

Rajat Dhawan

Tempe, AZ - 85281 | [LinkedIn](#) | 480-876-7946 | rajatdhawan96@gmail.com

WORK EXPERIENCE

On Semiconductors

Scottsdale, AZ

Data Analyst

August 2023 – December 2024

- Monitored and improved data quality for sales and supply chain datasets by developing Python-based workflows, collaborating with 12 stakeholders to validate and generate error reports for anomaly detection, resulting in a 5% increase in data accuracy
- Designed and developed comprehensive BI dashboards using Snowflake and Tableau to provide insights on project performance across business divisions during Quarterly Business Reviews, resulting in an 8% improvement in project schedule KPIs
- Programmed complex SQL queries & data pipelines utilizing common table expressions (CTEs), joins, & analytical functions to track metrics, & historical trends, collaborating with data engineers to integrate SQL in Airflow creating daily views for dashboards
- Built a Gen AI chatbot leveraging RAG architecture and large language models like OpenAI GPT 4.0 to efficiently retrieve domain-specific SOPs and historical project information, improving program managers' productivity and decision-making
- Implemented a Random Forest Regressor to predict project cycle time using TF-IDF, one-hot encoding, and ensemble techniques, enabling accurate milestone duration forecasting for new product development across multiple business units

HCL Software

Sunnyvale, CA

Data Science Consultant

December 2022 - May 2023

- Optimized marketing spends by analyzing channel clicks and sales data using SQL and Python, identifying critical trends and recommending a 23% budget reallocation from low-ROI to high-ROI channels
- Designed and developed interactive dashboards in Tableau to track key metrics such as ROI, CAC, and CPC across marketing channels, enabling stakeholders to make data-driven decisions on budget allocation and performance optimization
- Designed a Customer Identity Resolution method by building a Natural Language Processing pipeline in python utilizing Fuzzy Matching and Cosine Similarity techniques to cluster customer's PII data across 4 digital channels, achieving an accuracy of 88%

Suzuki Motor Corporation

Delhi, IN

Data Scientist

July 2018 – July 2022

- Led \$1M project to develop automated reporting of key performance metrics for EV & Hybrid telematics data by building analytic frameworks (ETL) in Python & MATLAB, which enabled team to manage, clean and process big data efficiently
- Improved product metric by 8% and resolved system-level issues by analyzing vehicle performance trends across different traffic patterns and drive modes through weekly and monthly technical Power BI reports to support global teams in hypothesis testing
- Reduced performance report generation time by 10% by developing a user-friendly graphical user interface (UI) in Python & MATLAB, to enable teams to choose projects, analysis task & dates. This led to streamlined processes & increased efficiency
- Achieved 23% reduction in data processing time by migrating analysis workflow to open-source technologies like Python, bringing down cost of analysis by \$10K annually

PROJECTS & INTERESTS

Retail Product Recognition using Computer Vision

March 2023 – May 2023

- Successfully developed convolutional neural network (CNN) DL pipelines using pre-trained models like ResNet50, InceptionV3 and VGG16 to accurately identify ten different retail products, achieving an overall accuracy of 93% on the test set
- Designed a Stream lit interface (UI), utilizing DL model to generate bill for detected items, enhancing customer experience

Fraud Detection using Predictive Analytics

September 2022 – October 2022

- Performed predictive modeling using Decision Trees and Random Forest classifiers for detecting fraud in Auto Insurance claims, utilizing hyper-parameter tuning methods such as Random Search and Grid Search to optimize machine learning models
- Enhanced training data quality by One Hot Encoding categorical features & handling imbalanced classes (90:10) with SMOTE

Predicting Homesite Insurance Quotes

September 2022 – October 2022

- Predicted the probability that a customer would buy a quoted insurance plan, using different classification methods in Python
- Built an ensemble prediction (one-layer-stacking) model, using Decision Tree, Random Forest, Support Vector Machines, Multi-Layer Perceptron and K-Nearest Neighbors classifiers accomplishing 90% + accuracy

EDUCATION

Arizona State University

Tempe, AZ

MS in Business Analytics (GPA: 4.0) - Distinction

August 2022 – May 2023

Coursework : Database Management, Predictive & Statistical Analysis, Data Mining, Text Mining, Generative LLMs, Deep Learning

Manipal University

Jaipur, IN

Bachelor of Technology Electronics & Communication Engineering (CGPA: 7.69)

July 2014 – June 2018

SKILLS & INTERESTS

Languages : SQL, Python (NumPy, Pandas, Scikit-learn, Plotly, NLTK, SpaCy, Tensorflow, Keras), MATLAB,

Tools : Advanced Excel, PowerPoint, Snowflake, Tableau, Power BI, AWS, Azure ML Studio, QlikView, MS Office Suite, JIRA

Technical Skills : Data Modeling, Data Warehouse, Data Visualization, Statistics, Business Intelligence, Machine Learning, ETL

Awards : Quarterly Excellence Award for Euro-EV Project, Beta Gamma Sigma, Lean Six Sigma Green Belt