A Dual-Branch Fake News Detection Framework Using BERT and Knowledge Graph Embeddings

Mohamed Ahmed Mansour Mahmoud
Email: mohammmmmmmmed852963@gmail.com

***Abstract*—The rapid spread of misinformation across online platforms poses serious social and political risks. Text-only classifiers capture linguistic cues but struggle to assess factual consistency. We propose a *dual-branch fake news detection framework* that integrates a *text branch* powered by BERT with a *knowledge branch* based on TransE embeddings. News articles are converted into factual triples, embedded in a knowledge space, and fused with BERT semantics via a lightweight MLP classifier. On the LIAR dataset, our current implementation reports BERT (text only) at 0.6425 accuracy and 0.6797 F1, while the Fusion model achieves 0.6339 accuracy and 0.7715 F1, indicating that factual grounding substantially improves F1 even when accuracy is comparable. We release an end-to-end pipeline and a Gradio GUI for real-time predictions.**

***Index Terms*—Fake news detection, NLP, BERT, Knowledge graphs, TransE, Misinformation**

I. Introduction

The ubiquity of social platforms has enabled misinformation to travel faster and farther than most fact-checking pipelines can handle. Automated detection systems are therefore essential, yet text-only methods primarily rely on stylistic and contextual features that can be adversarially imitated. Fact-checked reasoning requires access to external knowledge and a representation that captures factual consistency beyond prose.

Knowledge graphs (KGs) encode factual triples (h, r, t) and thus provide a complementary signal to text semantics. However, tightly coupling KG retrieval and reasoning inside a transformer can be heavy and brittle. We argue for a pragmatic middle ground: a *dual-branch* approach that processes text and knowledge independently and fuses the two at the representation level.

Contributions.

- **Architecture:** a simple, efficient dual-branch model combining BERT semantics with TransE knowledge embeddings.
- **End-to-end pipeline:** triplet extraction, TransE training, representation fusion, and a Gradio GUI for interactive use.
- **Evidence:** on LIAR, fusion improves macro-F1 over text-only baselines, indicating that factual grounding counters stylistic bias.

II. Related Work

Text-only detection. Pretrained transformers (BERT [1], RoBERTa [2], DeBERTa [3]) form strong baselines across rumor detection and stance classification. Datasets such as LIAR [4] and FakeNewsNet [5] standardize evaluation; nevertheless, text-only models can be misled by style or topical bias.

KG embeddings. TransE [6] models relations as translations; ComplEx [7] and RotatE [8] extend expressivity. Fact verification with KGs (e.g., FEVER [9]) often relies on retrieval and symbolic reasoning.

Knowledge-augmented transformers. KnowBERT [10], ERNIE [11], K-BERT [12], KEPLER [13], and KG-BERT [14] inject entity knowledge into PLMs through attention, joint objectives, or pretraining—powerful but complex. Our design is deliberately lean: learn KGE separately, then fuse with text via an MLP.

III. Methodology

A. Text Branch (BERT)

Given tokenized article x, BERT yields contextual states; we use the `[CLS]` vector as $\mathbf{h}_{\text{text}} \in \mathbb{R}^{d_t}$. Optionally, we apply dropout and a linear projection to match d_k.

B. Knowledge Branch (TransE)

For each article, a REBEL-style extractor yields triples $\mathcal{T}(x) = \{(h_i, r_i, t_i)\}$. TransE learns entity/relation embeddings $(\mathbf{e}_h, \mathbf{r}, \mathbf{e}_t) \in \mathbb{R}^{d_k}$ using the margin-ranking loss:

$$\mathcal{L} = \sum_{(h,r,t)\in\mathcal{T}} \left[\gamma + \|\mathbf{e}_h + \mathbf{r} - \mathbf{e}_t\|_p - \|\mathbf{e}_{h'} + \mathbf{r} - \mathbf{e}_{t'}\|_p\right]_+ . \tag{1}$$

We aggregate entities per article via mean pooling to $\mathbf{h}_{\text{kg}} \in \mathbb{R}^{d_k}$ (zeros if no entities are present).

C. Fusion and Classification

We concatenate $\mathbf{z} = [\mathbf{h}_{\text{text}}; \mathbf{h}_{\text{kg}}]$ and pass through a shallow MLP:

$$\hat{y} = \operatorname{softmax}\big(\mathbf{W}_2\, \sigma(\mathbf{W}_1\mathbf{z} + \mathbf{b}_1) + \mathbf{b}_2\big). \tag{2}$$

We optimize cross-entropy with optional class weights and weight decay.

D. Training Pipeline

IV. Experimental Setup

A. Datasets

We use LIAR [4] and reference FakeNewsNet [5]. Table I shows LIAR statistics.

This work was carried out under the supervision of **Professor Ramakrishna**. Code and assets: https://github.com/x50MANSOUR50x/Dual-Branch-Fake-News-Detection-Framework.

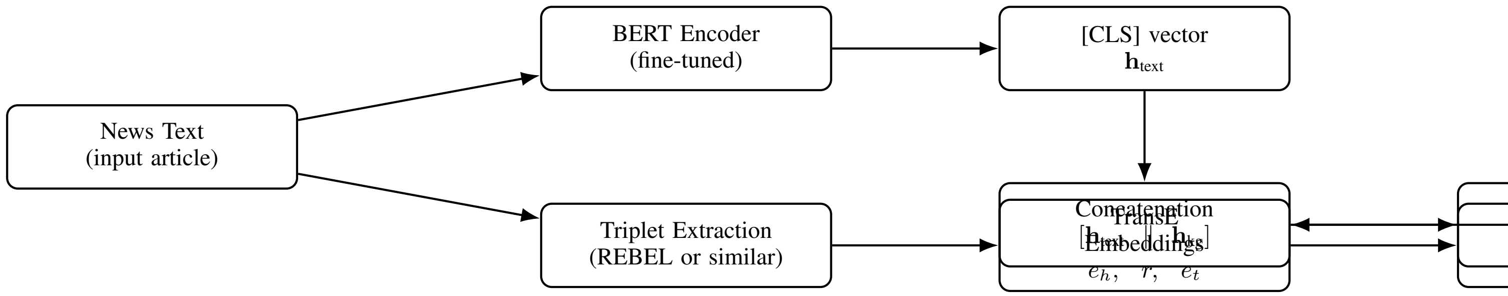

Fig. 1. Proposed dual-branch architecture. The text branch encodes each article with BERT to obtain the [CLS] vector $\mathbf{h}_{\text{text}}$. The knowledge branch extracts triples, embeds them with TransE, mean-pools entity embeddings to $\mathbf{h}_{\text{kg}}$, and we fuse $[\mathbf{h}_{\text{text}} \| \mathbf{h}_{\text{kg}}]$ with a shallow MLP for FAKE/REAL.

Algorithm 1 End-to-end Training

Train TransE on all extracted triples $\mathcal{T}$ to obtain entity embeddings.
for each article x **do**
    Encode x with BERT to obtain $\mathbf{h}_{\text{text}}$.
    Compute $\mathbf{h}_{\text{kg}}$ by mean-pooling entity vectors for entities in $\mathcal{T}(x)$.
    Concatenate $\mathbf{z} = [\mathbf{h}_{\text{text}}; \mathbf{h}_{\text{kg}}]$ and update the MLP via cross-entropy.
end for

TABLE I
LIAR DATASET STATISTICS.

Split	# Instances	Avg. tokens
Train	10,269	35
Valid	1,284	35
Test	1,266	35

B. Preprocessing

We lowercase, remove URLs, and keep punctuation. For triple extraction we keep relations with confidence $\geq \tau$ (default 0.5) and map surface forms to KG entities with a simple heuristic resolver; unresolved entities are dropped. Articles without valid entities get $\mathbf{h}_{\text{kg}} = \mathbf{0}$.

C. Implementation Details

PyTorch + HuggingFace Transformers. BERT-base uncased, max length 128, batch size 32, AdamW with lr 2×10^{-5} (text) and 1×10^{-3} (TransE), 5–10 epochs. Fusion MLP hidden size 256, dropout 0.2. For KG-only baselines we train a logistic regression over pooled entity vectors. Inference and demos are served via a Gradio GUI.

D. Baselines and Metrics

Baselines: (i) BERT (text only); (ii) TransE (KG only). Metrics: Accuracy and macro-F1.

TABLE II
PERFORMANCE ON LIAR TEST SET. BEST IN **BOLD**.

Model	Accuracy	F1 (macro)	Notes
BERT (text only)	0.6425	0.6797	–
TransE (KG only)	–	–	val_loss = 0.3785
Fusion (Ours)	0.6339	**0.7715**	–

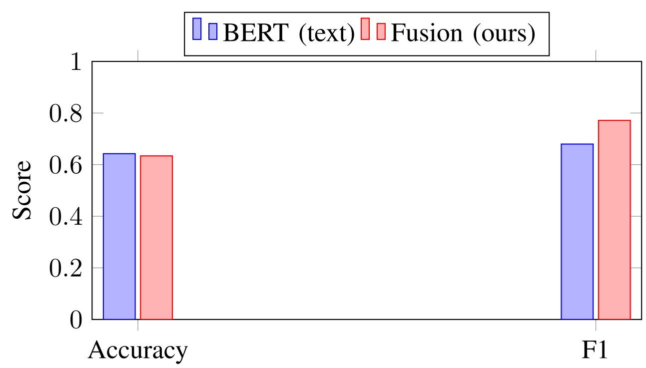

Fig. 2. BERT vs. Fusion on LIAR.

V. RESULTS AND DISCUSSION

A. Visualization

B. Analysis

Fusion notably boosts macro-F1, suggesting that KG information helps balance minority labels where textual style alone is insufficient. Accuracy is similar because decision thresholds favor majority classes; macro-F1 exposes improvements on harder cases. When triples are sparse or noisy, the fusion branch gracefully degrades to text-only behavior via dropout and learned weights.

C. Ablations

D. Error Analysis

We categorize typical errors into (i) satire/irony where semantics look plausible but world knowledge contradicts the claim; (ii) entity linking mistakes that send the KG branch to the wrong node; and (iii) temporal drift (facts true in the past).

TABLE III
ABLATION TEMPLATE (FILL WITH YOUR RUNS).

Setting	Accuracy	F1
Mean pooling (default)	0.6339	0.7715
Max pooling	–	–
Gated fusion	–	–
Higher triple threshold	–	–

Future improvements include temporal KGs and robust entity disambiguation.

E. Scalability

The knowledge branch is lightweight: TransE training is linear in #triples and inference pools a handful of entity vectors per article. The fusion MLP adds negligible parameters compared to BERT, keeping latency close to text-only inference.

VI. BROADER IMPACT AND LIMITATIONS

Impact. The framework can assist fact-checkers and educators by providing rapid signals for suspicious content. **Limitations.** Triplet extractors may hallucinate or miss facts; KGs are incomplete and time-sensitive. Domain shift (memes, multimodal content) remains challenging. **Future work.** Richer KGE models (RotatE/ComplEx), temporal and multilingual KGs, retrieval-augmented evidence, and explanation modules that highlight influential triples and tokens.

VII. REPRODUCIBILITY

We provide code, models, and a GUI. For deterministic runs, fix random seeds, freeze package versions, and log configuration. Results in Table II were obtained with the repository at commit time of writing.

VIII. CONCLUSION

We presented a practical dual-branch framework that fuses BERT semantics with TransE knowledge embeddings for fake news detection. On LIAR, fusion improves macro-F1 over a strong BERT baseline while maintaining comparable accuracy. The approach is simple, efficient, and compatible with real-time usage via a GUI.

APPENDIX A
TRIPLET EXTRACTION DETAILS

We use a REBEL-style model with beam search $k{=}4$ and confidence threshold $\tau{=}0.5$. Entities are canonicalized via lowercasing and punctuation stripping; unresolved mentions are dropped. For future work, we plan to add a lightweight linker against Wikidata aliases.

APPENDIX B
GUI AND DEPLOYMENT

The Gradio interface wraps the fusion model with preprocessing hooks for tokenization and triple extraction. We cache KGE lookups to minimize latency and expose an optional "evidence" panel with top triples used.

REFERENCES

[1] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. NAACL*, 2019.

[2] Y. Liu *et al.*, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv:1907.11692, 2019.

[3] P.-S. He *et al.*, "DeBERTa: Decoding-enhanced BERT with Disentangled Attention," in *Proc. ICLR*, 2021.

[4] W. Y. Wang, ""Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection," in *Proc. ACL*, 2017.

[5] K. Shu, D. Mahudeswaran, S. Wang, and H. Liu, "FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information," *Big Data*, 8(3):171–188, 2020.

[6] A. Bordes, N. Usunier, A. Garcia-Durán, J. Weston, and O. Yakhnenko, "Translating Embeddings for Modeling Multi-relational Data," in *Proc. NeurIPS*, 2013.

[7] T. Trouillon *et al.*, "Complex Embeddings for Simple Link Prediction," in *Proc. ICML*, 2016.

[8] Z. Sun *et al.*, "RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space," in *Proc. ICLR*, 2019.

[9] J. Thorne *et al.*, "FEVER: a Large-scale Dataset for Fact Extraction and VERification," in *Proc. NAACL*, 2018.

[10] M. Peters *et al.*, "Knowledge Enhanced Contextual Word Representations," in *Proc. EMNLP*, 2019.

[11] Y. Sun *et al.*, "ERNIE: Enhanced Language Representation with Informative Entities," arXiv:1905.07129, 2019.

[12] W. Liu *et al.*, "K-BERT: Enabling Language Representation with Knowledge Graph," in *Proc. AAAI*, 2020.

[13] X. Wang *et al.*, "KEPLER: A Unified Model for Knowledge Embedding and Pre-trained Language Representation," *TACL*, 2021.

[14] L. Yao *et al.*, "KG-BERT: BERT for Knowledge Graph Completion," arXiv:1909.03193, 2019.