Risk Governance Playbook

Model Risk Management & Governance policy guidelines

Contents

I. Overview and Objective

By definition, model risk is “the potential for adverse consequences from decisions based on incorrect or misused model outputs for reporting and making credit decisions”. Failure to take appropriate steps to mitigate this risk has resulted in financial and reputational losses for banks. There is a need for effective Model Risk Management (MRM) policy because the bank also uses quantitative analyses and models in many aspects of financial decision making.

The purpose of this document is to provide comprehensive guidance on effective model risk management policy across all aspects of the model lifecycle (sound model development, implementation, validation, and on-going use of models), controls and compliance, model risk quantification and senior management oversight.

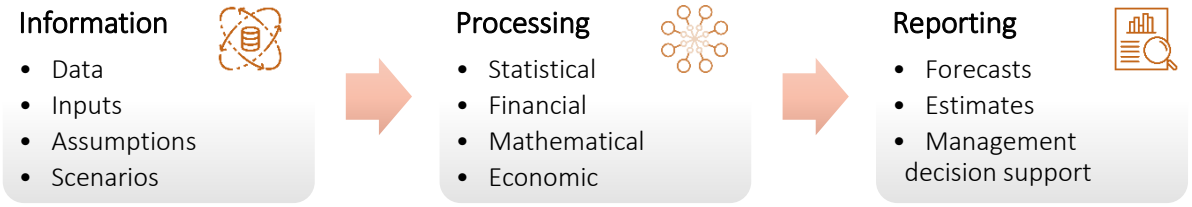
This document would serve as a handbook for the team(s), including the practitioners, to leverage the best practices as described. It is supplemented by a set of templates for effective implementation of the policy. This document leverages SR 11-07 guidelines and other industry best practices used by banks for Model Risk Management.

II. Definition of a Model

The term model refers to “a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques, and assumptions to process input data into quantitative estimates”.

A model consists of three components:

1. an information input component, which delivers assumptions and data to the model;
2. a processing component, which transforms inputs into estimates; and
3. a reporting component, which translates the estimates into useful business information.



Examples of Models:

1. Models meeting this definition might be used for analyzing business strategies, informing business decisions, identifying and measuring risks, valuing exposures, instruments or positions, conducting stress testing, assessing adequacy of capital, managing client assets, measuring compliance with internal limits, maintaining the formal control apparatus of the bank, or meeting financial or regulatory reporting requirements and issuing public disclosures.
2. Models working in areas including collections, staffing optimization, and marketing.
3. All new model development including non-scoring models such as decision trees and rule-based decision making algorithms.

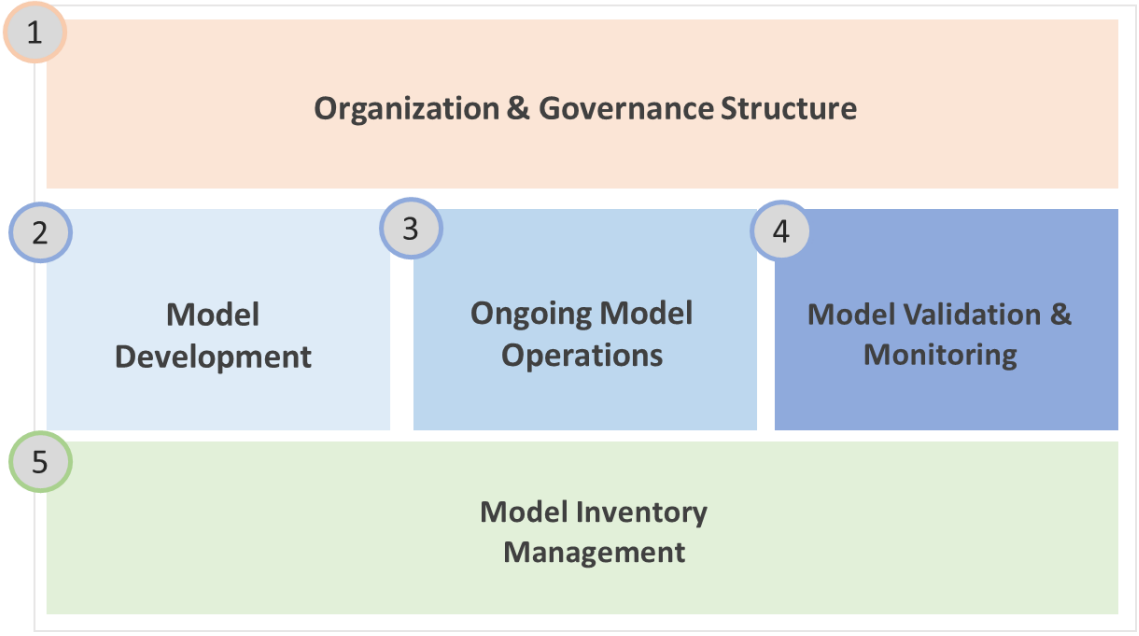
Examples of processes that are not models:

1. Any process missing one of the three steps is not a model.
2. Any process that uses discretion based and non-data based mechanism is not a model.

III. MRM Framework

The MRM Framework is supported by five key pillars as mentioned below:

1. Organization & Governance
2. Model Development
3. Ongoing Model Operations
4. Model Validation & Monitoring
5. Model Inventory Management



Key Pillars of MRM Framework executed through the governance structure

1. Organization & Governance

The critical components of robust organization and governance structure for MRM are defined as follows:

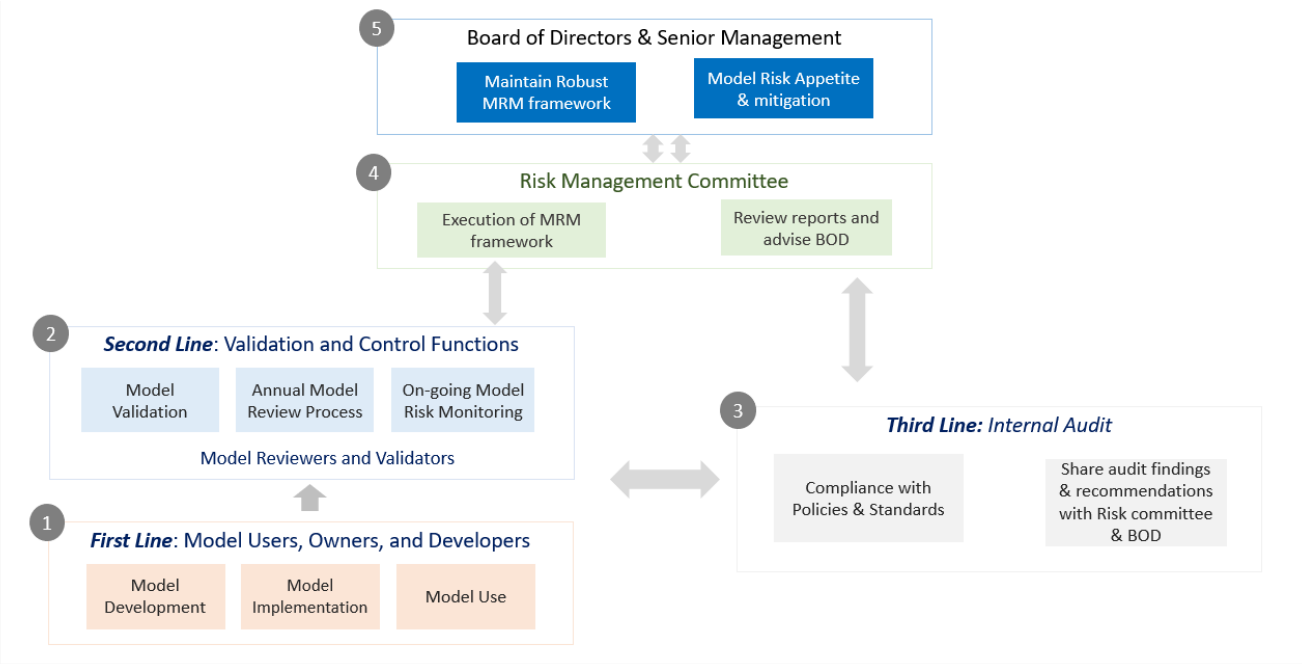
1.1. Effective MRM Organizational Design includes:

- Independent model validation function to provide effective challenge to the model development process and to identify conditions for use.
- Effective model audit procedure with a frequency of at least once a year. The frequency is higher during any adverse macro-economic situation (e.g. COVID-19).
- Transparency of model associated risk(s) through reports to the Board of Directors on a periodic basis.

1.2. Model Risk Assessment and Risk Appetite Principles include:

- Qualitative and quantitative assessment of Model Materiality and Model Risk Rating at an overall portfolio level and at individual model level.
- Model Risk Appetite statement based on solid risk governance foundation.

Effective MRM Organizational Structure



The entire process is carried out through functions with following implications:

1. First line of defense: Business Sponsor and Model Developers.
2. Second line of defense: Model Validation Unit.
3. Third line of defense: Internal and External Audit.
4. Board of Directors and Senior Management.

- **First & Second Lines of Defense:** The day-to-day management of inherent model risk is delegated to the first and the second lines of defense. This includes activities like Model

Development, Ongoing Model Operations, Inventory Management and Ongoing Monitoring.

- **Third line of defense (Internal Audit):** Internal Audit team owns the responsibility to ensure compliance with policies and standards. They should audit on a regular frequency (6-12 months for high risk assessment models) and report findings to the Risk Committee and the Board of Directors.

- **Risk management committee** should objectively look at risks reported by the second and the third line of defense to create an actionable plan for risk mitigation. This may include setting up of new infrastructure, hiring new talent in the bank, capability building through vendor project outsourcing etc.

- **The Board of Directors** is responsible for the approval of the MRM framework and reviewing regular reports on the implementation of the MRM Policy.

Apart from the above, external resources or vendors may be commissioned by banks to supplement internal capabilities for model risk monitoring, validation and review, compliance functions, or other activities in support of internal audit or other lines of defense.

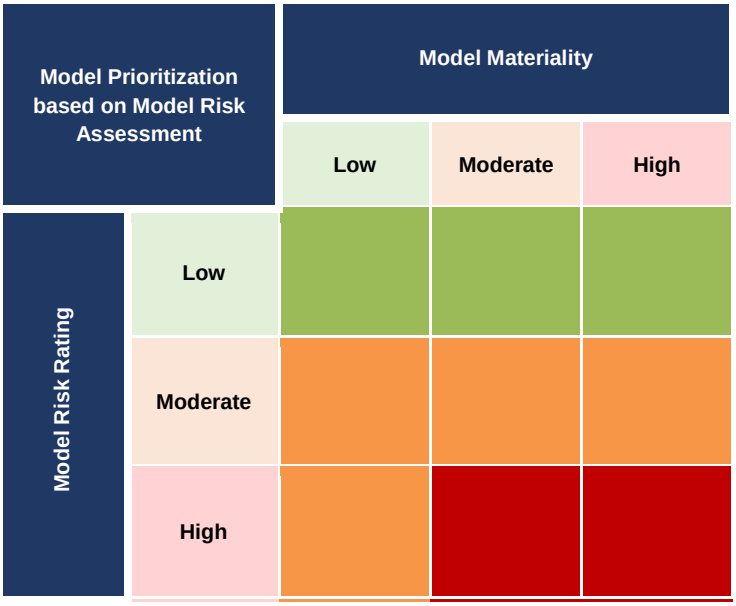
Approval process for first time implementation of models:

1. All models are to be reviewed by the MRM Reviewer and the Business Sponsor before implementation. The approval by the MRM Reviewer and the Business Sponsor serves as a basic criteria for all models.
2. All regulatory models and high Materiality models should go through review by the Risk Committee before their first time implementation. This is an extra layer of security for the high impact models.

Model Classification & Prioritization Based on Materiality and Risk Rating

A classification and prioritization procedure for models is essential along with a comprehensive Model Inventory. The level of review applied to each model is aligned with the risk posed by the model. A more frequent and more extensive review is applied to the most complex and high impact models.

A dual matrix is used to define the classification and prioritization of models based on the Model Materiality and Model Risk Rating. Materiality and its calculation are defined in the section 1.a.i.2. The Model Risk Rating and the components of quantification are explained in the section 1.a.i.1.a.i.1.



These are the recommended steps for the models based on the Red Amber Green (RAG) status in the matrix:

RAG	Recommended Steps*
	On Track
	Model Recalibration & Risk Band Change Analysis
	Retire and Redesign Model

**These actions are indicative but not limitative.*

Materiality Calculation Methodology

Materiality is the impact or consequences, particularly monetary, that model error has on the financial institution. It is determined by financial institutions in accordance with their internal assessments of risk appetite, risk capacity and risk profile, having regard to capital, liquidity and earnings at the entity level. Assessment of model Materiality helps in prioritizing actions and in rolling out the MRM framework.

Materiality is calculated by considering the following three factors:

1. Model Usage.
2. Exposure Factor.
3. Operational Factor.

The measurement for these factors in the case of classification models is as below:

- **Model Usage:** The measurement of this factor is dependent on the category of use of the model. The categories are defined as:

  Category A
  - External Solvency models
  - Regulatory models, like IFRS-9 Accounting models
  - Anti-Money Laundering models

Category B

- Credit Risk Models (like Origination Models, Behavioral models, Situational Models, etc)
- Fraud Models
- Pricing Models
- Other management models
- Collection Risk Models

Category C

- Marketing Models
- Operational Models
- Deposit Product Models

For Category A and B models, Exposure factor is used and for Category C models, like marketing models, Operational factor is used to finalize the Materiality status.

- **Exposure Factor:** The measurement of the Exposure Factor is based on the bank's risk or actual $ loss (annualized) that come under the purview of the model.

$$The\ Loss\ Incidence\ Rate\ = \frac{Volume\ of\ customers\ with\ loss}{Total\ volume\ of\ population}$$

The loss incidence rate in case of Credit Risk models is called as default rate, for Fraud Risk models, it is fraud incidence rate, for AML models, it is AML incidence rate, etc).

$$\boldsymbol{Exposure\ Factor = Volume\ of\ population \times Loss\ Incidence\ Rate \times Average\ Balance}$$

As an example, the illustration below shows the calculation of Exposure Risk for Credit Risk models:

Model	Volume	Annual Default Rate%	$ Average Balance	Exposure Risk $
ORG1	10000	4.2	420	17,640,000

Volume × Default Rate × Average Balance

- **Operational Factor:** This factor is based on the measurement of the eligible population the model processes, multiplied by the operational costs of the model.

$$\boldsymbol{Operational\ Factor = Volume\ of\ population(to\ be\ targeted) \times Channel\ cost\ per\ account\ (\$)}$$

Model	Volume (to be targeted)	Channel cost per account ($)	Operational Exposure ($)
Channel Optimization Model 1	10000	0.2	2,000

Volume targeted × Channel cost

Materiality Calculation Steps

Materiality calculation based on model usage is to be carried out in following steps:

1. If model is in Category A (Regulatory models), then tag it as 'High Materiality'.
2. If model is in Category B, then calculate the Exposure factor for each model, divide the exposure factor by the technical capital for the bank, and rank order models into High, Medium, and Low Materiality.
3. If model is in Category C, then calculate Operational factor for each model, divide the operational factor by the technical capital for the bank and rank order models into High, Medium, and Low Materiality.
4. For models in Category B and Category C, as a good practice, the models should be rank ordered on the basis of the Exposure or Operational factor divided by technical capital of the bank, and categorized equally as High, Medium and Low Materiality (33% in each). However, it is a good practice to review the entire inventory before assigning materiality.

Model Materiality – best practice for implementation:

The Materiality of the models can be categorized into High, Medium, or Low Materiality through following two possible ways:

1. **Overall enterprise level Materiality calculation:** All the models for the bank are taken into account while identifying the High, Medium or Low Materiality models. This view is helpful for senior level executives and one down to understand the enterprise level risk for models.
2. **Department level Materiality calculation**: This approach uses identification of High, Medium or Low Materiality models for each department, for example, marketing, risk, etc., so that the department head can focus on the high impact models under their purview.

Model Tiering & Frequency of Monitoring based on Materiality:

The tiering for the models is defined on the basis of the following RAG Status:

RAG	Comment	Frequency of Monitoring	Levels of Approvals
	Low Materiality	6 months	Least
	Moderate Materiality	3 to 6 months	Least
	High Materiality	3 months	Multiple

Model Risk Rating Based on Industry Standard Benchmarks

The evaluation of the risk rating is dependent on multiple factors that are defined in detail under the Model Validation for classification models. The factors considered are:

1. Model Discrimination.
2. Rank Ordering.

3. Population Stability Index.
4. Characteristic Stability Index.

Benchmarks for All Models:

- **Model Discrimination:** Ability of the model to accurately discriminate between event and non-event. Model Discrimination is defined through various elements: KS Statistic, AUC Score, Gini Score, Divergence, Accuracy, Precision, Recall, and F1 Score.
  The thresholds for assigning RAG status for these is defined as:

KS Statistic: Origination Models		
KS	**% Change**	**Status**
Validation > Development: KS < 30%		Amber
Validation > Development: KS >= 30%		Green
KS >= 30%	>=15%	Amber
	< 15%	Green
KS < 30%	>= 15%	Red
	5-15%	Amber
	< 5%	Green

KS Statistic: Behaviour Models		
KS	**% Change**	**Status**
Validation > Development: KS < 40%		Amber
Validation > Development: KS >= 40%		Green
KS >= 40%	>=15%	Amber
	< 15%	Green
KS < 40%	>= 15%	Red
	5-15%	Amber
	< 5%	Green

AUC Score		
AUC	**% Change**	**Status**
Validation > Development: AUC < 80%		Amber
Validation > Development: AUC >= 80%		Green
Validation <= Development	>=20%	Red
	10-20%	Amber
	< 10%	Green

Gini Score		
Gini	**% Change**	**Status**
Validation > Development: Gini < 60%		Amber
Validation > Development: Gini > 60%		Green
Validation <= Development	>=20%	Red
	10-20%	Amber
	< 10%	Green

Divergence	
Validation > Development	Green
% Change <= 10%	Green
10% < % Change <= 20%	Amber
% Change > 20%	Red

Accuracy	
Validation > Development	Green
Accuracy >= 75%	Green
60% <=Accuracy < 75%	Amber
Accuracy < 60%	Red

Precision	
Validation > Development	Green
Precision >= 85%	Green
70% <=Precision< 85%	Amber
Precision < 70%	Red

Recall	
Validation > Development	Green
Recall >= 80%	Green
65% <=Recall < 80%	Amber
Recall < 65%	Red

F1 Score	
Validation > Development	Green
F1 Score >= 82%	Green
67% <=F1 Score < 82%	Amber
F1 Score < 67%	Red

The overall thresholds defined for the Model Discrimination Component are

KS Statistic: Origination Models		
KS	**% Change**	**Status**
Validation > Development: KS < 30%		Amber
Validation > Development: KS >= 30%		Green
KS >= 30%	>=15%	Amber
	< 15%	Green
KS < 30%	>= 15%	Red
	5-15%	Amber
	< 5%	Green

KS Statistic: Behaviour Models		
KS	**% Change**	**Status**
Validation > Development: KS < 40%		Amber
Validation > Development: KS >= 40%		Green
KS >= 40%	>=15%	Amber
	< 15%	Green
KS < 40%	>= 15%	Red
	5-15%	Amber
	< 5%	Green

- **Rank Ordering:** Rank order breakage is a sign of deterioration in model performance in case of classification models. The Rank Ordering of a model refers to its ability to yield accurate point predictions for the output variable of predictive models.
  The threshold definition for Rank Ordering is as below:

Rank Ordering	
No Rank Order Breakage	Green
Rank ordering breakage in bottom 70% deciles	Amber
Rank ordering breakage in top 30% deciles (risky deciles)	Red

- **Population Stability Index:** Standard measure for stability of population on which model was developed versus the latest cohort for validation. High PSI indicates shift in population profile and may require model redevelopment.
  The threshold definition for Population Stability Index is as below:

Population Stability	
Threshold <= 0.10	Green
0.10 < Threshold <= 0.25	Amber
Threshold > 0.25	Red

- **Characteristic Stability Index:** CSI is a measure to examine how individual model characteristics or variables may have changed in distribution between the development cohort and the latest validation cohort.

  The threshold definition for CSI for a single variable is as below:

Characteristic Stability Index	
Threshold <= 5.0	Green
5.0 < Threshold <= 10.0	Amber
Threshold > 10.0	Red

  The threshold definition for the overall CSI component:

Characteristic Stability	
Otherwise	Green
More than 30% variables have amber CSI status and less than 30% variables have red CSI status	Amber
More than 30% variables have red CSI status	Red

Overall Risk Rating for the models on the basis of benchmarks:

Model Risk Rating	
Otherwise	Green
Either model discrimination is amber or population stability is red	Amber
Either model discrimination is red or model discrimination is amber and population stability is red	Red

Conceptual Soundness Annual Review:

The conceptual soundness is measured on two components. The MRM Reviewer should carry out a deep dive review of the conceptual soundness and share recommendations with the Model Development and Model Validation team on the following components:

- **Model Design Methodology:** Review the model creation procedure starting from the data considered for the model to the implementation. The following factors should be considered to review the model design:

  1. Data Quality
  2. Portfolio Considered
  3. Event Definition Methodology
  4. Selection of variables
  5. Statistical Techniques used
  6. Implementation Strategy

- **Stress-testing of Models:** Evaluate if the model would provide optimal results in changing macro-economic environment and if data used during development is a reasonable representation of the bank's portfolio and the market conditions.

Risk Appetite Statement

Risk Appetite should be created by the Risk Committee, reviewed and approved by the Board of Directors, and implemented across the organization. The Risk Appetite is governed through two key mechanisms:

1. The **Risk Appetite Framework** defines the risk capacity, the risk appetite, the risk limits and the risk profile along with the processes, governance bodies and features of how bank addresses risk appetite as a part of its ongoing business.
2. **Periodic Risk Appetite Statements** define the desired level of risk commensurate with return and growth targets and in line with the corporate strategy and stakeholder objectives.

The examples for Overall Risk Appetite Statement and Model Risk Appetite Statement are given below. They are only for illustration purposes. The Risk Committee needs to develop these statements and define the quantitative concepts to be rolled out for implementation to the relevant stakeholders post approval from Board of Directors.

Overall Risk Appetite Statement

"Group's risk policy is designed to achieve a moderate risk profile for the Bank through: prudent management and a responsible banking business model aimed at generating value, return adjusted to principles and recurring earnings, diversified by class of assets, portfolios and customers, keeping a medium/low risk profile in each country of Bank's operations and supported by long-term relationships with customers."

Model Risk Appetite Statement

"Model risk is the potential for adverse consequences from incorrect or misused model outputs and reports using these outputs. Model risk can lead to financial loss, poor business or strategic decision making, or damage our reputation. The term 'model' refers to a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques, and assumptions to process input data into quantitative estimates.

Model risk is managed across Risk, Capital & other models and Model Risk Appetite is aligned to the Group's qualitative statements, ensuring that Model Risk Management is embedded as part of risk and control process.

The management of model risk includes:

- *Performing robust independent model validation that provides effective challenge to the model development process and includes identification of conditions for use, methodological limitations that may require adjustments or overlays, and validation findings that require remediation;*
- *Establishing a strong Model Risk Management and Governance Framework, including senior forums for monitoring and escalation of model risk related topics;*
- *Creating Bank-wide model risk related policies, aligned to regulatory requirements with clear roles and responsibilities for key stakeholders across the model life cycle; and*
- *Providing an assessment of the model risk control environment and reporting to the Management Board and Risk Committee on a periodic basis.*

We have been implementing improvements in our model risk management framework, and significant progress has been achieved in 2020 across the above for all the credit risk processes:

Management has determined to instill a robust and consistent model risk management framework in order to ensure that the framework effectively identifies, monitors and measures risk. The effort focuses on board oversight, model identification and model inventory, standardized policy and governance framework, model development practices, and effective and rigorous model validation."

2. Model Development

Key components of Model Development for MRM are as follows:

- **Standardized model development process**, across the model development stages – from modeling problem definition to model pipeline creation
- **Standardized template** to document the process
- **Approval by the MRM Reviewer and the Business Sponsor on model insights**, before proceeding with implementation

The following process should be socialized across the development teams and key stakeholders:

2.1. Modelling problem definition:

Modeling problem definition leverages a combination of discussions with the business stakeholders, review of existing dashboards, and preliminary data exploration to define the problem scope, success metric(s), event definition, etc. for the model to be built. For example, for building an attrition prevention model, modeling team should review existing attrition trends (overall, by product, by segments, etc.) to identify the high priority areas/models that would further help in defining the framework for the model building exercise.

- Establish the business objective and ensure that the scope of the solution is clearly aligned with the business problem.
- Formulate the business objective as a modelling objective to translate the problem into the first step of the solution.
- Select a simple, observable and attributable success metric from the model such that it is the output of the model and the dependent variable.
- Ensure the technical soundness of the techniques selected in terms of ease of implementation and robust methodology. As an example, if the implementation system cannot use machine learning model output and can only use a set of 5-10 rules, then problem should be scoped as a segmentation or an optimal rule identification.

2.2. Data Requirement Assessment:

- Identify the data sources required for the model and adjudge the quality of the available data.

- Involve people from the relevant business domains, information technology and data analysts to assess internal support for the model.
- Establish the infrastructural requirements for proper handling and maintenance of the data sources and identify the connections with the code.
- Post model build, re-check the data quality of significant attributes to ensure only data with high quality has been used for the final model.

2.3. Modelling Process:

- Establish a standardized process aligned across the organization for the model development teams.
- Run test datasets close to the real environment to identify any gaps in the predictions of the model against decided benchmarks.
- Identify the limitations and assumptions used in the model and document them with the other processes.
- Develop a control process to constitute the benchmarks for the ultimate production model and identify the necessary objective criteria for the selection.
- Assign the appropriate systems for implementation of the models and ensure efficient integration of models into the deployment environment.

DATA PRE-PROCESSING	• Data Preparation: Examine the data and locate the inconsistencies – like, missing value, outliers, etc. Analyse and apply the appropriate technique to resolve inconsistencies. • Data Transformations: Analyse the data to identify opportunities to transform variables to log or exponential values for more accurate impact analysis. • Data Operations: Apply aggregation or formatting to the variables to ensure the data is easier to understand. E.g. filtering, binning, merging, renaming, etc.
EXPLORATORY DATA ANALYSIS	• Evaluate the final dataset and arrive at hypotheses aligned with the problem statement. • Test the hypotheses using techniques like correlation, pivot tables etc. to drill down to the dataset that will provide the most accurate predictions.
MODELING FRAMEWORK	• Based on the historical data, identify the optimal definition of event, use relevant cohorts/time periods of data with no anomalies. • Select the most optimal statistical modeling technique.
FEATURE ENGINEERING	• Reduce the number of variables using techniques like PCA, factor analysis, etc. • Transform the variables based on domain understanding and create new variables by combining variables or adding relevant features from a pre-developed Feature Store. • Examine the dataset against the independent variable to identify any class imbalance problems and select the appropriate sampling technique.

CONTROL PROCESS	• Develop a baseline model that serves as a benchmark for the rest of the models. The benchmark model is not required to capture the complex relationships between variables. • To avoid overfitting or underfitting of the data on the model, use a robust test vs train split technique.
HYPERPARAMETER TUNING	• Training the model on the dataset iteratively can lead to problems like overfitting. The other side of the coin is underfitting. • Regularise the models using hyper parameter tuning to improve the model performance. • Simple hyper parameters include number of training steps, learning rate, initialisation values and distribution.
MODEL TESTING	• Run the model using sample dataset to test the generalisation of the model. • Evaluate the performance of the sample model runs against benchmarks to ensure a high quality production code and document the results.

2.4. Model Pipeline Creation:

- Focus on the system infrastructure requirements to set up the model pipeline and align it with the specifications of the models.
- Run a sample model to ensure the robustness of the connections in the pipeline. Identify the bugs and assign them to the right parties.
- Document the iterations of the runs to benchmark against the required performance of the pipeline and the model.

Standardized template for Model Documentation:

For effective Model Documentation, the steps are as follows:

- Create an official and approved standardized template to be used across the organization for all models.
- In the document, detail the rationale, assumptions, process of development, derivations, tests, and other analyses supporting the usage, maintenance and implementation for each individual model. This document should be up-to-date and approved by the relevant stakeholders.

A standardized template for the model development document is detailed as follows:

Model Objective & Scope	This component includes: • Model Objective • Product and Portfolio Used • Need for the Modelling Exercise • Existing Model (if any)
Modelling Process	This component accounts for: • Data Pre-processing Techniques • Exploratory Data Analysis

	• Modelling Framework: This further includes:
	  • Event Creation Methodology
	  • Statistical Technique Used
	  • Assumptions and Limitations
	• Feature Engineering Techniques
	• Control Process
	• Hyperparameter Tuning
	• Model Testing & Evaluation: Results and Final Model Selection

Implementation Plan	This component includes the specifications of the model pipeline and the details of the model implementation (how will the model be implemented, when will it be implemented, etc.)
Ongoing Monitoring and Governance Plan	This component includes the suggested monitoring frequency for the model.

3. Ongoing Model Operations

Key components of Ongoing Model Operations for MRM are as follows:

- **Approval by Business Sponsor** on campaign (e.g. customers to be targeted etc.) using model with the customer level data and the implementation code saved in a folder.
- **Quality check of code by the model implementation team** in terms of selection of the right period, and targeting base.

3.1. Model Implementation process controls:

Model implementation process is carried out through the cycle below and would hold more relevance for Marketing Models where the output of models is used for campaign targeting:

1. Email from Business Sponsor to the modelling team describing the campaign objective and model use for creation of campaign base (Owner: Business Sponsor).
2. Code run by modelling team with results in terms of eligible base based on model shared with the Business Sponsor over email (Owner: Model Ops).
3. Approval of final targeting universe by the Business Sponsor over email before campaign implementation (Owner: Business Sponsor).

Review of the business and campaign implementation framework, volume of customers, and the product offered is the Business Sponsor's responsibility while the due diligence in terms of error free code implementation based on business requirements is the responsibility of the Model Ops team.

The email from Model Ops Team to the Business Sponsor should contain the following:

1. Results of the sample
2. Period of the data used in the sample run
3. Volumes targeted in the sample
4. Any assumptions or limitations during implementation

4. Model Validation & Monitoring

Model validation includes the set of processes and activities intended to verify that models are performing as expected, in line with their design objectives and business uses. Effective validation helps ensure that the models are sound. It also identifies potential limitations and assumptions, and assesses their possible impact.

The frequency for ongoing validation of models in general is 6 months and 3 months for high impact models. Model monitoring should be more frequent in case of internal change in targeting criteria or any uncertain/adverse environment conditions, like COVID-19.

Key components of Model Validation and Monitoring for MRM policy are as follows:

- **Statistical soundness check for the model** based on the industry standard benchmarks for the model (model discrimination, population stability, and rank ordering). The CSI measurement and independent variables data quality should also be reviewed in case of deteriorated model performance.
- **Conceptual soundness check** to ensure best methodology, technique etc. is used for the model. Validators should identify any potential gaps and share challenger methodology.
- **Control framework on on-going usage of model** including monitoring frequency, review of model validation reports by senior management etc**.**
- **Business Impact Analysis** to ensure model's performance results & business impact is incremental for the business.

4.1. Statistical Soundness:

- Following statistical metrics are represented in the model validation report:
  - Model Discrimination: Ability of model to discriminate between event (possible defaulter) & non-event (creditworthy customer).
  - Population Stability: Standard measure for stability of population on which model was developed vs present performance cohort. High PSI indicates shift in population profile and may require model redevelopment.
  - Characteristic Stability: A measure to examine how individual model characteristics or variables may have changed in distribution during two time periods (development vs validation).
  - Rank Ordering: Ability to yield an accurate point prediction for the output variable. Rank order breakage in validation is a sign of deteriorating model performance for classification models. Ideally, the model should deliver accurate point predictions for models where the probability values of the models are directly used.
- These metrics are dependent on the nature of the model (classification or regression) and other key success factors involved in the performance of the model.

- Define performance thresholds for the metrics to assign them a “traffic lights” or “RAG” based system. The thresholds should be objective and based on industry benchmarks.

Component	Description	Measure
Model Discrimination	• The ability of the model to differentiate between events and non-events based on its input values, such as defaults and non-defaults. • For models that do not have a binary outcome, this can be measured similarly as a ‘High’/’Low’ indicator.	• Gini Coefficient • Kolmogorov-Smirnov (KS) Statistic • Area Under Receiver Operating Characteristic (AUC ROC) Curve • Confusion Matrix with Accuracy, Precision, Recall and F1 Score • Divergence
System Stability	• How different is the current data being scored by the model, compared to the data from the model development; is the model stable over time?	• Population Stability Index
Characteristic Stability	• How different is the distribution of the current population in each explanatory variable in the model compared to the population used for model development? • What impact does this have on model performance?	• Characteristic Stability Index
Rank Ordering	• Rank order breakage in validation is a sign of deteriorating model performance for classification models. • Ideally, the model should deliver accurate point predictions for models where the probability values of the models are directly used.	• Rank Ordering Breaks across deciles • Actual vs Predicted Curve

4.2. Conceptual Soundness:

This element involves assessing the quality of the model design and its efficacy in the given environment. It entails review of documentation and empirical evidence supporting the methods used and variables selected for the model.

- Review the model design methodology to identify any potential flaws in data quality used, ensure the model is implemented on the right customer base, usage of the appropriate statistical technique etc.
- Evaluate if the model would provide optimal results in changing macro-economic environment and if data used during development is a reasonable representation of the bank’s portfolio and market conditions.
- Challenge the model creation process used at time of development. For example, check if the right event is used as dependent variable etc.
- Assess the selection of variables and their impact on the outputs with a focus on the assumptions and limitations.

- Document the results and identify the limitations of the model. In case of significance of limitations in regular use of a model, develop a new approach.

4.3. Control framework on on-going validation of the model:

Such monitoring confirms that the model is appropriately implemented and is being used and is performing as intended.

- Regular monitoring frequency for the model(s) based on its nature, Materiality and modeling approaches used. For example, 3 months for High Materiality models and 6 months for Low and Medium Materiality models.
- Model level results of validation to be shared with Risk Committee and Board of Directors on a quarterly basis with any risk(s) for high priority models highlighted in the reports.
- Identify the potential areas where business overrides are made on top of model predictions and measure the risk associated with such judgment criteria.

4.4. Business Impact Analysis:

Business impact analysis should be used to identify the Model vs BAU performance, based on actual results of targeting or prediction. The analysis should be available for all models that are used for targeting. For any regulatory models used for point estimation, business impact analysis is not compulsory.

- **Identify the KPIs that were primary expected outcomes of the modeling exercise:**

  Examples are mentioned below:

  - For credit originations model, KPIs could be % defaulters, $ Credit Loss, Approval rate etc
  - For collections risk prediction model, KPIs could be % recovery, $ recovery, $ credit loss etc

- **Test (Model) Vs Control (No-model) Impact Measurement :**

  The net incremental impact on KPIs should be measured through test vs control analysis.

5. Model Inventory Management

A model inventory takes stock of the models used by an institution and establishes clear ownership of the maintenance and usage of the model.

- **Maintain a standard inventory of all models** in the organization based on a single definition into a repository that contains the model names, model documentation, production code, and validation reports.
- **Define the measurement of Materiality** of the models – based on the Model Materiality framework, define whether the model is High, Medium or Low Materiality – and based on the Materiality, **classify and prioritize the models** across the portfolio.
- **Workflow based MRM system** with clear responsibilities in-built in a technology platform, to enable smooth functioning of model inventory management.

Guidelines for establishing a model inventory include:

- Segregate the inventory building exercise by model category or specific bank function. Example segments may include the following :

  Credit Risk

  - Underwriting/Application Scoring Models
  - Account/Customer Behavior Scoring Models
  - Pricing Models

  Collections

  - Collections targeting
  - Collections Recovery

  Regulatory

  - Capital Planning Models such as PD, LGD and EAD
  - Stress Testing Models

- At time of Model Development, the Developer will send request for addition of model to the inventory after approval to the Model Validator.
- The Model Validator will add the model to the inventory and create the repository with the required documents.

The inventory for all models should include the following information for each listed model, documented and placed at single source for various stakeholders:

Model Name, Details, and Description
Model Stakeholders: Business Sponsor, Model Developer, Model Validator, Risk Manager
Business Unit the model belongs to
Risk Rating at every validation date
Model Documentation, Validation Reports and any related documentation
Relevant Dates: Date of Implementation, Last Date of Validation, Model Creation Date, Date of Onboarding on platform
Materiality of the model
Product Type the model deals with

IV. Glossary

Annual Model Review: Annual Model Review Process is an annual control checkpoint to take stock of entire model inventory and identify any opportunities. This definitely has some overlap with model validation but there are other components like annual audit findings, any implementation issues etc.

Characteristic Stability Index: A measure to examine how individual model characteristics or variables may have changed in distribution during two time periods: development vs validation.

Conceptual Soundness: This check ensures the use of the best methodology, and statistical technique in the development of the model. Validators should also identify any potential gaps and share challenger methodology as required.

Exploratory Data Analysis: Exploratory Data Analysis refers to the critical process of performing initial investigations on data so as to discover patterns, to spot anomalies, to test hypothesis and to check assumptions with the help of summary statistics and graphical representations.

Hyperparameter Tuning: In machine learning, hyperparameter optimization or tuning is the problem of choosing a set of optimal hyperparameters for a learning algorithm. A hyperparameter is a parameter whose value is used to control the learning process.

Materiality: Materiality is the impact or consequences, particularly monetary, that a model error has on the financial institution.

Model Discrimination: Ability of model to discriminate between event (possible defaulter) & non-event (creditworthy customer).

Model Inventory: Aggregation of all defined models used by an organization along with all the other required quantitative and qualitative information.

Model Repository: Aggregation of all defined models along with their related components like the production code, model documentation, validation reports, and archived data.

Model Risk Appetite: Risk Appetite denotes the set of explicit internal firm statements that capture total amount of risk that an organization is prepared to accept, tolerate, or be exposed to at any point in time. Risk appetite statements must in general be translated into concrete policies, practices and management frameworks. It is defined through the Risk Appetite Framework that contains the risk capacity, the risk limits, the risk appetite and the risk profile. These are captured in periodic Risk Appetite Statements.

Model Validation: Model Validation is the process of measuring the performance of each model against industry benchmarks.

Monitoring Schedule: The monitoring schedule is defined on the basis of the model tiering and depends on the prioritization of the model. The high impact models are validated every 3 months and the medium and low impact models are validated every 6 months.

Population Stability Index: Standard measure for stability of population on which model was developed vs present performance cohort. High PSI indicates shift in population profile and may require model redevelopment.

Production Code: The code that is ready for implementation is the production code. This type of code is modular, optimized, and compatible with the production ecosystem, versioning and enhanced readability.

Rank Ordering: Ability to yield an accurate point prediction for the output variable. Rank order breakage in validation is a sign of deteriorating model performance. Ideally, the model should deliver accurate point predictions for models where the probability values of the models are directly used.