IRITH CHATURVEDI

irith2004@gmail.com | (447)-902-2453 | Champaign, IL - 61820, USA | linkedin.com/in/irithchaturvedi

EDUCATION

University of Illinois Urbana Champaign *(August 2022 — May 2026)*

Bachelor of Science in Information Science + Data Science, Minor in Computer Science and Statistics.
GPA — 4.0, James Honor's Scholar (2022, 2023, 2024, and 2025), Dean's List (2022, 2023, 2024, and 2025)

INFORMATION TECHNOLOGY SKILLS:

- ✓ Microsoft Office Suite
- ✓ Python (Pandas, NumPy, Seaborn, MATLAB, and Matplotlib)
- ✓ Django + Tkinter & HTML/CSS
- ✓ PyTorch, Sci-Kit Learn, PySpark.ml
- ✓ C++ and Java

- ✓ PySpark, Hadoop, Apache Hive and Iceberg
- ✓ SQL Server, NoSQL (Cosmos DB)
- ✓ R (Statistical Analysis, ML, and Interactive Dashboards)
- ✓ PyTest and Robot Framework (Test automation and scripting)

- ✓ REST API development, Amazon Web Services (AWS), Microsoft Azure
- ✓ Zena (Job), IDQE (Test Case), and Snakemake (Workflow automation)
- ✓ Tableau, Streamlit, Altair, Power BI
- ✓ Docker Containerization and VMs

WORK EXPERIENCE

Data Science Intern, COUNTRY Financial DigitaLab *(Champaign, IL: May 2025 - Present)*

➤ Collaborated in bi-weekly sprints as part of the Sales Compensation Agile Release Train, supporting data engineering efforts by writing PySpark and HQL scripts for large-scale data transformation and workflow automation via Zena. Integrated seamlessly into a 16-member team of senior engineers, tackling user defect stories and raising Change Requests/Orders through SDM for deployment.

➤ Enhanced CI/CD pipelines by developing automated test cases in SQL using IDQE to ensure data quality. Automated data migration from the Cloudera & Azure Data Lake to the Non-prod environment, maintaining consistency via Hive partition repair and validation procedures.

➤ Developed and deployed an end-to-end Test Case automation pipeline for Apache Iceberg, integrating GitLab CI/CD, Rally, and IDQE APIs. Engineered a modular data ingestion framework using Python and Pandas for dynamic data wrangling and transformation. This eliminated manual testing for database and table validation, resulting in a 90% reduction in QA effort and saving over 140+ man-hours per month.

Undergraduate Researcher, Biodiversity Genomics Lab, University of Illinois *(Champaign, IL: August 2023 - Present)*

➤ Workflow Automation and Computational Optimization using Nextflow, streamlining data modifications and renaming processes. Optimized phylogenetic analysis using Python-based parallelization and multi-threading, significantly reducing processing time by 60%.

➤ Data Analysis and Redundancy Resolution: Identified over 19,000 unique loci through advanced feature filtering and statistical analysis with R, enhancing the quality of evolutionary insights.

➤ Addressed reference paralog redundancy by implementing a novel weighted scoring system, improving alignment accuracy by 94%.

Course Assistant, (STAT 420) Department of Statistics, University of Illinois *(Champaign, IL: January 2025 – May 2025)*

➤ Evaluated and graded weekly assignments and Midterm exams for over 100 students. Conducted weekly office hours to guide students through the material taught in the course and homework problems.

➤ Collaborated with the instructor, TA, and fellow CAs in weekly meetings to discuss grading strategies and course structure.

Intern, 21st Century Dads Foundation *(Chicago, IL: June - August 2023 and May - August 2024)*

➤ Used Python Data Science libraries to perform Grant Research with the help of Foundation Directory Online and Trello to identify and analyze 15 potential grantmaking foundations. Raised over $120,000 for marketing and advertising from a single grantmaker.

➤ Developing and designing the 21st Century Dads Mobile App with the help of Subsplash, commanding a user base of 1,000 families.

➤ Conducted database management processes using SQL Server and AWS S3 buckets to handle both structured and unstructured data.

EXTRACURRICULAR ACTIVITIES

Theta Xi National Fraternity (Alpha Beta Chapter), Executive Board - Head of Scholarship and Recruitment *(April 2023 - Present)*

➤ Played a key role in planning and organizing the Speaker Series featuring industry leaders, leveraging partnerships with RSOs and academic departments to maximize outreach, attendance, and recruitment. Developed and implemented a strategic three-step plan to enhance chapter academic performance, increasing the average semester GPA of active members from 2.8 to 3.3.

Notable Projects

➤ *GitMyWiki – AI-Powered Virtual Assistant for Internal Documentation Retrieval: Azure AI Search (Vector Search), Azure OpenAI, REST API, RAG Pipeline, Python, Streamlit:* Built an award-winning solution (**1st place out of 23 teams at the COUNTRY Financial Hackathon**) to help employees query buried internal documentation using NLP. Leveraged a Retrieval-Augmented Generation (RAG) pipeline with Azure AI Search and Azure OpenAI to deliver accurate, summarized answers with direct links to GitLab Wiki and SharePoint files.

➤ *Assessing Ideal Northeastern States in the U.S. to Live In:* Python (Pandas, Seaborn, SKLearn ML Pipelines), API Integration, OpenRefine
Built a machine learning pipeline to recommend affordable housing in the U.S. Northeast, using demographic and geographic data and the Zillow Real Estate API. Delivered interactive spatial visualizations and predictive models based on family size, budget, and school proximity.

➤ *RemindMe – Android Mobile App:* Python (Kivy Material Design for mobile app development), SQL (pyodbc)
Developed an Android app combining six key utilities: Notes, Drawing, Alarms, Events, Lists, and a customizable Home screen into a single reminder platform, enhancing everyday organization and productivity.

HOBBIES AND INTERESTS

Singing, Reading, Travelling (Have been to over 40 countries), Soccer, Dancing, Blogging, History, Volunteering (HFH)

GitHub Portofolio: **www.github.com/IrithChaturvedi**   Website & Blog: **https://irithchaturvedi.github.io**