Biraj Bikram Pathak

Huntsville, AL | 256-468-8537 | birajbikrampathak@gmail.com | LinkedIn | Github | Portfolio

EDUCATION

University of Alabama in Huntsville — Aug 2024 - May 2026
Master of Science - Computer Science — *Huntsville, Alabama*
- **Relevant Coursework:** Data Structures and Algorithms, Machine Learning, Database System, Big Data & Cloud Computing

Tribhuvan University, IOE, Pulchowk Campus — Jan 2018 – May 2023
Bachelor of Engineering in Computer Engineering — Graduated with Distinction: 80.12% — *Kathmandu, Nepal*
- **Relevant Coursework:** Data Structures and Algorithms, Object-Oriented Programming, Distributed Systems, Discrete Mathematics, Probability and Statistics

SKILLS & CERTIFICATIONS

Programming: Python, SQL, JavaScript, HTML/CSS
Technologies: Django, Pandas, Scikit-Learn, PostgreSQL, Docker, Node.js, NumPy, Apache Airflow, Snowflake, AWS, Git
Certifications: CodePath Advanced Technical Interview Prep (June 2025 - August 2025), AWS Academy Data Engineering, Introduction to Data Engineering (IBM, Coursera)

EXPERIENCE

NASA-IMPACT, University of Alabama in Huntsville — Aug 2024 - Present
Graduate Research Assistant — *Huntsville, AL*
- Engineered ingestion pipelines and migrated around **4TB** of vendor data from **2** sources using **Airflow** and **AWS**.
- Accelerated data request workflows for CSDAP users by fabricating dynamic, user-centric forms with **Django**.
- Tested and modularized reusable unit tests for ingestion and forms using **Pytest**, boosting development time by **10%**.

LIS Nepal Pvt. Ltd — Feb 2024 - Jul 2024
Associate Software Engineer — *Lalitpur, Nepal*
- Optimized and deployed **ETL pipelines** with **Azure Data Factory** and **Snowflake** to reinforce data workflows.
- Prototyped a **proof of concept (PoC)** for Robling-based ETL solutions and collaborated within an **Agile Scrum** environment to remodel data warehousing pipelines.

extensoData, F1soft — Jul 2023 - Jan 2024
Data Trainee — *Lalitpur, Nepal*
- Modeled credit eligibility and assessed default risk for the *FoneLoan* system by analyzing customer behavior and mining historical **pre- and post-loan** data using **SQL**, **Python**, and **Pentaho**.
- Automated holiday-based pipeline scheduling by scraping national calendars with **Selenium** and **BeautifulSoup**, eradicating manual updates and enhancing operational efficiency by **20%**.

PROJECTS

SlideIt – AI Video Presentation Generator — *Python, BART, PyTorch, Flask, React*
- Engineered an AI-powered pipeline to generate video slide presentations from text, transforming lengthy articles into structured summaries with **ROUGE-L: 41.59** using BART-large on the CNN/Daily Mail dataset.
- Implemented K-Medoids clustering for sentence grouping and KNN for key sentence ranking; synthesized voice using Azure Cognitive Speech Services and assembled videos in under 3 minutes with **FFMPEG**.

Bank Marketing Dataset Classification — *Python, Scikit-learn*
- Modeled a classification pipeline using Random Forest to predict term deposit subscriptions with **89% accuracy**, and optimized it via SMOTE, PCA, and hyperparameter tuning.
- Assessed and compared classifier performance using cross-validation to prioritize optimal feature sets and algorithms.

EchoCraft - Custom Shell — GitHub
- Constructed a custom shell in **Python** with pipes, redirection, and history; built a lexical analyzer and parser from scratch, applying **OOP** concepts like abstraction and inheritance to modularize and enable future enhancements.

AWARDS AND EXTRACURRICULARS

- Outreach & Sponsorship Lead, HULT Prize at IOE   (December 2022 - February 2023)
- Coordinator in Football & Futsal Tournament, LOCUS 2021   (December 2020 - January 2021)