FAITH WAVINYA MUTINDA

Email: wavinya.faith@gmail.com ◇ **Phone:** +6592781184 ◇ LinkedIn

SUMMARY

As an experienced Data Scientist, I have developed strong expertise in the development of machine learning and natural language processing (NLP) solutions for diverse tasks. Enthusiastic about leveraging data and artificial intelligence (AI) technologies to create data-driven solutions that foster innovation.

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 and GPT, demonstrating their efficacy in diverse NLP applications, including named entity recognition, question answering, and text generation.
- Developed predictive machine learning models that achieved over 90% accuracy in identifying high-risk patients, earning recognition among top-performing models in a precisionFDA challenge.
- Delivered data-driven insights by conducting in-depth exploratory data analysis, and collaborated with cross-functional teams to transform business opportunities into strategic data science solutions.
- Ensured software excellence through rigorous end-to-end testing to ensure all functionalities were met and maintained 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 unstructured clinical notes, achieving a top 10 ranking in n2c2 shared task.
- Developed a transformer-based model that streamlined the meta-analysis process, achieving F1 score over 0.80 by efficiently extracting, analyzing, and visualizing core concepts from clinical trial articles, significantly reducing the time required for manual meta-analysis.
- Implemented a transformer-based model to compute the degree of semantic similarity in clinical texts, achieving a 0.904 correlation with human scores and effectively removing redundant information.

EDUCATION

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

Nara Institute of Science and Technology April 2019 – September 2020
Master's 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*

SKILLS

Programming	Python, SQL, R
Tools	PyTorch, TensorFlow, Pandas, NumPy, Scikit-learn, NLTK, spaCy, HuggingFace, Transformers, Matplotlib, Jupyter Notebook, Flask, OpenAI, Langchain, AWS, Git
Interests	Data Science, Machine Learning, NLP, Artificial Intelligence
Soft skills	Problem-solving, Collaboration, Teamwork, Communication

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).