Stanceformer: Target-Aware Transformer for Stance Detection

EMNLP 2022 Findings, Miami, Florida

Krishna Garg, Cornelia Caragea

University of Illinois Chicago

Stance Detection

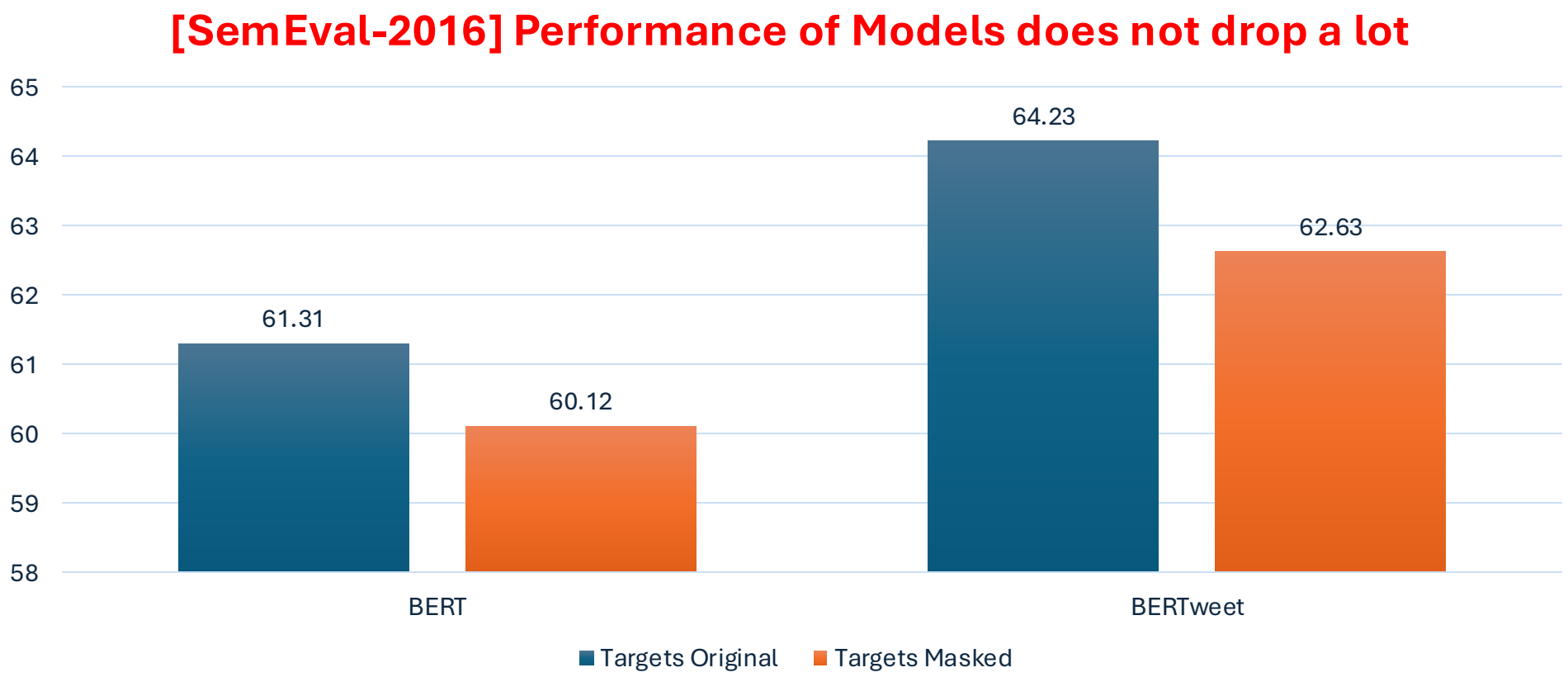
Tweet: a woman ?? wanting to be equal to a man ???! what montrosity is this

Target: feminist movement

Stance: FAVOR/ **AGAINST**/ NONE

Motivation

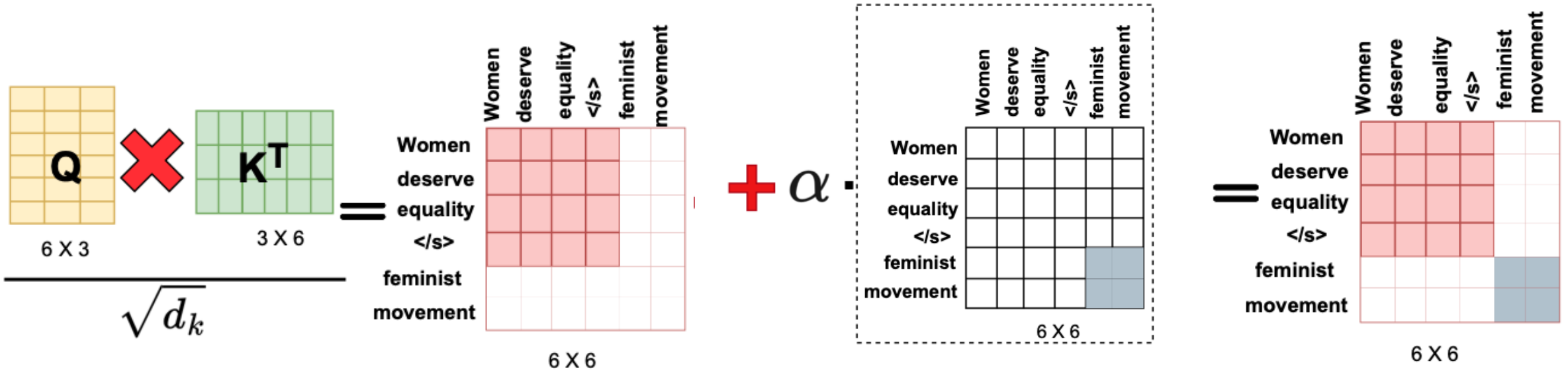
- What if the targets are hidden from the model?
  - Stance outputs remain the same -> Models tend to ignore the targets



Contributions

- Introduce Target Awareness Matrix
- Propose novel finetuning strategy Stanceformer
  - Improves traditional-BERT based as well as autoregressive LLM models
- First to finetune LLMs for Stance Detection
- Show generalization to Aspect-based Sentiment Analysis task

Architecture



Transformer

Target Awareness Matrix (M_{target})

Stanceformer

$$SA = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) * V$$

$$SA' = softmax\left(\frac{QK^T}{\sqrt{d_k}} + \alpha M_{target}\right) * V$$

Datasets

Dataset	#Train	#Val	#Test	Targets
SemEval-2016	2,160	359	1,080	Atheism, Feminist Movement, Hillary Clinton, Legalization of Abortion
COVID-19	4,533	800	800	Face Masks, Fauci, Stay at Home Orders, School Closures
P-Stance	17,224	2,193	2,157	Joe Biden, Bernie Sanders, Donald Trump
VAST-zero-shot*	13,477	1,019	1,460	drug addict, gun, constitutional right, etc.

	SemEval-2016				
	AT	FM	HC	LA	Avg.
BiCE†	64.88	57.93	58.81	60.86	57.23
CNN-based†	66.76	58.83	57.12	65.45	58.31
TAN‡	59.33	55.77	65.38	63.72	59.56
CrossNet	-	-	-	-	-
BERT‡	68.67	61.66	62.34	58.60	59.09
TGA-Net	-	-	-	-	-
BERT	65.19	55.95	63.01	61.08	61.31
-» Stanceformer	64.86	57.49	64.70	62.48	62.38
BERT-variant	68.15	60.06	65.77	62.93	64.23
-» Stanceformer	69.99	61.84	66.65	65.56	66.01
WS-BERT	70.38	63.20	71.33	62.99	66.98
-» Stanceformer	**72.01**	64.41	73.39	63.96	68.44
Closed-source LLM					
GPT-3.5 [0-shot]	24.92	69.41	73.27	57.94	56.38
Open-source LLM					
Llama-2-7b-chat [0-shot]	17.34	48.37	53.09	36.67	38.87
Llama-2-7b-chat-finetune	44.49	44.56	56.79	45.42	47.81
-» Stanceformer	49.13	48.40	55.11	40.51	48.29
Llama-2-13b-chat [0-shot]	36.92	58.18	73.78	57.01	56.47
Llama-2-13b-chat-finetune	66.11	68.64	**78.13**	67.45	70.08
-» Stanceformer	67.16	**71.43**	74.76	**73.98**	**71.83**

Baselines

1. Non-BERT variants (BiCE, CNN-based, TAN, CrossNet)
2. BERT-based models
    1. BERT
    2. WS-BERT
    3. BERT-variant (BERTweet for SemEval2016 & PStance, Covid-Twitter-BERT for Covid19, PT-HCL for VAST)
3. Autoregressive Models
    1. Closed-Source LLM [GPT-3.5]
    2. Open-Source LLM
        1. Llama-2-7b-chat [0-shot]
        2. Llama-2-13b-chat-finetune

	SemEval-2016				
	AT	FM	HC	LA	Avg.
BiCE†	64.88	57.93	58.81	60.86	57.23
CNN-based†	66.76	58.83	57.12	65.45	58.31
TAN‡	59.33	55.77	65.38	63.72	59.56
CrossNet	-	-	-	-	-
BERT‡	68.67	61.66	62.34	58.60	59.09
TGA-Net	-	-	-	-	-
BERT	65.19	55.95	63.01	61.08	61.31
-» Stanceformer	64.86	57.49	64.70	62.48	62.38
BERT-variant	68.15	60.06	65.77	62.93	64.23
-» Stanceformer	69.99	61.84	66.65	65.56	66.01
WS-BERT	70.38	63.20	71.33	62.99	66.98
-» Stanceformer	**72.01**	64.41	73.39	63.96	68.44
Closed-source LLM					
GPT-3.5 [0-shot]	24.92	69.41	73.27	57.94	56.38
Open-source LLM					
Llama-2-7b-chat [0-shot]	17.34	48.37	53.09	36.67	38.87
Llama-2-7b-chat-finetune	44.49	44.56	56.79	45.42	47.81
-» Stanceformer	49.13	48.40	55.11	40.51	48.29
Llama-2-13b-chat [0-shot]	36.92	58.18	73.78	57.01	56.47
Llama-2-13b-chat-finetune	66.11	68.64	**78.13**	67.45	70.08
-» Stanceformer	67.16	**71.43**	74.76	**73.98**	**71.83**

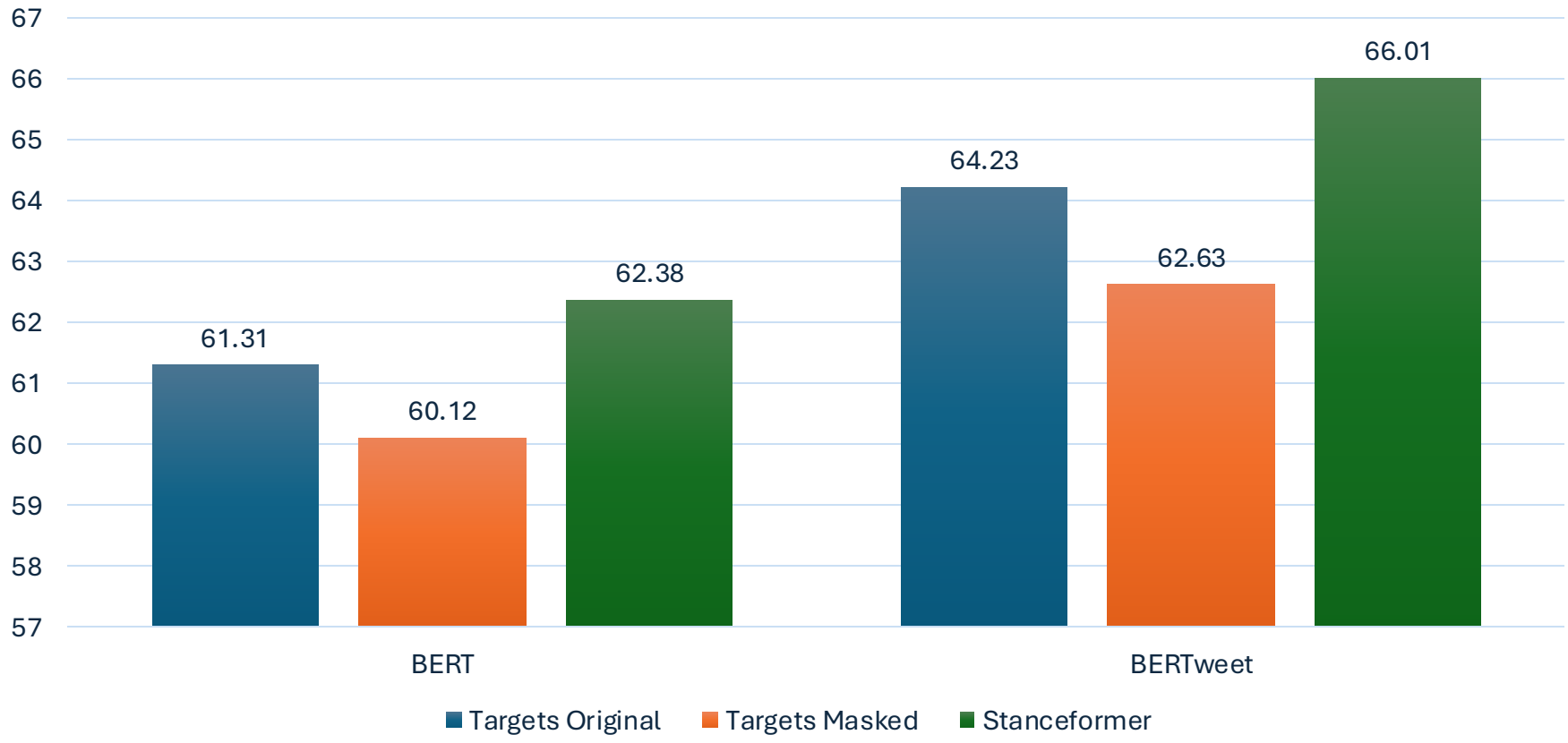
Results

1. Stanceformer outperforms BERT-based models
2. Stanceformer outperforms Llama-based models
3. Llama-2-13b outperforms Llama-2-7b
4. Finetuned LLMs outperform zero-shot
5. LLM models often do not outperform BERT-based models [Watch out for Challenges with LLMs]

	SemEval-2016					Covid19					PStance			
	AT	FM	HC	LA	Avg.	mask	fauci	home	school	Avg.	trump	biden	sanders	Avg.
BiCE†	64.88	57.93	58.81	60.86	57.23	56.70	63.00	64.50	54.80	59.75	77.15	77.69	71.24	75.36
CNN-based†	66.76	58.83	57.12	65.45	58.31	59.90	61.20	52.10	52.70	56.48	76.80	77.22	71.40	75.14
TAN‡	59.33	55.77	65.38	63.72	59.56	54.60	54.70	53.60	53.40	54.08	77.10	77.64	71.60	75.45
CrossNet	-	-	-	-	-	-	-	-	-	66.16	48.60	47.10	40.80	45.50
BERT‡	68.67	61.66	62.34	58.60	59.09	-	-	-	-	68.71	79.19	76.02	73.59	76.27
TGA-Net	-	-	-	-	-	-	-	-	-	69.09	-	-	-	77.66
BERT	65.19	55.95	63.01	61.08	61.31	71.13	72.52	77.60	61.77	70.76	79.81	79.34	76.61	78.59
-» Stanceformer	64.86	57.49	64.70	62.48	62.38	71.40	72.36	78.66	64.19	71.65	79.87	81.13	75.65	78.88
BERT-variant	68.15	60.06	65.77	62.93	64.23	79.23	81.25	83.16	**85.32**	82.24	80.92	81.24	75.91	79.36
-» Stanceformer	69.99	61.84	66.65	65.56	66.01	81.49	83.43	**87.49**	80.23	83.16	82.75	81.44	77.57	80.59
WS-BERT	70.38	63.20	71.33	62.99	66.98	82.59	82.48	84.53	81.09	82.67	84.97	82.86	79.97	82.60
-» Stanceformer	**72.01**	64.41	73.39	63.96	68.44	**85.10**	**83.79**	85.44	81.86	**84.05**	**85.35**	**83.96**	**80.57**	**83.30**
Closed-source LLM														
GPT-3.5 [0-shot]	24.92	69.41	73.27	57.94	56.38	76.90	73.03	72.81	50.96	68.42	79.80	79.65	77.77	79.07
Open-source LLM														
Llama-2-7b-chat [0-shot]	17.34	48.37	53.09	36.67	38.87	43.84	38.92	31.26	26.25	35.06	67.33	68.38	69.03	68.25
Llama-2-7b-chat-finetune	44.49	44.56	56.79	45.42	47.81	63.84	62.99	57.07	60.65	61.14	72.00	67.96	65.57	68.51
-» Stanceformer	49.13	48.40	55.11	40.51	48.29	68.00	73.25	55.75	63.62	65.15	78.89	73.54	72.63	75.02
Llama-2-13b-chat [0-shot]	36.92	58.18	73.78	57.01	56.47	42.31	38.03	51.75	21.08	38.29	64.10	78.19	73.46	71.92
Llama-2-13b-chat-finetune	66.11	68.64	**78.13**	67.45	70.08	62.34	66.48	60.02	47.75	59.15	76.62	71.88	68.44	72.31
-» Stanceformer	67.16	**71.43**	74.76	**73.98**	**71.83**	64.79	65.77	62.97	62.63	64.04	79.10	77.31	70.54	75.65

VAST-zero-shot	Pro	Con	Neu	All
BiCond†	44.6	47.4	34.9	42.8
CrossNet†	46.2	43.4	40.4	43.4
SEKT†	50.4	44.2	30.8	41.8
TPDG†	53.7	49.6	52.3	51.9
TOAD†	42.6	36.7	43.8	41.0
BERT†	54.6	58.4	85.3	66.1
TGA-Net†	55.4	58.5	85.8	66.6
BERT-GCN†	58.3	60.6	86.9	68.6
CKE-Net†	61.2	61.2	88.0	70.2
PT-HCL	56.1	62.8	87.9	68.9
-» Stanceformer	61.2	59.5	88.9	69.9
WS-BERT	57.0	**63.4**	**90.6**	70.3
-» Stanceformer	60.3	62.1	90.2	70.9
Closed-source LLM				
GPT-3.5 [0-shot]	**68.4**	63.2	81.9	**71.2**
Open-source LLM				
Llama-2-7b-chat [0-shot]	54.4	56.3	6.3	39.0
Llama-2-7b-finetune	49.5	34.0	46.5	43.4
-» Stanceformer	53.1	51.4	53.3	52.6
Llama-2-13b-chat [0-shot]	53.4	58.9	19.3	43.9
Llama-2-13b-finetune	49.0	27.7	48.2	41.6
-» Stanceformer	56.4	49.4	57.5	54.4

Stanceformer improves Target Awareness



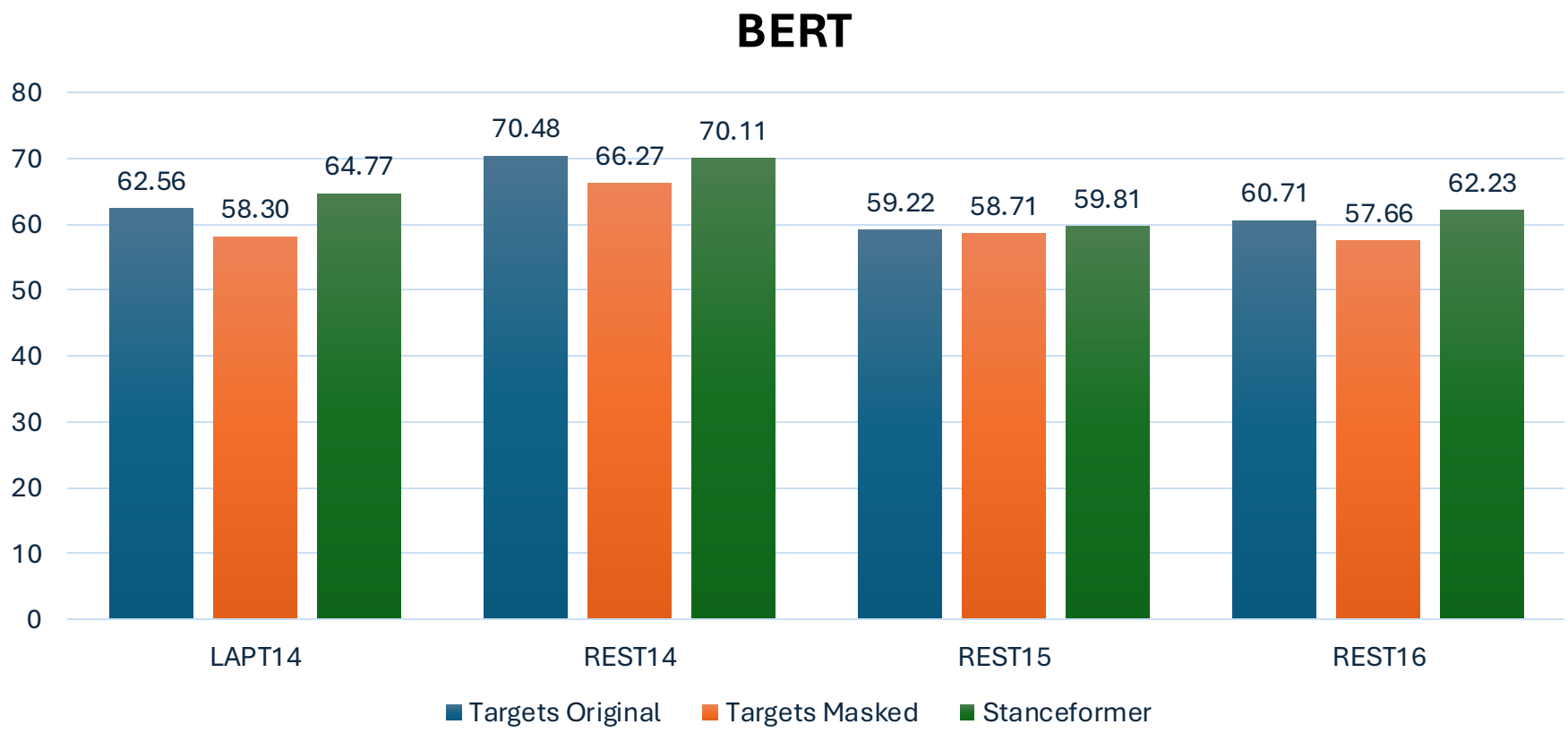
Samples

Text	Target	Ground Truth	BERTweet	GPT-3.5	Stanceformer
Remember, #God has it all worked out.	atheism	🔴	🟢	⚪	🔴
I'm not a feminist, I believe in equality of the sexes! THATS EXACTLY WHAT FEMINISM IS	feminist movement	🟢	⚪	🟢	🟢
Based on the long lines, I thought it was free burrito day at Pancheros but it was actually Hillary#ReadyForHillary	hillary clinton	🟢	🔴	🔴	🟢
@toby_dorena Pregnant people have more than heartbeats. They have feelings, and the ability to make decisions about their health.	legalization of abortion	🟢	🔴	🟢	🟢
Religions give its members an identity & without it, they cannot function. Feminists cannot function without feminism.	feminist movement	🔴	🔴	🟢	🟢

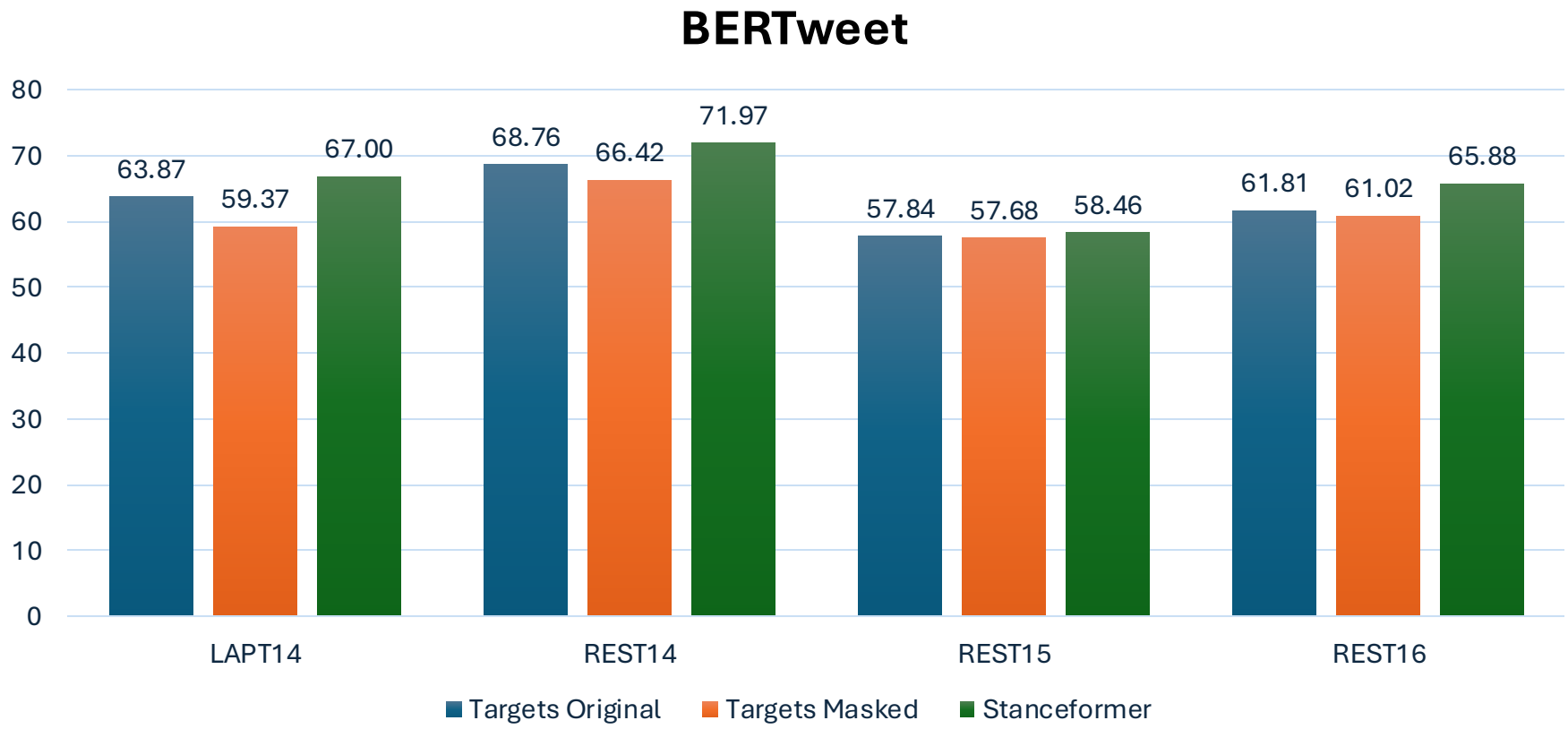
FAVOR: 🟢 **AGAINST:** 🔴 **NONE:** ⚪

Generalization to Aspect-based Sentiment Analysis

Stanceformer improves Target Awareness



Stanceformer improves Target Awareness



Challenges with LLMs

- Heavy computation cost
- High variance with different seeds and prompts
- Uncontrollable LLM outputs
  - Generate irrelevant or non-sensical text
  - Many *Abstain* cases (~15% with Llama-2 models)
- Inconsistent Evaluation
  - Accuracy depends on seed, quality of prompt, quality of regex for parsing
- Possible exposure to test set
- Not all LLMs are open-sourced

Conclusion

- Proposed Stanceformer to address the fundamental shortcoming
- Finetuned LLMs for the first time for Stance Detection
- Demonstrated improvements across Stance Detection datasets and various BERT models as well as Autoregressive LLMs
- Generalized to other domains such as Aspect-based Sentiment Analysis

Thank you