

ABHISHEK PAWAR

DATA SCIENTIST

 abhi626pawar@gmail.com

 +918605884412

 <https://www.linkedin.com/in/abhishek-pawar-3bb7a7112/>

SUMMARY

Results-driven Data Scientist with 5 years of experience designing and deploying machine learning, deep learning, and generative AI solutions across domains including sales forecasting, OCR, computer vision, and recommender systems. Adept at transforming raw data into actionable insights and building scalable ML pipelines. Strong cross-functional collaboration skills, combined with hands-on expertise in GenAI and advanced analytics. What sets me apart: I merge classical ML with emerging GenAI techniques to build business-ready solutions, including OCR systems and intelligent forecasting models, balancing innovation with impact.

PROFESSIONAL EXPERIENCE

- vConsturct pvt. Ltd., Pune (Data Scientist)**

Aug 2023 - Present

 - Delivered three GenAI assistants for Marketing, HR, Onsite teams, deployed bots via Azure with <2% downtime. Cut invoice processing time by 65%, saving ~\$100K/year. Enabled 60% faster info retrieval across 5+ departments.
 - Built an OCR pipeline for extracting key fields from construction invoices, improving accuracy by 40% and reducing manual entry.
 - Developed object detection models to identify and count elements in construction design drawings, streamlining quantity estimation.
- Keydabra (Senior Data Scientist)**

Aug 2021 - May 2022

 - Developed sales forecasting models improving accuracy by 18% and optimizing inventory planning.
 - Built a sentiment-aware product feedback system on customer reviews, influencing product roadmap decisions.
 - Created a computer vision-based recommendation engine, analyzing product images and user preferences to improve personalization.
 - Designed models to predict ecommerce conversion rates based on historical browsing and purchase behavior.
 - Collaborated with business teams to prioritize initiatives and mentored junior data scientists.
 - Increased ecommerce conversion rate by 12%. Reduced churn by 10% through model-driven engagement strategies. Contributed ~\$250K in incremental revenue via recommendation systems.
- Keydabra (Machine Learning Engineer)**

Nov 2020 - July 2021

 - Engineered a Customer Engagement Score (DXI) based on user activity data. Performed customer segmentation using K-Means and DBSCAN to tailor marketing efforts and retention strategies. Built data pipelines to transform web analytics into actionable datasets for ML modeling.
 - Partnered with marketing to create data-driven campaigns based on behavioral cohorts. Boosted campaign engagement by 15% via targeted audience segmentation. Enabled real-time customer scoring, improving remarketing precision.
- Globalstep pvt. Ltd., Pune (Test Engineer)**

June 2019 - Oct 2020

 - Designed and implemented automated testing pipelines for mobile and web-based gaming apps, using ML models to detect UI/UX bugs, rendering issues, and frame drops.
 - Utilized computer vision to validate game visuals against expected designs and catch anomalies missed by manual QA.
 - Collaborated with data scientists to deploy reinforcement learning agents for simulating complex player behavior and edge cases.
 - Built internal tools to cluster crash reports and gameplay logs, helping prioritize bugs based on impact and frequency.
 - Reduced manual testing hours by 40%, accelerating release cycles. Improved bug detection accuracy, leading to a 20% drop in post-launch issues.

EDUCATION

- B.E. in Computer Engineering** D.Y. Patil Institute of Technology, Pune

2017
- Diploma in Computer Technology** Government Polytechnic, Beed

2014

PROJECTS

- GenAI Assistant Bots (Marketing, HR, Onsite)**

Description: Created 3 specialized LLM-powered bots using internal SharePoint and Snowflake data:
 - Marketing Bot: Helps Getwork team(Marketing) to win contract bidding. Enabled natural language querying of SharePoint and Snowflake data to get past project insights.
 - HR Bot: Automated policy and document Q&A from the internal SharePoint HR portal.
 - Onsite Bot: Provided recommendations on best building frameworks and materials by querying structured SharePoint data.Tech Used: AzureOpenAI API, LangGraph, LangChain, FastAPI, Postgres, Snowflake, SharePoint API, Docker
- Construction Invoice OCR System**

Description: Developed an end-to-end OCR pipeline for extracting key fields (vendor, total, tax, etc.) from semi-structured construction invoices, replacing manual entry with automated data processing.
Tech Used: Python, OpenCV, Tesseract, FastAPI, Azure, Postgres
- Object Detection for Construction Design**

Description: Built a computer vision model to detect and count architectural components (e.g., beams, columns, fixtures) from engineering diagrams, supporting cost estimation and planning.
Tech Used: YOLOv8, OpenCV, PyTorch, LabelImg, Streamlit, Docker
- Digital Experience Index (DXI) – Customer Engagement Scoring System**

Description: Developed DXI dashboard to track customer engagement and conversions based on web activity.
 - Built a scoring system to quantify user engagement using web behavior data from Google Analytics and purchase history.
 - Engineered behavioral features and applied K-Means clustering to classify users into DXI tiers (Low, Medium, High).
 - Integrated DXI scores into CRM and marketing tools for targeted campaigns, boosting email open rates by 22%.
 - Delivered Tableau dashboards for cross-team visibility and decision-making.Tech: Python, Google Analytics API, Scikit-learn, SQL, K-Means, Power BI

TECHNICAL SKILLS

Machine Learning & AI: Scikit-learn, XGBoost, LightGBM, TensorFlow, PyTorch, OpenCV
GenAI & NLP: OpenAI API, LangChain, LangGraph, Hugging Face Transformers, LLM fine-tuning, Prompt Engineering
Data Science: Python, Pandas, NumPy, Matplotlib, Seaborn, SQL, Time Series Forecasting
Cloud & MLOps: AWS, Azure, Docker, Streamlit, FastAPI, Azure Dev Ops
Tools: Git, Jupyter, VS Code
Soft Skills: Analytical Thinking, Cross-Team Communication, Agile Collaboration, Problem Solving