

# Multi-view Counterfactual Contrastive Learning for Fact-checking Fake News Detection

The 14th International Conference on Multimedia Retrieval

ICMR 2024

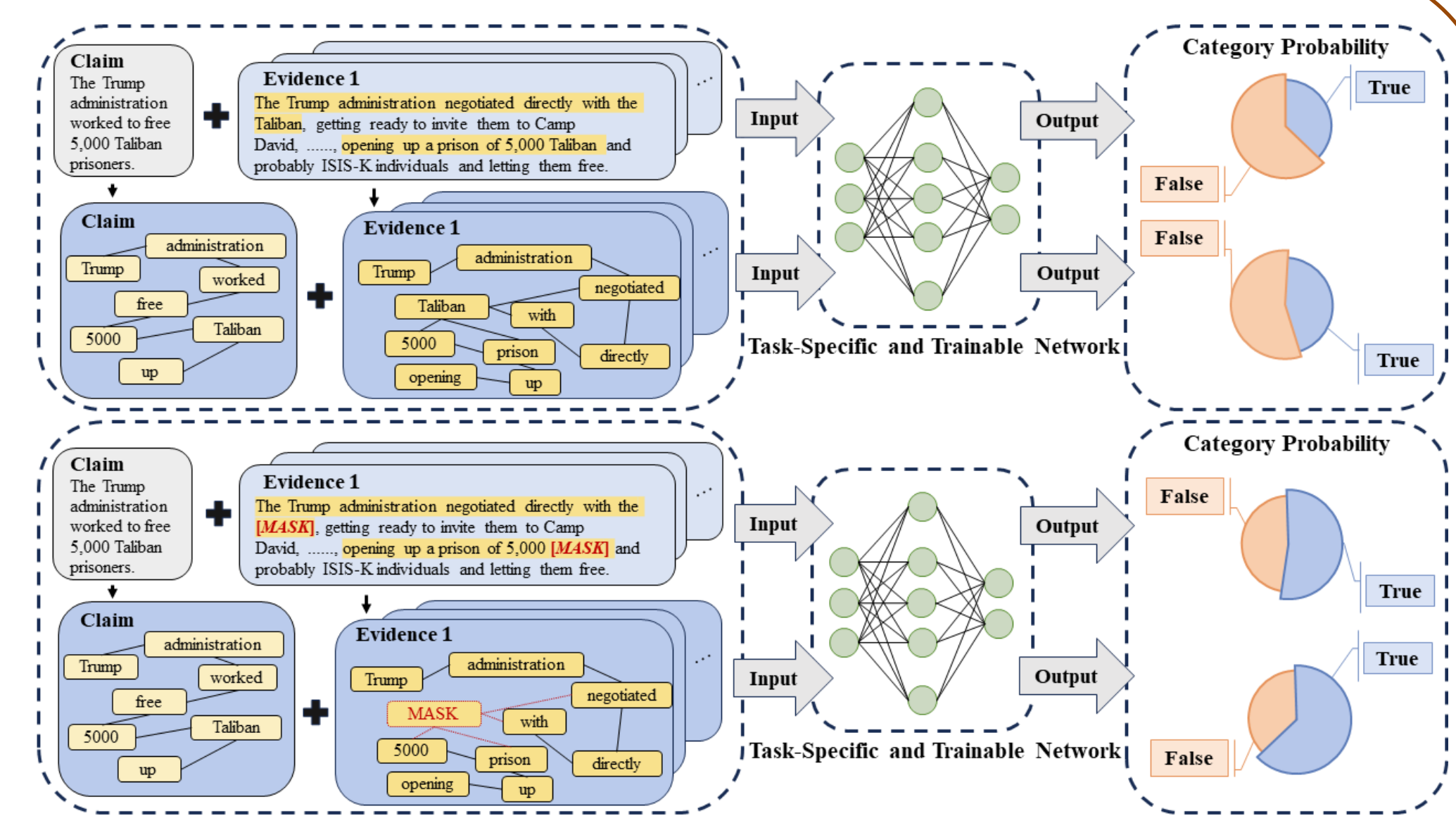
Yongcheng Zhang<sup>1</sup>, Lingou Kong<sup>2</sup>, Sheng Tian<sup>1</sup>, Hao Fei<sup>3</sup>, Changpeng Xiang<sup>1</sup>, Huan Wang<sup>1</sup> and Xiaomei Wei<sup>1\*</sup>

<sup>1</sup>Huazhong Agricultural University, <sup>2</sup>Wuhan University, <sup>3</sup>National University of Singapore

## 1 Introduction

### Contributions:

- Our framework tackles a significant challenge in fake news detection by synergistically harnessing text-based semantic and graph-based topological learning. We acknowledge the distinct approaches to identifying critical elements in these perspectives and strive to offer nuanced, causal insights into multi-view feature learning.
- Our framework integrates a text-based semantic view with a graph-based topological view, incorporating counterfactual reasoning to seamlessly merge diverse representation sets into a comprehensive multi-view analysis. This approach leverages supervised contrastive learning to sharpen the distinctions among multi-view features, enabling a precise assessment of the impact of critical elements across different perspectives on the feature learning of claims and evidence.
- Our framework undergoes experimental evaluation on real datasets, showing superior performance over current state-of-the-art methods in fact-checking fake news detection. Additionally, by examining critical elements from various angles within related evidence, our framework enhances the interpretability of category probability assessments.



An example about text-based semantic and graph-based topological learning.

## 2 Problem Definition

- Fact-checking fake news detection aims to classify observation instances into two predefined categories  $y = \{0, 1\}$  ("1" for False and "0" for True), which is achieved through a comprehensive analysis of the claim and related evidence. Each observation instance consists of a claim, denoted as  $c$ , and a set of evidence gathered from reliable website sources, represented as  $E = \{e_1, e_2, \dots, e_n\}$ , where  $n$  represents the number of evidence. The model predicts the probability of authenticity  $\hat{y} = f(c, E, \theta)$ , where  $f$  represents the prediction function, and  $\theta$  refers to the trainable parameters.

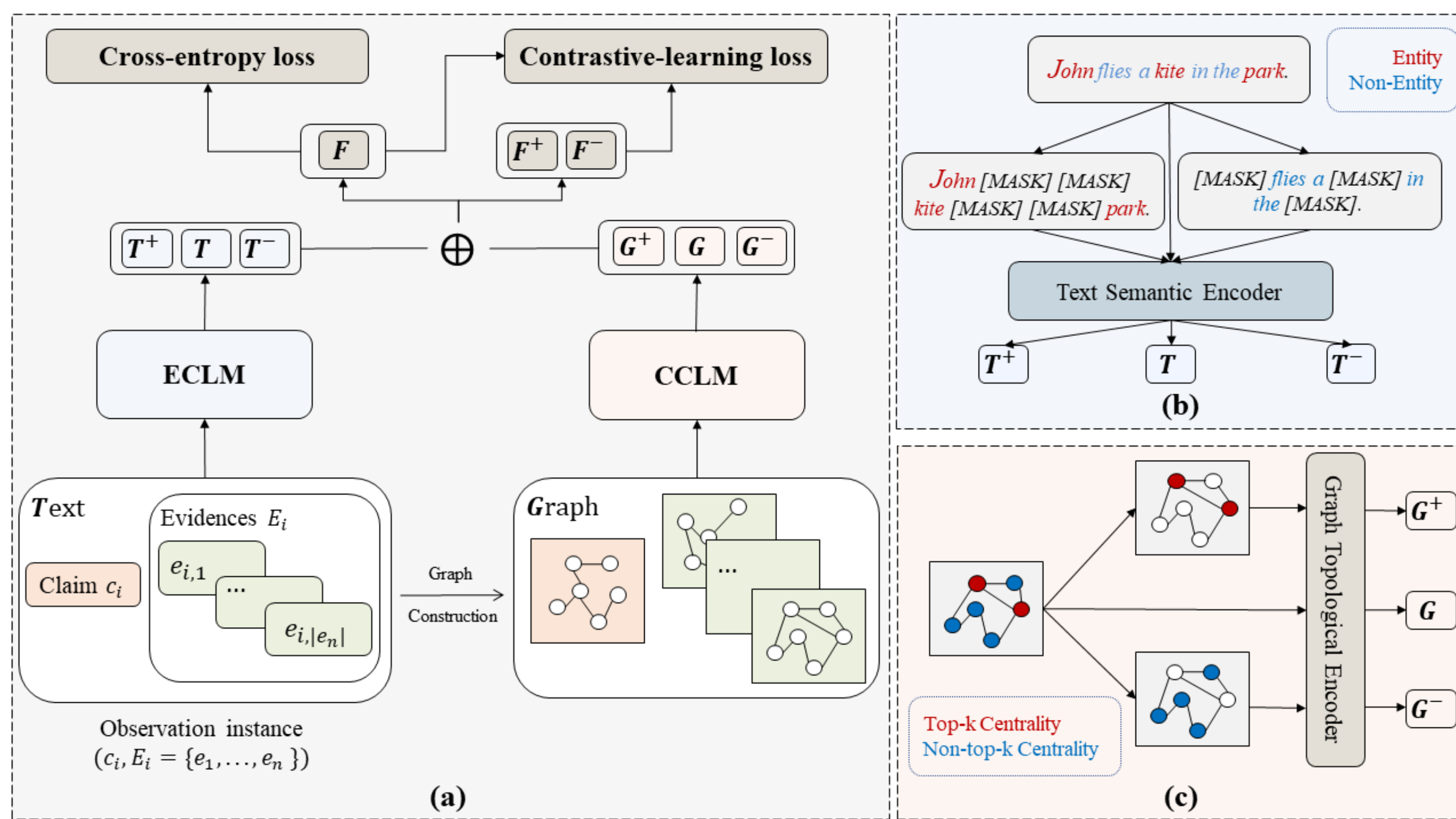
Claim	id_left	claim_text	claim_source
	1535	cut taxes 23 times mayor new york city	rufy giuliani
Evidences	id_right	evidence_text	evidence_source
	12705	the new york times site search navigation site advertisement ny region ...	nytimes.com
	12706	many of these proposals money is going to be coming off the baskets ...	politico.com
12707	been proud in cattaugaus county to convene several very successful ...	ny.gov	
Cred_label	True		

A sample fact-checking scenario involving a claim and its corresponding evidence.

## 3 Methodology

### Overall Framework:

This paper presents MCCL, an innovative multi-view classification framework for fact-checking fake news detection. As illustrated in the figure, the framework's architecture consists of three core components: Entity-view Counterfactual Learning Module (ECLM), Centrality-view Counterfactual Learning Module (CCLM) and Prediction and Contrastive Learning Module.



The MCCL overview for a claim and its related evidence.

### Entity-view Counterfactual Learning

- Text-based Semantic Representation Extraction.
- Text Counterfactual Instance Formulation.

### Centrality-view Counterfactual Learning

- Graph-based Topological Representation Extraction.
- Graph Counterfactual Instance Formulation.

### Prediction and Contrastive Learning

- Multi-view Feature Fusion and Prediction Loss.
- Counterfactual Contrastive Learning Loss.

### Training Objective

## 4 Experiments

### Dataset:

Dataset	# True	# False	# Evidence
Snopes	1164	3177	29242
PolitiFact	1867	1701	29556

Quantitative information of Snopes and PolitiFact.

### Ablation Study:

Method	Snopes		PolitiFact	
	F1-Ma	F1-Mi	F1-Ma	F1-Mi
MCCL	<b>0.812</b>	<b>0.859</b>	<b>0.705</b>	<b>0.706</b>
w/o ECL	0.799	0.846	0.692	0.696
w/o CCL	0.796	0.841	0.688	0.692
w/o CL	0.803	0.850	0.694	0.695
w/o ECL&CCL	0.791	0.837	0.686	0.689
w/o CCL&CL	0.787	0.832	0.685	0.687

### Model Comparison:

The performance comparison between MCCL and its variants.

Method	Snopes								PolitiFact							
	F1-Ma	F1-Mi	F1-T	P-T	R-T	F1-F	P-F	R-F	F1-Ma	F1-Mi	F1-T	P-T	R-T	F1-F	P-F	R-F
LSTM	0.621	0.719	0.430	0.484	0.397	0.812	0.791	0.837	0.606	0.609	0.618	0.632	0.613	0.593	0.590	0.604
TextCNN	0.631	0.720	0.450	0.482	0.430	0.812	0.799	0.826	0.604	0.607	0.615	0.630	0.610	0.592	0.591	0.604
BERT	0.621	0.716	0.431	0.477	0.407	0.810	0.793	0.830	0.597	0.598	0.608	0.619	0.599	0.586	0.577	0.597
DeClarE	0.725	0.786	0.594	0.610	0.579	0.857	0.852	0.863	0.653	0.652	0.675	0.667	0.683	0.631	0.637	0.625
HAN	0.752	0.802	0.636	0.625	0.647	0.868	0.876	0.861	0.661	0.660	0.679	0.676	0.682	0.643	0.650	0.637
EHAN	0.784	0.828	0.684	0.617	0.768	0.885	0.882	0.890	0.676	0.679	0.689	0.686	0.693	0.655	0.675	0.636
MAC	0.786	0.833	0.687	0.700	0.686	0.886	0.886	0.887	0.672	0.673	0.718	0.675	0.735	0.643	0.676	0.617
CICD	0.789	0.837	0.691	0.632	<b>0.775</b>	0.893	0.890	0.895	0.682	0.685	0.702	0.689	0.714	0.657	0.691	0.629
GETRAL	0.800	0.846	0.705	0.721	0.694	0.895	0.890	0.902	0.691	0.694	0.723	0.687	<b>0.764</b>	0.660	0.708	0.629
MCCL	<b>0.812</b>	<b>0.859</b>	<b>0.718</b>	0.754	0.714	<b>0.906</b>	<b>0.893</b>	<b>0.921</b>	<b>0.705</b>	<b>0.706</b>	<b>0.728</b>	<b>0.711</b>	0.752	<b>0.681</b>	<b>0.710</b>	<b>0.658</b>

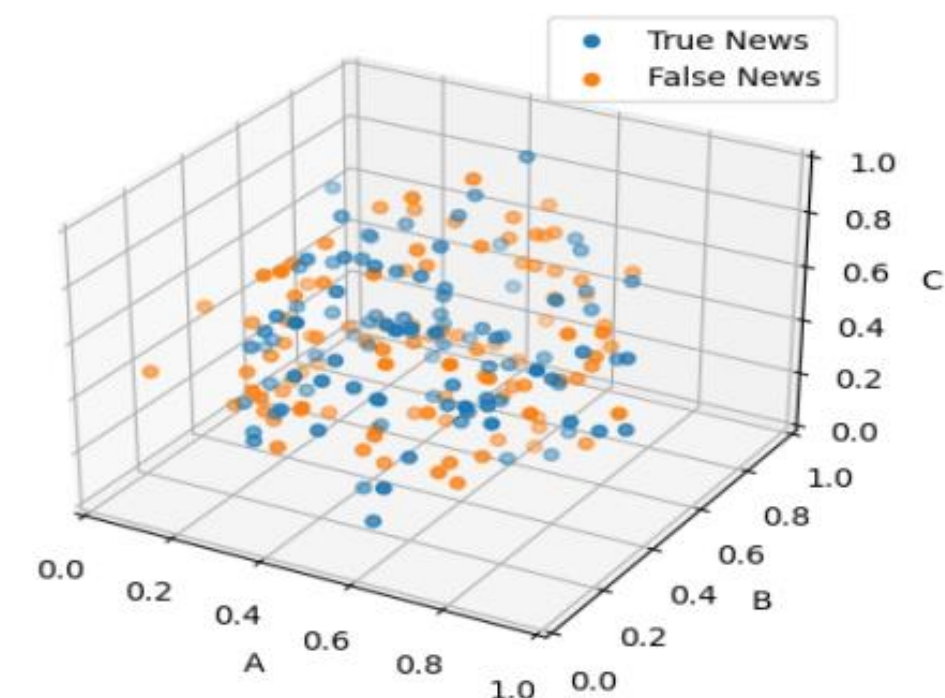
The model comparison results.

### Visualization of Counterfactual Reasoning:

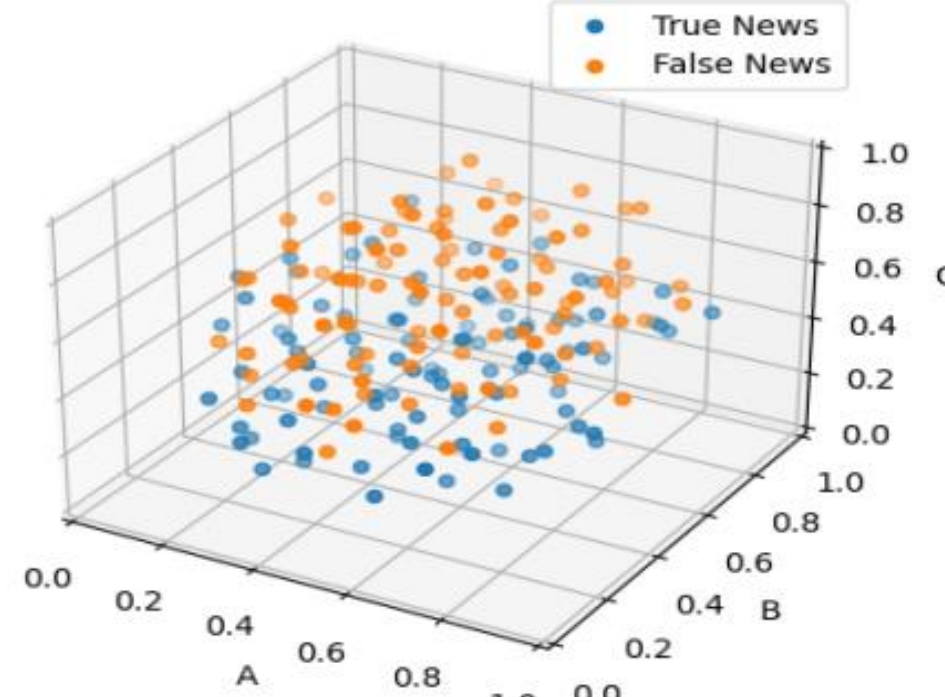
Claim	cut taxes 23 times mayor new york city
ECL	the new york times site search navigation site navigation site mobile navigation advertisement ny region bloomberg plans to cut new york s property taxes by diane 172007 mayor michael r bloomberg is proposing to cut property taxes by roughly 5 percent and eliminate the city sales tax on clothing and footwear as new york enjoys the bounty from its booming economy and real estate market city officials said yesterday mr bloomberg plans to outline his proposals today in his annual address to the city council the property tax cuts which would be in addition to an existing 400 annual
CCL	the new york times site search navigation site navigation site mobile navigation advertisement ny region bloomberg plans to cut new york s property taxes by diane 172007 mayor michael r bloomberg is proposing to cut property taxes by roughly 5 percent and eliminate the city sales tax on clothing and footwear as new york enjoys the bounty from its booming economy and real estate market city officials said yesterday mr bloomberg plans to outline his proposals today in his annual address to the city council the property tax cuts which would be in addition to an existing 400 annual

Visualization of critical elements in color.

### Visualization of Contrastive Learning:



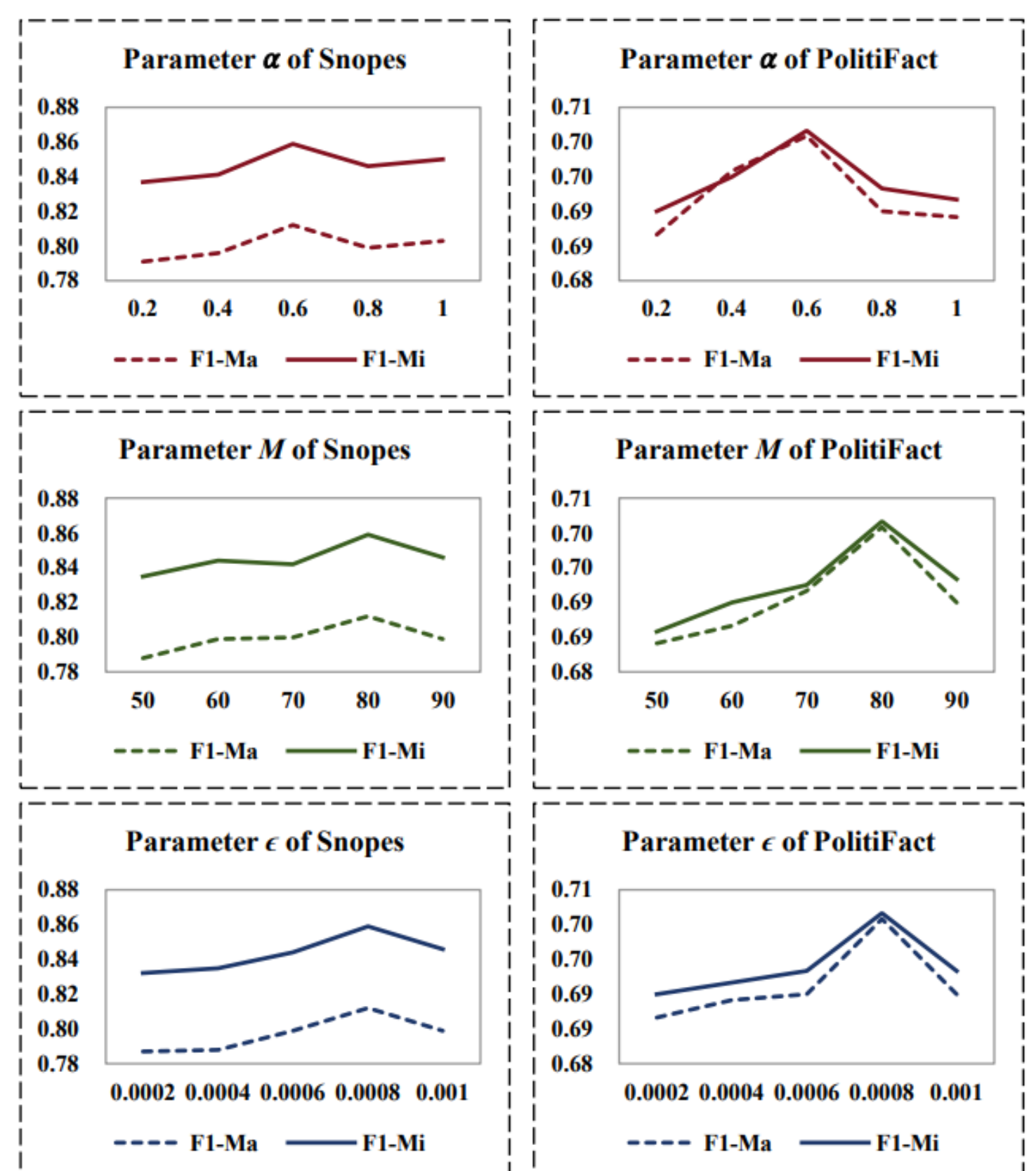
(a) The multi-view feature distribution of the same observation instance before counterfactual contrastive learning.



(b) The multi-view feature distribution of the same observation instance after counterfactual contrastive learning.

Distribution of multi-view representation vectors.

### Parameter Analysis:



The influence of parameters in the certainty algorithm on framework performance.