Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims

Zeyu Zhang, Jacob Daniel Devasier, Damian Jimenez, Sai Sandeep Naraparaju, Haiqi Zhang, Zhengyuan Zhu, Chengkai Li

The University of Texas at Arlington, TX, USA

{zeyu.zhang, zhengyuan.zhu,
haiqi.zhang,jacob.devasier,damian.jimenez,sxn1362}@mavs.uta.edu, cli@uta.edu

We present **ClaimBuster** [1, 2, 3], an AI model designed to automatically identify claims worth fact-checking. Built using a human-labeled dataset of factual claims from U.S. general election debate transcripts [4], ClaimBuster has been operational since 2016, providing real-time support for fact-checkers during live coverage of primary and general election debates for U.S. presidential elections. As illustrated in Figure 1, for each debate we analyze, the transcript is posted on the ClaimBuster website, with each sentence annotated with its ClaimBuster score. Scores closer to 1 indicate high check-worthiness, while scores closer to 0 indicate low check-worthiness. Fact-checkers and citizens can use ClaimBuster during the live coverage of an event to help them discover claims that need more caution to trust. Additionally, we developed a monitoring system that has been actively posting high-score claims on the ClaimBuster Twitter account by monitoring news channels such as MSNBC, as shown in Figure 2.

ClaimBuster has been widely adopted by researchers and fact-checkers. For example, the Duke Reporters' Lab (DRL) used ClaimBuster to generate daily email alerts for professional fact-checkers, highlighting the most check-worthy claims from TV program transcripts and social media. These alerts have contributed to at least 33 claims featured in 30 articles by prominent fact-checking outlets, including *The Washington Post* [5]. Users can easily call ClaimBuster's API by submitting a text snippet to assess its check-worthiness. ClaimBuster's API has been called over 219 million times: by our internal projects, the DRL, and other users. The API currently has 362 registered users ranging from academics to professional journalists across several countries, and we regularly see new registrations. Furthermore, we developed **ClaimPortal**, an online platform that tracks political factual claims by analyzing tweets from politicians, reporters, and media organizations. ClaimPortal streamlines the fact-checking process, enabling users to efficiently identify and verify relevant content.

For the 2024 U.S. presidential election, we collected and analyzed debate transcripts using ClaimBuster. In a case study, we examined two CNN articles covering the presidential debates held on June 27 [1] and September 10, 2024 [2]. Of the 71 claims fact-checked by CNN from these debates, the average ClaimBuster score was 0.69, significantly higher than the average score of 0.41 for all sentences in the debates. Table 1 and 2 summarize the number of sentences in the debate transcripts within specific ClaimBuster score ranges and how many of these sentences were fact-checked in the corresponding CNN articles. For instance, the first row in Table 1 shows that 59 sentences have ClaimBuster scores between 0.8 and 1, 9 of which were fact-checked in the CNN article. Notably, the majority of claims fact-checked by CNN have ClaimBuster scores above 0.6. This analysis highlights ClaimBuster's ability to effectively identify factual claims worth verifying.

[1] https://www.cnn.com/2024/06/27/politics/fact-checking-the-cnn-presidential-debate

[2] https://www.cnn.com/2024/09/10/politics/fact-check-debate-trump-harris

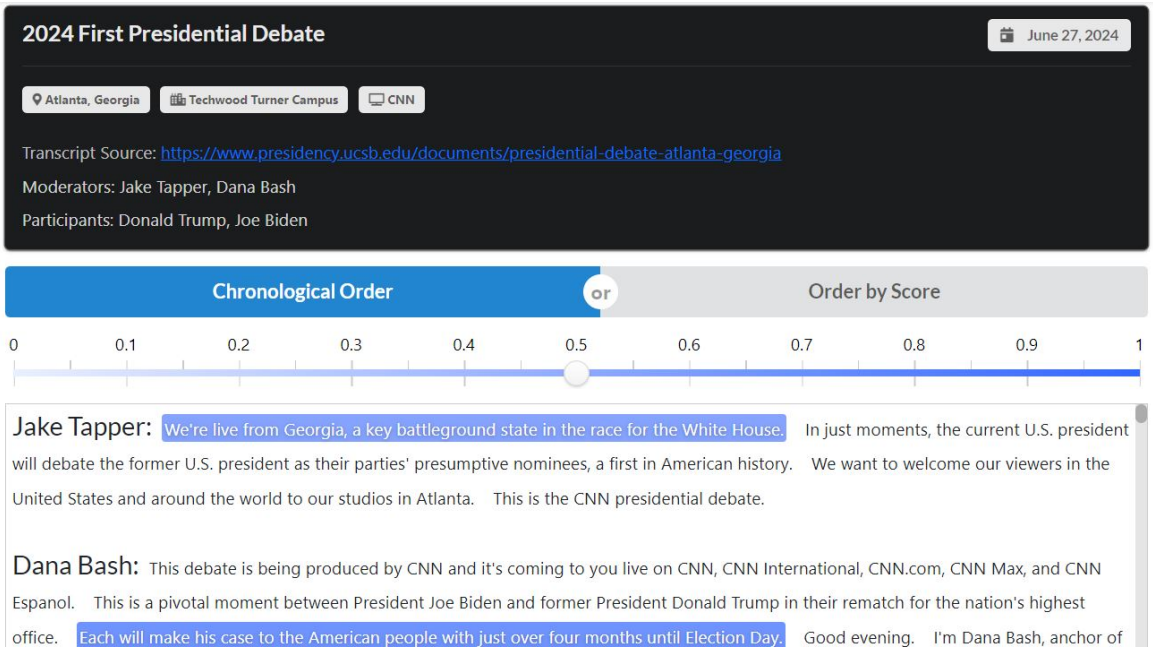

Figure 1: Scored Debate Transcript

Figure 2: ClaimBuster Twitter Account

Table 1: Statistics of the Presidential Debate on June 27, 2024

Score Range	Sentence Count	Sentences Covered by CNN Article
0.8 - 1	59	9
0.6 - 0.8	216	18
0.4 - 0.6	438	9
0.2 - 0.4	539	1
0 - 0.2	167	0

Table 2: Statistics of the Presidential Debate on September 10, 2024

Score Range	Sentence Count	Sentences Covered by CNN Article
0.8 - 1	50	11
0.6 - 0.8	187	12
0.4 - 0.6	372	10
0.2 - 0.4	651	1
0 - 0.2	215	0

References

[1] N. Hassan, G. Zhang, F. Arslan, J. Caraballo, D. Jimenez, S. Gawsane, S. Hasan, M. Joseph, A. Kulkarni, A. K. Nayak *et al.*, "Claimbuster: The first-ever end-to-end fact-checking system," *Proceedings of the VLDB Endowment*, vol. 10, no. 12, pp. 1945–1948, 2017.

[2] N. Hassan, F. Arslan, C. Li, and M. Tremayne, "Toward automated fact-checking: Detecting check-worthy factual claims by claimbuster," in *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 2017, pp. 1803–1812.

[3] K. Meng, D. Jimenez, J. D. Devasier, S. S. Naraparaju, F. Arslan, D. Obembe, and C. Li, "Gradient-based adversarial training on transformer networks for detecting check-worthy factual claims," *ACM Trans. Intell. Syst. Technol.*, vol. 15, no. 6, Nov. 2024.

[4] F. Arslan, N. Hassan, C. Li, and M. Tremayne, "A benchmark dataset of check-worthy factual claims," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, 2020, pp. 821–829.

[5] D. Funke, "This washington post fact check was chosen by a bot," 2018.