Summarization of German Court Rulings

Sebastian Moser, 10.5.2021, Guided Research

Chair of Software Engineering for Business Information Systems (sebis)
Faculty of Informatics
Technische Universität München
wwwmatthes.in.tum.de

Outline
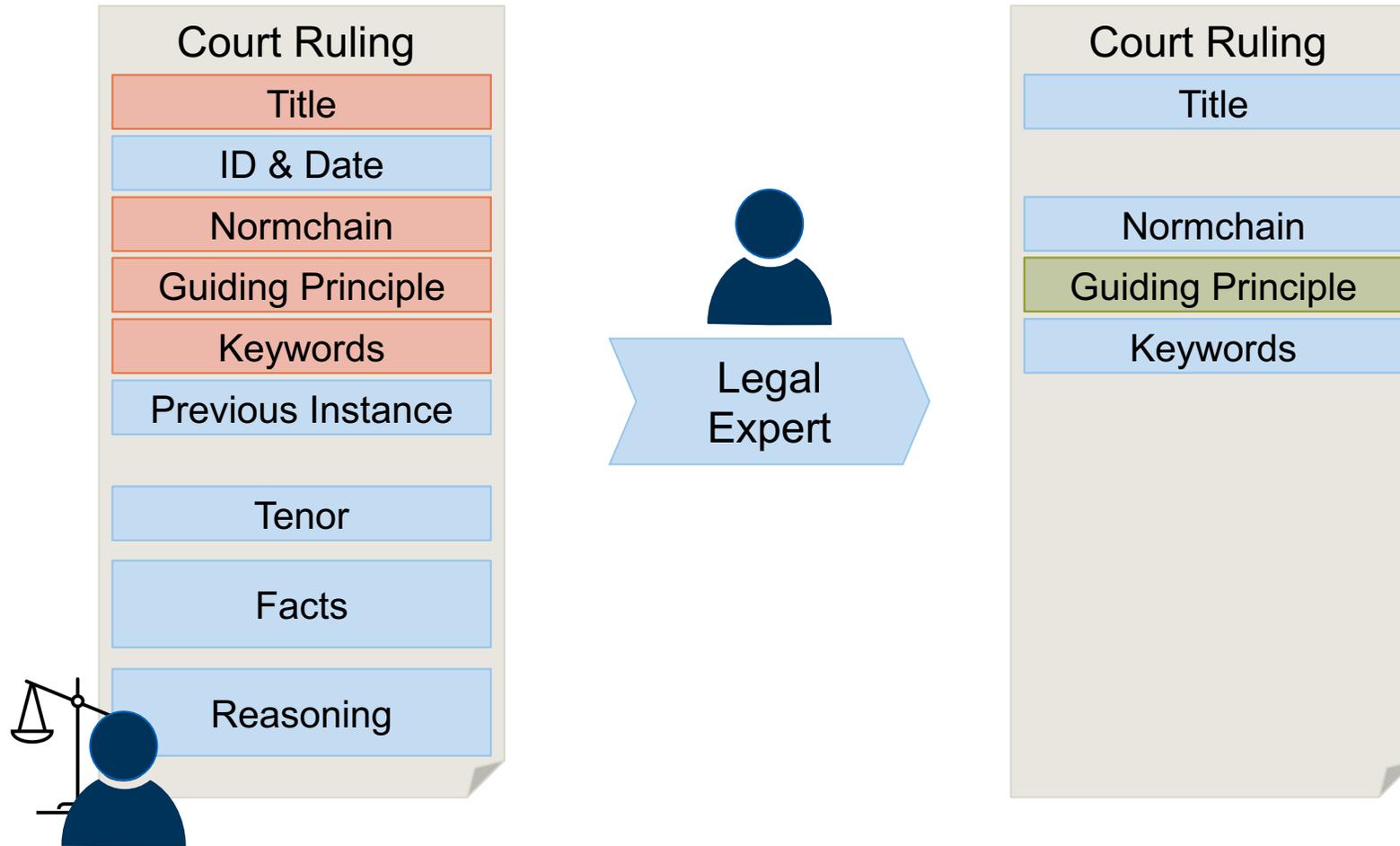
Motivation

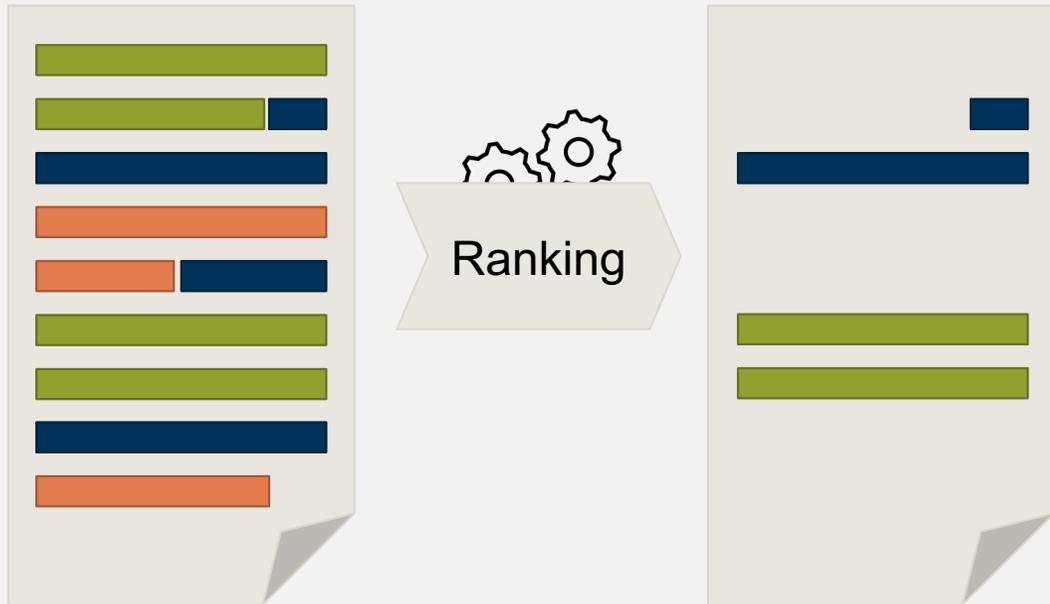Summarization & Research Questions

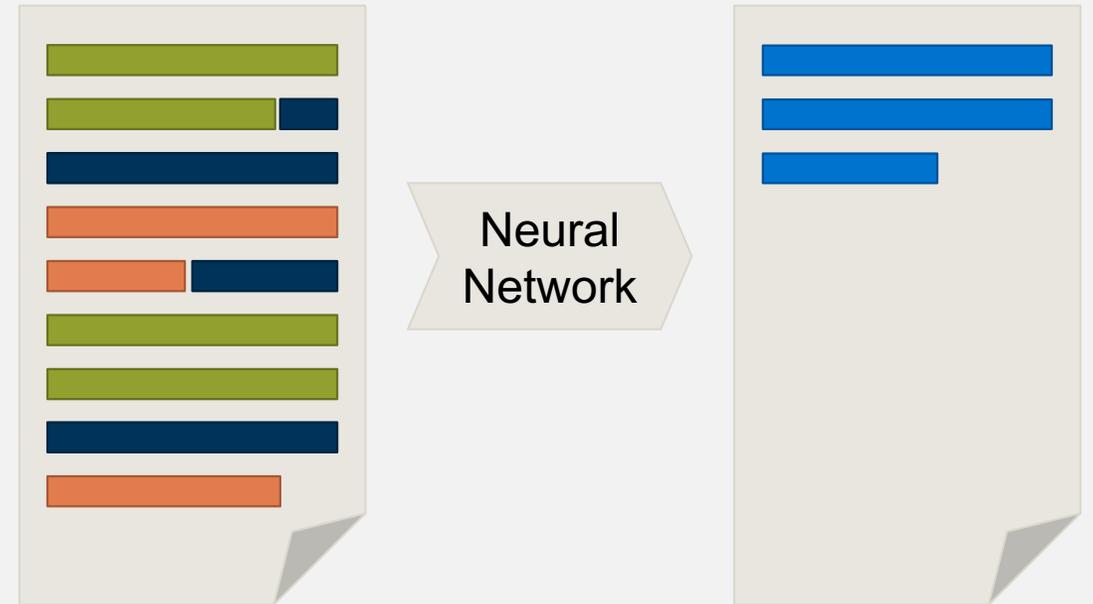Literature Research

Dataset

Data Pipeline

Methods & Evaluation

Limitations & Future Work

Motivation

Summarization

Extractive Summarization

Ranking

Abstractive Summarization

Neural Network

[1] [2] [3]

Research Questions

R1 How to generate high-quality summaries for German verdicts: Is extractive summarization sufficient or abstractive summarization necessary?

R2 How to apply summarization approaches to long, German legal documents?

Outline

Motivation

Summarization & Research Questions

Literature Research

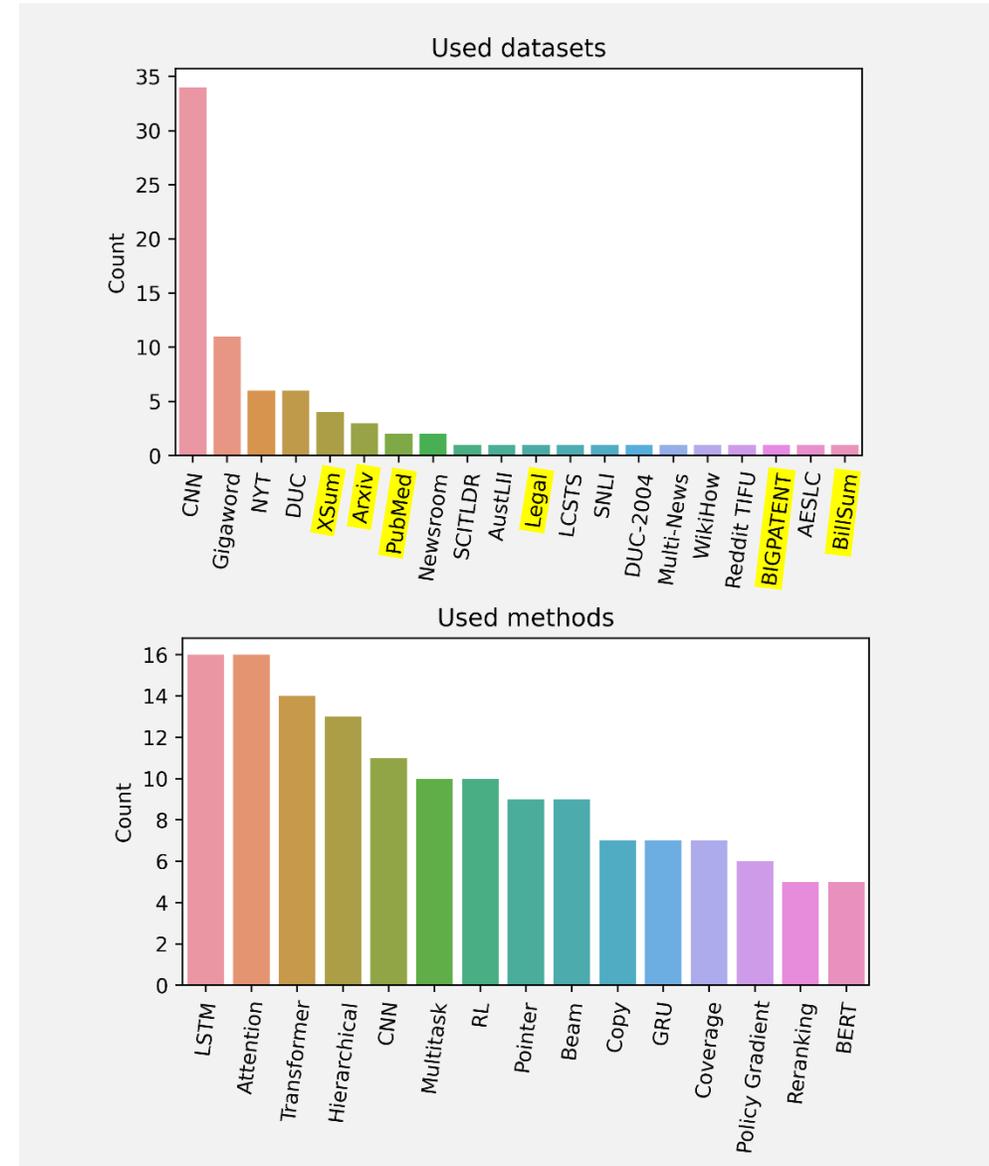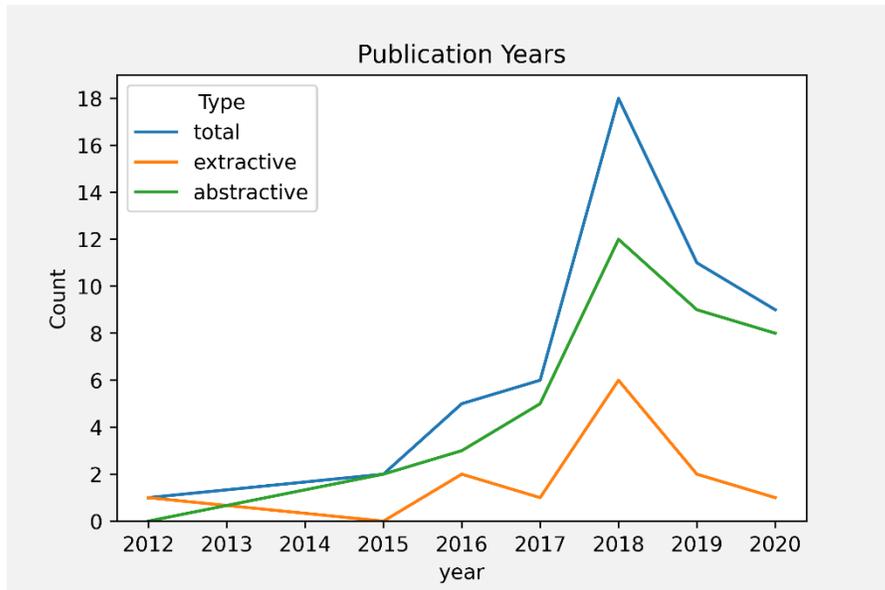Dataset

Data Pipeline

Methods & Evaluation

Limitations & Future Work

Literature Research

Outline

Motivation

Summarization & Research Questions

Literature Research

Dataset

Data Pipeline

Methods & Evaluation

Limitations & Future Work

Dataset - Acquisition

Justiz-ONLINE
Justizportal Nordrhein-Westfalen
justiz.nrw.de

[4] Se ✓ → HTML 158k → Custom Processing → Sentence Segmentation

BAYERN.RECHT
gesetze-bayern.de

[4] Se ✓ → XML 35k → Custom Processing → Sentence Segmentation

ottoschmidt

XML 72k → Custom Processing → Sentence Segmentation

JSON

Title	
ID & Date	
Normchain	
Norms	
Guiding Principle	
Keywords	
Previous Instance	
Tenor	
Facts	
Reasoning	

Dataset - Acquisition

Justiz-ONLINE
Justizportal Nordrhein-Westfalen

justiz.nrw.de

JSON
158k

BAYERN.RECHT

gesetze-bayern.de

JSON
35k

ottoschmidt

JSON
72k

Filtering → Manual Processing → Post Processing

JSON
100k

Dataset - Statistics

Dataset	#Docs	Avg. Length	Avg. Summary	Compression	
CNN	92k	656	43	6.6%	[5]
Daily Mail	219k	693	52	7.5%	
NYT	655k	530	38	7.2%	
PubMed	133k	3016	203	6.7%	
ArXiv	215k	4938	220	4.5%	
Newsroom	1212k	751	45	6.0%	[6]
BigPatent	1341k	3573	117	3.3%	
XSum	227k	431	23	5.3%	[7]
Legal	100k	2422	75	3.1%	

Verdict Publication Dates

Number of Previous Instances

Normchain

Novelty ≈ "Percentage of text only found in the summary"
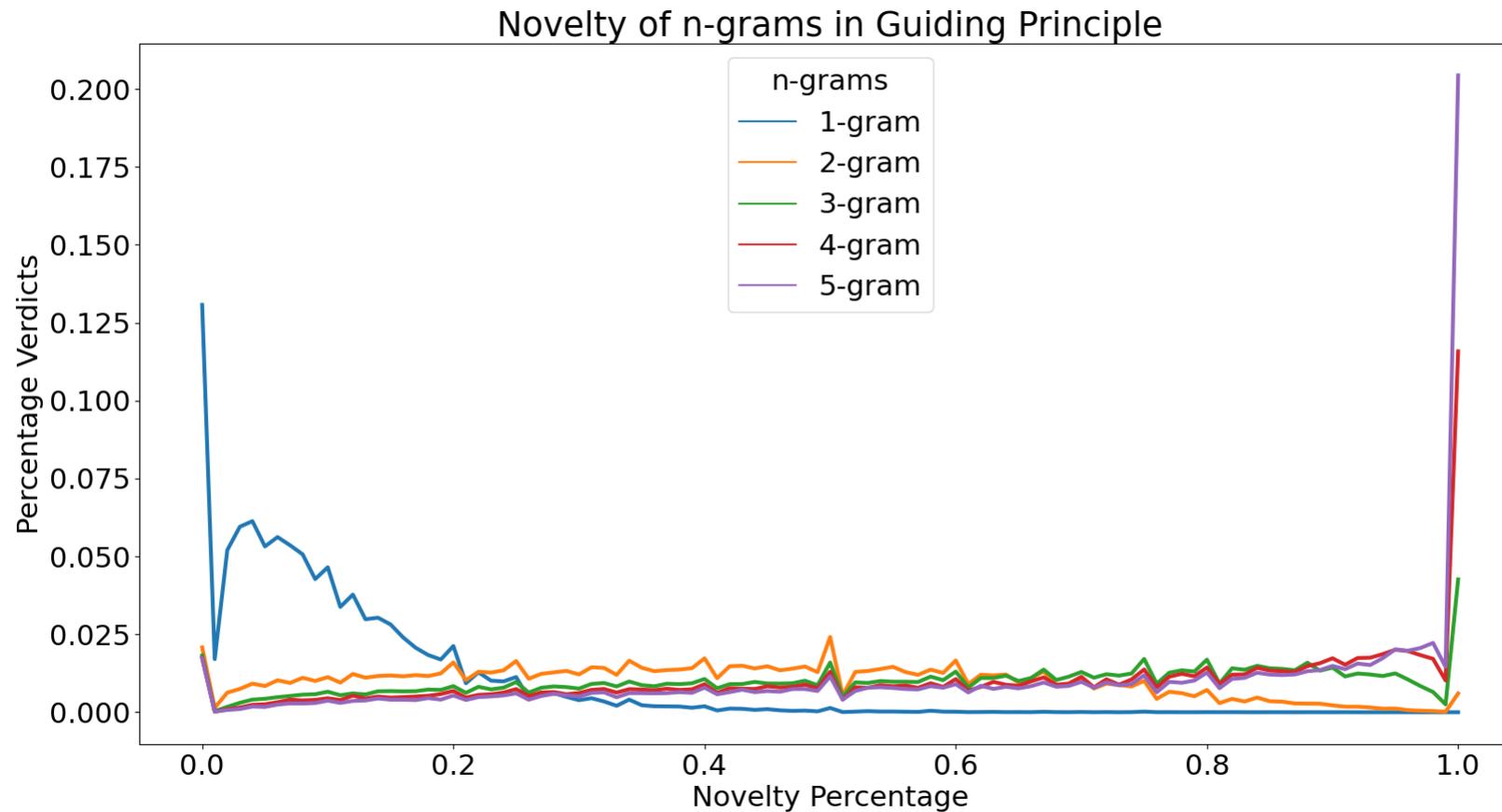
Novelty of n-grams in Guiding Principle
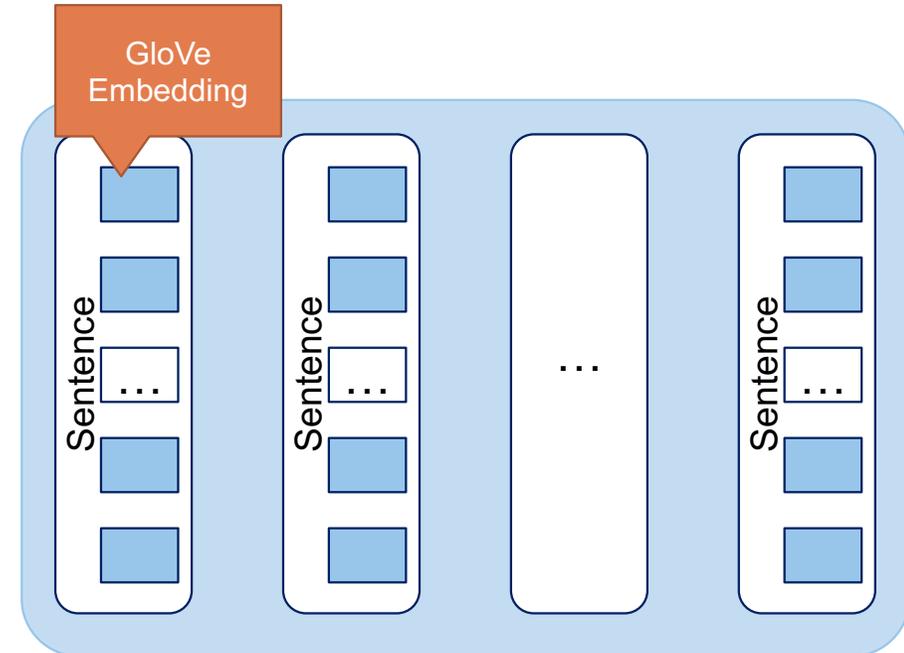
[8] [9] [6]

Outline

Motivation

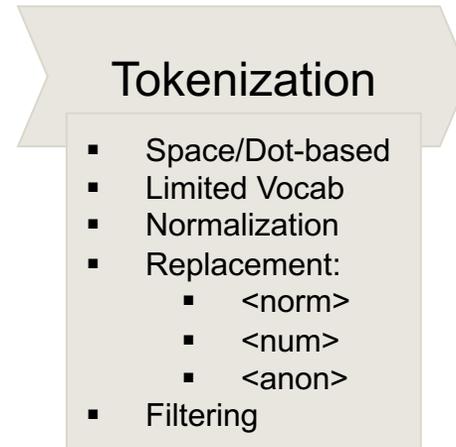Summarization & Research Questions

Literature Research

Dataset

Data Pipeline

Methods & Evaluation

Limitations & Future Work

Data Pipeline

JSON

- Title
- ID & Date
- Normchain
- Norms
- Guiding Principle
- Keywords
- Previous Instance

- Tenor
- Facts
- Reasoning

Tokenization

- Space/Dot-based
- Limited Vocab
- Normalization
- Replacement:
  - \<norm\>
  - \<num\>
  - \<anon\>
- Filtering

GloVe Embedding

Sentence Sentence … Sentence

Methods - Extractive

Cross Sentence Encoder
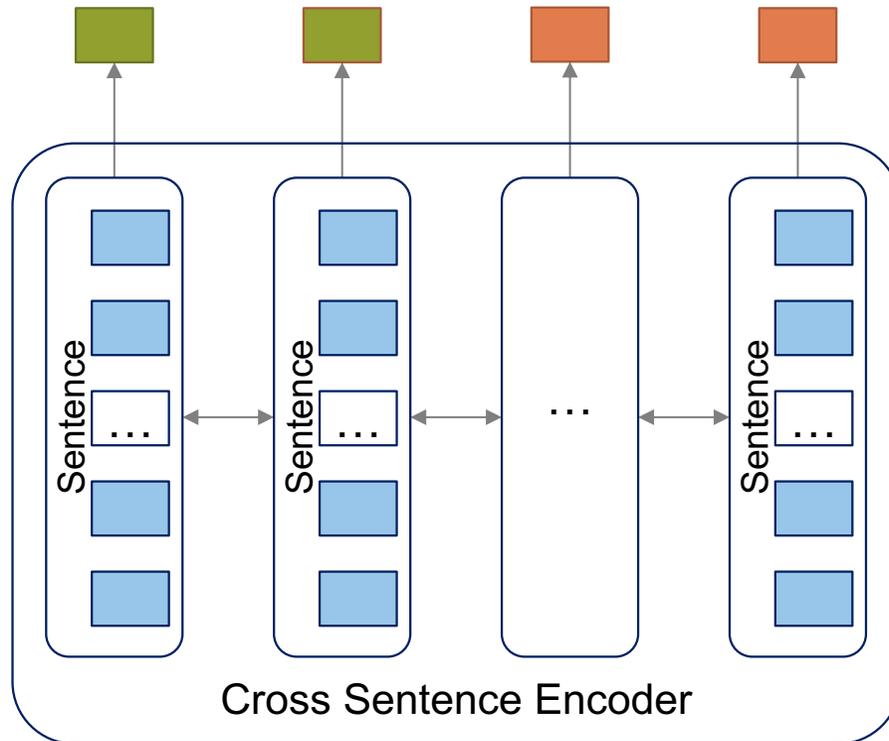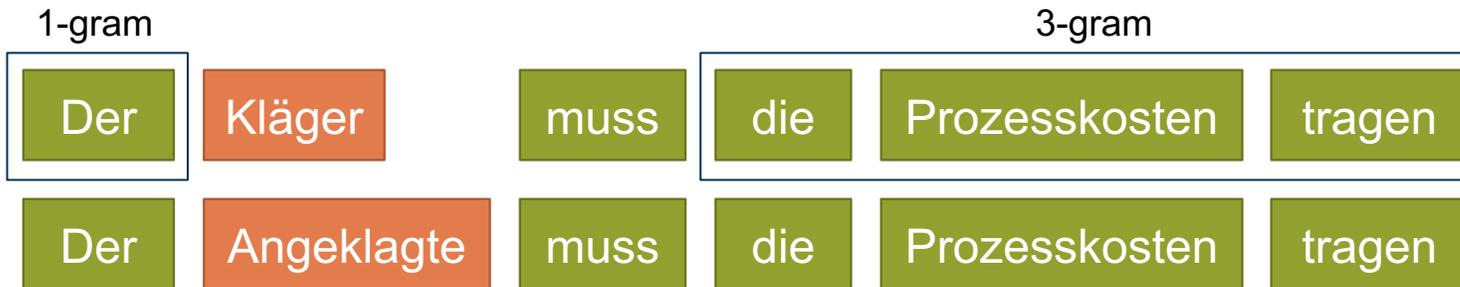
- Sentence Extraction
- Information Flow [10] [11] [12]
  - Word → Sentence: CNN, RNN or Attention
  - Sentence ↔ Sentence: None, CNN or RNN
  - Sentence → Prediction: FFN + Sigmoid
- Label approximation
  - Greedy [10]
  - One-to-One
- Optimization: Classification Loss

- Baseline
  - Lead-f-r [10]
  - Random [7]
  - Oracle [7]

Methods - Abstractive
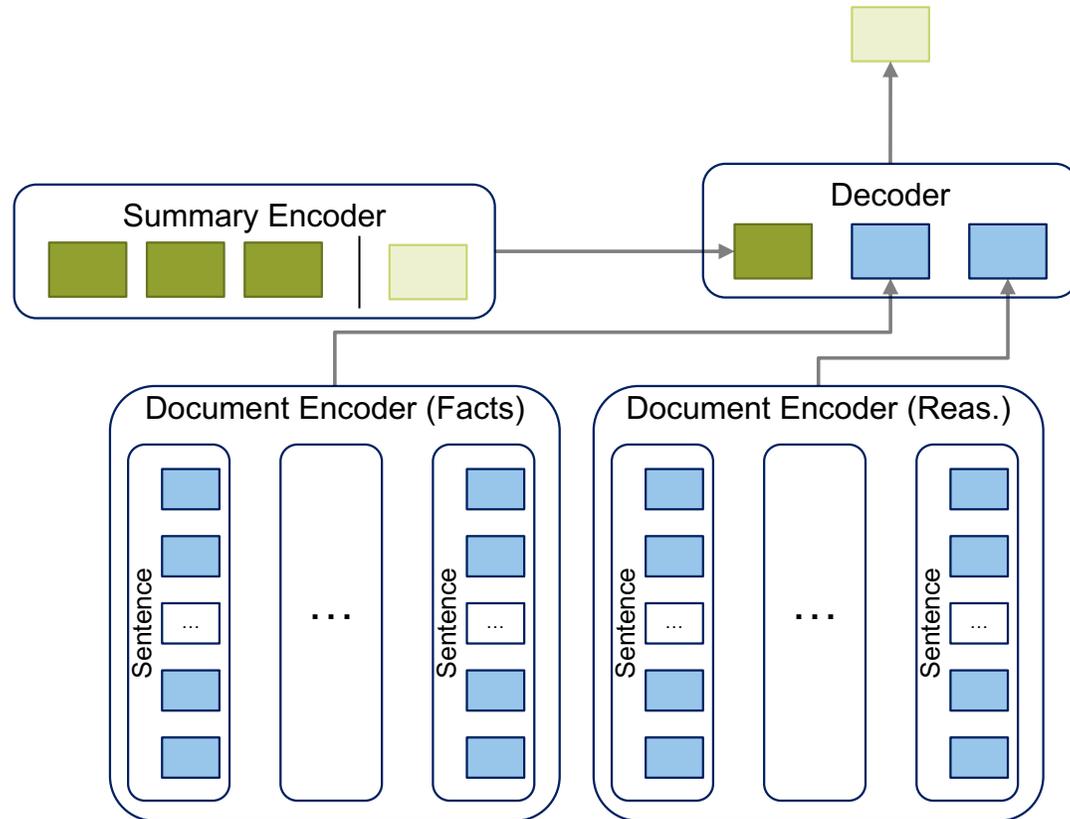
- Word-for-Word text generation
- Information Flow [13] [14] [15] [16]
    - Word → Sentence: Attention
    - Sentence → Document: RNN
    - Previous Words → Embedding: RNN
    - Decoder: FFN
- Optimization: Log-Probability correct word

- Baseline
    - Pointer [6]

Evaluation - Rouge

$$\text{Rouge-N} = \frac{\sum_{n\text{-}gram \in Summary} Match(n\text{-}gram)}{\sum_{n\text{-}gram \in Summary} Count(n\text{-}gram)} \quad [17]$$

1-gram

Der	Kläger

3-gram

muss	die	Prozesskosten	tragen

| Der	Angeklagte	muss	die	Prozesskosten	tragen

| Model	ROUGE-1	ROUGE-2	ROUGE-L	Paper / Source	Code
PEGASUS (Zhang et al., 2019)	47.21	24.56	39.25	PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization	Official
BART (Lewis et al., 2019)	45.14	22.27	37.25	BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension	Official
BertSumExtAbs (Liu et al., 2019)	38.81	16.50	31.27	Text Summarization with Pretrained Encoders	Official
T-ConvS2S	31.89	11.54	25.75	Don't Give Me the Details, Just the Summary!	Official
PtGen	29.70	9.21	23.24	Don't Give Me the Details, Just the Summary!	Official
Seq2Seq	28.42	8.77	22.48	Don't Give Me the Details, Just the Summary!	Official
PtGen-Covg	28.10	8.02	21.72	Don't Give Me the Details, Just the Summary!	Official
Baseline : Extractive Oracle	29.79	8.81	22.66	Don't Give Me the Details, Just the Summary!	Official
Baseline : Lead-3	16.30	1.60	11.95	Don't Give Me the Details, Just the Summary!	Official
Baseline : Random	15.16	1.78	11.27	Don't Give Me the Details, Just the Summary!	Official

Evaluation

Method	Greedy Labels			One-to-One Labels		
Word - Sentence	Rouge-1	Rouge-2	Rouge-L	Rouge-1	Rouge-2	Rouge-L
Attention – X	25.91	8.66	18.07	28.04	10.67	20.21
CNN – X	25.67	8.13	17.90	28.28	10.95	20.64
RNN – X	18.10	2.80	12.08	20.49	4.71	13.72
Att – CNN	23.98	6.62	15.94	18.31	2.98	12.04
Att – RNN	25.07	7.65	17.11	**30.50**	**13.06**	**22.95**
CNN – CNN	18.31	2.98	12.04	29.70	12.30	22.21
CNN – RNN	22.98	5.80	15.15	29.63	12.18	22.02
RNN – CNN	18.20	2.97	12.14	21.17	5.16	14.18
RNN – RNN	22.13	5.96	15.42*	22.05	5.88	15.30*
Mean	22.26	5.72	15.09	25.35	8.65	18.14
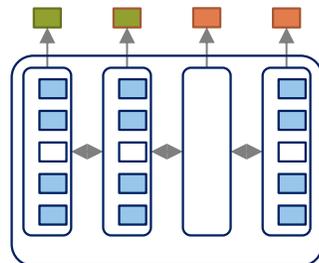
Method	Rouge-1	Rouge-2	Rouge-L
Baseline	12.37	0.77	9.96
Guided	12.37	0.77	9.96
Template	10.86	0.65	8.32

Evaluation
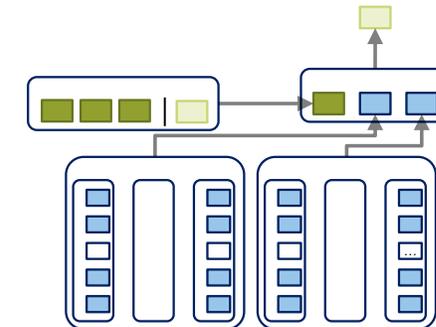
Method	Rouge-1	Rouge-2	Rouge-L
Oracle – One-to-One	55.12	40.39	51.38
Oracle – Greedy	37.90	26.10	35.16
Lead-3-0	11.80	1.52	7.75
Lead-0-3	18.44	2.88	12.38
Lead-3-3	19.69	3.18	12.42
Random Sentences	19.95	3.90	13.12
Pointer	20.80	5.27	15.98
Extr.: Att – RNN	**30.50**	**13.06**	**22.95**
Abstr.: Baseline	12.37	0.77	9.96
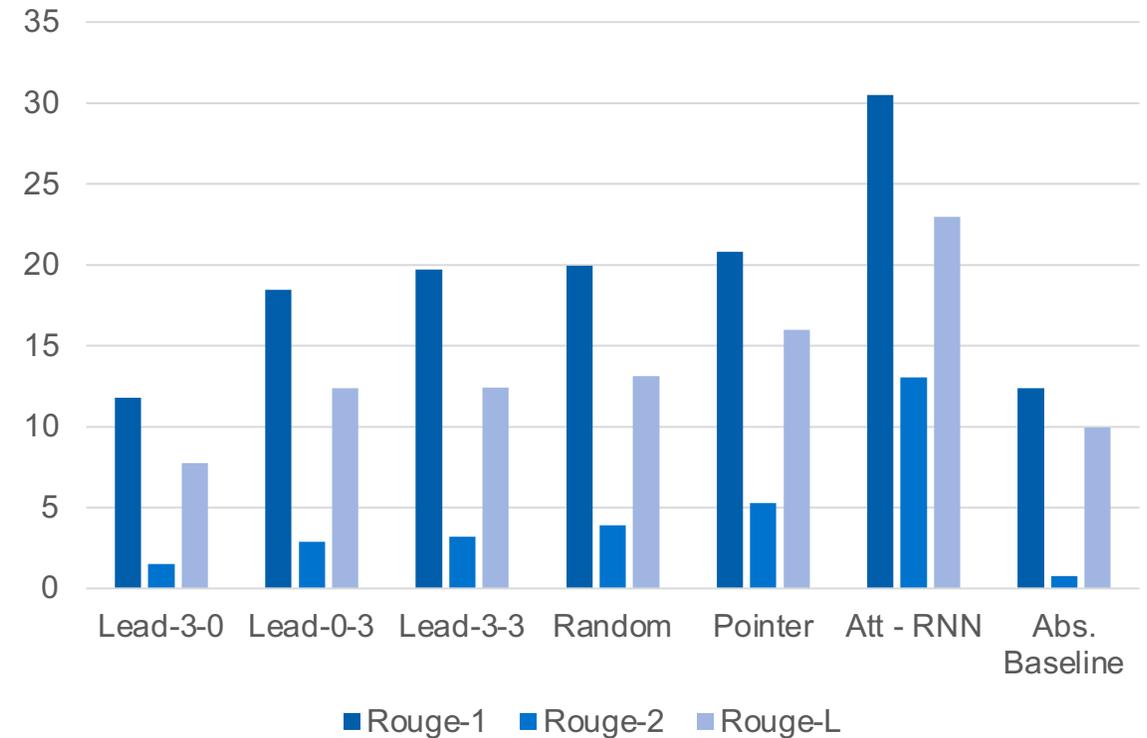
Outline

Motivation

Summarization & Research Questions

Literature Research

Dataset

Data Pipeline

Methods & Evaluation

Limitations & Future Work

Conclusion

Limitations

- Limited vocabulary + Normalization
- No qualitative evaluation
- Newer & more sophisticated Neural Network architectures

Summary

- Literature research
- Dataset with 100k court rulings
- Data Pipeline for German legal documents
- Extractive + Abstractive summarization systems

Future

Guiding Principle

regressprozess wegen einer versäumten unterhaltsabänderungsklage hat der klagende mandant dem gericht den sachverhalt vorzutragen den er als schuldner dem familiengericht im ausgangsverfahren unterbreitet hätte und aus dem sich schlüssig ergeben muss dass der titulierte unterhalt wegen unvorhergesehener wesentlicher veränderung der tatsächlichen undoder rechtlichen verhältnisse hätte herabgesetzt werden müssen

Extractive Summary

die darlegungs und beweislast im regressprozess richtet sich dann grundsätzlich nach der darlegungs und beweislast im ausgangsverfahren die rechtskraft eines unterhaltsurteils kann nicht beliebig durchbrochen werden sondern zum schutz von rechtssicherheit und klarheit gem <norm> nur dann wenn sich die tatsächlichen undoder rechtlichen verhältnisse die dem titel zugrunde liegen in unvorhergesehener weise so verändert haben dass der nach den veränderten verhältnissen zu zahlende unterhalt wesentlich unter oder über dem titulierten unterhalt liegt in diesem sinne wesentlich sind in der regel nur solche verhältnisse die zu einer veränderten unterhaltshöhe von wenigstens <num> führen

B. Sc.

Sebastian Moser

Technische Universität München
Faculty of Informatics
Chair of Software Engineering for Business
Information Systems

Boltzmannstraße 3
85748 Garching bei München

Tel   +49.89.289.   17132
Fax   +49.89.289.17136

wwwmatthes.in.tum.de

References

[1] Mehdi Allahyari, Seyedamin Pouriyeh, Mehdi Assefi, Saeid Safaei, Elizabeth Trippe, Juan Gutierrez and Krys Kochut. (2017). Text Summarization Techniques: A Brief Survey. International Journal of Advanced Computer Science and Applications (IJACSA). 8. 397-405. 10.14569/IJACSA.2017.081052.

[2] Oleksandra Klymenko, Daniel Braun and Florian Matthes. (2020). Automatic Text Summarization: A State-of-the-Art Review. In Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS 2020) - Volume 1. 648-655. 10.5220/0009723306480655.

[3] Hui Lin and Vincent Ng. (2019). Abstractive Summarization: A Survey of the State of the Art. In Proceedings of the AAAI Conference on Artificial Intelligence. 33. 9815-9822. 10.1609/aaai.v33i01.33019815.

[4] https://www.selenium.dev/

[5] Arman Cohan, Franck Dernoncourt, Doo Soon Kim, Trung Bui, Seokhwan Kim,Walter Chang, and Nazli Goharian. 2018. A discourse-aware attention model for abstractive summarization of long documents. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), New Orleans, Louisiana. Association for Computational Linguistics.

[6] Eva Sharma, Chen Li, and Lu Wang. 2019. BIGPATENT: A large-scale dataset for abstractive and coherent summarization. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy. Association for Computational Linguistics.

[7] Shashi Narayan, Shay B. Cohen, and Mirella Lapata. 2018. Don't give me the details, just the summary! topic-aware convolutional neural networks for extreme summarization. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium. Association for Computational Linguistics.

References

[8] See, Abigail & Liu, Peter J. & Manning, Christopher D. (2017). Get To The Point: Summarization with Pointer-Generator Networks. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 1073-1083. Association for Computational Linguistics. Vancouver, Canada. 10.18653/v1/P17-1099.

[9] Yen-Chun Chen and Mohit Bansal. 2018. Fast abstractive summarization with reinforce-selected sentence rewriting. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Melbourne, Australia. Association for Computational Linguistics.

[10] Ramesh Nallapati, Feifei Zhai, and Bowen Zhou. 2017. Summarunner: A recurrent neural network based sequence model for extractive summarization of documents. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI'17, page 3075-3081. AAAI Press.

[11] Jianpeng Cheng and Mirella Lapata. 2016. Neural summarization by extracting sentences and words. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Berlin, Germany. Association for Computational Linguistics.

[12] Aishwarya Jadhav and Vaibhav Rajan. 2018. Extractive summarization with SWAP-NET: Sentences and words from alternating pointer networks. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Melbourne, Australia. Association for Computational Linguistics.

[13] Ramesh Nallapati, Bowen Zhou, Cicero dos Santos, Caglar Gulcehre, and Bing Xiang. 2016. Abstractive text summarization using sequence-to-sequence RNNs and beyond. In Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning, Berlin, Germany. Association for Computational Linguistics.

[14] Sebastian Gehrmann, Yuntian Deng, and Alexander Rush. 2018. Bottom-up abstractive summarization. In Proceedings of the 2018 Conference on Empirical
Methods in Natural Language Processing, Brussels, Belgium. Association for Computational Linguistics.

References

[15] Wei Li, Xinyan Xiao, Yajuan Lyu, and Yuanzhuo Wang.2018. Improving neural abstractive document summarization with structural regularization. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium. Association for Computational Linguistics.

[16] Sumit Chopra, Michael Auli, and Alexander M. Rush. 2016. Abstractive sentence summarization with attentive recurrent neural networks. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, California. Association for Computational Linguistics.

[17] Lin, Chin-Yew. (2004). ROUGE: A Package for Automatic Evaluation of Summaries. In Text Summarization Branches Out. 74-81. Association for Computational Linguistics. Barcelona, Spain.

Literature Research

Used baselines