

# KK Thuwajit

608-440-4120 | [thuwaajt@wisc.edu](mailto:thuwaajt@wisc.edu) | [linkedin.com/in/kkuroma](https://www.linkedin.com/in/kkuroma) | [github.com/kkuroma](https://github.com/kkuroma) | [personal website](#)

## EDUCATION

### University of Wisconsin Madison

Madison, WI

*Bachelor of Science in Computer Science and Mathematics, GPA 4.00 Dean's List*

*Aug 2022 – Dec 2025*

- Selected courses: Graduate machine learning, Deep learning in computer vision, Big data system, Stochastic processes

## EXPERIENCE

### Machine Learning Engineer Intern

May 2024 – Aug 2024

*American Family Insurance*

*Madison, WI*

- Developed a **custom invoice extractor for bulk** achieving over 90% on F1-score
- Proposed a training strategy for a *LayoutLM* document extractor using *PyTorch* and *Huggingface* to support invoice question and answering via natural language, outperforming benchmarks
- Integrated *DocumentAI* and *Gemini 1.5* with a custom extractor for irregular-patterned documents
- Built a **user intent classification** model, enhancing internal searches with a starting accuracy of 88%
- Designed a query classification strategy using *word frequency* and *n-grams* from *BigQuery* statistics to develop a *graphical Bayesian network* classifier with an improving accuracy over time

### Research Assistant

May 2023 – Present

*Computer Sciences, University of Wisconsin Madison*

*Madison, WI*

- Developed *textual exclusion*, a **personalized text-to-image generation technique** via *denoising diffusion models* to address generating out-of-distribution images, outperforming Google's DreamBooth on prompt-adherence

### Research Assistant

Oct 2022 – Present

*Waisman Center, University of Wisconsin Madison*

*Madison, WI*

- Developed an **MRI denoising algorithm**, accelerating scan time from 9 to 2 minutes
- Proposed a novel *dual U-Net regularizer* paradigm to enforce quantitative T1 value consistency
- Utilized *low-rank reconstruction* and *distributed GPU training* to address computational limitations

### Machine Learning Engineer Intern

Oct 2021 – Feb 2022

*NXPO (Higher Education Science Research Policy Council)*

*Bangkok, Thailand*

- Acquired and analyzed 1M+ publications to **identify emerging scientific fields** for research funding
- Developed a technique combining *word2vec* and *contrastive learning* to convert publications into vectors for analysis
- Derived a *generalized logistic regression* expression for emerging number of publications over time

### Research Assistant

May 2020 – May 2023

*Information Science and Technology, VISTEC*

*Rayong, Thailand*

- Developed a novel *multi-scale CNN seizure detector* for EEG signals with 95% accuracy
- Adapted the technique for a **respiratory rate estimator** for real-time monitoring via wearable devices
- Published** these findings as first and corresponding authors on several IEEE journals

## PROJECTS

### Personal Website | *Next.js*, *TypeScript*, *TailwindCSS*

2023

- Created an informative website with interactive elements using the *Next.js* framework

### Unofficial Implementation of UNIT-DDPM | *Python*, *PyTorch*, *NumPy*, *Matplotlib*

2022

- Implemented the paper on **unpaired image-to-image** translation via *denoising diffusion models*
- Translated mathematical equations into **PyTorch** implementations and *Matplotlib* visualizations of the paper's results

### Anime Recommendation Bot | *Python*, *BeautifulSoup*, *TensorFlow*, *Discord.py*, *Flask*, *Ngrok*

2021

- Developed an *cosine similarity* based **anime recommendation system** using trained vector representations of the synopsis on MyAnimeList, using the anime genres and ratings similarity as objective for **contrastive learning**
- Deployed the system as a Discord bot using **Flask** and **Ngrok**, ensuring near-perfect uptime while used in 10+ servers
- Enabled **chat functionality** on the bot using generative text models trained on anime dialogues, predating ChatGPT.

## TECHNICAL SKILLS

**Languages:** Python, SQL, Java, JavaScript, TypeScript, Bash, HTML/CSS, LaTeX

**Data Science:** NumPy, SciPy, Pandas, Matplotlib, Plotly, Seaborn

**Machine Learning:** PyTorch, TensorFlow, Scikit-Learn, HuggingFace, DocumentAI, LangChain

**Databases:** PySpark, Cassandra, Hadoop Ecosystem, Kafka, MySQL, BigQuery, VertexAI

**Developer Tools:** Git, Docker, Google Cloud Platform, VS Code, Tmux, Nohup