Data Scientist

Name: Sangeeta Nahar
Current Location: Bengaluru, INDIA
Mobile: +919582817667, +919560269508
Email: naharsangeeta24@gmail.com
LinkedIn

Summary

Experienced in Data Science with a Master's degree in Mathematics and over 2 years of relevant industry experience. Proficient in predictive modeling, data processing, and natural language processing. Expertise in machine learning, statistics and Tableau for impactful data visualization.

Professional Experience

- **Brainalyst, Gurgaon, INDIA**
  Senior Data Scientist, July'23 – Present
  Experienced Senior Data Scientist in Ed Tech, Driving transformative data initiativets for enriched Learning. Expertise in content development and training, delivering impactful projects..

- **ArchLynk, Bengaluru, INDIA**
  Applied Matrials(Client): Data Scientist, July'21 – June'23
  Data Analysis, Predictive Modeling, Natural Language Processing, Tableau Data Visualization

- **University of Delhi, North-South Campus**
  Guest Faculty of Support Vector Machine, Jan'20 – May'21
  Taught courses on SVM and Fluid Dynamics, enhancing understanding of Advanced Machine Learning Algorithms

- **Mind Shaper Technologies Pvt. Ltd, Gurugram, INDIA**
  Associate Mathematics Content Developer, Robotics Trainer, June'18-Jan'20
  Mathematics and Statistics - related content, used C programming, and gained problem-solving skills

Project Description

Brainalyst Pvt. Ltd | Jul'23 - Present
Ed – Tech | Senior Data Scientist

Adaptive Learning Platform Enhancement

- Implemented machine learning algorithms for an adaptive learning platform.
- Developed curriculum mapping strategies to enhance personalized learning experiences.
- Engineered a user analytics dashboard for real-time insights into student performance.
- Conducted A/B testing for iterative improvements and user engagement enhancements.
- Successfully deployed and evaluated the adaptive learning system, demonstrating improved learning outcomes.
- Created and implemented a user analytics dashboard for real-time monitoring and optimization of the adaptive learning platform.

ArchLynk, Bengaluru, INDIA
IT Services and IT Consulting
Client - Applied Materials | Data Scientist

DDLT (Demand During Lead Time) Forecasting | Jul'21 – Jun'23

- Led the development and implementation of a periodic model refresh using linear regression with quadratic terms, resulting in improved accuracy of revenue forecasts.
- Analyzed cross-correlation of capacity data with WFE revenue to identify leads and lags with strong correlation, enabling more precise forecasting.
- Generated over 5500 models with 15 independent variables and applied filtering criteria to select the best-performing models, improving Adjusted R-squared values.
- Conducted outlier analysis and smoothing techniques to optimize the model's performance and ensure accurate predictions.
- Utilized @RISK Palisade for risk analysis and implemented Monte Carlo simulation to assess the likelihood of different outcomes, enhancing decision-making processes.

NSR(New Service Requests) Risk Classification | Sep'21 – Dec'21

- Mitigate risks to timely completion of systems ac cording to manufacturing timelines caused by NSRs.
- Previously, over 100 hours per week were spent on identifying high-risk items and delays caused by expedites and schedule further extended resolution time of high-risk NSRs by 10x.
- Implemented an automated NSR Risk classification model to reduce manual effort.
- Model executes every 6 hours, syncing with Tableau dashboard.
- Led team in harmonizing the AP (Automated Process) for sales entities.

Multiskilled Workforce Allocation | Jul'21 – Sep'21

- Implemented a greedy approach to allocate systems based on headcount per platform, ensuring 100% coverage and eliminating orphaned systems awaiting installation.
- Considered constraints such as qualified iCE allocation, chamber details, and availability, while prioritizing local iCE allocation.
- Decision variables included iCE capability, location, and availability, optimizing system allocation.
- Overcame challenges including cost minimization, inclusion of unqualified iCEs, data integration from multiple resources, and accounting for quarantine details and iCE churn.
- Utilized Python and libraries such as Pandas and NumPy for automation, achieving a solve duration of approximately 10 minutes.
-

Obsolescence and Commit Dashboard | Dec'21- Jun'23

- Managed obsolescence and diminishing manufacturing sources and material shortages(DMSMS) to maintain system availability and minimize Life-Cycle Costing(LCC).
- Developed an obsolescence management strategy, tracked component life cycles, and identified alternatives.
- Utilized Tableau for data analysis, visualization, and decision-making support.

Real Time Materials Dashboard | Dec'23- Jun'23

- Streamlined prototype part delivery process, enhancing speed, and quality.
- Strengthened supplier partnerships through active engagement in the Real Time Materials (RTM) program.
- Improved delivery speed and quality, accelerating innovation cycles.
- Optimized data analysis and visualization techniques, facilitating informed decision-making.

Master's Thesis, Indian Institute of Technology, Bhubaneswar
Classification of Data of Imbalance Problem and efficient classification under ROC Curve

- Explored various machine learning and statistical modeling techniques including supervised learning, decision trees, naïve Bayes, and support vector machines (SVMs).
- Demonstrated the effectiveness of these techniques in analyzing data and making accurate predictions.
- Contributed to the field of data analysis by addressing the challenges posed by imbalanced datasets and improving classification performance using the ROC curve.

Consultant profile
Data Scientist
Name: Sangeeta Nahar

Technical Skills

- **Programming**: Python, SQL, Matlab, Basic C, Basic R, Latex
- **Software Tools**: Risk Palisade, Tableau, Power Bi, Excel
- **Machine Learning Algorithms**: Liner regression with L1 regularization, Logistic regression with L1 regularization, Bossted trees, Neural networks, Random Forest, GLM Models, Naïve Baiyes, K-means Clustering, KNN classification, Support Vector Machine (SVM) , Gradient Boosting, Ensemble Methods etc.

Education

- **Postgraduate studies: 2015-2017 | Indian Institute of Technology (IIT), Bhubaneswar| India**
  Master of Science, Mathematics
- **Undergraduate studies: 2011-2014 | Sri Venkateshwara College, University of Delhi | India**
  Bachelor of Science, Mathematics Honors

Professional Certification(s)

- Data Mining, IIT Kharagpur, 15 Feb - 9 April' 21: National Programme on Technology Enhanced Learning (NPTEL)
- Data Science Math Skills: Coursera Python for Data Science, IIT Madras,18 Jan- 12 Feb' 21: NPTEL
- Time Series Analysis and Forecasting using Python: Udemy
- Introduction to @RISK. Monte Carlo Simulation addin on Excel:Udemy
- Tableau Essential Training: LinkedIn
- Python + SQL + Tableau: Integrating Python, SQL, and Tableau, LinkedIn
- Advanced SQL for Data Science: Time Series, LinkedIn
- Algorithmic Trading and Stocks Essential Training, LinkedIn

Achievements and Appreciations

- **Appreciation for OBS Dashboard Support:** Applied Materials
- **Qualified Joint Admission Test for M.Sc.(JAM)**: Feb 2015
- **Qualified Joint CSIR-UGC Junior Research Fellowship and Eligibility for Lectureship (JRF - NET)**: June 2018
  Wrote highly appreciated thesis: March 2017, IIT Bhubaneswar

Languages

English – Absolute fluency | **Hindi** – Native language