What's important in a text? An extensive evaluation of linguistic annotations for summarization

Markus Zopf, Teresa Botschen, Tobias Falke, Benjamin Heinzerling, Ana Marasovic, Todor Mihaylov, Avinesh P.V.S, Eneldo Loza Mencia, Johannes Fürnkranz and Anette Frank
`https://www.aiphes.tu-darmstadt.de`

AIPHES
ADAPTIVE INFORMATIONSAUFBEREITUNG AUS HETEROGENEN QUELLEN

RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

Heidelberger Institut für
Theoretische Studien

HITS

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Suppose you are a journalist…

TECHNISCHE UNIVERSITÄT DARMSTADT

AIPHES

The **U.S. Congress** will certify the results on January 6, 2017

Donald Trump won the election and will become the 45th president

There are rumors about **Russian interferences** in the elections

Automatic summarization = reduce text length while preserving most important information

1. make text shorter: $|\mathrm{Y}| < \sum_{i=1}^{n} |\mathrm{X}_i|$
2. don‘t add content: $\mathrm{Y} \subseteq \bigcup_{i=1}^{n} \mathrm{X}_i$
3. maximize a utility: $\mathrm{Y}^* = \underset{\mathrm{Y}}{\operatorname{argmax}} \; \boldsymbol{u}(Y)$

Suppose you are a journalist…

TECHNISCHE UNIVERSITÄT DARMSTADT

AIPHES

The **U.S. Congress** will certify the results on January 6, 2017

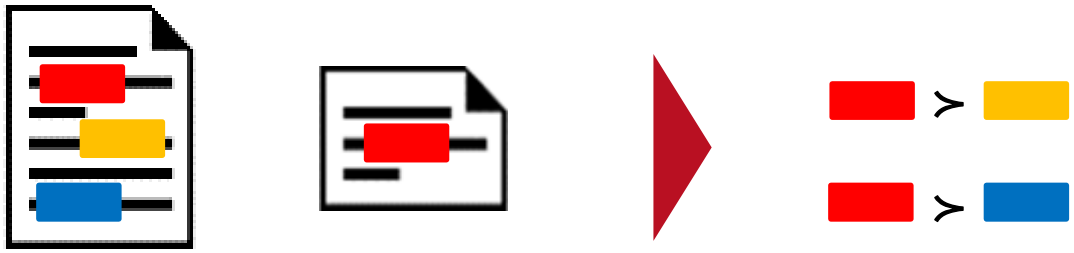
Donald Trump won the election and will become the 45th president

There are rumors about **Russian interferences** in the elections

?

Estimating information importance is a key challenge in automatic summarization

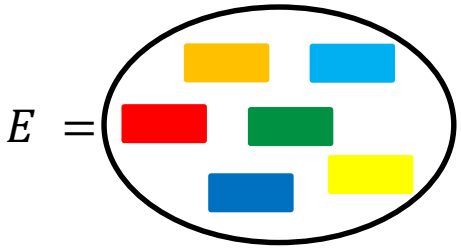
Promoted elements are preferred over not promoted elements



document X^i summary Y^i

- promoted elements $P^i = X^i \cap Y^i = \{$ $\}$
- not promoted elements $N^i = X^i \setminus P^i = \{$, $\}$

$$\mathrm{u}( ) = \frac{1}{|E|} \sum_{e \in E} \frac{n(\succ e)}{n(\succ e) + n(\succ e)}$$



$$E =$$

$$\mathrm{v}(x_j \in X^i) = \frac{1}{|s|} \sum_{e \in S} v(e)$$

Experimental Setup

TECHNISCHE UNIVERSITÄT DARMSTADT

AIPHES

Annotations under Investiation

- Unigrams, bigrams, trigrams
- Chunks - parts of sentences with specific grammatical meaning
- Concepts detected with open information extraction
- Verb stems - {killing, killed} → kill
- FrameNet frame annotations
- Connotation frames - subjective roles and relationships
- Discourse Relation Senses, e.g. *causation*, *contrast*, or *concession*

Data used

- DUC 2004, TAC 2008, and TAC 2009 summarization dataset

Rankings for importance estimation evaluation

$x_i \in X$ $\quad Y$

$$\text{Precision} = \frac{x_i \cap Y}{|x_i|}$$

$$\text{Recall} = \frac{x_i \cap Y}{|Y|}$$

P	R
x_4	x_1
x_2	x_7
x_3	x_4
...	...

Dataset Generation

P	R
x_4	x_1
x_2	x_7
x_3	x_4
...	...

target

P	R
x_3	x_5
x_8	x_1
x_4	x_4
...	...

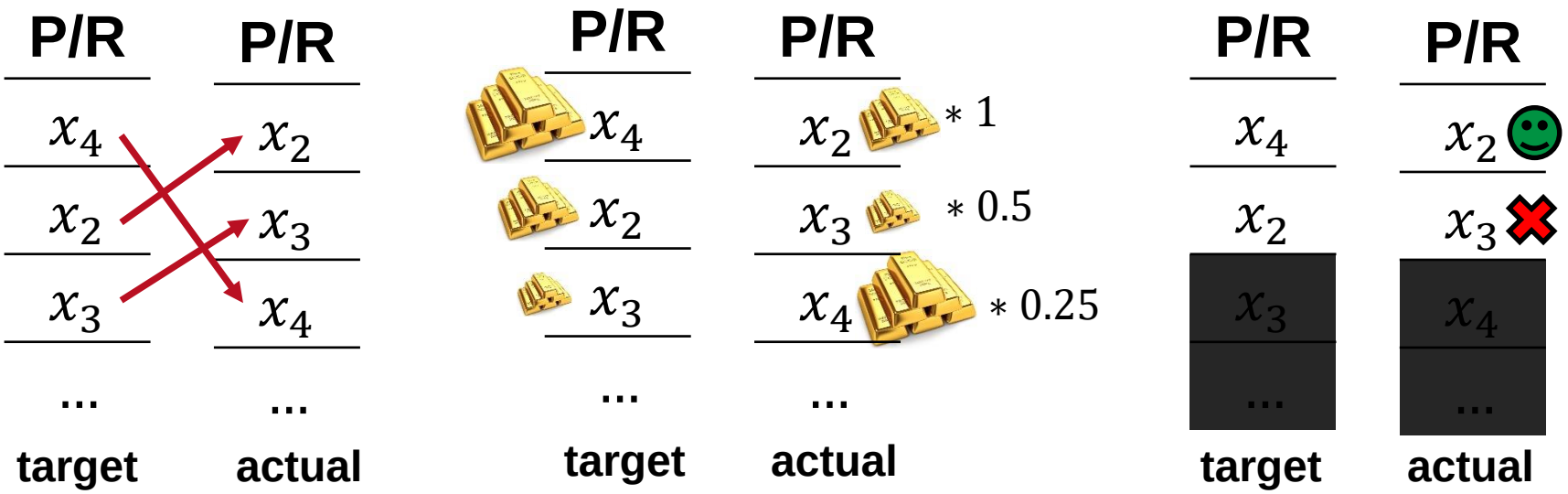
actual

Evaluation

We measure the distance to the target ranking with three different measures

TECHNISCHE UNIVERSITÄT DARMSTADT

AIPHES



Kendall's Tau

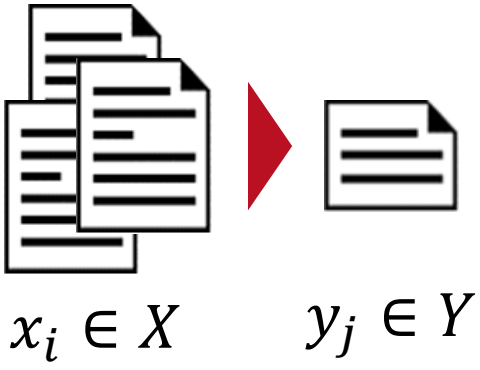
NEW **nDCRS**

precision@k

Preference prediction for importance estimation evaluation

TECHNISCHE UNIVERSITÄT DARMSTADT

AIPHES



$x_i \in X$ $y_j \in Y$

$x_i \prec y_j$

Dataset Generation

s_i, s_j ?

$s_i \succ s_j$

$s_i \prec s_j$

Evaluation

Ranking on unseen test data shows that simple annotations perform best

TECHNISCHE UNIVERSITÄT DARMSTADT

	Kendall's Tau		nDCRS		precision@k	
	P	R	P	R	P	R
bigram	**.306**	**.539**	**.253**	**.863**	**.253**	**.424**
cf-effect-object	-.051	.269	.083	.687	.083	.230
cf-state-subject	-.054	.284	.083	.697	.083	.234
chunk-concepts	**.175**	**.367**	**.206**	**.773**	**.206**	**.298**
concepts-string	.106	.193	.146	.639	.146	.179
concepts-sim	.093	.225	.135	.669	.135	.225
connotation-frames	-.011	.335	.089	.739	.089	.267
entity-importance	-.076	-.060	.107	.510	.107	.193
entity-links	**.135**	.264	**.169**	.709	**.169**	.261
entity-type-coarse	.031	.100	.138	.582	.138	.155
entity-type-corenlp	.075	.358	.132	.766	.132	**.316**
entity-type-figer	.122	.272	.165	.709	.165	.243
entity-type-fine	.117	.269	.163	.708	.163	.236
FN-frames	.027	**.383**	.107	**.772**	.107	.297
FN-frames-nounsOnly	.116	**.474**	.133	**.836**	.133	**.364**
FN-frames-verbsOnly	.010	.209	.096	.639	.096	.186
sentiment-annos	.068	.215	.148	.673	.148	.222
discours-rel	.011	.234	.133	.646	.133	.174
trigram	**.172**	.366	**.186**	.760	**.186**	.241
unigram	**.300**	**.654**	**.260**	**.913**	**.260**	**.515**
verb-stem	.042	.250	.114	.671	.114	.215

Simple annotations do not perform best at preference prediction on unseen test data

TECHNISCHE UNIVERSITÄT DARMSTADT

	DUC 2003	DUC 2004	TAC 2008	TAC 2009	average
bigram	**0.573**	0.538	0.415	0.445	0.493
cf-effect-object	0.538	0.520	**0.663**	**0.743**	**0.616**
cf-state-subject	0.548	0.439	0.420	0.512	0.480
chunk-concepts	**0.641**	**0.613**	**0.556**	0.602	**0.603**
concepts-string	0.513	0.429	0.371	0.382	0.424
concepts-sim	0.520	0.468	0.438	0.473	0.475
connotation-frames	0.551	0.556	0.546	0.592	0.561
entity-importance	**0.597**	**0.634**	**0.655**	**0.658**	**0.636**
entity-links	0.510	0.450	0.370	0.364	0.424
entity-type-coarse	0.512	0.487	**0.664**	**0.695**	**0.590**
entity-type-corenlp	**0.582**	**0.608**	0.551	**0.616**	0.589
entity-type-figer	0.495	0.487	0.453	0.408	0.461
entity-type-fine	0.497	0.490	0.456	0.405	0.462
FN-frames	0.474	0.497	0.515	0.496	0.496
FN-frames-nounsOnly	0.521	0.537	0.531	0.539	0.532
FN-frames-verbsOnly	0.490	0.487	0.468	0.507	0.488
sentiment-annos	0.430	0.402	0.353	0.356	0.385
discours-rel	0.550	**0.608**	**0.628**	**0.604**	**0.598**
trigram	0.373	0.285	0.210	0.254	0.281
unigram	**0.617**	**0.601**	0.530	0.553	0.575
verb-stem	0.497	0.517	0.515	0.500	0.507

What's important in a text? An extensive evaluation of linguistic annotations for summarization

TECHNISCHE UNIVERSITÄT DARMSTADT

AIPHES

Summary


Information importance estimation is a key problem in summarization


we investigated a wide range of annotations to replace bigrams


annotations close to surface work well for ranking input sentences


they do not work well for preference prediction

Markus Zopf, Teresa Botschen, Tobias Falke, Benjamin Heinzerling, Ana Marasovic, Todor Mihaylov, Avinesh P.V.S, Eneldo Loza Mencia, Johannes Fürnkranz and Anette Frank

https://www.aiphes.tu-darmstadt.de