**Harinath Reddy**

**+1 972-924-9390 | madhu@viratsolutions.com**

Versatile professional seeking a challenging role that leverages a unique blend of data engineering and data science skills to drive insightful decision-making and innovation. Over 12 years of experience in analyzing, developing, and processing large-scale applications using big data ecosystems, data management and visualization tools. Data-driven decision making and holds strong problem-solving skills, always excited to learn new things and contributing to Data science community for the past 10 years.

**PROFESSIONAL SUMMARY:**

* Experience with the data warehousing platforms Amazon Redshift, Azure, OCI and Google Cloud Platform.
* Advanced skills in machine learning and statistical modeling for predictive and prescriptive analytics.
* Experience with popular machine learning libraries such as scikit-learn, Keras, TensorFlow, and PyTorch.
* Experience in preparing ad-hoc reports and visual story telling.
* Good Expertise in ingesting, processing, exporting, and analyzing Terabytes of structured and unstructured data.
* Hands-on experience in Database tuning and Query tuning.
* Experience in building computer vision models (CNNs), and Deep Neural Networks on real world applications.
* Practical experience in MLOps and enhancing the existing platform for effective CI/CD.
* Utilized Databricks notebooks for interactive data exploration and model development.
* Proficient in Python, Java, and Scala for data manipulation, analysis, and software development.
* Knowledge of common database management systems like MySQL, Microsoft SQL Server, and Oracle Database.
* Experienced in the design, development, and Implementation ofBig data applications.
* Strong statistical modelling and analytical skills and able interpreting the research results to clients.
* Involved and contributed to all phases of the Software Development Life Cycle (SDLC) in large-scale enterprise software using Object-Oriented Analysis and Design.
* Possess strong skills in application programming and system programming using Python on Windows and LINUX platforms using principles of Object-Oriented Programming (OOPS) and Design Patterns.
* Strong analytical and problem-solving skills and the ability to follow through with projects from inception to completion.
* Ability to work effectively in cross-functional team environments, excellent communication and interpersonal skills.
* Motivated and highly detail oriented, always eager to learn new things and keep up to date with the domain facelifts.

**TECHNICAL SKILLS**

Programming Languages C#, Python, R.

Web Technologies/Frameworks JavaScript, Bootstrap, HTML/CSS, XML, Vue-JS, Python-Flask.

Web/ Cloud Services Amazon Web Services, Google Cloud Platform, Oracle Cloud Infrastructure, Heroku Dynos, Azure.

Databases SQL Server, MongoDB, MySQL, and PostgreSQL.

ETL Tools Amazon Redshift, SSIS.

Scripting Languages Shell, Perl, PHP.

Visualization Tools Tableau, Grafana, Power BI.

Operating Systems Windows server 2012/2016, Linux, Debian, Ubuntu, CentOS.

Source (version) Control GIT, TFS, SVN, Bitbucket, Jenkins, Gerrit, Bitbucket.

**PROFESSIONAL EXPERIENCE:**

**Client: Charter Communication, Stamford, CT December 2020 - Present**

**Role: Sr Data Analytics Engineer**

* Analyzed large amounts of data in parquet format on S3 buckets using Spark SQL.
* Performed statistical analytics and helped research scientists in Data Modelling.
* Built predictive and speech recognition models to improve teams performance and effeciency.
* Configured and setup Azure Hybrid connection to pull data from SAP systems.
* Implement continuous integration and deployment (CI/CD) pipelines for Databricks Jobs.
* Used Kafka and Spark Streaming for processing transaction messages.
* Writing Complex SnowSQL queries in snowflake cloud data warehouse to retrive and maintain business analytics reports.
* Migrated on-premises data to the cloud using tools like DBT (Data Build Tool) and Snowflake.
* Expertise in creating, debugging, optimization of ETL Data Pipelines using Apache Airflow to load data into Data
* Stored and retrieved data from data-warehouses using Amazon Redshift.
* Worked on scheduling jobs using Airflow scripts using python.
* Created data visualization dashboards and reports using Power BI and Tableau.
* Developed various test case scenarios based on rules provided by the business using Spark-Scala APIs.
* Worked closely with Business analysts and Enterprise architects to understand the rules provided by the business.

**Client: UPS, Fairfax, VA April 2017 - November 2020**

**Role: Sr Data Analytics Engineer**

* Engineered Programmatic Advertising Data and API integration solutions as part of self-service SPMP (SmartPrivate Marketplace) for campaigning and measuring media investments across inventory channels.
* Designed and implemented ETL/Data Pipelines to ingest disparate datasets from external SSP/DSP partners intoa Data Lake hosted in AWS cloud.
* Used AWS Glue for the data transformation, validation, and data cleansing.
* Analyzed and Transformed SQL scripts into Pyspark SQL scripts for optimized and faster performance.
* Designed and developed data architecture for OLAP cubes and data marts using star schema-based data models.
* Created scripts to read CSV, JSON and parquet files from S3 buckets in Python and load into AWS S3
* Developed and optimized ETL Data Pipelines using technologies like AWS Data Pipelines, Airflow,
* Unix Shell Scripts, Python modules, Athena queries, AWS Lambda Functions, ECS tasks, JSON config, AWS DMS, Marketplace Qlik Attunity Replicator, and others.
* Implemented Data Quality Control by testing and implementing database constraints.
* Developed SSIS ETL and migrated data from legacy systems into new database structures.
* Developed testing procedures and workbooks to compare counts between source and destination for data loads.
* Migrating API, Storage and Query Engines from AWS Beanstalk, S3, Athena to Azure counterparts API Services.

**Client: Innova Solutions, PA September 2014 - March 2017**

**Role: Data Analyst**

* Designed and deployed high performance data pipelines for Data Lake and Analytical applications.
* Created AWS Athena tables and queries for ad hoc data analysis.
* Developed Informatica Power Center mappings, workflows to convert legacy SSIS ETL
* Executed Data Analysis and Data Visualization on survey data using Tableau Desktop as well as compared corrrespondent’s demographics data with Univariate Analysis using Python (Pandas, NumPy, Seaborn, Sklearn, and Matplotlib).
* Built customized interactive reports, worksheets, and dashboards on streaming applications using Tableau.
* Used version control tools Git to update the project with team members.
* Working with an agile methodology to ensure delivery of high-quality work with monthly iteration.
* Performed all necessary day-to-day Git support for different projects, Responsible for maintenance of the Git-repositories, and the access control strategies.

**Value Labs, India August 2011 - February 2012**

**Role: Application/Java Developer**

* Designed a system and developed a framework using J2EE technologies based on MVC architecture.
* Involved in the iterative/incremental development of project applications. Participated in the requirement analysis and design meetings.
* Designed and Developed UI using JSP by following MVC architecture.
* Designed and developed Presentation Tier using the Struts framework, JSP, Servlets, TagLibs, HTML, and JavaScript.
* Designed the control which includes Class Diagrams and Sequence Diagrams using VISIO.
* Used the STRUTS framework in the application. Programmed the views using JSP pages with the struts tag library, Model is a combination of EJB’s and Java classes, and web implementation controllers are Servlets.
* Generated XML pages with templates using XSL. Used JSP and Servlets, EJBs on server side.

**EDUCATION AND CERTIFICATIONS:**

Bachelor of Technology in Computer Science from GITAM University, India. 2011

Masters in Computer Science Engineering from George Mason University. 2012-14