

# Towards Understanding and Answering Comparative Questions

---

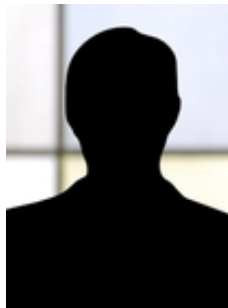
February 21–25, 2022



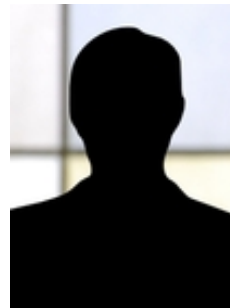
Alexander  
Bondarenko<sup>1</sup>



Yamen  
Ajjour<sup>1</sup>



Valentin  
Dittmar<sup>1</sup>



Niklas  
Homann<sup>1</sup>



Pavel  
Braslavski<sup>2,3</sup>



Matthias  
Hagen<sup>1</sup>

<sup>1</sup>Martin-Luther-Universität Halle-Wittenberg, <sup>2</sup>Ural Federal University, <sup>3</sup>HSE University

[www.webis.de](http://www.webis.de)

# Towards Understanding and Answering Comparative Questions

## Motivation

- ❑ Simple comparisons: “Did Messi or Ronaldo score more goals in 2021?”
- ❑ Life-changing and highly subjective: “Is it better to move abroad or stay?”
- ❑ For big decisions, 80% of Americans rely on online research [Turner & Rainie; 2020].
- ❑ 3% of search engine’s questions are comparative [Bondarenko et al.; WSDM’20].
- ❑ 50% of these comparative questions are non-factual [Bondarenko et al.; WSDM’20].

# Towards Understanding and Answering Comparative Questions

## Motivation

- ❑ Simple comparisons: “Did Messi or Ronaldo score more goals in 2021?”
- ❑ Life-changing and highly subjective: “Is it better to move abroad or stay?”
- ❑ For big decisions, 80% of Americans rely on online research [Turner & Rainie; 2020].
- ❑ 3% of search engine’s questions are comparative [Bondarenko et al.; WSDM’20].
- ❑ 50% of these comparative questions are non-factual [Bondarenko et al.; WSDM’20].

## Contributions

- ❑ Dataset: comparative questions with objects, aspects, and answers’ stances.
- ❑ Classifiers for comparative and subjective comparative questions.
- ❑ Identifying objects, aspects, and predicates.
- ❑ Stance detector for answers.

<https://github.com/webis-de/WSDM-22>

# Towards Understanding and Answering Comparative Questions

Dataset Webis-CompQuestions-22

Is a **cat** or a **dog** a **better** **friend**?

object 1                      object 2                      predicate                      aspect

Pro obj. 1: *Cats can be quite affectionate and attentive, and thus are good friends.*

Pro obj. 2: *Cats are less faithful than dogs.*



Source: <https://pixabay.com/images/id-2606759/>

# Towards Understanding and Answering Comparative Questions

Dataset Webis-CompQuestions-22

Is a **cat** or a **dog** a **better** **friend**?

object 1                      object 2                      predicate                      aspect

Pro obj. 1: *Cats can be quite affectionate and attentive, and thus are good friends.*

Pro obj. 2: *Cats are less faithful than dogs.*

- 31,000 questions, 3,500 comparative, 1,690 subjective from MS MARCO, Google Natural Questions, Quora, Stack Exchange.
- 950 answers (text passages) with 4 stance labels from Stack Exchange.



Source: <https://pixabay.com/images/id-2606759/>

# Towards Understanding and Answering Comparative Questions

## Comparative Question Classification

- Cascading ensemble recalls 71% of comparative questions at prec. of 1.0.
  1. 10 rules: e.g., “Is a cat **\_or\_** a dog a better **\_JJR\_** friend?” Recall 54%.
  2. Feature-based: Logistic regression with word 4-grams Recall 62%.
  3. Neural: RoBERTa, BART, SBERT for representations + DNN Recall 69%.
  4. Averaging the classifiers’ decision probabilities Recall 71%.
- Operating points (probability thresholds) chosen for precision of 1.0.
- Remove comparative questions after each classifiers’ group: more sophisticated classifiers for more difficult cases.
- 10-fold cross-validation.

# Towards Understanding and Answering Comparative Questions

## Parsing Comparative Questions

Direct:            **Is a cat or a dog a better friend?**  
                              object                    object                    predicate                    aspect

Indirect:            **What pet is the best friend?**  
                              object                                    predicate                    aspect

Without aspect:   **Who is better, a cat or a dog?**  
  predicate                    object                    object

# Towards Understanding and Answering Comparative Questions

## Parsing Comparative Questions

Direct:           **Is a cat or a dog a better friend?**  
                          object                   object           predicate           aspect

Indirect:         **What pet is the best friend?**  
                          object                                   predicate           aspect

Without aspect: **Who is better, a cat or a dog?**  
  predicate           object                   object

- 10-fold cross-validation.
- Baseline: BiLSTM with 300-dimensional GloVe embeddings [Arora et al.; CIKM'17].

	F1 scores			
Classifier	Object	Aspect	Predicate	None
BiLSTM	0.80	0.52	0.85	<b>0.98</b>
RoBERTa	<b>0.93</b>	<b>0.80</b>	<b>0.98</b>	0.94



# Towards Understanding and Answering Comparative Questions

## Answer Stance Detection

Is a **cat** or a **dog** a better friend?  
object 1                      object 2

Pro obj. 1: *Cats can be quite affectionate and attentive, and thus are good friends.*

- ❑ 4 labels: pro object 1, pro object 2, neutral, no stance.
- ❑ RoBERTa and Longformer for representations + DNN and logistic regression.
- ❑ RoBERTa and Longformer with sentiment prompts.
- ❑ Masking comparison objects.

# Towards Understanding and Answering Comparative Questions

## Answer Stance Detection

Is a **OBJECT 1** or a **OBJECT 2** a better friend?

Pro obj. 1: *OBJECT 1 can be quite affectionate and attentive, and thus are good friends.*

- ❑ Most effective classifier RoBERTa.
- ❑ Identifying subjective questions: F1 0.95.
- ❑ Comparison objects are masked in questions and answers.
- ❑ Add a sentiment prompt: *OBJECT 1 is better.*
- ❑ Input: *OBJECT 1 is better* [SEP] ANSWER.
- ❑ Highest accuracy on 4 labels (pro object 1 / 2, neutral, no stance) 0.63.

# Towards Understanding and Answering Comparative Questions

## Conclusions

- ❑ Dataset: comparative questions with objects, aspects, and answers' stances.
- ❑ Classifiers for comparative questions, objects, aspects, and predicates.
- ❑ Stance detector for potential answers.

# Towards Understanding and Answering Comparative Questions

## Conclusions

- Dataset: comparative questions with objects, aspects, and answers' stances.
- Classifiers for comparative questions, objects, aspects, and predicates.
- Stance detector for potential answers.

## Future Work

- Matching comparison objects in questions and answers.
- Improving the stance detection of comparative answers.

# Towards Understanding and Answering Comparative Questions

## Conclusions

- Dataset: comparative questions with objects, aspects, and answers' stances.
- Classifiers for comparative questions, objects, aspects, and predicates.
- Stance detector for potential answers.

## Future Work

- Matching comparison objects in questions and answers.
- Improving the stance detection of comparative answers.

SIGIR Student Author Registration Award.

*thank you!*

# References

1. Arora, Agrawal, Goyal, and Pathak. Extracting Entities of Interest from Comparative Product Reviews. Proceedings of CIKM 2017. ACM, 1975–1978.
2. Bondarenko, Braslavski, Völske, Aly, Fröbe, Panchenko, Biemann, Stein, and Hagen. Comparative Web Search Questions. Proceedings of WSDM 2020. ACM, 52–60.
3. Turner and Rainie. 2020. Most Americans Rely on Their Own Research to Make Big Decisions, and That Often Means Online Searches. Retrieved at <https://pewrsr.ch/2V07bQn>.