Sourav Roy

Jersey City, NJ | (646)229-1562 | souravroy.official93@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Data Science professional with over 7 years of extensive experience in data analytics, machine learning and A/B testing, actively driving significant business impact, identifying data-driven opportunities, and mentoring junior analysts. Demonstrated ability to communicate, collaborate cross-functionally and utilize management skills to deliver solutions in a fast-paced environment.

PROFESSIONAL EXPERIENCE

Staff Analyst, Product Specialist

EXL Services   Jul 2022 - Present

- **Revamped fraud detection model** using PySpark for Zelle and Check Deposits by introducing **15+ innovative features**, resulting in **reduction** of over **$4M in annual fraud losses** and **customer friction**.
- Partnered with cross-functional stakeholders, including product and engineering teams, to **integrate near real-time data** that enhances analytics and accelerates fraud detection, potentially **reducing future losses** by **over $50M**.
- **Hypothesize and conduct A/B testing** for customer care center call journey using SQL and Python, leading to **increased upsell opportunities by 8%** and **saved** by over **20 hours per day** for call center agents.
- Designed and **optimized ETL pipelines** to enable seamless **end-to-end data flow**, while **single-handedly** developing a **Tableau Dashboard** to **monitor KPI's** related to fraud, reducing reporting time and significantly improving the speed and accuracy of data-driven decision-making.

Lead Analyst, Customer Analytics Team

EXL Services   Dec 2020 - Jul 2021

- Developed and implemented a Python-based **Natural Language Processing (NLP)** system to analyze customer feedback, enabling more **informed product decisions** and reducing manual labor by over 90% through actionable insights from **unstructured data**.
- Utilized **unsupervised learning** in Python to segment customers based on performance across different sales channels, which led to approximately a **5% increase in sales**.

Senior Analyst

Kantar   Apr 2018 - Nov 2020

- **Identified gaps** in user experience on both the website and app for an entertainment company; **hypothesized and executed A/B testing**, resulting in a **8% surge in free trial to paid conversions** and an **5% improvement in subscriber retention**.
- Utilized **causal inference techniques** like **Propensity Score Matching (PSM)** and **Difference-in-Differences (DiD)** to assess product launch and campaign **impact**, uncovering **key behavior drivers** and delivering insights that optimized product strategy and marketing spend by 7%.
- Acted as a **tech-lead for the project** with a team of 2 data scientists, developing a supervised **customer lifetime value (LTV) model** using R and SQL, enabling targeted marketing strategies and increasing average customer LTV by 6% down the line.
- Maintained existing **data pipelines** supporting reports, dashboards, and models while **identifying bottlenecks** and **implementing enhancements**, improving data flow efficiency, accuracy, and scalability to ensure seamless delivery of business insights.

Business Analyst

Kantar   Jun 2016 - Apr 2018

- Performed **deep-dive analysis** on customer engagement patterns with R and Python for the largest US luxury jewelry brand, **identifying key trends** that helped with targeted ads leading to **5%** increase repeat purchases.
- Designed and implemented **A/B testing** to evaluate the effectiveness of various marketing email designs for cross-selling, resulting in **~3% revenue boost**.
- Created various Tableau dashboards with **key performance indicators (KPI's)** to provide stakeholders with insights into **customer engagement** and **product health that enhanced decision-making accuracy**.

SKILLS

- **Programming Languages:** Python, R, SQL
- **Data Manipulation & Visualization:** Pandas, NumPy, Matplotlib, Seaborn, Dplyr, Ggplot2, Tableau, Looker
- **Functional Expertise:** Metric Definition & Tracking, A/B Testing & Design, Causal Inference, ML Model Engineering, Stakeholder Communication
- **Statistics & Machine Learning:** Linear Regression, Logistic Regression, Clustering, Decision Trees, Random Forest, XGBoost, Natural Language Processing, Matrix Factorization, Scikit-learn, Statsmodel
- **Tools & Cloud Platform:** Jupyter Notebook, AWS, Google Cloud Platform, DataBricks, RStudio, GIT

EDUCATION

Master of Science in Business Analytics | University of California, San Diego   2022
Bachelor of Technology in Electrical Engineering | National Institute of Technology, Durgapur   2016

ACADEMIC PROJECTS

Gamification for Masterclass App: Product Enhancement for better Customer Engagement

- Identified an opportunity to increase user engagement and user retention by introducing gamification to Masterclass courses. Figma was used for feature design and the prototype was utilized for user testing. The age bracket of 18-30 responded positively to the gamification concept, expressing interest in increased engagement with the app.

Intuit Quickbooks targeted campaign: Python and Machine Learning Models

- Developed a highly profitable upsell campaign using targeted emails to encourage customers to upgrade to QuickBooks version 3. An ensemble model of Neural Network, Logistic Regression, and XGBoost was the best predictor and fetched a profit of $450K.