Limitations, Future Work, and Conclusion

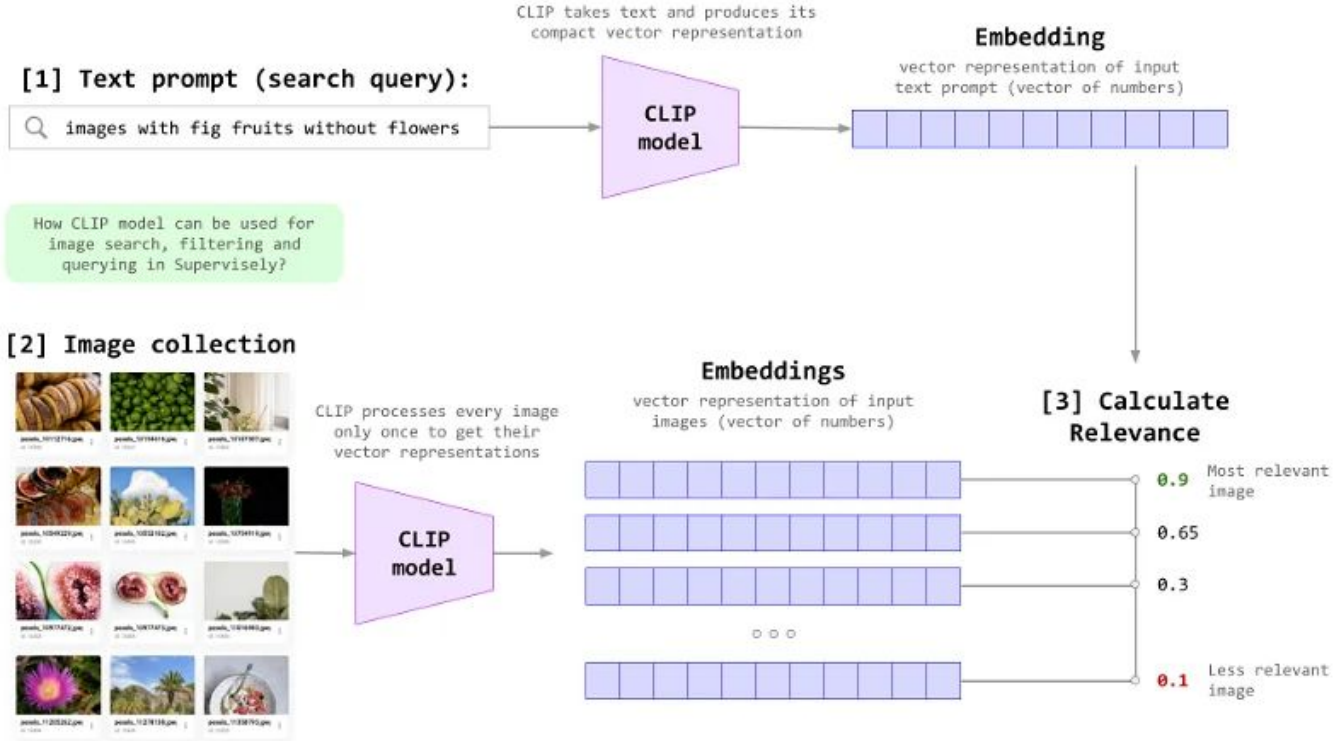
Maastricht University
Maastricht Law and Tech Lab

UNIL | Université de Lausanne

Limitations

- **Assumption:** Vendor handles Amsterdam007 and amsterdam007 belongs to one entity.
- **LLMs:** The use of chatGPT based automated systems for drafting advertisements can eliminate the individual stylometric features.
- **Architectural limitation:** Bigger architecture can lead to better performances.
- **Content-control AA:** Identifying stylometric features within a trade category may provide better representations. However, it requires either handcrafted trade labels or a classifier to identify trade categories from the text advertisements. The latter can lead to bottleneck issues.
- **Scalability:** Zero-shot capabilities of these approaches needs to be tested. Requires collecting of data from a diverse set of websites.
- **XAI:** Different frameworks lead to different explanations. A qualitative reliability test needs to be performed.
- **Environmental Impact:** Training transformers-based architecture leads to huge carbon emission and power consumption. Knowledge-Distillation to smaller architectures can help.
- **Ethical and Privacy Consideration:** Application of AA to sensitive applications (connecting HT vendors on escort markets) and processing of personal data could lead to harm on individuals.

Future Research



Conclusion

- What is Authorship Attribution? How can it be applied to DarkNet Markets
- Guidelines to Responsible Authorship Approaches
- Data Analysis and Traditional Stylometric Approaches.
- Authorship Identification as the Classification Task: BoW, traditional NNs, and Contextualized models.
- Authorship Verification as the Ranking Task: FAISS similarity
- Model Explainability
- Limitations and Future Research

VendorLink: An NLP approach for Identifying & Linking Vendor Migrants & Potential Aliases on Darknet Markets

Association for Computational Linguistics
ACL 2023

arXiv SCAN ME

GitHub SCAN ME

Maastricht University
Maastricht Law and Tech Lab

German Research Center for Artificial Intelligence

IDTraffickers: An Authorship Attribution Dataset to link and connect Potential Human-Trafficking Operations on Text Escort Advertisements

EMNLP 2023

arXiv SCAN ME

GitHub SCAN ME

Maastricht University
Maastricht Law and Tech Lab

<b> Bashpole Software, Inc.
Attention to Your Cause