Jonathan Kabala Mbuya

GMU NLP, Computer Science
George Mason University, Fairfax, VA, US

✉ jmbuya@gmu.edu
☎ +1-313-288-5390
in www.linkedin.com/in/jkabala

EDUCATION

- **PhD, Computer Science** *August 2020 – June 2025 (expected)*
  George Mason University, Fairfax, VA, US
  Advisor: Antonios Anastasopoulos
  Co-Advisor: Dieter Pfoser

- **MSc., Computer Science** *August 2020 – December 2022*
  George Mason University, Fairfax, VA, US
- **BSc., (honors), Computer Science** *August 2013 – June 2017*
  Africa University, Zimbabwe

RESEARCH INTEREST

Machine learning and Deep Learning with applications in Natural Language Processing, Geospatial Analysis, and Speech

PROFESSIONAL EXPERIENCE

- **Academic Research Partner** *January 2024 - Present*
  Expedition Technology Inc., Herdon, VA, US
  ○ **Research Project**: Work on innovative approaches for trajectory anomaly detection.

- **Graduate Research Assistant** *January 2021 - Present*
  George Mason University, Fairfax, VA, US
  ○ **Applying Language Modeling Techniques on Trajectory Data**: Develop techniques to apply language modeling techniques to various problems on trajectory data, including trajectory anomaly detection and trajectory generation.
  ○ **Speech Translation for Low Resource Languages**: Develop techniques to allow speech translation for languages with scarce data.

- **Research Scientist Intern** *May 2023 - August 2023*
  Adobe, San Jose, CA, US
  ○ **Research Project**: Worked on bias detection and mitigation for text-to-image generative models.

- **Graduate Teaching Assistant** *August 2020 - May 2023*
  George Mason University, Fairfax, VA, US
  ○ **Analysis of Algorithms**: Provide learning assistance to graduate students learning Analysis of Algorithms.
  ○ **Teaching Programming**: Taught undergrad students in Computer Science object-oriented programming in Java.
  ○ **Class Logistics**: Help students with homework, projects, and exam preparation during office hours.

- **Research Engineer Intern** *May 2022 - November 2022*
  Adobe, Remote, VA, US
  ○ **Research Project**: Worked on a model-agnostic machine learning interpretability framework that quantifies the importance of high-level semantic features individually and based on interaction with other features.

- **IT Solutions Delivery Developer** *December 2017 - July 2020*
  Tenke Fungurume Mining, Democratic Republic of the Congo
  ○ **Software Development**: Developed the company's custom applications to accommodate business needs.
  ○ **Data Analysis**: Developed reports using SSRS, SSAS, and PowerBI for various company reporting needs.
  ○ **Data Integration**: Developed data interfaces between internal systems as well as external vendors to retrieve or transmit data for reporting purposes using ETL expertise and tools.

Publications

[1] Liming Zhang, **Jonathan Mbuya**, Dieter Pfoser, End-to-end Trajectory Generation - Contrasting Deep Generative Models and Language Models. ACM Transactions on Spatial Algorithms and Systems [under review]

[2] **Jonathan Mbuya**, Dieter Pfoser, and Antonios Anastasopoulos, Trajectory Anomaly Detection with Language Models. In Proceedings of the 32nd International Conference on Advances in Geographic Information Systems (Seattle, Washington) (SIGSPATIAL '24). Association for Computing Machinery

[3] Ohi Dibua, **Jonathan Mbuya**, Mackenzie Austin, Kushal Kafle, A Sobol Variance Decomposition Based Method for Understanding the Impact of Image Regions on Model Outputs. Preprint

[4] **Jonathan Mbuya**, and Antonios Anastasopoulos, GMU systems for the IWSLT 2023 dialect and low-resource speech translation tasks. In Proceedings of the 20th International Conference on Spoken Language Translation (IWSLT 2023), pages 269–276, Toronto, Canada (in-person and online), July 2023. Association for Computational Linguistics.

Programming Skills

❒ **Languages**: Python, Java, C#, JavaScript ❒ **Frameworks**: Pytorch, Keras, Scikit-learn, Tensorflow, .Net Framework ❒ **Database**: T-SQL, SQL Server ❒ **Technical Writing**: LaTeX

Relevant Courses

Theory and Applications of Data Mining (A), Machine Learning(A-), Analysis of Algorithms (A), Large Scale Optimization for Machine Learning (A), Advanced NLP (A) Deep Learning Specialization (Andrew Ng, Coursera), Mathematics for Machine Learning Specialization(Coursera)

Honors and Awards

- Distinguished graduate teaching assistant, 2023, Computer Science Department, George Mason University.
- Best computer science graduating student, 2017, Africa University.
- Best engineering student, 2017, Africa University.