

# VANDANA SINHA

Hillsboro Oregon  
(971)371-8926

linkedin.com/in/vandana-sinha  
vandanasinha7@gmail.com

## CAREER SUMMARY

---

Accomplished data professional with 6+ years of progressive experience, seeking a Senior Data Scientist role. Excelled in architecting AI solutions and advanced data models, significantly improving performance metrics across various industries. Proficient in Python, TensorFlow, and Scikit-learn, adept in federated learning, bias mitigation, and predictive analytics.

## Areas of Expertise

Data-Driven Decision Making – Advanced AI/ML Techniques – Diagnostic, Predictive, and Prescriptive Analytics – Statistical and Machine Learning Methods – Data Wrangling and Programming – Algorithm Development – Data Quality Improvement & ETL Tool Integration – Cross-Functional Collaboration – Mentorship and Team Development – Deep Learning and Neural Networks – Natural Language Processing – Computer Vision – Big Data Analytics – Time Series Forecasting – Marketing Analytics – Financial Modeling – Database Management – Cloud Computing – Data Visualization and Reporting – Agile Methodologies – Software Development – Quantitative Analysis – Digital Marketing Optimization

## PROFESSIONAL EXPERIENCE

---

### Intel Corporation Folsom – California, United States

May 2022 - Dec 2022

#### *Solution Architect*

- Drove a 3% improvement in customer satisfaction and a **5% increase in solution performance** through advanced data modeling (linear regression, decision trees) and Python libraries (Scikit-learn, NumPy), revealing deep insights into customer behavior.
- Spearheaded integration of cutting-edge AI technologies like deep learning (CNNs, RNNs) and transformers (BERT, GPT) into hardware platforms, boosting Xeon product loyalty and customer engagement, and aligning cross-functional strategies via TensorFlow and PyTorch frameworks.
- Orchestrated collaborative problem-solving across teams by applying data-driven decision-making techniques (hypothesis testing with SciPy, A/B testing with Stats Models) to address challenges with tailored solutions, reduce grievances, and elevate solution efficacy.
- Recognized for exceptional problem-solving capabilities by pioneering new processes leveraging object detection (YOLO, Faster R-CNN) and NLP (spaCy, NLTK), **reducing customer complaints by 20%**.

### CRISIL Global Research Mumbai – Maharashtra, India

Sep 2017 - Aug 2021

#### *Senior Data Scientist*

- Engineered advanced data models using Python (Scikit-learn, XGBoost), SQL, and R for transaction monitoring, leveraging ensemble techniques (Random Forests, Gradient Boosting) to detect fraudulent activities, **reducing incidents by 12%** within eight months through quantitative analysis.
- Developed data-driven algorithmic trading strategies by applying time series forecasting methods (ARIMA, Prophet) in Stats Models and Neural Prophet, driving profit increase by adapting predictive analytics to market dynamics and influencing business strategy.
- Optimized financial **forecasting accuracy by 8%** through statistical modeling using regression analysis (linear, logistic) in Scikit-learn to identify market trends and risks, enabling strategic planning and adjustments.
- Architected marketing mix models by synthesizing diagnostic insights from exploratory data analysis, predictive modeling (regression, clustering), and prescriptive analytics using operations research for data-driven decision-making and solution shaping.
- **Cleaned 10 GB of unstructured** financial data to eliminate campaigns based on spending patterns with expenditure above 15% and reach less than 10% using Python and Excel (pivot, VLOOKUP).
- Built SQL scripts using BigQuery (GCP) to create views and reports for in-depth analysis, increasing productivity by 18%.
- Created logistic regression, tree-based ensemble models, XGboost, and neural network to forecast churn using customer insights.

### Freelance – Mumbai, Maharashtra, India

Aug 2011 - Aug 2017

#### *Senior Data Warehouse & Digital Marketing Analyst*

- Executed data cleansing and wrangling techniques in Informatica PowerCenter ETL tool, improving data quality by 16% and translating complex datasets into actionable insights for data-driven decision-making.
- Optimized digital marketing campaigns leveraging statistical methods (A/B testing using Python stats models) and machine learning models (logistic regression in Scikit-learn) for strategic targeting and budget allocation, boosting **click-through rates by 15%** and ad efficiency by 20%.
- Engineered automated reporting dashboards and visualizations in BI tools like Tableau, enhancing **data accessibility by 10%** and enabling real-time tracking of critical KPIs.
- Led cross-functional reporting by consolidating diverse stakeholder insights, guiding hypothesis formulation using statistical tests (t-tests, ANOVA), and validating performance optimization strategies through research.

- Mentored and upskilled junior data scientists on data wrangling using Pandas, programming in Python, and deploying machine learning models, increasing project efficiency and data quality metrics.
- Managed a cross-functional team to execute a rebranding campaign, achieving a 50% increase in social media engagement and a 35% growth in online sales within three months.
- Analyzed unstructured data from customer feedback systems using econometrics techniques and presented findings to the executive team; actionable insights led to a 10% increase in customer satisfaction scores over six months.

#### Hewlett Packard – Bangalore, Karnataka, India

Aug 2007 - Jul 2011

##### Data Engineer

- Streamlined real-time issue resolution for order processing by architecting data mapping strategies using Unix shell scripting, Informatica PowerCenter ETL workflows, and Oracle SQL/PLSQL, meeting stringent deadlines.
- Engineered an innovative inventory optimization algorithm by leveraging advanced programming techniques in Java and Python, resulting in a **17% improvement in system performance** over six months.
- Implemented integration of Informatica PowerCenter ETL tools into an enterprise data warehouse, performing data profiling, cleansing, and loading of partner sales and inventory **data within eight weeks**, reducing downtime and boosting efficiency.
- Recognized as a pivotal contributor for spearheading development and automation initiatives through Unix shell scripting for job scheduling and Java for application development, **HP's market share increased by 10%**.

#### EXTRA CURRICULAR

---

#### San Jose State University – California, United States

Jun 2023 - Present

##### Data Science Research Assistant

- Architected high-performance federated learning framework using PySyft and TensorFlow Federated for privacy-preserving medical data analysis, processing 10x more data per second than existing solutions.
- Innovated transformative AI solutions by leveraging diffusion models like Stable Diffusion and DALL-E, resulting increase in generated image quality and a 12% reduction in design iteration cycles.
- Integrated bias mitigation algorithms (Adversarial Debiasing, Calibrated Multiclass Predictions) and trustworthy AI techniques to improve fairness, safety and responsibility while reducing harmful outputs.
- Spearheaded user studies leveraging human-computer interaction principles to evaluate the real-world impact and utility of developed AI solutions.

#### University of North Texas Denton – Texas, United States

Jan 2022 - May 2023

##### Teaching Assistant

- Optimized data-driven education models through regression techniques (linear, logistic) and clustering algorithms (K-Means, DBSCAN) in Scikit-learn, enhancing student satisfaction by 12% and assignment performance by 14%.
- Leveraged statistical analysis using SciPy and Pandas to devise personalized tutoring programs based on student patterns, decreasing failing grades and improving retention.
- Pioneered adoption of gamification using Reinforcement Learning (Q-Learning) and flipped classroom techniques powered by interactive Jupyter notebooks, boosting engagement and exam scores.

#### EDUCATION

---

##### Master of Science, Computer Engineering – GPA 4.0/4.0

Graduation Date – May 2023

University of North Texas Denton – Texas, United States

Coursework: Deep Learning, Big Data, Neural Networks, Natural Language Processing, Computer Algorithm

##### Master of Science, Information Science

Graduation Date – May 2007

Birla Institute of Technology – Mesra Ranchi, Jharkhand, India

Coursework: Artificial Intelligence, Data Structures & Algorithms, Computer networks, MATLAB, DBMS

#### TECHNICAL PROFICIENCIES

---

- **Programming:** Python (Pandas, SciPy, NumPy, Scikit-learn, Keras, Stats Model, Matplotlib, TensorFlow, PyTorch, Flask), Git, R
- **Databases:** Microsoft SQL Server, Oracle 11g, MySQL
- **Machine Learning:** Linear regression, Data analysis, Clustering, SVM, Logistic regression, Random Forest, Predictive modeling, Time series forecasting, Statistical modeling, Hypothesis testing, Data modeling
- **Data Reporting:** SAP BO, Power BI, Tableau, Google Data Studio
- **Tools:** Confluence, Microsoft Office, JIRA, Jupyter, RStudio, Google Cloud Platform (GCP), AWS, Snowflake, Google Analytics, Google BigQuery, Spark, Agile
- **Soft Skills:** Communication, Collaboration, Problem-Solving, Critical Thinking, Adaptability, Leadership, Teamwork