# KOTESWARA CHARAN TEJA AYODHYA

**DATA ENGINEER**

**Location:** CT | **Phone:** (203)-679-9147 | **Email:** ayodhyakoteswara@gmail.com | LinkedIn |GitHub


## SUMMARY


Accomplished Data Engineer with around 5 years of experience, adept at designing and implementing end-to-end data pipelines to streamline data processing workflows. Proficient in Python, SQL, and R, with expertise in managing diverse databases including MYSQL, MS SQL Server, Oracle, MongoDB, and PostgreSQL. Skilled in utilizing libraries like NumPy, Pandas, and Scikit-Learn for data manipulation and analysis, alongside Matplotlib for visualization. Experienced in cloud technologies such as AWS and Azure, and proficient in Big Data tools like Hadoop, Apache Spark, and Hive. Proficient in ETL processes using tools like Informatica PowerCenter, Talend, Informatica Intelligent Cloud Services (IICS), and Snowflake. Expertise in SDLC, Agile, and Waterfall methodologies, with proficiency in version control using Git and GitHub, and comfortable working with IDEs like PyCharm, Jupyter Notebook, Visual Studio, and Anaconda.


## TECHNICAL SKILLS

| | |
|---|---|
| **Programming Language:** | Python, SQL, R. |
| **Databases:** | MYSQL, MS SQL Server, Oracle, MongoDB, PostgreSQL, Teradata. |
| **Libraries:** | NumPy, Pandas, Scikit-Learn, Matplotlib. |
| **Cloud Technologies:** | AWS, Azure**.** |
| **Big Data Tools:** | Hadoop, Apache Spark, Hive, Pig, MapReduce. |
| **Visualization Tools:** | Tableau, Power BI, MS Excel. |
| **ETL Tools:** | Informatica Intelligent Cloud Services (IICS), Snowflake, Informatica PowerCenter, Talend, SSIS. |
| **Methodologies:** | SDLC, Agile, Waterfall. |
| **Version Control Tools:** | Git, GitHub. |
| **IDES:** | PyCharm, Jupyter Notebook, Visual Studio, Anaconda. |

## EDUCATION

**University of Bridgeport** — **Bridgeport, USA**
Master of Science in Computer Science. — Sept 2022 – Nov 2023

**Prasad V Potluri Siddhartha Institute of Technology** — **Vijayawada, India**
Bachelor of Technology in Information Technology. — Jun 2015 – Apr 2019

## PROFESSIONAL EXPERIENCE

**McKesson** — **CT**
**Data Engineer** — **Feb 2024 - Current**

- Designed and implemented robust data pipelines using Informatica Intelligent Cloud Services (IICS), streamlining ETL processes for over 30 diverse data sources, reducing data processing time by 40%.
- Managed and optimized large-scale data storage systems, including Snowflake and Teradata, enhancing data retrieval efficiency and supporting complex analytics applications.
- Developed and maintained real-time data processing solutions with Apache Spark, reducing system response times by 50% and enhancing operational efficiency for key business processes.
- Design, implement, and optimize scalable data models and transformations using DBT and Snowflake, resulting in a 50% reduction in transformation time and a 30% increase in data accuracy
- Developed a Pandas-powered Python GUI tool for automated data analysis across prod, UAT, and dev, saving 40% in manual effort, increasing data consistency by 25%, and enabling dynamic reporting for informed decision-making.
- Administered MySQL and MongoDB databases, enhancing schema design and performance tuning, which led to a 35% increase in query efficiency and supported 500 million+ data transactions.
- Engineered and maintained ETL pipelines on Azure utilizing Azure Data Factory, Azure Blob Storage, and Azure Functions, optimizing data processing and transformation workflows for enhanced efficiency.
- Utilized Scikit-Learn and Matplotlib for data analysis, creating 50+ interactive dashboards and reports in Tableau and Microsoft Excel that influenced strategic business decisions.

**Accenture Solutions Private Limited** — **India**
**Data Engineer / Application Development Analyst** — **Oct 2021 - Sep 2022**

- Delivered a 40% increase in end-to-end implementation execution by analyzing business cases and refining version control strategies.
- Spearheaded the successful migration of Salesforce data to Oracle, reducing data processing time by 30% and enhancing system scalability.
- Excelled in developing efficient Informatica Cloud mappings, achieving a 50% increase in data processing efficiency and a 75% reduction in replication errors, enabling organizations to make informed decisions and drive success.
- Leveraged ServiceNow's service catalog for efficient resource allocation, resulting in a 15% cost reduction in the Informatica IICS project implementation.
- Executed seamless integration of Informatica IICS on Linux, enhancing system stability by 15% and reducing maintenance efforts by 25%.
- Employed Snowflake's fully-managed cloud data platform, which offered scalable computing power to handle variable workloads

without the need for physical hardware management.
- Drove a 30% reduction in query response time and a 20% increase in system efficiency through expert database performance tuning and optimization.
- Improved query performance by using Snowflake's automatic clustering to reorganize data into optimally clustered formats, which significantly reduced query times and increased overall system performance.
- Enhanced data integrity and system responsiveness by streamlining data flows and storage solutions across multiple platforms, leading to improved performance benchmarks.
- Led a team in the strategic overhaul of network infrastructure, which improved data transmission rates by 20% and reduced system downtime by 35%.
- Implemented a data backup strategy reducing loss incidents by 50% and sped up recovery. Orchestrated system updates for seamless rollout, boosting user adoption by 25%.

**Cognizant Technology Solutions Private Limited** **India**
**Data Engineer / Programmer Analyst** **Jul 2019 - Sep 2021**

- Executed efficient data integration solutions within Informatica IICS, which led to a 25% improvement in ETL performance achieved by optimizing PostgreSQL database queries.
- Employed streamlined project workflows utilizing Jira, resulting in a 20% reduction in task completion time for the Informatica IICS integration project.
- Implemented cost-effective storage solutions on AWS S3, resulting in a 25% reduction in overall project infrastructure expenses for Informatica IICS.
- Leveraged Integrated Informatica Cloud Services (IICS) and SQL Server Integration Services (SSIS) for advanced data integration tasks, significantly improving data quality and accessibility.
- Developed and automated end-to-end data pipelines within Snowflake using its native support for tasks and streams.
- Automated repetitive tasks and data cleaning processes using Python scripts, leading to a 40% reduction in manual workload and ensuring data consistency across the platform.
- Enhanced data processing in Informatica Intelligent Cloud Services (IICS) with a 30% efficiency boost through the utilization of Optimized Transformations.
- Spearheaded the development of a scalable data pipeline architecture on Azure, leveraging a comprehensive suite of Azure technologies including Azure Databricks, Blob Storage, Functions, Data Factory, and Synapse Analytics, to enable seamless analytics and insights generation.
- Collaborated with cross-functional teams to design and implement data integration strategies, ensuring alignment with business objectives and requirements.
- Conducted performance tuning exercises for Informatica IICS workflows, identifying and resolving bottlenecks to enhance system efficiency.
- Implemented automated monitoring and alerting mechanisms within Snowflake, enabling proactive identification and resolution of issues.

**Espire** **India**
**Jr. Data Engineer** **May 2018 - Jun 2019**

- Developed Python scripts to automate data quality checks, reducing data error rates by 30% and enhancing the integrity of datasets used for critical analysis.
- Implemented complex SQL queries for data aggregation and joins, enabling detailed analysis that supported increased report accuracy and decision-making efficiency.
- Developed and maintained PostgreSQL databases, increasing data retrieval speeds by 25% and improving database security compliance by 40%.
- Optimized SQL queries and database performance, identifying and resolving bottlenecks that resulted in a 30% faster data processing and a 20% reduction in load times.
- Utilized NumPy and Pandas to efficiently manipulate large arrays of numerical data and structured datasets, respectively, resulting in enhanced accuracy of computed metrics, a 25% improvement in data exploration efficiency, and a 20% increase in stakeholder satisfaction.
- Designed and implemented scalable data solutions on AWS EMR and RDS, utilizing auto-scaling to handle workload fluctuations, reducing operational costs while maintaining data availability.
- Leveraged Hadoop for distributed storage and processing of large datasets, enhancing data processing capabilities and facilitating advanced analytics initiatives.
- Utilized Power BI's DAX to create complex calculated columns, which enhanced report detail and accuracy, influencing strategic decisions that improved project outcomes by 15%.
- Managed and analyzed large datasets in MS Excel, identifying trends and KPIs that led to a 10% improvement in market analysis accuracy and supported strategic planning.

## PROJECTS

**Image Classification System Successfully Using Python**

- Developed an image classification system using Python and Flask, achieving a 20% improvement in accuracy, and catering to a user base of over 1,000,000 individuals (about the population of Delaware).

## CERTIFICATION

Informatica Cloud - Data Integration certification.