

YASH KANTHARIA

Jersey City, NJ 07306 | ykantharia27@gmail.com | 551-344-8949
yashkantharia.github.io | linkedin.com/in/yashkantharia | github.com/yashkantharia | kaggle.com/yashkantharia

SKILLS

Programming Skills: Python, SQL, Java, JavaScript, C, C++, HTML, CSS | **Cloud:** Amazon Web Services, Google Cloud Platform, Azure
Libraries: Pandas, NumPy, SciPy, Scikit-learn, BigQuery, Matplotlib, Seaborn, Boto3, Statsmodels, TensorFlow, PyTorch, Beautiful Soup, Requests, Selenium, StreamLit, Unsloth, Flask, Keras, Prophet
Theoretical Knowledge: Exploratory Data Analysis, Tableau, Web Scrapping, RESTful APIs, Advanced SQL, NLP, LLM Fine-Tuning, ChatBot, Docker, Kubernetes, NoSQL, S3, EC2, Redis, Lambda, Snowflake, DBT, Times Series Forecasting, Generative AI, Classification, Regression, Clustering, Android, Unity, Git, PostgreSQL, Firebase, JIRA, Splunk, Google AI Studio, TCP, DNS, CDN, FTP, SCP, Webhooks, JSON
Account Management: Solutions Architect, Sales Engineering, Client Facing, Escalation Management, Root Cause Analysis, Upselling, QBRs

WORK EXPERIENCE

- | | | |
|--|---------------|---------------------------|
| Data Engineer Coach , Data Engineer Academy | New York, NY | November 2024 - Present |
| <ul style="list-style-type: none">Coaching clients with topics including Python, SQL, PySpark, Snowflake, DBT, Tableau, AWS, Azure, Data Modeling, Data Analytics, and System Design to support their career aspirations. | | |
| Data Scientist , Filmy AI | New York, NY | August 2024 - Present |
| <ul style="list-style-type: none">Developing a job-to-user matching system using context vectors with Doc2Vec, aiming to improve job search accuracy by 40%.Implementing a Natural Language-to-SQL querying mechanism using NLP and AI to enable querying through simple Natural Language.Fabricated an LLM-powered chatbot using Python, StreamLit, and Gemini API to assist users with career guidance and tailoring resumes.Architected and managed cloud infrastructure on AWS, leveraging RDS, EC2, Lambda, S3, and IAM to deploy a scalable website.Configured custom domains in DNS with SSL certificates to secure websites via HTTPS, ensuring best practices for web security. | | |
| Data Engineer , Fractal Analytics | Mumbai, India | August 2021 – August 2022 |
| <ul style="list-style-type: none">Engineered ETL and ELT data pipelines using Python (NumPy, Pandas, Scikit-learn), PostgreSQL, SCP transfers, REST APIs, and Azure: Logic Apps and Blob Storage, resulting in an 80% reduction in execution time and minimizing data processing costs by 20%.Automated large-scale data ingestion for AI models using Python scripts, AWS, Dremio, and SQL, cutting operational costs by 40%.Orchestrated analytical dashboard to monitor employee performance indicators and key sales metrics for Finance and CPG clients.Formulated scalable data infrastructure solutions on AWS and Azure, reducing redundancy by 30% and increasing efficiency by 20%. | | |
| Senior Customer Success Engineer , CleverTap | Mumbai, India | July 2019 – August 2021 |
| <ul style="list-style-type: none">Drove global customer success and ROI for enterprise client portfolio including market leaders, achieving a 97% CSAT score through campaign design, data-driven analytics, troubleshooting, debugging, and resolving 15,000+ Ad Tech queries.Created a modular Java-based Android application with CleverTap and Firebase SDK to facilitate Demos, Root Cause Analysis, Training, and Testing, with 90% adoption across teams.Collaborated with marketing teams to define campaign strategies, architect engagement flow, and formulate conversion tracking metrics using tools like CleverTap, Google Analytics, and Firebase Analytics, improving their campaign performance by an average of 15%.Analyzed and improved JIRA support tickets by utilizing Latent Dirichlet Allocation (LDA), keyword analysis, word clouds, and heatmaps, driving targeted workflow improvements and achieving a 15% reduction in average issue resolution time.Championed pre-sales by demonstrating Proof of Value and estimated ROI to prospective technical and non-technical stakeholders.Facilitated as a Subject Matter Expert for Data Protection Laws (GDPR and CCPA), REST API integrations, and Mobile Marketing channels including Push Notifications, Email, SMS, Web Push Notifications, Web Popups, Webhooks, CDNs, AWS, GCP, and WhatsApp.Managed and mentored a team of seven, enhancing productivity by 30% through collaboration and knowledge sharing. | | |

EDUCATION

- | | |
|--|---------------|
| Pace University , Seidenberg School of Computer Science and Information Systems | New York, NY |
| Master of Science (MS) in Computer Science GPA: 3.83 | May 2024 |
| University of Mumbai | Mumbai, India |
| Bachelor of Engineering (BEng) in Information Technology | June 2019 |

PROJECTS (More projects available on [GitHub](#))

- | | |
|--|---------------|
| Fine-tuning LLM for English-to-French Translation Python, Sloth, Gemma 2b, PyTorch, Seq2Seq, LORA. | May 2024 |
| <ul style="list-style-type: none">Enhanced English to French sentence translation quality using the Gemma 2b Large Language Model with Unsloth, outperforming a baseline Seq2Seq model (PyTorch). | |
| WebScraper API Python, Selenium, Flask, BeautifulSoup, HTML, DOM, Pandas. | October 2020 |
| <ul style="list-style-type: none">Built a Flask API with Selenium, BeautifulSoup, Gensim, and basic authentication for capturing metadata, icons, contact details, a screenshot of the home page, and a website summary into a CSV file. | |
| Real-time Uninstall Tracking for Android Applications Java, Firebase Analytics, Node.js, REST APIs, Android. | December 2019 |
| <ul style="list-style-type: none">Integrated a Java-based Android app with CleverTap and Firebase Analytics SDK for real-time uninstall tracking, achieving near 100% accuracy and eliminating reliance on unreliable silent push notifications. | |

PUBLICATION

- | | |
|---|-------------|
| Success Prediction using Random Forest, CatBoost, XGBoost, and AdaBoost for Kickstarter Campaigns | August 2019 |
| Link: https://ieeexplore.ieee.org/document/8819828 | |