Is Everything in Order?
A simple way to Order Sentences

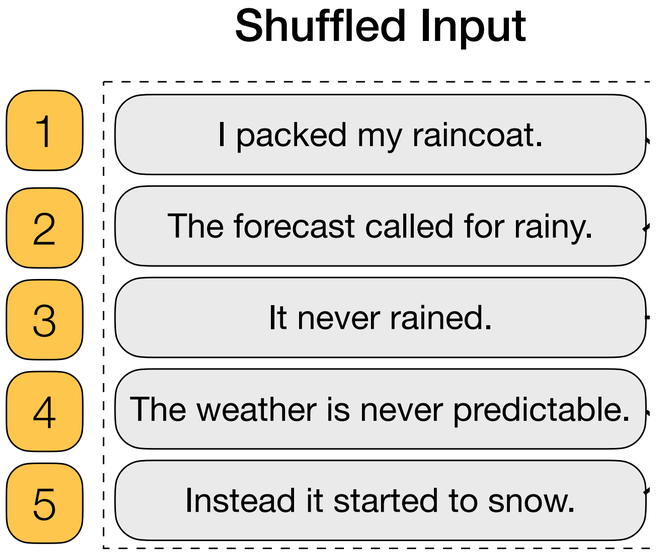
Somnath Basu Roy Chowdhury* , Faeze Brahman*
Snigdha Chaturvedi

EMNLP 2021

Motivation

Sentence Ordering Task:

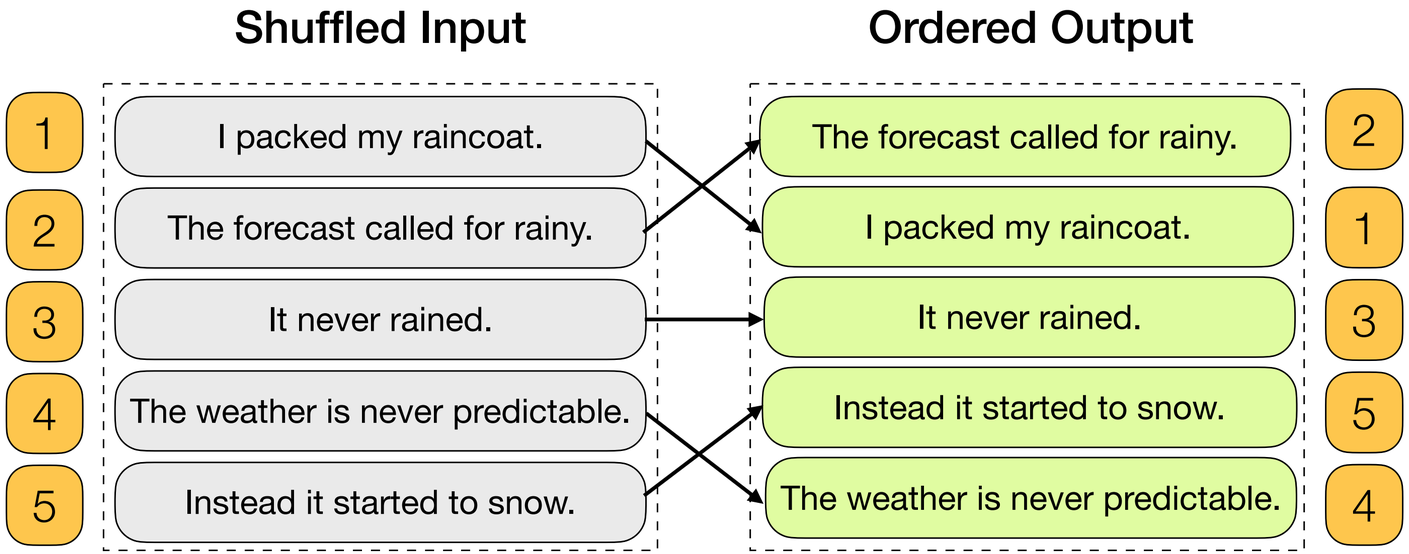
- Organizing a shuffled set of sentences into a coherent text



Motivation

Sentence Ordering Task:

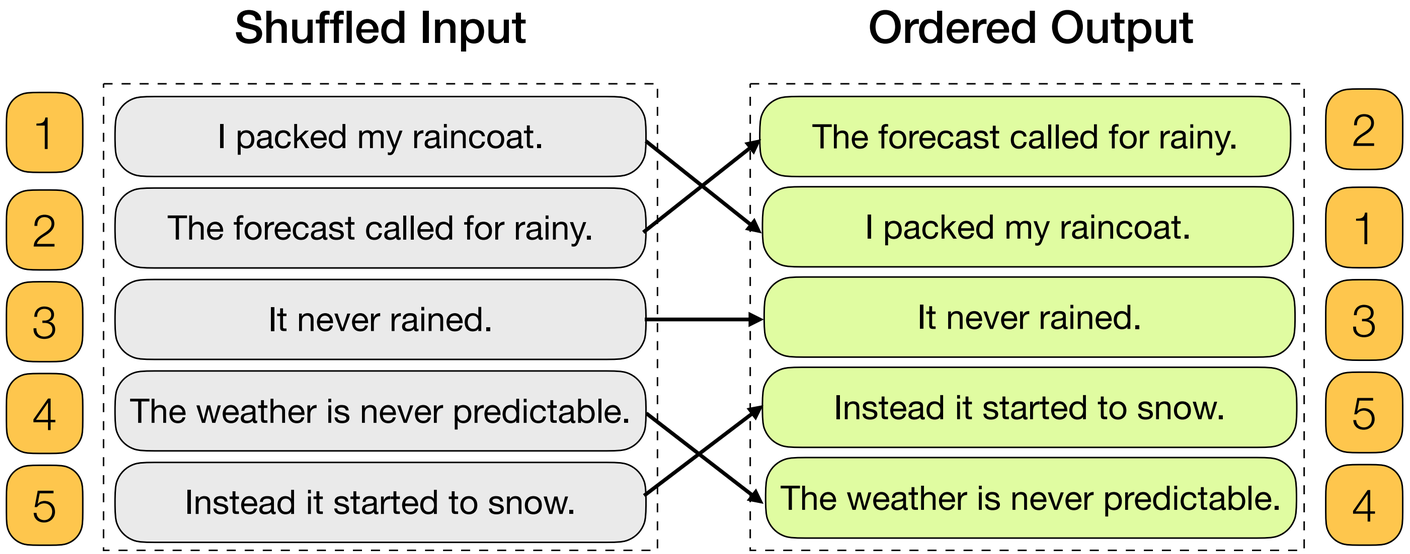
- Organizing a shuffled set of sentences into a coherent text



Motivation

Sentence Ordering Task:

- Organizing a shuffled set of sentences into a coherent text
- Requires understanding of causal and temporal relations.
- Applications in NLG, QA, etc.



Prior Works

- Pointer networks for Pairwise Ranking (Gong et al., 2016, Logeswaran et al., 2018a, Cui et al., 2018, Yin et al., 2019, 2020)
- Solving a ranking problem (Chen et. Al, 2016)
- Constraint solving + topological sorting (Prabhumoye et al., 2020)
- SOTA: Novel Pointer Decoder with Deep relational module (Cui et al., 2020)

Prior Works

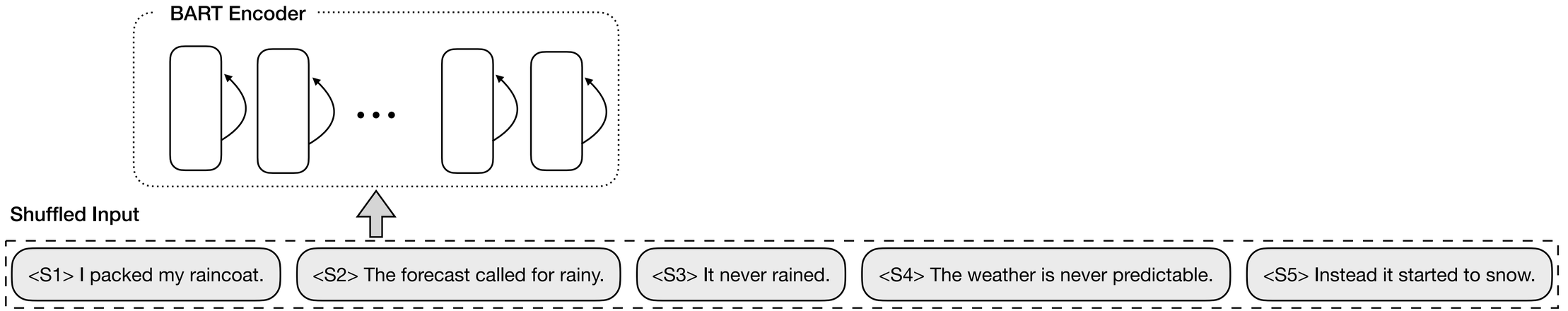
- Pointer networks for Pairwise Ranking (Gong et al., 2016, Logeswaran et al., 2018a, Cui et al., 2018, Yin et al., 2019, 2020)
- Solving a ranking problem (Chen et. Al, 2016)
- Constraint solving + topological sorting (Prabhumoye et al., 2020)
- SOTA: Novel Pointer Decoder with Deep relational module (Cui et al., 2020)

Prior Works

- Pointer networks for Pairwise Ranking (Gong et al., 2016, Logeswaran et al., 2018a, Cui et al., 2018, Yin et al., 2019, 2020)
- Solving a ranking problem (Chen et. Al, 2016)
- Constraint solving + topological sorting (Prabhumoye et al., 2020)
- SOTA: BERSON, proposed a novel Pointer Decoder with Deep relational module (Cui et al., 2020)

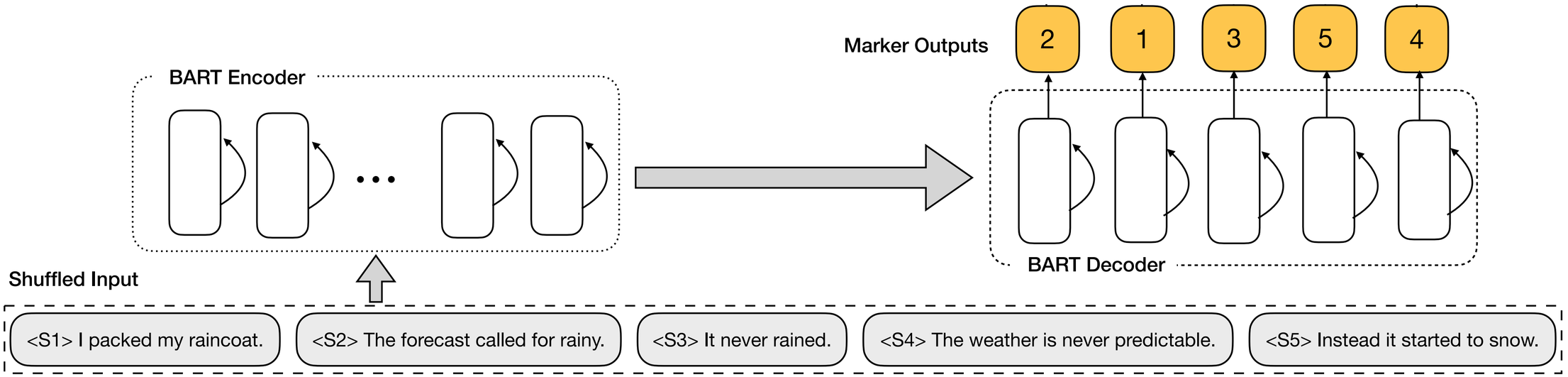
Re-order BART (Re-Bart)

- Solve the task as a conditional text-to-marker generation problem



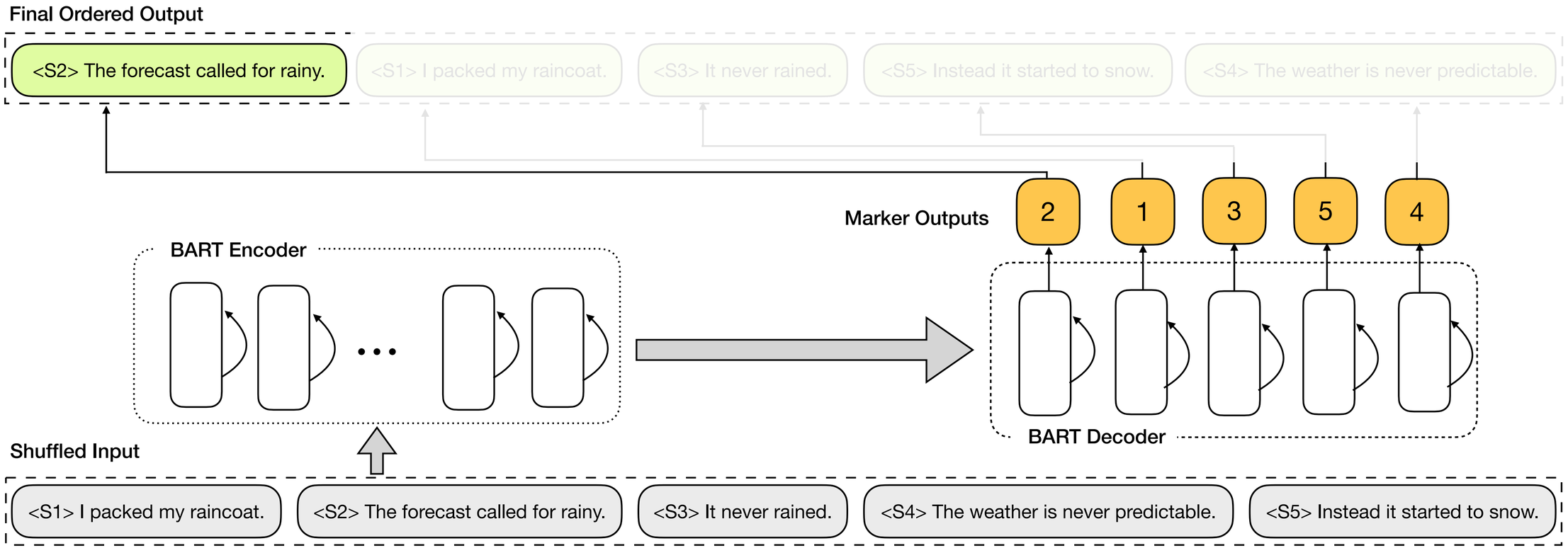
Re-order BART (Re-Bart)

- Solve the task as a conditional text-to-marker generation problem



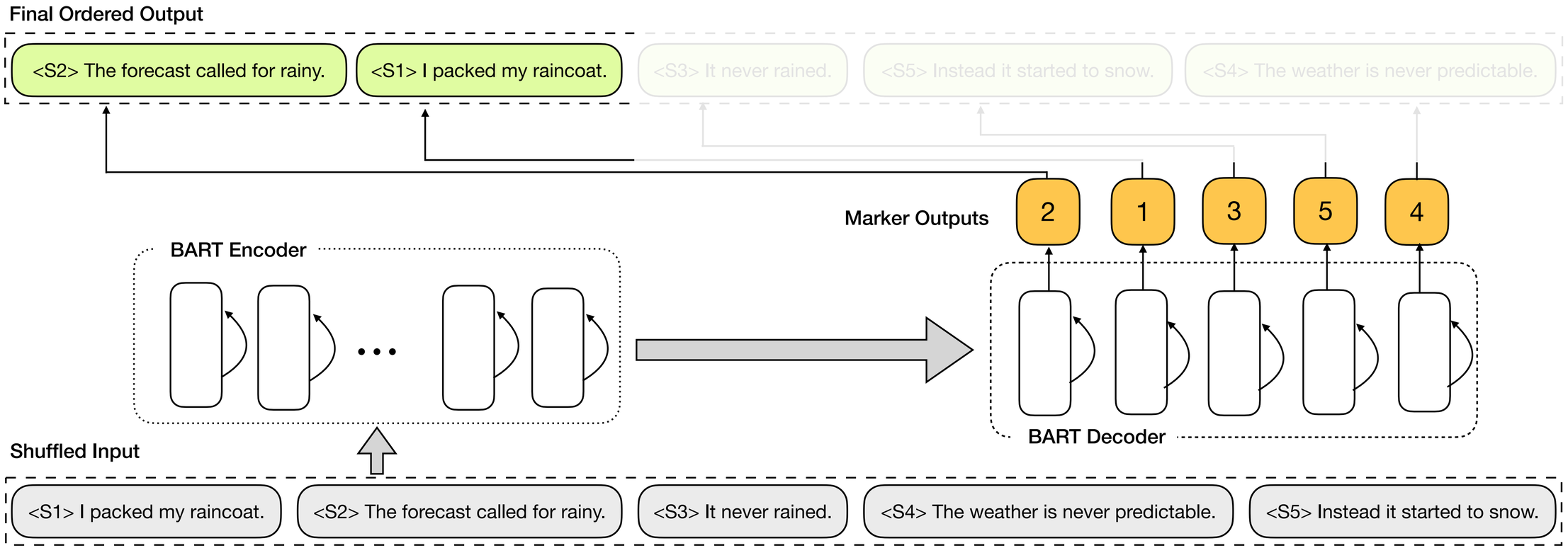
Re-order BART (Re-Bart)

- Solve the task as a conditional text-to-marker generation problem



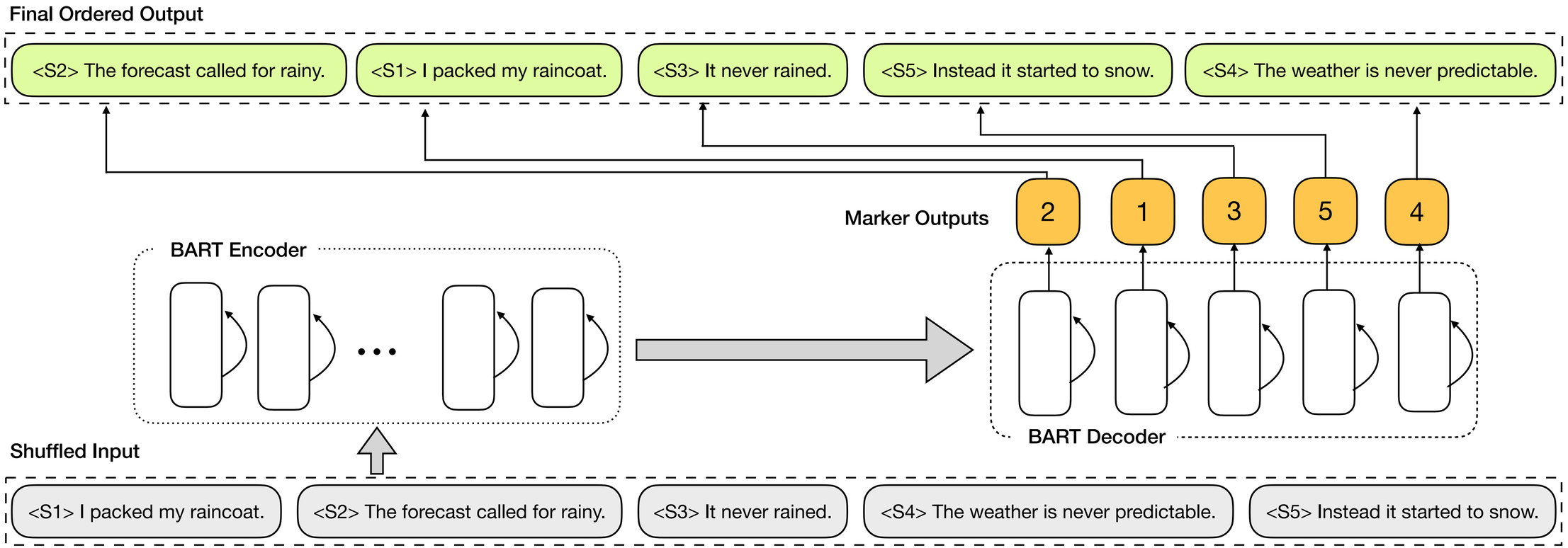
Re-order BART (Re-Bart)

- Solve the task as a conditional text-to-marker generation problem



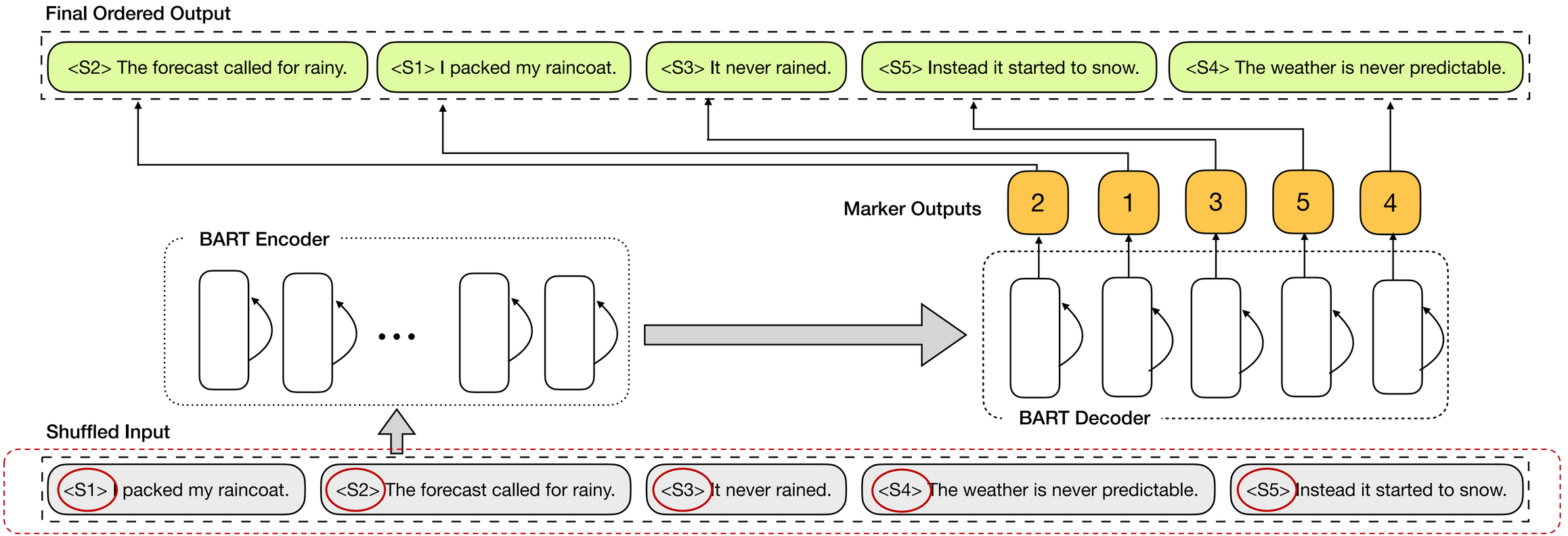
Re-order BART (Re-Bart)

- Solve the task as a conditional text-to-marker generation problem



Re-order BART (Re-Bart)

- Solve the task as a conditional text-to-marker problem



Re-order BART (Re-Bart)

What are the advantages of text-to-marker setup over text-to-text setup?

- Less susceptible to neural degeneration.
- Text-to-text often generate tokens that are not part of the input.

Re-order BART (Re-Bart)

What are the advantages of text-to-marker setup over text-to-text setup?

- Less susceptible to neural degeneration.
- Text-to-text often generate tokens that are not part of the input.

Datasets

Paper Abstracts

- NeurIPS
- AAN
- ACL
- NSF Research Awards
- arXiv

Narratives

- ROCStoris
- SIND
- Wikipedia Movie Plots

Evaluation Metrics

- **Accuracy:** The fraction of output sentence positions predicted correctly
- **Perfect Match Ratio:** The fraction of sentence orders exactly matching with the correct order
- **Kendall's Tau:** The correlations between predicted and gold order

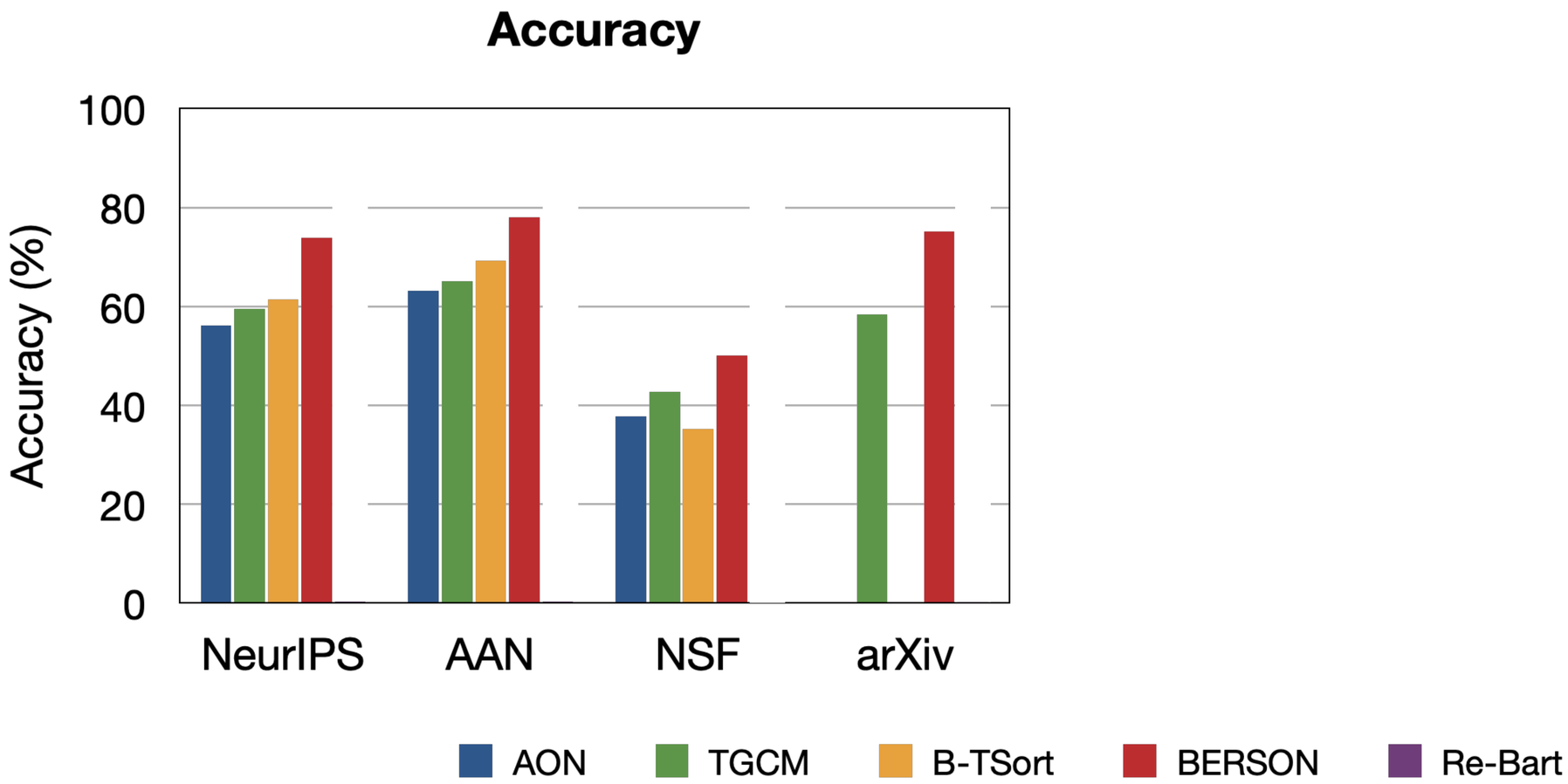
$$\tau = 1 - {2\,(\#invesrions)}\Big/{\binom{n}{2}}$$

Results

Paper Abstracts

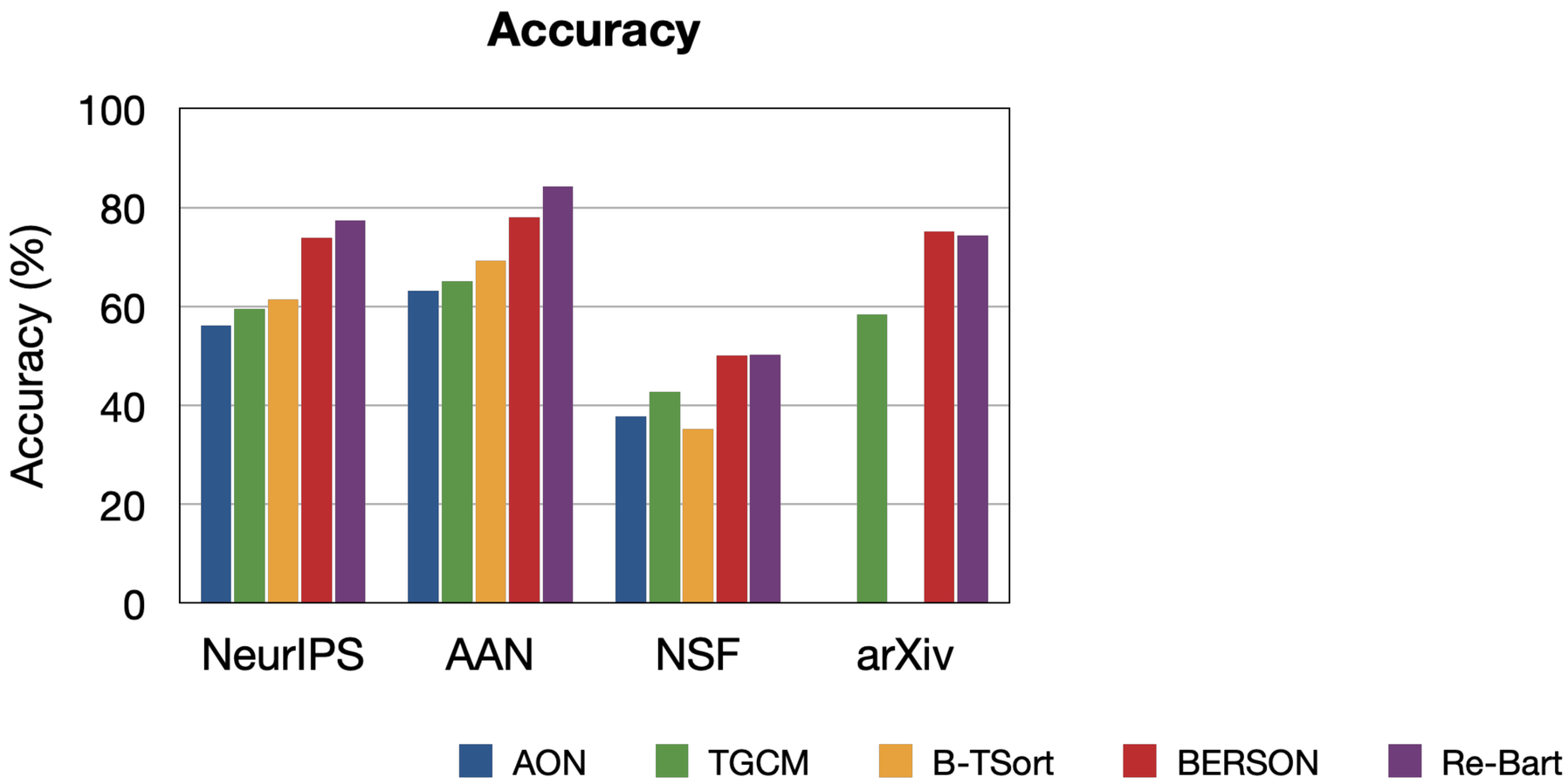
Results

Paper Abstracts



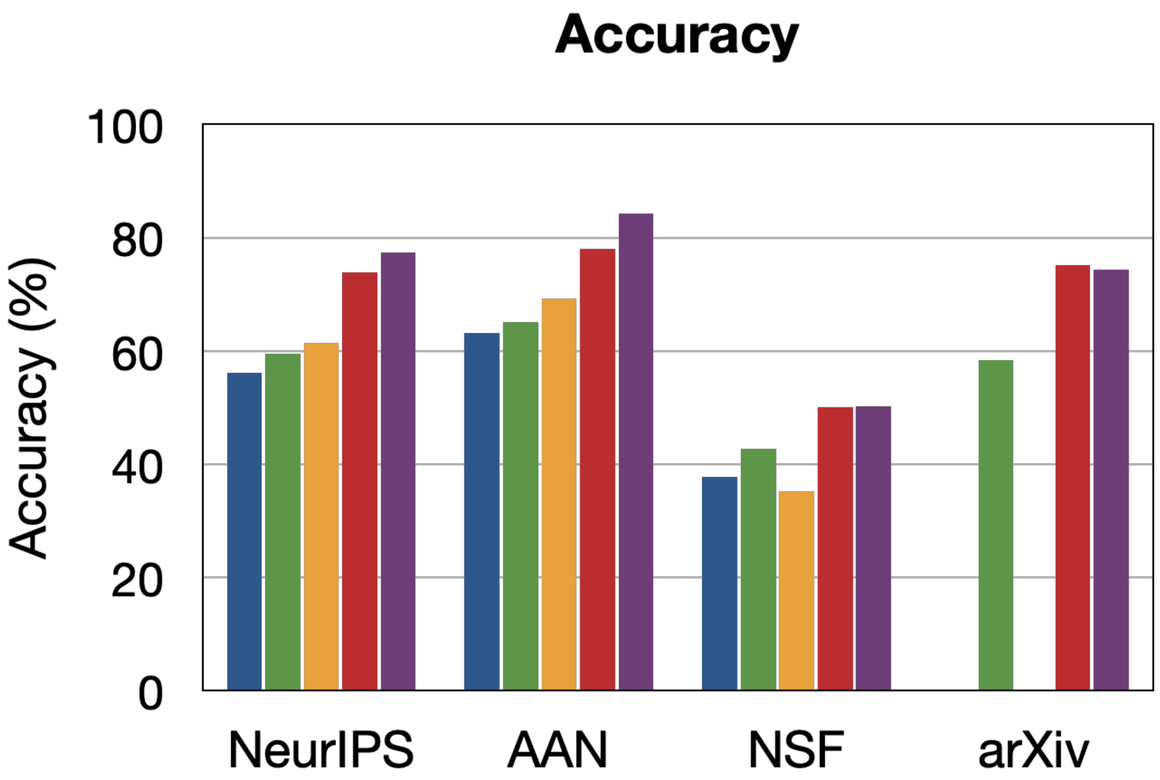
Results

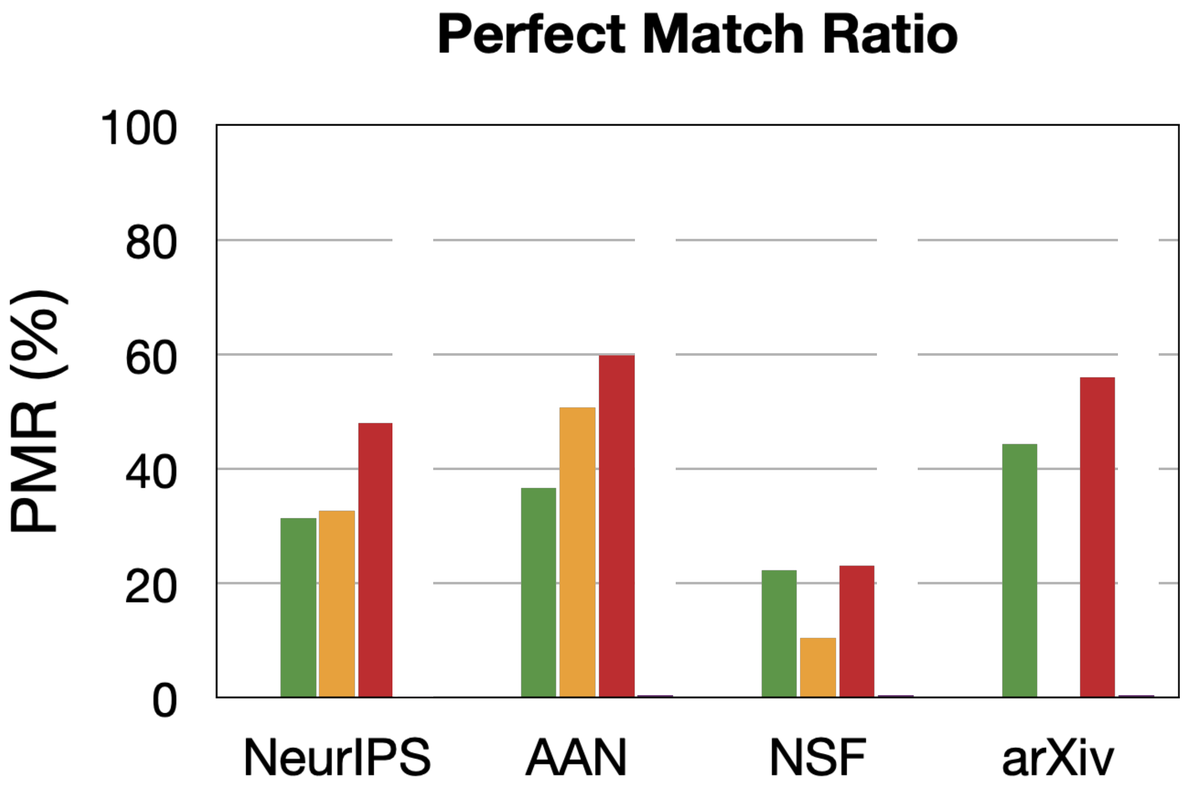
Paper Abstracts



Results

Paper Abstracts

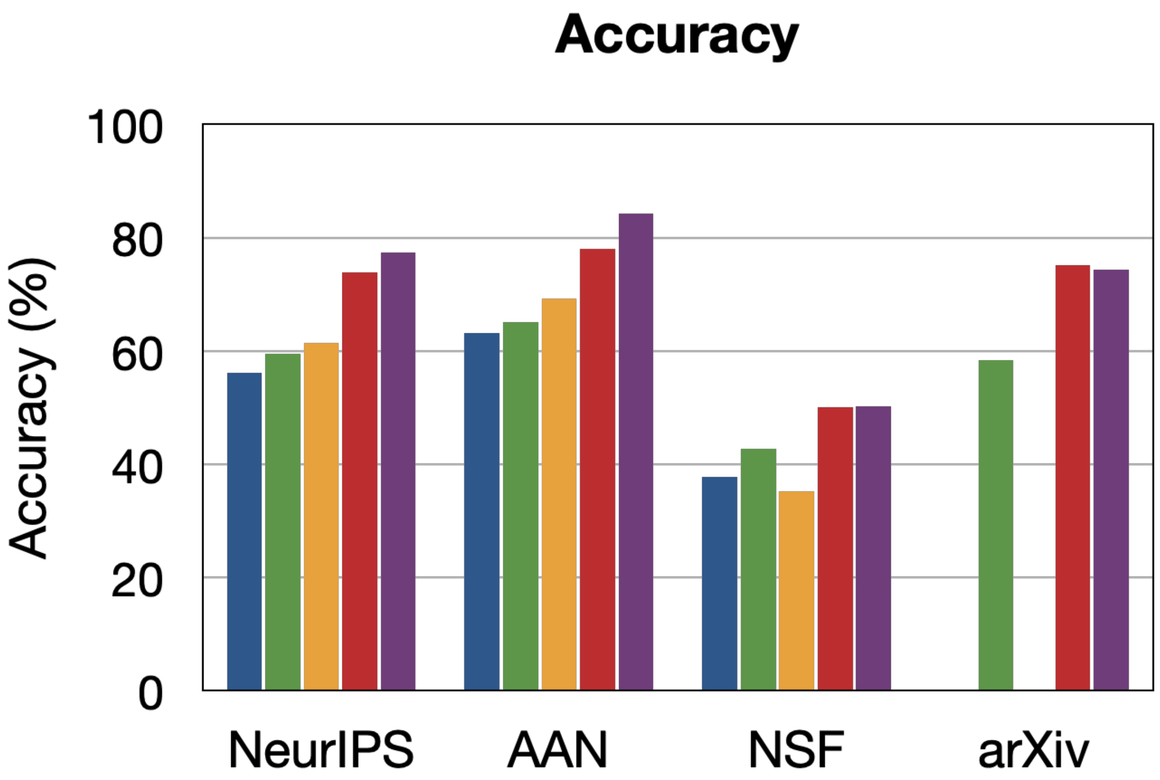


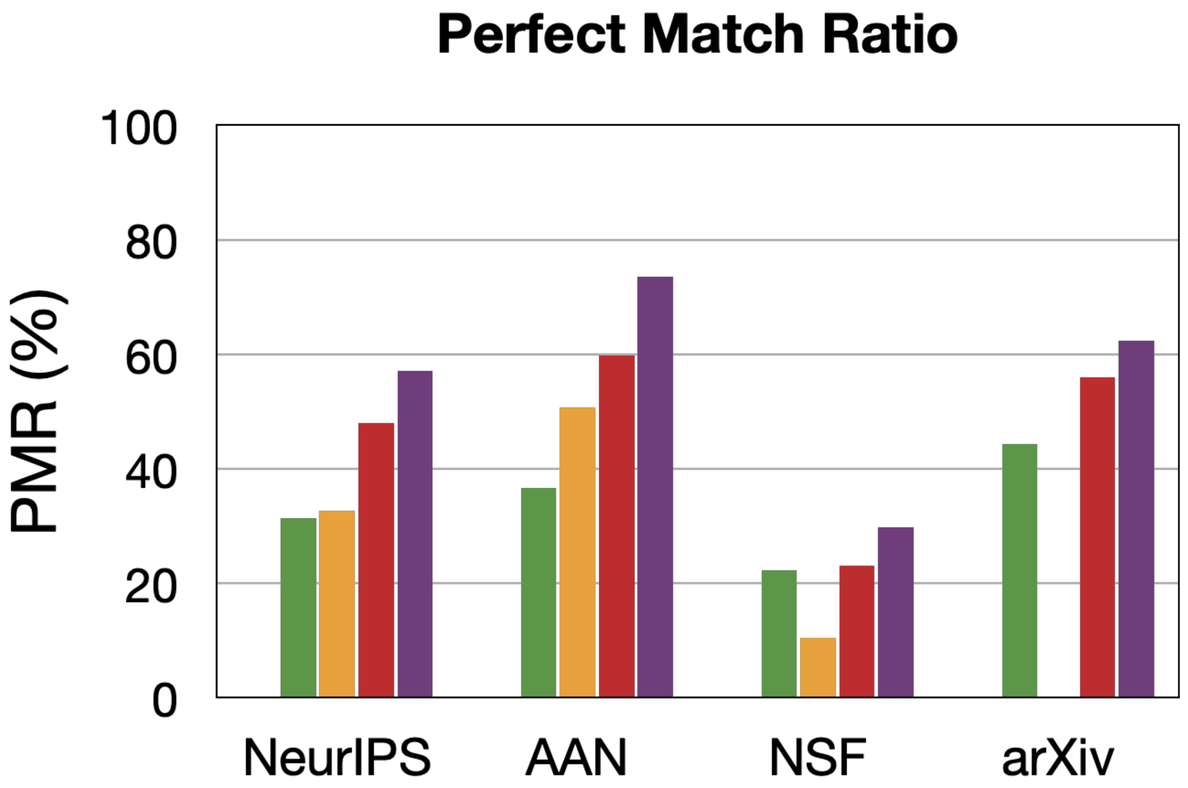


AON TGCM B-TSort BERSON Re-Bart

Results

Paper Abstracts

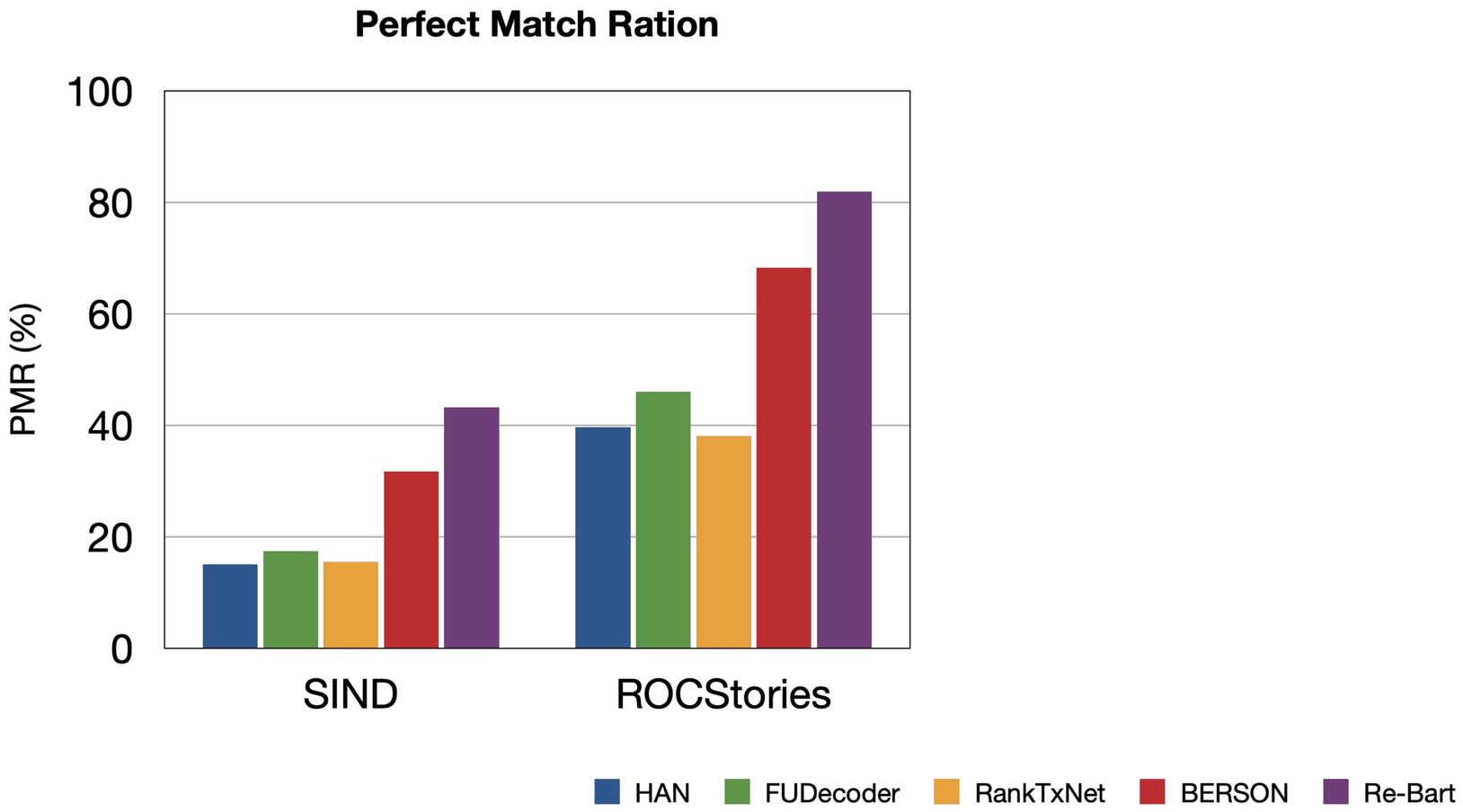




AON TGCM B-TSort BERSON Re-Bart

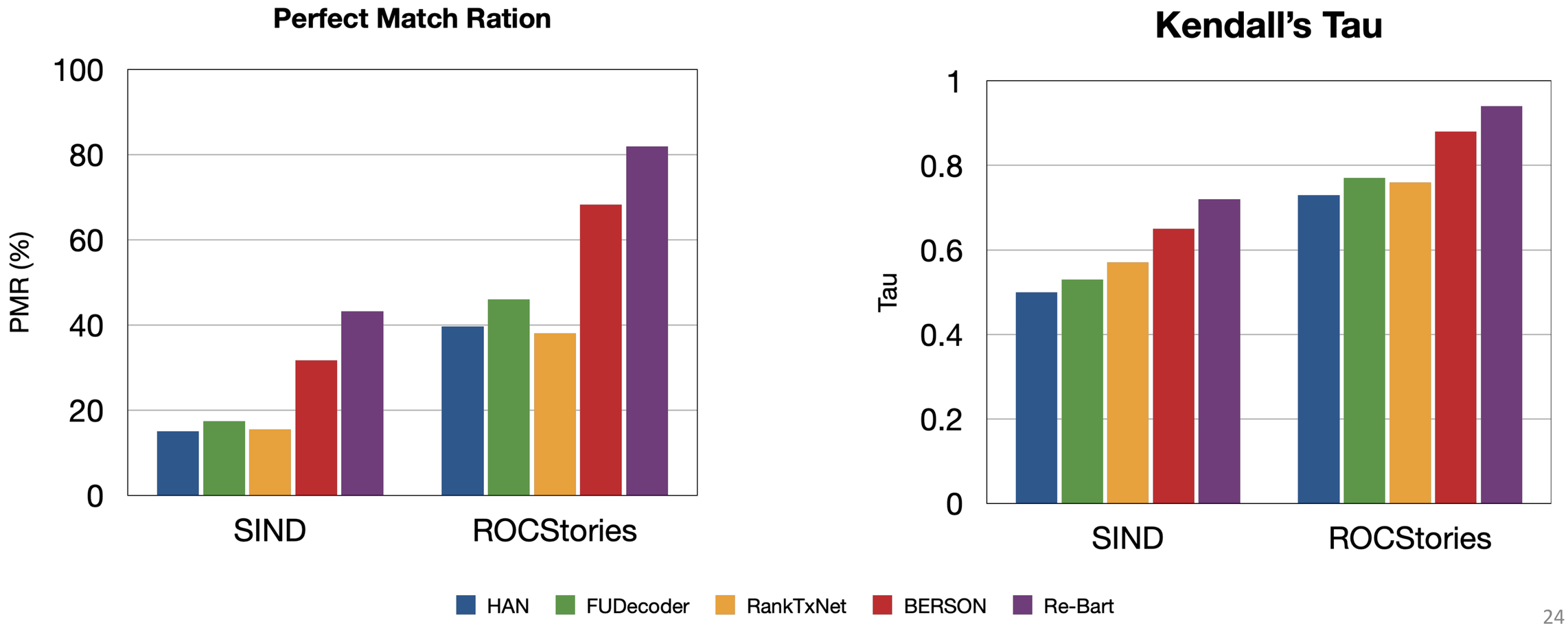
Results

Paper Abstracts



Results

Paper Abstracts

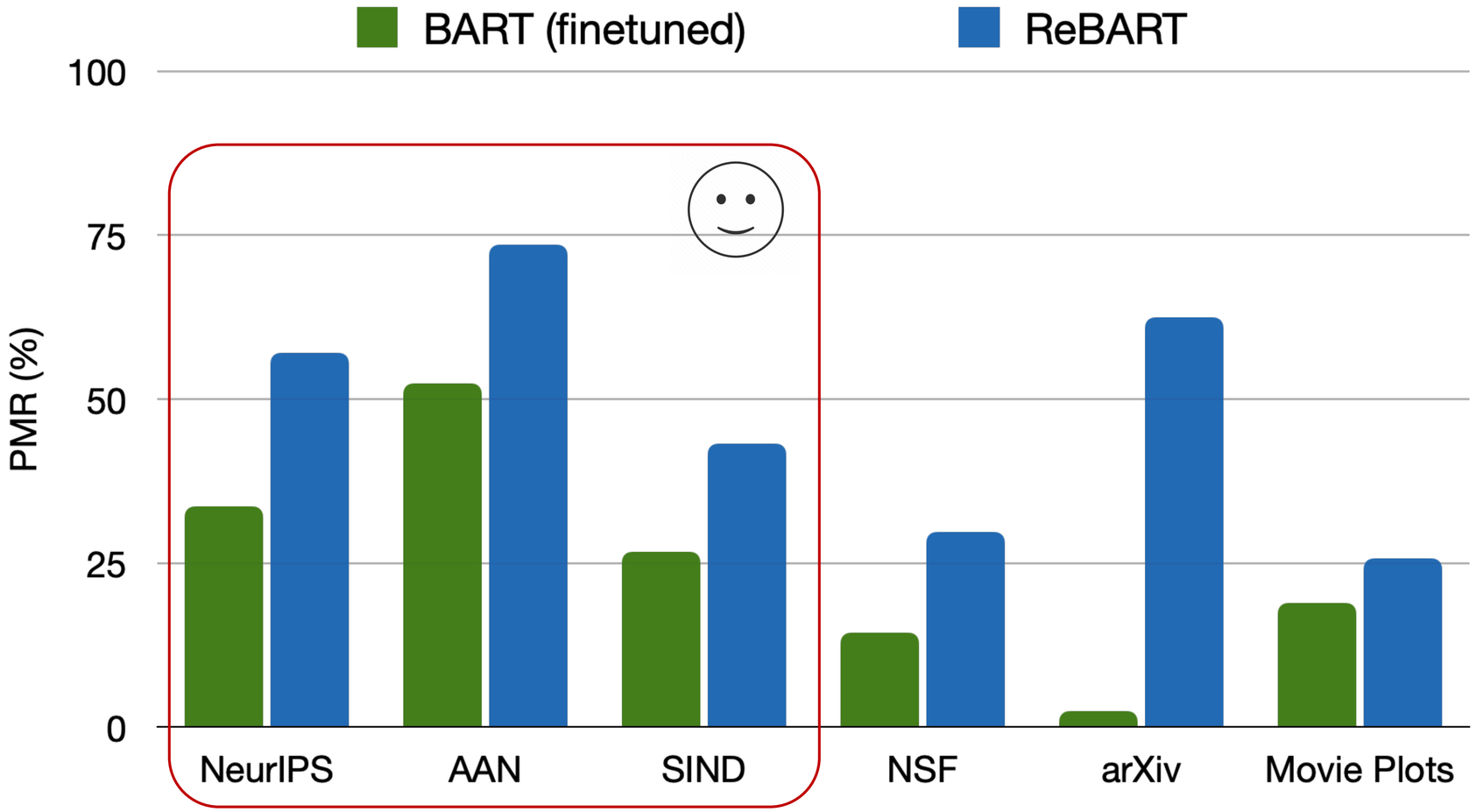


Results

How does text-to-text framework perform?

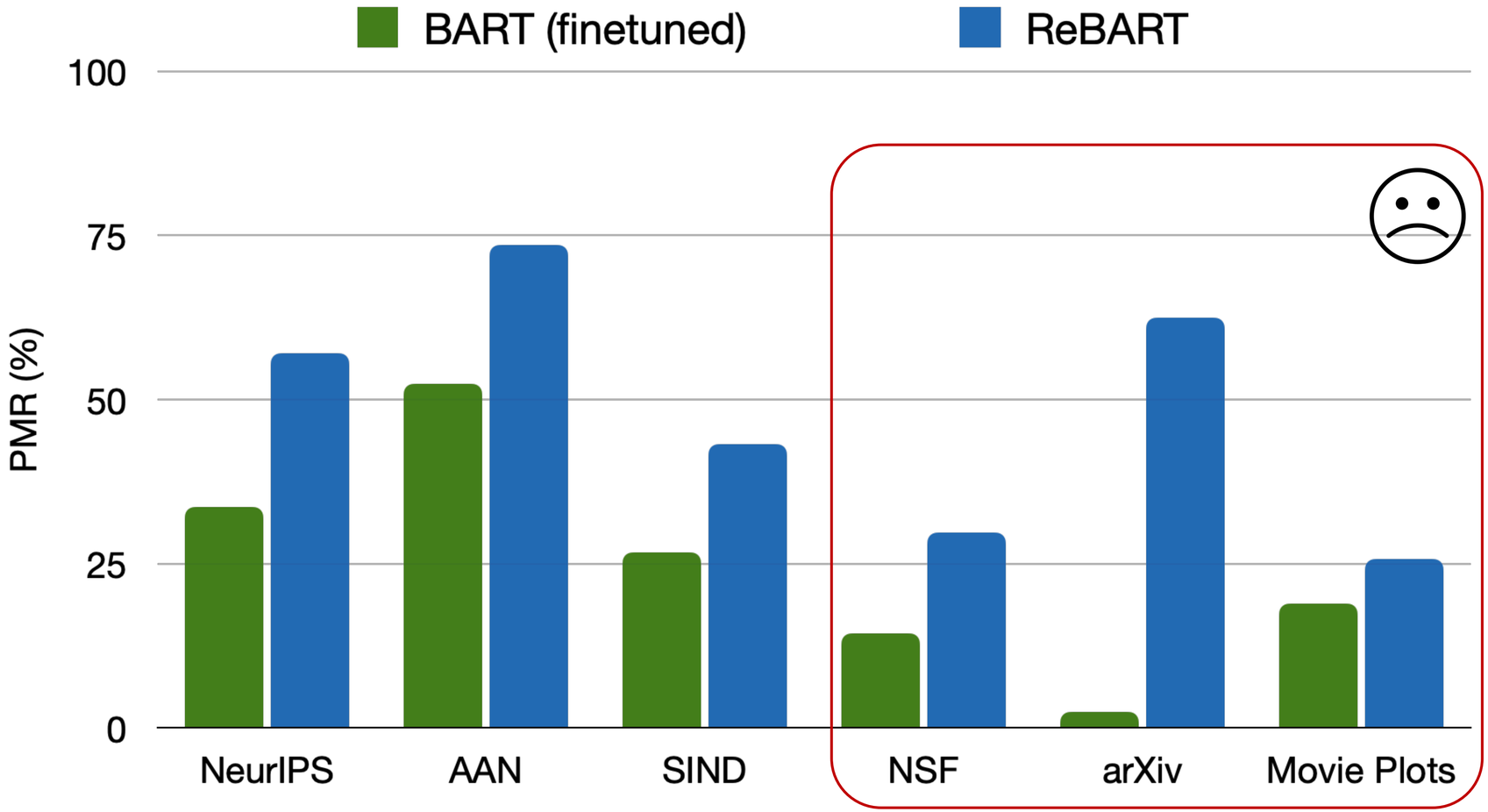
Results

How does text-to-text framework perform?

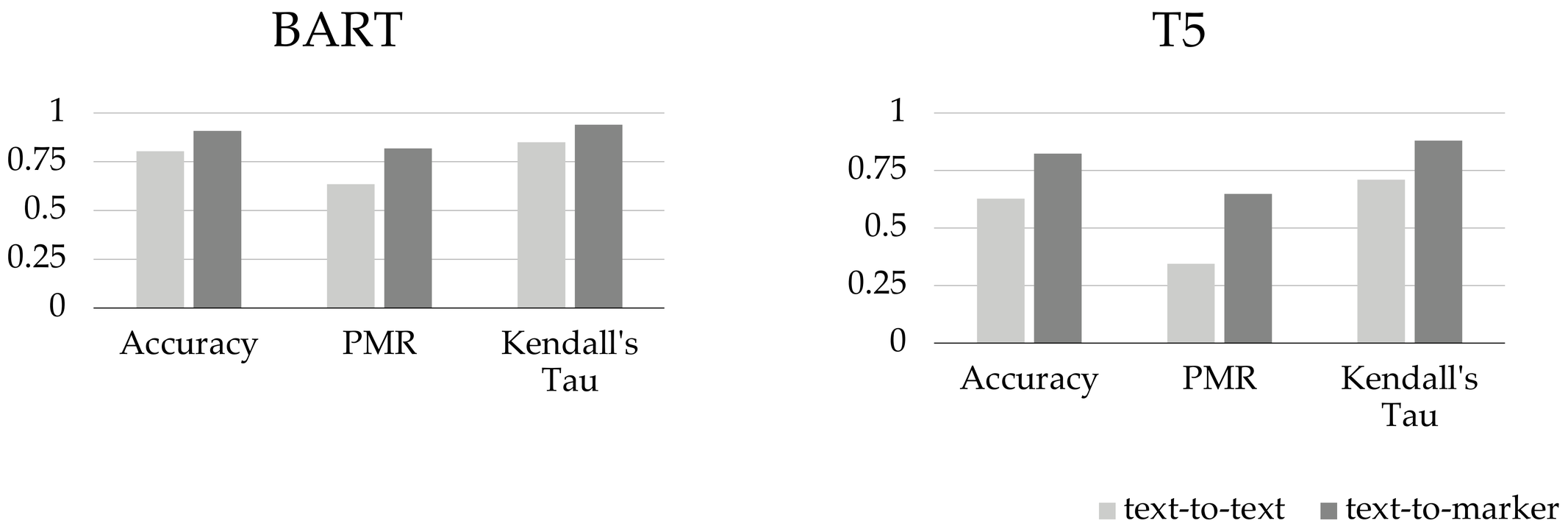


Results

How does text-to-text framework perform?

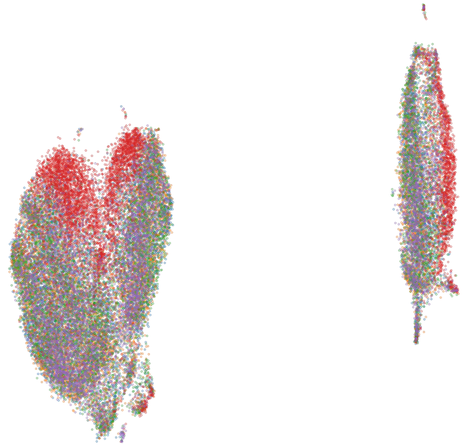


Results - BART vs T5



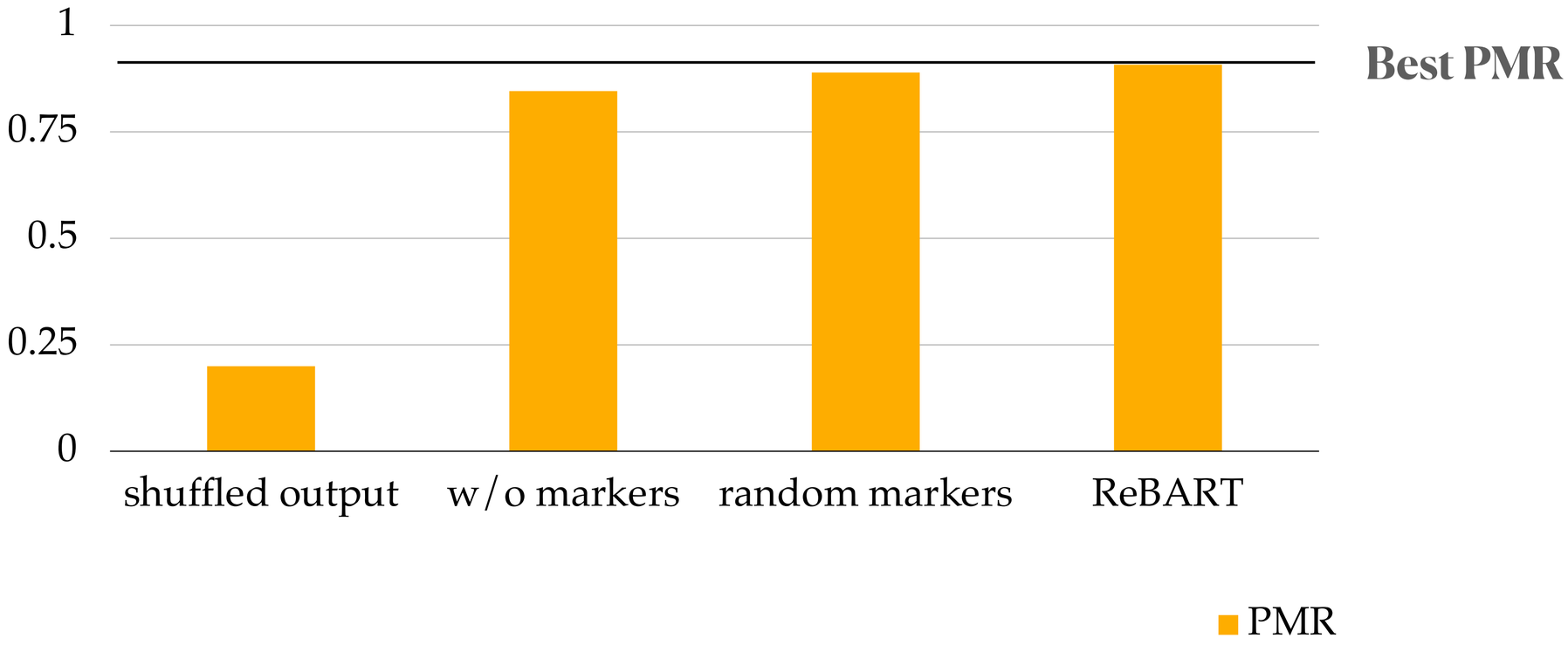
Results - BART vs T5

BART embeddings



T5 embeddings

Ablations


1
0.75
0.5
0.25
0
shuffled output
w/o markers
random markers
ReBART
Best PMR
PMR

Analysis - Effect of Shuffling

Analysis - Effect of Shuffling

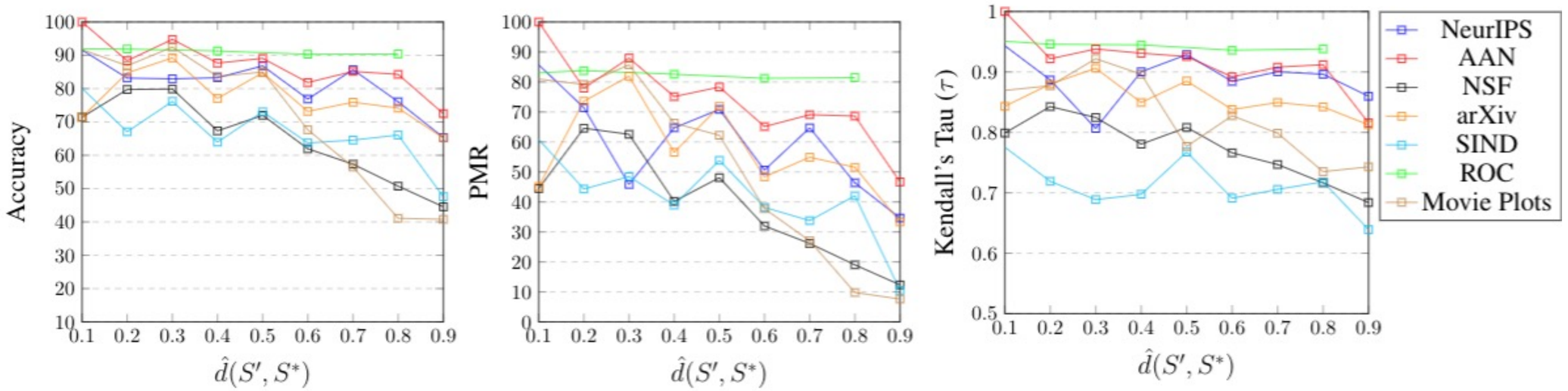
Degree of shuffling $d(S^*, S')$: Minimum number of swaps required to reconstruct ordered sequence S^* from shuffled input S'

$$\hat{d}(S^*, S') = \frac{d(S^*, S')}{|S^*|}$$

Analysis - Effect of Shuffling

Degree of shuffling $d(S^*, S')$: Minimum number of swaps required to reconstruct ordered sequence S^* from shuffled input S'

$$\hat{d}(S^*, S') = \frac{d(S^*, S')}{|S^*|}$$



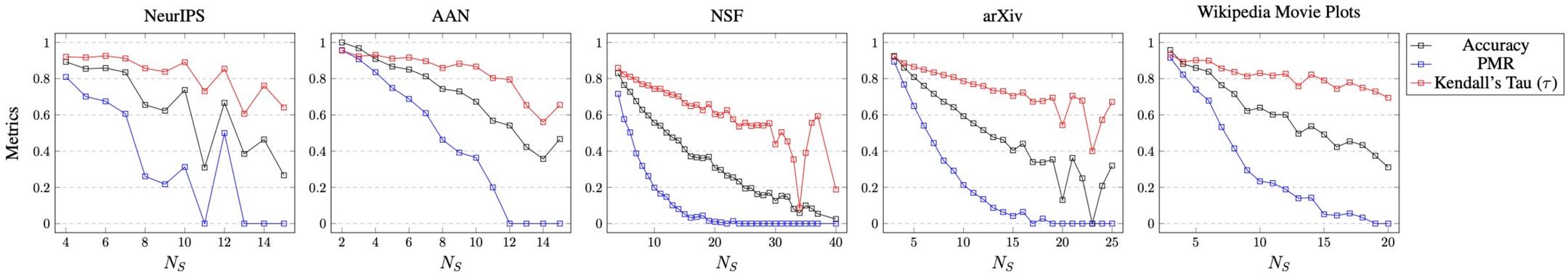
Analysis - Effect of Input Length

Analysis - Effect of Input Length

Investigate the effect on Re-BART's performance varies with change in number of input sentences in the sequence

Analysis - Effect of Input Length

Investigate the effect on Re-BART's performance varies with change in number of input sentences in the sequence



Analysis - Effect of Sentence Position

Analysis - Effect of Sentence Position

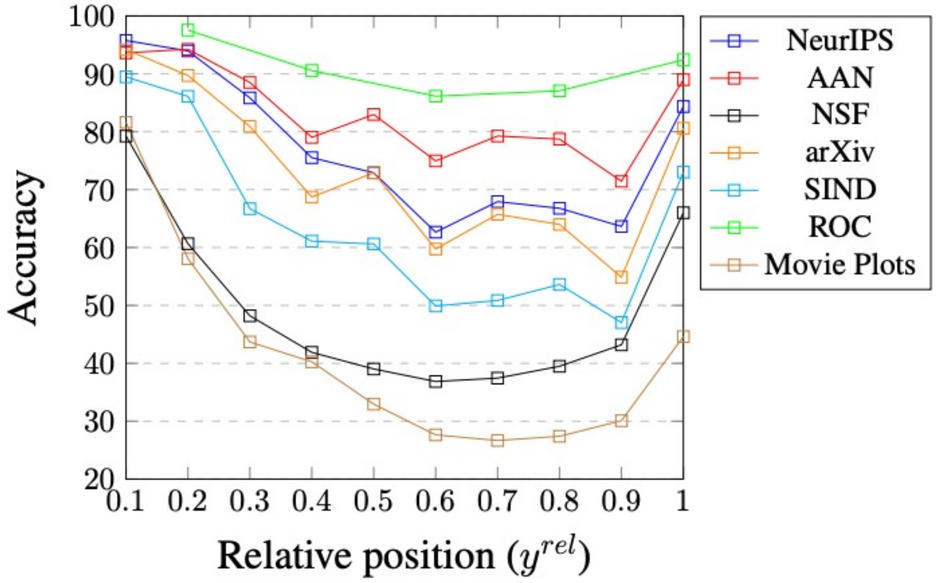
Investigate if Re-BART predict sentences at certain positions better than others

$$y^{rel} = \frac{y_i}{|S|}$$

Analysis - Effect of Sentence Position

Investigate if Re-BART predict sentences at certain positions better than others

$$y^{rel} = \frac{y_i}{|S|}$$



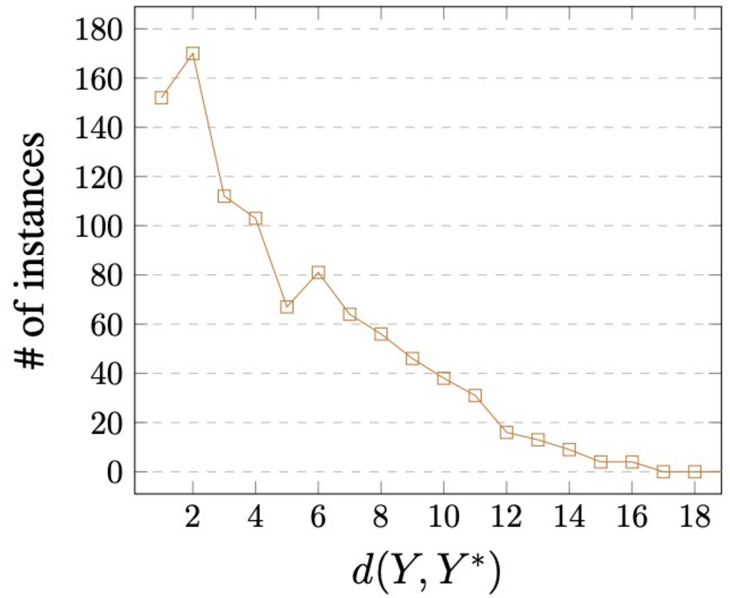
Analysis - Prediction Displacement

Analysis - Prediction Displacement

For incorrect predictions Y, investigate how far it was from the correct prediction Y^*. Computed as the minimum number of swaps $d(Y, Y^*)$.

Analysis - Prediction Displacement

For incorrect predictions Y, investigate how far it was from the correct prediction Y^*. Computed as the minimum number of swaps $d(Y, Y^*)$.



Analysis - Sentence Displacement

Analysis - Sentence Displacement

Evaluate how Re-BART’s performance varies based on shuffled sentence’s distance from its original position. Relative distance from correct position

Analysis - Sentence Displacement

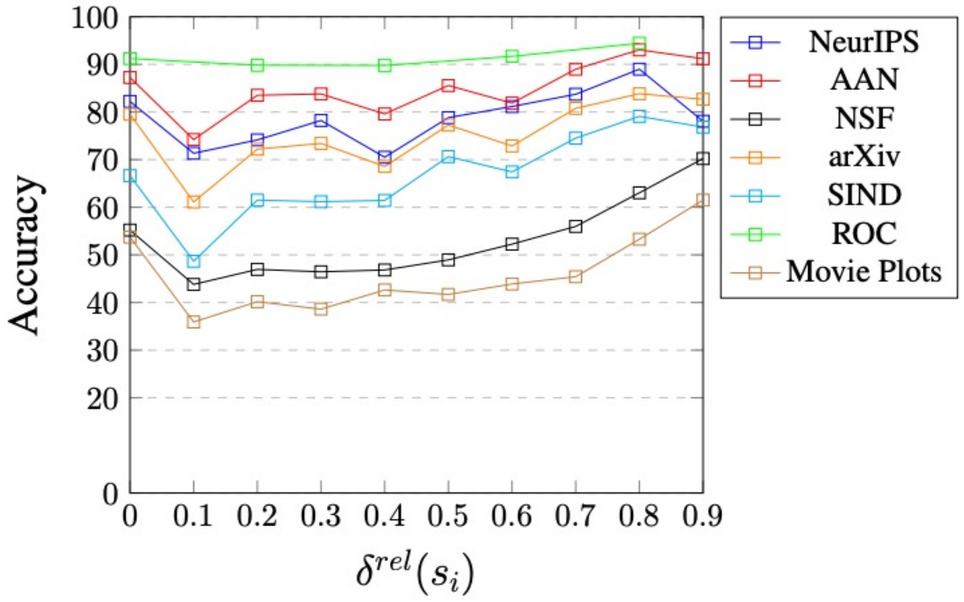
Evaluate how Re-BART's performance varies based on shuffled sentence's distance from its original position. Relative distance from correct position

$$\delta^{rel}(s_i) = \frac{|i-j|}{|S^*|} \text{ s.t. } s_i = s'_j$$

Analysis - Sentence Displacement

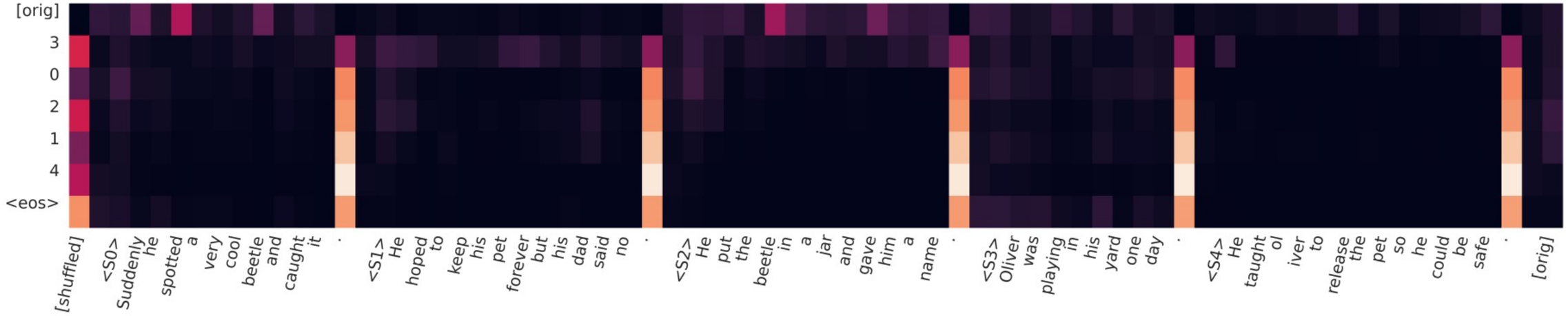
Evaluate how Re-BART's performance varies based on shuffled sentence's distance from its original position. Relative distance from correct position

$$\delta^{rel}(s_i) = \frac{|i-j|}{|S^*|} \text{ s.t. } s_i = s'_j$$

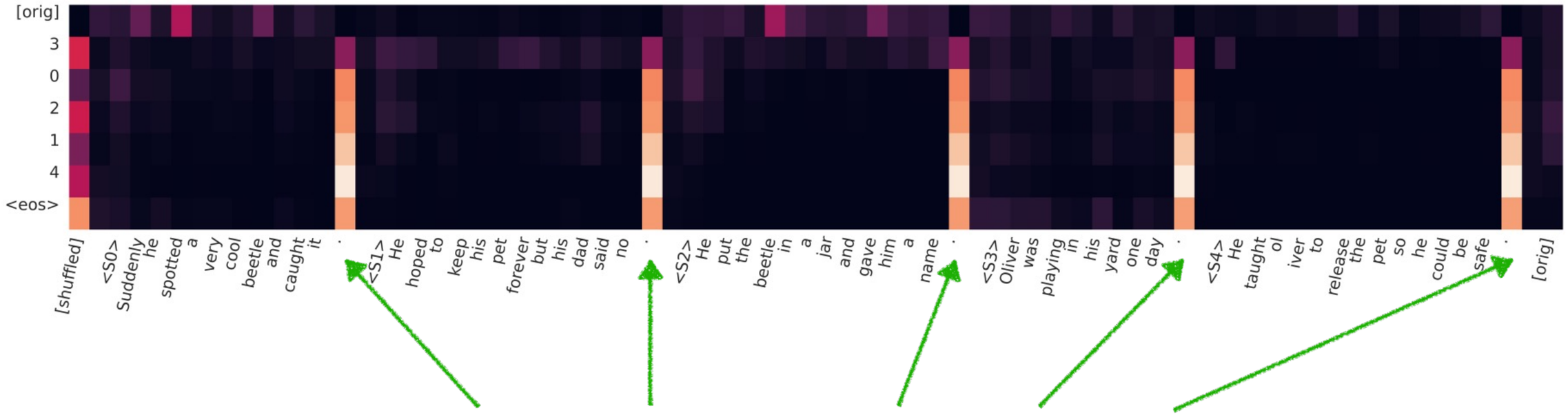


Cross-Attention Visualization

Cross-Attention Visualization



Cross-Attention Visualization



Attention is concentrated near the sentence markers

Takeaways

★ We formulate Sentence Ordering as a conditional text generation task and introduce Re-BART.

Takeaways

★ We formulate Sentence Ordering as a conditional text generation task and introduce Re-BART.

★ We show that our text-to-marker setup is better than a vanilla text-to-text framework as it reduces the prediction space

Takeaways

★ We formulate Sentence Ordering as a conditional text generation task and introduce Re-BART.

★ We show that our text-to-marker setup is better than a vanilla text-to-text framework as it reduces the prediction space

★ Re-BART achieves SOTA on 7 datasets and shows excellent performance in zero-shot setups

Takeaways

★ We formulate Sentence Ordering as a conditional text generation task and introduce Re-BART.

★ We show that our text-to-marker setup is better than a vanilla text-to-text framework as it reduces the prediction space

★ Re-BART achieves SOTA on 7 datasets and shows excellent performance in zero-shot setups

★ Re-BART is sensitive to certain factors like degree of shuffling and number of input sentences

Takeaways

★ We formulate Sentence Ordering as a conditional text generation task and introduce Re-BART.

★ We show that our text-to-marker setup is better than a vanilla text-to-text framework as it reduces the prediction space

★ Re-BART achieves SOTA on 7 datasets and shows excellent performance in zero-shot setups

★ Re-BART is sensitive to certain factors like degree of shuffling and number of input sentences

@SomnathBrc
@faeze_brh

fabrahman/ReBART

somnath@cs.unc.edu
fbrahman@ucsc.edu