

Prasanth Sagar Kottakota

San Jose, CA, 95192 | (917) 518 5271 | @website | @mail | @linkedin | @github | @kaggle



PROFESSIONAL SUMMARY

A results-oriented Software Engineer poised to graduate in May 2024 with an MS in CS. Demonstrates expertise in software development, data analysis, QA, and cloud technologies, with a strong focus on delivering impactful outcomes.

EDUCATION

Syracuse University, NY, USA **MS in Computer Science** | GPA : 3.93/4.00 *August 2022 - May 2024*
NITK Surathkal, KA, IN **B.Tech in Computer Science and Engineering** | GPA : 3.55/4 *August 2016 - May 2020*

AWARDS AND CERTIFICATIONS

AWS Solution Architect Associate | [link](#) | Amazon Web Services *March 2024*
Syracuse University Graduate Award (Fall 2023 and Spring 2024) | Syracuse University *August 2023*
Data Analytics Professional | [link](#) | Google *June 2022*

TECHNICAL SKILLS

Languages : Python | R | C | C++ | SQL | PHP | HTML | CSS | Javascript | Typescript | Haskell | Bash.
Frameworks/Lib : Scikit-Learn, Keras, PyTorch, Pandas, nltk, matplotlib, Apache Spark, Apache Kafka, Git, Django.
Data Tools : MySQL, AWS, S3, Athena, Redshift, Sagemaker, Glue, Snowflake, Apache Airflow, Databricks, K8s, BigQuery.
Data Visualization : Advanced Excel, Tableau, Looker, Seaborn, MS Office.
Others : Data Modeling, ETL, Big Data, ML, Automation, Debugging, RCA, Agile (Scrum/Kanban), Design Patterns.

EXPERIENCE

Graduate Teaching & Research Assistant | Syracuse University, Syracuse, NY *August 2023 - Present*

- Coordinating with Prof J.J Waclawski, Lectured a class of over 60 students, focusing on software implementations.
- Research involves collecting data from cattle herds using 20+ accelerometers. The data aim to detect patterns in cattle behavior related to health, productivity, herd direction, social cohesion, aiming to increase milk production.
- Engineered robust data pipelines leveraging **AWS infrastructure** and **Apache Spark**, employing **Lambda** architecture for real-time and batch processing of time-series data.

Software Engineer | McAfee, Bangalore, India *August 2020 - July 2022*

- Created and executed over 50 complex **SQL** queries involving **sub queries, stored procedures, triggers, views**.
- Designed and implemented 20+ **Airflow** pipelines tailored to team requirements, encompassing both inbound and outbound data pipelines, trigger rules, sensors, and pipeline orchestration.
- Improved dashboard performance by optimizing query pipelines and data processing for multiple attributes, using **Python, SQL, Apache Kafka, Apache Spark**, and visualization libraries to achieve a 20% load time reduction.
- Automated **ETL** processes, model training, & evaluation via scalable end-to-end data infrastructure pipelines in **Azure Databricks & Azure Data Factory**, leveraging **Spark clusters**, saving 120 labor hours monthly.

Machine Learning Researcher | NITK Surathkal, Mangalore, India *January 2019 - June 2019*

- Utilized CMU DSL-StrongPasswordData.csv for fraud detection, analyzing keystroke dynamics from 51 individuals.
- Explored various ML algorithms (**LSTM, Random Forest, Logistic Regression, KNN, SVM, etc.**) on keystroke dynamics data. **ARIMA** outperformed others with 69.8% accuracy in fraud detection.

Data Analyst | CampusHatt, Indian Institute of Technology Kanpur, India *May 2018 - July 2018*

- Executed **SQL queries** across **5+ data sources** to extract, clean, and analyze data, facilitating the identification of trends and patterns for actionable business insights.
- Used Excel for dynamic data manipulation, visualization, and dashboard creation, aiding decision-making.

PROJECTS

Reddit-Based Sentiment Analysis of Israel-Palestine Conflict | [link](#), Data Mining, NLP *August 2023 - December 2023*

- Employing the **Python Reddit API Wrapper (PRAW)**, I gathered **50,000+** comments and metadata on conflict.
- Preprocessed data and used **VADER sentiment analysis** on Reddit posts to categorize them into 3 sentiments, improving sentiment trend comprehension.

Wearable Device-Based Security: ML Authentication | [link](#), Data Mining, Machine Learning *January 2023 - April 2023*

- Expanded internship project to analyze authentication data from wearable devices, encompassing gyro, magnet, and accelerometer, data from 34 individuals.

McAfee CyberInsight Pipeline | Python, SQL, Apache airflow, Snowflake, K8's *January 2022 - July 2022*

- Utilizes **Snowflake** and **Apache Airflow** for efficient ETL processes, extracting feature information and monitoring firewall attacks from **40,000+** McAfee customers' servers.
- Deployed **Kubernetes** to host **20+ versions** of McAfee software in containers. Integrated Apache Airflow with **Apache File Sensor** for extracting debug files from customer servers. Leveraged Kubernetes for seamless deployment and data extraction, driving product enhancement initiatives.