
PROFESSIONAL EXPERIENCE

Data Analyst

August 2023 – Present

*Onsemi**Scottsdale, AZ*

- Led the development and implementation of automated Python scripts to identify data quality issues across 12 large scale data sources (>10M records), resulting in a 5% improvement in data quality for supply chain systems & reducing manual intervention.
- Designed and developed comprehensive BI dashboards using Snowflake and Tableau to track project schedules for New Product Development projects, resulting in an 8% improvement in project schedule KPIs
- Built a chatbot leveraging RAG architecture and large language models like OpenAI GPT 4.0 to efficiently retrieve domain-specific SOPs and historical project insights, improving program managers' productivity and decision-making
- Implemented a machine learning algorithm achieving 93% accuracy in predicting schedules for new product development using key features impacting project schedules, optimizing planning for projects across 3 business units

Data Science Intern

December 2022 – May 2023

*HCL Software**Sunnyvale, CA*

- Designed a Customer Identity Resolution method by building a Natural Language Processing pipeline utilizing Fuzzy Matching, TFIDF and Cosine Similarity techniques to cluster customers across 4 digital channels, achieving an accuracy of 88%

Data Scientist

July 2018 – July 2022

*Suzuki Motor Corporation**Gurgaon, India*

- 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, regions, and drive modes through weekly & monthly technical reports to support global teams in product development
- Programmed complex SQL queries utilizing multiple joins, nested queries, analytical functions to extract, manipulate and analyze data from databases to track metrics, seasonal trends, and provide actionable insights for product
- 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 assisting team to migrate analysis workflow to open-source technologies like Python, bringing down cost of analysis by \$10K annually

EDUCATION

Arizona State University, W. P. Carey School of Business

August 2022 – May 2023

*Master of Science in Business Analytics (MSBA) / CGPA : 4.0/4.0 (with Distinction)**Tempe, AZ*

Relevant Coursework: Database Management, Predictive & Statistical Analysis, Supply Chain Quality Management, Decision Modeling, Data Mining, Text Mining & Generative LLM, Marketing Analytics, Deep Learning

Manipal University Jaipur

July 2014 – June 2018

*Bachelor of Technology (Electronics & Communication Engineering) / CGPA: 7.69/10**Jaipur, India*

PROJECTS & INTERESTS

Retail Product Recognition using Computer Vision (Python, Tensorflow, Keras)

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.
- Enhanced experience for customer by developing a Streamlit interface (UI), utilizing DL model to generate bill for detected items

Fraud Detection using Predictive Analytics (Python, MS Excel)

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 (Python, MS Excel)

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

SKILLS, CERTIFICATIONS & ACTIVITIES

- **Languages:** SQL (MySQL), 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 at Maruti Suzuki for Europe EV, Beta Gamma Sigma, Lean Six Sigma Green Belt