FAITH WAVINYA MUTINDA

Email: wavinya.faith@gmail.com ◇ **Phone:** +6592781184 ◇ **Github:** `https://github.com/faith-wm`◇

SUMMARY

Results-driven professional with a proven track record in research and implementation of machine learning and natural language processing (NLP) models for diverse applications, including classification, named entity recognition, and text generation. Seeking a role in a fast-paced environment where I can leverage my technical and problem-solving skills to contribute to the development of cutting-edge AI solutions.

EDUCATION

Nara Institute of Science and Technology October 2020 – March 2023
PhD in Information Science and Engineering *Nara, Japan*

Nara Institute of Science and Technology April 2019 – September 2020
Master in Information Science and Engineering *Nara, Japan*

Osaka University April 2017 – March 2019
Research Student (Big Data Engineering Laboratory) *Osaka, Japan*

Kenyatta University May 2011 – December 2015
BSc Telecommunication and Information Technology *Nairobi, Kenya*

RELEVANT EXPERIENCE

Biofourmis Singapore PTE. Limited April 2023 – October 2023
Data Scientist (Research & Development) *Singapore*

- Conducted extensive research and evaluation of large language models (LLMs) like BERT, Llama, and ChatGPT, demonstrating their efficacy in diverse NLP applications, including named entity recognition (NER), automatic medical coding, question answering and code generation.
- Developed predictive machine learning models that achieved over 90% accuracy in identifying high-risk patients, earning recognition among top-performing solutions in the precisionFDA challenge.
- Ensured software excellence through rigorous end-to-end testing and detailed technical documentation of system specifications

Nara Institute of Science and Technology April 2019 – March 2023
Graduate Researcher *Japan*

- Fine-tuned BERT-based models to successfully extract medications and their contexts from clinical notes, achieving a top 10 ranking in the N2C2 shared task.
- Developed a transformer-based model to support automation of clinical domain meta-analysis process by extracting, structuring, analyzing, and visualizing core concepts from clinical trials articles.
- Implemented a deep-learning model for computing the degree of semantic similarity in English and Japanese clinical texts, aiding in the removal of redundant information.

SKILLS

Languages	Python, R, SQL, JavaScript, HTML, CSS, JavaScript
Tools	PyTorch, Pandas, NumPy, Scikit-learn, NLTK, spaCy, HuggingFace, Transformers, Flask, OpenAI, CUDA, Docker, Git
Interests	Machine Learning, Deep Learning, NLP, Data Science, AI, Software development
Soft skills	Problem-solving, Time management, Teamwork

PUBLICATIONS

- **Mutinda, F.W.,** Liew K., Yada, S., Wakamiya, S., & Aramaki, E. (2022). Automatic Data Extraction to Support Meta-Analysis Statistical Analysis: A Case Study on Breast Cancer. BMC Medical Informatics and Decision Making.
- **Mutinda, F. W.**, Liew K., Yada, S., Wakamiya, S., & Aramaki, E. (2022). PICO Corpus: A Publicly Available Corpus to Support Automatic Data Extraction from Biomedical Literature. In Proceedings of the First Workshop on Information Extraction from Scientific Publications. Asia-Pacific Chapter of the Association for Computational Linguistics.
- **Mutinda, F.W.,** Yada, S., Wakamiya, S., & Aramaki, E. (2022). AUTOMETA: Automatic Meta-Analysis System Employing Natural Language Processing. MEDINFO 2021: One World, One Health–Global Partnership for Digital Innovation.
- **Mutinda, F.W.,** Yada, S., Wakamiya, S., & Aramaki, E. (2021). Semantic Textual Similarity in Japanese Clinical Domain Texts Using BERT. Methods of Information in Medicine.
- **Mutinda, F.W.,** Nigo, S., Shibata, D., Yada, S., Wakamiya, S., & Aramaki, E. (2020). Detecting Redundancy in Electronic Medical Records Using Clinical BERT. The Association for Natural Language Processing.
- **Mutinda, F.W.,** Nakashima, A., Takeuchi, K., Sasaki, Y., & Onizuka, M. (2019). Time Series Link Prediction Using NMF. IEEE International Conference on Big Data and Smart Computing (BigComp).