# BEFIKIR BOGALE

✉ bbogale@vols.utk.edu  	in linkedin.com/in/befikir

## Education

| | |
|---|---|
| **University of Tennessee** | **Aug 2024 - Present** |
| PhD in Computer Science | Knoxville, TN |

| | |
|---|---|
| **University of Tennessee** | **Aug 2020 - May 2024** |
| Bachelor of Science in Computer Science (GPA: 3.94) | Knoxville, TN |

## Experience

**Global Computing Laboratory - University of Tennessee**  	**Aug 2024 – Present**
*Graduate Research Assistant*  	*Knoxville, TN*

- Developing a methodology to analyze performance variations across different compilers and compiler options using the RAJA Performance Suite.
- Utilizing LLNL-hosted supercomputers to conduct research and performance analysis.
- Collaborating with LLNL, participating in weekly meetings, and deliver milestone reports and software tools to enhance performance optimization efforts.

**Lawrence Livermore National Laboratory**  	**May 2024 – Aug 2024**
*Computing Graduate Scholar*  	*Livermore, CA*

- Developed a cluster-based methodology to assess performance portability using the RAJA Performance Suite on Intel Xeon and NVIDIA V100 platforms.
- Utilized Intel's top-down pipeline and advanced clustering techniques to categorize computational kernels based on performance characteristics.
- Collected and analyzed profile metrics to identify performance bottlenecks, with a focus on memory-bound kernels, achieving accurate speedup predictions.
- Collaborated with LLNL and the University of Tennessee to implement and validate the methodology across 73 RAJAPerf kernels, confirming its effectiveness in characterizing performance differentials.

**Global Computing Laboratory - University of Tennessee**  	**Oct 2022 – May 2024**
*Undergraduate Research Assistant*  	*Knoxville, TN*

- Developed software for several high-performance computing projects with C/C++, Python, etc. to achieve optimal performance optimizations.
- Improved the reproducibility and scalability of high-performance computing applications through the development of containerized images through the Apptainer and Docker platforms.
- Contributed to the ANACIN-X project, which focuses on characterizing non-determinism in MPI applications, showcasing a an understanding of parallel computing and a commitment to addressing complex challenges in scientific computing.
- Led the development of Jupyter notebooks utilized by researchers for tutorials and enhanced project accessibility.
- Collaborated with Los Alamos National Laboratory, Argonne National Laboratory, and Lawrence Livermore National Laboratory on various projects.

**Los Alamos National Laboratory**  	**June 2023 – Aug 2023**
*Parallel Computing Intern*  	*Los Alamos, NM*

- Developed proxy applications on the acceleration of X-ray transport simulations on supercomputers by implementing targeted performance optimizations, focusing on on-node efficiency.
- Ensured code portability by leveraging Kokkos , allowing effortless migration across GPU and CPU architectures and facilitating deployment on diverse high-performance computing clusters.
- Implemented and fine-tuned thread team policies to enhance parallel processing capabilities, resulting in a streamlined and scalable simulation framework.
- Successfully integrated vectorization techniques, significantly enhancing computational throughput for X-ray transport simulations.

## Research Projects

**Explaining Performance Across Compilers and Compiler Options** | *Python, Bash, Thicket*    **Aug 2024 - Present**

The project aims to develop a performance analysis methodology to explain variations in HPC application performance across different compilers and compiler options.

- Develop a methodology to analyze performance variations across different compilers and compiler options using RAJAPerf and tools like Thicket and Hatchet.
- Refine the methodology to be applicable to benchmarks more representative of standalone HPC applications, using open-source benchmarks such as AMG2023, Kripke, Lagos, and Remhos.
- Apply the refined methodology to the multiphysics code MARBL to evaluate its effectiveness in explaining performance differences across compilers and optimization levels.

**Thicket** | *C/C++, Python, Kokkos*    **Jan 2024 - Present**

Thicket is an open-source Python toolkit for Exploratory Data Analysis (EDA) of multi-run performance experiments that enables an understanding of optimal performance configuration for large-scale application codes.

- Analyzing caliper profile outputs for the MARBL multiphysics simulation library designed for exascale systems.
- Developing a scoring mechanism for ranking caliper profiles to determine optimal performance configurations.

**Checkpointing using GPU-Accelerated Data De-Duplication** | *C/C++, Python, Kokkos*    **Aug 2023 - May 2024**

Library developed with the Kokoos portability abstraction that is used for deduplicating memory using a variety of methods with varying complexity and performance.

- Implemented a comparison algorithm that reduces the number of computations in the comparison of merkel trees between checkpoints.
- Validated the performance of Kokkos hash tables in the context of merkel tree creation, find, and insertion at varying configurations.

**ANACIN-X (NSF 1900888)** | *C/C++, Python*    **Oct 2022 - May 2024**

ANACIN-X advances the reproducibility study of HPC applications through an open source modular framework for automatic measurement, analysis, and visualization of non-determinism and root causes of non-determinism in MPI applications.

- Created containerized images of ANACIN-X for multiple system architectures using Apptainer which are used by scientists as part of a tutorial apparatus in tackling non-determinism.
- Developed Jupyter notebooks that integrate multiple benchmark applications to demonstrate the practicality of ANACIN-X in identifying and visualizing non-determinism in MPI applications.

## Development Projects

**Checkpointing on Evolutionary Neural Architecture Search** | *Python, Jupyter, Slurm*

- Explored storage-efficient checkpointing methods in Evolutionary Neural Architecture Search (ENAS) to address the challenge of data scarcity for post-hoc analysis due to storage constraints.
- Developed and tested two checkpointing methods, Full and Basic, with Basic Checkpointing optimizing storage by capturing only significant model changes while maintaining prediction accuracy.
- Demonstrated the potential of incremental checkpointing to enhance memory efficiency in ENAS workflows, laying the foundation for further research into advanced checkpointing strategies and improving explainability.

**F.R.E.D: Friendly Robot to Ease Dementia** | *DART, Python, HTML/CSS, JavaScript*

- Collaborated with the NOVI team to develop the FRED mobile application, aimed at improving the lives of individuals with Alzheimer's Disease and Related Dementias (ADRD) and their caregivers.
- Focused on integrating essential features such as task management, personal reminders, and AI-driven social interaction to enhance user engagement and alleviate isolation.
- Leveraged the capabilities of the existing FRED robot to create a user-friendly, cross-platform mobile app tailored to ADRD patients' needs.

**Binary Classification of Income Using Census Data** | *Python, Jupyter*

- Explored different machine learning algorithms in predicting household income from census data.
- Utilized data preprocessing, visualization, and systematic model selection with hyperparameter tuning to find the optimal algorithm.
- Identified RBF SVM as the optimal algorithm, achieving 84% training accuracy and 82% testing accuracy, highlighting martial status and education as key elements on income prediction.

**Trip Planner** | *HTML/CSS, Python, Flask*
- Developed an online outdoor trip planning website that allows users to prepare for trips easily.
- Implemented diverse features such as scheduling, mapping, meal planning, and grocery shopping list functionalities, enriching the platform and providing users with a comprehensive suite of tools to enhance their outdoor adventure planning.

## Publications

[1] Alexander, M., Bhowmick, S., **Bogale, B.**, Diaz, G., Elster, A., Ellsworth, D. A., Hernandez, C. J. B., Jaffe, E., Marquez, J., Melton, A., Pandey, A., Parete-Koon, S., Tan, N., Taufer, M., Vergara, V., Whitnah, L., and Thiruvathukal, G. K., "EduHPC Lightning Talk Summary," in *Proceedings of the SC '23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*, ser. SC-W '23, Denver, CO, USA: Association for Computing Machinery, 2023, pp. 374–378, ISBN: 9798400707858. DOI: 10.1145/3624062.3625542. [Online]. Available: https://doi.org/10.1145/3624062.3625542.

[2] O. Pearce, J. Burmark, R. Hornung, **Befikir Bogale**, I. Lumsden, M. McKinsey, D. Yokelson, D. Boehme, S. Brink, M. Taufer, and T. Scogland, "RAJA Performance Suite: Performance Portability Analysis with Caliper and Thicket," in *Proceedings of the 2024 International Workshop on Performance, Portability, and Productivity in HPC (P3HPC)-Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis (SC24)*, Atlanta, GA, USA: IEEE Computer Society, Nov. 2024.

[3] N. Tan, K. Assogba, J. Asworth, **Befikir Bogale**, M. M. Rafique, F. Cappello, M. Taufer, and B. Nicolae, "Towards Affordable Reproducibility Using Scalable Capture and Comparison of Intermediate Multi-Run Results," in *Proceedings of the 25th ACM/IFIP International Middleware Conference (Middleware)*, Hong Kong, Chiana: ACM, Dec. 2024.

## Posters

[4] **Befikir Bogale**, O. Pearce, T. Scogland, and M. Taufer, "Cluster-Based Methodology for Characterizing the Performance of Portable Applications," presented at the 33rd ACM/IEEE International Conference for High Performance Computing and Communications Conference (SC), IEEE Computer Society, Nov. 2024.

## Talks

| | |
|---|---|
| **Supercomputing 24 Student Research Competition** | **Nov 2024** |
| *Cluster-based Methodology for Characterizing the Performance of Portable Applications* | *Atlanta, GA* |
| **Computing Scholar Symposium** | **Aug 2024** |
| *Cluster-based Methodology for Characterizing the Performance of Portable Applications* | *Livermore, CA* |
| **Los Alamos National Laboratory Student Symposium** | **Aug 2023** |
| *Accelerating X-Ray Transport Simulations* | *Los Alamos, NM* |

## Volunteering

| | |
|---|---|
| **Hack4Impact** | **Aug 2023 – Dec 2023** |
| *Programmer* | *University of Tennessee* |

- Reached out to and developed productivity software for local non-profit organizations.
- Developed an inventory and outreach website in HTML, CSS, and NodeJS.
- Improved the productivity of a non-profit to reach out to supporters and manage donation inventory through the digitization of their inventory bookkeeping.

## Relevant Coursework

- Data Structures and Algorithms
- Systems Programming
- Software Engineering
- Parallel Computing
- Machine Learning
- Cloud and Edge Computing

## Technical Skills

**Languages**: C/C++, Python, Bash, HTML/CSS, DART
**Developer Tools/Libraries**: OpenMP, CUDA/HIP, MPI, Kokkos, Apptainer, Docker, Flask, Jupyter, Caliper
**Misc**: Linux, Slurm, LSF, Flux, Spack, Git