Bhanu Krishna Motamarri
Data Engineer

F1 OPT Visa | Open to Relocate | bhanu.m@my-itcareer.com | +1 6177089878| LinkedIn | GitHub
| Massachusetts |

SUMMARY

- Data Engineer with Around 4 years of experience in building data-intensive applications, tackling challenging architectural and scalability problems.
- Proficiency in complete Software Development Life Cycle SDLC for projects using methodologies like Agile and Waterfall.
- Knowledge of using various packages in Python like pandas, NumPy, seaborn, SciPy, Matplotlib, and sci-kit-learn.
- Well-versed in handling complex data ecosystems, including Hadoop, Pig, Hive, and Apache Spark, to enable efficient data processing and analysis.
- Skilled in ETL processes using SSIS and cloud technologies like AWS, GCP and Snowflake for seamless data integration.
- Proficient in creating impactful visualizations with Tableau, Power BI, and SSRS to communicate insights effectively.
- Proven knowledge of writing SQL Queries, Stored procedures, functions, packages, and triggers using a database such as MongoDB and MySQL.

EDUCATION

Master of Science in Data Analytics Engineering
Northeastern University, Boston, USA

Bachelors in computer science and engineering
BML Munjal University Delhi, India

SKILLS

Methodology: SDLC, Agile, Waterfall
Programming Language: R, Python, SQL
IDE's: PyCharm, Jupiter Notebook, Google Collab
Big Data Ecosystem: Hadoop, MapReduce, Hive, Apache Spark, Pig
ETL Tools: SSIS, Snowflake
Cloud Technologies: AWS, Azure, GCP
Packages: NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow
Reporting Tools: Tableau, Power BI, SSRS
Database: MongoDB, MySQL
Other Tools: Git, MS Office, Medallia

EXPERIENCE

Ethos Farm | Nov 2023 - Current | Data Engineer

- Follow an Agile methodology and involve daily Scrum meetings, sprint planning, showcases, and retrospectives.
- Leverage PyCharm and Jupyter Notebook to develop and test code, ensuring high-quality and error-free data processing and analysis.
- Working with the big data ecosystem, including Hadoop, MapReduce, Hive, Apache Spark, and Pig, to process and analyze large datasets efficiently.
- Reducing data integration time by 25% through streamlined ETL processes, leading to faster access to accurate and up-to-date data.
- Employing packages like NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, and TensorFlow for data manipulation, analysis, and machine learning tasks.
- Collaborating with cross-functional teams to migrate on-premises data to GCP, ensuring seamless data integration and accessibility.
- Improving stakeholder engagement by 20% through clear and visually appealing data visualizations, aiding in data-driven decision-making.
- Creating impactful visualizations using Tableau, and SSRS to communicate data insights effectively to stakeholders.
- Performing end-to-end Architecture and implementation assessment of various AWS services like Amazon EMR, Redshift, and S3.
- Achieve consistent cross-platform compatibility, resulting in 70% operational efficiency and reduced development time.

Northeastern University | Oct 2022 - Aug 2023 | Research Assistant

- Conducted statistical analysis on Emergency Medical Services data for 240 counties, contributing to research published in a peer-reviewed journal.
- Improved existing system in identifying nearest rescue locations using nearest neighbor search algorithms.
- Collaborated with research team to visualize detailed maps and informative histograms to improve interpreted patterns.

LARSEN TOUBRO DEFENCE | Jan 2020 - Mar 2021 | Data Engineer

- Collaborated with R&D to address business problems and reduced maintenance costs of Naval Ships by 30%.
- Implemented Apache Spark-based ETL with MySQL and Redis to process 85% gigabytes of data to clean, validate, and transform the raw data.
- Utilized Predictive Modeling techniques in Python for anomaly detection extending ship lifetime and saving $1M in service costs.
- Developed Plotly Dash dashboards to decrease time to insight by 15% and reduce maintenance downtime by 20% in the Naval ship's infrastructure.

Blue Light IT Solutions | Jan 2019 - Dec 2019 | Data Engineer

- Worked with relative ease with different working strategies like Waterfall methodology.
- Implemented a data pipeline using Spark and Hive to ingest customer behavioral data into the Hadoop platform to perform user behavioral analytics.
- Operated on migrating MapReduce programs into Spark transformations using Spark and Scala.
- Involved in Data Architecture and Application Design using Cloud and Big Data solutions on AWS, and Microsoft Azure.
- Generated property list for every application dynamically using Python modules like NumPy, Matplotlib, seaborn, and pandas.
- Leveraged GCP such as Google Big Query for data warehousing, Google Dataflow for data processing, and Google Cloud Storage for data storage.
- Developed dashboards and visualizations to help business users analyze data as well as provide data insight to upper management with a focus on Microsoft products like SQL Server Reporting Services (SSRS) and Power BI.
- Designed and developed SSIS Packages to import and export data from MS Excel, SQL Server, and Flat files.
- Used Git repository hosted on a cloud platform.

CERTIFICATIONS

AWS (Cloud Practitioner), Power BI (Power BI Data Analyst)