

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 | yashkantharia.medium.com

SKILLS

Programming Skills: Python, SQL, Java, JavaScript, C, C++, HTML, CSS | **Data Security & Compliance:** IAM, GDPR, CCPA, HIPAA

Data Analytics: Advanced SQL, Tableau, Snowflake, BigQuery, Matplotlib, Seaborn, Google Analytics, Power BI, Excel

Cloud Tools: Amazon Web Services, Google Cloud Platform, Azure, Kubernetes, Docker, Databricks, DBT, Airflow, Firebase

API & Integrations: RESTful APIs, OAuth, JWT, SAML, Webhooks, Event-Driven Architecture, SDK Integration & Debugging, Postman, Swagger

Machine Learning & AI: TensorFlow, PyTorch, NLP, LLM Fine-Tuning, ChatBots, Classification, Regression, Clustering, Time Series Forecasting, GPT, Gemini, Llama, BERT, TFIDF, Word2Vec, Doc2Vec, RAG, GAN, Vertex AI

WORK EXPERIENCE

Customer Success Manager, Data Engineer Academy | B2C | New York, NY

November 2024 - Present

- Onboarding and coaching new clients to the program with a customized learning journey with Data Technologies, including **Python, SQL, PySpark, Snowflake, DBT, Tableau, AWS, Azure, Data Modeling**, and **System Design**, leading to a 97% client success rate.
- Conducted** live **Python and Data Modeling** doubt sessions and projects, enhancing student understanding and data engineering skills.

Data Scientist, Filmy AI | B2C | New York, NY

August 2024 - Present

- Developing a **job-to-user matching system** using **context vectors** with **Doc2Vec**, aiming to improve job search accuracy by 40%.
- Fabricated an **LLM-powered chatbot** using **Python, StreamLit**, and **Gemini API** to assist 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.

Senior Customer Solutions Architect, Fractal Analytics | B2B | Mumbai, India

August 2021 – August 2022

- Engineered end-to-end **ETL and ELT data pipelines** using **Python (NumPy, Pandas, Scikit-learn)**, **PostgreSQL**, **SCP transfers**, **RESTful APIs**, and **Azure** cloud services, resulting in an **80% reduction** in **execution time** and **minimizing data processing costs** by **20%**.
- Optimized SQL procedures** to transform the data by **reducing redundant reads and updates**, and **employing partitions** and **indexing** to reduce the read time overhead, **reducing the time to run the processes by 45%**.
- Gathered client requirements** and **mapped use cases** to **business KPIs** by collaborating with **technical** and **non-technical stakeholders**. Translated insights into actionable tasks and **delivered verified solutions** with **proof of concepts**, reducing project delivery time by 15%.
- Implemented testing breakpoints** and **fail-safes** to validate data integrity and ensure **100% accuracy** and **reliability** of published data.
- Led technical pre-sales efforts**, designing tailored solutions that addressed business needs, demonstrated value through POCs, and secured a **\$120K deal** and a **3-year contract extension** for two clients.
- Automated ETL process** using **Python scripts**, **AWS**, **Azure**, **Dremio**, and **SQL**, reducing manual effort in periodic data update by **90%**.
- Orchestrated **Analytical Dashboards** for performance indicators and key sales metrics for **Finance** and **CPG** clients, boosting platform usage to 97%.

Senior Customer Success Engineer, CleverTap | B2B | Mumbai, India

July 2019 – August 2021

- Led** global customer success initiatives for enterprise clients, achieving a **97% CSAT score** by optimizing campaign strategies, conducting data-driven analytics, and resolving **15,000+ AdTech queries**.
- Developed** an **Android app** with CleverTap SDK and Firebase Analytics, enabling client demos, issue reproduction, feature testing, and internal training, accelerating debugging efficiency. This increased the team's bug reporting efficiency by 30%.
- Mastered presales consulting**, integrating **proof-of-concept solutions**, and **debugging SDK integrations** with **client developers** to ensure seamless implementation. Presented **product demos** to stakeholders, **achieving a 100% deal closure rate**.
- Onboarded 150 clients** with Platform training, **SDK integration**, and **debugged issues** across platforms, including **Android, iOS, React Native, Ionic, Unity**, and **Cordova**, and web technologies like **HTML, CSS, XML, JavaScript**, and **RESTful APIs**.
- Collaborated with SDK teams to develop **custom ad-hoc solutions**, securing a **\$150K deal with a global gaming client** by addressing key integration challenges with Android X and Unity SDK integrations.
- Partnered** with marketing teams to **architect ad campaign engagement flows** and define conversion tracking using **CleverTap, Google Analytics**, and **Firebase Analytics**, boosting campaign performance by 15%.
- Served** as a **Subject Matter Expert** on **GDPR, CCPA**, and mobile marketing channels (Push Notifications, Email, SMS, Web Push, Webhooks, CDNs, AWS, GCP, WhatsApp), guiding clients on compliance and best practices.
- Managed and mentored** a team of seven, improving overall team productivity by **30%** through structured training, collaboration, and knowledge-sharing initiatives.

EDUCATION

Pace University

Master of Science (MS) in Computer Science

New York, NY

May 2024

University of Mumbai

Bachelor of Engineering (BEng) in Information Technology

Mumbai, India

June 2019

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

Fine-tuning LLM for English-to-French Translation | *Python, Sloth, Gemma 2b, PyTorch, Seq2Seq, LORA*

- *Enhanced English to French sentence translation quality using the Gemma 2b Large Language Model with Unsloth, outperforming a baseline Seq2Seq model (PyTorch).*

Data Pipeline Automation | *Python, SQL, Azure Blob Storage, REST APIs*

- *Implemented Python scripts to automate periodic data extraction from a private network's SQL database, executing complex calculations, formatting data into CSV or Parquet files, uploading to Azure Blob Storage, and triggering an HTTP API.*

WebScraper API | *Python, Selenium, Flask, BeautifulSoup, HTML, DOM, Pandas*

- *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*

- *Integrated a Java-based Android app with Firebase cloud functions, REST APIs, CleverTap SDK, and Firebase Analytics SDK for real-time uninstall tracking, achieving nearly 100% accuracy and eliminating reliance on unreliable silent push notifications.*

Vector Representations of Words using Glove and Word2Vec | *Python, Glove, Word2Vec, Pandas, Seaborn, matplotlib*

- *Explored semantic relationships between words using pre-trained Glove and Word2Vec embeddings. Visualized word vectors and performed vector arithmetic to analyze analogies and similarities.*

Cat vs Dog Classification with Transfer Learning | *Python, Pandas, Keras, PyTorch, MobileNetV2, Tensorflow, Seaborn, matplotlib*

- *Developed an image classification model using transfer learning with MobileNetV2, comparing the performance of TensorFlow (Keras) and PyTorch frameworks. Evaluated standard classification metrics and TensorBoard, visualizing key performance insights.*

Power Consumption Prediction and Analysis for New York using FBProphet | *Python, Pandas, Numpy, Scikit-learn, FBProphet, Seaborn*

- *Analyzed time series data of power consumption in New York City boroughs to identify trends and seasonality. Built a forecasting model using FBProphet to predict future power consumption, visualize results, and evaluate model accuracy.*

PUBLICATION

Success Prediction using Random Forest, CatBoost, XGBoost, and AdaBoost for Kickstarter Campaigns

August 2019

Link: <https://ieeexplore.ieee.org/document/8819828>