VAIBHAVI KUNDLE

Tempe, Arizona| vkundle@asu.edu | +1(623)283-7322 | linkedin.com/in/vaibhavi-kundle/

EDUCATION

Master of Computer Science **Aug 2024 - May 2026**
Arizona State University, United States **(CGPA-4.0/4.0)**
Relevant Courses: Data Processing at Scale, Cloud Computing, Data Mining and Visualization.

Bachelor of Technology in Information Technology **Aug 2018 - Jun 2022**
K. J. Somaiya College of Engineering, Mumbai, India **(CGPA-3.6/4.0)**
Relevant Courses - Data Structures and Algorithms, Data Analytics, Database Management, Machine Learning.

TECHNICAL SKILLS

- **Programming Languages:** Python, SQL, JavaScript, HTML, CSS, C.
- **Libraries**: Numpy, Pandas, Keras, Scikit-learn, Matplotlib, OpenCV, SciPy.
- **Tools/Technologies:** Airflow, Hive, PySpark, Presto, MySQL, Tableau, AWS, Hadoop, Django, Docker, Git.

PROFESSIONAL EXPERIENCES

Meta | Data Engineer Intern **May 2025– Aug 2025**

- Architected and deployed **6+ Hive data sources** from scratch, designing schemas, keys, and dimensional models to track metadata accuracy across **13+ signals and 5+ critical Instagram ad events**, strengthening ad delivery reliability and revenue protection.
- Designed end to end ETL pipelines using Python, SQL, Hive, Presto, and Airflow to process petabyte scale ad interaction logs, implementing partitioning, bucketing, and dimensional modeling to cut query latency by ~40%.
- Implemented an automated data reliability layer with 15+ SQL checks, null monitoring, and alerting proactively detect data regressions and maintain high availability, correctness, and SLA compliance across analytics pipelines.
- Developed 3+ UniDash dashboards tracking event coverage, metadata health, and compute-to-revenue ROI, uncovering a 30% signal logging gap that contributed to revenue leakage and enabling data-driven logging fixes and resource reallocation to improve signal completeness and infrastructure efficiency.
- Partnered with product, data science, and infrastructure teams to translate ad delivery requirements into scalable, production-grade data platforms, influencing architectural trade-offs, performance optimizations, and downstream reliability.

ZS Associates | Data Engineer **Jul 2022 - May 2024**

- Engineered ETL workflows using SQL to generate and publish **~100** accurate monthly actionable business insights, driving measurable impact on client's sales performance.
- Reduced the risk of incorrect insights and system failures by **70%**, ensuring **100%** data accuracy for downstream analytics by developing Python and Pyspark scripts to manage data checks on Hadoop HDFS.
- Cut processing time **from 60 mins to 20 mins** by optimizing SQL queries and improving stability and efficiency of the insights processing system.
- Leveraged Python to create a reusable template of the complete data processing architecture, reducing manual configuration by **60%**, accelerating delivery timelines and ensuring consistency across multiple client projects.
- Automated software testing suites, reducing manual efforts by **45%** and accelerating the testing process by **2x**.
- Led the implementation and management of the Insights module for multiple projects, ensuring seamless collaboration, data-gathering, test setups, and operational handovers.

ZS Associates | Data Engineer Intern **Jan 2022–Jun 2022**

- Improved data integrity by performing **40+** quality checks, optimizing SQL queries in AWS Redshift and creating interactive **Tableau** dashboards.
- Contributed to data management solutions by developing pipelines using **SQL** and **Excel** to generate insights for clients. Led operational runs and client demos, resolving **70%** of issues in real-time and creating other internal documentation.
- Awarded the **Rising Star** award for exceptional performance and contributions to the Insights team, showcasing strong learning curve.

ACADEMIC PROJECTS

Video Captcha Proposition based on VQA, NLP, Deep Learning and Computer Vision

- Developed a software which acts as a video CAPTCHA for user authentication using ML and AI techniques.
- Integrated sentiment analysis while utilizing CNN and LSTM models, ensuring **98%** response validation accuracy.
- Authored and Published work on the project at the 2022 IEEE-ICAST Conference.