SAI KRUTHIK REDDY PADURU

Plainsboro, NJ – Open to relocation | +1 (732) 666-3372 | kruthikreddi@gmail.com | LinkedIn | GitHub | Portfolio

EDUCATION

The University of Texas at Dallas   **July 2022 – May 2024**
Master of Science, Business Analytics, Artificial Intelligence   *Richardson, TX*
Major: Data Science, Minor: Applied Machine Learning
Coursework: Applied Machine Learning, Deep Learning, Predictive & Prescriptive Analytics, Big Data, Web Analytics, Time Series Analysis, and SQL/NoSQL Databases. Focused on hands-on projects using Python, R, SQL, and modern tools for BI, data engineering, and advanced analytics

National Institute of Technology Puducherry   **July 2017 - May 2021**
Bachelor of Technology in Computer Science and Engineering   *Puducherry, India*

TECHNICAL SKILLS

Languages: Python, SQL, Java, C, C++, Bash, HTML, CSS, JavaScript, TypeScript, Ruby on Rails
Databases: MySQL, PostgreSQL, MongoDB, Cassandra, Snowflake, SQLite, Oracle
Cloud Technologies: AWS (Lambda, EC2, S3, RDS, CloudWatch), Azure DevOps, GCP, Terraform, Ansible, OpenShift
Frameworks: Django, Flask, FastAPI, Angular, React.js, Node.js, Docker, Kubernetes, PyTorch Lightning, torchrun, fast.ai, Django-CMS, Tailwind CSS
Libraries & Tools for Data & ML: NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras, PyTorch, SpaCy, NLTK**,** HuggingFace Transformers, Boto3, SQL Alchemy, BeautifulSoup, Selenium, Pickle, Plotly Dash
Visualization Tools: Tableau, Power BI, Looker Studio, Plotly Dash**,**Maestro
DevOps & CI/CD: Git, Jenkins, GitHub Actions, GitLab CI, SVN, Maven, Docker, Kubernetes, Airflow, Slurm Manager
Monitoring & Logging: BigPanda, ELK Stack (Elasticsearch, Logstash, Kibana)
Project & Task Management: Jira, Asana, Trello
IDEs & Dev Tools: PyCharm, Eclipse, VS Code, NetBeans, Anaconda, Sublime Text, Postman
Operating Systems: Linux (Ubuntu, RedHat), Windows, Unix
Machine Learning & AI: Regression, Classification, Clustering, CNN, RNN, NLP, Time Series Forecasting, GPT-4, Gemini AI, OpenAI APIs

WORK EXPERIENCE

ConnX AI   *November 2024 – Present*
*Data Analyst – **Technologies:** (Docker, Django, BigPanda, ELK Stack (Elasticsearch, Logstash, Kibana), Plotly Dash)*   *Plainsboro, NJ*
- Built interactive visualizations using **Plotly Dash** and integrated them into a **Django** web application using **django-plotly-dash**, enabling seamless access to data insights for stakeholders.
- Developed geospatial dashboards with **Plotly Leaflet** and **Pandas**, transforming large datasets into real-time, client-facing visual interfaces.
- Designed customizable dashboards with filtering, drill-downs, and user input, enhancing decision-making and engagement through rich interactivity.
- Utilized **BigPanda** for large-scale application monitoring, implementing alert logic and correlation rules to reduce alert noise and detect root causes efficiently.
- Automated incident workflows by integrating **BigPanda** with **ServiceNow**, streamlining ticket creation, prioritization, and SLA-based escalation.
- Conducted detailed **Root Cause Analysis (RCA)** using alert logs, historical metrics, and application traces, driving improvements in system reliability.
- Fetched real-time metrics from multiple monitoring systems via APIs, ingesting them into the **Elastic Stack** for centralized observability and performance tracking.
- Built and deployed **Kibana dashboards** for anomaly detection, log analysis, and SLA trend monitoring, helping reduce MTTR and improve system uptime.
- Collaborated with cross-functional teams to define alert thresholds, custom tags, and dashboard KPIs, ensuring monitoring strategies aligned with service goals.

Intelligenie Software Private Limited   *January 2024 – May 2024*
*Data Engineer– **Technologies:** (PyTorch DDP, Distributed DL ,Airflow, Spark, Snowflake, AWS, GCP, Jenkins, MongoDB)*   *Alpharetta, FL*
- Developed and maintained scalable data pipelines using Apache Airflow for orchestration and Spark for distributed data processing.
- Engineered robust ETL workflows for structured and semi-structured data using Hadoop, Kafka, and Snowflake to support analytics and reporting.
- Automated infrastructure provisioning and deployment pipelines using Terraform, Ansible, and Docker, improving CI/CD velocity and reproducibility.
- Managed containerized applications in production using Docker and Kubernetes, ensuring high availability and fault tolerance.
- Designed, optimized, and maintained relational and NoSQL databases including PostgreSQL, MySQL, MongoDB, Cassandra, and Oracle.
- Deployed and monitored data services and pipelines in the cloud using **AWS (Lambda, EC2, S3, RDS),** GCP, and Azure DevOps..
- Collaborated in version-controlled environments using Git, GitHub Actions, and Jenkins to support CI/CD and agile delivery cycles.
- Performed root cause analysis and system tuning using logs, metrics, and performance data gathered from Redis, Slurm Manager, and OpenMP tools.

MSys Technologies   *December 2021 – August 2022*
Software Engineer – Data Science ()   *Bangalore, India*
- Applied machine learning algorithms like **Logistic Regression**, **Random Forest**, and **XGBoost** to analyze datasets and build predictive models.
- Preprocessed and cleaned data using **Pandas** and **NumPy**, ensuring quality inputs for training and evaluation.
- Built data visualizations with **Matplotlib**, **Seaborn**, **Tableau**, and **Power BI** to explore trends and communicate results clearly.
- Worked with **PostgreSQL** to query and manage structured data for analysis and modeling tasks.
- Developed basic deep learning models using **TensorFlow** and **Keras** for image and text classification projects.
- Performed Natural Language Processing (NLP) tasks like sentiment analysis using **SpaCy**, **NLTK**, and **HuggingFace Transformers**.
- Deployed simple models using **AWS Lambda** and experimented with **SageMaker** for hands-on cloud deployment experience.
- Collaborated with team members using **Git** for version control and managed tasks using tools like **JIRA**.
- Participated in agile-style workflows, contributing to data projects from cleaning and modeling to visualization and reporting.

The 10x Academy **June 2021 – July 2022**

*Data Analyst Intern–(**Python, SQL, Power BI,Google Sheets, Pandas, NumPy, Matplotlib, MySQL, MongoDB, Excel**)* *Hyderabad, India*

- Conducted in-depth data analysis on university student performance using **Python**, **SQL**, and **Power BI**, resulting in a 50% reduction in resource workload.
- Analyzed performance data for 1,000+ applicants using **Python**, **SQL**, and **Power BI**, helping leadership prioritize resources and reduce manual workload by 50%.
- Designed an organizational chart using **Tableau** and **Looker Studio** to visualize reporting structure across 5 teams, streamlining onboarding and team alignment.
- Designed and implemented interactive dashboards using **Power BI**,**Google Sheets** to visualize performance trends and enable data-driven decisions.
- Performed web scraping, predictive analytics, and data modeling with **Pandas**, **NumPy**, **Matplotlib** to extract actionable insights.
- Managed data in **MySQL**,**MongoDB**; utilized **Excel** and **AWS** for data handling and dashboard deployment.
- Collaborated in an agile environment using **Git**, **JIRA**, and **Asana** to manage tasks, track analytics requests, and deliver insights efficiently.

PROJECTS

BeatBuddy: Constructed a multimodal search engine that accepts queries via text and audio inputs. Implemented **COLBERT** for text-based search and K-Nearest Neighbors (**KNN**) algorithm for song recommendations.

LEADERSHIP

- Worked as group leader for Innovation Incubation center (2021).
- Executive Supervisor and Led National Institute Startup Policy NISP(2020-21).
- Functioned as Executive Member in Swach Bharat Committee at NIT Puducherry (2017-2020).