

Vashisht D. Brahmhatt

Ahmedabad, India • vashishtdb24@gmail.com • +91 9979981224 • [VASHISHT BRAHMBHATT](#) • [vdbonweb](#)

SUMMARY

Data Scientist and Product Engineer with experience in AI driven automation and data centric product development. Delivered a Python based AI content automation pipeline that reduced batch creation time from 6-7 hours to 2 hours, freeing 20 hours per week for the content team. Built an end to end data preparation and modeling pipeline employing statistical validation, feature engineering, and ML models (KNN, Random Forest, Logistic Regression) that enhanced prediction reliability. Combines strong engineering foundation with product insight to create scalable, data focused solutions.

EDUCATION

Indian Institute of Technology Madras

Chennai, India

B.S. in Data Science

Present

- Concentrations: Data Science and Artificial Intelligence

G.L.S University

Ahmedabad, India

B.Tech in Computer Science and Engineering

Present

- Concentrations: System Design and Machine Learning

SKILLS

Programming Languages: Python, Java, C++, C, SQL, JavaScript, HTML/CSS, Linux Scripting

Machine Learning & AI: Machine Learning, Artificial Intelligence, Deep Learning, LLMs, Langchain, RAG Pipelines, Statistical Validation, Feature Engineering, KNN, Random Forest, Logistic Regression, Pandas, Matplotlib, Jupyter Notebooks

Data & Databases: Database Management Systems, Data Structures & Algorithms, Data Pipelines, Supabase, Pinecone

Cloud & DevOps Tools: Google Cloud Services, AWS, Oracle Cloud, Git, PyCharm

Web Development Frameworks: Node.js, React.js, Flask

Core Computer Science & Analytics: System Design, Object Oriented Programming, Objects & Design, Computer Organization & Programming, Modern Application Development, Business Analytics, Statistics

EXPERIENCE

Product Management and Engineering Intern

Jun 2025 – Oct 2025

Clan

New York, United States

- Designed a Python based AI content automation pipeline that reduced end to end content creation time from 6-7 hours to 2 hours per batch, saving 20+ hours per week for the content team.
- Optimized backend workflows and databases to manage 750+ video content pieces, improving delivery reliability and creator onboarding efficiency.
- Managed database systems of 100+ education creators and optimized the platform to deliver 750+ content pieces

Data Science Research Intern

May 2025 – Jun 2025

Prodigy Infotech

Pune, India

- Engineered an end to end Python data pipeline in Jupyter Notebook, leveraging Pandas for preprocessing, statistical validation, and feature engineering; integrated baseline machine learning models (KNN, Random Forest, Logistic Regression) and achieved a 15% lift in prediction accuracy on benchmark datasets.
- Implemented multiple imputation strategies and model interpretability analyses using scikit learn; automated hyperparameter tuning and documented results in Matplotlib visualizations, enabling data driven decision making for business analytics.
- Orchestrated code versioning with Git and applied object oriented programming principles to modularize the modeling framework, reducing code duplication by 25% and facilitating reproducible research.

PROJECTS

Financial Management Platform | [Link](#)

- Built a full stack personal finance platform to track income, expenses, budgets, and savings with data driven insights and real time analytics.
- Designed data pipelines to analyze spending patterns, and generate AI assisted financial insights.

Agency Workflow Automation

- Workflow automation solution with retrieval methods, Langchain, automating lead management for clients and reducing manual onboarding time by 50%
- Developed using Python, Flask, Pinecone and Supabase. Integrated industry leading LLMs for generative processing

ACHIEVEMENTS

- ColrnX - Co-founder: Co founded a social collaborative platform designed to match project creators with contributors, helping users discover real projects, form teams, and collaborate effectively. Received backing from SSIP (Startup Support and Innovation Promotion) under Govt. of Gujarat.

CERTIFICATIONS

- Google Analytics
- Oracle Cloud Associate
- Google Cloud badges
- Programming and Data Science Foundations

PUBLICATIONS

Review Paper

Authors: Author

Published a review paper reviewing data preprocessing pipelines in Machine Learning on ScienceOpen