



# INDIAN INSTITUTE OF TECHNOLOGY BOMBAY

## MATERIALS MANAGEMENT DIVISION

**Ref No. (PR No.1000049254)**

**(RFQ No. 3000014714)**

### Training an LLM for Indian Banking

#### Scope of Work:

Knowdis will fine-tune an LLM for Indian banking in consultation with IITB faculty and SBI data scientists. The work involves as following:

- Collecting and filtering data that is relevant to Indian finance and legal domain.
- Surveying existing LLMs for financial domain.
- Understanding latest practice of training LLMs, and adopting best practice for this training.
- Pre-training 2--3 LLMs of varying sizes.
- Choosing a relevant set of supervised tasks.
- Evaluating public LLMs and currently trained LLMs on those tasks.
- Task-specific fine-tuning on the tasks.
- Evaluating the trained model.

#### Deliverables:

- Fine-tuned model weights (or API endpoint access if weights are proprietary).
- Evaluation framework + benchmark reports (accuracy, hallucination rate, compliance checks).
- Documentation (architecture, training process, data sources, hyperparameters)

#### SLA:

- An initial 3--8B parameter model pre-trained by month 3
- Fine-tuning on a suite of finance tasks by month 5
- Accuracy on Indian finance tasks significantly better than general-purpose models
- Inference speed comparable to public models of comparable size
- A first release of a trained model within 6 months.
- Project duration 6 months

#### Payment Terms:

- Payment within 30 days from the date of submission of invoice subject to satisfactory certification from the Indentor. Payment will be made on monthly basis.