

VINEET SHUKLA

Data Science & Machine Learning | AWS Cloud

Lucknow, Uttar Pradesh, India | +91 92369 50937 | vineetshukla.work@gmail.com | huggingface.co/Pandaisop
vineet-shukl.xyz | github.com/Vineet-shukl | linkedin.com/in/vineet-shukla-0b1b67313 | credly.com/users/vineet-shukla.54a43186

SUMMARY

Third-year B.Tech CSE student seeking Data Science, Machine Learning, and AWS Cloud internships. Builds end-to-end ML apps in Python: a deepfake-speech classifier on a fine-tuned Wav2Vec 2.0 model (PyTorch, Hugging Face) served as a live REST API, and a TensorFlow/Keras CNN for facial emotion recognition, plus LLM agent apps. Core strength in AWS: hands-on labs across 23 AWS services and a production serverless application on 10+ services defined entirely in CloudFormation. Credly-verified IBM data and AI credentials, 54 Google Cloud skill badges, and 5th national rank among 40,000+ in the India AI Impact Buildathon.

SKILLS

Programming & Databases: Python, SQL, JavaScript, TypeScript, HTML/CSS

Machine Learning & Deep Learning: Regression, Classification, Clustering (decision trees, random forest, K-means), scikit-learn, deep learning, CNNs, TensorFlow/Keras, PyTorch, Hugging Face Transformers (Wav2Vec 2.0), model fine-tuning, audio processing (librosa)

Data Science & Analytics: Pandas, NumPy, SciPy, statistics, probability, data cleaning, data analytics, Jupyter Notebook, BigQuery ML

Data Visualization: Matplotlib, Seaborn, Tableau Desktop, IBM Watson Studio, Streamlit

Generative AI & LLMs: LLMs (Groq/Llama, Google Gemini), RAG, AI agents (Pydantic AI), prompt engineering, IBM Granite

AWS Cloud: EC2, VPC, S3, RDS, Lambda, API Gateway, CloudFront, Route 53, ACM, Cognito, DynamoDB, EventBridge, IAM, CloudWatch, CloudTrail, SNS, WAF, Elastic Load Balancing, Auto Scaling, SSM Parameter Store, CloudFormation (IaC), serverless

DevOps & Platforms: CI/CD, GitHub Actions, AWS CodePipeline, Git, GitHub, pytest, Google Cloud, Firebase, Hugging Face Spaces

Web & APIs: FastAPI, Django, Flask, REST APIs, React, Pydantic

EXPERIENCE

Cloud Computing & DevOps / Networking Intern | SRDT Technologies

Jul 2026

- Completed hands-on labs on 23 AWS services across security and networking, compute and storage, high availability and databases, serverless, monitoring, and DevOps, documenting each with setup guides and proof screenshots in a public repository.
- Built a highly available web tier in a custom VPC with public and private subnets across two Availability Zones: EC2 behind an Application Load Balancer with an Auto Scaling group (1 to 4 instances) and an RDS MySQL database.
- Configured monitoring and protection with a CloudWatch CPU alarm and SNS email alerts, multi-region CloudTrail audit logging to S3, and an AWS WAF rate-limit rule; set up CI/CD with CodePipeline and GitHub Actions using a least-privilege IAM deployer.
- Delivered MemeGuru end to end as the internship project with phased CloudFormation stacks, a PRD, and deployment runbooks.

Academic Intern, Data Analytics with AI | AICTE – IBM SkillsBuild (BharatCares)

Aug 2026

- Completed the data analytics track covering data concepts, the data science landscape, data tools and languages, and cleaning, refining, and visualizing data in IBM Watson Studio and Tableau Desktop.
- Completed hands-on modules on Generative AI for data analytics, including summarizing and classifying data with IBM Granite.

Python Full Stack Intern | SRDT Technologies

- Developed backend web features using Python frameworks, including Django.
- Integrated frontend interfaces with backend services and database operations.

PROJECTS

VoiceGuard – Deepfake Speech Classifier | Live demo | GitHub

Python, PyTorch, Hugging Face Transformers, Wav2Vec 2.0, librosa, SciPy, NumPy, FastAPI, pytest

- Built a classifier that detects AI-generated (deepfake) speech using a fine-tuned Wav2Vec 2.0 (XLSR-53) backbone with an attentive statistics pooling head and MLP, scoring multiple overlapping segments of each clip to catch partial deepfakes.
- Engineered a forensic analysis engine (spectral smoothness, temporal cadence, formant transitions, phase and synthesis artifacts) plus audio-quality profiling (SNR, clipping, silence ratio).
- Designed a decision-fusion layer that raises confidence when the neural and forensic engines agree and flags disagreements for manual review; served via a FastAPI endpoint returning a label, confidence score, and explanation (live on Hugging Face Spaces).

CampusMind – Facial Emotion Recognition (Hackathon Team Project) | GitHub

Python, TensorFlow/Keras, CNN, NumPy, Django

- Designed a Keras CNN for 7-class facial emotion recognition (FER2013 classes, 48x48 grayscale): three convolution blocks with batch normalization, max pooling, and dropout, plus a dense softmax head (Adam, categorical cross-entropy).
- Set up image augmentation (rotation, shift, zoom, flip) and a prediction script for the team's Django mental-health platform.

VoteWise India – AI Election Assistant | Live | GitHub

Python, Google Gemini API, Firebase Cloud Functions, Firestore, GitHub Actions, pytest

- Built a non-partisan assistant that answers Indian voters' questions on registration, eligibility, EPIC cards, EVMs, and election schedules in all 22 official Indian languages.
- Designed a three-layer answer pipeline (partisan-keyword guard, Firestore response cache with 24-hour TTL, then Gemini) grounded by a deterministic rules engine over Election Commission of India data.
- Exposed REST endpoints for chat, eligibility, and timelines with per-user rate limiting (30 requests/hour via atomic Firestore transactions), locked-down Firestore rules, and strict CORS/CSP headers; automated deploys with GitHub Actions.

MemeGuru – Serverless Application on AWS | Live | GitHub

AWS Lambda, API Gateway, S3, CloudFront, Route 53, ACM, Cognito, DynamoDB, EventBridge, CloudWatch, SNS, SSM, CloudFormation

- Architected a serverless backend on 10+ AWS services defined 100% in phased CloudFormation stacks; costs about \$0.50/month.
- Secured it with Cognito email-verified accounts, a JWT-protected API, least-privilege IAM per Lambda, private S3 behind CloudFront Origin Access Control, and SSM SecureString secrets; monitored with a nightly canary, SNS alerts, and a \$5 billing alarm.

PhishGuard – LLM Phishing Email Classifier | GitHub

Python, Pydantic AI, Groq (Llama 3.3 70B), Streamlit, FastAPI, Jupyter Notebook

- Built an AI agent that classifies raw emails as phishing or legitimate and returns a schema-validated threat report (confidence, severity, attack type, red flags, spoofed brand, recommended action).
- Equipped the agent with URL-extraction, typosquatting, and header-mismatch tools; shipped a Streamlit UI and FastAPI endpoint.

Legalogy – AI Legal Assistant | Live | GitHub

React, TypeScript, Vite, Tailwind CSS, Google Gemini API, Firebase

- Built a web app that explains contract clauses in plain English and Hindi and audits them against 6+ Indian statutes, returning structured JSON with statute citations and falling back to an offline rule-based engine.

CERTIFICATIONS

IBM SkillsBuild (Credly-verified, credly.com/users/vineet-shukla.54a43186)

- **Data Fundamentals; Getting Started with Data; Generative AI Essentials: Using LLMs to Work with Data; Make Agentic AI Work for You** (all Aug 2026); **AI Literacy** (Sep 2026)
- **AICTE | IBM SkillsBuild Academic Internship – Data Analytics with AI | BharatCares** (Aug 2026)
- **Courses (Aug 2026):** Introduction to Generative AI; Introduction to Large Language Models; Introduction to Retrieval Augmented Generation; The Rise of Multiagent Systems; Summarizing and Classifying Data Using IBM Granite; Supercharge Your Data Analytics with Generative AI; Introduction to Tableau Desktop; Clean, Refine, and Visualize Data with IBM Watson Studio

Google Cloud (54 skill badges, Oct – Dec 2025, Credly-verified)

- **Data & Analytics (12):** BigQuery, BigQuery ML predictive modeling, data warehousing, streaming analytics, Bigtable, Dataplex, Looker
- **AI, ML & Generative AI (12):** ML APIs, Vertex AI prompt design, Gemini multimodal RAG, Vision, Natural Language, Speech
- **Serverless & App Development (12):** Cloud Run, Cloud Run functions, Firebase, Eventarc, API Gateway, Apigee X, App Engine
- **DevOps & Operations (7):** CI/CD pipelines, DevOps workflows, secure software delivery, Cloud Monitoring and Logging
- **Security & IAM (4):** cloud security fundamentals, service accounts and IAM roles, Security Command Center, data protection
- **Networking, Compute & Kubernetes (4):** VPC networking, Network Connectivity Center, Kubernetes deployments, Compute Engine
- **Workspace & No-Code (3):** AppSheet, Apps Script, collaboration and productivity workflows

LEADERSHIP & ACTIVITIES

Technical Head | GeeksforGeeks Campus Body, SRMCEM | Aug 2026 – Present

- Built the chapter's mailer studio and appointment-letter/certificate PDF generator: gfg-srmcem-suite.vercel.app

Event Management Coordinator | GeeksforGeeks Campus Body, SRMCEM | Nov 2025 – Aug 2026

- Organized technical events and programming initiatives for the campus developer community.

Campus Ambassador | Google Student Ambassadors (India), Lucknow | May 2026 – Present

PR & Outreach Coordinator | AlgoZenith SRMCEM | Oct 2025 – Apr 2026

- Led outreach initiatives and promoted technical events and hackathons across campus.

5th Rank Nationally | India AI Impact Buildathon (Team Byte-Bandit), among 40,000+ participants

EDUCATION

Bachelor of Technology, Computer Science and Engineering

Sep 2024 – May 2028 (expected)

Shri Ramswaroop Memorial College of Engineering and Management (SRMCEM), Lucknow (AKTU) | CGPA: 7.54/10

Class XII (75.2%, 2024), Class X (85.33%, 2022) | Evergreen Public School, Ghazipur