VIVEKANAND SINGH

DATA SCIENCE · MACHINE LEARNING · DATA ENGINEERING

SINGHVIVEKANAND376@GMAIL.COM
VIVEKANAND-SINGH.PAGES.DEV
GITHUB.COM/VIVEKANANDSINGH1
LEETCODE.COM/U/VIVEKANANDSINGH1
DHARWAD, KARNATAKA · FROM VARANASI, UP

First-year B.Tech student in Data Science and AI at IIIT Dharwad. I build data pipelines and dashboards, three of them live, and I'm training a terrain segmentation model on NASA JPL data. Looking for a data science or ML internship.

EDUCATION

Indian Institute of Information Technology (IIIT) Dharwad — 2026 – 2030
B.Tech in Data Science and Artificial Intelligence · first year — DHARWAD, KARNATAKA

PROJECTS

GH Pulse · Batch ETL over GitHub Archive — LIVE · CODE

PYTHON · HTTPX · DUCKDB · PARQUET · POSTGRESQL · PYTEST · GITHUB ACTIONS · STREAMLIT

- Built a batch ETL pipeline that streams GitHub Archive's 8,784 hourly gzipped JSON files a year, with retries on failed downloads.
- Unnested 100M+ raw events into Parquet with DuckDB and aggregated repo stats, contributor counts and event-type breakdowns.
- Loaded results into PostgreSQL with idempotent upserts and an audit-log row for every run; 15+ tests, CI on every push, and a live Streamlit dashboard on top.

Mars Rover Terrain Segmentation · In progress — 2026

PYTHON · PYTORCH / TENSORFLOW · AI4MARS (NASA JPL)

- Building a segmentation model that labels every pixel of Curiosity rover images as soil, bedrock, sand or big rock, so a rover can tell where it is safe to drive.
- Training on AI4Mars, NASA JPL's terrain dataset labelled by thousands of volunteers.

NetWatch · Live attack map from server logs — LIVE · CODE

PYTHON · D3 · LINUX · SYSTEMD · UFW · CADDY · GEOIP

- Built a collector that follows firewall drops and failed SSH logins from the systemd journal, masks every source IP and geolocates it with an offline database.
- Wrote detection rules for port scans (10+ closed ports in 10 s) and SSH brute force (5+ failed logins in 5 min), with a cooldown per source.
- Deployed on a VPS as a sandboxed systemd service with a 200 MB memory cap, behind Caddy HTTPS with a strict CSP; 10 tests. A D3 world map draws each event live.

Messenger · Private, invite-only chat — LIVE · CODE PRIVATE

REACT · VITE · TAILWIND CSS · NODE.JS · EXPRESS · SOCKET.IO · POSTGRESQL · PRISMA · ZOD · DOCKER · CADDY

- Built a real-time chat app with direct and group chats, group admin roles, friend requests, blocking, typing indicators, presence and read receipts.
- Encrypted every message at rest with AES-256-GCM using a fresh IV per message; argon2id passwords, emailed one-time codes at sign-up, and rotating, revocable refresh tokens.
- Validated every request body with Zod, rate-limited auth routes, and deployed with Docker Compose behind Caddy HTTPS.

Vajra P1 · Service pistol design concept — 2026

BLENDER · PYTHON (BPY) · CYCLES

- Wrote a design brief and five engineering schematics for a modular service pistol sized from DRDO DIPAS Indian soldier hand data (68 mm trigger reach vs 72 mm on the Glock 17).
- Modelled it parametrically in Blender with a Python script, iterated through 14 versions, and rendered a 15 s Cycles reveal; wrote a one-page pitch against the in-service pistol.

SKILLS

USED IN THE PROJECTS ABOVE

LANGUAGES	Python, SQL, JavaScript, C++
DATA	DuckDB, Parquet, PostgreSQL, Streamlit, D3
BACKEND	Node.js, Express, Socket.io, Prisma, Zod, PostgreSQL
FRONTEND	React, Vite, Tailwind CSS
INFRA AND TESTING	Docker, Linux, systemd, Caddy, GitHub Actions, pytest, Git
SECURITY	AES-256-GCM, argon2id, JWT, CSP, firewall logging
3D	Blender with Python scripting, Cycles

PROBLEM SOLVING

27 LeetCode problems solved, 12 of them hard, mostly in C++ plus SQL, JavaScript and Python. Solutions on GitHub