



Indian Institute of Bank Management

GUWAHATI · ESTABLISHED 1980

DECIDING BEYOND INTUITION

Programme on Data Science and AI/ML Applications in Banking

A model that predicts is analytics, a model that changes a decision is value, and a model the bank can explain is one it can defend.

A three-day classroom programme for bankers who commission, use or oversee models, taught with the Reserve Bank of India, senior practitioners and IIBM faculty.

October 28–30, 2026

Programme Note · IIBM Guwahati

Sponsored by the Reserve Bank of India, NABARD, State Bank of India, Punjab National Bank, Union Bank of India, UCO Bank, Central Bank of India and Indian Bank

CONTENTS

03	Objective
04	Background and rationale
06	Where the readiness gaps cluster
07	What has changed
08	What the Governor asked of banks
09	Faculty and perspective
11	The ground the programme covers
12	The laboratory datasets
14	The algorithms taught, and where
15	Programme schedule
18	How the laboratories and cases are used
19	Programme particulars

PARTICIPANTS

Officers from Compliance, Risk, Credit, Fraud Risk Management, Internal Audit, Treasury, Operations, Digital Banking and Analytics verticals. No coding or statistical background assumed.

DATES AND VENUE

October 28–30, 2026 · IIBM Guwahati

DURATION

Three-day classroom programme · 13 sessions · 3 laboratories

20.8%

of surveyed regulated entities were using or developing AI

127 of 612 entities. Source: RBI FREE-AI Committee Report, August 2025, Department of Supervision survey

67%

expressed interest in exploring AI use cases

The gap between intent and capability is what this programme addresses

15%

of AI-using entities reported using interpretability tools

Only 10% had bias mitigation protocols and 18% maintained audit logs

26

recommendations across 6 pillars and 7 sutras

RBI FREE-AI Committee Report, released 13 August 2025

11

categories of regulated entity covered by the draft Model Risk Management guidance

Draft Guidance on Regulatory Principles for Model Risk Management, 24 June 2026

Rs 48,021cr

amount involved in bank frauds reported in FY 2025-26

Across 10,114 cases, of which advances accounted for Rs 40,774 crore. Source: RBI Annual Report 2025-26

OBJECTIVE

To equip bankers to specify, interrogate, use and govern data science and AI/ML models in the decisions their own function makes.

The programme is designed for the officer who will commission a model, act on its output, audit it or answer for it — not for the specialist who builds it. It assumes no background in programming or statistics, and it begins from the banking decision rather than from the algorithm: whom to lend to and at what price, which alert to investigate first, and which control gap an inspection will find.

Two things make this urgent rather than merely interesting. The first is that adoption in the Indian financial sector is still shallow — roughly a fifth of surveyed regulated entities were using or developing AI at the time of the RBI's FREE-AI survey, concentrated in large institutions and mostly in simple rule-based form. The second is that the supervisory architecture has moved faster than institutional capability: the FREE-AI framework in August 2025, draft Model Risk Management guidance in June 2026 and draft Data Governance guidance in July 2026 together establish an expectation of board-approved model governance, independent validation, explainability and lifecycle control that most banks are not yet staffed to meet.

The sessions move from the statistical and analytical foundations, through the principal banking applications — credit, early warning, fraud and financial crime — into Generative AI, and finally into model risk governance and supervisory expectation, culminating in a group presentation in which each team must present both a model and the control framework around it.

“Not whether the model is accurate, but whether the bank can explain it, monitor it and answer for it.”

The question supervision is moving towards — model governance is now an enterprise obligation, not a technology matter



Khanapara campus, Guwahati

Adoption is concentrated at the top of the system, and thin everywhere else.

The RBI's own survey found that four-fifths of regulated entities were not using AI at all, and that where it was used, it was mostly simple.

Two surveys were administered for the FREE-AI Committee, by the Department of Supervision and the Fintech Department, covering banks, NBFCs, ARCs and All India Financial Institutions representing close to ninety per cent of sector asset size. Of 612 entities surveyed, 127 were either using or developing AI. Adoption was driven by large public and private sector banks and the larger NBFCs, and even there took the form of simpler rule-based models or early-stage exploration rather than production machine learning.

Where adoption sits

ENTITY CATEGORY	REPORTED AI ADOPTION
All surveyed entities	20.8 per cent (127 of 612)
NBFCs	27 per cent of 171 surveyed
Urban co-operative banks — Tier 2 and Tier 3	below 10 per cent
Urban co-operative banks — Tier 1	nil
Asset reconstruction companies	nil

Source: RBI FREE-AI Committee Report, August 2025. Surveyed entities represented approximately 90 per cent of sector asset size.

The gap is widening by ownership as well as by size. A separate Reserve Bank study, applying text-mining to the annual reports of 32 Indian commercial banks, found that **AI-related disclosure in private sector banks rose roughly six-fold between FY 2015-16 and FY 2022-23, against a little over three-fold for public sector banks.**

Source: RBI Bulletin, October 2024, “How Indian Banks are Adopting Artificial Intelligence?” The measure is disclosure intensity rather than deployed capability; the study finds asset size and capital adequacy to be the significant drivers.

What it is being used for

APPLICATION	SHARE OF SURVEYED ENTITIES
Customer support	15.6 per cent
Credit underwriting	13.7 per cent
Sales and marketing	11.8 per cent
Cybersecurity	10.6 per cent

Generative AI interest was strong but largely confined to small-scale pilots, most often internal chatbots. Source: as above.

The exposure the programme exists to reduce

Analytics capability is not an efficiency question alone. The decisions it touches are the same decisions that generate the sector's largest losses. In FY 2025-26 the amount involved in reported bank frauds reached Rs 48,021 crore, the highest in three years, even as the number of cases fell to 10,114 from 23,722. Frauds in the advances category alone accounted for Rs 40,774 crore — 84.9 per cent of the total — while card, internet and digital payment frauds fell away almost entirely, to Rs 29 crore across 293 cases.

REPORTED BANK FRAUDS	FY 2023-24	FY 2024-25	FY 2025-26
Number of cases	35,800	23,722	10,114
Amount involved	Rs 11,013 cr	Rs 32,803 cr	Rs 48,021 cr
Of which, advances	Rs 8,917 cr	Rs 30,367 cr	Rs 40,774 cr
Card, internet and digital payments	Rs 1,452 cr	Rs 517 cr	Rs 29 cr
Public sector banks — amount	Rs 8,092 cr	Rs 23,617 cr	Rs 35,709 cr
Private sector banks — amount	Rs 2,667 cr	Rs 8,927 cr	Rs 11,399 cr

Source: RBI Annual Report 2025-26, on which basis the earlier years are restated. Covers frauds of Rs 1 lakh and above reported during the year, which may have occurred earlier. Rs 30,199 crore of the FY 2025-26 figure comes from 314 legacy cases reported afresh following the Supreme Court judgement of 27 March 2023; the concentration in the loan book does not.

Two movements in that table matter more than the headline. First, **public sector banks now carry close to three-quarters of the value — Rs 35,709 crore of Rs 48,021 crore — from 5,418 cases**, against Rs 11,399 crore for private sector banks. Second, the losses have moved decisively out of the payment rails and into the loan book: digital payment fraud has fallen from Rs 1,452 crore to Rs 29 crore in two years, while advances have risen more than fourfold. The analytics burden has moved with it, from transaction surveillance alone towards underwriting quality, early warning and post-sanction monitoring.

The two failures this creates work for

Credit decisions and financial crime detection are precisely the domains in which analytics is claimed to add most value, and precisely the domains in which the sector's losses are concentrated. A bank that cannot build, read and challenge an early warning model is dependent on a vendor's assertion that one works. A bank whose transaction monitoring generates alerts it cannot risk-rank is spending investigator time in proportion to alert volume rather than to risk.

The problem is rarely that the model does not exist. It is that no one in the room can say what it is doing, on what data, or when it last worked.

Why the laboratories in this programme are built rather than demonstrated

Among the entities that do use AI, the governance safeguards are the weakest part of the picture.

The FREE-AI surveys asked the 127 entities reporting AI use what controls they had around it. The answers, set out below in the order of severity, define the syllabus of the governance sessions and are the heads against which participants will be examined in the group exercise.

SAFEGUARD	ENTITIES REPORTING IT
Real-time model performance monitoring Detecting degradation while the model is still making decisions, rather than at the next review	14 per cent
Interpretability and explanation tools The ability to state why a particular customer was declined, flagged or scored as they were	15 per cent
Bias mitigation protocols Acting on unfair outcomes, as distinct from testing for them	10 per cent
Audit logs A reconstructable record of what the model decided and on what inputs	18 per cent
Data or model drift monitoring Detecting that the population has moved away from the one the model was trained on	21 per cent
Validation for bias and fairness Reported mostly at early development stages only	35 per cent
Board-level oversight of AI Some form of board oversight, concentrated in large banks	one-third

Source: RBI FREE-AI Committee Report, August 2025. Most entities had no dedicated policy for training AI models and relied on existing IT, cybersecurity and privacy policies.

Read against the draft Model Risk Management guidance of June 2026 — which contemplates a board-approved framework, an enterprise model inventory, risk-based tiering, independent validation and continuous monitoring — these figures describe a capability gap rather than a compliance lapse. Closing it is a training problem before it is a technology problem.

Four developments in twelve months have moved model governance from good practice to regulatory expectation.

1

The FREE-AI framework was published

13 AUGUST 2025

The Committee chaired by Dr Pushpak Bhattacharyya set out seven sutras — trust as the foundation, people first, innovation over restraint, fairness and equity, accountability, understandable by design, and safety, resilience and sustainability — organised into two sub-frameworks and six pillars, carrying twenty-six recommendations with timelines and action owners. The report is advisory, but it explicitly anticipates conversion into binding instruments.

2

Draft Model Risk Management guidance was issued

24 JUNE 2026 · COMMENTS CLOSED 24 JULY 2026

The draft applies to eleven categories of regulated entity and, once final, would replace the 2002 guidance on credit risk models. Its definition of a model is deliberately broad, extending to statistical models, scoring algorithms, business rule engines, AI/ML and Generative AI systems, third-party models and material spreadsheets that influence business decisions. It contemplates a board-approved framework, an enterprise model inventory as single source of truth, risk-based tiering with Risk Management Committee approval for high-risk models, independent validation, and continued accountability for outsourced and vendor models.

3

Draft Data Governance guidance followed

15 JULY 2026 · COMMENTS CLOSED 17 AUGUST 2026

Drawing on BCBS 239, the draft requires a board-accountable Data Governance Framework covering the full data lifecycle, with defined roles, single-source-of-truth architecture, metadata and lineage management, data classification, measurable data quality processes and controls over third-party data sharing. Model quality is downstream of data quality, and the two drafts are best read together.

4

The loss profile shifted to the loan book

FY 2025-26

Fraud amount reached Rs 48,021 crore, the highest in three years, even as case numbers more than halved to 10,114. Advances accounted for Rs 40,774 crore — 84.9 per cent of the total — while card, internet and digital payment fraud fell away to Rs 29 crore. Public sector banks carried Rs 35,709 crore of the amount. The analytics burden has moved correspondingly from transaction surveillance alone towards underwriting quality, early warning and post-sanction monitoring.

Both drafts were open for public consultation during 2026 and final instruments had not been issued at the time this note was prepared. Programme content is updated to the position obtaining at the date of delivery.

What the Governor asked of banks — and what this programme does about it.

Delivering the inaugural address at FIBAC 2026 in Mumbai, Governor Shri Sanjay Malhotra placed artificial intelligence at the centre of the decade, and accountability at the centre of artificial intelligence.

He described AI as the force that will define this decade of Indian banking as decisively as liberalisation defined the 1990s and digitalisation the 2010s, and framed the central bank's position as treating it a capability to be harnessed responsibly rather than merely a risk to be contained. Earlier technologies multiplied energy and connectivity; this one multiplies intelligence, and so reaches directly into how a bank assesses risk, serves customers, prices capital and organises itself. The choice he put to the industry was whether a bank shapes its own AI journey, or lets that journey shape the bank by default.

“The model decided” can never be an acceptable answer

The Governor, on responsibility for an AI-driven lending decision — to a customer, an auditor or the Reserve Bank

He then set out the risks in seven heads: explainability of black-box decisions; bias and exclusion inherited from historical lending data; concentration in a narrow set of models and vendors; third-party dependency; data privacy; cyber and adversarial vulnerability; and the erosion of human accountability. Final responsibility must remain with the bank rather than a vendor or an algorithm, and fairness is a design requirement rather than a compliance checkbox. Success will belong not to the institutions that adopt most aggressively, but to those that understand what they have deployed and can answer for it.

What he asked banks to do, and where the programme teaches it

Maintain a complete inventory of AI and models in use	Day 3, model risk governance — the enterprise model inventory as single source of truth
Adopt board-approved policies with clear accountability	Day 3 governance session, and the closing group presentation
Be able to explain a lending or fraud decision	Day 2 builds a scorecard where every point traces to a named bin; Day 3 covers explainability and its limits
Stress-test models before and after deployment	Laboratory 2 stresses the scorecard against a shifted population; Day 3 covers drift monitoring
Keep a human in the loop where failure would harm	Day 2 on alert triage, and Day 3 on the failure modes of generative systems

Inaugural address at FIBAC 2026, organised by FICCI and the Indian Banks' Association, Mumbai, August 2026. Paraphrased from the address as reported; the full text is on the Reserve Bank's website.

The method, the application and the supervisory test — in one classroom

The programme is delivered jointly by the regulator, senior practitioners from banks' analytics and risk functions, and IIBM faculty.

The distinguishing feature is that the same model is examined from three positions in a single room: participants build it themselves in the laboratory, hear from practitioners what happened when an equivalent model met production data at scale, and then hear from the supervisor what would be asked of it in an inspection. A model that survives all three readings is one the participant can take back and defend internally.

Sessions are sequenced so that no application is taught before the data it depends on, and no model is taught before the governance that would be applied to it. The laboratories are cumulative — the dataset assembled on the first day is the dataset scored, monitored and validated for the remainder of the programme.



Classroom programme · IIBM Guwahati

Faculty and the perspective each brings

FACULTY	SESSIONS LED	PERSPECTIVE OFFERED
Reserve Bank of India Department of Supervision, Department of Regulation and Fintech Department	Foundations and the AI landscape; model risk governance and the FREE-AI framework; supervisory expectations	The regulator's own reading of what constitutes an adequate model control, how model findings are framed in an inspection, and what remediation is regarded as sufficient
Bank credit risk units	Credit risk analytics and the anatomy of a scorecard	How model output is actually used at the sanctioning desk, and where analytical and judgemental decisions part company
Bank risk and fraud risk management	Early warning signals and collections analytics; fraud analytics and mule account detection	Implementation reality at scale — what reached production, what did not, and the actual false positive burden carried by the investigation team
Industry and academia Technology practitioners and academic faculty	Generative AI and large language models in banking	What the technology can and cannot do, its failure modes in a regulated environment, and an independent view free of vendor interest
IIBM faculty	Statistical and analytical foundations; all three laboratories; facilitation and assessment	Continuity across the three days, structuring and facilitation, and drawing out the supervisory and business learning from each laboratory

Session-wise faculty allocation is confirmed once the programme dates are agreed with the participating institutions. Where external faculty are unavailable for a session, it is delivered by IIBM faculty without change to the syllabus.

A note on the teaching approach

Half of the engagement time is spent with participants working on data rather than listening to it being described. This is a deliberate choice: the capability the sector reports lacking is not awareness of what AI can do, which is widespread, but the ability to specify a problem, judge a dataset, read a performance measure and identify the point at which a model should not be trusted. That capability is not transferable by lecture.

Term size reflects the classroom time each idea receives, not its importance in the abstract.



13

3

3

synthetic banking datasets, built at IIBM

Three banking datasets, built at IIBM, with the answer concealed inside each one.

All data is synthetic. No record derives from any bank's live systems and no row corresponds to a real customer, account or transaction.

Generated programmatically at IIBM, they can be circulated freely, reused across cohorts and given to participants to take away. Each is built so the signal a competent team should find is genuinely present and genuinely recoverable — but not so cleanly that the exercise is trivial — and each carries deliberate traps of the kind that defeat real projects.

1 · Retail credit risk — Laboratories 1 and 2

Size	13,500 applications × 41 fields, plus a 1,235-row defective extract for the data quality audit
Contents	Seven retail and MSME products; applicant demographics and employment, facility structure, bureau attributes, banking behaviour, and a full twelve-month performance window
Target	90+ days past due within 12 months on book. Bad rate 10.95 per cent
Traps built in	2,739 rejected applications with no performance, forcing the reject inference question; bureau score blank for 11.5 per cent new-to-credit applicants, which is a population and not a gap; and two post-decision columns that will inflate any model that leaks them

2 · Fraud analytics and mule account detection — Laboratory 3

Size	Two linked files — 3,004 savings accounts × 17 fields, and 117,940 transactions × 12 fields over 92 days
Contents	Account opening dates, KYC mode, existing risk categorisation, declared occupation and income, address pincode and registered device; and a transaction log carrying channel, direction, amount, counterparty, device, IP address and location
Concealed inside	64 mule and collector accounts across 5 rings — 2.13 per cent of the book — carrying the live signature: opened in a tight window through OTP-based eKYC by young low-income customers, dormant, then a burst of inbound credits from many unrelated payers in amounts below reporting thresholds, at night, passed to a collector within hours and cashed out through ATMs in another state
The point	The bank's own KYC risk grading rates almost every one of them Low. Behavioural monitoring has to find what the categorisation missed

3 · Customer profiling — the closing group exercise

Size	15,000 liability customers × 38 fields
Contents	Four quarters of average balances, product holdings across seven categories, digital and branch channel engagement, salary and pension credits, complaints and service history, and compliance hygiene
Two targets	Six behavioural archetypes held back for the clustering exercise, and attrition in the following quarter at a 9.13 per cent base rate
Trap built in	One segment follows an agricultural cash cycle — high, lean, lean, harvest peak. Any team that triggers on a falling balance will misclassify it entirely and send retention offers to customers who were never leaving. Seasonality is not decline

Ground truth — the mule flags, the ring identifiers and the true segment labels — is withheld until the evaluation step, so that teams score their own work against something they have not seen. The notebooks quarantine these columns automatically, so a participant working from the supplied file cannot use them by accident.

How the datasets are prepared and maintained

Generation	Written and seeded at IIBM. Changing the seed produces a statistically equivalent but different pack, so that worked solutions do not circulate between batches. Each generator prints a signal check on completion — bad rate and baseline Gini, mule precision at a given queue depth, clustering agreement — confirming a regenerated pack still teaches what it should before it reaches a classroom
Formats supplied	CSV for the laboratory environment, and a workbook holding the data dictionary and a sample of each file for participants working in a spreadsheet or in the event of a connectivity failure
Notebooks	Supplied in both R and Python, with identical steps and identical conclusions. The R notebooks run on a stock installation with no package downloads, which matters on a locked-down bank laptop
Taken away	Datasets, completed notebooks with worked solutions, a model documentation template, a model validation checklist and a use-case prioritisation matrix

Benchmark results are recorded for each dataset so that faculty can tell quickly whether a team has gone wrong. A result materially better than the benchmark almost always indicates leakage rather than insight, and is the more instructive finding of the two.

Every method is introduced through the banking decision it serves, and judged against what that decision costs if it is wrong.

Performance is assessed throughout on KS, Gini, ROC-AUC, precision and recall at a fixed alert budget, and the information criteria. The programme does not teach the newest method, but the one that matches the consequence of the decision. R is the working environment throughout.

METHOD	APPLIED TO	WHY THIS METHOD, AND WHAT IT TEACHES
Descriptive and inferential statistics data types, cross-section, panel and time series	Day 1 foundations	What kind of data a bank actually holds, and which questions each kind can and cannot answer. The prerequisite for everything that follows
Supervised and unsupervised learning classification, regression, clustering	Day 1 foundations	The map of methods, anchored to banking examples, so that participants can place any vendor claim on it
Weight of Evidence binning and Information Value	Variable selection for the credit scorecard	Handles missing values as a category, linearises the relationship with the log-odds and caps outliers. The IV table also shows how little independent signal geography and demographics carry — a fair-lending result as much as a statistical one
Logistic regression	Application and behavioural scorecards; probability of default	Still preferred for regulated credit decisioning: every point awarded traces to a named bin of a named variable, so the bank can tell a declined applicant why and an inspecting officer how
Anomaly and outlier detection unsupervised, on transaction streams	Ranking savings accounts by behavioural abnormality	Needs no labels — mule typologies mutate faster than a supervised model can be retrained. Its contamination parameter is a capacity decision disguised as a hyperparameter
Graph and network analytics connected components, shared-attribute edges	Identifying mule rings and the collector account within each	Anomaly detection produces a ranked list of suspects; the graph produces a case, with a money trail and an articulable relationship
Deep learning and natural language processing	Day 1 landscape; Day 3 generative systems	What these families do, where banks are actually using them, and why they are harder to govern than a scorecard
Large language models with retrieval-augmented generation	Sourced search over circulars; document extraction; drafting	Grounding an answer in a retrieved passage rather than in the model's own recall. Participants deliberately induce a wrong answer and design the control that would have caught it
Explainability techniques variable contribution and permutation importance	Interrogating the models built in the laboratories	Separating predictive drivers from actionable ones. A model whose signal is all of the first kind forecasts accurately and changes nothing
Drift and stability monitoring population stability index, KS on scored populations	Post-deployment monitoring	Detecting that the population has moved away from the one the model was trained on — the safeguard only 21 per cent of AI-using entities reported having

PROGRAMME SCHEDULE

October 28–30, 2026 · IIBM Guwahati. Tentative; the laboratory sequence should not be reordered.

The dataset assembled in Laboratory 1 is used in every subsequent laboratory. Where a guest faculty is unavailable, the session is delivered by IIBM faculty without change to coverage.

DAY 1 — WEDNESDAY, 28 OCTOBER · FOUNDATIONS AND THE FIRST MODEL		
TIME	SESSION AND COVERAGE	FACULTY
SESSION 01 1 session (1.30 hour)	Basic concepts of statistics Different types of data — cross-section, panel and time series, and what each can and cannot be asked. What AI, machine learning, deep learning and natural language processing actually are, and where each is applied in banks. Theoretical foundations for everything that follows.	IIBM / RBI
SESSION 02 1 session (1.30 hour)	Why data-driven decision making, and why now A functional map of analytics use cases across credit, deposits, fraud and financial crime, and which models the programme will use. A working understanding of the AI/ML techniques and algorithms to be applied, and the basic foundations of R . Analytical foundations for bankers — no mathematics assumed Descriptive, diagnostic, predictive and prescriptive analytics through banking questions. Supervised and unsupervised learning; classification, regression, clustering and time series. The model lifecycle from problem framing to retirement. Training, validation and test sets, and why a model that performs well in development fails in production — overfitting, target leakage and population drift. Reading performance honestly: confusion matrix, precision and recall, ROC-AUC, KS, Gini and the information criteria.	IIBM
SESSION 03–04 2 sessions (1.30 – 2 hour)	Building your own model Laboratory 1 — from raw records to a decision-ready customer view Feature engineering, the least glamorous and most valuable skill in banking analytics. Teams build a customer view from four separate files and derive behavioural features: balance volatility, salary credit regularity, instalment bounce count, channel mix, days since last branch contact, bureau vintage and enquiry velocity. Output: the base dataset used in every subsequent laboratory. Worked solution supplied.	IIBM

DAY 2 — THURSDAY, 29 OCTOBER · CREDIT, EARLY WARNING AND FINANCIAL CRIME

TIME	SESSION AND COVERAGE	FACULTY
SESSION 05–06 2 sessions (1 hour each) OR 1 session (1.30 – 2 hour)	Credit risk analytics — the anatomy of a scorecard Application and behavioural scorecards. Defining the target, the performance window, sampling, exclusions and reject inference. Weight of Evidence and Information Value for variable selection. Logistic regression and why regulated credit decisioning still prefers it to more complex methods. Converting output into score bands and probability of default. Setting cut-offs — the trade-off between approval rate, expected loss and business volume — and using the same model for risk-based pricing and straight-through processing.	IIBM / Bank credit risk
SESSION 07 1 session (1.30 – 2 hour)	Laboratory 2 — build, validate and defend a retail scorecard Teams bin the variables, compute information values, fit the model, produce the scorecard points table and measure KS and Gini on a holdout sample. They then stress it: what happens to performance if the geography mix shifts, or if bureau data is unavailable for a segment. Assessment: each team defends its scorecard before a mock credit committee that questions the variable choices, the cut-off and the exclusions.	IIBM
SESSION 08 1 session (1.30 – 2 hour)	Early warning signals, delinquency prediction and collections analytics Designing an EWS framework that combines rule-based triggers with a predictive model. Behavioural indicators of incipient stress in retail and MSME accounts — utilisation ramp, cheque returns, GST filing gaps, current account churn and interrupted salary credits. Predicting slippage into SMA-1 and SMA-2 and prioritising limited collections capacity by expected recovery rather than by outstanding amount. Vintage curves, roll-rate matrices, concentration and scenario-based stress.	Bank risk / IIBM
SESSION 09–10 2 sessions (1.30 hour each)	Fraud analytics, financial crime and mule account detection Why static rules generate high false positives and miss adaptive fraud. Anomaly and outlier detection on transaction streams. Graph and network analytics to surface mule rings — shared devices and IP addresses, beneficiary overlap, layering chains and pass-through velocity. Risk-scoring alerts so that investigator effort follows risk, and tuning thresholds against alert-to-report conversion. Laboratory 3: run an anomaly detection routine on a savings transaction log seeded with mule patterns, draw the beneficiary network to identify the ring, and write the red flags that would support an escalation.	Bank fraud risk / IIBM

DAY 3 — FRIDAY, 30 OCTOBER · GENERATIVE AI, GOVERNANCE AND PRESENTATION

TIME	SESSION AND COVERAGE	FACULTY
SESSION 11 1 session (1.30 – 2 hour)	Generative AI and large language models in banking What a large language model is, and more importantly what it is not. Retrieval-augmented search over Master Directions, internal circulars and product manuals so that staff receive a sourced answer rather than a plausible one; extraction and verification of KYC and financial documents; drafting of credit appraisal notes and audit observations; summarisation of investigation files. The failure modes that matter in a regulated environment — hallucination, prompt injection, leakage of customer data to third-party models and non-reproducibility of output. Demonstration: participants deliberately induce a wrong answer and design the control that would have caught it.	Industry
SESSION 12 1 session (1.30 – 2 hour)	Model risk governance, responsible AI and supervisory expectations The seven sutras and six pillars in operational terms. Enterprise model inventory, risk-based tiering, independent validation, approval, monitoring, periodic review and retirement. Explainability techniques and their limits. Testing for bias and adverse impact in lending. Three Lines of Defence applied to models; Board and Senior Management oversight; incident reporting; audit of AI systems. Third-party and vendor model accountability, which outsourcing does not transfer. Obligations under the DPDP Act, 2023. Exercise: teams complete a model validation checklist against the scorecard built in Laboratory 2 and identify its three weakest control points.	RBI / IIBM
SESSION 13 1 session (2 hours)	Presentation from groups Each team presents the model it has built and the control framework around it, before faculty acting as the risk management committee of a board. Teams are questioned on their assumptions rather than their arithmetic.	IIBM

Session numbering is for reference and does not fix the order of delivery. Where two options are shown, the programme committee may run either two sessions of one hour or a single session of ninety minutes to two hours.

One dataset, carried across three days, examined from a different position each time.

Participants do not see three unrelated demonstrations. They build one customer dataset on the first day and then spend the rest of the programme scoring it, monitoring it, breaking it and validating it.

Before the programme

Two short cases and a one-page glossary are circulated a week in advance. At the time of nomination each participant is asked to submit one decision in their vertical that could be improved with data; these submissions shape the closing group exercise, so it is drawn from the room rather than from a textbook.

During the laboratories

Participants work in pairs in a laboratory environment. Every laboratory has a worked solution, and a spreadsheet version of each exercise is held in reserve against connectivity failure. Datasets are synthetic or fully anonymised, generated at IIBM, and are not drawn from any participant bank's live systems.

At the close

The programme culminates in the group-based exercise, requiring participants to integrate the concepts covered during the three days and present their assessment, observations and recommended actions to faculty acting as a board risk management committee.

What each group is asked to produce

- » The business problem stated as a decision, with the measure by which success would be recognised
- » The data required, its likely quality defects, and the lawful basis for using it
- » The analytical method proposed, and why a simpler method was rejected
- » Model performance expressed in terms a business committee can act on, not only in technical measures
- » The validation, monitoring and explainability approach, mapped to the three lines of defence
- » The failure modes, and what would be done when — not if — the model degrades
- » Implementation path, indicative cost and expected benefit, stated as a business case

What participants take away, and how it is assessed

Each participant leaves with the datasets and completed notebooks including worked solutions, a model documentation template, a model validation checklist, and a use-case prioritisation matrix for their own vertical. A short diagnostic at the start of Day 1 and again at the close of Day 3 gives the Institute a measurable learning gain, and a completion certificate is issued subject to full attendance and participation in the group exercise. Participants are invited to submit a one-page analytics proposal for their own function within thirty days, with a brief online review session at the sixty-day mark to sustain momentum and give the Institute evidence of application on the job.

PROGRAMME PARTICULARS

Programme on Data Science and AI/ML Applications in Banking

Participants

Officers from Compliance, Risk, Credit, Fraud Risk Management, Internal Audit, Treasury, Operations, Digital Banking and Analytics verticals

Venue

IIBM Guwahati

Duration

Three-day classroom programme · 13 sessions
· 3 laboratories

Prerequisite

None — no coding or statistical background assumed. Participants to carry a laptop

Date

October 28–30, 2026

Coordinators

Shri Arindam Chakraborty, MoF · +91 93118 57688

Ms Monisha Thapa · +91 93659 73843

Nominations to programme@iibm.ac.in

FEE:

All-inclusive for the entire programme — programme fee, accommodation, food and course material.

Sponsor Banks/ institutions of IIBM viz. RBI, NABARD, SBI, UCO Bank, Union

Bank of India, Indian Bank, Central Bank of India and Punjab National Bank: **Rs. 15,930/-**

Non-Sponsored Banks (Commercial Banks, SFBs, FIs etc.): **Rs. 16,992/-**

RRBs & Co-operatives: **Rs. 12,390/-**

Urban Co-operative Banks: **NIL (NER Region); Others: Rs. 6,195/-**



Indian Institute of Bank Management

Jawahar Nagar, Khanapara, Guwahati 781022, Assam

iibm.ac.in · Programme on Data Science and AI/ML Applications in Banking — Programme Note