Zurich Universitiy of Applied Sciences

Potential and Limitations of Commercial Sentiment Detection Tools

Fatih Uzdilli

joint work with **Mark Cieliebak** and **Oliver Dürr**

03.12.2013 @ ESSEM'13

About Me

Zurich Universitiy of Applied Sciences

Fatih Uzdilli

Institute of Applied Information Technology (InIT)

ZHAW, Winterthur, Switzerland

Email: uzdi@zhaw.ch, more about me: home.zhaw.ch/~uzdi

Research Interest

Information Retrieval, Machine Learning, Sentiment Analysis

Background

Software Engineer, Social Media Monitoring, Search Technologies

Abstract

Zurich Universitiy of Applied Sciences

Evaluation of 9 commercial sentiment tools on approx. 30'000 short texts.

Best commercial tools have accuracy of only 60%.

Combining all tools using Random Forest improved the accuracy.

Motivation

Zurich Universitiy of Applied Sciences

- **Scientific results** for sentiment detection:

«*very good performance:* ***> 80% accuracy***»

- Blog posts about **commercial tools**:

«*very poor quality,* ***unusable***»

Motivation

Zurich Universitiy of Applied Sciences

- **Scientific results** for sentiment detection:

«*very good performance: > 80% **accuracy***»

Really?

- Blog posts about **commercial tools**:

«*very poor quality, **unusable***»

How good is commercial Sentiment Detection?

source: http://www.commute.com/images/schools_evaluation.jpg

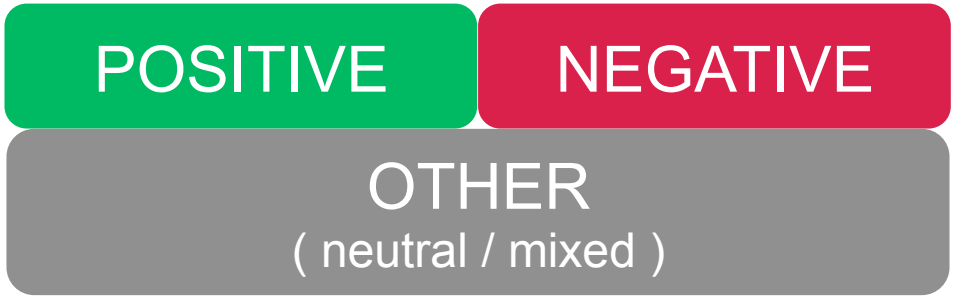
Is there potential for improvement?

source: http://3.bp.blogspot.com/-u3acK_WjaLU/ULYv51mHEhI/AAAAAAAAARY/ DIZqOfxuswc/s1600/IcebergQ1.jpg

Evaluation Setup

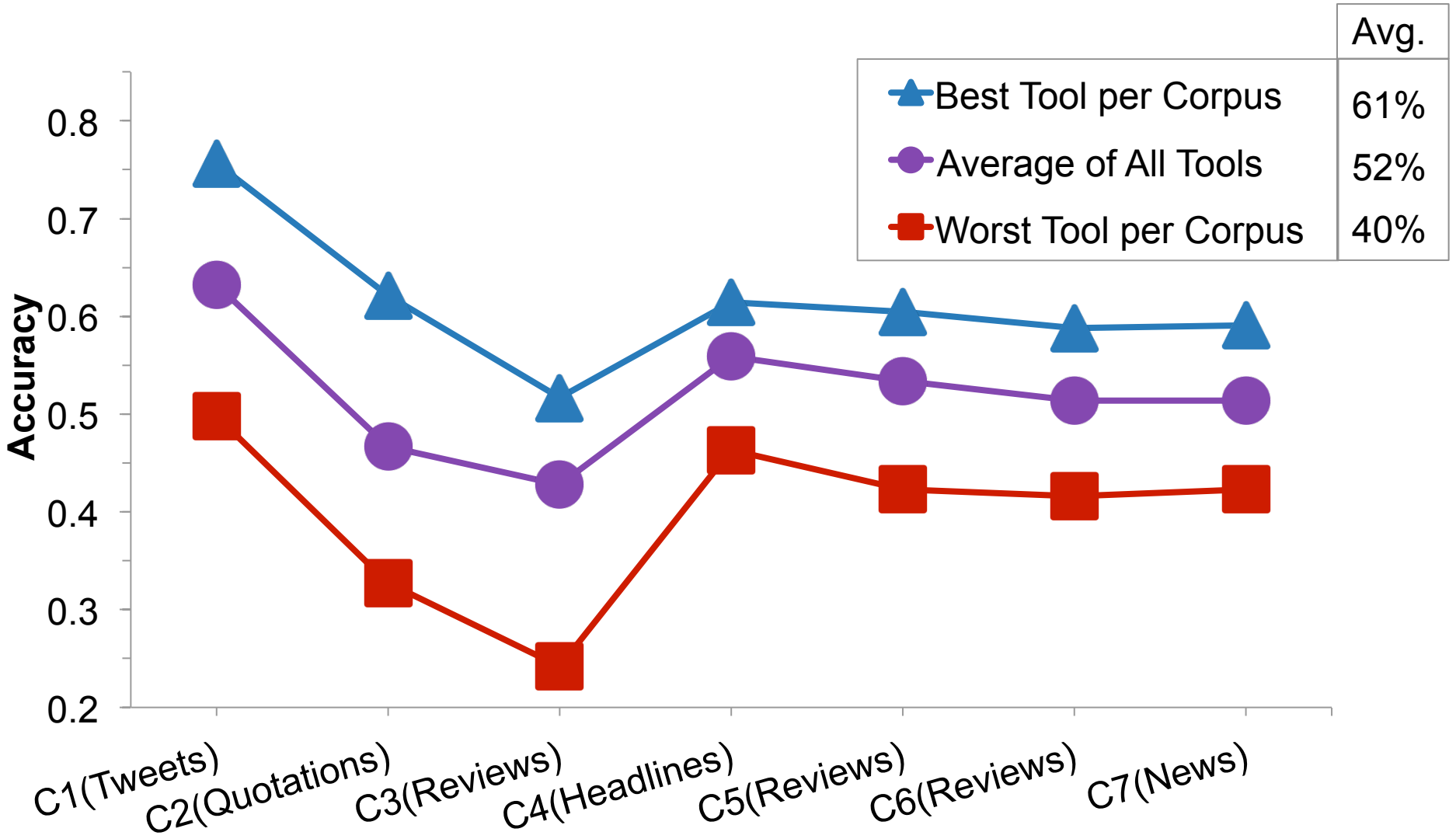
Zurich Universitiy of Applied Sciences

- 7 Public Text Corpora
  - Single Statements
  - Different Media Types
    - Tweet, News, Review, Speech Transcript
  - Total: 28653 Texts
- 9 Commercial APIs
  - Stand-alone
  - Free for this evaluation
  - Arbitrary Text



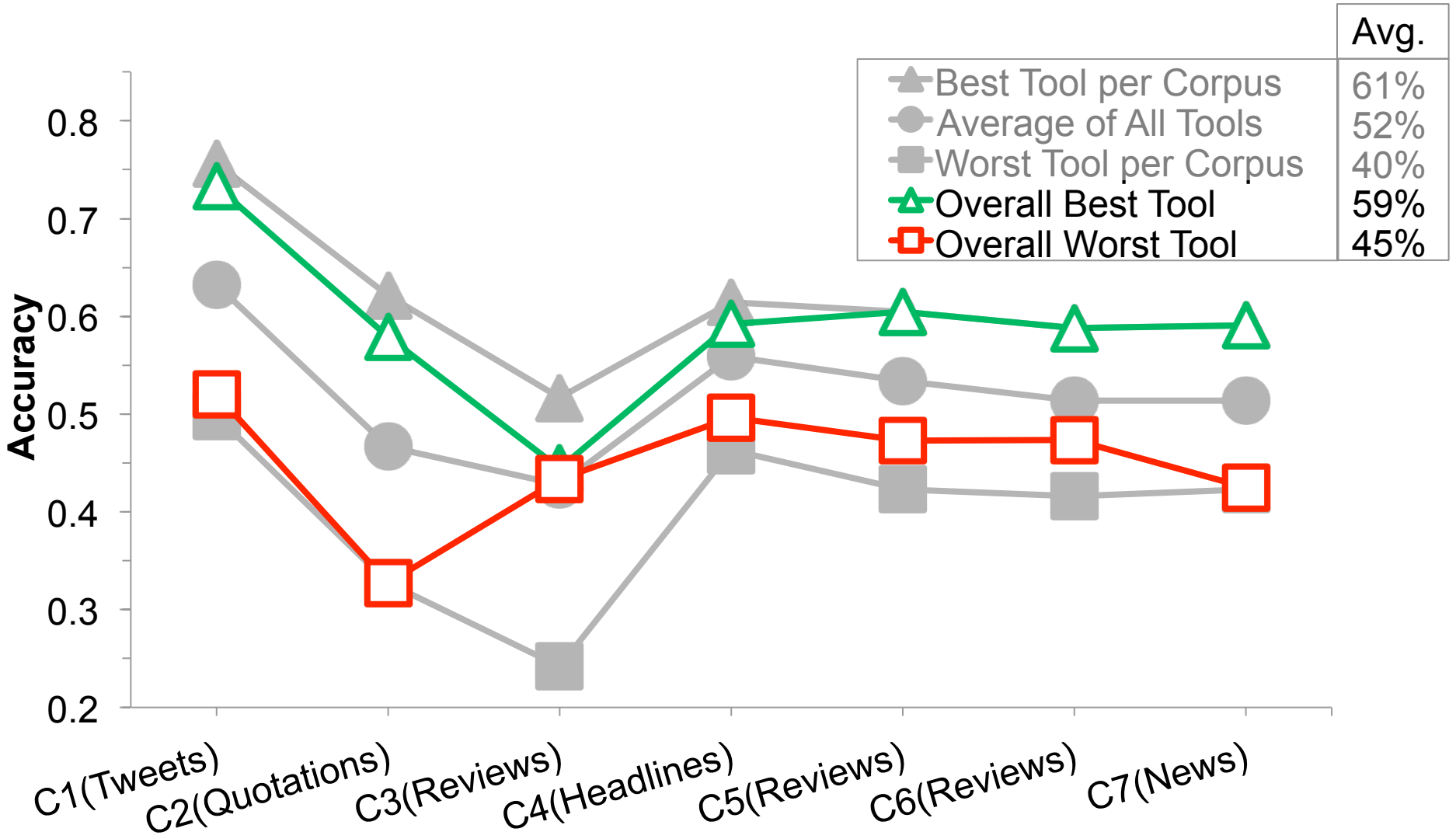
Tool Accuracy

Zurich University of Applied Sciences



Tool Accuracy

Zurich University of Applied Sciences



Further Findings

Zurich Universitiy of Applied Sciences

- Longer texts are hard to classify
- Corpus annotations might be erroneous

Can a Meta-Classifier do better?

Zurich Universitiy of Applied Sciences

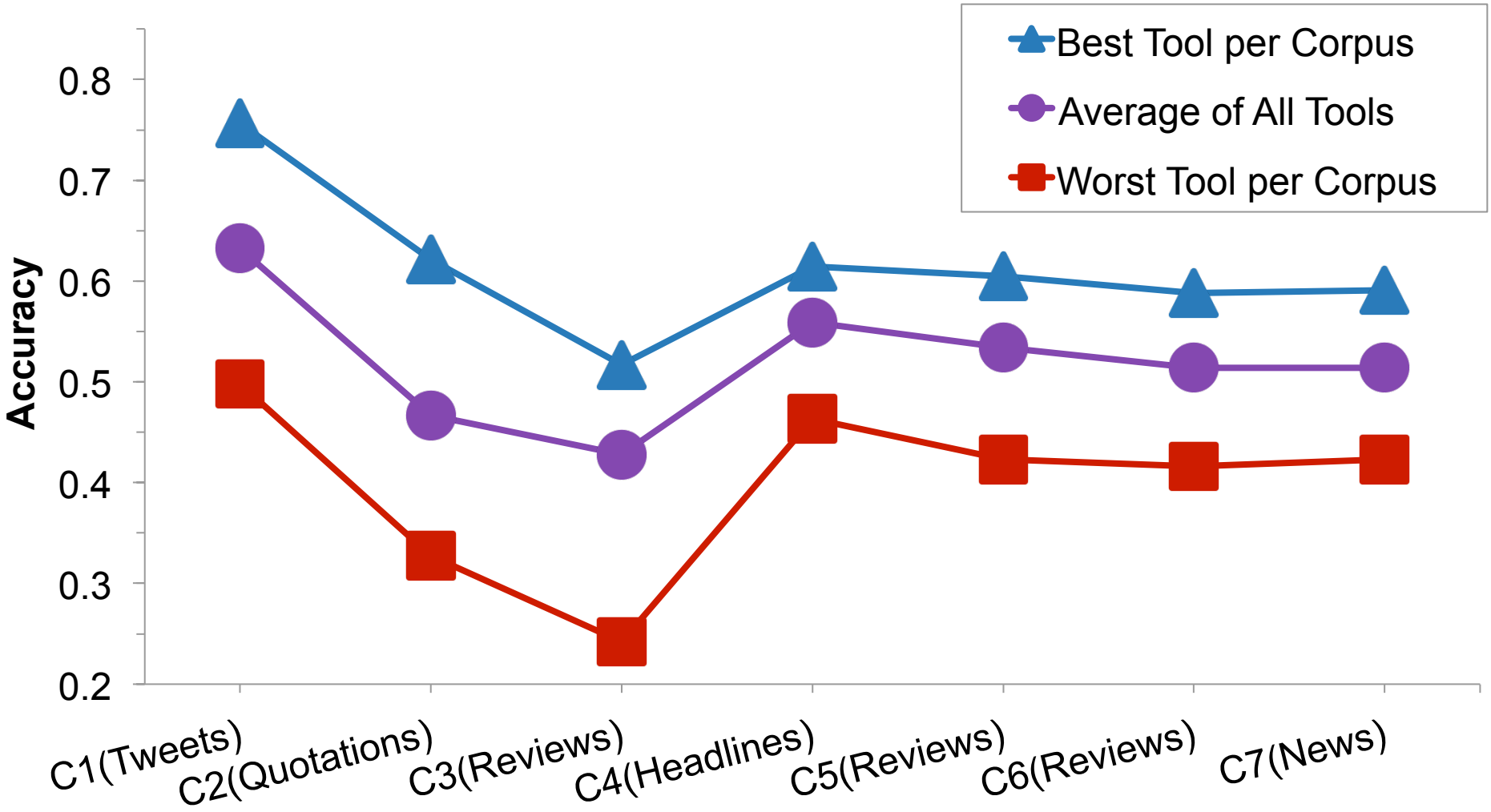
- 1st Approach: **Majority Classifier**
  - Sentiment with most votes chosen

Illustration:

	api1	api2	api3	api4	api5	api6	api7		Majority
Text 1	+	+	-	o	-	+	o		+
Text 2	-	+	+	-	-	-	-		-
Text 3	-	o	+	+	+	+	-		+
Text n	o	o	+	o	-	o	o		o

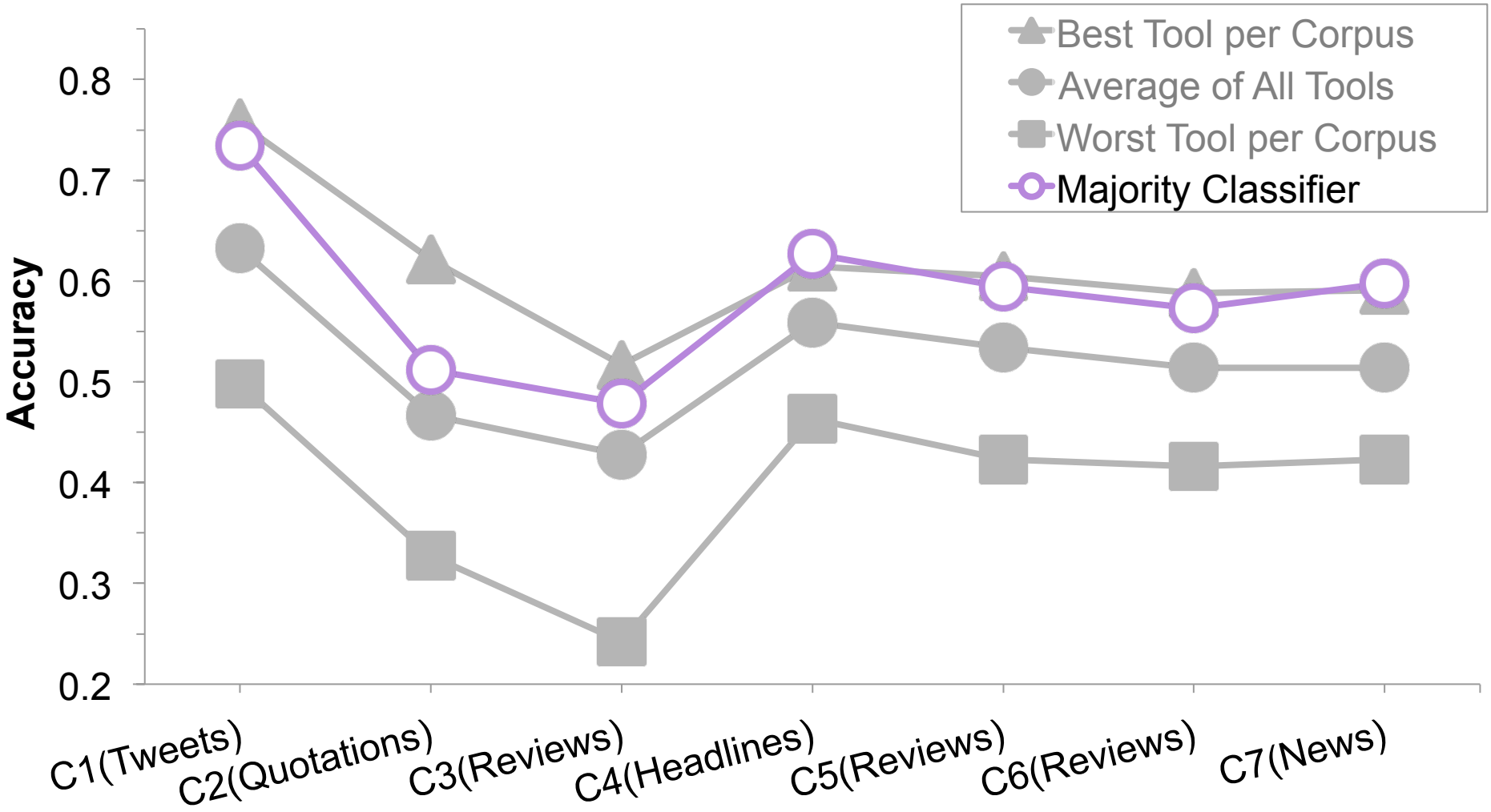
Tool Accuracy

Zurich Universitiy of Applied Sciences



Majority Classifier beats Average

Zurich University of Applied Sciences



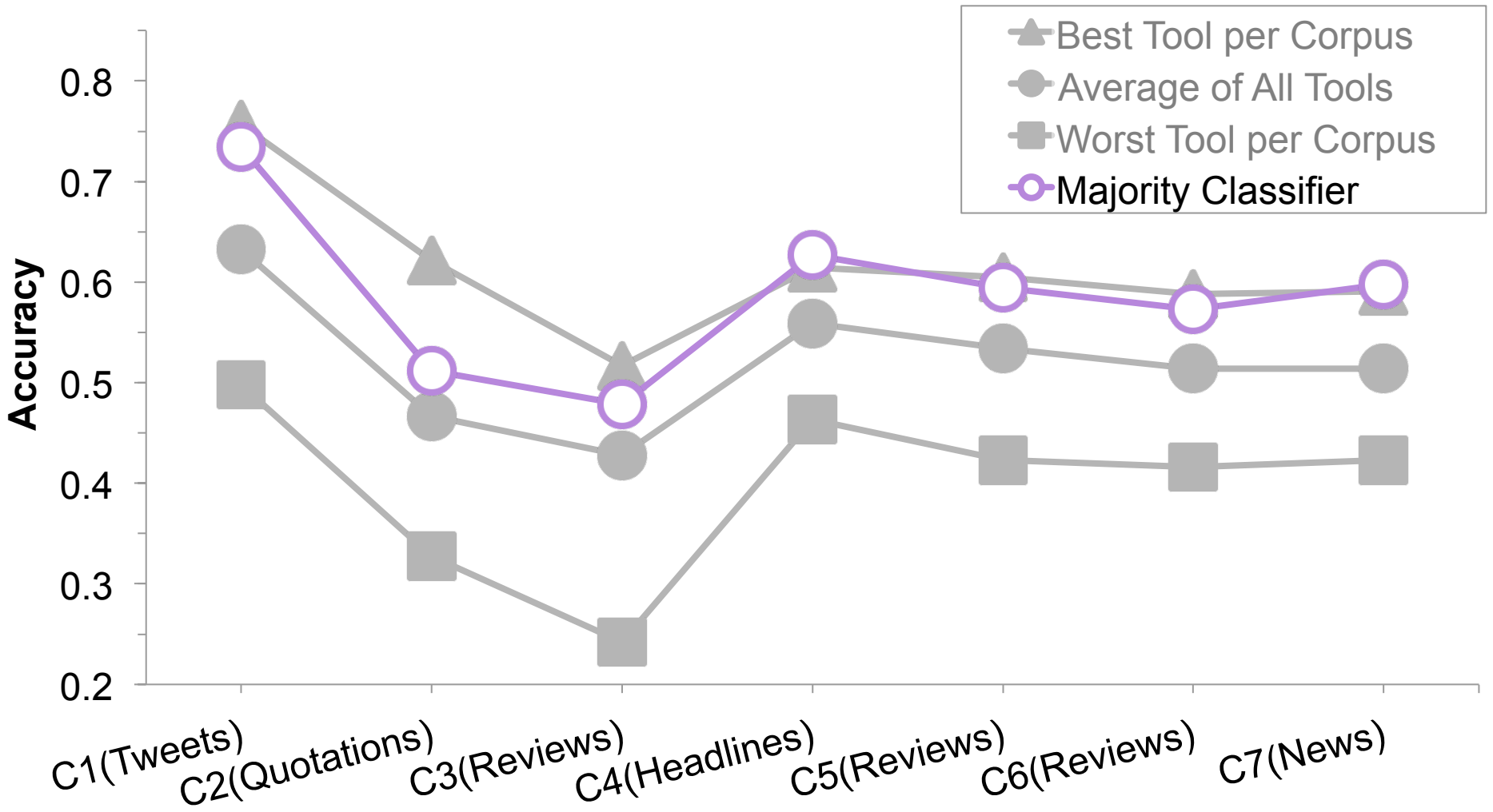
2nd Approach: Random-Forest

Zurich Universitiy of Applied Sciences

	api1	api2	api3	…	api n	annotation		Random Forest Classifier
Text 1	+	-	+	…	o	+	Train	
Text 2	-	+	o	…	+	-	Train	
Text 3	-	o	-	…	+	-	Train	
Text 4	+	o	+	…	-	+	Train	
Text 5	+	o	+	…	o	o	Train	
Text 6	+	o	o	…	-	o	Train	
Text 7	+	-	+	…	o	*unknown*	Predict	+
Text 8	+	+	o	…	-	*unknown*	Predict	+
Text 9	o	-	+	…	o	*unknown*	Predict	o

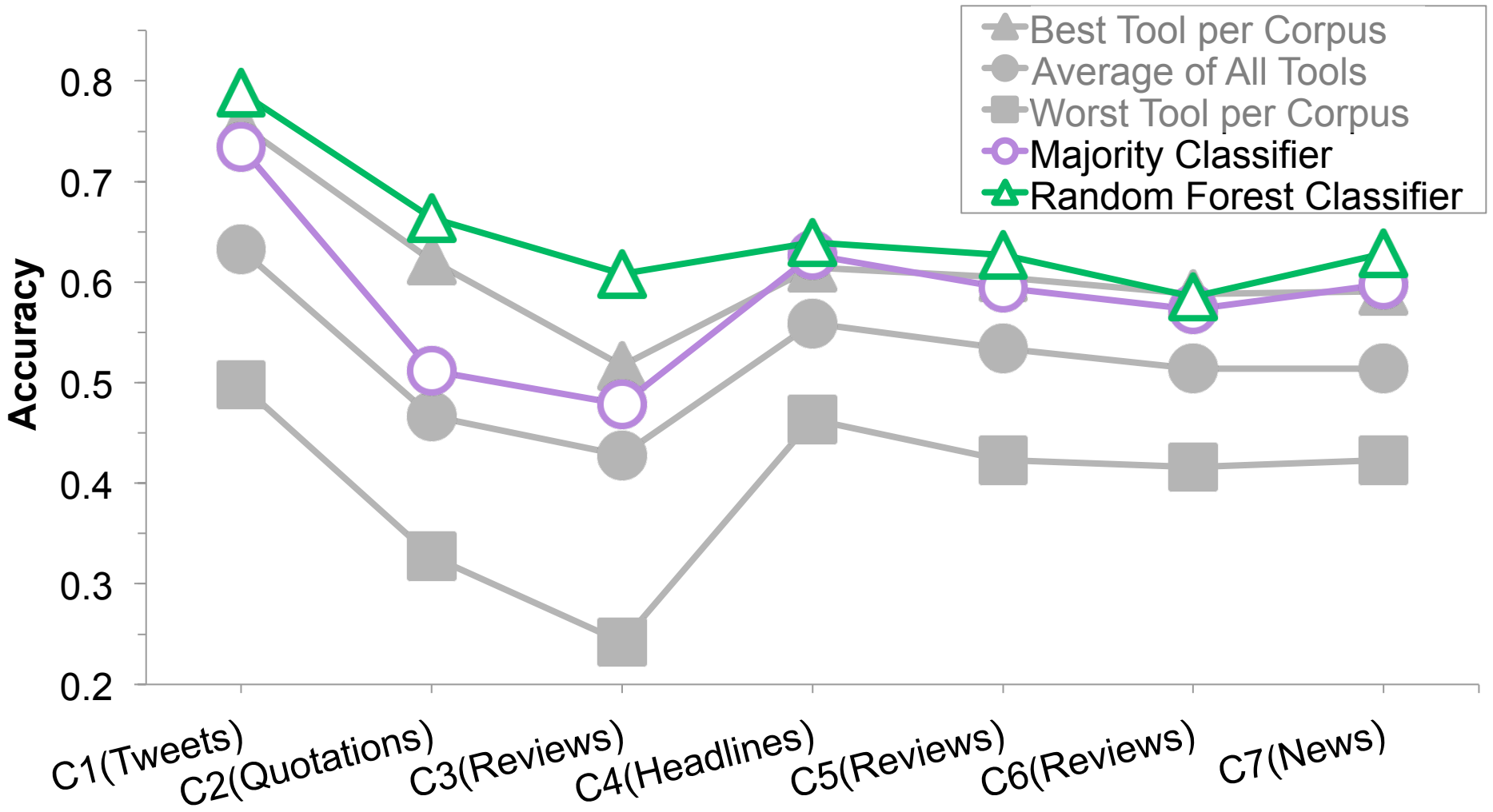
Before Random Forest

Zurich University of Applied Sciences



Random Forest Beats Best Single Tool

Zurich University of Applied Sciences



Summary

Zurich Universitiy of Applied Sciences

- Best Tool: 59% Accuracy
- Random Forest combination: Up to 9% improvement