Automatic Fact-Checking and Reasoning

Mitra Mohtarami

Spoken Language Systems
Massachusetts Institute of Technology

April 15, 2020

ROOF OF PUNGGOL WATERWAY TERRACES COLLAPSE

Claim (24/03/2018):

– *US military closes its base in Qatar*

FAKE NEWS!

> **U.S. Central Command** ✔
> @CENTCOM
>
> The US is not leaving Incirlik Air Base in Turkey, nor is the US leaving Al Udeid AB, Qatar. These reports are false and without merit.
>
> 6:56 PM - Mar 25, 2018
>
> ♡ 1,040   💬 1,250 people are talking about this   ⓘ
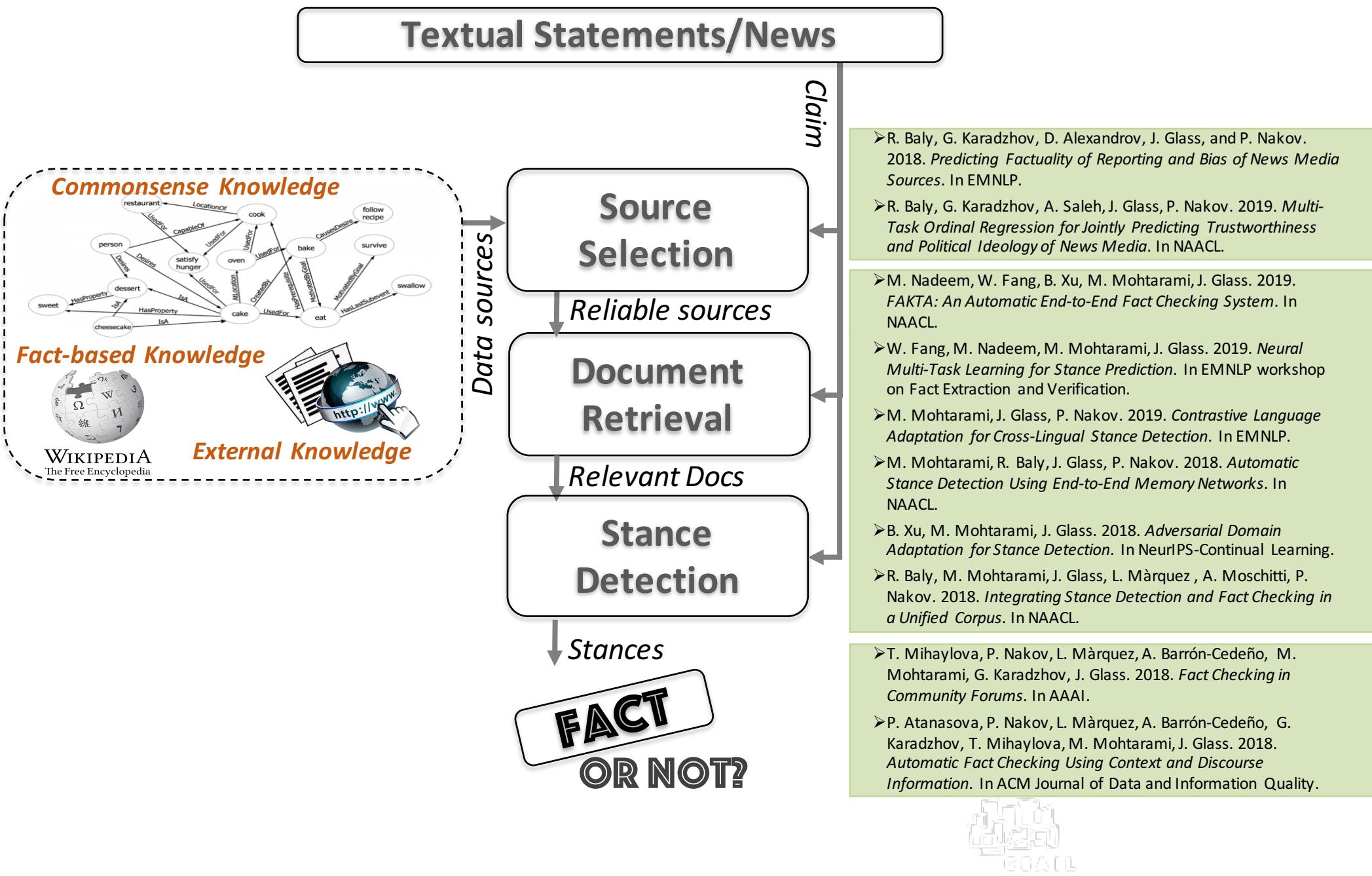
Fact-Checking Process

Textual Statements/News

Claim

Commonsense Knowledge

Fact-based Knowledge

WIKIPEDIA
The Free Encyclopedia

External Knowledge

http://www.

Data sources

Source Selection

Reliable sources

Document Retrieval

Relevant Docs

Stance Detection

Stances

FACT OR NOT?

➢ R. Baly, G. Karadzhov, D. Alexandrov, J. Glass, and P. Nakov. 2018. *Predicting Factuality of Reporting and Bias of News Media Sources*. In EMNLP.

➢ R. Baly, G. Karadzhov, A. Saleh, J. Glass, P. Nakov. 2019. *Multi-Task Ordinal Regression for Jointly Predicting Trustworthiness and Political Ideology of News Media*. In NAACL.

➢ M. Nadeem, W. Fang, B. Xu, M. Mohtarami, J. Glass. 2019. *FAKTA: An Automatic End-to-End Fact Checking System*. In NAACL.

➢ W. Fang, M. Nadeem, M. Mohtarami, J. Glass. 2019. *Neural Multi-Task Learning for Stance Prediction*. In EMNLP workshop on Fact Extraction and Verification.

➢ M. Mohtarami, J. Glass, P. Nakov. 2019. *Contrastive Language Adaptation for Cross-Lingual Stance Detection*. In EMNLP.

➢ M. Mohtarami, R. Baly, J. Glass, P. Nakov. 2018. *Automatic Stance Detection Using End-to-End Memory Networks*. In NAACL.

➢ B. Xu, M. Mohtarami, J. Glass. 2018. *Adversarial Domain Adaptation for Stance Detection*. In NeurIPS-Continual Learning.

➢ R. Baly, M. Mohtarami, J. Glass, L. Màrquez, A. Moschitti, P. Nakov. 2018. *Integrating Stance Detection and Fact Checking in a Unified Corpus*. In NAACL.

➢ T. Mihaylova, P. Nakov, L. Màrquez, A. Barrón-Cedeño, M. Mohtarami, G. Karadzhov, J. Glass. 2018. *Fact Checking in Community Forums*. In AAAI.

➢ P. Atanasova, P. Nakov, L. Màrquez, A. Barrón-Cedeño, G. Karadzhov, T. Mihaylova, M. Mohtarami, J. Glass. 2018. *Automatic Fact Checking Using Context and Discourse Information*. In ACM Journal of Data and Information Quality.

Source Selection

- Wikipedia
- Mainstream media sources (Media Bias/ Fact Check - MBFC)
  - high-factuality =1,300; low-factuality=700; mix-factuality=500

Name	URL	Factuality	Twitter Handle	Wikipedia page
Associated Press	http://apnews.com	*Very High	@apnews	~/wiki/Associated_Press
NBC News	http://www.nbcnews.com/	High	@nbcnews	~/wiki/NBC_News
Russia Insider	http://russia-insider.com	Mixed	@russiainsider	~/wiki/Russia_Insider
Patriots Voice	http://patriotsvoice.info/	Low	@pegidaukgroup	N/A

⇓ MBFC

Least Reliable Source   *Most Reliable Source*
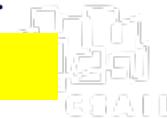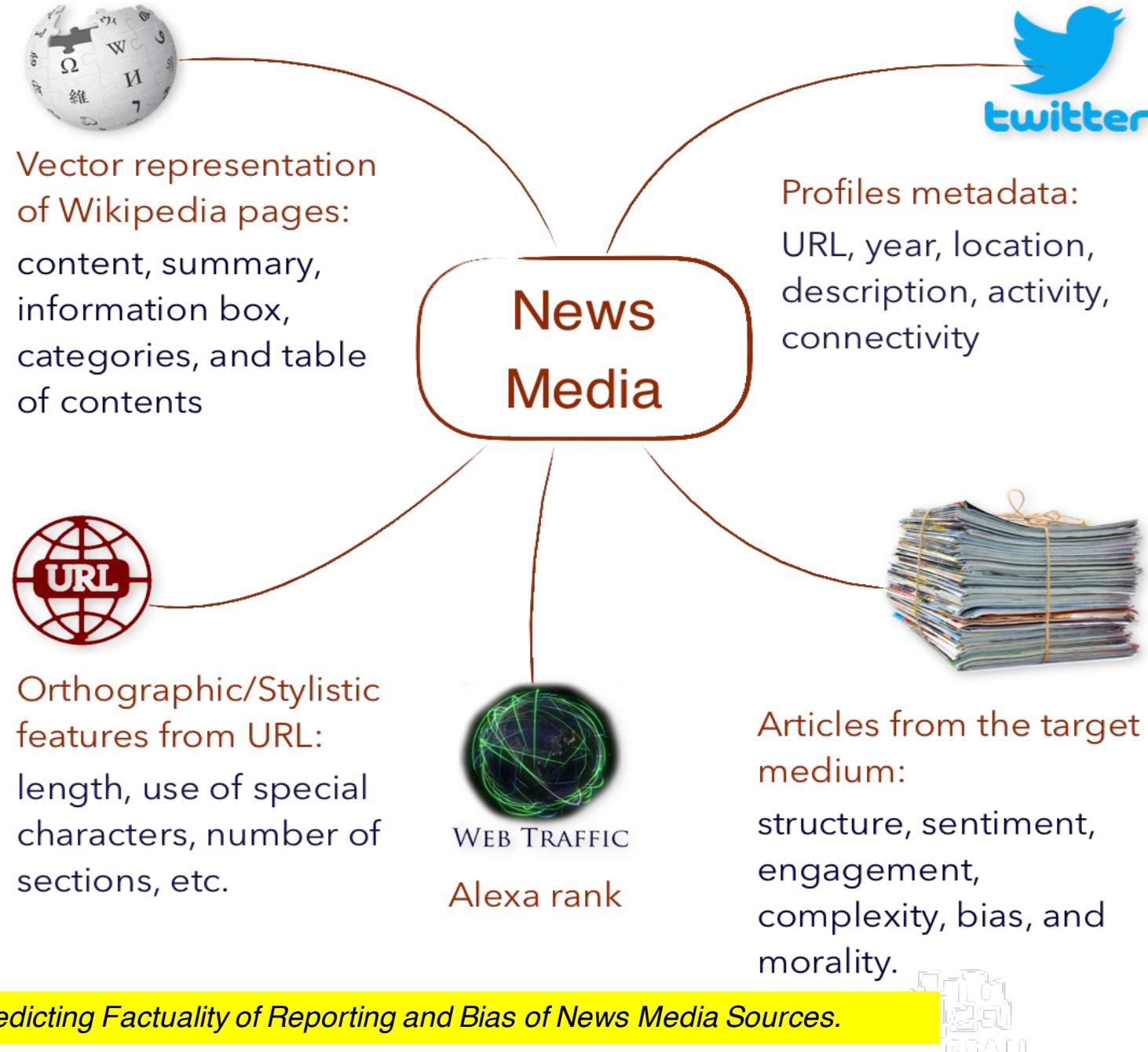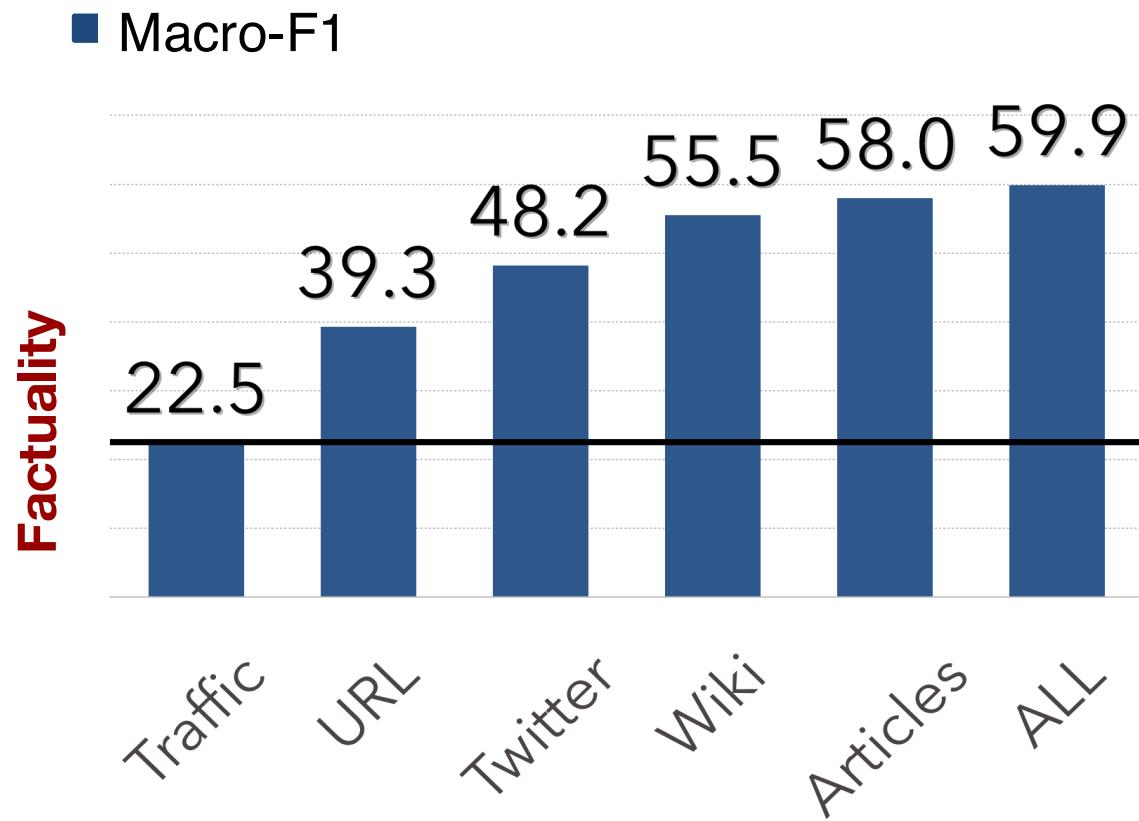
Modeling Media Factuality

Vector representation of Wikipedia pages:
content, summary, information box, categories, and table of contents

Profiles metadata:
URL, year, location, description, activity, connectivity

News Media

Orthographic/Stylistic features from URL:
length, use of special characters, number of sections, etc.

WEB TRAFFIC

Alexa rank

Articles from the target medium:
structure, sentiment, engagement, complexity, bias, and morality.

Modeling Media Factuality

Document Retrieval

Textual Claim → Document Retrieval (Search Engine) → Query Generation → Re-ranking Model

➤ Fact Extraction and VERification (FEVER) ~ 145K claims

R@1

Model	R@1
BM25	28.8
Lucene	43.1
Lucene (QG)	40.7
Google (QG)	56.6
Lucene (Re-rank)	62.4
Google (Re-rank)	57.8

Stance Detection

The task aims to identify the relative perspective of a document or piece of text with respect to a claim, typically modelled using labels such as *agree*, *disagree*, *discuss*, and *unrelated*!

US shifts Qatar base to Saudi

KSA GDN Online Desk

Claim: *US military closes its base in Qatar*

AGREE

The US installations at the Qatari base are being shipped to the Prince Sultan air base in central Saudi Arabia

WASHINGTON: The US has started shifting its air base in Qatar to Saudi Arabia, it was revealed yesterday.

This follows a meeting between Saudi Crown Prince Mohammed bin Salman, currently visiting the US, and President Donald Trump at the White House on Tuesday, a prominent Israeli intelligence website has reported.

US military denies leaving bases in Turkey, Qatar

US Central Command says all such 'reports are false and without merit'

home > Turkey, world, todays headlines, americas   25.03.2018

Claim: *US military closes its base in Qatar*

DISAGREE
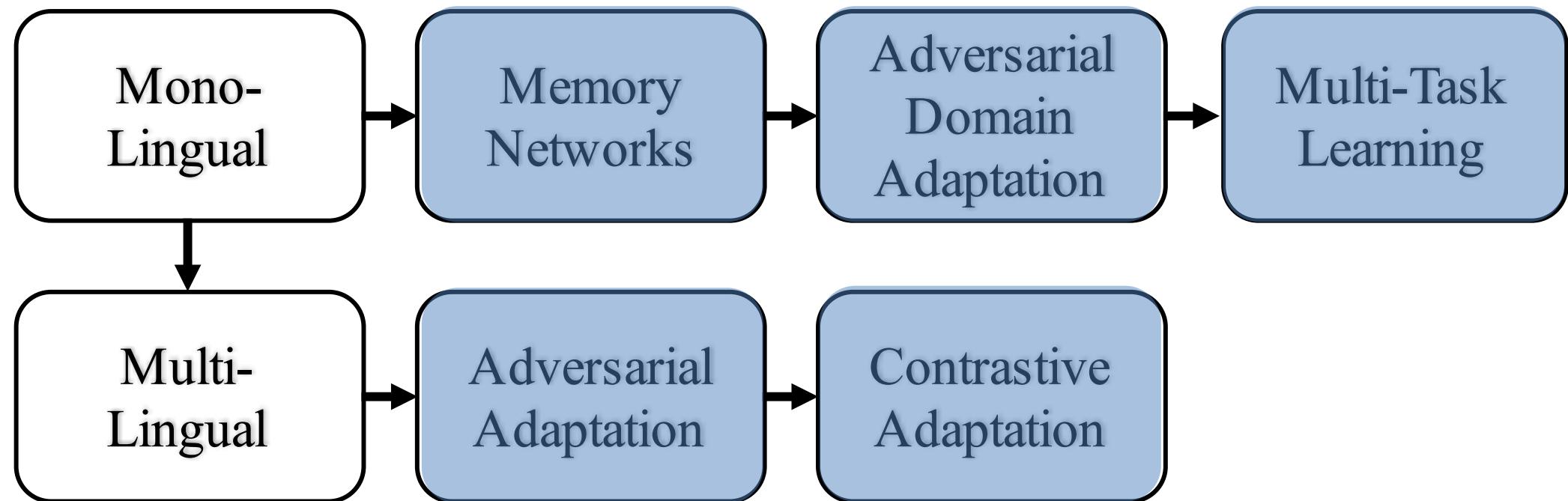
By Kasim Ileri

WASHINGTON

The U.S. Central Command (CENTCOM) on Sunday denied reports it was leaving the Incirlik base in southern Turkey's Adana province and the Al Udeid base in Qatar.

"The U.S. is not leaving Incirlik Air Base in Turkey, nor is the U.S. leaving Al Udeid AB, Qatar. These reports are false and without merit," CENTCOM posted on its official Twitter page.

Stance Detection


```
Mono-Lingual → Memory Networks → Adversarial Domain Adaptation → Multi-Task Learning
Mono-Lingual → Multi-Lingual → Adversarial Adaptation → Contrastive Adaptation
```

NAACL'2018: *Automatic Stance Detection Using End-to-End Memory Networks.*
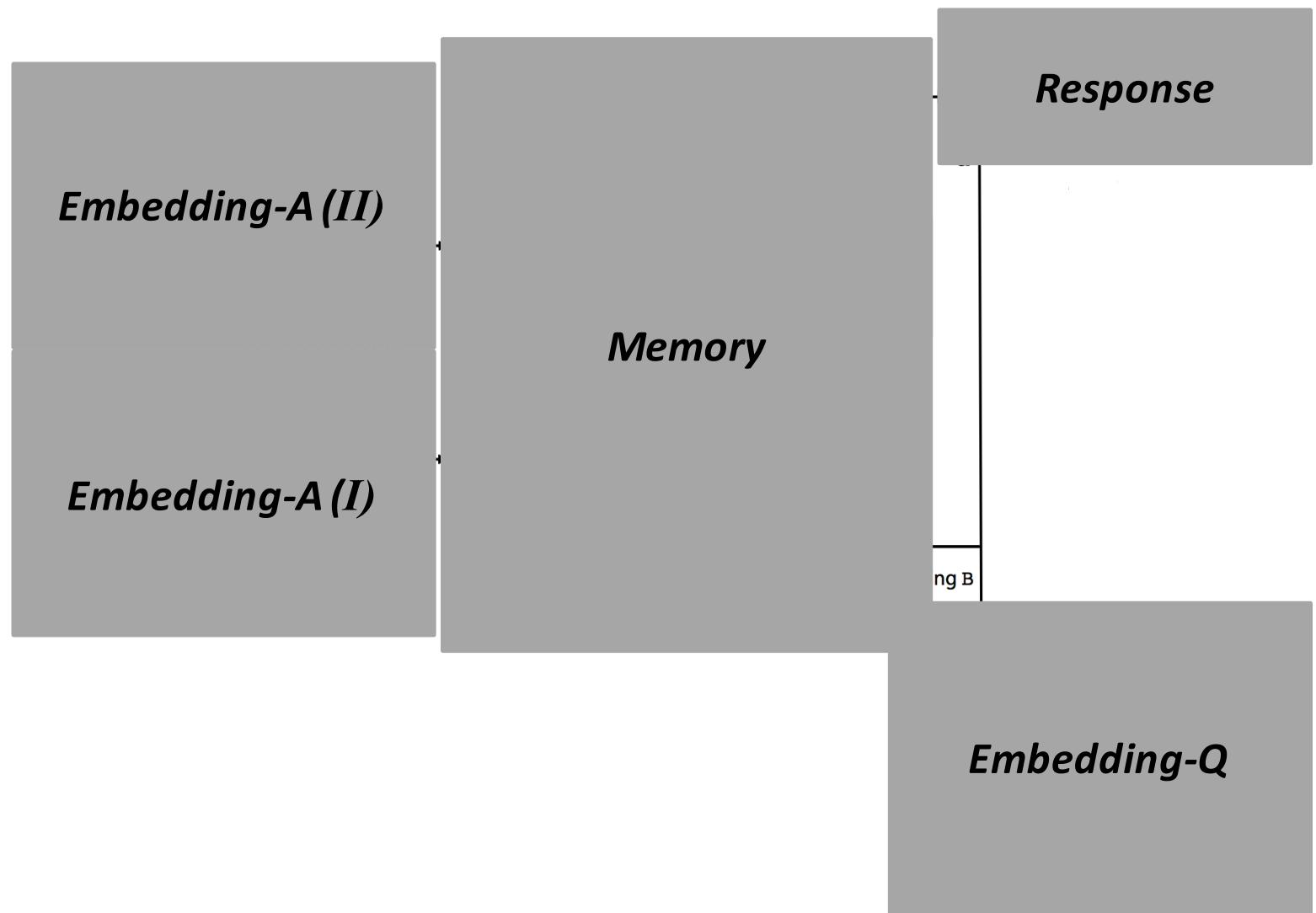
NeurIPS-CL'2018: *Adversarial Doman Adaptation for Stance Detection.*

EMNLP-FEVER'2019: *Neural Multi-Task Learning for Stance Prediction.*
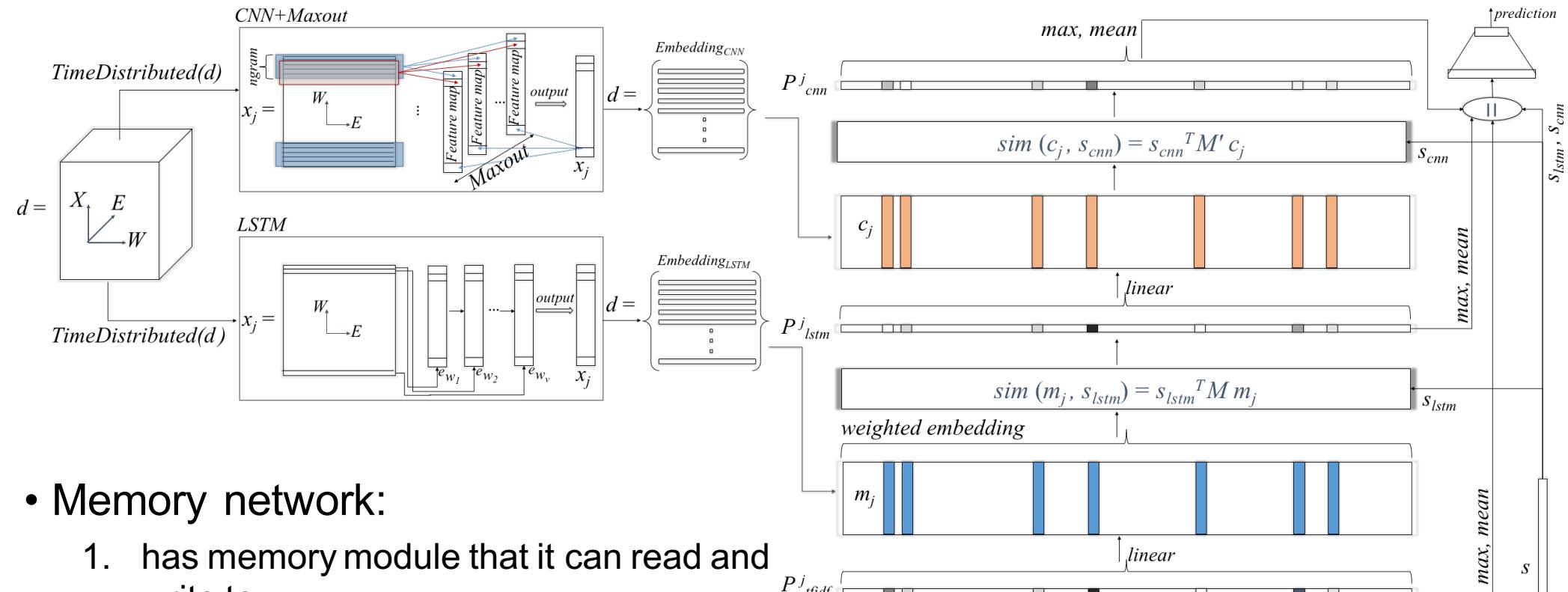
EMNLP'2019: *Contrastive Language Adaptation for Cross-Lingual Stance Detection.*

End-to-End Memory Networks

Embedding-A (II)

Embedding-A (I)

Memory

Response

ng B

Embedding-Q

End-to-End Memory Networks
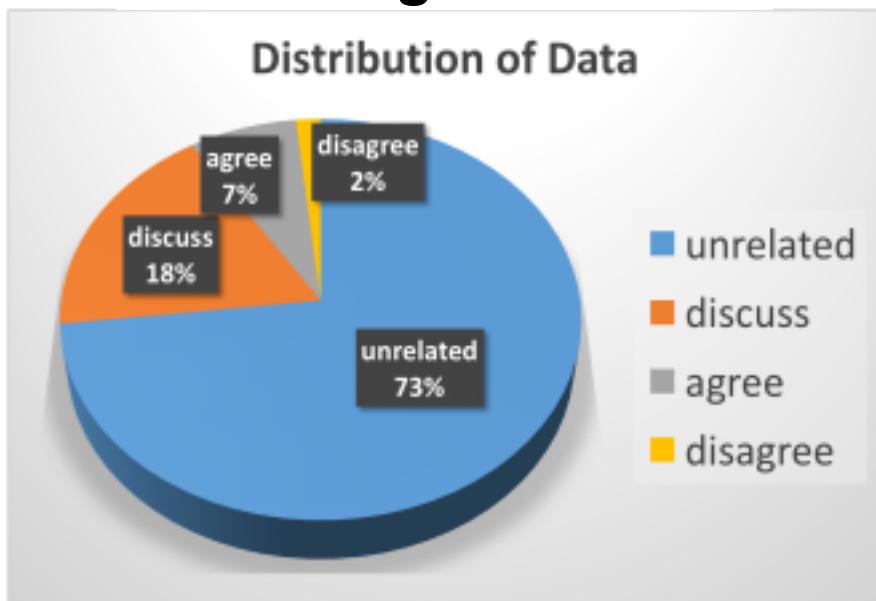
- Memory network:
  1. has memory module that it can read and write to
  2. incorporates reasoning with attention/*similarity-matrix* over memory.
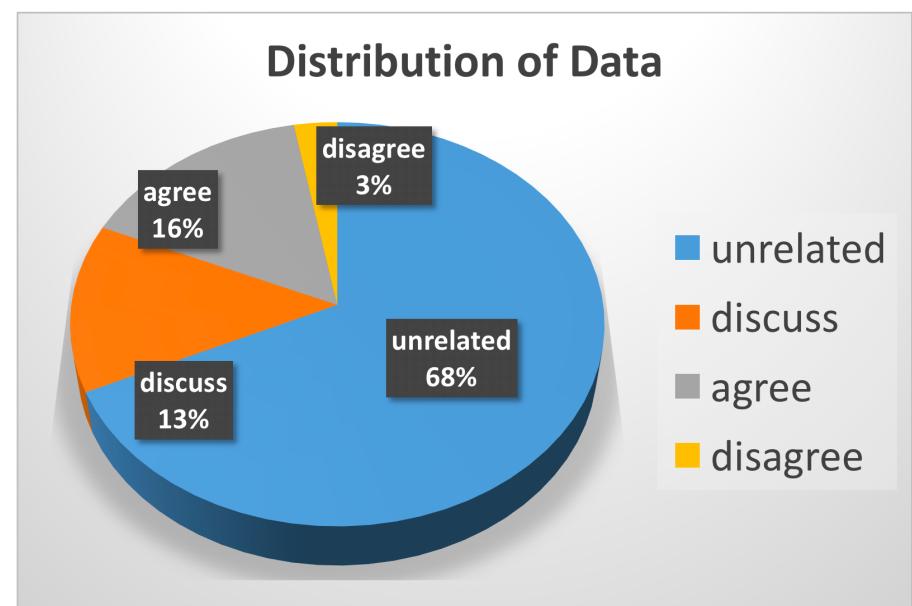
Datasets

➢ Total *English* stances ~ 50K

➢ Total Arabic stances ~ 3K

English

Distribution of Data

- unrelated 73%
- discuss 18%
- agree 7%
- disagree 2%

Arabic

Distribution of Data

- unrelated 68%
- discuss 13%
- agree 16%
- disagree 3%
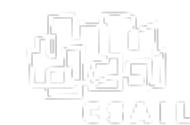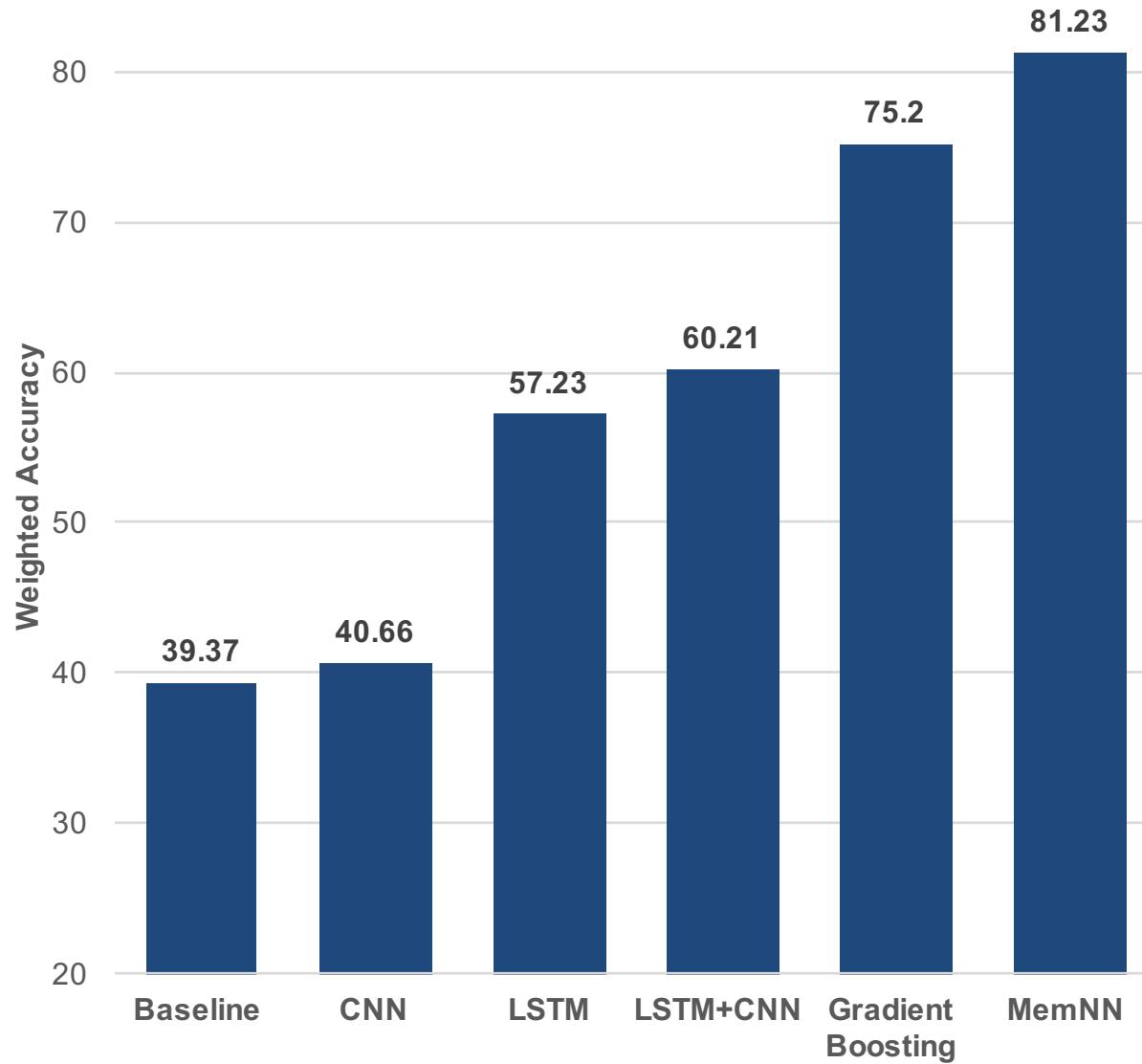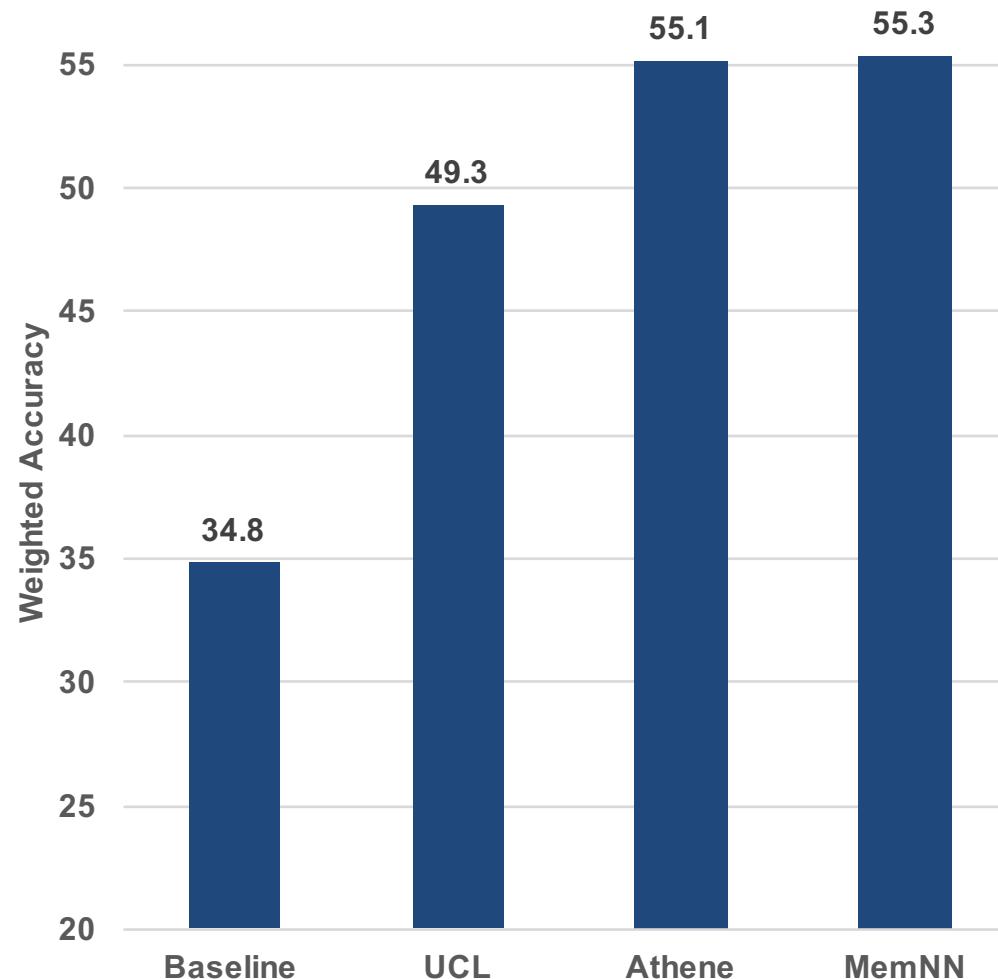
Results on *English* Data
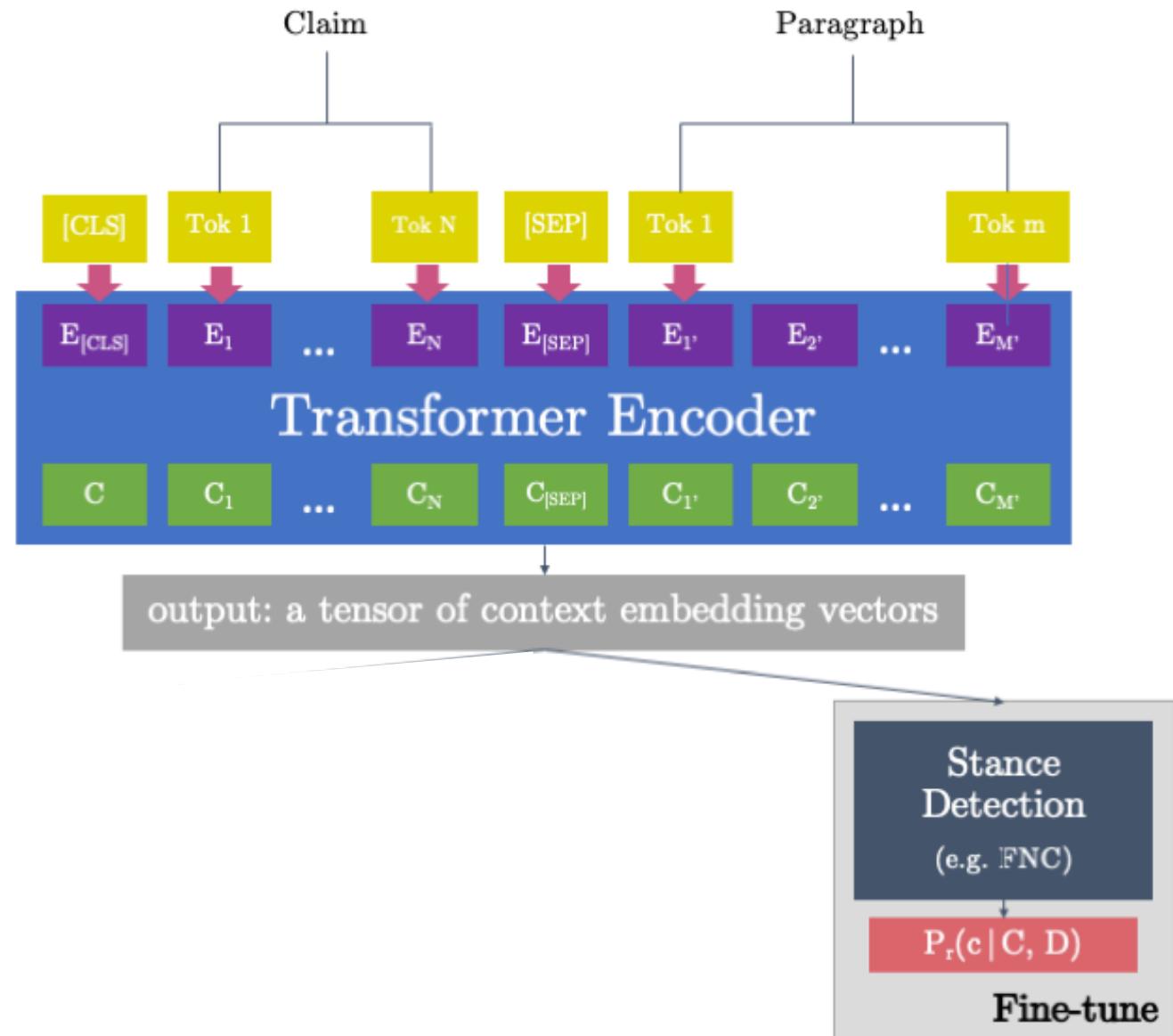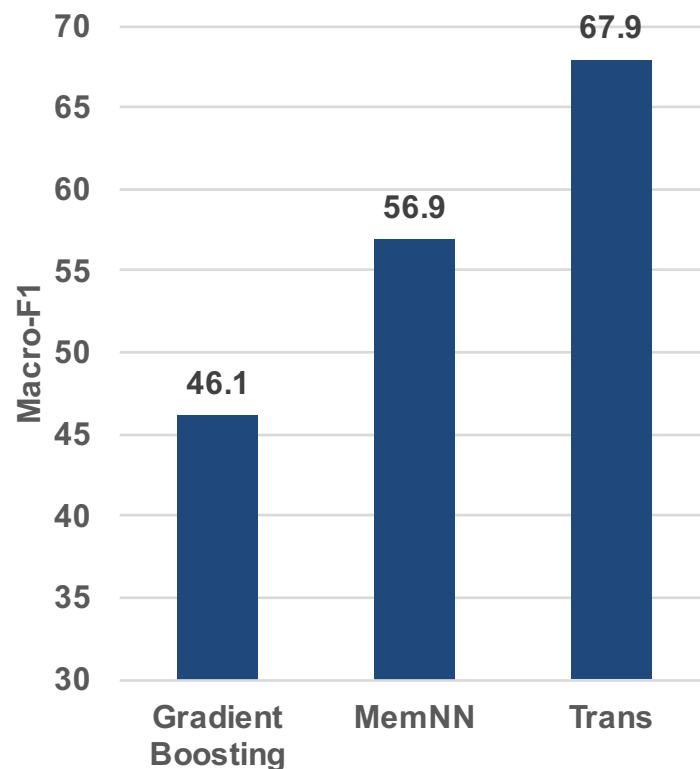
# Results on *Arabic* Data
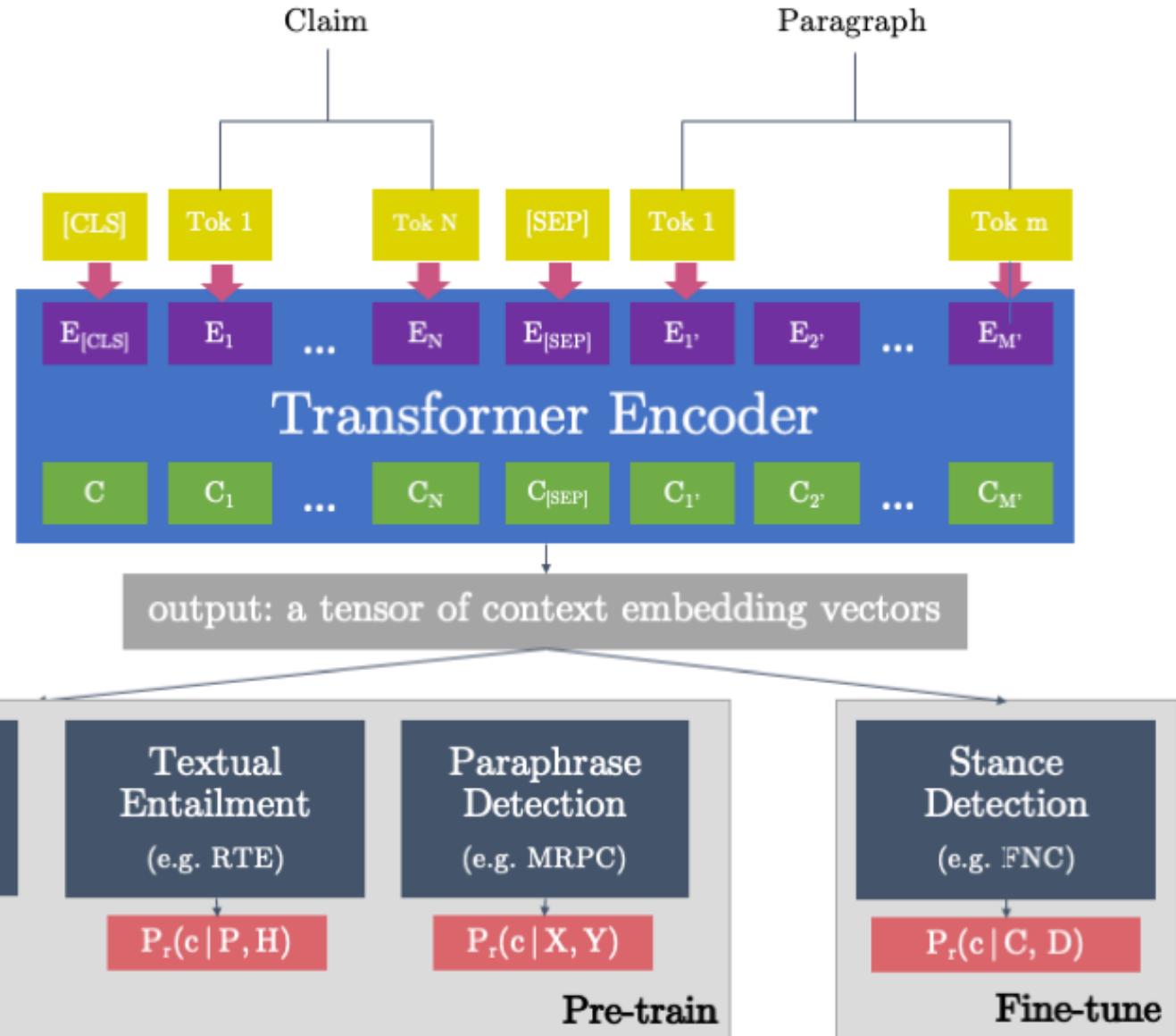
Neural Multi-Task Learning

- ➤ Target Data
  - ➤ Fake News Challenge (FNC)
 - ➤ Total instances ~ **50K**
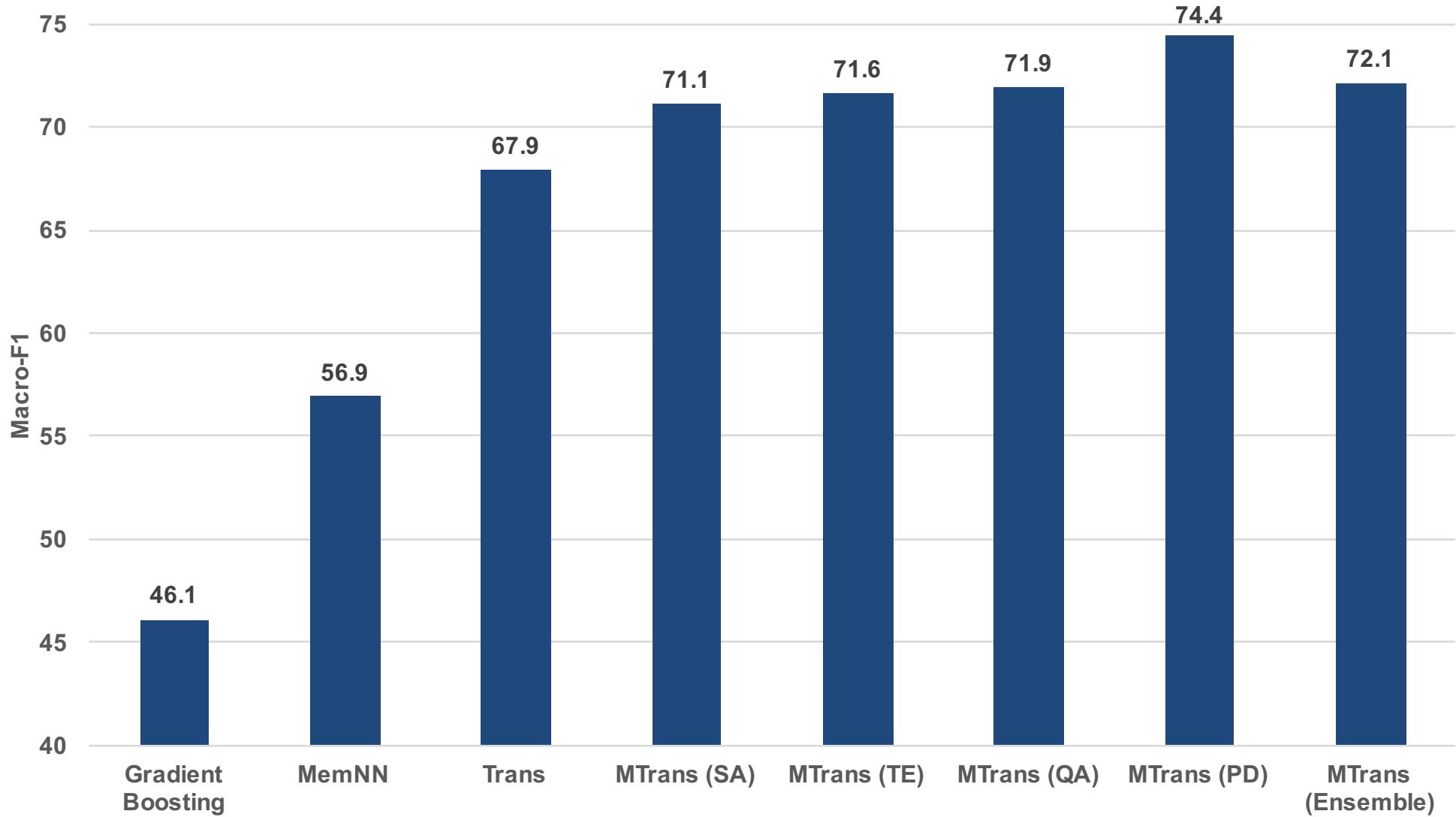
Neural Multi-Task Learning
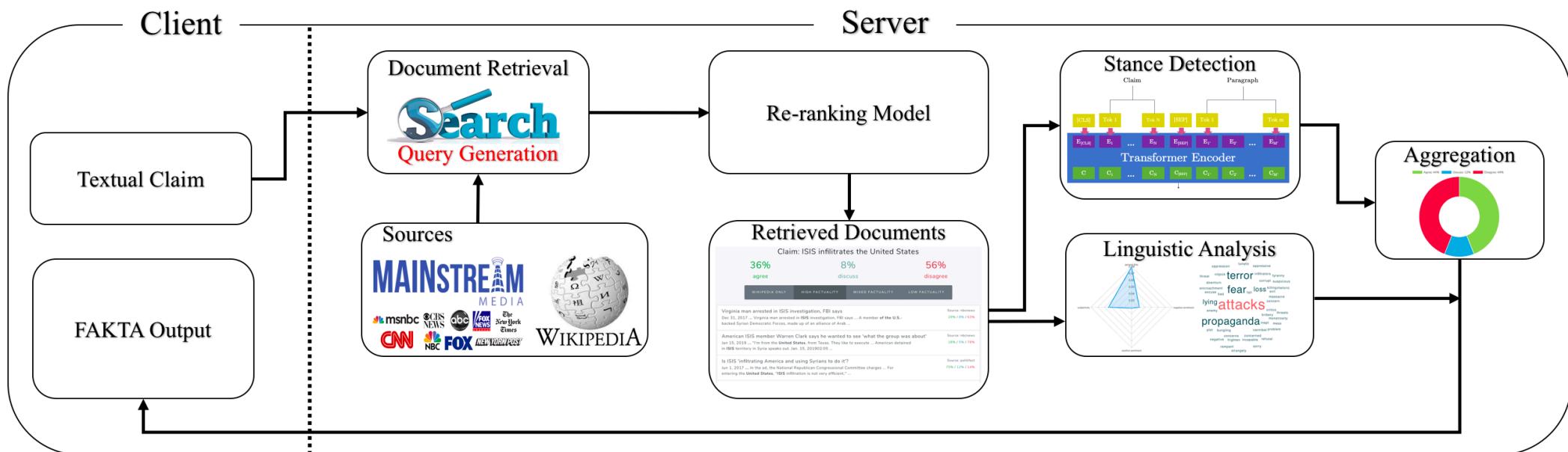
- ➢ Target Data
  - ➢ Fake News Challenge (~50K)
- ➢ Sentiment Analysis (SA)
  - ➢ IMDB (~50K), SST (~67K)
- ➢ Question Answering (QA)
  - ➢ QNLI (~108K)
- ➢ Textual Entailment (TE)
  - ➢ SNLI (~549K), MNLI(~393K), RTE(~2.5K)
- ➢ Paraphrase Detection (PD)
  - ➢ QQP (~364K), MRPC (~3.7K)

Claim Paragraph

| [CLS] | Tok 1 | Tok N | [SEP] | Tok 1 | | Tok m |

| $E_{[CLS]}$ | E_1 | ... | E_N | $E_{[SEP]}$ | $E_{1'}$ | $E_{2'}$ | ... | $E_{M'}$ |

Transformer Encoder

| C | C_1 | ... | C_N | $C_{[SEP]}$ | $C_{1'}$ | $C_{2'}$ | ... | $C_{M'}$ |

output: a tensor of context embedding vectors

Sentiment Analysis (e.g. SST)	Question Answering (e.g. QNLI)	Textual Entailment (e.g. RTE)	Paraphrase Detection (e.g. MRPC)	Stance Detection (e.g. FNC)
$P_r(c\|X)$	$P_r(c\|S,Q)$	$P_r(c\|P,H)$	$P_r(c\|X,Y)$	$P_r(c\|C,D)$

Pre-train **Fine-tune**

Neural Multi-Task Learning

FAKTA: An Automatic End-to-End Fact Checking System

Demo: *https://fakta.mit.edu*

Conclusion

Textual Statements/News

Claim

Commonsense Knowledge

Fact-based Knowledge

WIKIPEDIA
The Free Encyclopedia

External Knowledge

Data sources

Source Selection

Reliable sources

Document Retrieval

Relevant Docs

Stance Detection

Stances

FACT OR NOT?

➢ R. Baly, G. Karadzhov, D. Alexandrov, J. Glass, and P. Nakov. 2018. *Predicting Factuality of Reporting and Bias of News Media Sources*. In EMNLP.

➢ R. Baly, G. Karadzhov, A. Saleh, J. Glass, P. Nakov. 2019. *Multi-Task Ordinal Regression for Jointly Predicting Trustworthiness and Political Ideology of News Media*. In NAACL.

➢ M. Nadeem, W. Fang, B. Xu, M. Mohtarami, J. Glass. 2019. *FAKTA: An Automatic End-to-End Fact Checking System*. In NAACL.

➢ W. Fang, M. Nadeem, M. Mohtarami, J. Glass. 2019. *Neural Multi-Task Learning for Stance Prediction*. In EMNLP workshop on Fact Extraction and Verification.

➢ M. Mohtarami, J. Glass, P. Nakov. 2019. *Contrastive Language Adaptation for Cross-Lingual Stance Detection*. In EMNLP.

➢ M. Mohtarami, R. Baly, J. Glass, P. Nakov. 2018. *Automatic Stance Detection Using End-to-End Memory Networks*. In NAACL.

➢ B. Xu, M. Mohtarami, J. Glass. 2018. *Adversarial Domain Adaptation for Stance Detection*. In NeurIPS-Continual Learning.

➢ R. Baly, M. Mohtarami, J. Glass, L. Màrquez, A. Moschitti, P. Nakov. 2018. *Integrating Stance Detection and Fact Checking in a Unified Corpus*. In NAACL.

➢ T. Mihaylova, P. Nakov, L. Màrquez, A. Barrón-Cedeño, M. Mohtarami, G. Karadzhov, J. Glass. 2018. *Fact Checking in Community Forums*. In AAAI.

➢ P. Atanasova, P. Nakov, L. Màrquez, A. Barrón-Cedeño, G. Karadzhov, T. Mihaylova, M. Mohtarami, J. Glass. 2018. *Automatic Fact Checking Using Context and Discourse Information*. In ACM Journal of Data and Information Quality.

Questions!