

Shagun Kala

San Francisco, CA | kala.shagun@gmail.com | +1 (669)-292-4656
[Portfolio Website](#) | [LinkedIn](#) | [Medium](#) | [Kaggle](#) | [GitHub](#)

EDUCATION

University of San Francisco
Master of Science in Data Science

July 2023 - June 2024 (Expected)
San Francisco, CA

Malaviya National Institute of Technology (MNIT Jaipur)
Bachelor of Technology

July 2015 - May 2019
Rajasthan, India

SKILLS

- Python, SQL, Machine Learning, Deep Learning, NLP, Transformers, Large Language Models (LLM), Ensemble Models, Tableau, PowerBI
- Apache PySpark, SparkSQL, ETL, SparkML, MongoDB, Databricks, Azure, AWS (EC2, EMR, S3), Databricks, Flask, Docker, Hadoop
- Pytorch, TensorFlow, NumPy, Pandas, Scikit-Learn, Matplotlib, Seaborn, SciPy, Git, SVN, Selenium, A/B Testing, Statistical Modelling

BUSINESS EXPERIENCE

Upwork | Marketing Data Science Intern | San Francisco, USA

Oct 2023 – Present

Client Churn Prediction

- Utilized Apache Spark on Databricks for in-depth analysis of massive datasets, extracting insights from payments, client interactions, and freelancer activities (~40M rows each).
- Improved client churn definitions, enhancing accuracy for nuanced behaviors and reactivations.
- Developed a robust framework using Survival Analysis and Machine Learning to predict non-subscription client churn.
- Equipped the marketing team with urgency and incentive metrics, enabling targeted campaigns.

Accenture | Data Science Analyst | Gurugram, India

May 2021 – June 2023

Accenture's Responsible AI Toolkit and RAI Assessment

- Contributed to the development of Accenture's Responsible AI toolkit to build ethical AI solutions.
- Implemented fairness tests and added explainability packages to enhance transparency and fairness in AI models.
- Developed RAI algorithmic assessment methodology, leveraging EU Artificial Intelligence Act.
- Screened 5000+ AI projects for RAI compliance across 7 RAI principles.

Go-To-Market Strategy: Sales Motion Recommendation

- Recommended tailored sales approaches for different customer segments using client's Installed Base data (~1M rows).
- Employed a data-driven approach to refine sales strategy utilizing a Random Forest model.
- Achieved a significant 13% boost in customer engagement and overall business growth.

Customer Acquisition and Cross-Sell Product Research

- Conducted in-depth cross-sell opportunity analysis on client's sales data (200+ features).
- Implemented an intelligent propensity scoring system to recommend broadband customer leads using a regression model.
- Improved client's customer acquisition rate by 150%.

Capgemini | Senior Analyst | Pune, India

July 2019 – April 2021

Seller Segmentation Analysis

- Introduced seller health evaluation framework, employing K-means clustering for segmentation based on KPIs.
- Created a next best action recommender using regression models.
- Resulted in personalized recommendations and enhanced seller performance.

Image Pattern Detection

- Developed an image template matching algorithm on aircraft sectional drawings, saving 200+ man hours for the client.

Python Scripting

- Created a HPC solver program to optimize modular approach, improving runtime performance from 32 to 7 seconds.

ACADEMIC PROJECTS

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

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

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

- Forecasted the Nifty50 stock index prices with LSTM and Facebook Prophet model using Twitter News Sentiment Polarity.
- Performed model quantization by converting Float32 weights to Float16, reducing the prediction time by 1/10th.

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.