

INTRODUCTION

Multiple technical documents and their versions

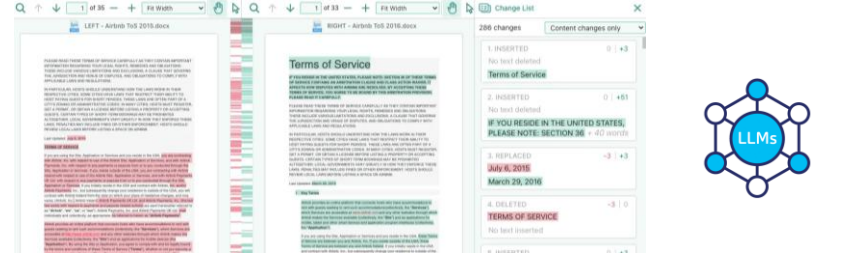


Manual processing to extract useful information

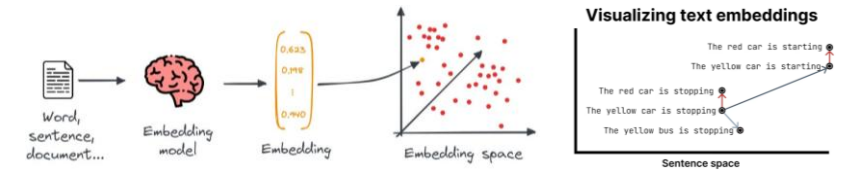


OBJECTIVES

Document versions comparison using LLMs

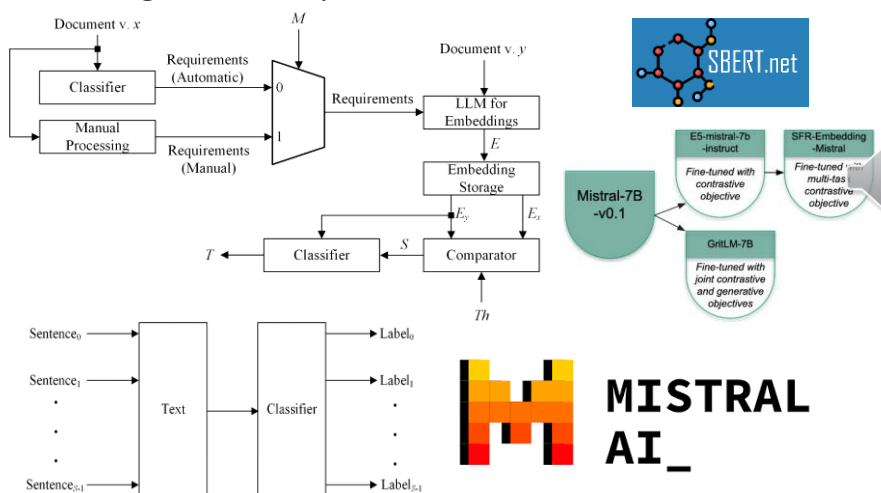


Comparing semantics by embedding vectors



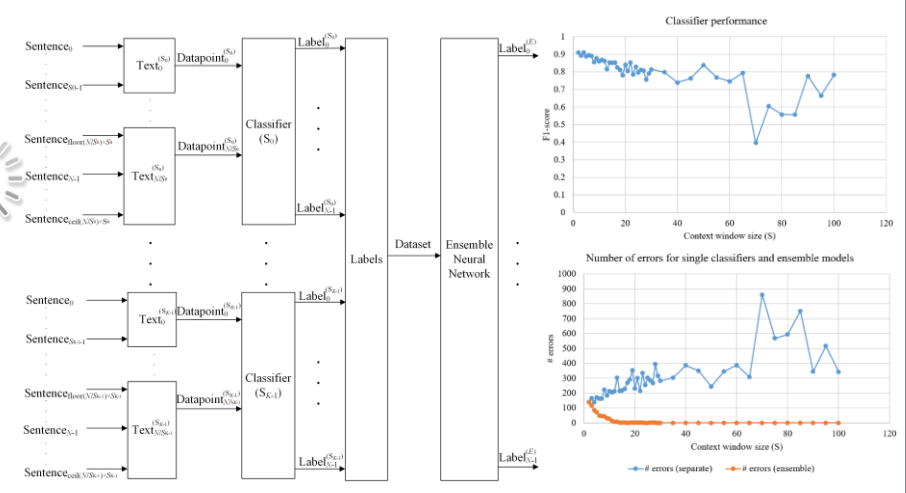
RESULTS

Embedding vectors analysis and classification

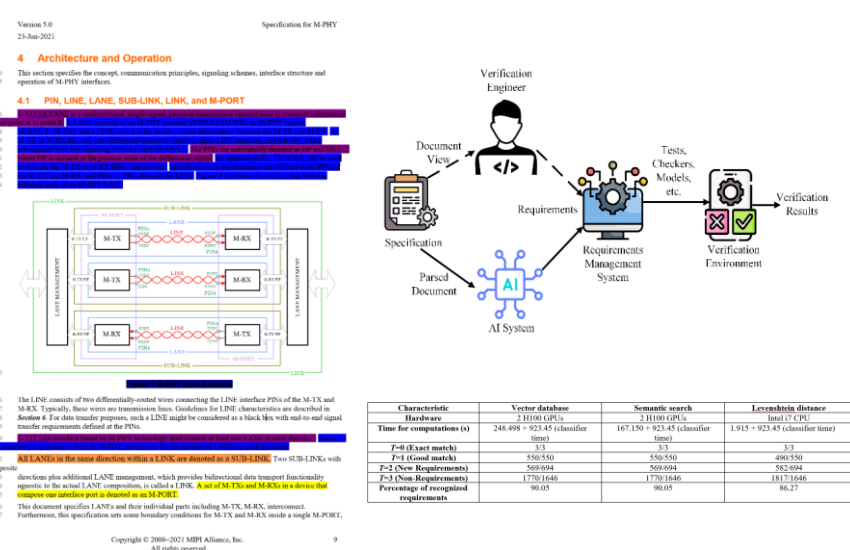


RESULTS

Ensemble model for classification



CONCLUSIONS



REFERENCES

- Ivanov, V., et al. "Extracting software requirements from unstructured documents", *International Conference on Analysis of Images, Social Networks and Texts*, Springer International Publishing, pp. 17-29, 2021.
- Arora, C., et al. "Advancing requirements engineering through generative AI: Assessing the role of LLMs", *Generative AI for Effective Software Development*, Cham: Springer Nature Switzerland, pp. 129-148, 2024.
- Reimers, N., Gurevych, I., "Sentence-bert: Sentence embeddings using siamese bert-networks.", *arXiv preprint arXiv:1908.10084*, 2019.
- Ganai, M. A., et al. "Ensemble deep learning: A review." *arXiv preprint arXiv: 2104.02395*, 2022.
- Zhang, Q., et al. "Document Parsing Unveiled: Techniques, Challenges, and Prospects for Structured Information Extraction." *arXiv preprint arXiv:2410.21169*, 2024.
- MIPI Alliance "MIPI M-PHY Specification", available: <https://www.mipi.org/specifications/m-phy> (2024).
- Reimers, N. "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.", *arXiv preprint arXiv:1908.10084*, 2019.
- NVIDIA Corporation "NVIDIA H100 Tensor Core GPU", available: <https://www.nvidia.com/en-us/data-center/h100/> (2024).



 Hugging Face

 Lightning AI

UNSTRUCTURED

Acknowledgements

SK hynix memory solutions Poland
224 Juliusza Slowackiego,
Gdansk, 80-298

SK hynix memory solutions Poland