Siddhesh Sheth

✉ shethsiddhesh268@gmail.com   📞 8127784657   📍 Bloomington, IN, USA   in LinkedIn   ⬛ GitHub   ➤ Portfolio

EDUCATION

Master of Science in Computer Science   Aug 2022 – May 2024
Indiana University Bloomington | GPA: 3.97   Bloomington, United States

Bachelor of Engineering in Computer Science & engineering   Jul 2017 – Jul 2021 | Pune, India
Pune University | GPA: 4

PROFESSIONAL EXPERIENCE

Database Developer   Mar 2023 – present | Bloomington, USA
Indiana University, Bloomington
- Spearheaded the design and maintenance of SQL database schemas, yielding a notable 20% reduction in data storage footprint alongside a 15% improvement in retrieval speed.
- Engineered and executed complex SQL queries, stored procedures, and triggers, leveraging ETL methodologies for data transformation. This led to a 25% reduction in manual task completion time and a 30% increase in operational efficiency.
- Sales data from transactions were visualized using PowerBI dashboards, facilitating sector-specific profit analysis and informing optimization strategies leading to a 15% increase in profits and a 20% reduction in operational costs through restructuring efforts.

Full Stack Engineer Intern   May 2023 – Aug 2023 | Indianapolis, USA
Scotfin Tech
- Conceptualized and executed visually compelling web pages through Tailwind CSS and HTML5; handled asynchronous client requests using Ajax, leading to a significant 40% surge in user traffic; proficiently mitigated bounce rates by 25%.
- Conducted a comprehensive analysis of trends within the tax resolution industry with SEMrush and optimized local search visibility using SEO Toolkit. Then, using standard Javascript, refined the website to ensure close to 100% throughput.
- MySQL queries improved audits. Predefined schemas boosted data extraction by 30% for tax receipts, maintaining ACID properties, while normalization cut processing time by 20%, enabling faster access.

Software and Data Engineer   Aug 2021 – Aug 2022 | Pune, India
Accenture
- Oversaw the migration, updates, and backups of AT&T databases for Financial Billing operations through SSMS, guaranteeing no data loss and a 15% reduction in recovery time.
- Orchestrated the seamless transfer of packages, overseeing 100+ SQL Server databases with Azure DMS. Established streamlined handling protocols, resulting in a 20% reduction in migration time.
- As the lead RSA Archer Administrator, directed access roles for 100+ clients, addressing 50 queries/week. Introduced logical frameworks, diminishing recurring issues tied to governance, risk, and compliance by 75%.

PROJECTS

Netflix Trend Analysis and Recommendation System   Dec 2023 – Jan 2024
- The movie recommendation system, with TF-IDF vectors and diverse algorithms, achieved 95% precision in classifying movie types.
- Conducted EDA and visualized the top 10% of popular genres and released year patterns using Tableau and PowerBI.
- Created a Flask web application featuring 2 separate pages for trend analysis and movie recommendations.

Real-Time Weather Data ETL Automation with Apache Airflow on AWS   Jul 2023 – Aug 2023
- Orchestrated an ETL process for real-time weather data using Apache Airflow, emphasizing DAG design and Operator utilization.
- Optimizing workflows and enhancing data pipeline with AWS Glue reduced manual intervention on AWS by 30%.
- Integrated AWS Athena and Kinesis with real-time data for swift data insights, aiding weather forecasting and decision-making.

Analyzing Stock Market data in Real-Time using Kafka-driven ETL.   Jun 2023 – Jul 2023
- Performed extraction of real-time stock market data in .csv format and then launched an EC2 VM instance for the Linux environment.
- Instantiated a zookeeper, created a topic within a producer, and delivered it to the consumer and S3 via a cluster using Kafka Connect.
- Leveraged Apache Hive to enhance business intelligence processes, resulting in a 20% improvement in data analysis efficiency.

Literacy Education Nexus ⬏   Jan 2023 – Apr 2023
- Analyzed SOFI and World Bank data with Python and Tableau to visualize literacy education's link to global development.
- Utilized linear regression algorithms to forecast key parameters like GDP, employment, and school enrollment trends until 2040.
- Curated PowerBI visualizations to advocate education as a human right by showcasing its positive impact on economic growth.

SKILLS

Databases (PostgreSQL, Oracle, MongoDB, DynamoDB, NoSQL, RedShift, BigQuery, Firebase, Snowflake, Cassandra),
Programming Languages (Python, R, Scala, SQL, T-SQL, MySQL, C++, C, Java, HTML,), **Cloud** (AWS, Microsoft Azure, Google Cloud),
OS (Windows, Linux, MacOS, Unix), **Data Engineering Tools** (Py Spark, Kafka, Airflow, Hive, Hadoop, Data Lake, Data Warehouse),
Platforms & Security (RSA Archer, SecureID, SSMS, Jupyter, Apache, Anaconda, VS Code, Eclipse, Google Colab, Shell, Bash, CLI, VAPT),
Other Tools and Technologies (HTML, CSS, ASP.NET, JavaScript, TypeScript Bootstrap, RESTful API, Docker, Kubernetes, Databricks,
CI/CD, AWS Kinesis, Lambda, Glue, Rekognition, Quicksight, Tableau, Power BI, SEMRush, Git, JIRA, Kanban, Flask, SDLC, Agile, Salesforce
CRM, Django, Node.js Express.js, React.js, Angular.js Flutter)

CERTIFICATES AND PUBLICATIONS

- **Amazon Web Services Certified Cloud Practitioner** ⬏ (57H7X4627EQEQ8W8),
- **Google Data Analytics Professional Certificate** ⬏ (QYPU4KC8V8CF),
- **CNSS Certified Network Security Specialist from ICSI, UK** ⬏ (18504724),
- **Review paper "Stock Prediction using Sentiment Analysis"** ⬏ ,
- **A research paper on "Breast Cancer Detection using Machine Learning" was also published in springer.** ⬏