

# Dheeraj Kumar | Senior Software Engineer

## | AI/ML Systems

• +91 99885-36242 • digitaldk.in@gmail.com • debugwithdheeraj.com  
in digitaldk • Dheeraj-Bhandari • debugwithdheeraj • Bathinda, Punjab, India

## Professional Summary

Senior Software Engineer with 5+ years building scalable AI/ML systems, leading engineering teams, and architecting full-stack applications. Expertise in machine learning infrastructure, distributed systems, microservices, and DevOps automation. Spearheaded autonomous ML platforms and AI customer support systems serving 100+ enterprises.

## Technical Skills

**Languages:** JavaScript, TypeScript, Python, Java, SQL, Bash, HTML5, CSS3

**Frontend:** React.js, Next.js, Vue.js, Redux, Bootstrap, Material-UI, Vite

**Backend:** Node.js, Express.js, Django, Flask, FastAPI, RESTful APIs, GraphQL, Microservices

**AI/ML:** ML Pipelines, LLM Integration, Model Deployment, Fine-tuning, Vector DBs, RAG, OpenAI, TensorFlow

**Databases:** MongoDB, PostgreSQL, MySQL, Redis, DynamoDB, Pinecone, Elasticsearch

**Cloud/DevOps:** AWS (EC2, Lambda, S3, EKS, RDS, SQS, Cognito), Docker, Kubernetes, CI/CD, GitHub Actions

## Professional Experience

### NeuralTalk AI

*Founder & CEO, Acquired by investors*

Punjab, India

Apr 2025–Present

- Founded privacy-first AI support platform serving 100+ companies, reducing support costs by 60%
- Built scalable RAG system with Pinecone achieving sub-30-second response times
- Engineered full-stack platform (React.js, Node.js, Python) integrated with WhatsApp, Slack, Discord, Telegram
- Architected zero-data-retention system handling 200+ daily inquiries with zero breaches
- Integrated SuperTokens, Paddle payments, AWS SES, HubSpot, Notion, Google Sheets, Zendesk APIs

### Neo

*Lead Software Engineer*

Remote

Apr 2024–May 2025

- Directed team of 5+ developers building autonomous ML platform, reducing deployment time by 70%
- Architected backend services for ML orchestration using Python, FastAPI, distributed processing
- Improved system performance by 40% through job scheduling, caching, async processing with Redis
- Developed microservices handling 10,000+ concurrent ML workflows with fault tolerance
- Implemented CI/CD with Docker, Kubernetes, GitHub Actions reducing deployment from hours to minutes
- Mentored engineers, established coding standards, guided technical decisions and architecture

### Monster API

*Lead Software Engineer*

New Delhi, India

May 2023–May 2025

- Developed Monster API NPM package ([npmjs.com/package/monsterapi](https://npmjs.com/package/monsterapi)) enabling 1,000+ developers
- Contributed to open-source Portkey AI Gateway (GitHub PR #280)
- Built flagship AI Playground serving as primary customer acquisition tool
- Created GPT Assistant for LLM fine-tuning automating model deployment workflows
- Planned Docker/K8s architecture achieving 99.9% uptime and horizontal scalability
- Implemented Redis caching increasing speed by 6x, handling 50,000+ daily API requests
- Automated workflows using GitHub Actions increasing efficiency by 5x
- Worked with AWS: Lambda, DynamoDB, EC2, EKS, SQS, RDS, Cognito

### Amazon

*Machine Learning Data Associate*

Chennai, India

Aug 2020–Aug 2022

- Recognized as Top Performer with 2 awards for ML training data projects
- Spearheaded POC enhancing team performance by 25%
- Managed quality enhancement projects improving accuracy by 15%
- Streamlined workflows reducing processing time by 30%

## Education

### Lovely Professional University

Jalandhar, India

MCA, CGPA: 8.0/10.0

2020–2022

### Masai School

Remote

Full Stack Development Bootcamp

2022

### Punjabi University

Patiala, India

BSc Agriculture, CGPA: 7.0/10.0

2016–2020

## Achievements & Certifications