Pujan Dave

Data Scientist

✉ pujan.dave9@gmail.com ☎ 8401486488 ◉ Ahmedabad 🌐 https://www.pujandave.me

Summary

Technical Project Manager & Lead Data Scientist with **8+ years of experience in Data Science and Machine Learning** and **14+ years in the software industry**, specializing in delivering scalable **AI, Generative AI, and healthcare technology solutions**. Proven expertise in **leading cross-functional teams, managing end-to-end project delivery, and aligning AI initiatives with business objectives** to drive measurable impact in clinical operations, revenue optimization, and customer experience.

Strong background in **model development, training, validation, MLOps, LLM fine-tuning, and multi-agent AI systems**, combined with hands-on experience in **cloud deployment (OCI, AWS, Azure), stakeholder management, agile execution, and technical program leadership**. Adept at translating complex business problems into **production-ready AI products** that improve efficiency, accuracy, and decision-making at scale.

Profiles

Linked In
Linked In

Github
GitHub

Professional Experience

OMNIMD   Ahmedabad
Technical Project Manager   2024/10/12 - Present

- Designed and executed end-to-end **model training pipelines**, including data preprocessing, feature engineering, hyperparameter tuning, and distributed training for scalable healthcare AI solutions.
- Implemented robust **model validation frameworks** using cross-validation, hold-out strategies, and performance benchmarking across multiple clinical and operational datasets.
- Established **evaluation metrics and monitoring protocols** (accuracy, F1, ROC-AUC, hallucination detection for LLMs, and drift monitoring) to ensure production-grade reliability and regulatory alignment.
- Conducted **A/B testing and offline/online experimentation** to compare model variants and continuously improve prediction quality and decision-making performance.
- Automated **retraining and continuous learning workflows** integrated with OCI pipelines to maintain model freshness and adapt to real-time healthcare data streams.

https://omnimd.com/

Inferenz   Ahmedabad
Lead Data Scientist   2024/01/23 - 2024/09/31

- Lead the end-to-end design, development, and deployment of data science products, improving client satisfaction and business performance.
- Manage and mentor a team of 6+ data scientists and engineers, overseeing agile workflows, task assignments, and ensuring project deadlines are met.
- Spearheaded AI and ML solution development aligned with client goals, resulting in 20% improvement in project delivery times and a 15% increase in revenue.
- Drove innovation in AI solutions, implementing scalable machine learning pipelines to support ongoing client needs.

https://www.inferenz.ai

ANALYTICS VIDHYA

Senior Data Scientist   2023/05/01 - 2024/01/20

- Conducted training need analysis and developed curricula for corporate and in-house training.
- Performed post-training evaluations and implemented feedback to improve programs.
- Monitored session progress, prepared reports, and updated Data Science course content.
- Acted as a mentor for Internshala, providing guidance, assessments, and doubt-clearing sessions.

https://www.analyticsvidhya.com/

TRANSORG ANALYTICS

Specialist Data Analyst   2022/03/21 - 2023/05/01

- Led customer analytics projects, including predictive ML models (Lead Scoring, Customer Lifetime Value, Churn), which enhanced BMW's targeted marketing by 35%.
- Developed ML algorithms with over 80% accuracy, integrating models with CRM portals and delivering actionable insights for customer segmentation.

https://transorg.com/

I TRIANGLE TECHNOLABS

Data Scientist   2018/11/11 - 2022/03/21

- Managed a team of data scientists to design, deploy, and enhance NLP and ML models, achieving a 25% improvement in model performance across client platforms.
- Championed end-to-end development for cutting-edge machine learning models, from requirements gathering through deployment, increasing customer retention rates by 20%.

https://itriangletechnolabs.com/

I TRIANGLE TECHNOLABS

Devlopment Team Lead   2013/03/11 - 2018/03/11

- I held full responsibility for both in-house and client-based web development assignments, demonstrating my commitment to excellence in every project.
- By assisting esteemed clients and securing top ratings for the company, I take pride in my contributions to achieving exceptional outcomes.

https://itriangletechnolabs.com/

IMOBDEV TECHNOLOGY

Senior Software Engineer   2012/03/11 - 2013/03/11

- I had the privilege of being part of an exceptional web development team, collaborating closely to design and develop websites for diverse clients of the company.

https://www.imobdevtech.com/

IMOBDEV TECHNOLOGY

Software Engineer   2010/02/11 - 2012/03/11

- As a fresher with a bit of prior knowledge in web development, I enthusiastically embraced the opportunity to work on web development and database management tasks.
- Through dedication and effort, I successfully crafted some remarkable websites.

https://www.imobdevtech.com/

Education	**C.U. SHAH COLLEGE OF ENG & TECH**	BACHALOR OF ENGINEERING • 7.9
	Computer Science	2005-06-06 - 2009-06-06

https://www.cushahuniversity.ac.in/

Projects

AI Scribe Agent (Clinical Documentation Automation) 2024/10/12 - 2025/05/06

Designed and deployed an **AI-powered medical scribe agent** to automatically convert doctor–patient conversations into structured clinical documentation.

- Implemented **speech-to-text, LLM-based clinical understanding, and structured EHR chart generation** (HPI, ROS, Assessment, Plan, medications, labs, and follow-ups).
- Integrated **clinical summarization, coding suggestions, and encounter comparison** to improve provider efficiency and documentation accuracy.
- Enabled **real-time editing, compliance validation, and EHR integration**, reducing physician documentation burden.
- Built scalable deployment using **OCI AI Services, Kubernetes (OKE), and secure healthcare data pipelines**.

Impact:

- Reduced physician documentation time by **60%+** and increased patient interaction time.
- Improved **clinical note completeness and coding accuracy**, supporting better compliance and reimbursement.
- Enhanced **provider productivity and patient throughput** across clinics.

Generalized Gen AI E-commerce Chatbot Feb 2024 - Sep 2024

- Developed a Generative AI chatbot for product-related queries in e-commerce.
- Utilized Snowflake Cortex for development and deployed the chatbot as a native app in the Snowflake Marketplace.
- Used Mistral 7B LLM model for embedding and chatbot building.
- The chatbot assists users with product recommendations, order status updates, and general inquiries, company information and FAQ, live agent transfer, domain related information.
- Integrated advanced Prompt Engineering techniques and LLM models to enhance user interaction and satisfaction.

Shopping Assistance Chatbot Feb 2024 - Aug 2024

Developed a Generative AI chatbot for user assistance, leveraging the GPT-3.5 Turbo version as the language model. The chatbot, named Personic, serves as a shopping assistant, providing customers with various support functions:

- Product fetching
- Order status updates
- General and domain-specific information
- Organization details
- Add-to-cart functionality
- Feedback support

Additionally, the chatbot can transfer conversations to a live agent when necessary, ensuring a seamless customer service experience.

Recommendation System   Jan 2024 - Jun 2024

- Create Recommendation System Engine for Gifting E Commerce company.
- Create Similarity Search Engine, Personalize search engine, Collaborative filtering, Hybrid machine for recommendation, Apriori algorithm.
- Resolve Recommendation, cold start, Up selling, Cross selling issue.
- Used Matrix Factorization, Boltzman machine, GRU for recommendation algorithm

Customer Lifetime Value   2022-08-08 - 2023-04-26

• Estimate the total revenue a customer is expected to generate for a business throughout their entire relationship with the company. • This metric helps businesses identify their most valuable customers and allocate resources effectively to maximize profitability and customer retention. Solution : • Create a equation that can find CLV of customer with help of churn rate. • We use BG/ NBD Model for this solution and test it on 6 months of data. • Created API using ask and integrated with the application.

Tools used: Python, Probabilistic Model , Classification Algorithm, etc.

Customer Segmentation   2022-08-02 - 2023-01-30

• Categorize customers into distinct groups based on shared characteristics and behaviors, allowing businesses to tailor their strategies and offerings to meet each segment's unique needs effectively. Solution : • Create different clusters of customer validate them and named them depend on their characteristics Used RFM analysis. • Used K - Means clustering method and validate it with Silhouette score. • Created API using ask and integrated with the application. Tools used: Python, Unsupervised Algorithms, RFM Analysis, Azure ML pipeline, etc.

Lead Scoring Model   2022-03-22 - 2022-08-31

Create ML model for an automobile company, enabling them to prioritize potential customers based on their likelihood of conversion. Solution = • Built classification probabilistic model. • Used random forest among other classification model. • Got more than 80% accuracy. • Created API using ask and integrated with the CRM portal. Tools used: Python, Classification algorithms, Azure ML pipeline, etc.

Creat a chatbot for internet provider company.   2020-11-02 - 2022-02-01

• To Create a virtual CSR system that can help customers with no time queue. Solution : • Collect samples of Frequently Asked Questions and Answers from the organization, add a few more from online sources. • Also used word2vec model for more vocabulary and word meaning. • Used RNN, LSTM,SEQ2SEQ in order to create a model. • Test it, and improved it with 85%. • Created API and integrated with the application and Whats-app. • Test on more than 500 parallel Query handling. Tools used: Python, RASA, NLP, Word2Vec, RNN, LSTM, pandas, NumPy, etc.

Restaurant billing software Speech to text conversion   2019-02-01 - 2021-05-01

• To create a restaurant billing software so customers can give food orders verbally (as In Speech) and it will convert to text and generate bills too. Solution : • Built State-of-the-Art in-house speech recognition for the Indian accent English language for a company that has more than 30000+ client, and have deployed the same in a different restaurant for order taking through speech. • Take sample recording from all the different states of India as tone and language very for better model performance. • Used Lebrosa library and CNN for deep learning. • Got more than 80% accuracy. • Created API using ask and integrated with the application. Tools used: Python, CNN, Librosa, pandas, NumPy, etc. libraries.

AI Front Desk Agent (Virtual Patient Engagement & Operations)   2025/01/25

Developed a **multilingual AI front desk agent** to automate patient communication and clinic administrative workflows.

- Automated **appointment booking, rescheduling, reminders, insurance pre-verification, and FAQ handling** via voice and chat interfaces.
- Implemented **natural language conversation orchestration using LLMs and intent management** to simulate human-like receptionist interactions.
- Integrated with **hospital management systems, calendars, and CRM/EHR platforms** for real-time updates and scheduling.
- Designed scalable conversational infrastructure supporting **high concurrent patient interactions**.

Impact:

- Reduced **call center workload and administrative overhead by 40%+**.
- Improved **appointment conversion rates and patient response time**, enhancing patient experience.
- Enabled **24/7 patient engagement** without proportional staffing increase.

AI Revenue Cycle Manager (RCM Intelligence & Automation)   2025/06/2 -2025/09/25

Architected an **AI-driven Revenue Cycle Management platform** to optimize billing accuracy, denial reduction, and financial performance.

- Built ML and rule-based intelligence for **claim validation, coding accuracy checks, denial prediction, and payment forecasting**.
- Leveraged **LLMs for clinical-to-billing summarization, documentation gap detection, and automated appeal drafting**.
- Implemented **financial analytics dashboards** for AR aging, denial trends, reimbursement optimization, and KPI monitoring.
- Integrated with **payer workflows, billing systems, and EHR data pipelines** for end-to-end automation.

Impact:

- Reduced **claim denial rates** and accelerated **reimbursement cycles** for healthcare providers.
- Increased **revenue realization and billing accuracy**, improving overall financial performance.
- Provided **data-driven visibility into RCM operations** for leadership decision-making.

References

Vishad Dubey
Managing Director of Transorg

Bhavik Darji
Managing Director of I- Triangle Technolabs

Gayatri Thakkar
CEO and Founder
https://www.inferenz.ai

Interests

Listening Music

Tennis

Travelling

Technical Skills *

- **Programming & Data Engineering**
  Python (Pandas, NumPy, Scikit-Learn, TensorFlow, PyTorch), SQL, Spark, REST APIs, Data Pipelines, Feature Engineering

- **Machine Learning & Deep Learning**
  Regression, Classification, Clustering, Ensemble Models, CNN, RNN, LSTM, Seq2Seq, Recommendation Systems, Time-Series Modeling, Probabilistic Models

- **Generative AI & NLP**
  LLMs (GPT, BERT, Mistral), Prompt Engineering, RAG Pipelines, Fine-Tuning, Embeddings, Vector Databases, Multi-Agent AI Systems, Conversational AI, Clinical NLP

- **Healthcare AI & Analytics**
  AI Scribe, Clinical Documentation Automation, Revenue Cycle Management (RCM) Intelligence, Patient Engagement AI, Medical Coding Assistance, Compliance-aware AI

- **MLOps, Deployment & Architecture**
  FastAPI, Flask, Streamlit, Docker, Kubernetes, CI/CD, Model Monitoring, Experiment Tracking, A/B Testing, Drift Detection, Scalable AI Architecture

- **Cloud & Platforms**
  Oracle Cloud (OCI Data Science, OKE, AI Services, Functions, Object Storage, Streaming), AWS, Azure ML, Snowflake, Git, JIRA, Agile/Scrum

---

Languages

English
Fluent
● ● ● ● ●

Hindi
Fluent
● ● ● ● ●

Gujarati
Native
● ● ● ● ●