Siddhesh Sheth

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

SUMMARY

Inquisitive software and data engineer with a couple of years of corporate experience proficiently managing client applications and databases, committed to enhancing businesses with unwavering diligence.

PROFESSIONAL EXPERIENCE

Database Developer   Mar 2023 – May 2024 | 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.
- Proficient in extracting and cleaning 100% of data using **openpyxl** on computer inventory Excel sheets and SQL for advanced querying.
- Utilized **PowerBI dashboards** to visualize sales data, driving sector-specific profit analysis, thus achieving a **15% profit increase.**

Graduate Teaching Assistant   Aug 2023 – Dec 2023 | Bloomington, USA
Indiana University, Bloomington
- Direct hands-on 1.5-hour sessions focusing on configuring routers and switches and implementing protocols like **TCP/IP** twice a week.
- Supervised troubleshooting exercises and conducted mock reviews, showcasing my exemplary **leadership and communication** skills.
- Assisting over **250 students** weekly with topics including network topologies, routing algorithms, subnetting, and OSI model layers.

Full Stack Engineer Intern   May 2023 – Aug 2023 | Indianapolis, USA
Scotfin Consulting Co.
- Conceptualized and executed visually compelling web pages through **Tailwind CSS and HTML5**; handled asynchronous client requests using **Ajax/JS**, leading to a significant **40%** surge in user traffic; proficiently mitigated bounce rates by **25%**.
- Conducted an in-depth analysis of trends in the tax resolution industry with **SEMrush** and optimized search using SEO Toolkit by **60%**.
- Improved **MySQL** schemas boosted data extraction for tax receipts, maintaining ACID, while normalization enabled **20%** 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** Admin, directed access roles, addressing at least 50 queries/week and reducing issues tied to **GRC by 75%**.

Data Analytics Engineer   Nov 2020 – Apr 2021 | Mumbai, India
Ticket On Go
- Led the implementation of AWS cloud infrastructure for analytics, utilizing Amazon Redshift and EMR for scalable data processing.
- Continuously capture and store ticketing data as it is generated in real-time via Amazon Kinesis Data Streams.
- Designed and executed SQL-based pipelines to ensure reliable insights for the Ticket On Go project.

SKILLS

Databases — PostgreSQL, Oracle, MongoDB, DynamoDB, NoSQL, RedShift, BigQuery, Firebase, Snowflake, Cassandra, Azure Data Factory, **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** — ASP.NET, RESTful API, Docker, Kubernetes, Databricks, CI/CD, AWS Kinesis, Lambda, Glue, Rekognition, IoT IP, Excel, TensorFlow, version control, FPGA, Pytorch, Quicksight, Tableau, Power BI, JIRA, Kanban, Flask, SDLC, Agile, Salesforce CRM

EDUCATION

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

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.

CERTIFICATES AND PUBLICATIONS

Amazon Web Services Certified Cloud Practitioner ⧉ — 57H7X4627EQEQ8W8

Google Data Analytics Professional Certificate ⧉ — QYPU4KC8V8CF

A review paper on "Literacy Education Nexus" ⧉

A research paper on "Breast Cancer Detection using Machine Learning". ⧉