SciDataCon 2025



Contribution ID: 289

Type: Poster

LETNER: Label-EfficienT Named Entity Recognition for Cyber Threat Intelligence

Monday 13 October 2025 19:10 (20 minutes)

With the rise of cyber threats, automating Named Entity Recognition (NER) in open-source documents is crucial for Cyber Threat Intelligence (CTI). However, cybersecurity NER models face challenges in maintaining large annotated datasets due to the ever-evolving threat landscape. To address this, we introduce LETNER, a label-efficient NER framework that balances performance and annotation demands. LETNER features Span-CNN-Gate, a convolutional gating module that enhances span-based entity representation, and integrates metric learning to effectively capture entity-span relationships in a shared metric space, improving adaptability in low-resource settings. We also propose a systematic evaluation framework for label efficiency in supervised NER models. Experimental results demonstrate that LETNER achieves state-of-the-art label efficiency, significantly reducing annotation costs while maintaining high performance. On a complex CTI dataset with 21 fine-grained entity classes, LETNER outperforms the widely adopted Flair NER framework by 11.8\% in F1 score while using only 10\% of the training data.

Primary author: WANG, Yue (Centre for Data Science, Queensland University of Technology)

Co-authors: ZHANG, Duoyi (The Queensland University of Technology); BASHAR, Md Abul (The Queensland University of Technology); NAYAK, Richi (The Queensland University of Technology)

Presenter: WANG, Yue (Centre for Data Science, Queensland University of Technology)

Session Classification: Poster Session

Track Classification: SciDataCon Persistent Themes: Data Science and Data Analysis