

Shagun Kala

+1 (669)-292-4656 | San Francisco, CA 94107 | kala.shagun@gmail.com | www.linkedin.com/in/shagunkala
[Portfolio Website](#) | [Medium](#) | [Kaggle](#) | [GitHub](#)

EDUCATION

University of San Francisco

Master of Science (MS) in Data Science

Malaviya National Institute of Technology (MNIT Jaipur)

Bachelor of Technology

July 2023 - June 2024

San Francisco, CA

July 2015 - May 2019

Rajasthan, India

SKILLS

Programming/ Visualization: Python, R, SQL, NoSQL, HTML, Linux, Tableau, PowerBI, Looker, Excel, PowerPoint, Airflow, Cron
Machine Learning: Regression, Classification, Unsupervised, Deep Learning, NLP, Transformers, Large Language Models (LLM)
Big Data/Database: PySpark, SparkSQL, ETL, Spark, MongoDB, Flask, Hadoop, Git, SVN, Hive, Scala, Snowflake
Cloud/ MLOps: Azure Synapse, AWS (EC2, EMR, S3), GCP, Databricks, Docker, Kubernetes
Libraries and framework: Pytorch, TensorFlow, NumPy, Pandas, Plotly, Scikit-Learn, Matplotlib, Seaborn, Spacy, Scipy, Selenium
Web Experimentation: Causal Analysis, A/B Testing, Hypothesis testing, Statistics, Quantitative Analysis, Experimental Design
Marketing Analytics: Predictive Modeling, Cohort Analyses, Funnel Optimization, Multi-touch Attribution, Market Mix Modeling

BUSINESS EXPERIENCE

Upwork | Marketing Data Scientist | San Francisco, USA

Oct 2023 – Present

Client Churn Prediction

- Analysed 200 million rows large dataset using Apache Spark for extracting insights from payments, client and freelancer interactions
- Improved the client churn definition to encompass 90% of diverse client behaviours and reactivations through exploratory data analysis
- Developed a ML framework by training Random Survival Forest and Decision Tree classifier for predicting non-subscription client churn
- Conducted A/B testing and improved 8% client retention by recommending urgency KPI and incentives for tailored marketing campaigns

Accenture | Data Scientist | Gurugram, India

May 2021 – June 2023

Hewlett Packard: Go-To-Market and Sales Motion Strategy

- Designed Cross-sell and Upsell sales strategy for HP's cloud product utilizing ML, yielding a 13% growth in customer engagement
- Proposed tailored sales strategies for diverse segments by training a Gradient Boosting model with 80% Recall on B2B Installed base data
- Achieved 7% reduction in customer acquisition costs through strategic sales optimization while enhancing customer lifetime value metric

Vodafone: Customer Acquisition and Cross-Sell Product Research

- Built a Machine Learning based propensity scoring system to generate leads, improving broadband customer acquisition rate by 150%
- Engineered 50+ features and trained Random Forest regression model with Log Loss metric, reducing customer acquisition cost by 8%
- Conducted in-depth cross-sell opportunity analysis for Mobius product leveraging 400+ sales data attributes for effective lead generation
- Collaborated with client's cross-functional team on defining KPIs, resulting in an in-depth solution revolutionizing the cross-sell strategy

Accenture: Responsible AI (RAI)- Ethics and Governance

- Led the Responsible AI initiative by implementing tools and techniques to promote ethical, transparent and trustworthy development of AI
- Developed RAI toolkit with fairness, robustness, and explainability tests, employing statistical methods to reduce algorithmic bias by 25%
- Assessed compliance of 5000+ AI projects over 7 RAI principles by creating 4-step algorithmic assessment methodology based on EU AI Act

Capgemini | Senior Analyst | Pune, India

July 2019 – April 2021

eBay: E-Commerce Seller Segmentation Analysis

- Created a seller performance analytical framework using K-means clustering, assigned the seller health scores based on 20+ business KPIs
- Implemented dashboard using a decision tree-based next best action recommender, resulting in 4% increase in seller revenue in 6 months

Airbus: Image Pattern Detection

- Developed a scalable image pattern-matching algorithm using OpenCV for aircraft sectional drawings, saving 200+ hours in fault detection

Peugeot: Python script optimization

- Engineered a HPC solver program using python and modular approach, significantly enhancing runtime performance from 32 to 7 seconds

DEPLOYMENT PROJECTS

Newsgroup BERT Classifier – End to End production deployment using MLOps tools – [Medium Blog](#) | [GitHub Repository](#)

- Fine-tuned a BERT transformer model to classify news articles across 20 categories, deployed the application using Flask
- Built an automated CI/CD pipeline to deploy the commits on the AWS EC2 cloud instance using docker and GitHub Actions

Stock Market Forecasting using News Sentiments – Detailed Time Series Analysis – [Medium Blog](#) | [GitHub Repository](#)

- Forecasted the Nifty50 stock index prices with LSTM neural network and Facebook Prophet model using Twitter News Sentiment Polarity
- Conducted anomaly detection and model quantization, reducing prediction pipeline runtime by 90% for production deployment

ACHIEVEMENTS & AWARDS

- Received 'Wizard@Work' STAR award in 2022 for outstanding contribution in the Customer Acquisition Project at Accenture
- 4X Certified in Azure Cloud Platform; Microsoft Azure Data Scientist Associate
- 2X Kaggle Expert; Active data science blogger on Medium having 30k+ viewers