Niharika Ganji

Bloomington, IN, USA | nganji@iu.edu | linkedin.com/in/niharikaganji/ | niharikaganji9.github.io/

EDUCATION

Master of Science, Data Science Aug 2022 – Dec 2023
Indiana University, Bloomington, IN GPA: 3.80/4.0

- ***Relevant Coursework*:** Data Mining, Advanced Database Concepts, Applied Machine Learning, Statistics
- ***Graduate Teaching Assistant – Fall 2023:*** Applied Algorithms & Creativity, Innovation, Technology

Bachelor of Technology Aug 2018 - May 2022
Indian Institute of Technology, Indore, MP GPA: 3.53/4.0

- ***Research Publications*:** ICDSA 2022 - CoviIS: A Real-Time Covid Help Information System using Digital Media
- ***Lead Positions:*** Aeromodelling Club, Debating Society, Avana IIT Indore

WORK EXPERIENCE

Data Analyst May 2021 – Aug 2021
Kaikili Inc. *Prosper, TX*

- Designed and implemented robust data pipelines to automate streaming data from SQL databases, resulting in a 60% reduction in data latency while ensuring efficient data access.
- Leveraged advanced statistical analysis techniques, including data wrangling, aggregation, correlation, and modeling, to identify patterns in customer service usage, driving a 10% increase in demand.
- Collaborated with stakeholders to develop visualization dashboards, using Tableau, to measure KPIs driving business intelligence, reducing manual analysis time by 70%.

Research Data Analyst Mar 2021 – May 2022
IIT Indore *Indore, MP*

- Engaged in research in data science and natural language processing, collaborating on over 6 projects in 1 year, with a focus on exploring novel techniques in linguistic analysis.
- Leveraged analytical skills to transform raw data into meaningful insights, using BigQuery facilitating informed decision-making processes for enhanced organization efficiency.
- Collaborated with multidisciplinary teams to collect, analyze, and interpret large datasets from different data sources, contributing to the development of predictive models with an accuracy of up to 96%.

PROJECTS

Covid Information System | *Pandas, Sklearn, NLTK, Sentiment Analysis* May 2021 - Mar 2022

- Consolidated real-time Covid information from approximately 9 websites and tools to develop a user-friendly Covid analytics platform, providing critical information during emergencies.
- Reduced response time for first responders by 40% through the efficient extraction of Covid emergency information from Twitter, facilitating timely emergency management.

Loan Default Prediction, Univ.AI | *Python, Tableau, Classification, Feature Engineering* Feb 2021 - Apr 2021

- Conducted statistical analysis on the demographic behavior of 1,00,000 customers to develop a highly accurate predictive model with a precision rate of 92%.
- Enhanced accuracy by 9% through the implementation of the optimal boosting algorithm, leveraging effective feature selection of the categorical features of customers.

Credit Card Fraud Detection | *Python, Sklearn, Anamoly Detection, Machine Learning* Apr 2021 - Jun 2021

- Achieved 95% accuracy rate in credit card fraud detection by deploying an end-to-end pipeline, utilizing machine learning techniques including logistic regression, random forest, and anomaly detection.
- Improved performance accuracy to 98.5% by optimizing the model through subsampling the negative class and employing feature engineering techniques.

TECHNICAL SKILLS

- **Programming Languages:** Python (Numpy, Pandas, Keras, Sklearn, Matplotlib), R, SQL
- **Databases:** MySQL, PostgreSQL
- **Tools:** Excel, Tableau, Power BI, BigQuery, Google Cloud, AWS (S3, E2, RDS, SageMaker), Hadoop
- **Data Science Methods:** Data Collection, Data Pre-processing, Exploratory Data Analysis, Data Visualization
- **Research Interests**: Machine Learning, Natural Language Processing
- **Certifications:** IBM Data Analyst Professional Certificate, IBM Data Scientist Professional Certificate