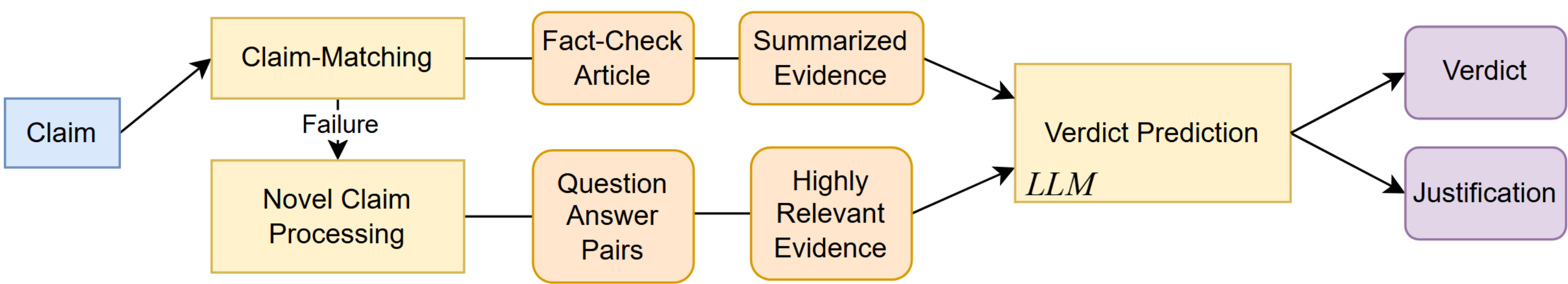




System Architecture

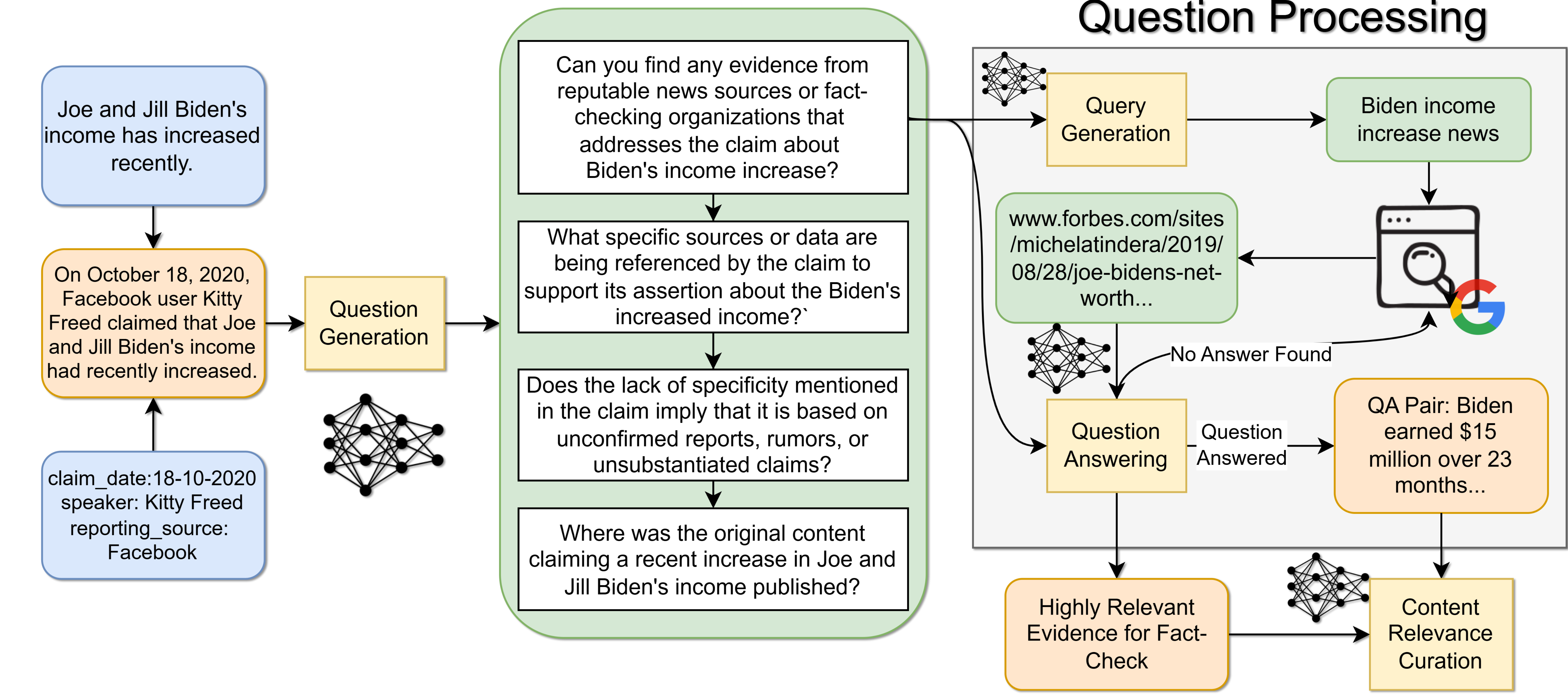


Evidence retrieval extracts information to fact-check a given claim.

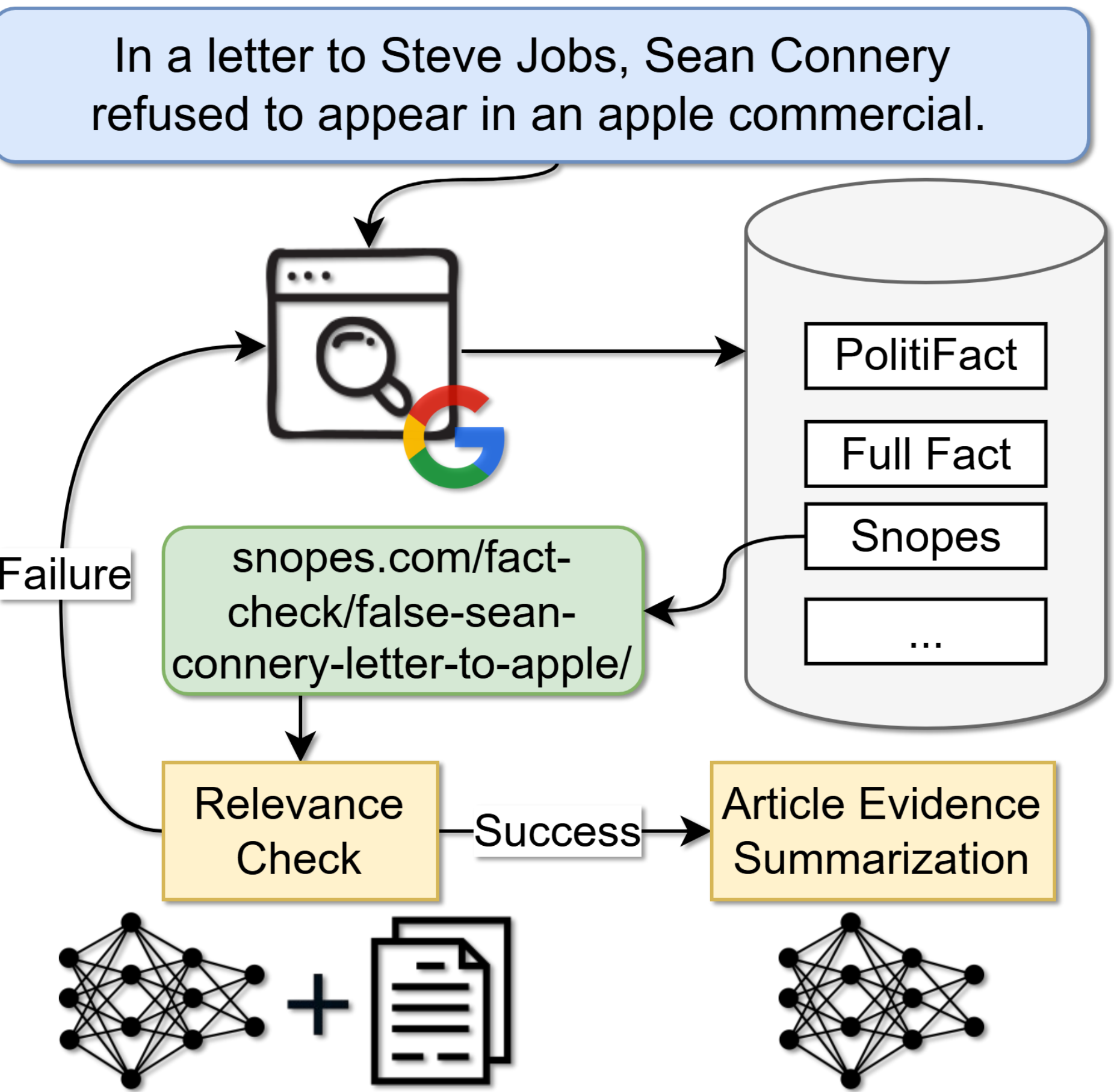
- **Claim-Matching** – Previous fact-check articles for a claim.
- **Novel Claims** – Reasoning with evidence from Internet.

Fact Verification evaluates truthfulness of the claim using collected evidence.

Novel Claim Processing



Claim Matching



Model Selection

- Used only **small LLMs**.
- Base **Qwen 2.5-7B** (4-bit quantized) for most tasks
- **LoRA-Finetuned** for fact verification

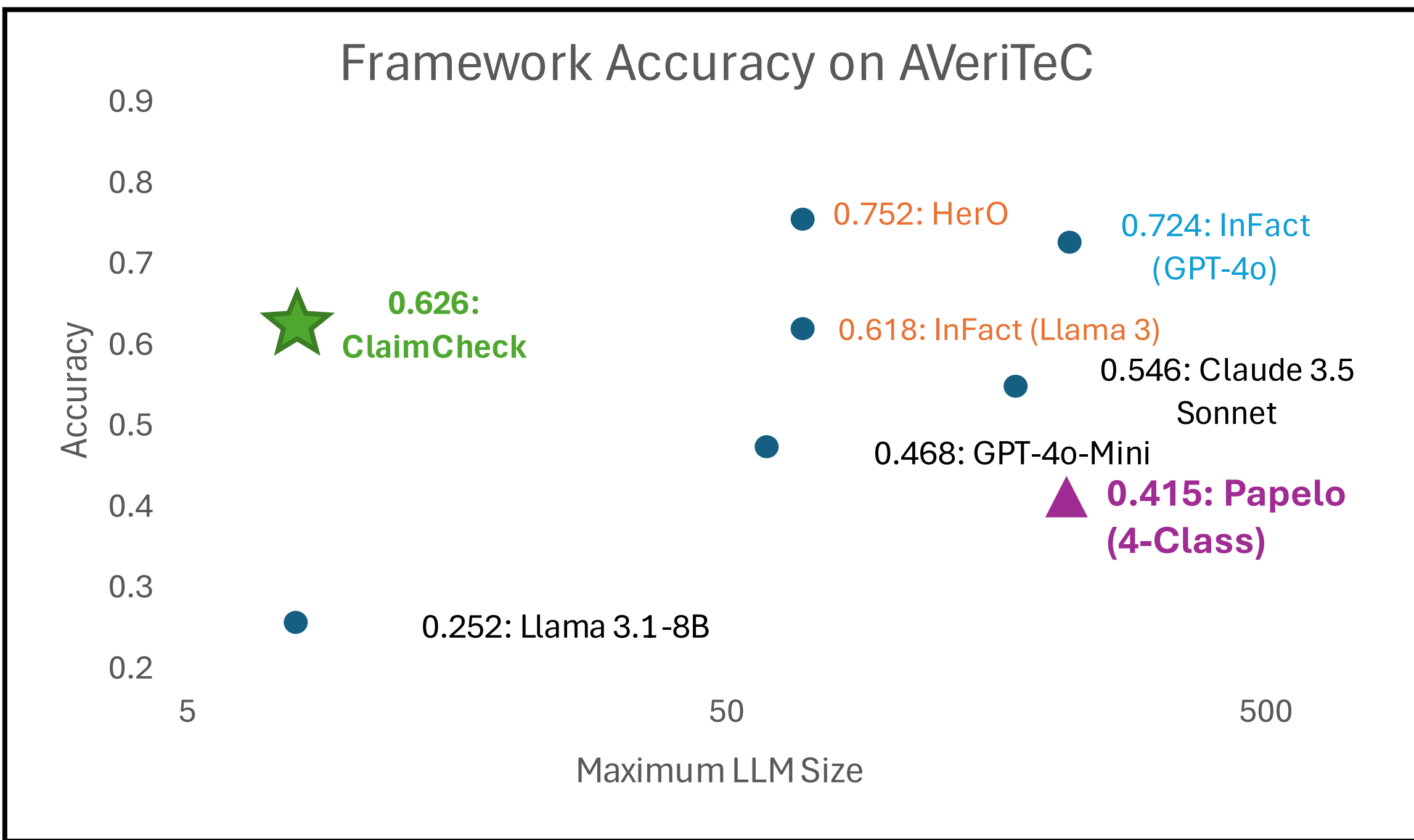
Methodology

- Web search via Google Search API **before** claim date
- Included many trusted fact-checking organizations
- **Fact-checking report** is generated by:
 - Analyzing top fact-check articles to determine relevance
 - Extracting **important evidence** from fact-check using LLM

Claim Matching Analysis

- 15.8% of claims could be fact-checked using Claim Matching
- **+13.3%** accuracy advantage over verdict prediction model.

Results



Evidence Retrieval Analysis

- Found evidence for **98% of claims** in dataset
- After LLM curation, only 70% of claims retrieved **relevant** evidence
- 95% of questions were **answerable** using evidence

Verdict Prediction

- **Exact match** accuracy on predicted Verdict
- Most work relies on **large/proprietary** LLMs
- Also rely on curated evidence list from AVeriTeC
- **+21%** over best **online** retrieval system (**Papelo**)
- Finetuning small LLMs significantly improves their fact-checking capabilities, similar to 70B+ models

Metric	Percentage
Claims with Evidence	98.0%
Claims with Evidence (After Content Relevance Curation)	69.6%
Questions Answered	94.9%
Fact-Check Articles Matched	15.8%
Claim Matching Accuracy	75.9%

Error Analysis

Claim	Predicted Verdict	LLM-generated Justification	Error
Even if your mail-in ballot is postmarked prior to the Election day but arrives after the Election day it will be thrown out.	N/A Ground-truth: Cherry-picking	The article discusses potential risks associated with following Trump's advice during the election, particularly concerning long lines at polling places. Experts note that such behavior could exacerbate existing infrastructure challenges posed by the pandemic, indicating that this scenario is plausible and thus supports the idea of chaos on Election Day.	No verdict generated.
Illegal drug, border crossings, and human smuggling activities have decreased in areas where barriers are deployed in the United States.	Refuted Ground-truth: Not Enough Evidence	This claim is refuted. According to the QA Pairs provided it seems that there was no evidence found for this claim.	Incorrect understanding of verdict definitions.
Donald Trump said that \$15 an hour is too much for essential workers	Refuted Ground-truth: Cherry-picking	The claim states that Joe Biden commented on Donald Trumps statement. There was no proof of this in the QA pairs, Therefore it's refuted.	Incorrect claim interpretation.
There is a scam involving Santander customers being sent fake bank cards.	Refuted Ground-truth: Not Enough Evidence	Q&A 1 gives no evidence for the claim so it is refuted.	Uses incorrect verdict.