Tushar Gupta

Data Scientist at WITHIN
New York City Metropolitan Area

Summary

A full-stack data scientist with four years of experience in strategy consulting, data analysis, and product incubation. I have managed and deployed end-to-end data science projects for clients from multiple industries including BFSI, Hospitality, E-Commerce, and Telecom. I am a proactive problem-solver and a structured thinker with an immense passion for diving into complex data and pulling out insightful and actionable information to drive customer success.

I have worked on all the phases of an analytics project pipeline including:
• Building of machine learning and deep learning models
• Creation of interactive data visualization and self-service reporting portals
• Productionizing of data science solutions on-premises and on-cloud
• Development of real-time and batch data pipelines (ETL)
• Integration, cleaning, and profiling of structured, semi-structured, and unstructured data
• Deployment of distributed Big Data infrastructures
• Design, development and management data-driven products such as an AutoML Platform and a mobile-based Visualization Engine

I have graduated with an MS in Business Analytics from the McCombs School of Business, The University of Texas at Austin.

My technical competencies include:
• Programming Languages & Technologies: Python (scikit-learn, TensorFlow, Keras), R, Java, SAS, Spark (PySpark), Hive, Sqoop, Kafka, Beam
• ML/DL: Linear, Logistic Regression, Decision Trees, Random Forest, Bagging, Boosting, Clustering, PCA, MLP, CNN, RNN
• Visualization: Tableau, PowerBI, SAS VA
• Platforms: Cloudera Data Hub, Google Cloud Platform, Amazon Web Services, Docker, Kubernetes
• Databases: MySQL, PostgreSQL, MS SQL Server, Oracle DB

• Operation Systems: Linux (Ubuntu, CentOS, RHEL), Windows, Macintosh

I like enhancing and validating my existing skillset through MOOCs and have completed:
• Machine Learning by University of Washington
• Data Engineering with Google Cloud
• Machine Learning with TensorFlow by Google Cloud
• Neural Network and Deep Learning by deeplearning.ai

If you love to talk about data and technology just hit me up at tushar.gupta@utexas.edu or drop a message saying hi.

_____

Experience

WITHIN
Data Scientist
August 2020 - Present (1 month)
New York, United States

NetApp
Data Science Intern
January 2020 - May 2020 (5 months)
Austin, Texas, United States

- Predicted and prioritized top leads for cloud migration, helping streamline efforts of the sales team, achieving a lift of 1.7
- Built descriptive profiles of the prospective leads to highlight distinctive attributes, enabling the marketing team to develop customized creatives and perform targetted campaigning

TransOrg Analytics
4 years

Data Science Lead
April 2018 - May 2019 (1 year 2 months)
Gurugram, India

For a leading e-commerce company
- Managed a 3-member team to develop a CLTV model for ~20MM customers using PySpark for a leading Indian e-commerce company, improving repurchase rate by 9% and customer churn by 3% using personalized upsell/ cross-sell recommendations

- Technology Stack: Python, SQL

For a global luxury hotel chain
- Assisted the development of customer 360 to personalize post-purchase enrollment offers, status tiers upgrades for the loyalty program of $8B Indian luxury hotel chain, increasing enrollment rate by 25% and average customer spend by $200
- Designed Python based data pipeline to streamline data for downstream consumption, created interactive dashboards for self-service and enterprise analytics eliminating the need for standard excel reporting, reducing man-hours by 40%
- Performed sentiment analysis on guest feedback data across all hotels, identified primary issues, and recommended solutions, improving Net Promoter Score from 7.8 to 8.6
- Technology Stack: Python, SQL, PowerBI

Data Science Specialist
April 2016 - March 2018 (2 years)
Gurugram, India

AI solution for Expense Report Validation for a global financial services company
- Extracted merchant and transaction-related information from digital invoices using OCR detection and regular expression techniques
- Analyzed Google reviews of merchants to discover the true identity of merchants and flag illegitimate ones
- Detected unauthorized items in the invoices to invalidate the transactions
- Productized the solutions so as to integrate with other tools and solutions
- Technology Stack: Python, SQL, Google Vision API, Google Translate API, Google Places API

Big Data Architecture Strategy for US-based State Fund
- Formulated data integration, governance, persistence, and visualization strategy
- Developed application security, user roles, and access control grid to track server activity
- Documented the complete architecture deployment strategy
- Technology Stack: Cloudera Data Hub, Informatica Big Data Management, Tableau

Product Development

- Led the development of the cloud-based AutoML product to automate feature engineering, model training, and model selection
- Designed and deployed framework for on-the-fly visualization engine on Android
- Planned product go-to-market strategies with senior leadership
- Assisted sales and marketing team with collaterals and sales pitches
- Technology Stack: AWS, Spark, Python, Django Framework, HTML, CSS, SQL, Java, XML

Data Analyst
June 2015 - March 2016 (10 months)
Gurugram, India

For a leading Indian Payments Bank
- Architected and deployed big data clusters for ingestion, integration, profiling, and transformation for large volumes of structured and unstructured data
- Created data pipelines for synchronized data ingestion, both in batches and real-time, from multiple data sources
- Optimized number of collection agents for micro-lending business using a combination of Dijkstra and Traveling Salesman algorithms resulting in a 12% reduction in man-hours
- Built complex logics to compute business metrics and created SAS dashboards to help business analyze data visually and take action
- Technology Stack: Cloudera Data Hub, Spark, Flume, Python, SAS

For a Telecom Client
- Built a churn prediction models using random forest technique, achieved up to 85% recall in top three probability deciles for telecom client
- Technology Stack: R, SQL

Wipro Infotech
Computer Networking Intern
July 2014 - August 2014 (2 months)
Gurgaon, Haryana, India

- Assisted team in installation of Zimbra Mail Solution and Messaging System intending to centralize server system for client
- Supported network engineers with implementation of PAM at client location
- Handled installation of new servers at client data centers

Wipro Ltd. - India
Intern
July 2013 - August 2013 (2 months)

Gurgaon, India

_____

Education

The University of Texas at Austin - Red McCombs School of Business
Master of Science - MS, Business Analytics · (2019 - 2020)

Guru Gobind Singh Indraprastha University
Bachelor's Degree, Information Technology · (2011 - 2015)

Apeejay School
 · (2000 - 2011)