

Nikhil Yeminedi

Atlanta, GA — (404) 203-8442 — [✉ nikhilgoudyeminedi@gmail.com](mailto:nikhilgoudyeminedi@gmail.com) — [in](#) LinkedIn — [G](#) GitHub

Education

Masters in Computer Science

Georgia State University, Atlanta, Georgia, USA

GPA: 4.0/4.0

January 2025 - May 2027

Bachelor of Technology in Computer Science

Jawaharlal Nehru Technological University

GPA: 8.3/10

July 2018 — June 2022

Work Experience

Software Engineer - SCALES — Atlanta, GA — *May 2025 - Present* — [Link](#)

- Processed and disambiguated millions of US Federal Court (PACER) records, building entity resolution pipelines to generate unique IDs for judges, law firms, and legal parties. Presented a poster on designing and implementing a scalable data pipeline.
- Extracted key entities from federal docket text using spaCy NER models to support analysis of judicial patterns.
- Provided foundational data infrastructure that transformed raw federal case data into structured, analyzable datasets for legal and policy research, and presented outcomes through technical presentations and poster sessions to stakeholders.

Research Assistant – DICE Lab — Georgia State University — *January 2025 - May 2025*

- Co-authored and published research (NLPIR 2025) on a neuro-symbolic adverse event detection framework, where I developed and fine-tuned a BERT-based classifier with domain-specific preprocessing and semantic filtering.
- Build and implemented symbolic validation layer, engineering co-occurrence extraction and graph-based analysis (LLR, PMI, DSC) to construct interpretable knowledge graphs and AI-driven disaster informatics systems.

Data Engineer & Analyst - Tata Consultancy Services — *July 2022 - December 2024 (2.5 Years)*

Projects: Global Federated Data Mart (GFDM) & Tusker – Client: Visa

- Created and automated ETL pipelines by analyzing SQL scripts, preprocessing and optimized HQL workflows on Hadoop ecosystems, orchestrated via Jenkins CI/CD and Tuber DAG, while leading a team in workflow optimization.
- Built and optimized big data solutions using Hive, Sqoop, Oozie, Scala, Java, and MapReduce, while tuning Spark drivers and YARN logs for distributed performance and reliability.
- Utilized Postman, Jetty, and Tomcat for API testing, service validation, and request handling, ensuring seamless interaction between big data services and downstream applications.
- Streamlined data migration, wrangling, and transformation processes on Unix/Linux environments, leveraging cloud storage platforms (AWS, Azure, Cloudera) to enhance scalability, operational efficiency, and real-time analytics.

Web Developer - Steed 26 Solutions Pvt Ltd, Bengaluru, India — *March 2021 - June 2022 (1.3 Years)* Hospital Records Management System (HRMS)

- Designed and performed responsive, dynamic user interfaces operating HTML, CSS3, and JavaScript.
- Maintained code quality and deployment readiness through collaborative Agile sprints, including reviews and stand-ups.

Projects

Eviction Data Pipeline: Automated Court Case Processing & Demographic Analysis

Built an automated pipeline processing 200,000+ eviction cases, implementing fuzzy matching algorithms and Census API .

Adverse-Event Detection and Co-Occurrence Analysis for Disaster Tweets

Developed a scalable NLP Model to analyze co-occurrence patterns across a corpus of 50,000+ disaster-related tweets.

Hand gesture recognition System for Sign Language Using Deep Learning [Link](#)

Initiated an end-to-end deep learning application using CNNs, OpenCV, and Flask to translate sign language into text.

Finland Labs – Student Intern — Machine Learning, AI, IoT, AWS

Trained on ML, AI, and IoT using AWS; accomplished Cat & Dog Classification with CNNs, Home Automation on AWS.

Technical Skills

- **Programming & Databases:** Python, Scala, Java, SQL, HQL, Shell Scripting, JavaScript, MySQL, Oracle, RDBMS
- **Big Data, Cloud & DevOps:** Apache (Spark, Airflow DAGs), Hadoop Ecosystem, ETL/ELT Pipelines, AWS (S3, EC2), Microsoft Azure, Cloudera, Git, Docker, Jenkins, Jira, Postman, Tomcat, Jetty, IntelliJ, Jupyter Notebook, Agile (Scrum)
- **Machine Learning & Analytics:** Deep Learning (CNNs, RNNs), NLP (NER, BERT, spaCy), Power BI, Matplotlib, Advanced Excel. **Certifications:** AWS Cloud Practitioner, Moving Data into Hadoop (IBM), Python and Flask (Udemy).