

# DIVYA THAKUR

Data Science Educator · Data Automation Analyst · Electronics Engineer

## CONTACT

+91 9588813396  
ayvid.thakur@gmail.com  
Ajmer, Rajasthan, India  
divyathakur24.github.io/DivyaThakurPortfolio  
LinkedIn: divya-thakurr  
GitHub: DivyaThakur24  
YouTube: @DivyaThakur

## SKILLS

### Languages

Python · R · SQL · Java · HTML/CSS · Shell · Verilog · VHDL

### ML & AI

ML · Deep Learning · NLP · CNNs · RNNs · LSTM · PyTorch · TensorFlow · Scikit-learn

### Data Tools

Power BI · Tableau · Excel · Pandas · NumPy · Matplotlib · Seaborn · Spark · AWS S3

### Automation

n8n · FastAPI · Playwright · Selenium · Web Scraping · API Integration

### Teaching

Curriculum Design · Live Instruction · 1:1 Mentorship · LMS Platforms

### Gen AI Tools

Generative AI Studio · Groq Cloud AI · DALL-E-mini · Stable Diffusion · RAG Systems

### Electronics

VLSI · FPGA · Digital Design · C/C++ · Android (Java/Kotlin)

## EDUCATION

### PG Diploma, VLSI Design

CDAC ACTS, Pune · 73% (Grade A) 2014–2015

### B.Tech, Electronics & Comm.

Bhagwant University · CGPA 7.99 2010–2014

### Class 12, CBSE

Maheshwari Public School, Ajmer · 80% 2009

### Class 10, CBSE

Choithram School, Indore · 86.4% 2007

## ACHIEVEMENTS & SCHOLARSHIPS

- AMCAT: Programming 96.8% | English 97.7% | Quant 91.1%
- All India Rank 300 — C-CAT 2014
- College Rep, E-Cell IIT Bombay
- Udacity Top 10% Mentor
- Speaker — ISDC (PyTorch, Data Viz)
- 1 Million Women to Tech Mentor
- Intel Edge OpenVINO AI Challenge (2019–20)
- Facebook Secure & Private AI / PyTorch (2019)
- AWS Machine Learning Challenge (2019)
- Udacity Bertelsmann Data Science (2018–19)
- Google Udacity Android Scholarship (2018)

## CERTIFICATIONS

- AMCAT — Business Analyst, Data Processing Specialist, Software Engineer
- Tech Mahindra — Artificial Intelligence
- Full Stack Web Dev — FreeCodeCamp

## PROFESSIONAL SUMMARY

Versatile Data professional with **6+ years of experience** at the intersection of education, analytics, and automation — having taught Python, SQL, Power BI, Tableau, and Machine Learning to **50,000+ students** across Udacity, Unacademy, Testbook, and Eduonix. Proven ability to translate complex data concepts into measurable outcomes: from building end-to-end automation pipelines (n8n, FastAPI, Playwright, Selenium) at LabelNest to pioneering corporate training modules adopted by **3 tech startups**. Grounded in Electronics & Communication Engineering (B.Tech, CGPA 7.99) and VLSI Design (CDAC ACTS Pune), with deep expertise in statistical analysis, A/B testing, NLP, deep learning, and scalable ETL workflows. Equally comfortable in hands-on technical roles and cross-functional environments where data drives decisions.

## WORK EXPERIENCE

### Tutedude

May 2026 – Present

#### Data Analytics Mentor

- Mentoring students in Data Analytics covering Python, SQL, Power BI, and Tableau via live and recorded sessions
- Developing structured learning paths and project-based curricula aligned with industry requirements
- Providing 1:1 guidance and feedback to help students build job-ready data analytics skills

### LabelNest

Feb 2026 – Present

#### Data Automation Analyst – I

- Built Private Market Intelligence App for automated data aggregation using FastAPI and n8n AI Automation
- Developed AnnoNest — an annotation intelligence platform with automated scraping via Playwright and Selenium
- Architected Contact Intelligence App using automated enrichment pipelines

### LabelNest

Nov 2025 – Jan 2026

#### Data Automation & Engineering Intern

- Built end-to-end automation pipelines using n8n AI Automation, FastAPI, Python, and SQL
- Developed web scraping and browser automation workflows using Playwright and Selenium
- Automated data extraction, cleaning, validation, and structuring workflows

### SigmaTech AI

Feb 2025 – Mar 2025

#### Generative AI Intern

- Used free LLMs from Groq Cloud AI to generate high-quality, context-aware text on UPSC topics (polity, economy and ecology). Ensured coherence, creativity, and relevance to the prompt.
- Utilized pre-trained generative models like DALL-E-mini and Stable Diffusion to create images from text prompts.
- Developed a Retrieval-Augmented Generation (RAG) system integrating vector search with LLMs to deliver context-aware, fact-grounded AI responses. Improved accuracy and reduced hallucinations, strengthening skills in Python, NLP, and cloud deployment.
- Built a Markov chain text generator to explore statistical NLP models.

### Eduonix

Nov 2022 – Jul 2025

#### Data Science Educator

- Developed 50+ industry-aligned courses in Excel, Python, Power BI, and Machine Learning
- Reduced assessment time by 30% via scalable grading rubrics; corporate modules adopted by 3 tech startups

### Testbook

Jan 2022 – Aug 2022

#### Freelance Educator

- Co-developed Data Science Career Track: EDA with Python, SQL analytics, and Power BI visualizations

### Code Heroku

Mar 2019 – Mar 2020

#### Lead Instructor

- Designed 15+ live workshops on Data Science, ML workflows, and cloud pipelines (AWS, Spark)
- Created project-based curriculum achieving 92% student satisfaction

### Udacity

Mar 2019 – Mar 2020

#### Mentor

- Provided 1:1 mentorship for 80+ Data Analyst & Data Engineering Nanodegree students — Top 10% Mentor

### Unacademy

Nov 2017 – Present

#### Educator

- Produced 200+ tutorial videos; maintained 4.9/5 rating across 5,000+ reviews; 10K+ subscribers

### EduSmart

Feb 2018 – Feb 2019

#### Mentor

- Mentored 50+ students in Android Dev, Data Science, and ML; improved certification pass rates by 25%

### Proxy Notes

Sep – Nov 2017

#### Content Creator

- Designed Mathematics and Science curriculum for high school classes 9th–12th

### OZOCK (Futech Division)

Feb – Apr 2017

#### Web Designer / Developer

- Created logos, banners, and websites using Adobe Illustrator, Chart.js, Duda, and HubSpot CMS

# DIVYA THAKUR

Data Science Educator · Data Automation Analyst · Electronics Engineer

## CONTACT

+91 9588813396  
ayvid.thakur@gmail.com  
Ajmer, Rajasthan, India  
divyathakur24.github.io/DivyaThakurPortfolio  
LinkedIn: divya-thakurr  
GitHub: DivyaThakur24  
YouTube: @DivyaThakur

## SKILLS

### Languages

Python · R · SQL · Java · HTML/CSS · Shell · Verilog · VHDL

### ML & AI

ML · Deep Learning · NLP · CNNs · RNNs · LSTM · PyTorch · TensorFlow · Scikit-learn

### Data Tools

Power BI · Tableau · Excel · Pandas · NumPy · Matplotlib · Seaborn · Spark · AWS S3

### Automation

n8n · FastAPI · Playwright · Selenium · Web Scraping · API Integration

### Teaching

Curriculum Design · Live Instruction · 1:1 Mentorship · LMS Platforms

### Gen AI Tools

Generative AI Studio · Groq Cloud AI · DALL-E-mini · Stable Diffusion · RAG Systems

### Electronics

VLSI · FPGA · Digital Design · C/C++ · Android (Java/Kotlin)

## EDUCATION

### PG Diploma, VLSI Design

CDAC ACTS, Pune · 73% (Grade A) 2014–2015

### B.Tech, Electronics & Comm.

Bhagwant University · CGPA 7.99 2010–2014

### Class 12, CBSE

Maheshwari Public School, Ajmer · 80% 2009

### Class 10, CBSE

Choithram School, Indore · 86.4% 2007

## ACHIEVEMENTS & SCHOLARSHIPS

- AMCAT: Programming 96.8% | English 97.7% | Quant 91.1%
- All India Rank 300 — C-CAT 2014
- College Rep, E-Cell IIT Bombay
- Udacity Top 10% Mentor
- Speaker — ISDC (PyTorch, Data Viz)
- 1 Million Women to Tech Mentor
- Intel Edge OpenVINO AI Challenge (2019–20)
- Facebook Secure & Private AI / PyTorch (2019)
- AWS Machine Learning Challenge (2019)
- Udacity Bertelsmann Data Science (2018–19)
- Google Udacity Android Scholarship (2018)

## CERTIFICATIONS

- AMCAT — Business Analyst, Data Processing Specialist, Software Engineer
- Tech Mahindra — Artificial Intelligence
- Full Stack Web Dev — FreeCodeCamp

## KEY PROJECTS

### DATA AUTOMATION

#### ◆ Private Market Intelligence Workflow

Built automation pipelines converting unstructured private market data into structured intelligence using Python, FastAPI, n8n, Playwright, and Selenium.

#### ◆ Contact Intelligence Enrichment Pipeline

Architected automated contact enrichment workflows combining web scraping, API integration, and validation using n8n and FastAPI at LabelNest.

### PYTHON & STATISTICS

#### ◆ Google App Rating Analysis (41 ★, 68 forks)

Cleaned and analysed 10,841 Play Store records; category-wise EDA on ratings, reviews, installs, and pricing using Pandas, NumPy, and Matplotlib.

#### ◆ A/B Testing for E-Commerce

Analyzed conversion rates via hypothesis testing; recommended retaining original page using logistic regression (Pandas, Statsmodels).

#### ◆ Stock Market Analysis

Resampled and analysed historical stock data; calculated moving averages and volatility metrics using Pandas and Matplotlib.

### SQL

#### ◆ SQL Music Store Analysis (Chinook DB)

Extracted genre and customer insights using advanced SQL queries and subqueries; identified top-selling genres and high-value customers.

### MACHINE LEARNING

#### ◆ Fraud Detection Graph Database

Built Neo4j prototype modelling users, cards, devices, IPs, and merchants; Cypher queries to detect multi-hop suspicious patterns; automated with Python and Docker Compose.

#### ◆ Sentiment Analysis with RNN/LSTM

LSTM-based model for IMDB movie review classification using PyTorch deep learning sequence modelling.

### GEN AI

#### ◆ AnnoNest — Annotation Intelligence Platform

Developed automated annotation scraping and structuring workflows using Playwright and Selenium as part of LabelNest's AI data tooling.

#### ◆ Speech-to-Text Recognition

Leveraged Google AI API for 90–95% accurate conversion of speech and music excerpts to text.

### DATA LAKE & PIPELINES

#### ◆ Data Lake with Spark on AWS

Built ETL pipeline processing JSON logs in S3 using Spark; deployed on AWS cluster for Sparkify music streaming startup.

### OPEN SOURCE

#### ◆ Design Villa — Community Android App

Served as head team member on this collaborative Android app (Google Udacity Scholarship). Coordinated pull requests, reviewed and merged features, and helped shape the final polished release.

## STANDARDISED TEST SCORES

- GRE (Sep 2020): Quantitative Reasoning 160 · Verbal Reasoning 151 · Analytical Writing 3.0
- TOEFL (Nov 2020): 103 — Reading 24 · Listening 27 · Speaking 25 · Writing 27

## CONTINUING EDUCATION

- Data Analyst Nanodegree (Sep 2018 – Mar 2019): Python, data wrangling, A/B testing, regression, visualizations
- Data Foundations Nanodegree (Jun – Aug 2018): Excel, SQL, Tableau for data-informed decision-making
- Android Basics Nanodegree (Feb – Nov 2018): Android app development in Java; networking, SQLite, fragments