Tejasvi Venkata Krishna Sarraju

Hyderabad, India
(+91) 9989787697
svktejasvi@gmail.com

Summary

- 12+ years of experience in IT, including onsite/client locations.
- Persistent problem solver. Avid learner. Devoted to Data Science and Engineering.
- My recent roles: Data Scientist, Data Architect, Tech Lead, Senior ETL Developer.
- Designed enterprise solutions on cloud optimizing for performance, cost, development effort and scalability, while achieving business goals on time.
- Focused on generating business value by deriving actionable insights from raw datasets and building advanced analytics solutions from scratch.
- Have coded recently in Python. Coded previously in R, Java. Can quickly adapt to any programming language or tool.
- Used Databricks with PySpark to extract/process data in a distributed environment.
- Built variety of supervised and unsupervised Machine Learning and Deep Learning models. Experienced in end-to-end lifecycle of data wrangling, feature engineering, model selection, hyper parameter tuning and evaluation of the results.
- Operationalized the analytics pipelines in AWS cloud platform, primarily using SageMaker.
- Adopted DevOps strategies using Jenkins, GitHub.
- Experienced in multiple Time Series Analysis and Forecasting techniques.
- Used computer vision algorithms for skin texture analysis and PII data masking in facial images.
- Have strong understanding of descriptive and inferential statistics concepts.
- Created reactive web apps using R Shiny, to allow users to interact with the plots via browser.
- Worked extensively in Data Warehousing space. Strong in ETL/SQL solutions. Have exposure to multiple Informatica tools.
- Designed Enterprise Data warehouse and modelled Star schemas from scratch by adhering to Ralph Kimball methodologies.

Technical Profile

Certifications: AWS Certified Machine Learning - Speciality (Badge), Informatica PowerCenter Data Integration 10 Developer (ID: 067-000032), IZ0-051 Oracle Database 11g: SQL (ID: 249572474)
Analytics: Machine Learning, Deep Learning, Time Series Forecasting, Statistics, Computer Vision
Programming Languages: Python, R, Java, SQL
Cloud services: AWS: SageMaker, Lambda, S3, Glue, Kinesis, Athena, CloudFormation, EC2
Azure: Blob Storage, Search, App registrations, Graph API + SharePoint
Big data: Databricks, PySpark; **DBMS:** Oracle, Microsoft SQL Server
DevOps: Jenkins, GitHub
Visualisation: Python (matplotlib, seaborn), R Shiny
Other Tools: Informatica: PowerCenter (10x), Metadata Manager, DVO; Control-M; Power Designer
Operating systems: UNIX, Windows

Employment

Cognizant Technology Solutions   **Mar 2016 – till date | 5. years**
DLF Building, Hyderabad, India
<u>Senior Data Scientist</u>

Cognizant is one of the top 5 IT services company – listed in NASDAQ-100, Fortune 200. A highly reputed player in Analytics and Digital market, having customers all over the globe touching every industry/domain.

<u>**Client: Thermo Fisher Scientific | Domain: Life Sciences | Duration: 2.6 years**</u>

<u>Project summary</u>: Thermo Fisher Scientific has an annual revenue exceeding $25 billion. They provide products and services in the areas of Life Sciences, Industrial & Applied Sciences, Clinical & Diagnostics, Lab Solutions, and many more in Biotechnology. By leveraging advanced analytics, we reduced 45% of failures and prevented numerous performance issues, hence streamlining a complex BI environment.

<u>Contributions</u>:

Data Scientist & Engineer
- Leveraged Databricks platform to build CDC jobs and delta lakes sourcing supply chain data
- Designed solution architecture from scratch by integrating Databricks, AWS S3, Lambda, SageMaker, Glue and Athena services.
- Built various Machine Learning models to detect anomalies in the operations.
- Generated actionable insights from BI Operations Audit/Log data.
- Extensively performed EDA and Feature Engineering.
- Used Time Series algorithms (like SARIMA, Facebook prophet) for forecasting and detecting anomalies in job performance.
- Adopted DevOps strategy by building a CI/CD pipeline to deploy the end-to-end solution using GitHub, Jenkins, Databricks CLI and AWS CloudFormation.
- Periodically demonstrated the solution to Senior and Executive level management up to Vice President level

<u>**Client: Unilever| Domain: Clinical | Duration: 5 months**</u>

<u>Project summary</u>: Unilever is a global consumer goods company, headquartered in UK. We built a solution for Unilever's R&D department, end users being scientists and clinicians. The solution involved ability to store/search huge volumes of images that get generated over a time series during clinical trials, extract features from images and run Analytical pipelines.

<u>Contributions</u>:

Data Scientist & Engineer
- Generated features from images and masked certain areas of images for de-identification using computer vision algorithms.

- Gathered requirements from domain experts and built plots accordingly by implementing statistical models in R. Eg:
  - Show average change from baseline per product per feature. Highlight the products showing significant change at agreed critical p-values.
  - Display Top 5 panellists who showed the highest improvement per product.
  - Feature based trend to show product efficacy by age.
- Used R Shiny to render reactive plots into a browser and provided users the flexibility to interact with plots.
- Used multiple Python packages such as pandas and numpy for data transformations as required for Analytical pipelines.
- Programmatically uploaded images to Azure BLOB storage and associated metadata by fetching information from csv files and image names.
- Programmatically searched and read BLOBs by hitting Azure search service via REST APIs.
- Generated dummy images programmatically by rendering text and overlaying masked/unmasked images as required for testing.

Client: Novartis| Domain: Healthcare | Duration: 3 months

Project summary: Novartis business is primarily focussed on pharmaceuticals, generics, and oncology. Huge volumes of their data reside in HDFS and loaded via Spark. Analytics is performed using R, Tableau, and multiple other tools.

Contributions:

Data Engineer
- Write Spark SQL and Python code to load data from multiple sources into HDFS.
- Build a generic framework to use across all the jobs.
- Write shell scripts to submit multiple Spark jobs using the generic framework.
- Analyse the existing complex PL/SQL code and convert them into Spark SQL code.

Client: UBS (Swiss Bank) | Domain: IT & Infrastructure | Duration: 2.4 years

Project summary: UBS Group AG is a Swiss multinational investment bank and financial services company. Our team was responsible for developing data warehouse and reporting solutions for UBS management to analyse IT Service Management and Infrastructure data. Data is primarily extracted from ITSM (Service Now), HR and Asset management tools.

Contributions:

Data Architect
- Designed Star schemas for IT Service Management data.
- Built logical and physical models in Power Designer tool. Ensured that the models are always in sync with database schemas.
- Ensured that the data marts fit well into overall Enterprise DWH architecture.
- Proposed ETL designs to developers. Guided the team with ETL best practices.

- Ensured that end-to-end (Sources to reporting) data lineage is always maintained in Informatica Metadata Manager tool.
- Simplified the usage of Metadata tools to reduce documentation efforts for developers.

Project Lead/Senior Developer
- Implemented complex business logics by considering code quality.
- Tuned code iteratively at Informatica, database and system levels to meet 10 min SLA per load to manage huge impact/risk.
- Had been a POC and SME for all requirements/issues around Service Now project.
- Played a BA role and explained the business requirements to the developers.
- Proposed detailed, robust project plans with effort estimates to the client.
- Adhered to Agile methodology; used JIRA to manage sprints effectively.

Achievements: Won award for success of Service Now project under my leadership.
                 Got appreciated at Vice President Level at Thermo Fisher Scientific client

Oakton, A Dimension Data Company   Dec 2010 – Feb 2016 | 5.2 years
36, Jubilee Hills, Hyderabad, India
Data Warehouse Team Leader

Oakton is a 30-year-old Australian based leading IT consulting firm. Works on Australian government and commercial projects in Australian/New Zealand market. Oakton is currently owned by Dimension Data, an 8 billion USD global company.

Client: GE Capital ANZ | Domain: Banking | Duration: 5.2 years

Project summary: GE Capital runs complex data warehouse projects due to its huge customer base, variety of products and compliance needs. Our team handled end-to-end Data Warehouse solution including Dev, L3 support and Administration. Huge focus lies in data quality and performance to manage volumes generated by Credit card transactions and various applications dealing with personal, vehicle loans and insurance.

Contributions:

DWH Team Lead Offshore
- Worked closely with Onshore Team to discuss day to day priorities, status updates, escalations and improvements.
- Had driven highly complex projects by coordinating with multiple teams.
- Prepared shift roster and planned resourcing for enhancements, deployments & BAU tasks.
- Handled a team of 9 members by guiding the team and balancing the workload effectively.

Development
- Developed complex mappings while working with several types of sources/transformations/ targets.
- Spent significant efforts into performance tuning at Informatica/database/system levels.
- Experienced with complex data models and comprehensive audit framework.
- Included logic to build robust ETLs by maintaining dependencies.

- Had written SQLs to provide retrospective data fixes.

Production Support
- Monitored and fixed production ETL failures including 24x7 on call support.
- Debugged data inconsistencies while working directly with business.
- Performed Admin tasks across production and pre-prod environments, e.g., handling the Informatica outages, installing Informatica hotfixes, upgrading software.
- Built a pre-prod environment from scratch for production support team including: PowerCenter installation, full repository migration.

Onsite responsibilities at client location – GE Capital Australia & NZ (8 months in Melbourne)
- Worked with Business Analysts to understand issues in production and provided workarounds/ quick fixes and permanent solutions.
- Worked closely with Release teams while planning major deployments.
- Worked with Service Stability teams to drive technical and process improvements.
- Setup meetings and gather required support from multiple teams to solve complex issues.

Achievements: Received Bright New Talent award within 3 months of joining.
     Won best team award twice in my leadership.

SumTotal Systems, A Skillsoft Company Aug 2009 – Dec 2010 | 1.4 years
Mindspace Raheja IT Park, Hyderabad, India
Associate Engineer Trainee, Cloud Operations

Project summary: SumTotal Systems is the global leader in complete learning and talent management software that enables organizations to drive business strategy more effectively. Our team provided cloud services for 2000+ customers with end-to-end site administration.

Contributions:

Database & Windows Server Administration
- Monitored Transaction Log issues and SQL job failures, including ETLs. Took necessary actions to rectify and prevented them from re-occurring.
- Performed Backup and restoration of Production SQL server databases.
- Debugged blockings and other database bottlenecks that are impacting the applications.
- Handled Memory and CPU bottleneck issues.
- End-to-end site monitoring and server maintenance.
- Monitored and analysed performance graphs from VM Ware.
- Wrote batch scripts to help maintenance activities over 500 windows servers.
- Worked on automation tool called Active batch – Creating jobs, scheduling, and monitoring.

Achievements: Provided automated solutions and reduced the volume of alerts.
     Appreciated by management for taking up next level/complex tasks.

Education

B. Tech graduate in Computer Science Engineering, Guru Nanak Engineering College, JNTU, 2009.