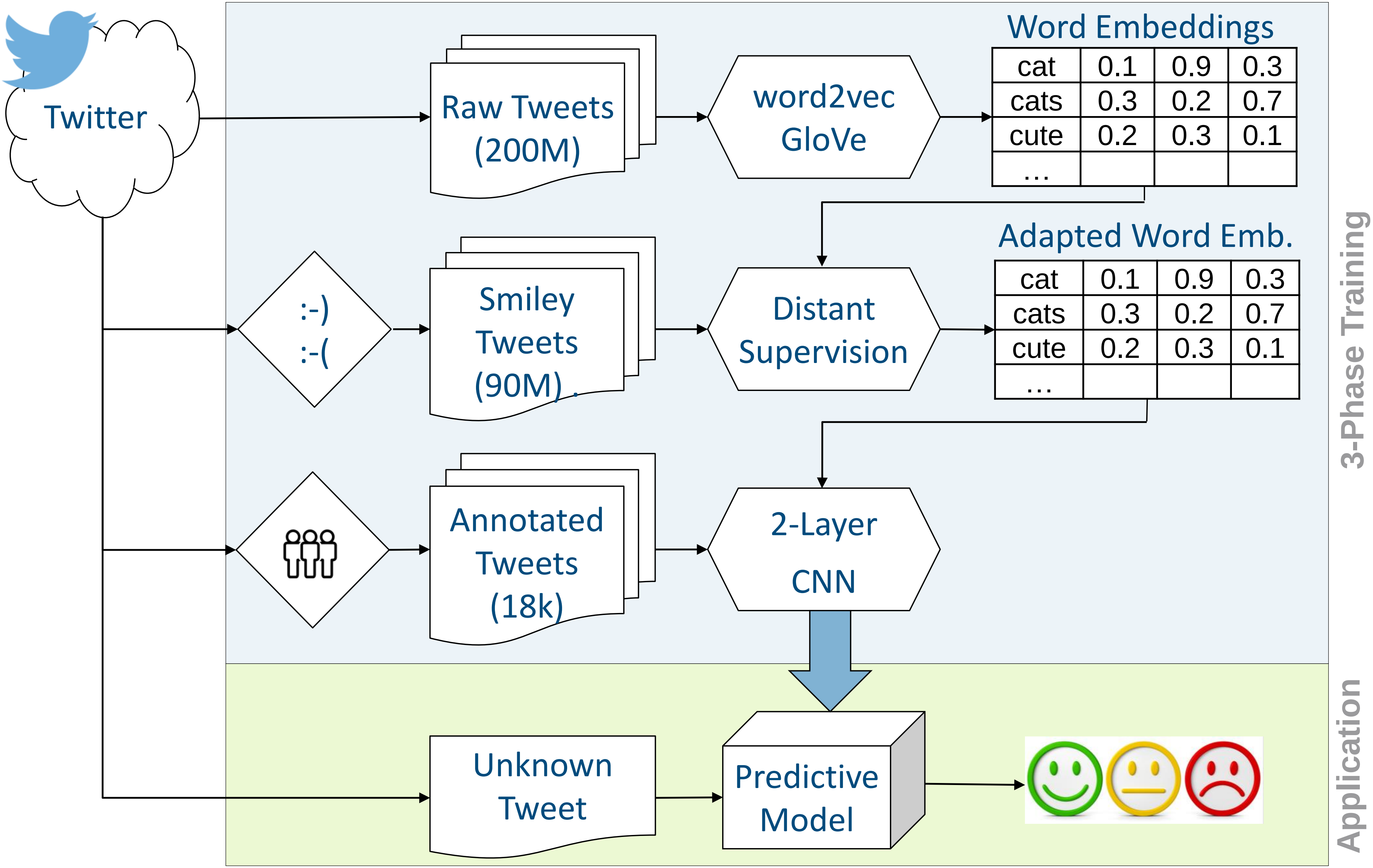


Red line: performance of SemEval winner from 2016 (SwissCheese)
if only trained on training data for each year

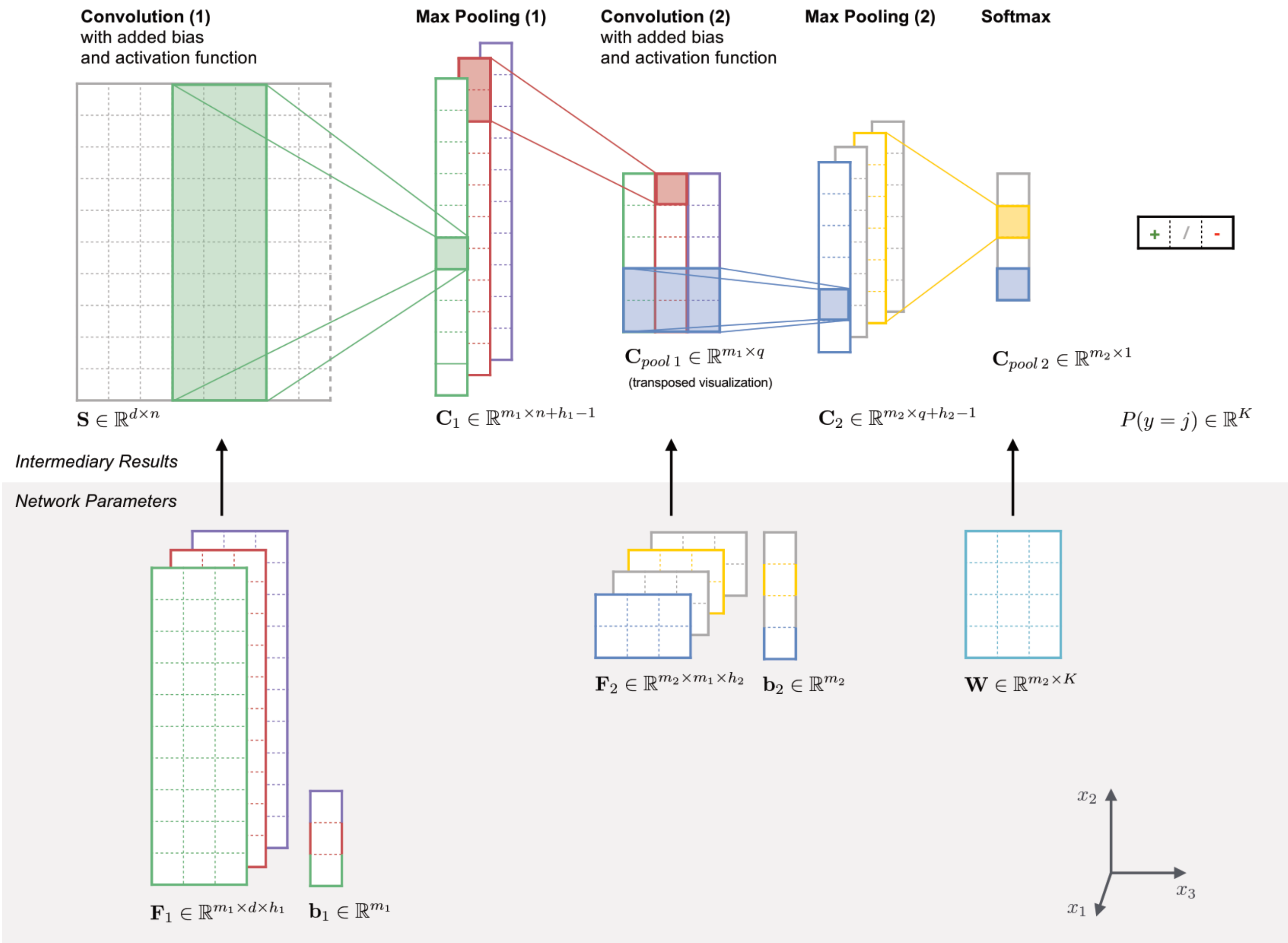
Zurich University
of Applied Sciences

A Shallow Dive into Technology

SwissCheese: 3-Phase Training with Distant Supervision

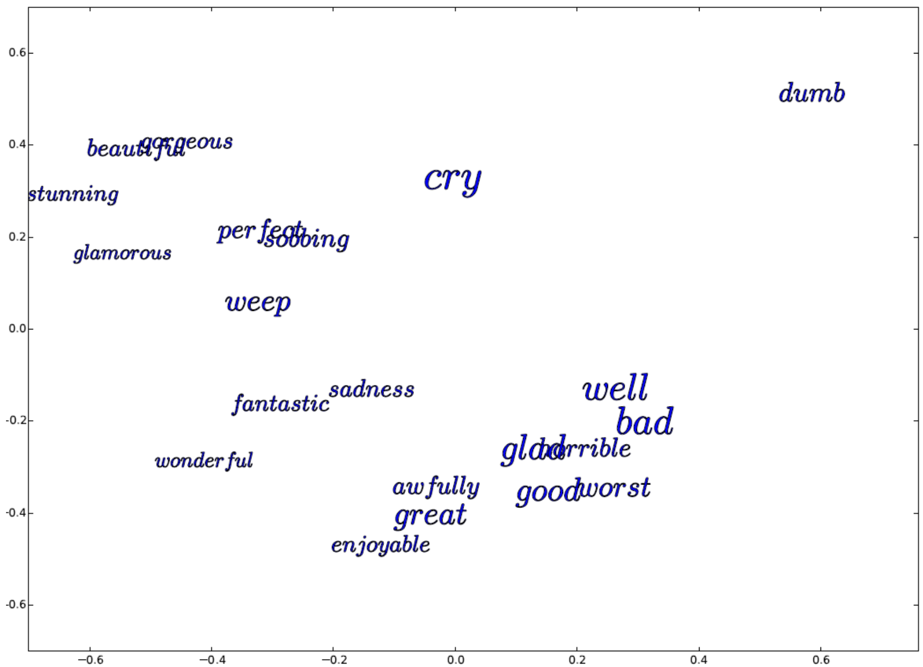


2-Layer Convolutional Neural Network for Sentiment Analysis

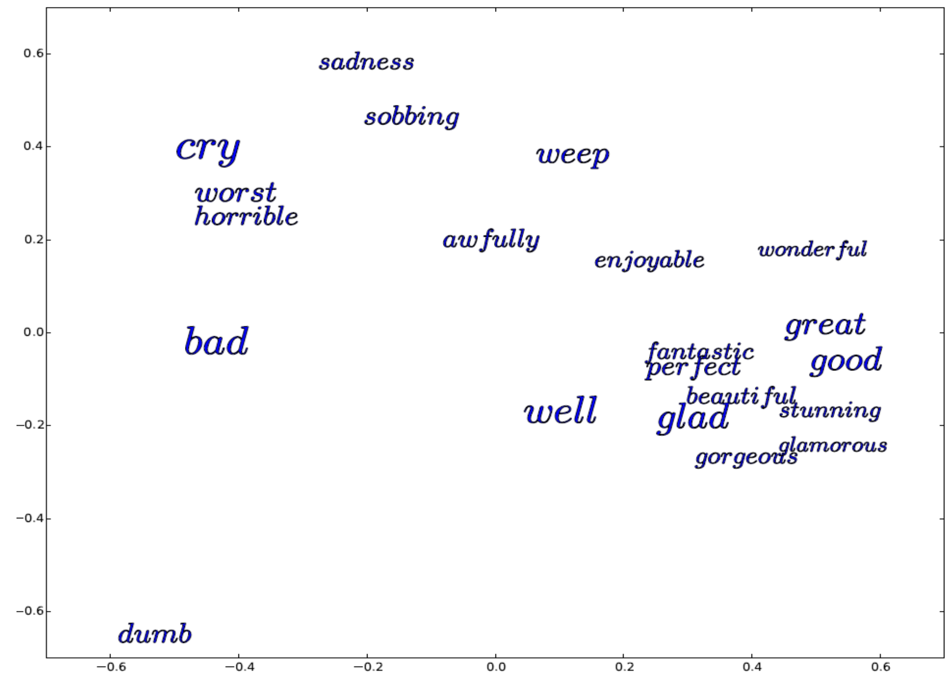


Zurich University
of Applied Sciences

Distant Phase rearranges Word Embeddings

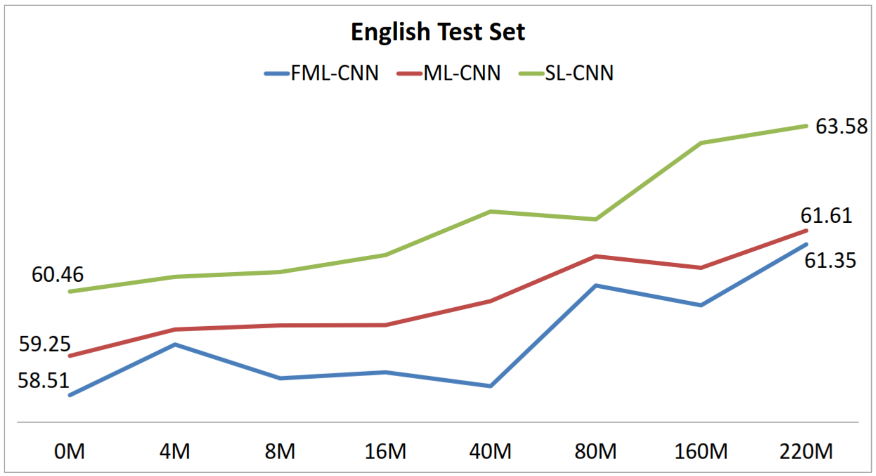


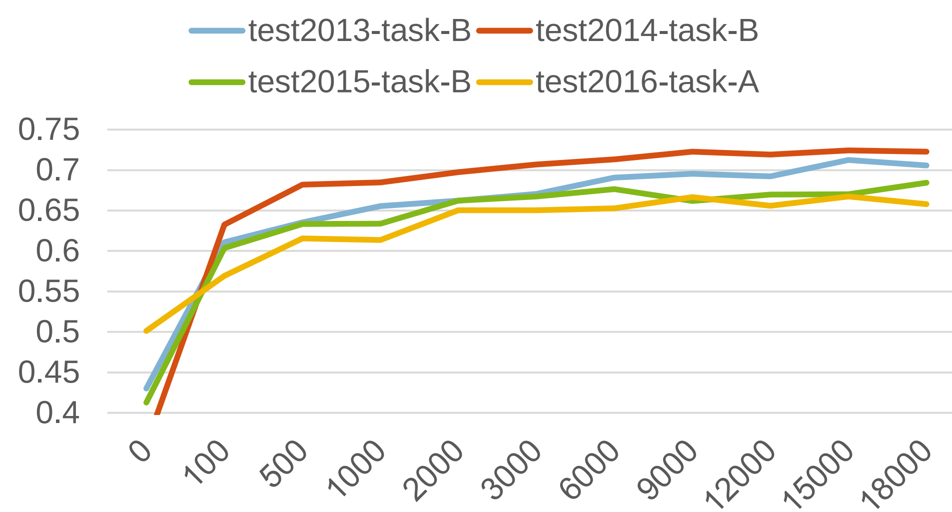
Before the Distant Phase



After the Distant Phase

The More Data, The Better!





Number of tweets in distant phase

Number of annotated tweets

Learn on Tweets, Classify News?

test / train	SemEval'13_tweets	MPQ_reviews	DIL_reviews	DAI_tweets	Union of All Test Data
SemEval'13_tweets	72.4	45.8	53.1	62.2	63.9
MPQ_reviews	62.2	54.1	40.9	57.8	58.7
DIL_reviews	57.3	36.8	55.1	48.5	52.9
DAI_tweets	67.9	37.7	50.4	70.8	60.4
Union of All Training Data	73.0	50.8	49.9	76.6	66.6

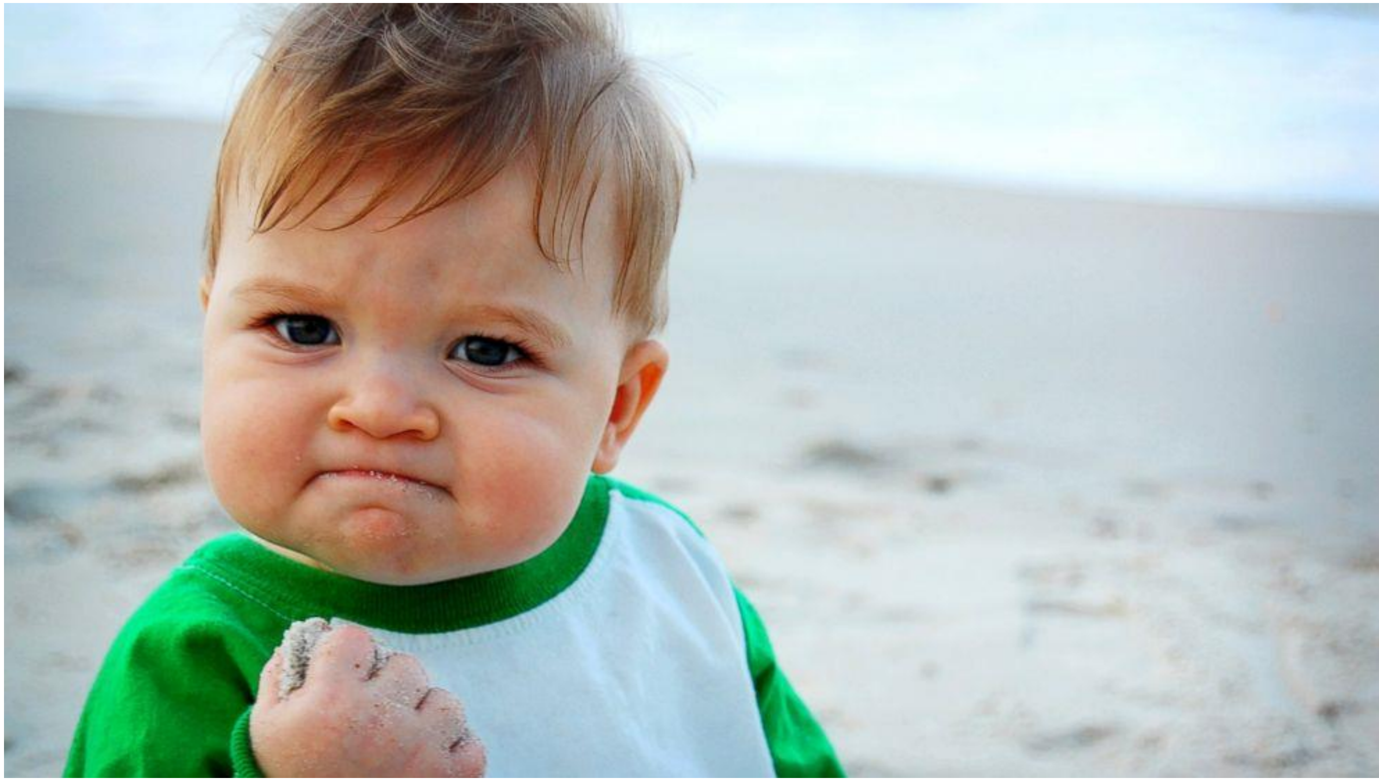
Measured in F1 score

Cross-Domain Performance of SemEval Winner 2016

Zurich University
of Applied Sciences

Sentiment for other Languages

Language	Available Data	Best Know Result (F1 Score)	Reference
German	10'000 Tweets	64.19	Deriu et al., 2016, WSDM (submitted)
Spanish	68'000 Tweets	71.1 (precision)	Villena-Roman et al., 2013, Procesamiento del Lenguaje Natural
Italian	7'000 Tweets	65.87	Deriu et al., 2016, WSDM (submitted)
Dutch	1'100 Tweets (labeled pos/neg)	88.33	Deriu et al., 2016, WSDM (submitted)
Arabic	1'100 Tweets	73.5	Salab et al., 2015, ANLP



We did it: Theory is over!

Do It Yourself:
Sentiment Analysis Tools and APIs

Big Players
- Google Prediction API
- IBM AlchemyAPI
- Microsoft Azure Text Analytics API

NLP Specialists
- RapidMiner
- Repustate
- Semantria
- SentiStrength
- SpinningBytes

Development Toolkits
- Natural Language ToolKit NLTK (Python)
- StanfordNLP (Java)

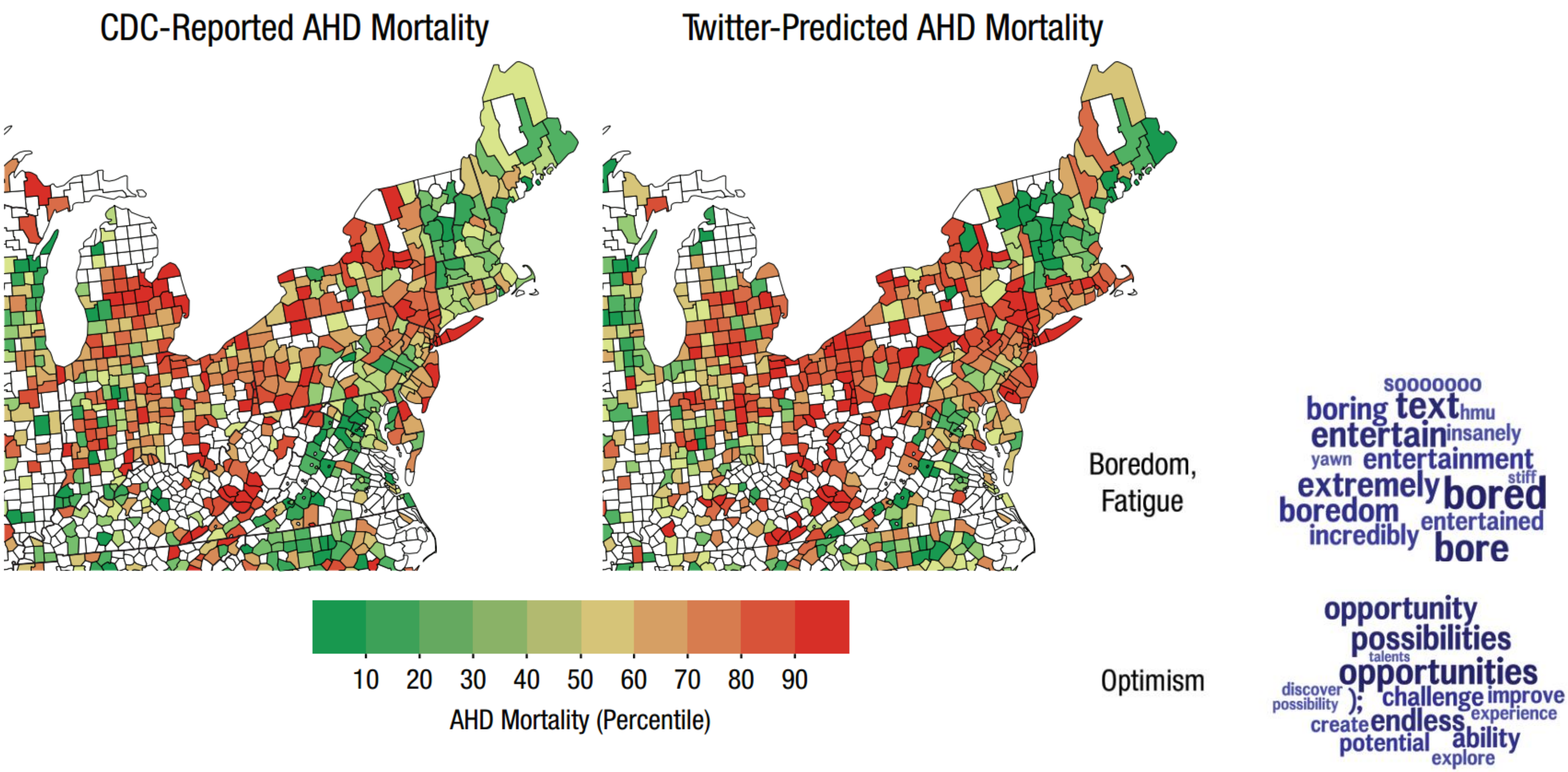
Understand Customer Reviews

Zurich University of Applied Sciences

Example: Aspect-based Sentiment Analysis for Hotel Reviews

Source: http://blog.aylien.com/aspect-based-sentiment-analysis-now-available-in/

Use Twitter to predict Heart Disease Mortality



Source: Eichstaedt et al., 2015: Psychological Language on Twitter Predicts County-Level Heart Disease Mortality

Zurich University
of Applied Sciences

"Cleantechness" of Company Products

Company Website

Automatic Classifier

Cleantech Topics

Disaster Prevention

Energy Transportation

Energy Production

Energy Efficiency

Mobility

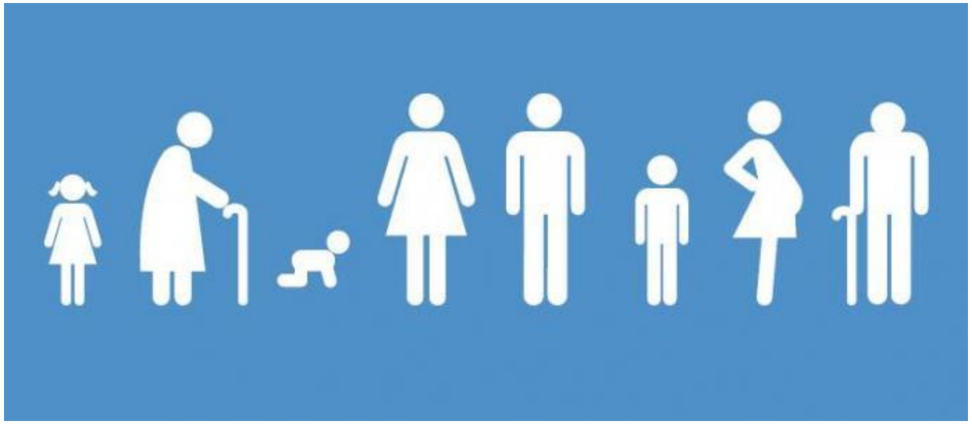
Air and Environment

Age and Gender of "Anonymous" Users

Goal: **Predict age (18-24, 25-34, 35-49, 50+) and gender (male/female) of Twitter users**

Results PAN 2015:

Age:	86%
Gender:	84%



Source: Rangel et al., 2015: Overview of the 3rd Author Profiling Task at PAN 2015

Talk in Short!

Sentiment Analysis

- approx. 70% F1-score
- the more data – the better
- has important application

Thanks!!

Mark Cieliebak

Zurich University of Applied Sciences (ZHAW)

Email: ciel@zhaw.ch, Website: www.zhaw.ch/~ciel

This presentation is based on joint work with:

- **Aurelien Lucchi, ETH**
- **Dominic Egger, ZHAW**
- **Fatih Uzdilli, ZHAW**
- **Jan Deriu, ZHAW**
- **Leon Derczynski, Univ. of Sheffield**
- **Martin Jaggi, EPFL**
- **Maurice Gonzenbach, ZHAW**
- **Valeria de Luca, ETH**