Mahendra Singh Thapa

Rochester NY mt8633@g.rit.edu (+1)-585-957-6971 mahenthapa

Research Interests

Advancing Artificial Intelligence through Bayesian principles.

Bayesian Inference, Hierarchical Bayesian Framework, Federated Learning, Deep Learning

Education

Rochester Institute of Technology (RIT) Rochester, NY
Ph.D. in Computing and Information Sciences *Aug 2023 – Present*

- **Relevant Courses:** Statistical Machine Learning, Deep Learning, Non-Convex Optimization for Modern Machine Learning

Institute of Engineering, Tribhuvan University Lalitpur, Nepal
Bachelor's Degree in Computer Engineering from Pulchowk Campus *Nov 2012 – Sep 2016*

- **Relevant Courses:** Probability and Statistics, Artificial intelligence

Publications

Mahendra Singh Thapa, Rui Li (2025). *Harnessing Heterogeneous Statistical Strength for Personalized Federated Learning via Hierarchical Bayesian Inference.* In Forty-second International Conference on Machine Learning (**ICML 2025**).

Ayush Shakya, Bijay Gurung, **Mahendra Singh Thapa**, Mehang Rai, Basanta Joshi (2017). *Music classification based on genre and mood.* In First International Conference on Computational Intelligence, Communications, and Business Analytics (**CICBA 2017**).

Patents

Mahendra Singh Thapa, Miran Ghimire, Rakesh Kumar Katuwal, Suresh Manandhar, and Sameer Maskey. *System for online course creation.* U.S. Patent Application 18/599,106, filed January 30, 2025.

Mahendra Singh Thapa, Miran Ghimire, Suresh Manandhar, Kiran Prajapati, Roshani Acharya, Purushottam Shrestha, and Aashish Pokharel. *Method of extracting information from an image of a document.* U.S. Patent Application 18/732,252, filed July 10, 2025.

Shushanta Pudasaini, Sunil Ghimire, Bijaya Khadka, **Mahendra Singh Thapa**, Giovanni Patrick Lemus, Rakesh Kumar Katuwal, and Sameer Maskey. *Method of automating collection and screening of resumes.* U.S. Patent Application 18/732,261, filed July 10, 2025.

Reseach Experience

Graduate Research Assistant *Aug 2023 - Present*
Rochester Institute of Technology (RIT) Rochester, NY

- Developed a Hierarchical Bayesian framework to model the relationship between the personalized and global models in the Federated Learning setup. Part of this work was published in ICML 2025.

Professional Experience

Founder & CTO *April 2023 - Dec 2024*
Skillprompt Nepal

- Led **cross-functional teams** (Engineers, Designers, QA, DevOps) to establish **internal processes**, develop **IT training curricula**, and deliver **client-centric technology solutions** that drove successful project outcomes.

Assistant Manager AI Research *Jan 2021 - March 2023*
Fusemachines Nepal

- Designed and implemented **transformer and reinforcement learning based solutions for geometric unwarping and document alignment** of high-resolution KYC images by **detecting keypoints, predicting displacement flow, and estimating homography matrix** to improve document processing accuracy.
- Extracted **course syllabus data from documents in a low-data regime** by combining visual and textual cues, applying **Infinite Mixture Models** for effective feature-based soft clustering.

AI Instructor *Sep 2019 - March 2023*
Fusemachines Nepal

- **Delivered Machine Learning, Deep Learning, and Computer Vision training programs** to over **70 fellows** across cohorts in Nepal and Bangladesh, covering topics such as supervised/unsupervised learning, CNNs, RNNs, Transformers, and reinforcement learning.
- **Achieved an average instructor rating of 4.65/5**; led **research paper discussion** and hosted a **Kaggle competition** to strengthen hands-on skills and technical insight.

Machine Learning Engineer *Oct 2016 - Dec 2020*
Fusemachines

- Developed and implemented a **reinforcement learning solution using the TD3 algorithm** to control engine torque and brake command in a simulated truck environment with continuous state and discrete action spaces. Led **experimentation and optimization of control policies** to improve vehicle performance.
- Built a **Named Entity Recognition (NER) system leveraging a BERT-based model**, driving improvements through **data analysis, visualization, and hyperparameter tuning** to enhance model accuracy and robustness.
- Developed a **CNN-based system for bone localization** on X-ray images by predicting bounding boxes. Responsible for **model architecture, training, and evaluation** to enhance detection accuracy.
- Built a **custom neural network-based demand forecasting model** for a food. Conducted extensive experimentation with standard and tailored approaches, integrating domain-specific insights through data analysis and visualization. **Achieved a 10% reduction in forecasting error** compared to expert-generated predictions.
- Built an **AI-powered sales assistant** that identifies and ranks follow-up emails for users. Led end-to-end development, including **data preprocessing, model training, and evaluation**, to streamline sales workflows.
- Developed an **AI assistant for account-based lead generation** by designing a ranking model to prioritize leads based on Ideal Customer Profile (ICP) criteria. Focused on **model design, feature engineering, and ranking optimization** to improve lead targeting accuracy.

Awards and Honors

RIT Ph.D. Scholarship/Assistantship. Financial Support at RIT to pursue a Ph.D. in Computing and Information Sciences via NSF and NIH Grants.

Ncell Scholarships and Excellence Awards. Awarded **twice** by Ncell in recognition of outstanding academic performance.

Full Scholarship at Tribhuvan University. Awarded by Nepal Government for achieving 153rd rank out of 13000 students in entrance examination (4% acceptance rate).

Skills

Programming Languages: Python (advanced), Java, C, C++

Deep Learning: PyTorch (advanced), TensorFlow, Keras

Machine Learning: scikit-learn, NumPy, Pandas, Matplotlib, Seaborn, Neptune, Tensorboard

Miscellaneous LaTeX, Git, Bash, Jira, Linux, AWS