

Rajashree Dahal

256-910-9092 | rajashreedahal4@gmail.com | [linkedin.com/in/rajashreedahal](https://www.linkedin.com/in/rajashreedahal) | github.com/RajashreeDahal4

Motivated hard working computer science graduate with more than 2 years of professional experience in python, data science, machine learning and Natural Language Processing with backend skills

EDUCATION

The University of Alabama in Huntsville (Advisor: Dr. Tatagatha Mukherjee) Huntsville, AL
Master of Science in Computer Science (Thesis), GPA 4.0 Aug. 2022 – May 2024
Thesis: Hierarchical MultiLabel Text Classification in Earth Science Datasets.

IOE Pulchowk Campus, Tribhuvan University Lalitpur, Nepal
Bachelor's Degree in Mechanical Engineering , 76.63% Nov. 2016 – Sept. 2021
Final Year Project: Computational Analysis of Mechanical Compatibility of GFR-PEEK and PU-Glass Fiber Bio-composites for Femur Implant.

EXPERIENCE

West Virginia University Morgantown, WV
Research Volunteer July. 2024 – Present

Skills: Python, Large Language Models, Openai, Langchain, Prompt Engineering, Thematic Analysis

- Implementing Large Language Models to explore its potential application as a Virtual Teaching Assistant. (NAACL 2025- Target Conference for paper submission)
- Using Large Language Models to explore facilitators and barriers to patient-provider communication regarding diabetes self-management using Thematic Analysis (Pedagogy in Health Promotion-Target journal for paper submission)

NASA IMPACT Huntsville, AL
Graduate Research Assistant–enhancing search capacity leveraging NLP Aug. 2022 – May 2024

Skills: Python, Transformers, SQL, Classification, NLP, Prompt Engineering, FineTuning, PyTorch, Langchain, LLMs, Quantization, numpy, pandas, matplotlib, FastAPI, scikit-learn, wandb, S3, SageMaker, Data Analysis, Git, API Gateway, Lambda, Docker, ECR, Airflow, Unit testing, Django

- **Acronym Disambiguation:** Achieved 98% recall and 97% precision in acronym disambiguation by implementing a BERT-based sequence classifier within the Science Discovery Engine (SDE) for precise identification in earth science datasets.
- **Automated Document Tagging:** Optimized XLNet model for SDE text classification, reducing inference time by 79%. Experimented on architecture of **inference system** using SageMaker, webapp ecosystem, and implemented self-managed Airflow in a cost-effective manner.
- **Chain of Prompt:** Developed a data-driven solution utilizing Chat GPT API for structured queries, CMR search for temporal and spatial data retrieval using LangChain conversational buffer window and prompt engineering.
- **Hierarchical Text Classification (Ongoing paper):** Introduced zeroshot approach for hierarchical classification of earth science datasets, enhancing recall@10 to 0.61 from 0.57. Deployed inference system via API endpoint.
- **Finetune Metadata Extraction:** Developed a metadata extraction pipeline using Larch and fine-tuned a gpt-3.5-turbo-1106 model for enhanced performance on Environmental Justice documents.
- **Data In Action (DIA):** Enhanced DIA web app stability by resolving 20+ software bugs through detailed code analysis, achieving reduction in critical issues.

ExtensoData Pvt Ltd Lalitpur, Nepal
Data Quality Assurance Engineer Oct. 2021 – Jun. 2022

Skills: Python, SQL, A/B testing, Airflow, Robot Framework

- Led end-to-end implementation of key projects, including Credit Scoring, Foneloan Decision Intelligence System.
- Prepared UAT Test case and Test plan according to business use case.
- Communicate accepted quantifiable metrics inefficiencies and assist in new metrics formulation.
- Conducted A/B testing, resulting in a 15% improvement in data accuracy for the Credit Scoring project.

PERSONAL PROJECTS AND COMPETITION

Skills: Python, Tensorflow, Image Segmentation, Logistic Regression, Ensemble Learning, ANN, Random Forest, Explainable AI, Data Analysis, Data Visualization, Feature Engineering, Git

Segmentation of HUBMAP Kidney Data

Jan. 2023 – Apr. 2023

- Implemented double U-net semantic segmentation model to identify glomeruli in human kidney tissue images on a Kaggle based dataset and experimented with PCA across datasets of different batch size.

Presumptive Diagnosis of Urinary System Diseases

Feb. 2023 – Apr. 2023

- Experimented with different machine learning algorithms which will perform the presumptive diagnosis of acute Nephritis of renal pelvis origin and acute Inflammation of urinary bladder showing 100% accuracy.

Feel the Rhythm by Unearthed

Jul. 2021 – Aug. 2021

- Ranked in the top ten position out of 254 participants.
- Implemented an explainable boosting classifier for workplace safety prediction in an imbalanced dataset, achieving a private dataset ROC-AUC score of 0.654 and conducting an interpretability assessment.

Looks Like Grains by Unearthed

Jun. 2021 – Jul. 2021

- Ranked in the top eight position out of 269 participants.
- Used pretrained EfficientNetB2 model to classify defects within oats, wheat and barley grains.

EXTRAS

Wintercup Intra College Basketball Tournament - Organizer

2020

- Planned, scheduled and managed basketball matches.
- Provided training and increased female participation by 25% by team and 29% by number than the previous year.

FabLab Humanitarian Design Challenge by DASSAULT SYSTEMES - Participant

Feb. 2020

- Iterated gilfo writing tool for Nepali market and final prototype was designed incorporating data findings from design thinking interviews. The product can help children with cerebral palsy to use basic tools like crayons and pencil with ease.

Society of Mechanical Engineering Students (SOMES) - Aero-Coordinator

2018-2019

HONORS AND AWARDS

- 2022-2024: **Graduate Research Assistant** (fully funded for 22 months long MS program).
- 2016-2021: **Institute of Engineering** Undergraduate Fellowship (fully funded for 4 years B.E. program).
- 2021: Nepal Engineer's Association (NEA) Bagmati Innovation Support Award-2021.
- 2019: **Gold Medal Winner** in oral presentation in National Young Scientist Conference (NYSC-2019).
- 2016: **Institute of Engineering Entrance Girls Topper.** (Mechanical Department)

SPORTS

- 2020: **Winner** in 'Super Six Cricket Tournament 2020' organized by Nepal Terai Bidhyarthi Navajagaran Sangh.
- 2020: **Second Position** in 'Wintercup Intra College Basketball Tournament 2020' organized by Pulchowk-A team.
- 2019: **Women of the Match (Final)** in 'Super Six Cricket Tournament 2019 organized by NTBNS'.
- 2018: **Second Position** in 'First Himalayan Intertechnical Basketball Championship 2018'.
- 2017: **First Position** in 'Third National Intertechnical Basketball Tournament 2017'.

REFERENCES

- **Assoc. Prof. Dr. Tatagatha Mukherjee**, Department of Computer Sciene, The University of Alabama in Huntsville
- **Assoc. Prof. Dr. Vineetha Menon**, Department of Computer Science, The University of Alabama in Huntsville

Personal Statement

Rajashree Dahal (rajashreedahal4@gmail.com)

Having been born in a traditional Nepali family in a three daughter family, I had to face a lot of social stigmas and I was constantly being made to remember since childhood that I am a girl and whatever savings was left of my parents would eventually go to my cousins given that they are boys and they represent our family lineage and the daughters will go to somebody else's home. My parents were determined to provide me with opportunities they never had, however, all they could provide was basic food and not being able to afford my education. Due to my excellent academic background, I was fully funded during my school education which would have resulted in a full stop in my undergraduate study had I not aced among the top students in the entrance examination which would be funded by the Government of Nepal as my parents were more fixated to invest in my marriage than in my education. The challenges intensified when my father resigned from his job due to serious health issues, right as I began my undergraduate studies. This was supposed to be a pivotal time for my academic foundation, but it brought significant financial strain.

I woke up every morning with my grandmother mentioning that the home environment has been a waste with granddaughters. I encountered stigma during my undergraduate studies when a fellow lab partner remarked that "a boy with a green card would make my life easier". This was the first harsh comment I faced outside of my family network, highlighting the persistent gender biases I continue to navigate. Despite these challenges, I turned out to be the only one among girls and boys from my school who happened to pursue higher education in the U.S. Yet, my academic as well as professional journey has been marked by challenges—whether it was being dismissed by authority figures who underestimated my voice or witnessing funds collected for my marriage instead of education. A pivotal moment in my academic journey came when a senior from the Computer Science Department demonstrated how she could predict reliable numbers from limited data to aid real-world decision-making. That same year, during my internship at CG Foods in Nepal, I saw firsthand how advanced techniques like risk-based assessments, optimization algorithms, and machine learning had surpassed the traditional methods still used in my country. These experiences made me realize I could make a greater impact using computer science, steering me toward the field for my graduate studies.

Personal Statement

Rajashree Dahal (rajashreedahal4@gmail.com)

During the COVID-19 lockdown, I took online courses in machine learning and computer vision, which led to a project detecting functional tissue units across different pipelines. I explored model architectures like U-net and double U-net and loss functions such as dice loss and Tversky loss. This project introduced me to AI applications in healthcare, including diagnosis, patient monitoring, and disease prevention. My goal is to pursue a PhD in Computer Science and, long-term, contribute to my country's health sector by improving human-computer interaction for motor disorders and enhancing medical dataset interoperability.

I earned my Master of Computer Science from the University of Alabama in Huntsville, specializing in subjects like Statistics, Big Data, AI, Database Systems, and Algorithms. As a Graduate Research Assistant at NASA-IMPACT, I worked on enhancing search capabilities for Earth science datasets. My projects ranged from novel approaches that became my thesis to products developed for the Science Discovery Engine (SDE) team, some with potential for open-source release. Over 21 months, I gained hands-on experience in Natural Language Processing, Langchain, Large Language Models, prompt engineering, text-based similarity search, and information retrieval. These projects honed my research and problem-solving skills in NLP applications. Following graduation, I joined West Virginia University as a research volunteer under Dr. Prashnna Gyawali, focusing on the application of LLMs as virtual teaching assistants and their use in qualitative analysis in health. These experiences have solidified my commitment to pursuing a PhD in Computer Science, with a research focus on NLP and machine learning.

In my quest for a program match, I was thrilled to learn that Columbia University has a vast faculty in multiple research areas as I plan to pursue a dissertation in the combination of multiple subfields of AI. At the Department of Computer Science, Dr. Noémie Elhadad's innovative work with CEHR-GPT, which utilizes Generative Pre-trained Transformers (GPT) to generate synthetic Electronic Health Records (EHR) while preserving crucial temporal dependencies, aligns with my research goals of improving machine learning models in clinical settings. Her focus on capturing patient visit sequences, inpatient durations, and time intervals ensures a more realistic representation of patient histories, addressing key limitations in existing EHR generation methods. I am particularly inspired by her integration of synthetic

Personal Statement

Rajashree Dahal (rajashreedahal4@gmail.com)

data into the OMOP Common Data Model, which enhances the accessibility and utility of such data in broader healthcare applications. Dr. Elhadad's commitment to privacy-preserving techniques also resonates with my desire to balance technological advancements with ethical data use in healthcare research. Having known her and read her publications for quite a while, working alongside her during my graduate study would make my experiences at Columbia University a challenging and fulfilling one.

Similarly, Dr. Smaranda Muresan's paper entitled 'Is ChatGPT a Better Explainer than My Professor?': Evaluating the Explanation Capabilities of LLMs in Conversation Compared to a Human Baseline', which highlights the ability of LLMs to generate effective explanations in conversational settings by comparing them to human explainers. I can see significant potential for further exploration in this area, which aligns closely with my current work I am doing as a research volunteer. I am exploring more on different roles of bot to answer the query based on the intent classification framework with an objective of introducing personalized virtual teaching assistants.

Likewise, Dr. Zhou Yu's research on conversational AI, particularly her work on constrained text generation and dialogue systems, deeply resonates with my interests in human-computer interaction and effective communication. Her focus on creating models that adapt to real-world dialogue constraints aligns with my goals of developing AI-driven systems that enhance user interaction and accessibility. I am excited about the opportunity to conduct research under the guidance of esteemed professors and be part of a university that embraces diverse research backgrounds. I look forward to contributing to the academic community through my research and sharing my culture along the way.

Being the only female in my undergraduate batch of my institution and the first in my family to pursue a Ph.D., I face constant pressure from family and society to focus on marriage and family life. This has made my journey challenging in ways that are difficult to explain to others, even among my close friends. Yet, these struggles have only strengthened my resolve. I come from a remote area of Nepal where few people, especially women, are expected to reach for the kinds of goals I am pursuing. I am currently working as an unpaid research volunteer which I find is very interesting and aligns closely to what kind of

Personal Statement

Rajashree Dahal (rajashreedahal4@gmail.com)

research I want to pursue for my PhD. However, the financial strain of not yet being self-sufficient weighs on me daily, especially as I rely on my cousin for accommodation in the US, which adds to both my concerns and theirs. This program is my final chance, but I firmly believe in my ability to push past these boundaries and prove that I can contribute meaningfully to the world. For me, this program represents much more than academic advancement—it's a way to show that even from my background, it is possible to achieve things that many believe are out of reach. I believe this higher education is not just a personal goal, but a powerful tool for revolutionizing lives and communities.

I am confident that the research and services that I do at Columbia University are going to touch the lives of many as I am planning to explore more on the field which has the potential to influence globally. I also want to claim that my background has not only qualified me technically but has also given me the right mindset for such a career prospect. With a sound academic background and a professional career in the Computer Science field behind me, I believe, I can stake a claim of the upper hand as far as my choices of subjects are concerned. From facing the societal stigma of being born in a three-daughter family, to walking barefoot for hours to reach school, and studying by candlelight during frequent power outages, I take great pride in how far I've come and remain resolute in my ambitions for the future. It will be difficult for me to continue my graduate studies without financial aid. Therefore, if you could endow me with admission to the course I am counting on an opportunity to get financial assistance in the form of RA/TA/GA based on my merit, I will feel a step closer to the goal of conducting research worthy of academic and experimental significance. Thank you very much for reviewing my application.