OSHEEN GUPTA

+1-669-226-8490 | ogupta@horizon.csueastbay.edu | **LinkedIn** | **GitHub** | **Website**

Education

MS Business Analytics (Data Science) – GPA: 3.80 – California State University, East Bay	Dec 2025
MBA Marketing – GPA : 3.60 – IMT Hyderabad	Mar 2020
BE Civil Engineering – GPA : 3.29 – Medicaps Institute of Science & Technology, Indore	Jun 2016

Experience

Data Science Analyst | Qlikmatrix | Delhi, India   Sept 2022 – Oct 2023

- **A/B testing:** Improved Google and Amazon ads performance through A/B testing on ad assets, which helped in increasing CTR by 32%.
- **Time-series forecasting:** Implemented ARIMA model in Python to forecast ad sales with 90%+ accuracy, supporting weekly optimization of ad campaigns and resource allocation.
- **Statistical Analysis:** Segmented customers by behavior and demographics using K-means clustering in Python to optimize campaigns, improving ROI by 14%.
- **BI Dashboard:** Developed Tableau dashboards to visualize ad performance, enhancing tracking efficiency and supporting data-driven decision-making.

Data Analyst | Amazon | Bangalore, India   May 2020 – Nov 2021

- **Leadership**: Led a team of 20 associates to prevent fraud, by identifying fraud patterns and implemented corrective strategies, reduced weekly chargebacks by an impressive 76%.
- **Predictive Modelling** : Identified key customer details in fraud patterns using logistic regression, enabling associates to take more accurate actions, leading to a 22% increase in fraud detection accuracy.
- **Rule-Based Logic:** Created an in-dashboard decision support tool using rule-based logic to guide fraud investigators through key actions—cut bad debt by 16% and increased operational speed by 10%.
- **Business Intelligence**: Created Tableau dashboards for weekly reports and presented actionable insights to leadership.

Projects

Companio – HackHayward | Winner - Best use of Perplexity/Sonar | [*GitHub*/*Devpost*]

- Created an AI voice journaling system that feels like a real conversation. Summarizes the highlights from user's day using Perplexity and stores them in Google Sheets.

CourtIQ – Women in AI RAG Hackathon | Stanford University | Winner - 2nd place | [*GitHub*/*Article*]

- Built an AI legal assistant with RAG, Milvus, and LLama to simplify legal research. Helps users understand legal precedents and find relevant case information faster.

Crop Variety Prediction| Datathon 1.0 at CSUEB | 2nd Runner-up | [*GitHub*]

- Engineered an end-to-end supervised learning pipeline (Random Forest/XGBoost) for crop variety prediction (80% accuracy), using SQL for data extraction, Tableau for EDA, and Python (Pandas, Scikit-learn) for preprocessing and modeling.

Predictive Modeling on Loan Approvals | CSU Business Analytics Competition | [*GitHub*]

- Developed a machine learning pipeline leveraging Random Forest and LightGBM, achieving 95% recall through threshold optimization and cost-benefit analysis.
- Enhanced loan risk predictions by identifying at-risk applications with high precision.

Data Science Job Market Analysis | [*GitHub*]

- Scraped 1,200+ data points using Selenium, applying text mining for data preparation and predictive modeling.
- Developed regression models (linear regression, decision trees) predicting data science salaries with 97.5% accuracy.

Research

OSCAR Research Scholar | California State University, East Bay | [*Badge*]   Oct 2024- Present

- Executed end-to-end analysis of 250+ AI startups—data curation, EDA, and log-log OLS regression using Python (Pandas, Seaborn)—to uncover VC investment trends, tech adoption (ML, GenAI), and scaling patterns in valuation and funding.

Skills

Technical skills : Python, SQL, R, Tableau, Power BI, Excel, TensorFlow, PyTorch, scikit-learn, NLP, Hadoop, Hive, Pig, Spark**,** AWS (EC2, S3, Redshift), Azure.

Courses : Data Analytics, Data Engineering & Business Intelligence, Big Data Technology, Optimization for Analytics, Machine Learning, Database Management, Statistics, Text Analytics & Generative AI and Time Series Analytics.

Certifications : Intermediate Machine Learning by Kaggle, Certified Six Sigma Green Belt – KPMG