AI and Gen AI developments in Credit Risk Management

Matthew See,
MD, Head of Risk and Business Analytics, DBS

Shifts towards adopting digitalisation and AI in credit risk management…

Wave 1: Process Optimisation

- **Digitising workflows**, e.g. STP decisioning, digital CX/EX
- **Automation**, e.g. RPA
- **Scorecard/model adoption,** e.g. decisioning, risk monitoring

Wave 2: AI-based Capabilities

- **ML techniques**, e.g. random forest, XGBoost, neural networks
- **Alternative data**, e.g. transaction, merchant, e-commerce
- **Early warning focus**, e.g. network analytics, sentiment analysis

Wave 3: Gen AI Applications

- **Credit analysis**, e.g. memo preparation, ESG
- **Documentary checks**, e.g. policy queries, facility agreement review
- **Others**, e.g. collections, risk surveillance

… comes with attendant challenges to navigate

Post-Covid Changes

- Behavioural, e.g. digital consumption, bureau
- Delinquency, e.g. “cliff” effect post govt support

Regulatory Developments

- Data localisation
- Governance on responsible data usage, AI/ML, GenAI

Internal Investments

- Process
- Data/Tech
- People

Data-Driven

DATA FIRST

MAXIMISING VALUE FROM DATA

Models to predict / detect

Insights / Evidence to improve

Analytics Capability

Accelerating and industrializaing A.I.

Culture & Curriculum

Building a data-driven mindset

Enable Data Usage

Enabling right data to the right people at the right time.

Tech Platform

Providing analytical data science and business intelligence (BI) capabilities under one platform

AI Industrialisation:
Transforming Tech, Process and People

ADA: Single data platform across the bank

- Modern hybrid architecture with advanced analytics capabilities
- Data management principles of discoverability and lineage

Common Processes *supporting rapid builds*

- ALAN: Central knowledge repository for all AI/ML use cases supporting rapid model builds
- Responsible Use of Data / PURE

Data Chapter: *One Team of Data Professionals*

- Comprehensive Employee Value Proposition for data professionals (700+ with 200 data scientists)
- Data staff federated across BUs/ SUs, working in MTJs

- **800+ Models in 350+ Use Cases**
- **Reduced Time to Value[1] from AI/ ML by >80% from 18 months to 2-3 months by end 2023**
- **> 9,000 staff trained since 2021**
- **Economic Outcomes[2] – 2022: $178million 2023: $370million**

[1] *Time to Value: end-to-end time to develop and deploy models*
[2] *Economic Outcomes: Incremental revenue, credit/ fraud losses saves and productivity gains*
[3] *Full year 2023 forecast*

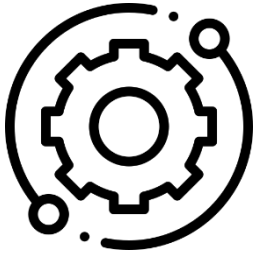
Culture & Curriculum: Over 9000 employees trained in our DataFirst curriculum since 2021

NON-TECHNICAL CURRICULUM

Focused on building a data-driven mindset, frame a problem and shape a solution

Audience

All Staff



TECHNICAL CURRICULUM

Hands-on training in advanced analytics, data science, data engineering & machine learning

Audience

Data Analyst

Data Scientist

Machine Learning Engineer

Some of our culture initiatives to drive data-driven mindset

Train 40,000 to understand data as a strategic asset and how to get started

Train leaders to understand importance of data and analytics to drive value

Create Data Heroes
to upskill employees with big data and data analytics skills

Bringing data into decision making processes

Making product and customer experience design data-driven

Building out data analytics & governance specialisation roles

Enable Data Usage: Responsible Data Use (RDU) is a focus areas for Data Management in DBS

Business Opportunities

Emerging Risks
Increased Regulatory Focus

5 KEY FOCUS AREAS FOR DATA MANAGEMENT

1

DATA ACCESS & SECURITY

Increase accessibility across the bank, whilst protecting it from theft and misuse

2

DATA OWNERSHIP

Establish clear data ownership to ensure we treat data as an enterprise asset

3

DATA UNDERSTANDING

Develop shared understanding of data assets to increase its discovery, utility and value

4

DATA QUALITY

Improve the quality and integrity of our data so we can use it with confidence

5

RESPONSIBLE DATA USE

PURE

Maximise value we can derive from data whilst ensuring the trust of our stakeholders

AI ML Governance

Scaling analytics and AI while building confidence and trust in model outputs and business insights

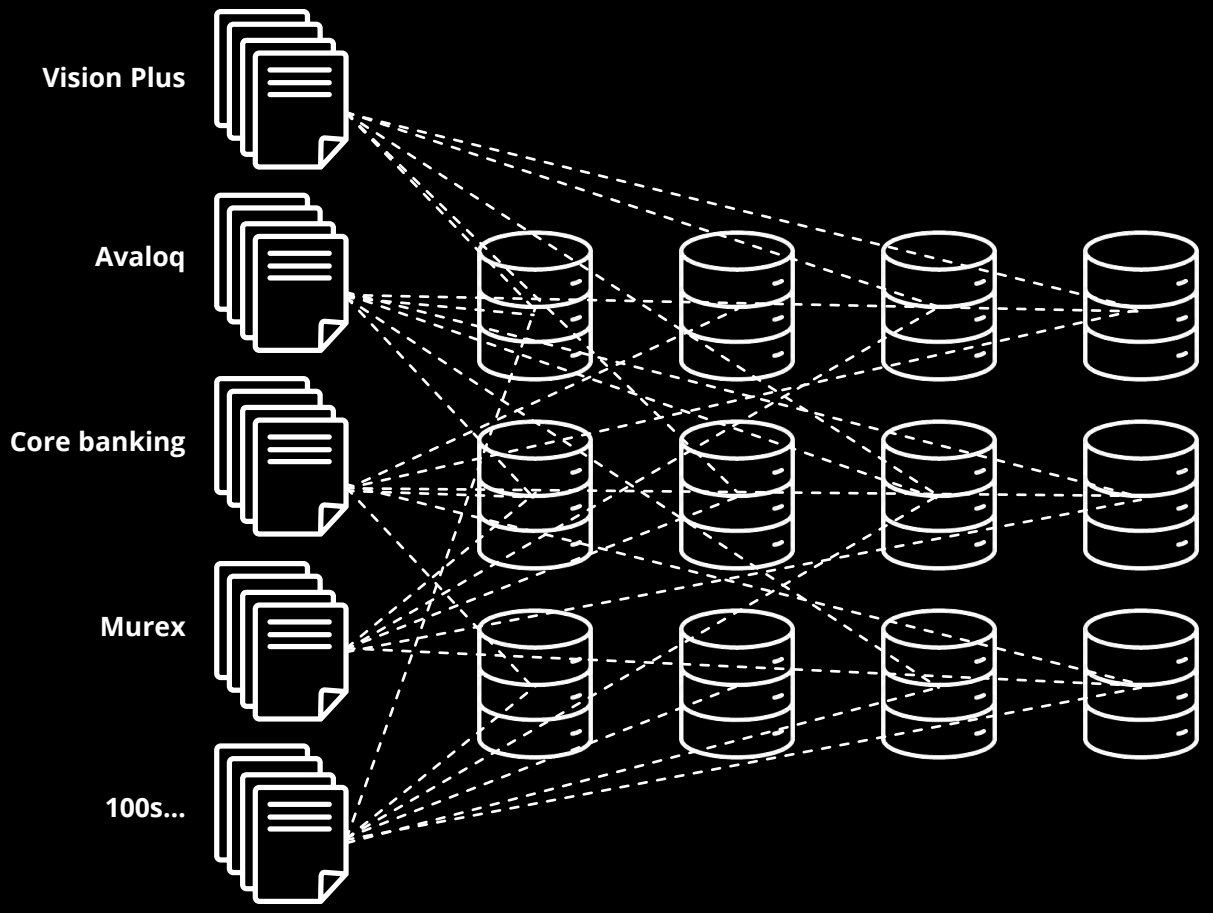
Provide guiding principles for Responsible Data Use

Tech Platform: Build the required features in our data tech platform

FROM	TO
Multiple & Siloed Data Marts	Unified data platform
Inability to handle large volume of Data	Scalable infrastructure
Long time to access data	Role based access
Tedious data discoverability	Services based data discoverability
Limited metadata information	Metadata driven ingestion
Limited support for analytics tools	Flexibility on analytics tools

Our data platform transforms the bank's data into a competitive differentiator

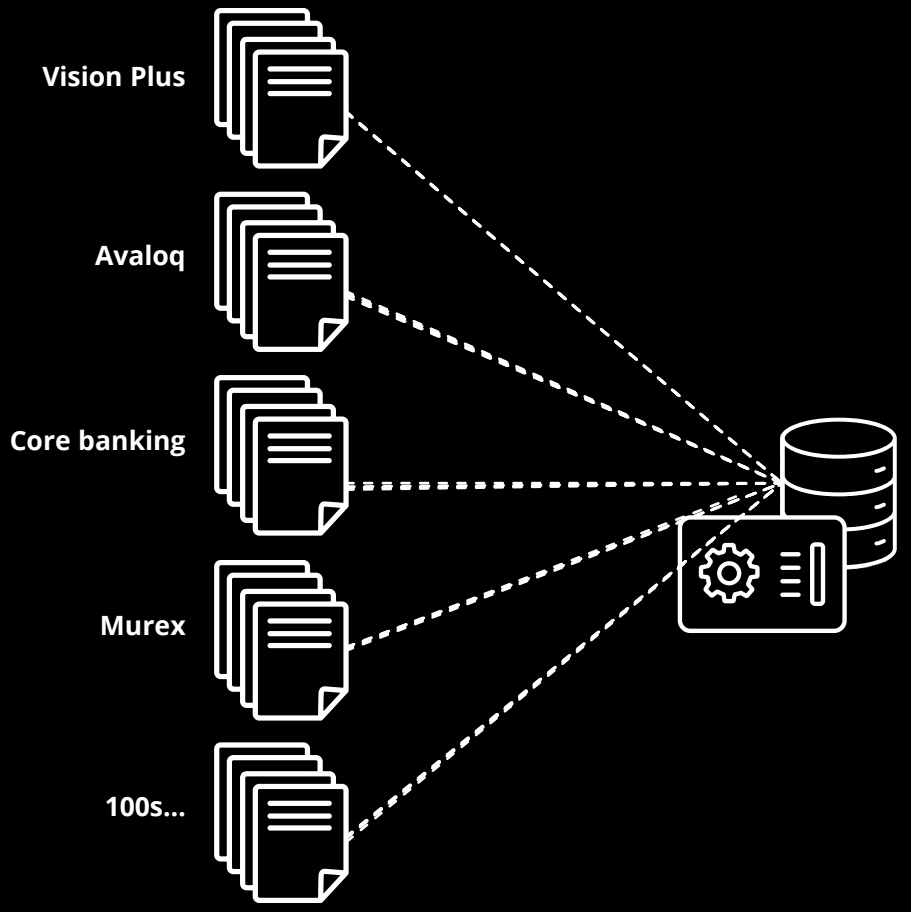
All data from Single Source of Truth



From 12 Data Warehouses
to a Single Source of Truth

Our data platform transforms the bank's data into a competitive differentiator

All data from Single Source of Truth



Volume
- 5.3 PB of data,
- 31k data sets
- Customer360 with 16,000 attributes

Variety
- Structured
- Text
- Video
- Logs

A Truly Big Data Platform

Velocity
- Batch
- Near real time
- Real time
- Streaming

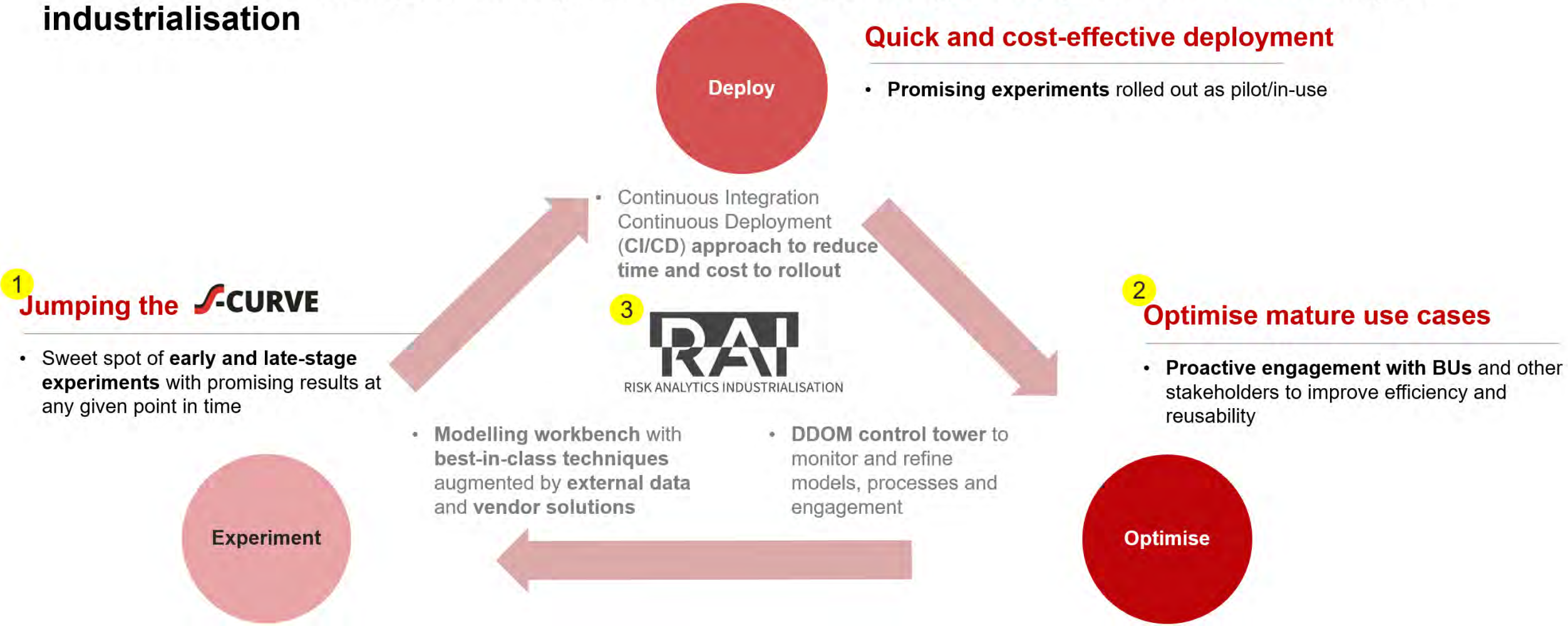
Pervasive adoption across the Bank

All Business Units and Markets

22,000
jobs

2,500
Data Professionals

Achieving scale and speed is key to deepening risk analytics capabilities enabled by AI industrialisation



Enablers

Data Architecture + **Model Governance**

Holistic, robust, promoting innovation and time-to-market

- All RMG models under risk governance with independent 2nd line of defence
- Address evolving Fairness, Ethics, Accountability, Transparency developments

AI / GenAI applications in credit risk management

Pre-underwriting → **Post-underwriting**

	Credit Assessment			Documentary Processes		Surveillance	Fulfilment
	Customer Targeting	**KYC/Credit Fraud**	**Credit Decisioning**	**Contract Generation**	**Documentary Reviews**	**Risk / Portfolio Monitoring**	**Remediation / Collections**
AI techniques	• Product propensity e.g. marketing nudges, supplier/buy network, • Credit pre-screening e.g. pre-emptive lending terms	• Rule-based triggers optimised with scorecards • Relational/transactional network analysis • Doc forgery, e.g. bank statements	• Application / Behavioural scorecards • Alternative data sources via aggregators / ecosystem players • Expert to statistical approach, risk grading	• Data extraction from docs (e.g. financial statements) using OCR and NLP • Analysis based on financial spreading using knowledge graphs and NLG • RPA to Intelligent Doc Processing (IDP)		• Reuse of behavioural scorecards • ML early warning system using network and news sentiment analysis • Portfolio stress testing incorporating macro-financials transmissions	• Collection scorecards • Recovery models • Channel/Timing recommendation engines • Skip-tracing leveraging triangulation and network analysis

Use Case on Credit Scoring: Using alternative data to enable underwriting to "underserved" segments

Problem

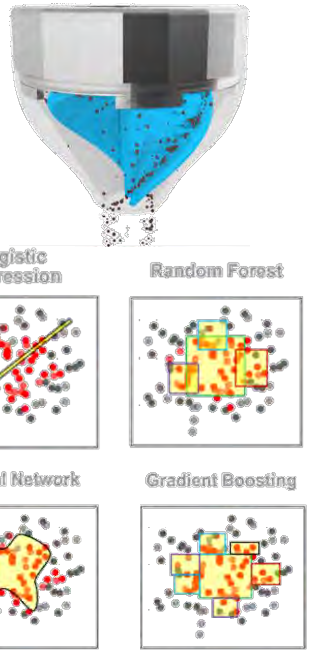
Large pool of underserved customers that do not have a presence on credit bureau

Can we develop alternate credit evaluation based on ecosystem partnership?

Approach

- Develop credit scorecard using alternative data
- ~500 input variables were considered

Analysis



Machine Learning models using techniques such as random forest, NN etc.

Outcome

Based on the model, there is a **significant lift** in the addressable customer base that is potentially attractive to us

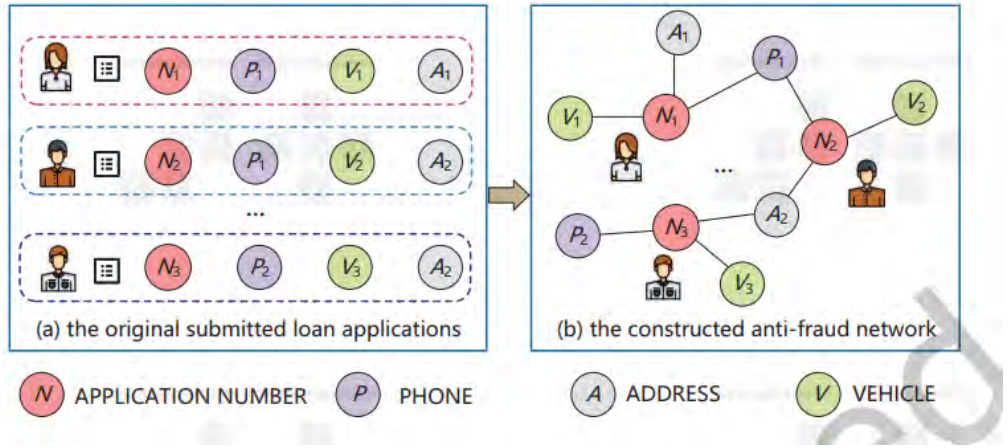
Use Case on Credit Fraud: Using network analysis to identify fraudulent credit applications

Problem Statement

- Current fraud databases and processes are **rule-based** and have a **narrow focus on individual assessment**
- Analysis on relational linkages are **limited to direct linkages** only
- Clustering analysis is done **manually** and painstakingly, which is not comprehensive nor efficient

Network Analysis

- Incorporate **broader associations** e.g. consumer and corporate
- Reflect **2nd degree connections** and **optimise flagging** across various associations
- **Automate clustering** for comprehensive and efficient analysis



Use Case on Doc Forgery Detection

Problem Statement

1. **Surge** of forged documents in loan applications
   - **Increased sophistication** aided by AI advancement

2. **Manual checks** are inefficient/ineffective
   - **Time-consuming** particularly with huge volumes
   - Require **well-trained personnel** to detect signs of forgery
   - **Inconsistency** and **complacency** over time

3. Loops back to borrower for **live verification**
   - Creates **friction** on customer journey
   - **Additional toil** on bank employees from outreaches
   - False positives may lead to **drop-offs** for genuine applications

Advertisement by loan syndicate

Advertisement on fake banks statements

Use Case on Early Warning: Using alternate data and multi-modal approach to early detect borrowers with deteriorating credit quality

Problem

Highlight **emerging credit risk** for portfolio reviews using alternate data

Develop an **automated method of reviewing portfolios** so that precious man hours can be freed up

Approach

Several data sources considered

- Financial statements
- Economic variables
- Industry drivers
- News data
- Satellite data
- Internal data (Payment)

Analysis

Developed **ensemble model** to predict impact of early signals on Industry growth rate

Early Signals

Economic variables | Satellite data | News data

Model 1 | Model 2 | Model 3

ENSEMBLE

Industry sales Growth forecast

- Simulate impact on company Ratios & Ratings
- Prioritise portfolio action

Outcome

Reduce bank losses through early detection

Manpower and time savings in Portfolio review

Forward looking portfolio shaping

Use Case on Early Warning (2/2): SME models through Covid

- Developed **forward-looking** monitoring models in **Singapore** leveraging AI/machine-learning techniques, high freq transactional data and big data
- Performance of both models shows that the algo was able to predict NPAs as early as 3 months before NPA

Data Challenges

1. **Default data** has yet to fully manifest
2. Public **bureau** & **loan repayment** data is distorted due to **moratorium effect**
3. **Poor financials** due to **lockdown**

Train monitoring models in an agile and iterative manner (refining against call feedback and early delinquency data) to mitigate credit losses via early identification and intervention

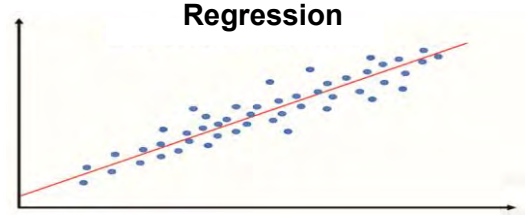
Cashflow Model

Use of **cash flow** and **transactional data** to assess customers with liquidity issues

Data Points

Cashflow Projections, Overdraft Excess, Credit & Debit Transactions, Returned Transactions (e.g. Cheques)

Regression



Score

CRANE Model

Data Points (~3K+)

Financial Ratios, Country Risk, Macroeconomic Variables, Market indices, Import / Export data, Remittance transactions, Corporate & Individual bank account

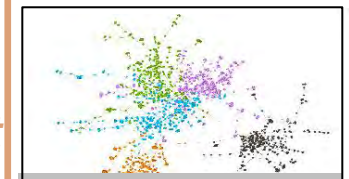
XGBoost
a modelling technique

Historical Performance Data (e.g. NPLs, delinquencies, negative watchlist)

Base model score

Expert overlay incorporates data-driven expert judgment from practitioners

Overlay score



Network effects capture contagion risk due to linkages to counterparties

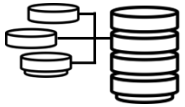
Score incorporating Network effects

Adjusting for any **Overrides** to reduce false positives

Final score

GenAI unlocks the power of unstructured data and drives outcomes in two ways

01 Improves existing use cases that leverage AI/ML



Convert **unstructured data** to structured data

Improved **prediction power** by leveraging additional information

Build truly **personalised messaging** (n=1)

Human-like conversational engagement

02 Enables new use cases

Persona-based productivity solutions (e.g. developers, CSOs)

Ease access to **knowledge base**

Search, insights and alerts (e.g., enterprise search)

Structured approach to enable GenAI use cases across the organisation to drive value

"Vertical"
Specific roles co-pilots

- Relationship Managers
- Customer Service Officers
- Operations Teams
- Software Developers
- Others…

"Horizontal"
Enable across the bank

- Enterprise Knowledge Base (EKB) -- DBS GPT

Enablers

- Value Capture Playbook
- AI Industrialisation 2.0
- Responsible AI Guardrails
- People & Change Management

AI / GenAI applications in credit risk management

Pre-underwriting → **Post-underwriting**

	Credit Assessment			Documentary Processes		Surveillance	Fulfilment
	Customer Targeting	**KYC/Credit Fraud**	**Credit Decisioning**	**Contract Generation**	**Documentary Reviews**	**Risk / Portfolio Monitoring**	**Remediation / Collections**
AI techniques	• Product propensity e.g. marketing nudges, supplier/buy network, • Credit pre-screening e.g. pre-emptive lending terms	• Rule-based triggers optimised with scorecards • Relational/transactional network analysis • Doc forgery, e.g. bank statements	• Application / Behavioural scorecards • Alternative data sources via aggregators / ecosystem players • Expert to statistical approach, risk grading	• Data extraction from docs (e.g. financial statements) using OCR and NLP • Analysis based on financial spreading using knowledge graphs and NLG • RPA to Intelligent Doc Processing (IDP)		• Reuse of behavioural scorecards • ML early warning system using network and news sentiment analysis • Portfolio stress testing incorporating macro-financials transmissions	• Collection scorecards • Recovery models • Channel/Timing recommendation engines • Skip-tracing leveraging triangulation and network analysis
Gen AI solutions	• Draft personalised content in customer communications via different channels e.g. email, chatbot, voicebot	• Adverse news scanning on applicant • Creating synthetic datasets for model training	• Structuring unstructured data to enrich features used in model training	• Credit memo generation – augment or from scratch using agentic approach • ESG risk assessment - extract from docs (e.g. annual report, sustainability report) and generate insights based on questionnaire • Doc reviews – check completeness, comparison of terms across docs (facility vs security) and policies, flag exceptions and assess materiality		• Adverse news scanning on portfolio • Credit-relevant insights from external (e.g. equity reports, industry reports) and internal sources (e.g. customer call reports, committee discussions)	• Enhancing efficiency/ effectiveness of collectors, e.g. transcribing and summarising calls, digital twin, virtual agent

GenAI can be applied meaningfully in customer engagement and extracting insights for enhancing Collection practices

Sharpening strategy | **Improving execution** | **Performance management**

Capability 1

Recovery models to prioritise calls based on size and propensity to repay

Channel/Timing recommendation engine to identify best channel/timing to contact customers

Objective: Improve customer contactability, and collection outcomes

Augment structured data sources to uplift model performance

Enabling human agents

Capability 2 - GenAI

Collector Assistant (Co-pilot) to provide live transcription, dynamic recommendations (including repayment solutions) and summarise call logs*

Objectives:
(1) Improve productivity, customer experience and collection outcomes;
(2) Create new data points for insights and analytics

Calls transcribed and structured to enrich KB to power other solutions

Automated enforcements

Capability 3

Deposit deduction
Automate deduction from available deposit balances to repay delinquent accounts

Objective: Reduce toil, improve timeliness and collection outcomes

Capability 2 - GenAI

Review and enhance
- Identify good/bad practices for training and personalized coaching
- Flag compliments/complaints
- Identify patterns/trends in delinquency

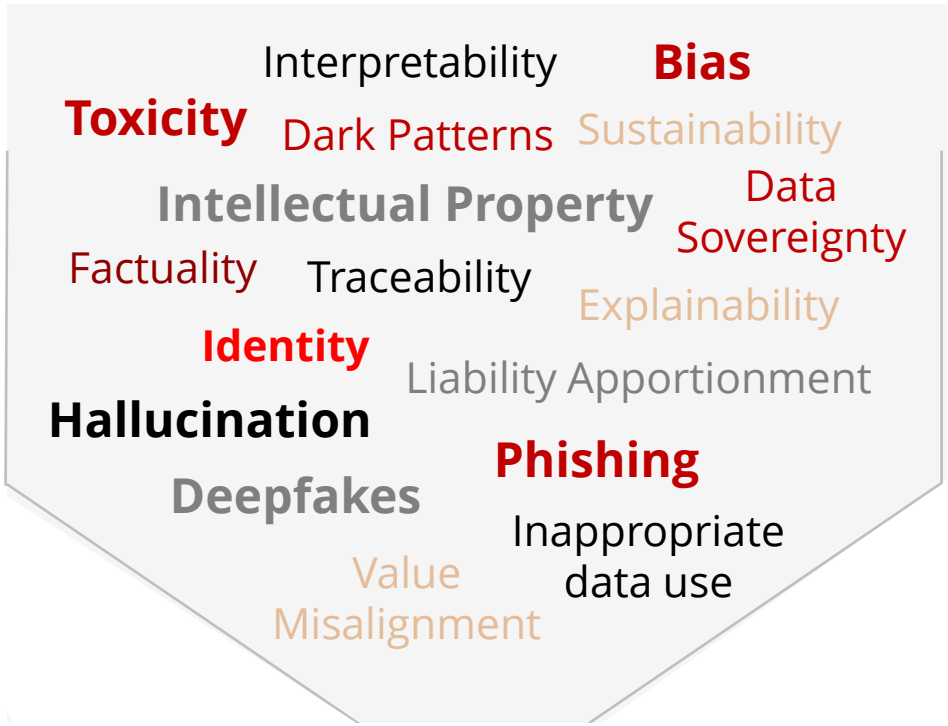
Objective: Enhance collection practices, strengthen risk and controls

Read across call logs structured in KB to extract insights

Customer Knowledge Base (*Harness structured and unstructured insights for personalized actions*)

Assessing GenAI for Risks

AI/ML risks we cover well today	
Accountability	Responsibility and governance for the outcomes and impacts of data & AI systems.
Stability	Soundness, robustness, and operational stability of the model or service.
Transparency	Human awareness, explainability, interpretability, and auditability of data & AI systems.
Fairness	Setting fairness objectives to help identify and address unintentional bias and discrimination.
Legal and Regulatory	Legal or regulatory obligations that need to be met or may be challenged by the use of AI.
Ethics	Responsible and ethical outcomes in the use of AI against a clearly defined set of core values and practices.
Privacy & Security	Protecting data & AI systems from unauthorised access, data loss and enabling privacy rights for PII related to data subjects.



3 areas of high focus

Data Confidentiality & Disclosure | Data Management | Cybersecurity

DBS

Thank You

DBS