

# NISHANT GUPTA

ONE PAGE SUMMARY

nishantgupta.me | imnishantg.com | github.com/imnishantg

## WORK EXPERIENCE

### Data Science Manager | ZS Associates Advanced Data Science (New Delhi) Track Lead

January 2020 - Present

Leading data science projects, products development, and work streams under newly formed Non-Healthcare (NHC) Analytics Space

- Working as individual contributor (IC role) in **Artificial Intelligence Center of Excellence (AI CoE)**, focused on medical imaging asset and expertise development
- Leading the development of AI-guided product innovation (product discovery) process for an American food & beverage company
- Leading design and development of models for omni-channel strategy for the customers (HCPs) of a pharmaceutical MNC
- Led 8-member team to design, develop, and deploy multi-model ecosystem under 'AI-Guided Selling' initiative through end-to-end automated model generation process (AutoFeat + AutoML) to enable B2B inside sales teams
- Developed AI Ops assets and expertise in-house to automate ML system lifecycle management and deployment

### Data Science Consultant | ZS Associates

January 2018 – December 2019 [24 Months]

Led 12+ member team to design and develop a portfolio of ML/AI products and systems for a Big-5 technology company

Previously led a 6-member team to create AI products in real-world evidence (RWE) & clinical trial optimization (CTO) domain under R&D Excellence practice area.

- Led 10+ member team to develop and enhance ML-model ecosystem, to identify quality leads from ~7M leads across 100+ countries prioritize ~700K leads across sales channels, resulting in \$70+M of increased revenue for client
  - Won Project Champion (Innovation of the Year) – 2019**
- Achieved 66% AUPRC in patient event prediction to identify 25,000+ naïve patients/year exposing opportunity worth \$150 MM/year
- Achieved 87+% accuracy (over client's 11%) through AI-enabled automated schema matching process to standardize clinical data to save \$8MM of direct cost
  - Won Project Champion (Innovation of the Year) – 2018 and showcased in SCDM 2018 Annual Conference**

### Data Science Associate Consultant | ZS Associates

January 2016 – December 2017 [24 Month]

Among the first (of ~15 in US and ~5 in India) batch that was laterally promoted to Advanced Data Science Track after a company-wide shortlisting process. Completed 5-month long Advanced Analytics Learning Program offered through ZS-INSOFE collaboration

- Created contracting tool based on patient-level forecast modeling that broke new grounds in forecasting patient level dynamics
- Automated discrepancy detection in clinical trial data with active learning component to achieve 10X reduction in manual review effort
  - The AI-enabled process was showcased in SCOPE Summit 2018**

### Research Analyst | Grail Research

October 2012 – March 2014 [18 Month]

Involved in projects across industry verticals such as Telecom, FMCG, Health, Tourism, Lifestyle, and Pharmaceuticals

## EDUCATION

Bachelor of Engineering  
Electronics and Communication

Netaji Subhas Institute of Technology  
(University Of Delhi)

2008 – 2012

72.0%

## SKILLS

Programming Language

PYTHON, R, MATLAB, BASH PROGRAMMING, SPARK

Deep Learning

PYTORCH, FASTAI, KERAS

Cloud Services

AZURE, GCP

Miscellaneous

SQL, EXCEL, VBA, GIT, LINUX, DOCKER

# NISHANT GUPTA

nishantgupta.me | imnishantg.com | github.com/imnishantg

## WORK EXPERIENCE

---

### Data Science Manager | ZS Associates

January 2020 - Present

#### Advanced Data Science (New Delhi) Track Lead

*Leading 35+ member team of Advanced Data Science (ADS) Track, New Delhi office*

*Leading data science projects, products development, and space management of newly formed **Non-Healthcare (NHC) Analytics** venture*

- Working as individual contributor (IC role) in **Artificial Intelligence Center of Excellence (AI CoE)**
  - **AI CoE** focuses on conducting applied AI research, developing frontier innovations for our clients in high-impact research areas
  - Leading the asset and expertise development for medical imaging based business problems
- Leading 3-member AI team in developing AI-guided product innovation (product discovery) for an American food & beverage company
  - Current process is long, expensive innovation cycle based on empirical experimentation and launch guided by market research
  - Objective is to uncover new product concept/profiles through AI, that would replace the prior process
  - Estimated sales forecast based on the product design and its market positioning is used as the success metric
- Leading 4-member DS team to design and develop models for omni-channel strategy for the customers (HCPs) of a pharmaceutical MNC
  - The objective is to recommend timing and message content for HCPs to maximize for customer impact and experience
  - The models to be designed includes patient event prediction, influence mapping, HCP segmentation, customer (channel) affinity and preference, and sequence optimization
- Leading the conceptualization and productization of ML/AI applications in pricing optimization problem in B2B environment
- Led an 8-member team to design and develop a multi-model ecosystem under 'AI-Guided Selling' initiative, focused on enabling inside sales teams across B2B companies
  - Includes model to determine churn/leakage, lifetime value, upsell/cross-sell/net new sell, new customer, purchase value, relationship management, opportunity conversion optimization, and forecast CHI (customer health index) for a customer
  - Models are developed through end-to-end automated model generation process (AutoFeaturization + AutoML) and supplemented through explanation model framework for faster and low-inertia adoption by end-users
  - The customizable AI Guided selling ecosystem is molded as a cloud-agnostic, hybrid-SaaS product
  - Data Science tech stack: Python, PySpark, Docker-Kubernetes, Cassandra, Kafka, etc.
- Developing AI/ML system DevOps as the in-house capability to automate ML system lifecycle management and deployment

### Data Science Consultant | ZS Associates

January 2018 – December 2019 [24 Months]

*Led 12+ member team to design and develop a portfolio of ML/AI products and systems for a Big-5 technology company*

*Previously led a 6-member team to create AI products in real-world evidence (RWE) & clinical trial optimization (CTO) domain under R&D Excellence practice area.*

- Developed propensity models to advance client's Go To Market (GTM) strategy of their product portfolio targeting 18K+ accounts
  - Achieved AUROC of 78% in predicting accounts' propensity to buy new products based on their current purchase and consumption patterns, enabling account executives to drive targeted sales efforts
- Led a 10+ member team to develop and enhance multiple ML-model ecosystem for a top-5 technology company, focused on optimizing the identification from ~7M leads across 100+ countries to prioritization of selected ~700K leads across sales channels
  - Evolved the ML system from a 2-model to an 8-model ecosystem
  - Achieved ~\$50M growth in annual revenue by integrating purchase value model in the ecosystem
  - Increased revenue by ~\$13M through a model that focuses on the sales team's effort for lead prioritization
  - Evaluated and added features from four new data sources, resulting in ~\$8M increase in annual revenue
  - Achieved AUROC of 84% on new propensity model developed for different sales channel (Partner Channel)
  - Streamlined hypotheses testing, feature engineering, and model experimentation for optimizing model ecosystem
  - Assessed impact of a model configuration change, corrected for model drift and self-fulfilling prophecy in the model ecosystem
  - Liaised with ML engineering team to streamline processes and successfully complete 7+ feature release cycles
  - Tech Stack : Python, Docker, Airflow, Spark, Azure Cloud Services (Databricks notebooks, DataLake, ML Services, DataFactory, Repos)
  - **Won Project Champion (Innovation of the Year) - 2019**

- Developed algorithm for predicting specific medical diagnosis in the next six months using patients' medical history and event data
  - Achieved AUPRC of 66% through 1-D CNN-RNN hybrid model to identify 25,000+ naïve patients/year
  - Exposed opportunity worth \$150 MM/year from highly imbalanced EHR data (~2% positive class)
- Designed AI-enabled automated schema matching process to convert raw clinical trial datasets of disparate schema to standard format
  - Achieved 87+% average accuracy with a runtime of ~2 mins/dataset, improving over 11% accuracy achieved by client's in-house team
  - Techniques used: Shingle-phrase mapping, semantic lookup dictionary, sentence2Vec, WordNet, SVM, XGBoost, neural networks
  - \$8MM of direct cost saved for client; Attracted \$800K worth of project work; **Won Project Champion (Innovation of the Year) - 2018**
  - Results were showcased in SCDM 2018 Annual Conference (World's Top Clinical Data Management Event)

## Data Science Associate Consultant | ZS Associates

January 2016 – December 2017 [24 Month]

*Among the first (of ~15 in US and ~5 in India) batch that was laterally promoted to Advanced Data Science Track after a company-wide shortlisting process*

- Completed 5-month long **Advanced Analytics Learning Program** offered through **ZS-INSOFE** collaboration
- Created over 12 hours of training material on data handling and management, machine learning, and visualization in R for employee learning roadmap and mentored ZS New Delhi employees and trainers
- Conducted over five advanced analytics sessions and competitions for 250+ employees as a key member of *AlgoRhythms*, an internal initiative to raise analytical quotient of employees
- Developed and facilitated case studies and hackathons with data science recruitment team for 700+ participants across premier institutes
- Developed search and optimization process to rebalance inclusion-exclusion criteria to ensure availability of sufficient patient pool for clinical trial enrollment [*Proof of Concept*]
- Automated discrepancy detection process in CDISC SDTM data to bolster CDR team's capabilities at US pharma company
  - R-Shiny based dashboard automates data ingestion, integration, anomaly detection, and report synthesis to achieve more than 10X reduction in CDR teams' effort on clinical trial data review
  - Included active-learning component that imbibes ground truth from CDR's (Clinical Data Reviewer) feedback to optimize performance
  - The process was showcased in SCOPE Summit 2018 (Summit for Clinical Ops Executives) by the ZS leadership
- Created a replicable, customizable disease KPI dashboard template for a Swiss based pharmaceutical company
  - Tech stack used: [*Frontend*] R, Shiny, htmlwidgets libraries, HTML, CSS, JavaScript; [*Backend*] Impala, SQL

## Business Analytics Associate | ZS Associates

April 2014 – December 2015 [21 Month]

*Involved in projects specific to managed care practice area in US healthcare market*

- Created contracting tool based on patient-level forecast modeling that helped US-based pharmaceutical company assess the effect of their contracting strategy on future patient flow
  - Model broke new grounds in forecasting patient level dynamics in specialty therapeutic market and garnered appreciation from ZS principals and client
- Developed portfolio rebating constructs to help UK-based pharmaceutical company improve share and access of their product portfolio in diabetes market while preserving the prices of their latest drugs
- Formulated market access strategy for a US-based pharmaceutical company to optimize the commercial performance of their therapeutic drug

## Research Analyst | Grail Research

October 2012 – March 2014 [18 Month]

*Involved in projects across industry verticals such as Telecom, FMCG, Health, Tourism, Lifestyle, and Pharmaceuticals*

- Studied the brand awareness and appeal of an African country as an international leisure destination for its tourism ministry
  - Led 3-member team in brand tracking study, which includes managing questionnaire fielding, data quality, and output
  - Supervised 26-member vendor's team and managed interviewers' training
- Assessed impact of global TV Ad campaign by an African country's tourism ministry, intended to showcase country as leisure destination
  - Carried out comparative analysis of viewers' impression vis-à-vis existing Ad concept through quantitative survey approach
- Analyzed promotional activities of distribution channels to help US-based lifestyle company optimize their brand positioning
  - Streamlined the tracking process of email promotional activities from 30+ distribution channels across geographies
- First Runner-Up in Training Case Study Competition: Evolution of Smart Grid Technologies in Developing Countries

## SKILLS

Programming Language	PYTHON   R   MATLAB   BASH PROGRAMMING
Deep Learning	PYTORCH   FASTAI   KERAS
Big Data Technologies	PYSPARK
Cloud Services	AZURE   GCP
Front-End Interface	RSHINY
Miscellaneous	SQL   EXCEL   VBA   GIT   LINUX   DOCKER

## EDUCATION

Bachelor of Engineering Electronics and Communication	Netaji Subhas Institute of Technology (University Of Delhi)	2008 – 2012	72.0%
AISSEE (CBSE Board): Class XII	Delhi Public School, R K Puram	2007	83.0% (90.33% in PCM)
AISCE (CBSE Board): Class X	Ryan International School, Delhi	2005	91.4% (1st Rank in Class)

## INDEPENDENT PROJECTS

### [B.Tech Project] Simulation of Indoor OWC System

- Designed and simulated a fully customizable, high speed optical wireless communication system, which uses lighting sources as communication devices, in MATLAB
  - Simulation demonstrated the system viability and advantage of LED's simultaneous use for lighting & communication purpose

## INDEPENDENT COURSEWORK

Deep Learning	Neural Networks and Deep Learning, Improving Deep Neural Networks, Structuring Machine Learning Projects, Convolutional Neural Networks, Sequence Models	Deeplearning.ai	16 weeks
	CS231n   Convolutional Neural Networks for Visual Recognition	Stanford	16 lectures
	FastAI Course 1 (2019)	USF	7 lectures
Data Science	CS109: Introduction to Data Science	Harvard University	15 weeks
	Introduction to Data Science	University of Washington	8 weeks
	The Data Scientist's Toolbox, Getting and Cleaning Data, R Programming, Exploratory Data Analysis, Reproducible Research, Statistical Inference, Regression Models, Practical Machine Learning, and Developing Data Products	John Hopkins University	36 weeks
	The Analytics Edge	MIT	12 weeks
	Introduction to Big Data with Apache Spark	UC Berkeley	4 weeks
	Scalable Machine Learning	UC Berkeley	4 weeks
Machine Learning	Machine Learning	Stanford	10 weeks
	Statistical Learning	Stanford	10 weeks
Data Mining	Pattern Discovery in Data Mining	UIUC	4 weeks
	Mining Massive Datasets	Stanford	10 weeks
Networks	Networks, Crowds, and Markets	Cornell University	10 weeks
Game Theory	Introduction to Game Theory	Yale University (OYC)	24 lectures
	Game Theory	Stanford	7 weeks
	Game Theory II: Advanced Applications	Stanford and UBC	6 weeks
Statistics	'Descriptive', 'Probability', and 'Inferential' statistics	UC Berkeley	15 weeks
	Statistics 110	Harvard University	35 lectures