Applying Automatic Text Summarization for Fake News Detection

LREC 2022 | Philipp Hartl & Udo Kruschwitz

Challenge

VS

"All the News That's Fit to Print"

The New York Times

Late Edition

Today, mostly sunny, warm for early November, high 72. **Tonight,** clear, mild, low 54. **Tomorrow,** sunshine, patchy clouds, remaining warm, high 71. Weather map, Page 21.

VOL. CLXX . . . No. 58,871 © 2020 The New York Times Company NEW YORK, SUNDAY, NOVEMBER 8, 2020 $6.00

BIDEN BEATS TRUMP

RACE IS FINALLY CALLED AFTER RECORD TURNOUT; CHAOTIC TERM ENDS WITH RARE INCUMBENT LOSS

Harris Will Become the Country's First Female Vice President

By JONATHAN MARTIN and ALEXANDER BURNS

Joseph Robinette Biden Jr. was elected the 46th president of the United States on Saturday, promising to restore political normalcy and a spirit of national unity to confront raging health and economic crises, and making Donald J. Trump a one-term president after four years of tumult in the White House.

Mr. Biden's victory amounted to a repudiation of Mr. Trump by millions of voters exhausted with his divisive conduct and chaotic administration, and was delivered by an unlikely alliance of women, people of color, old and young voters and a sliver of disaffected Republicans. Mr. Trump is only the third elected president since World War II to lose re-election, and the first in more than a quarter-century.

The result also provided a history-making moment for Mr. Biden's running mate, Senator Kamala Harris of California, who will become the first woman to serve as vice president.

With his triumph, Mr. Biden, who turns 78 later this month, fulfilled his decades-long ambition in his third bid for the White House, becoming the oldest person elected president. A pillar of Washington who was first elected amid the Watergate scandal, and who prefers political consensus over combat, Mr. Biden will lead a nation and a Democratic Party that have become far more ideological since his arrival in the capital in 1973.

He offered a mainstream Democratic agenda, yet it was less his policy platform than his biography to which many voters gravitated. Seeking the nation's highest office a half-century after his first campaign, Mr. Biden — a candidate in the late autumn of his career — presented his life of setback and recovery to voters as a parable for a wounded country.

Appearing Saturday night before supporters at a drive-in rally in Wilmington, Del., and speaking against the din of enthusiastic honking, Mr. Biden claimed the presidency and called on the country to reunite after what he described as a toxic political interlude.

"Let this grim era of demonization in America begin to end here and now," he said.

Without addressing Mr. Trump, the president-elect spoke directly to the president's supporters and said he recognized their disappointment. "I've lost a couple of times myself," he recalled of his past failures to win the presidency, before adding: "Now let's give each other a chance."

In a statement earlier in the day, Mr. Trump insisted

Continued on Page 12

BIDEN TRUMP NOT CALLED

Kamala Harris and Joseph R. Biden Jr. addressed the nation Saturday night from Wilmington, Del., after the election was called. Page 11.

https://arstechnica.com/tech-policy/2020/11/as-trump-falsely-claims-victory-twitter-and-facebook-counter-misinformation/
https://www.nytimes.com/2020/11/08/learning/lesson-of-the-day-biden-wins-presidency-ending-four-tumultuous-years-under-trump.html

Fake News Detection

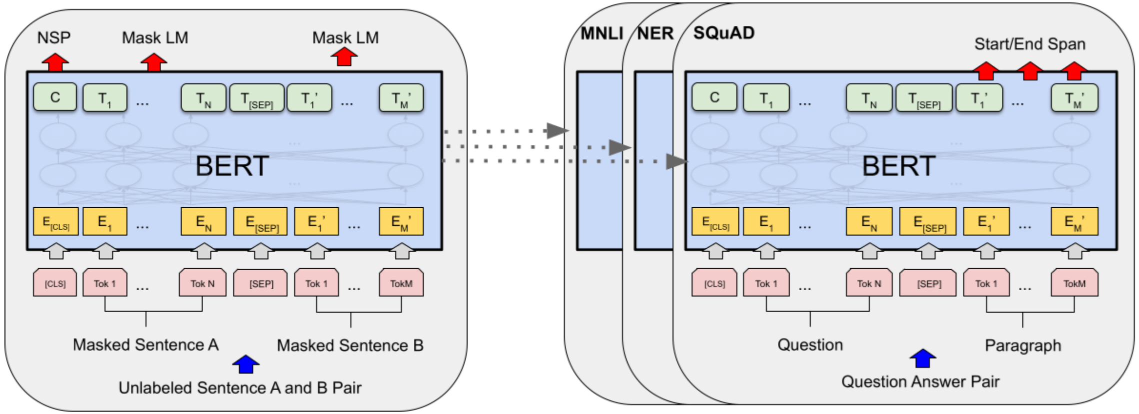
Content-based Identification [1, 2]

Feedback-based Identification [3, 4]

Intervention-based Solutions [5, 6]

[1] Zhou, X., Jain, A., Phoha, V. V., & Zafarani, R. (2020). Fake News Early Detection: A Theory-driven Model. Digital Threats: Research and Practice, 1(2), 1-25. https://doi.org/10.1145/3377478
[2] Zhou, X., Wu, J., & Zafarani, R. (2020). SAFE: Similarity-Aware Multi-modal Fake News Detection. In H. W. Lauw, R. C.-W. Wong, A. Ntoulas, E.-P. Lim, S.-K. Ng, & S. J. Pan (Eds.), Advances in Knowledge Discovery and Data Mining (pp. 354-367). Springer International Publishing. https://doi.org/10.1007/978-3-030-47436-2_27
[3] Shu, K., Cui, L., Wang, S., Lee, D., & Liu, H. (2019). dEFEND: Explainable Fake News Detection. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 395-405. https://doi.org/10.1145/3292500.3330935
[4] Shu, K., Wang, S., & Liu, H. (2019). Beyond news contents: The role of social context for fake news detection. WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining, 312-320. https://doi.org/10.1145/3289600.3290994
[5] Jooyeon Kim, Behzad Tabibian, Alice Oh, Bernhard Schölkopf, and Manuel Gomez Rodriguez. 2018. Leveraging the Crowd to Detect and Reduce the Spread of Fake News and Misinformation. In Proceedings of the 11th ACM International Conference on Web Search and Data Mining (WSDM 2018).
[6] Yiangos Papanastasiou. 2017. Fake news propagation and detection: A sequential model. SSRN (2017)

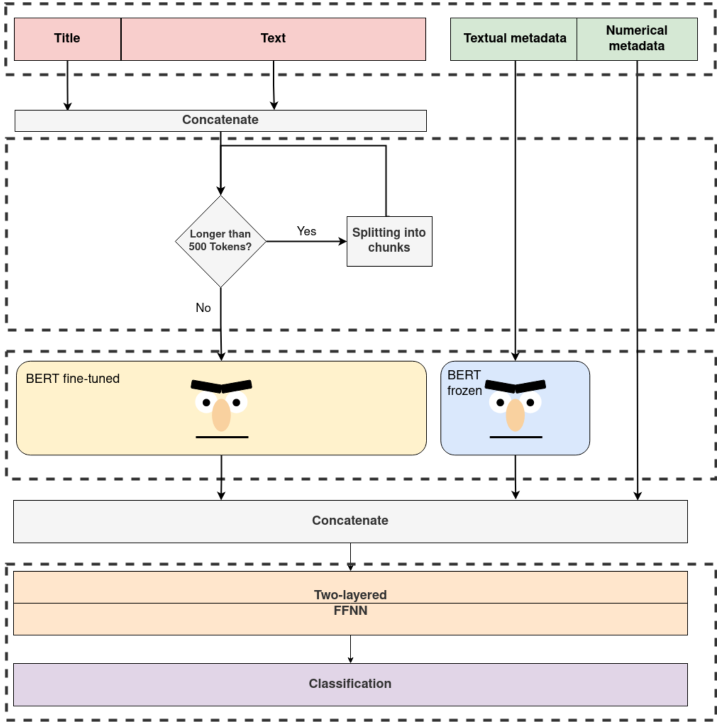
Transfer Learning



Pre-training Fine-Tuning

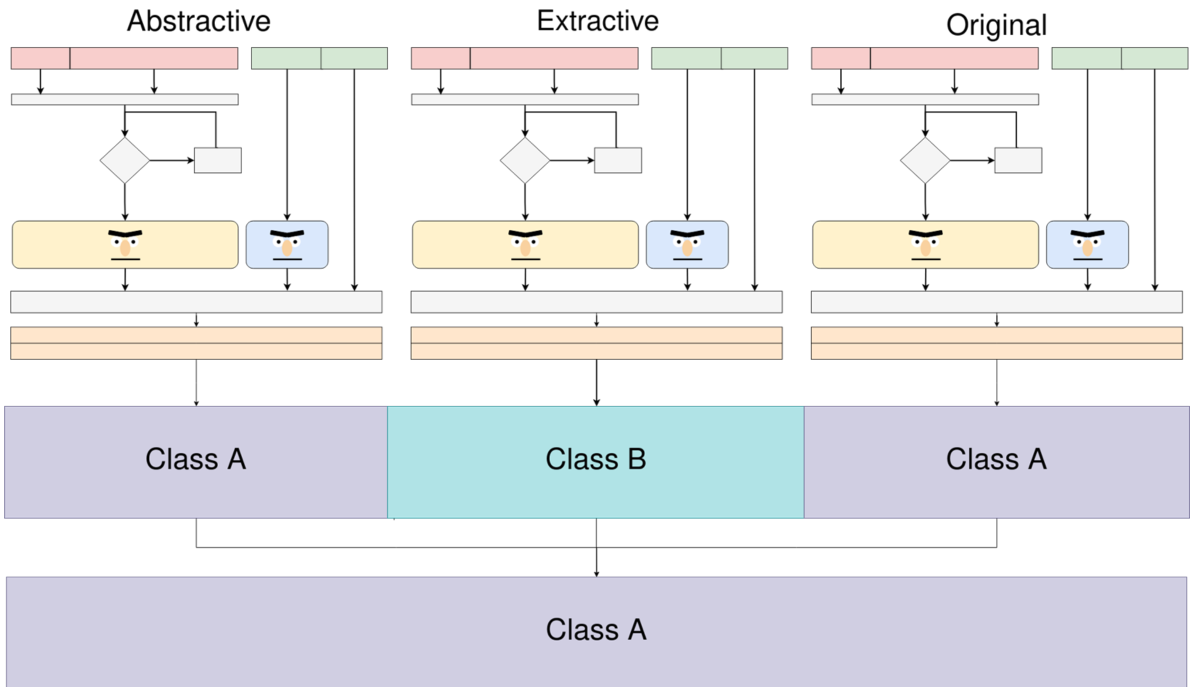
Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019, June). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1(Long and Short Papers)(pp. 4171-4186). Minneapolis, Minnesota: Associationfor Computational Linguistics.

CMTR-BERT

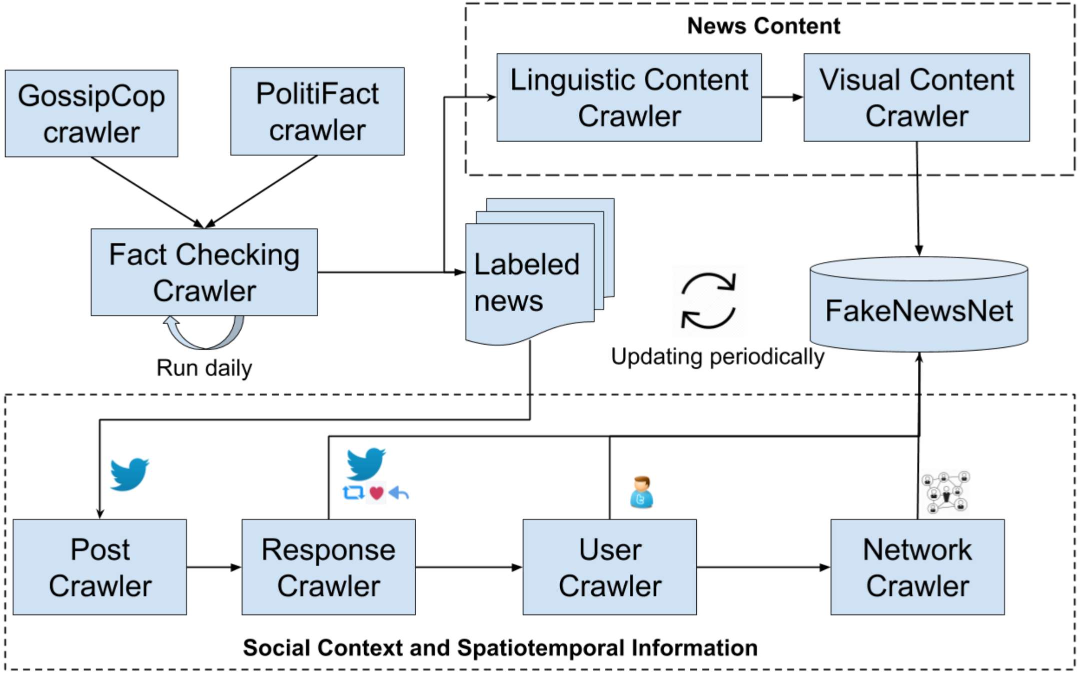


Title
Text
Textual metadata
Numerical metadata
Concatenate
Longer than 500 Tokens?
Yes
Splitting into chunks
No
BERT fine-tuned
BERT frozen
Concatenate
Two-layered FFNN
Classification

CMTR-BERT Ensemble



Datasets (1)



Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2020). FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media. Big Data, 8(3), 171-188. https://doi.org/10.1089/big.2020.0062

Datasets (2)


CLEF 2021 BUCHAREST
CLEF2021 - CheckThat! Lab
Detecting Check-Worthy Claims, Previously Fact-Checked Claims, and Fake News

Performance (1)

Dataset	Metric	BERT	SAFE	dEFEND	CMTR-BERT O	CMTR-BERT A	CMTR-BERT E	CMTR-BERT C	CMTR-BERT
Politifact	Accuracy	0.924	0.874	0.904	0.950	0.948	0.950	0.912	**0.956**
	Precision	0.934	0.889	0.902	0.953	0.945	0.941	**0.977**	0.958
	Recall	0.897	0.903	**0.956**	0.936	0.943	0.952	0.827	0.947
	F1	0.914	0.896	0.928	0.944	0.943	0.946	0.895	**0.952**
Gossipcop	Accuracy	0.863	0.838	0.808	0.960	0.956	0.958	0.957	**0.963**
	Precision	0.741	0.854	0.729	0.926	0.917	0.924	0.859	**0.936**
	Recall	0.666	0.937	0.782	0.908	0.898	0.899	**0.985**	0.910
	F1	0.701	0.895	0.755	0.917	0.907	0.911	0.918	**0.923**

Performance (2)

Dataset	Metric	NoFake	NLP & IR@UNED	BERT-baseline	BERT	CMTR-BERT O w/o context	CMTR-BERT A w/o context	CMTR-BERT E w/o context
CT-FAN 21	Accuracy	**0.853**	0.528	0.316	0.453	0.461	0.441	0.480
	Precision	-	-	0.128	0.424	0.446	0.422	**0.467**
	Recall	-	-	0.251	0.402	0.414	0.384	**0.435**
	F1-macro	**0.838**	0.468	0.124	0.395	0.406	0.373	0.428

Key findings

- CMTR-BERT is able to achieve state-of-the-art results
- Contextual information is a key aspect of Fake News
- Single text representations do not consistently bring advantage → Combination does the trick
- It remains unclear to what extent the input transformation influences the performance

Thank you for your attention!