Howard University-AI4PC at SemEval-2025 Task 7: Crosslingual Fact-Checked Claim Retrieval-Combining Zero-Shot Claim Extraction and KNN-Based Classification for Multilingual Claim Matching

Suprabhat Rijal and **Saurav K. Aryal**
EECS, Howard University
Washington, DC 20059, USA
saurav.aryal@howard.edu

Abstract

SemEval Task 7 introduced a dataset for multilingual and cross lingual fact checking. We propose a system that leverages similarity matching, KNN, zero-shot classification and summarization to retrieve fact-checks for social media posts across multiple languages. Our approach achieves performance within the expected range, aligning with baseline results. Although competitive, the findings highlight the potential and challenges of zero-shot methods, providing a foundation for future research in cross lingual information verification.

1 Introduction

The rapid spread of misinformation on social media platforms has highlighted the urgent need for automated fact-checking systems that can operate across multiple languages. Multilingual and cross lingual fact checking pose significant challenges, as they require systems to verify claims in diverse linguistic contexts while maintaining high accuracy and scalability (Ngueajio et al., 2025; Washington et al., 2021). To address these challenges, SemEval Task 7 introduced a benchmark dataset designed to evaluate the performance of such systems, providing a platform for advancing research in this domain (Peng et al., 2025).

Our proposed approach to cross lingual fact checking leverages zero-shot classification and summarization techniques. Using zero-shot methods, our system retrieves fact checks associated with social media posts across multiple languages without requiring language-specific training data. This strategy enables the system to bridge the gap between high-resource and low-resource languages, promoting more equitable access to fact-checking tools (Aryal et al., 2023b,d). The zero-shot framework allows the system to generalize across languages, making it scalable and adaptable to diverse linguistic contexts.

Our results demonstrate that the proposed system achieves performance within the expected range, aligning with baseline results reported in prior work. Although the results are competitive, they also reveal the inherent challenges of zero-shot methods, particularly in claim classification tasks. These findings validate the feasibility of zero-shot approaches for cross lingual fact-checking and provide a foundation for future research.

To support reproducibility and further research, we have released the code for our system. The code can be found in Appendix A.

2 Background

2.1 Task Summarization

The task comprised two distinct tracks: Monolingual (Track 1) and Cross Lingual (Track 2). In the Monolingual track, the objective was to match social media posts with fact checks written in the same language across multiple languages. In contrast, the Cross Lingual track required matching social media posts with fact-checks across different languages, addressing the challenge of verifying claims in multilingual contexts. Our work aims to tackle Track 2.

2.2 Dataset

This dataset contains social media posts, each including images with extracted text (OCR) and accompanying captions, presented in both the original language and English translations. These posts were paired with fact checks that also contained claims in the source language and their English translations. The development phase covered eight languages, while the testing phase introduced two additional unseen languages.

2.3 Related Work

Cross lingual claim matching, a critical task in misinformation detection and fact checking, has

Proceedings of the The 19th International Workshop on Semantic Evaluation (SemEval-2025), pages 1777–1782
July 31 - August 1, 2025

received relatively limited attention in the research community, as evidenced by recent surveys (Panchendrarajan and Zubiaga, 2024). Existing approaches have predominantly relied on embedding-based representations of claims to retrieve relevant fact-checks using similarity metrics, as exemplified by the authors of the MultiClaim dataset. Furthermore, the use of fine-tuned multilingual models, such as mBERT (Devlin et al., 2019), for claim classification—as explored in the MMTweets framework(Singh et al., 2024)—has shown promise but has not yet achieved consistent performance in cross lingual settings. Existing literature indicates that claim extraction techniques using LLMs yield promising results when applied to fact-checking tasks (Sundriyal et al., 2023). Given these limitations, our work seeks to experiment with alternative approaches, combining semantic similarity and classification-based methods while incorporating zero-shot techniques to explore their potential for improving cross lingual claim matching.

3 System Overview

Our cross lingual claim matching system is composed of three main components: Text Translation, Knowledge Base Creation and Fact-Check Retrieval.

3.1 Text Translation

Both the social media post information (OCR text and caption text) and verified fact checks are translated into English. For this step, we utilize the translated social media OCR text and caption text provided by the MultiClaim Dataset which uses Google Translate API for translation.

3.2 Knowledge Base Creation

To enable efficient claim retrieval, we first generate vector embeddings for all fact-checked claims using a pre-trained language model (all-MiniLM-L6-v2). These embeddings capture the semantic representations of the claims and are stored in a vector database. The database facilitates cosine-similarity-based searches, allowing for the identification of semantically similar claims during the retrieval process.

3.3 Fact-Check Retrieval

The retrieval process involves the following steps:

- **Claim Extraction:** A large language model, Deepseek-r1:14B (DeepSeek-AI et al., 2025) quantization Q4_K_M, is employed to extract the core claim from the translated social media post. This is achieved using carefully designed prompts that ensure accurate and concise claim identification.

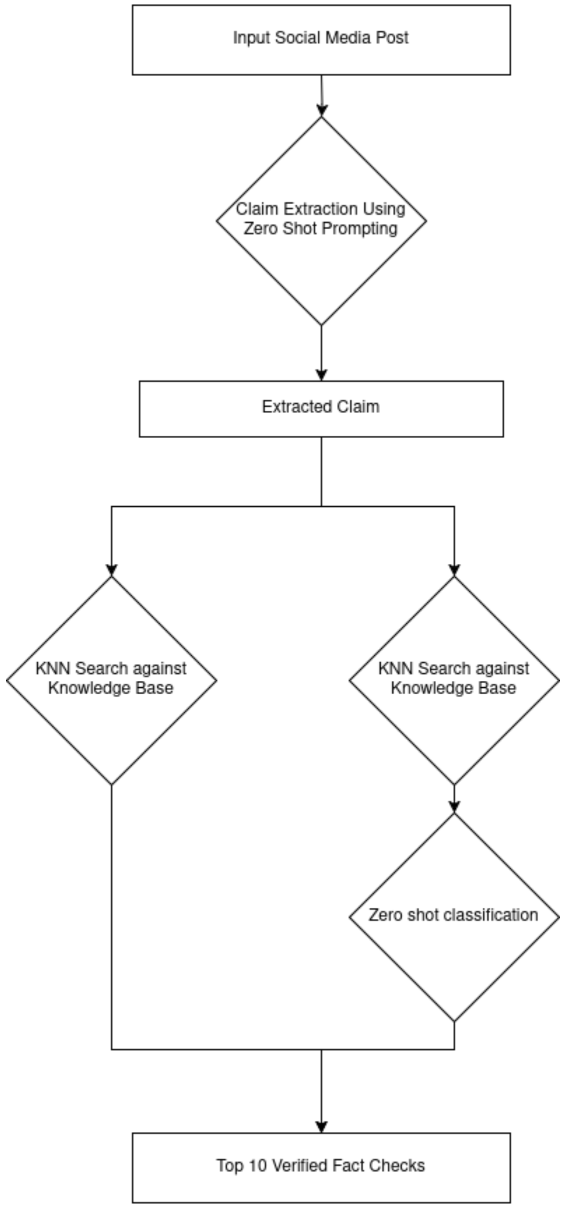

Figure 1: Crosslingual Fact Check Retrieval Pipeline

- **Knowledge Base Search:**To provide 10 fact checks for each query as required by the task, we implemented two distinct approaches:
  - **KNN Search:** The top-10 fact-check embeddings are retrieved solely based on semantic similarity to the social media claim.
  - **Zero-shot classification:** In this approach, we first retrieve the top-10 fact-

check embeddings based on cosine similarity. For each retrieved fact-check, a large language model is prompted to classify whether the social media claim and the fact-check claim are semantically equivalent. If fewer than 10 matches are identified, the process iteratively retrieves the next 20 closest vectors and repeats the classification until 10 claims are matched. This hybrid approach combines embedding-based retrieval with LLM-driven classification, significantly reducing the search space and computational overhead by avoiding classification inference over the entire "verified fact checks" search space.

4 Experimental Setup

4.1 Open Source Software used

- **Embedding Function:** To generate semantic embeddings for the claims, we utilized the `sentence-transformers` (Reimers and Gurevych, 2019). library with the `all-MiniLM-L6-v2`(Wang et al., 2020) model. This model was chosen for its efficiency and effectiveness in capturing semantic representations of text, enabling robust similarity-based retrieval (Wang et al., 2020).

- **Vector Database:** For storing and querying the embeddings, we employed `chromadb` as the vector database.

- **Hosting LLM:** The distance function we used was the standard `cosine similarity` provided by chromadb due to its invariance to magnitude.

- **Embedding Function:** The large language models (LLMs) used in our system were hosted locally using `ollama`, ensuring low-latency access and control over model configurations. We employed the structured output functionality in ollama to ensure the large language model (LLM) generated precise JSON-formatted responses.

4.2 LLM Setup for Claim Generation

For claim extraction from social media posts, we used the `deepseek-r1:14b` model with quantization `Q4_K_M`. The following prompt was designed to guide the model in generating concise and accurate claims:

```
Task: Generate a concise and
    accurate claim made by a
    social media post.

Input:
    OCR Text: {ocr}
    Social Media Caption: {text}

Output:
JSON
{
  "claim": [The claim made in the
       social media post, based on
       both the image text and the
       caption]
}

Guidelines:
    - The claim should be stated
      objectively and avoid
      subjective language.
    - If multiple claims are
      present, focus on the most
       prominent one while
      mentioning others.
    - Rephrase opinions or
      sentiments as formal
      statements.
    - Ensure grammatical
      correctness and
      professional tone.
JSON
{
  "claim": "97% of scientists
     agree climate change is real
     ."
}
```

4.3 LLM Setup for Prompt Classification

For zero-shot classification, we used the `deepseek-r1:7b` model with quantization `Q4_K_M` and the following prompt:

```
Task: Determine whether a fact-
    checked claim (Claim A) can
    verify or contradict another
    claim (Claim B).
```

```
Input:
    Claim A: {
        claim_to_check_against}
    Claim B: {claim}

Output:
JSON
{
  "can_fact_check": boolean
}

Guidelines:
    - Identify key entities and
        central themes in both
        claims.
    - Compare the claims for
        overlap, contradiction, or
        support.
    - Provide a logical
        explanation for whether
        Claim A can fact-check
        Claim B.
    - Return a boolean value
        indicating the result.
```

This prompt was designed to ensure the model could accurately assess the relationship between claims, enabling effective zero-shot classification for fact-check retrieval.

5 Results

The performance of our system was evaluated using the Success@10 (`S@10`) metric, which measures whether at least one associated fact check is successfully retrieved within the top 10 results for a given social media post.

5.1 Knowledge Base Search Results

The results for the two Knowledge Base Search methods—KNN Search and KNN + Zero-shot Classification—are presented below:

Method	S@10
KNN Search	**0.59**
KNN + Zero-shot Classification	0.47

Table 1: Results for Knowledge Base Search methods using Test dataset

The similarity search method achieves a S@10 score of **0.59** while KNN + Zero-shot classification achieves a score of **0.47** suggesting that zero-shot claim classification is ineffective.

Since we did not train any models, the dev dataset was not used.

5.2 Overall Task Leaderboard

The results for the cross lingual track are summarized in the leaderboard below:

Rank	Team Name	S@10
19	UPC-HLE	0.6385
20	JU_NLP	0.619
21	**Howard University - AI4PC**	**0.59225**
22	Word2winners	0.55425

Table 2: Leaderboard results for Crosslingual Track

Our system achieved an average (`S@10`) score of **0.59225**, securing the 21st position on the leaderboard. This result demonstrates the effectiveness of our approach in addressing the challenges of cross lingual claim matching. While there is room for improvement, our methodology shows promise, and further refinements are expected to enhance performance in future iterations.

6 Limitations and future work

Although our system combines KNN with zero-shot classification to reduce the search space, it remains computationally intensive. As highlighted in the results, simple KNN outperformed zero-shot classification, underscoring the limitations of zero-shot methods for claim matching in this context. Additionally, the translation process may lead to the loss of certain linguistic nuances and information (Sapkota et al., 2023; Aryal et al., 2023c), further contributing to the system's reduced effectiveness. Looking ahead, we aim to address these limitations by exploring multilingual embedding models and large language models (LLMs) that are better suited for cross-lingual tasks (Aryal and Prioleau, 2023; Aryal et al., 2023a). We also plan to fine-tune models specifically for claim extraction and claim matching to improve accuracy and efficiency. These refinements are expected to enhance the system's performance and scalability in multilingual fact-checking scenarios.

7 Conclusion

In this paper, we presented a system for cross lingual fact-checked claim retrieval, addressing the challenges posed by SemEval Task 7. Our approach leveraged a combination KNN search and zero-shot classification. The results demonstrated that simple similarity search methods, such as KNN, outperformed zero-shot classification for the claim matching task, achieving an average Success@10 score of 0.59225 and securing the 21st position on the leaderboard. While this performance is competitive, it highlights the potential for further refinement, particularly in improving the precision of claim classification. Future work will focus on optimizing the classification process, and exploring more robust embedding techniques.

Acknowledgements

This research project was supported in part by the Office of Naval Research grant N00014-22-1-2714. The work is solely the responsibility of the authors and does not necessarily represent the official view of the Office of Naval Research.

References

Saurav Aryal and Howard Prioleau. Howard university computer science at semeval-2023 task 12: A 2-step system design for multilingual sentiment classification with language identification. In *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)*, pages 2153–2159, 2023.

Saurav K Aryal, Howard Prioleau, and Surakshya Aryal. Sentiment analysis across multiple african languages: A current benchmark. *arXiv preprint arXiv:2310.14120*, 2023a.

Saurav K Aryal, Howard Prioleau, Surakshya Aryal, and Gloria Washington. Baselining performance for multilingual codeswitching sentiment classification. *Journal of Computing Sciences in Colleges*, 39(3): 337–346, 2023b.

Saurav K Aryal, Howard Prioleau, Gloria Washington, and Legand Burge. Evaluating ensembled transformers for multilingual code-switched sentiment analysis. In *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 165–173. IEEE, 2023c.

Saurav K Aryal, Ujjawal Shah, Howard Prioleau, and Legand Burge. Ensembling and modeling approaches for enhancing alzheimer's disease scoring and severity assessment. In *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 1364–1370. IEEE, 2023d.

DeepSeek-AI, Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, Xiaokang Zhang, Xingkai Yu, Yu Wu, Z. F. Wu, Zhibin Gou, Zhihong Shao, Zhuoshu Li, Ziyi Gao, Aixin Liu, Bing Xue, Bingxuan Wang, Bochao Wu, Bei Feng, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, Damai Dai, Deli Chen, Dongjie Ji, Erhang Li, Fangyun Lin, Fucong Dai, Fuli Luo, Guangbo Hao, Guanting Chen, Guowei Li, H. Zhang, Han Bao, Hanwei Xu, Haocheng Wang, Honghui Ding, Huajian Xin, Huazuo Gao, Hui Qu, Hui Li, Jianzhong Guo, Jiashi Li, Jiawei Wang, Jingchang Chen, Jingyang Yuan, Junjie Qiu, Junlong Li, J. L. Cai, Jiaqi Ni, Jian Liang, Jin Chen, Kai Dong, Kai Hu, Kaige Gao, Kang Guan, Kexin Huang, Kuai Yu, Lean Wang, Lecong Zhang, Liang Zhao, Litong Wang, Liyue Zhang, Lei Xu, Leyi Xia, Mingchuan Zhang, Minghua Zhang, Minghui Tang, Meng Li, Miaojun Wang, Mingming Li, Ning Tian, Panpan Huang, Peng Zhang, Qiancheng Wang, Qinyu Chen, Qiushi Du, Ruiqi Ge, Ruisong Zhang, Ruizhe Pan, Runji Wang, R. J. Chen, R. L. Jin, Ruyi Chen, Shanghao Lu, Shangyan Zhou, Shanhuang Chen, Shengfeng Ye, Shiyu Wang, Shuiping Yu, Shunfeng Zhou, Shuting Pan, S. S. Li, Shuang Zhou, Shaoqing Wu, Shengfeng Ye, Tao Yun, Tian Pei, Tianyu Sun, T. Wang, Wangding Zeng, Wanjia Zhao, Wen Liu, Wenfeng Liang, Wenjun Gao, Wenqin Yu, Wentao Zhang, W. L. Xiao, Wei An, Xiaodong Liu, Xiaohan Wang, Xiaokang Chen, Xiaotao Nie, Xin Cheng, Xin Liu, Xin Xie, Xingchao Liu, Xinyu Yang, Xinyuan Li, Xuecheng Su, Xuheng Lin, X. Q. Li, Xiangyue Jin, Xiaojin Shen, Xiaosha Chen, Xiaowen Sun, Xiaoxiang Wang, Xinnan Song, Xinyi Zhou, Xianzu Wang, Xinxia Shan, Y. K. Li, Y. Q. Wang, Y. X. Wei, Yang Zhang, Yanhong Xu, Yao Li, Yao Zhao, Yaofeng Sun, Yaohui Wang, Yi Yu, Yichao Zhang, Yifan Shi, Yiliang Xiong, Ying He, Yishi Piao, Yisong Wang, Yixuan Tan, Yiyang Ma, Yiyuan Liu, Yongqiang Guo, Yuan Ou, Yuduan Wang, Yue Gong, Yuheng Zou, Yujia He, Yunfan Xiong, Yuxiang Luo, Yuxiang You, Yuxuan Liu, Yuyang Zhou, Y. X. Zhu, Yanhong Xu, Yanping Huang, Yaohui Li, Yi Zheng, Yuchen Zhu, Yunxian Ma, Ying Tang, Yukun Zha, Yuting Yan, Z. Z. Ren, Zehui Ren, Zhangli Sha, Zhe Fu, Zhean Xu, Zhenda Xie, Zhengyan Zhang, Zhewen Hao, Zhicheng Ma, Zhigang Yan, Zhiyu Wu, Zihui Gu, Zijia Zhu, Zijun Liu, Zilin Li, Ziwei Xie, Ziyang Song, Zizheng Pan, Zhen Huang, Zhipeng Xu, Zhongyu Zhang, and Zhen Zhang. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning, 2025. URL https://arxiv.org/abs/2501.12948.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1423. URL https://aclanthology.org/N19-1423/.

Mikel K Ngueajio, Saurav Aryal, Marcellin Atemkeng, Gloria Washington, and Danda Rawat. Decoding fake news and hate speech: A survey of explainable ai techniques. *ACM Computing Surveys*, 57(7):1–37, 2025.

Rrubaa Panchendrarajan and Arkaitz Zubiaga. Claim detection for automated fact-checking: A survey on monolingual, multilingual and cross-lingual research. *Natural Language Processing Journal*, 7:100066, 2024. ISSN 2949-7191. doi: https://doi.org/10.1016/j.nlp.2024.100066. URL https://www.sciencedirect.com/science/article/pii/S2949719124000141.

Qiwei Peng, Robert Moro, Michal Gregor, Ivan Srba, Simon Ostermann, Marian Simko, Juraj Podroužek, Matúš Mesarčík, Jaroslav Kopčan, and Anders Søgaard. SemEval-2025 task 7: Multilingual and crosslingual fact-checked claim retrieval. In *Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval-2025)*, Vienna, Austria, July 2025. Association for Computational Linguistics.

Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019. URL http://arxiv.org/abs/1908.10084.

Hrishav Sapkota, Saurav Keshari Aryal, and Howard Prioleau. Zero-shot classification reveals potential positive sentiment bias in african languages translations. *ICLR Tiny Papers*, 2023.

Iknoor Singh, Carolina Scarton, Xingyi Song, and Kalina Bontcheva. Breaking language barriers with mmtweets: Advancing cross-lingual debunked narrative retrieval for fact-checking, 2024. URL https://arxiv.org/abs/2308.05680.

Megha Sundriyal, Tanmoy Chakraborty, and Preslav Nakov. From chaos to clarity: Claim normalization to empower fact-checking. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 6594–6609, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.findings-emnlp.439. URL https://aclanthology.org/2023.findings-emnlp.439/.

Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. Minilm: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 5776–5788. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

Gloria J Washington, GiShawn Mance, Saurav K Aryal, and Mikel Kengni. Abl-micro: Opportunities for affective ai built using a multimodal microaggression dataset. In *AffCon@ AAAI*, pages 23–29, 2021.

A Appendix

The code is available at https://github.com/suprabhatrijal/semeval_task_7