

# Krishna Sruthi Velaga

[ksvelaga@gmail.com](mailto:ksvelaga@gmail.com) | [LinkedIn](#) | [Github](#) | [Website](#) | [Google Scholar](#) | [Towson, MD](#)

## Summary

First-year PhD student working on adaptive model compression and efficient parameter encoding in federated edge learning. Previously, I have built instrumented testbeds and conducted empirical studies on quantized large language models for smart home automation, analyzing latency, energy consumption, and thermal characteristics. I am looking for research internships in federated learning and distributed machine learning for Summer 2026.

## Experience

**Doctoral Research Assistant**, Department of Computer and Information Sciences, Towson University Aug 2025 – Present

- Reviewed 241 papers on Edge AI, mapping applications, data sources, models, and infrastructure across five urban domains.
- Develop communication-efficient model compression techniques for federated edge learning across heterogeneous devices.
- Design space-efficient parameter encoding schemes to minimize communication overhead in federated model aggregation.

**Doctoral Teaching Assistant**, Department of Computer and Information Sciences, Towson University Aug 2025 – Present

- Grade weekly programming assignments for 60 students in COSC 237, providing feedback on algorithm design approach.
- Evaluate labs for 30 students in COSC 350, focusing on networking protocols and system design.
- Hold weekly office hours, providing individualized support to help students debug code, understand concepts.

**Master's Teaching Assistant**, Department of Computer and Information Sciences, Towson University Aug 2024 – May 2025

- Analyze curriculum data from 428 universities to evaluate NIST standard adoption, supporting new course development.
- Guide 60+ undergraduates per semester in lab sessions to reinforce programming concepts and support learning outcomes.
- Assist students with debugging and problem-solving to improve lab assignment completion and technical confidence.

**Lab Administrator**, Department of Computer and Information Sciences, Towson University Aug 2023 – May 2025

- Provide technical support to 1,700+ students and faculty by resolving software, hardware, and network issues.
- Assist faculty by troubleshooting technical problems in labs and classrooms, ensuring uninterrupted instruction.
- Configure lab environments through software installation and PC setup to streamline academic workflows.

## Publications

C=Conference, J=Journal

**[C.2] Krishna Sruthi Velaga**, Anik Mallik, and Yifan Guo, “How Energy is Consumed by LLM-enabled Smart Home Assistant Systems on Low-Cost Devices: An Empirical Study”, *Proceedings of the IEEE Consumer Communications & Networking Conference (CCNC)*, 2026. [Accepted]

**[J.1] Krishna Sruthi Velaga**, Yifan Guo, and Wei Yu, “Edge AI for Smart Cities: Foundation, Challenges, and Opportunities” *Smart Cities*, 2025 [Paper]

**[C.1] Krishna Sruthi Velaga**, Yifan Guo “Optimizing Large Language Models Assisted Smart Home Assistant Systems at the Edge: An Empirical Study”, *AI4WCN AAAI* 2025 [Demo] [Paper]

## Education

**PhD in Information Technology, Computer Science Track**

Aug 2025 – Dec 2028

Towson University, Towson, MD

Advisor: Dr. Yifan Guo

GPA: 4.00/4.00

**Master of Science, Computer Science, Data Science Track**

Aug 2023 – May 2025

Towson University, Towson, MD

GPA: 3.97/4.00

**Bachelor of Technology, Computer Science and Engineering**

Aug 2019 – Jun 2023

Vasireddy Venkatadri Institute of Technology (VVIT), Guntur, AP, India

GPA: 3.63/4.00

## Projects

**LLM-enabled Smart Home Assistant Systems on Low-Cost Devices: An Empirical Study** [Paper] [Demo]

Tools: Python, LoRA, PyTorch, Raspberry Pi 4B/5, Home Assistant, HV Power Monitor

- Built a testbed with quantized LLMs on Raspberry Pi 4/5 with Home Assistant and thermal/power instrumentation.
- Collected multi-source raw data across 4 configurations accounting including log files, thermal readings, and power data.
- Synchronized heterogeneous timestamps to get fine-grained insights into performance, energy use, and thermal trade-offs.

- Reduced on-device response time by 82% (from 45.1s to 7.9s) through entity optimization, enabling faster responses.
- Improved intent recognition by fine-tuning TinyHome, TinyHome-Qwen, and StableHome models on synthetic datasets.
- Profiled workload distribution and identified inference as the dominant contributor (> 80%) to total energy consumption.
- Conducted thermal analysis; Raspberry Pi 5 showed a 44% lower temperature rise, supporting stable sustained workloads.
- Evaluated battery performance with 3,000–5,000 mAh packs at multiple health levels: Pi 4B provided longer battery life but slower responses, while Pi 5 delivered faster responses but shorter battery life.

## 911 Call Analysis of Detroit and New York City [\[Code\]](#)

Tools: *Python, TensorFlow, Matplotlib, NumPy, Pandas, Scikit-learn, Statsmodels*

- Analyzed 7M+ 911 call data from Detroit and NYC to uncover patterns in call volume, response time, and incident types.
- Built predictive models (SARIMAX, XGBoost, DBSCAN, K-Means) achieving ~85% accuracy in incident forecasting.
- Applied inferential statistics and clustering to identify peak crime times and delays in emergency response allocation.

## Predicting Ride Fares with Reinforcement Learning

Tools: *Python, NumPy, Pandas, Scikit-learn, Matplotlib*

- Processed NYC Taxi trip records with feature engineering (distance, time, passenger count, weather, and location data).
- Leveraged Q-Learning, DQN to predict optimal ride fares, formulating fare estimation as a state-action-reward problem.
- Visualized spatial and temporal pricing trends, revealing surge patterns in high-demand zones and times.

## Malicious URL Detection using Machine Learning [\[Code\]](#)

Tools: *Python, BeautifulSoup4, GoogleSearch-Python, Scikit-learn*

- Developed a Chromium-based browser extension to alert users about malicious websites, enhancing online safety.
- Engineered 24 handcrafted URL features across lexical, host-based, and content-based categories for model training.
- Increased classification accuracy from 91.8% to 97.3% by integrating content-based features in the Random forest model.

## Skills

---

**Languages:** Python, Java, SQL, C++, Bash, L<sup>A</sup>T<sub>E</sub>X

**Libraries:** Scikit-learn, TensorFlow, PyTorch, Keras, Pandas, NumPy, Matplotlib, Seaborn

**Edge Computing:** Raspberry Pi 4/5, Power Monitoring Devices

**Technologies:** Docker, Git, Home Assistant, Virtual Machines

## Leadership

---

**Vice Chairperson**, Association for Computing Machinery Student Chapter, VVIT

- Collaborated with chapter officers to organize workshops, seminars, and other events weekly to enhance member engagement
- Managed event logistics during the VIVA-VVIT 2021 Annual Fest, ensuring smooth operations and successful outcomes.