

N-22/2/2021-NeGD-Part(1)
National e-Governance Division
Digital India Corporation
Electronics Niketan, 6, CGO Complex Lodhi Road,
New Delhi – 110003
Website: www.negd.gov.in / www.dic.gov.in

Web Advertisement
13.10.2025

The National e-Governance Division (NeGD) is an independent business division under the Digital India Corporation, Ministry of Electronics and Information Technology. NeGD has been playing a pivotal role in supporting MeitY in Programme Management and implementation of e-Governance projects and initiatives undertaken by various Ministries/ Departments, both at the Central and State levels.

NeGD has been spearheading several innovative initiatives under the aegis of the Digital India Programme. Those have been developed keeping the vision areas of Digital India at the core- providing digital infrastructure as a core utility to every citizen, governance and services on demand and in particular, digital empowerment of the citizens of our country; some of these initiatives include DigiLocker, UMANG, Poshan Tracker, OpenForge Platform, API Setu, National Academic Depository, Academic Bank of Credits, Learning Management System.

It has myriad roles and responsibilities from supporting Central Line Ministries and State Government Departments on e-Governance projects, reviewing State Action Plans, offering support in technology management, strategy formulation & implementation of Emerging Technologies viz. AI, Blockchain, GIS etc., to facilitating digital diplomacy with focus on Indian startups and products

NeGD has been a leader in implementation and execution of a gamut of pilot/ infrastructure/ technical/ special projects and support components to framing core policies, project appraisals, R&D, and guiding /conducting assessments, undertaking activities for building capacities of both Government officials and] other stakeholders, and creating mass awareness about schemes and services under the Digital India Programme.

NeGD is currently inviting applications for the following positions purely on Contract basis initially for a period of 3 years which is further extendable as per the requirement of the project.

S. No	Position	Vacancy
1	Full Stack Developer (AI/ML & Web Applications)	01
2	AI/ML Engineer	01
3	Data Scientist (AI & ML)	01

* The maximum age limit shall be 55 years on the closing date of receipt of application.

** The place of posting shall be in New Delhi but transferable to project locations of NeGD as per existing policy of NeGD/DIC.

Screening of applications will be based on qualifications, age, and relevant experience. NeGD reserves the right to fix higher threshold of qualifications and experience for screening and limiting the number of candidates for interview. Only shortlisted candidates shall be invited for interviews. NeGD reserves the right to not to select any of the candidates without assigning any reason thereof.

The details can be downloaded from the official website of DIC and NeGD, viz. www.dic.gov.in, www.negd.gov.in.

Eligible candidates may apply ONLINE: <https://ora.digitalindiacorporation.in/>

Last date for submission of applications will be: 26.10.2026

1. Full Stack Developer (AI/ML & Web Applications)

Role Summary

The Full Stack Developer will design, build and deploy secure, scalable web applications with integrated AI and machine learning capabilities. The role requires strong full-stack engineering across frontend, backend, data and cloud layers, together with practical AI/ML engineering competence: integrating models across commercial, cloud and self-hosted providers, building retrieval and agent workflows, and adapting and serving models where required.

Educational Qualification

- B.Tech./B.E., M.Tech., MCA or M.Sc. in Computer Science, Information Technology, AI/ML or a related discipline (**Must have**)
- Certifications are not mandatory; demonstrated work, as listed under Minimum Evidence at Screening, shall carry greater weight
- Candidates with strong verifiable work (production applications or significant open-source contributions) will be considered in lieu of formal qualification

Experience

- 4+ years of professional full-stack development (5+ years preferred); **minimum 2 years building and deploying AI- or ML-powered features in production applications**
- Demonstrable end-to-end ownership of at least one production web application, covering frontend, backend, database, deployment and monitoring
- **Demonstrable experience delivering at least one LLM or ML-powered application to end users** (such as a retrieval-augmented assistant, a document-processing workflow or an agent), including its evaluation and monitoring
- Hands-on experience with Python and TypeScript in production, and with React-based frontends
- Experience deploying containerised applications on cloud platforms through CI/CD pipelines
- Experience fine-tuning or self-hosting models, working in regulated sectors, or building Indian-language applications is a plus

Key Responsibilities

Application Development

- Design and develop responsive, accessible web applications using React and Next.js with TypeScript
- Develop backend services and APIs in Python (FastAPI) and Node.js, with clear contracts (OpenAPI or GraphQL) and versioning
- Design data models and optimise queries across relational and NoSQL databases, and implement caching and background job processing
- Translate functional requirements into well-structured, maintainable and production-ready code

AI & LLM Integration

- Integrate LLMs in a provider-agnostic manner across commercial APIs, cloud AI platforms and self-hosted open-weight models, through an abstraction or gateway layer, so that models can be switched or compared without application changes
- Build LLM-powered features, including streaming conversational interfaces, retrieval-augmented generation with source citations, structured extraction and tool-calling agents using the Model Context Protocol (MCP)
- Implement retrieval pipelines covering document ingestion and parsing, chunking, embeddings, hybrid vector search and re-ranking
- Implement routing, caching, rate limiting and fallbacks across providers to control quality, cost and latency

Model Integration & Serving

- Integrate and serve pre-trained models (such as OCR, speech, classification and embedding models, including Indic language models) behind production APIs
- Fine-tune or adapt models where required (for example, LoRA fine-tuning of an LLM or embedding model), in coordination with data scientists
- Build evaluation suites for AI features (golden datasets, regression tests and LLM-based evaluation) and run them in CI as release gates

Security, Privacy & Responsible AI

- Implement authentication and authorisation (OAuth 2.0, OpenID Connect and role-based access control), and secure handling of APIs, secrets and personal data
- Apply OWASP Top 10 and OWASP Top 10 for LLM Applications controls, including defences against prompt injection and data leakage
- Apply data-minimisation practices in compliance with applicable data protection law, including for data sent to external model providers, and design human review for consequential AI outputs

Quality, Deployment & Operations

- Practise test-driven development with unit, integration and end-to-end tests, and maintain adequate coverage
- Containerise applications and manage deployments through CI/CD pipelines on Kubernetes or serverless platforms, using infrastructure as code
- Instrument applications with logging, metrics and tracing, including LLM traces, token usage and cost, and participate in incident resolution
- Ensure accessibility compliance (WCAG 2.2 and GIGW) and support for Indian languages in user interfaces

Collaboration & Mentoring

- Participate in Agile delivery, code reviews and architecture discussions
- Mentor junior developers, and document designs, APIs and operational runbooks

Technical Competencies

- **Frontend:** TypeScript, React, Next.js (App Router and server components), Tailwind CSS, component libraries (shadcn/ui or equivalent), TanStack Query; accessibility and internationalisation
- **Backend:** Python (FastAPI, Pydantic, asyncio) and Node.js (NestJS, Fastify or Express); REST with OpenAPI, GraphQL; WebSockets and server-sent events for streaming; background jobs (Celery or BullMQ)
- **Databases & Caching:** PostgreSQL (with pgvector), MongoDB, Redis; ORMs (SQLAlchemy, Prisma or Drizzle); schema migrations and query optimisation
- **LLM Providers & Frameworks:** commercial APIs (OpenAI, Anthropic, Google Gemini), cloud AI platforms (Azure AI Foundry, AWS Bedrock, Google Vertex AI), Indian providers (Sarvam, Bhashini) and self-hosted open-weight models via vLLM or Ollama; gateways such as LiteLLM; LangGraph, LlamaIndex or equivalent multi-provider frameworks; structured outputs, tool calling and MCP
- **Retrieval & Vector Search:** embeddings, chunking strategies, hybrid search and re-ranking; vector databases (pgvector, Qdrant, Milvus or Weaviate); document parsing (Docling or vision-language OCR models)
- **Applied ML:** PyTorch and Hugging Face Transformers for inference and fine-tuning (PEFT); scikit-learn; ONNX Runtime and quantised models; pre-trained models for OCR, speech (IndicConformer, Whisper) and Indic NLP (AI4Bharat)
- **AI Evaluation & Observability:** promptfoo, DeepEval or RAGAS in CI; MLflow; Langfuse or Arize Phoenix; OpenTelemetry, Prometheus and Grafana

- **Security:** OAuth 2.0, OpenID Connect, JWT, Keycloak or equivalent identity providers; OWASP Top 10 and OWASP Top 10 for LLM Applications; secrets management; guardrails for LLM inputs and outputs
- **Testing, DevOps & Cloud:** pytest, Vitest or Jest, Playwright; Git, GitHub Actions or GitLab CI, Docker, Kubernetes and Helm, Terraform; AWS, Azure or GCP managed services
- **Data Pipelines (desirable):** Airflow or Dagster; Kafka or other message queues

2. AI/ML Engineer

Role Summary

The AI/ML Engineer will develop, train, fine-tune, optimise and deploy machine learning, deep learning and large language models for text, document, image and speech applications. The role requires depth in modern model architectures and training methods, rigorous evaluation, and the engineering discipline to deliver models that are efficient, reliable and secure in production.

Educational Qualification

- B.Tech./B.E., M.Tech./M.S. or M.Sc. in Computer Science, AI/ML, Data Science, Electronics or a related discipline (**Must have**)
- Advanced degree (M.Tech./M.S./Ph.D.) or published work in machine learning, NLP, computer vision or speech is highly desirable
- Certifications are not mandatory; demonstrated work, as listed under Minimum Evidence at Screening, shall carry greater weight
- Candidates from non-traditional backgrounds with strong verifiable work (published models, competition results or significant open-source contributions) will be considered in lieu of formal qualification

Experience

- 4+ years of hands-on experience developing and deploying AI/ML models (6+ years preferred); **minimum 3 years training or fine-tuning deep learning models that were deployed in production**
- Demonstrable end-to-end delivery of at least two models, from data preparation through training, evaluation, optimisation and deployment, with measured outcomes
- **Demonstrable hands-on experience fine-tuning LLMs or other foundation models on domain data** (supervised fine-tuning, parameter-efficient fine-tuning or preference optimisation), with documented evaluation against a baseline
- Practical experience in at least two of the following: NLP and LLM applications, computer vision, document AI and OCR, speech recognition or synthesis, and predictive models on structured data
- Experience optimising and serving models on GPUs to meet production latency and throughput targets
- Experience in regulated sectors or with Indian-language data is a plus

Key Responsibilities

Model Development & Training

- Develop, train and fine-tune machine learning and deep learning models for use cases such as text classification, summarisation, entity extraction, document understanding, image matching, object detection and speech recognition
- Select architectures and pre-trained models suited to the task, the data and the deployment constraints, starting from strong baselines and preferring transfer learning to training from scratch
- Build training data pipelines covering preprocessing, feature engineering, data augmentation and synthetic data where appropriate, with versioned datasets
- Run reproducible, tracked experiments, including hyperparameter tuning, and perform systematic error analysis to improve accuracy, robustness and latency

LLM Adaptation & Applications

- Fine-tune open-weight LLMs and embedding models through supervised fine-tuning, parameter-efficient methods (LoRA, QLoRA), preference optimisation (DPO) and

reinforcement learning with verifiable rewards (GRPO and its variants), where these outperform prompting

- Distil larger models into smaller, cost-efficient models suitable for production
- Build retrieval-augmented and agentic systems with sound retrieval design, structured outputs and tool calling, integrated across commercial, cloud and self-hosted model providers through an abstraction layer
- Engineer the context supplied to models (retrieved content, memory, tool outputs and instructions), and set reasoning and token budgets for reasoning models against cost and latency targets
- Choose between prompting, retrieval, fine-tuning and smaller specialised models based on accuracy, cost, latency and data-governance requirements, and document the rationale

Computer Vision, Document & Speech AI

- Develop document-understanding pipelines combining OCR, layout analysis and vision-language models for forms, scanned documents and handwritten records
- Develop and fine-tune vision models for classification, detection, segmentation and image similarity, including embedding-based matching
- Develop and fine-tune speech recognition and synthesis models, including for Indian languages and code-mixed speech

Inference Optimisation & Serving

- Optimise models through quantisation (including FP8 and FP4 formats), pruning, distillation and compilation (ONNX, TensorRT), and apply serving optimisations such as continuous batching, prefix caching and speculative decoding
- Benchmark accuracy, latency, throughput, memory and cost per request on CPU and GPU against defined service-level targets
- Deploy models with production serving frameworks, including autoscaling, versioning and rollback, in coordination with the MLOps team
- Monitor deployed models for accuracy drift, latency, GPU utilisation and cost, and feed production failures back into training and evaluation data

Evaluation & Quality

- Define task-appropriate metrics (such as F1, AUC, BLEU, chrF, ROUGE, WER, CER and mAP), and build golden datasets, benchmarks and regression suites that run as release gates
- Evaluate LLM and retrieval systems for groundedness, hallucination and retrieval quality, using LLM-based evaluation calibrated against human labels
- Evaluate agentic systems on trajectories, tool-call correctness and task completion, and diagnose failures from execution traces
- Conduct subgroup analysis across languages, demographics, document types and acoustic conditions, and work with QA on acceptance thresholds

Responsible AI, Safety & Privacy

- Apply guardrails and test models for prompt injection, jailbreaks, leakage of personal data and unsafe outputs, in coordination with AI safety specialists
- Measure and mitigate performance disparities across cohorts, and document models through model cards, dataset documentation and evaluation records
- Apply anonymisation and data minimisation in training and inference in compliance with applicable data protection law, and verify the licences of models and datasets used

Collaboration & Mentoring

- Work with data scientists, data engineers, MLOps and application teams to integrate models into production applications through well-defined APIs
- Critically assess emerging models and research for practical value, review the work of junior engineers, and review model deliverables from external vendors

Technical Competencies

- **Programming:** Python (expert level); SQL; Git; C++ or Rust for performance-critical components (desirable)
- **Deep Learning Frameworks:** PyTorch (primary), JAX or TensorFlow; Hugging Face Transformers, Datasets and Accelerate
- **Model Architectures:** transformers, mixture-of-experts and reasoning models; CNNs and vision transformers; conformer-based speech models; embedding and contrastive models
- **LLM Training & Adaptation:** Hugging Face TRL and PEFT, Unsloth or Axolotl; verl or OpenRLHF for reinforcement learning; supervised fine-tuning, LoRA and QLoRA, DPO, GRPO and reinforcement learning with verifiable rewards; distillation and synthetic data generation
- **GPU & Distributed Training:** mixed-precision training, FlashAttention, working knowledge of CUDA and GPU profiling; PyTorch FSDP or DeepSpeed for multi-GPU training (desirable)
- **Inference Optimisation & Serving:** quantisation (GPTQ, AWQ, FP8 or NVFP4); continuous batching, prefix caching and speculative decoding (such as EAGLE-3); vLLM, SGLang, TensorRT-LLM or NVIDIA Triton Inference Server; ONNX Runtime; llama.cpp for CPU and edge deployment
- **Vision, Document & Speech:** OpenCV; detection and segmentation models; vision-language models for document parsing (such as PaddleOCR-VL, Docling or the Qwen-VL family); speech recognition (IndicConformer, Whisper or newer open models) and speech synthesis
- **Indic Language AI:** AI4Bharat models (IndicConformer, IndicTrans2, IndicBERT, Indic Parler-TTS), Sarvam models and Bhashini services; handling of code-mixed and transliterated text
- **LLM Applications & Agents:** retrieval-augmented generation with hybrid search and re-ranking; vector databases (pgvector, Qdrant, Milvus or Weaviate); LangGraph or LlamaIndex; Model Context Protocol (MCP); context engineering; multi-provider integration through gateways such as LiteLLM
- **Evaluation:** task-specific metrics; lm-evaluation-harness, RAGAS, DeepEval, promptfoo or Inspect; agent trajectory and tool-use evaluation; human evaluation and inter-annotator agreement
- **MLOps & Platforms:** MLflow or Weights & Biases, DVC and model registries; Docker and Kubernetes; Ray (desirable); AWS SageMaker, Azure Machine Learning or Google Vertex AI; Langfuse or Arize Phoenix for LLM tracing
- **AI Safety & Privacy:** guardrail frameworks (NeMo Guardrails, Llama Guard), red-teaming tools (garak or PyRIT), anonymisation and working knowledge of differential privacy

3. Data Scientist (AI & ML)

Role Summary

The Data Scientist will design, develop and evaluate statistical, machine learning and AI solutions, including LLM applications, retrieval systems and deep learning models for text, image, speech and document data. The role requires sound judgement in selecting the appropriate technique, rigorous validation of predictive and generative models, and delivery of production-quality code.

Educational Qualification

- B.Tech./B.E., M.Tech./M.S. or M.Sc. in Computer Science, Statistics, Mathematics, Data Science, AI/ML or a related quantitative discipline (**Must have**)
- Advanced degree (M.S./M.Tech./Ph.D.) or published applied work in machine learning, NLP, computer vision or statistics is highly desirable
- Certifications are not mandatory; candidates with strong verifiable work (Kaggle Expert level or above, published research or significant open-source contributions) will be considered in lieu of formal qualification

Experience

- 4+ years in data science, machine learning or applied AI (6+ years preferred); **minimum 3 years developing models or AI systems deployed in production or used in operational decisions**
- Demonstrable end-to-end ownership of at least two models or AI systems, with measured outcomes
- **Hands-on experience with at least one LLM or deep learning application deployed for end users**, with documented evaluation
- Experience in at least two of the following: fraud or anomaly detection, forecasting, risk scoring, NLP and LLM applications, computer vision or document AI, speech, and causal inference
- Experience with large, heterogeneous datasets, including linkage of records across sources
- Experience in a regulated sector (financial services, healthcare, telecommunications or the public sector) or with Indian-language data is a plus)

Key Responsibilities

Problem Formulation

- Translate business requirements into well-defined analytical problems with a clear decision, baseline and evaluation metric
- Select the appropriate approach (rules-based, classical machine learning, deep learning, a fine-tuned model or an LLM) based on accuracy, cost, latency and explainability, and document the rationale
- Assess data feasibility, including label quality, leakage and bias, and advise where the data is insufficient for the objective

Machine Learning & Statistical Modelling

- Develop and validate models for classification, regression, ranking and forecasting, starting from robust baselines
- Develop fraud and anomaly detection models for imbalanced settings with limited labelled data
- Generate labels through weak supervision, active learning and LLM-assisted labelling, with human validation

Generative AI & Deep Learning

- Develop LLM-based solutions, including retrieval-augmented generation, document extraction, summarisation and tool-using agents

- Design retrieval components (chunking, embeddings, hybrid search and re-ranking), and adapt models through prompting, parameter-efficient fine-tuning and distillation
- Develop deep learning solutions for NLP, document understanding (OCR and vision-language models), vision and speech, including Indian-language and code-mixed data
- Optimise cost and latency, including the choice between open-weight and API-based models

Evaluation & Statistical Rigour

- Validate models using leakage-free schemes, and quantify uncertainty through prediction intervals, calibration and conformal prediction
- Evaluate generative AI systems through golden datasets, separate retrieval and generation metrics, hallucination checks, and LLM-based evaluation calibrated against human labels
- Conduct error and subgroup analysis across cohorts, languages and input types
- Design experiments and apply quasi-experimental methods to assess whether deployed models have improved outcomes

Production & Monitoring

- Develop tested, peer-reviewed, production-quality Python code and ensure all experiments are reproducible
- Hand over models to the MLOps team with defined monitoring metrics, drift thresholds and retraining triggers
- Monitor deployed systems for quality, drift, latency and cost, and incorporate failures into evaluation sets

Responsible AI & Privacy

- Measure and mitigate performance disparities across cohorts, and provide explanations for decisions affecting individuals
- Test LLM systems for prompt injection, jailbreaks and leakage of personal data, and design guardrails in coordination with security specialists
- Apply anonymisation and data minimisation in compliance with applicable data protection law, and prepare model cards and evaluation records

Communication & Mentoring

- Prepare clear reports for non-technical stakeholders, including model limitations and uncertainty
- Review the work of junior data scientists, and support the review of vendor model deliverables

Technical Competencies

- **Statistics & Causal Inference:** generalised linear and hierarchical models, Bayesian inference, hypothesis testing, bootstrap methods; A/B testing, difference-in-differences, synthetic control, regression discontinuity (DoWhy, EconML)
- **Machine Learning:** XGBoost, LightGBM and CatBoost; tabular foundation models (TabPFN); anomaly detection, clustering and graph methods; forecasting with statistical models and time-series foundation models (Chronos-2, TimesFM)
- **Generative AI & LLMs:** prompt design, structured outputs and tool calling; retrieval-augmented generation with hybrid search and re-ranking; vector databases (pgvector, Qdrant, Milvus or Weaviate) and FAISS; open-weight and API-based models; agent frameworks (LangGraph, LlamaIndex or equivalent) and the Model Context Protocol (MCP)
- **Model Adaptation & Serving:** LoRA or QLoRA fine-tuning and preference optimisation (DPO, GRPO) using Hugging Face TRL, PEFT or Unsloth; distillation; quantisation; efficient serving (vLLM or SGLang)
- **Deep Learning & Document AI:** PyTorch and Hugging Face Transformers; vision-language models for document parsing (such as PaddleOCR-VL, Docling or the Qwen-VL family); vision models for classification and detection; speech recognition (IndicConformer, Whisper or newer open models)

- **Indic Language AI:** AI4Bharat models (IndicConformer speech recognition for 22 Indian languages, IndicTrans2 translation, IndicBERT, Indic Parler-TTS), Sarvam models and Bhashini services; handling of code-mixed and transliterated text
- **Evaluation & Calibration:** RAGAS, DeepEval, promptfoo or Inspect; retrieval metrics (recall at k, MRR, nDCG); calibration and conformal prediction; SHAP and Fairlearn
- **Data, Engineering & Observability:** advanced SQL, PySpark, Polars or DuckDB; Python with pytest, Git, MLflow, DVC and Docker; LLM tracing and observability (Langfuse or Arize Phoenix, OpenTelemetry)
- **AI Safety & Privacy:** prompt-injection and data-leakage testing, guardrail frameworks (NeMo Guardrails, Llama Guard), anonymisation, working knowledge of differential privacy
- **Communication:** Matplotlib, Plotly, Power BI or Tableau; explaining model limitations to non-technical audiences

General Conditions applicable to all applicants covered under this advertisement

1. Those candidates, who are already in regular or contractual employment under Central / State Government, Public Sector Undertakings or Autonomous Bodies, are expected to apply through proper channel or attach a 'No Objection Certificate' from the employer concerned with the application OR produce No Objection Certificate at the time of interview.
2. NeGD reserves the right to fill all or some or none of the positions advertised without assigning any reason as it deems fit.
3. The positions are purely temporary in nature for the project of NeGD/DIC and the appointees shall not derive any right or claim for permanent appointment at NeGD/DIC or on any vacancies existing or that shall be advertised for recruitment by NeGD in future.
4. NeGD reserves the right to terminate the appointments of all positions with a notice of one month or without any notice by paying one month's salary in lieu of the notice period.
5. The designation of the selected candidates shall be mapped as per the existing designation policy of NeGD.
6. In case of a query, the following officer may be contacted:

HR Team

National e Governance Division, 4th Floor, Electronics Niketan, 6-
CGO, Complex Lodhi Road, New Delhi – 110003
Email: Negdhr@digitalindia.gov.in