A Survey on Automated Fact-Checking

**Zhijiang Guo***, **Michael Schlichtkrull*** and **Andreas Vlachos**

Introduction

- **Fact-checking** assesses and argues for the **factuality** of **claims** made in written or spoken language.
- **NLP** can play a major role in fact-checking:
  - **Searching** large collections of documents for evidence.
  - **Triaging** claims to identify the highest priority targets for professionals to debunk.
  - **Spotting** connections between pieces of evidence.
  - **Identifying** previously fact-checked claims.

Motivation

- Summarizes **up-to-date** research efforts, including the production of justifications.
- Introduces a **unified framework** for automated fact-checking.
- Documents and compares **datasets and models** across different approaches.
- Anticipates **key future challenges.**

Framework
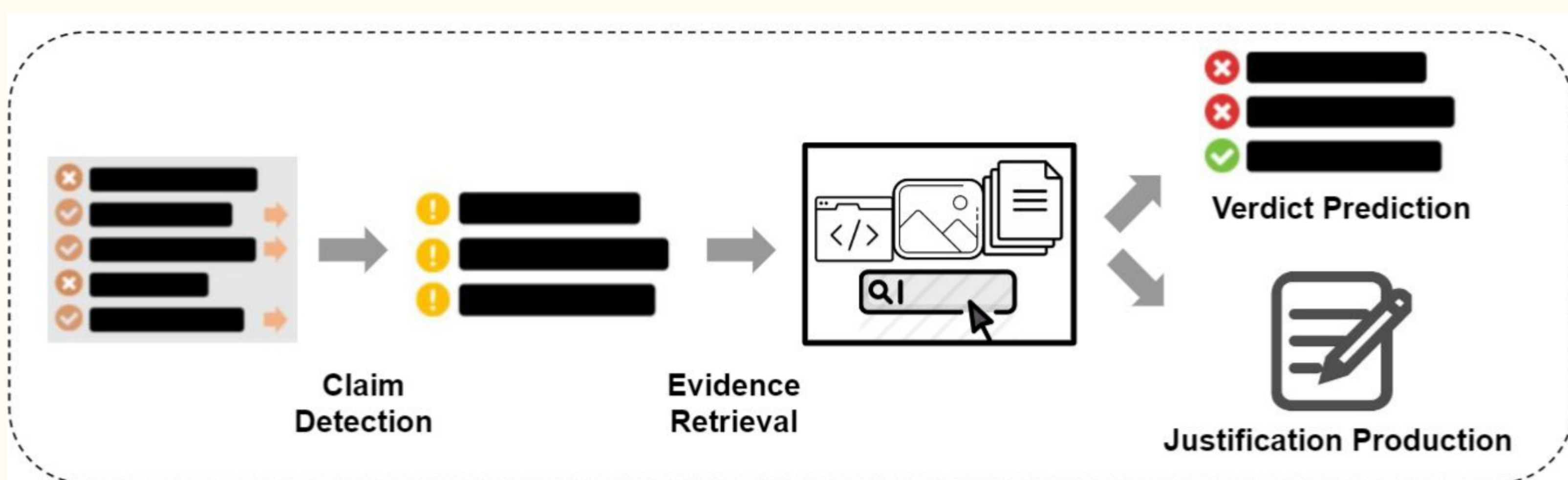
- **Claim detection** identifies claims that require verification.
- **Evidence retrieval** finds sources supporting or refuting the claim.
- **Claim verification** assesses the veracity of the claim based on the retrieved evidence. It can be decomposed into two parts:
  - **Verdict prediction,** where claims are assigned truthfulness labels.
  - **Justification production,** where explanations for verdicts are produced.
- Evidence retrieval and claim verification are sometimes tackled as a **single task** called **factual verification.**

Datasets & Models

- In this survey, we analyse and compare existing datasets and models extensively.

Stages	Type	#Datasets	#Models
Claim Detection	Check-worthy	14	18
	Checkable	3	
Factual Verification	Natural	25	37
	Artificial	17	

- Datasets are analysed along three axes aligned with the three stages of our framework: the **input,** the **evidence** used, and the **verdicts and justifications** that constitute the output.

Dataset	Input	#Inputs	Evidence	Verdict	Sources
Politifact (Vlachos and Riedel, 2014)	State	106	Text/Meta	5Class	Factcheck
Emergent (Ferreira and Vlachos, 2016)	State	300	Text	3Class	Emergent
PunditFact (Rashkin et al., 2017)	State	4,361	✗	2/6Class	Factcheck
Liar (Wang, 2017)	State	12,836	Meta	6Class	Factcheck
Snopes (Hanselowski et al., 2019)	State	6,422	Text	3Class	Factcheck
MultiFC (Augenstein et al., 2019)	State	36,534	Text/Meta	2-27Class	Factcheck
SciFact (Wadden et al., 2020)	State	1,409	Text	3Class	Science
Health (Kotonya and Toni, 2020b)	State	11,832	Text	4Class	Factcheck
X-Fact (Gupta and Srikumar, 2021)	State	31,189	Text	7Class	Factcheck
FEVER (Thorne et al., 2018a)	State	185,445	Text	3Class	Wiki
HOVER (Jiang et al., 2020)	State	26,171	Text	2Class	Wiki
VitaminC (Schuster et al., 2021)	State	488,904	Text	3Class	Wiki
TabFact (Chen et al., 2020)	State	92,283	Table	2Class	Wiki
InfoTabs (Gupta et al., 2020)	State	23,738	Table	3Class	Wiki
Sem-Tab-Fact (Wang et al., 2021)	State	5,715	Table	3Class	Wiki
FEVEROUS (Aly et al., 2021)	State	87,026	Text/Table	3Class	Wiki

- **Pipeline** and **joint modelling strategies** are also surveyed based on our proposed framework.

Research Challenges

- **Epistemology** (binary labelling too simplistic, untrustworthy and/or contradictory sources an inevitability, biased datasets)
- **Practice** (multilinguality, multimodality, and how to generate faithful justifications)
- **Application** (Early intervention and prebunking)

Resources

- Check out our **Github Repo** for details of **datasets and models,** or send us a pull request if your work is missing!