Model Documentation Template

Template Owner: Model Risk Management

Template Last Updated: July. 2024

**Template Statement**

Model Owners are ultimately responsible for ensuring that model development and maintenance is properly documented—comprehensive and transparent enough that the model development process can be replicated by a knowledgeable third party. Model Owners often employ the assistance of other parties, such as the Model Developers and the Model Implementation Team, when writing the model development documentation.

This template is designed to aid in the model development and maintenance documentation process for all models. The Model Owner should use judgment to determine if additional detail or elements not explicitly identified in this template should also be incorporated. The addition or removal of sections should be done at the discretion of the Model Owner, based on relevance to the model at hand and in accordance with the Model Risk Management Procedures.

Guidelines (indicated in blue text throughout the document) have been provided to guide the writer regarding desired contents and writing style. Please note that these are merely suggestions and should only be included if applicable.

**Documentation Scope**

The documentation requirements may vary depending on the model, for example depending on whether the model is internally- or externally-developed, as well as statistical or computational. Unless otherwise mentioned, this documentation requirement applies to all models.

|  |  |
| --- | --- |
| Model Information | |
| Info | Content |
| **Model ID** | *The ID number of the model starting with ‘MD\_’ in model inventory* |
| **Model Name** | *The exact model name in model inventory* |
| **Model Version** | *Version of the model* |
| **Model Owner** | *Model Owner name in model inventory* |
| **Model Owner Contact Information** | *Model Owner email* |
| **Model Sponsor of 1st LoD** | *Business Unit Model Sponsor name in model inventory* |
| **Model Approver of 1st LoD** | *The individual who provides the approval for the model to be submitted for MRM validation. The model approver can be Business Unit Model Sponsor or Model Owner’s manager. If Model Owner is VP level or above, Model Owner can also be the Model Approver.* |
| **Model Approval Date** | *The date of the model approval is obtained* |
| **Model Developer** | *Model Developer(s) name(s) in the model inventory* |
| **Model Validator of 1LoD** | *Internal validator name(s), if applicable.* |

|  |  |
| --- | --- |
| Document information | |
| Info | Content |
| **Document Author** | *The author(s) name(s) for this documentation* |
| **Document Approver of 1st LoD** | *The Approver of the model development documentation can be Model Owner or Business Unit Model Sponsor. If Model Owner delegate is Senior Director or above, Model Owner delegate can also be the document approver.* |
| **Document Approval Date** | *The date of the approval for the latest version of this documentation* |
| **Previous Version Approval Date** | *The approval date of the previous version of documentation, if applicable.* |

Complete the table below to record the history of changes to this document, as well as the update of version number. Complete a new row of the table for each major revision to the document, typically occurring if the model itself has changed, or the documentation has been modified. Judgment should be applied to determine whether a change is major or minor (i.e., V1.1 or V2.0, respectively). For the first draft of the model documentation, please list as ‘initial version’ with version number 1.0.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Change and Approval Control History | | | | |
| Version | Description of Modification (List of Sections Updated) | Changed by | Approved by | Date of Change |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Contents**

[Model Documentation Template 1](#_Toc175661462)

[1. Document Introduction 5](#_Toc175661463)

[1.1 Purpose 5](#_Toc175661464)

[1.2 Scope 5](#_Toc175661465)

[1.3 Terms and Descriptions 5](#_Toc175661466)

[2. Executive Summary 6](#_Toc175661467)

[2.1 Business Context and Requirements 6](#_Toc175661468)

[2.2 Objective of the Model 6](#_Toc175661469)

[2.3 Intended Use and Scope of Applicability 6](#_Toc175661470)

[2.4 Model Revenue Information 6](#_Toc175661471)

[2.5 Upstream/Downstream Models 6](#_Toc175661472)

[2.6 Model Overview 6](#_Toc175661473)

[2.7 Regulatory/Compliance/Legal Requirements 6](#_Toc175661474)

[2.8 Model History 7](#_Toc175661475)

[2.9 Model Ongoing Change Plan 7](#_Toc175661476)

[3. Development Data 8](#_Toc175661477)

[3.1 Development Data Set Overview 8](#_Toc175661478)

[3.2 Data Extraction 8](#_Toc175661479)

[3.3 Data Reconciliation and Quality 8](#_Toc175661480)

[3.4 Data Cleansing 9](#_Toc175661481)

[3.5 Data Sampling 9](#_Toc175661482)

[3.6 Development Data Testing 9](#_Toc175661483)

[3.7 Development Data and Processing/Testing Code 9](#_Toc175661484)

[4. Model Methodology 11](#_Toc175661485)

[4.1 Model Theory and Approach 11](#_Toc175661486)

[4.1.1 Model Framework Overview 11](#_Toc175661487)

[4.1.2 Model Segmentation 11](#_Toc175661488)

[4.1.3 Model Processing Logic 11](#_Toc175661489)

[4.2 Model Features/Variables 11](#_Toc175661490)

[4.3 Model Assumptions/Limitations 12](#_Toc175661491)

[4.4 Expert Consultation and Industry Research (if applicable) 13](#_Toc175661492)

[4.5 Model Methodology Testing 13](#_Toc175661493)

[4.6 Model Specification 13](#_Toc175661494)

[4.7 Model Development and Testing Code 14](#_Toc175661495)

[5. Model Output 15](#_Toc175661496)

[5.1 Sensitivity, Stability and Robustness Testing 15](#_Toc175661497)

[5.2 Calibration and Overlay Testing 15](#_Toc175661498)

[5.3 Benchmark Testing 15](#_Toc175661499)

[5.4 Bias Testing 15](#_Toc175661500)

[6. Model Implementation 16](#_Toc175661501)

[6.1 Model Implementation Processes and Procedures 16](#_Toc175661502)

[6.2 Production Input Data 16](#_Toc175661503)

[6.3 Model Implementation Testing 16](#_Toc175661504)

[6.4 Updating Production Model 17](#_Toc175661505)

[7. Model Ongoing Monitoring 18](#_Toc175661506)

[7.1 Ongoing Monitoring 18](#_Toc175661507)

[7.2 Business Continuity 19](#_Toc175661508)

[7.3 Model Access Management 19](#_Toc175661509)

[8. Appendix 20](#_Toc175661510)

[8.1 Supplemental Documents 20](#_Toc175661511)

[8.2 Other 20](#_Toc175661512)

# Document Introduction

* 1. Purpose

Describe the purpose of this document.

* 1. Scope

Describe the scope/coverage of this document.

* 1. Terms and Descriptions

Describe all acronyms used in this document.

|  |  |
| --- | --- |
| **Acronym / Terms** | **Full name / Description** |
|  |  |
|  |  |
|  |  |

1. Executive Summary
   1. Business Context and Requirements

Briefly describe the context of the business that uses the model. This should include Visa’s businesses as well as any applicable clients.

Describe the sources of business requirements and the process of gathering business requirements. Also, describe the coordination between all relevant functions to ensure the model meets the business requirements. Please include a summary of the business requirements in this section and include the actual business requirements documents in appendix.

* 1. Objective of the Model

Briefly describe the objective of the model (e.g. fraud detection, credit rating, etc.)

* 1. Intended Use and Scope of Applicability
* Briefly describe the intended uses of the model including how the model is intended to serve the line(s) of business, geography, and product(s) to which the model can be applied. Include whether the model is used for internal purposes, or the model / its output are provided to Visa’s clients.
* The scale of the model usage, including risk exposure and the model usage (transaction counts per day or per month, credit exposure of the transactions).
  1. Model Revenue Information

For all revenue-generating models, the potential expected annual direct revenue of the model must be provided. It is also strongly recommended to provide the potential expected annual indirect revenue, if applicable.

* 1. Upstream/Downstream Models

List **all** upstream models, whose outputs are used as the inputs of this model.

List any dependent models as per the knowledge of the model owner or the model development team.

If there is not any upstream model, please specify no upstream model. If there is not any downstream model, please specify no downstream model.

* 1. Model Overview

Provide a brief overview of the model, including model development history and model design.

Describe the process for the model development including key stakeholders, roles and responsibilities of the key stakeholders, governance of the development process (e.g. committees), timeline, etc.

* 1. Regulatory/Compliance/Legal Requirements

*Describe the pertinent regulatory or compliance requirements and the impact of these matters (e.g., Data Use Case Assessment (DUCA), Privacy Risk Assessment (PRA)), if applicable.*

*Provide evidence of legal approval, if applicable.*

* 1. Model History

*Provide a high-level description of history of changes and the timing for the changes, if possible, include the records for all previous versions. For the most recent change, provide description include details of changes made in, for example, model segmentations, model features and modeling methods.*

*Please also explain the rationale of related changes, if applicable.*

* 1. Model Ongoing Change Plan

*Describe any pre-scheduled change plan for future versions of model development, which is scheduled/seasonal model change regardless of model performance evaluation, if applicable.*

# Development Data

* 1. Development Data Set Overview

For each data input:

* Provide a name for each data input source.
* Provide a description of key data characteristics, including data structure (e.g., time series, panel data, unstructured, or structured tabular, etc.), level of granularity (transaction-level, account-level or merchant-level etc.), length of data available, length of data considered for modeling (and kept as part of testing), and rationale for choice of data versus other potential options (with a description of alternatives considered). If certain regulations require the use of a particular data type, describe those requirements, and discuss how the selected development dataset complies with the requirements, or explain why these requirements cannot be met.
* *Describe the data sources used, including internal vs. external, frequency of data updates to collect data from these sources (e.g., quarterly, monthly vs. weekly). In selecting data sources and data elements for inclusion in the modeling process, consideration is given to the:* 
  + *Applicability of the data to the model purpose*
  + *Integrity and accuracy of historical data*
  + *Expected future availability and reliability of the systems providing the data.*
  + *Expected future integrity and accuracy of the data across all of the channels and segments to which the model will be applied.*
  + *Compliance with applicable business and regulatory requirements - As a general rule, internal data must be used whenever possible. If internal data are not available, or external data are necessary to fill gaps in the data, it is especially critical to assess data quality, suitability, and relevance to the model for the external data.*
  + *If official centralized Company data sources cannot be utilized and alternative data sources are utilized instead, describe the controls must be in place to ensure data integrity.*

|  |  |  |
| --- | --- | --- |
| Model Input Data | | |
| Input name | Description | Source |
| ***XXX*** | *XXX* | *XXX* |
| ***XXX*** | *XXX* | *XXX* |

* 1. Data Extraction

Describe procedures performed to collect data from these sources. Include information on how the security and confidentiality of customer information was ensured, and where the development code is centralized, if applicable.

* 1. Data Reconciliation and Quality

Quantitative and qualitative analysis and checks should be performed to assess the **quality of data**, especially as it relates to accuracy, completeness, stability, alignment with strategy described above, outlier analysis, missing values, consistency across multiple data sources, and ability to pass certain logical tests. Elements of this section may include description of:

* Data quality checks, controls, and procedures applied to ensure data accuracy and completeness.
* Review of frequency distributions of data values, high, and low values (to ensure consistency with business intuition) – Include detailed tables with variable statistics in the Appendix (as applicable)
* Review of unusual values and outliers – Describe the rationale and adjustment approach if any adjustments or overrides were applied; include measurements of the impact of excluding/adjusting outliers.
* Review of period-over-period changes in key data elements – Discuss if there are any unusual variances identified and how these potential integrity concerns were resolved.
* Manual checks to ensure automated checks in place function appropriately.
  1. Data Cleansing
* Details of any applicable scrubbing rules (merging, cleaning, filtering, and segmenting) applied to the source data, along with empirical support for the use of any missing/replace data values performed during the creation of a **relevant and complete** dataset for modeling.
* Any exclusions (e.g., exclusions for missing values, indeterminates and low-quality labels, etc.) should be clearly stated and justified. Waterfall charts are recommended for data exclusions.
  1. Data Sampling

*Detail the sampling procedures/methodology, including the impact of the sampling specification on the model estimation (e.g. sample weighting) used during the creation of the development data sets.* *Additionally, include the rationale, process for splitting and time periods of training, testing and Out-of-Time validation of the development data, if applicable.*

* 1. Development Data Testing
* *Describe and summarize the tests that were performed on development data, for example (but not limited to), data exploratory analysis, overlaying histograms[[1]](#footnote-2), seasonality testing and testing on missing value imputations.*
* *Draw conclusions/opine on each testing results.*
* *Include testing results in the documentation and specify the location of the code for all tests conducted.*
  1. Development Data and Processing/Testing Code

Provide a link to the centralized location where the data processing/testing code are saved, and the location where the final development data is saved. Code should be clearly labeled and documented to support reproduction by someone other than the model developer. Please include the following:

* Link to data and code.
* Final data testing code [Required for reproducing testing results]
* Data sampling and/or cohort creation code
* Data engineering code used to create variables.
* Data engineering code used to create the training, out-of-sample and out-of-time datasets.
* Any other relevant documentation

Please note that it is the development team’s responsibility that the development code is **executable**, if given the same environment setting.

# Model Methodology

Document the model estimation technique/methodology (e.g. theory, formulas, etc.) including support for any deviations away from standard estimation techniques and evidence that model approach is reasonable and consistent with business requirements and user specifications. For instance, the reasoning behind choosing a particular methodology for model development, its alignment with the business objectives, any industry and academic research supporting the use of this methodology for similar objectives, and/or any constraints that hinder the development team from employing other potentially superior methodologies.

Include selected functional form/statistical distribution/simulation method and consider using diagrams to illustrate complex statistical concepts discussed.

* 1. Model Theory and Approach

Descriptions of each of the model functional specifications, theoretical structure, components, data flows, and assumptions/variables.

### Model Framework Overview

*Include an overview of the model architecture including a high-level process flow diagram of the model.*

### Model Segmentation

*Describe the model’s segmentations schema, if any. Include the rationale for the selection of the model segmentation (i.e., business justification, difference in characteristics and importance ranking of key features in each segmentation). Describe any alternative segmentation schema that has been considered.*

*Document the test scripts, and results for validating the accuracy of model components.*

### Model Processing Logic

*Describe the model’s components or sub-modules. Include their relation to the intended usage (i.e., how the output of the component is utilized) and rationale for the selection of the model structure. Detail the process flow diagram of the model framework and descriptions of each component.*

*Document the test scripts, and results for validating the accuracy of model components.*

* 1. Model Features/Variables

Document and identify for all model variables/parameters a description, creation rule, source data field(s), format, decoding rules, etc. Ensure all variables/parameters are documented and support provided for their reasonability/consistency relative to model objectives/specifications. Support should include:

* Starting long list of considered variables, describe the categories of variables considered and the criteria guiding their selection.
* Step-by-step process and criteria for narrowing down the long list. Include tables summarizing the various steps in the variable selection process.

Transformations applied to the variables, including the rationale for the selected transformations. If any statistical software was applied to select optimal variable transformations, describe the underlying procedures in detail.

* Any procedures applied to derive the considered variables (e.g., calculation of ratios from raw data inputs)
* Provide data dictionary that includes the name, description, and calculation logic for derived features.

Please be sure to specify both the target variables (dependent variable(s) being solved for, if any) and the input variables (independent variables that the model/function is pulling in).

* 1. Model Assumptions/Limitations

*Document the key model assumptions/limitations that are considered for developing and utilizing the model. Include support for the reasonability/consistency with applicable business requirements or similar industry models. For limitations, indicate any remediation actions taken.*

* List all assumptions made in the model in the table below.
* Where applicable, link this to mitigants and controls in place
* Note: **Technically, all analytical assumptions need to be tested.**

**Expert Assumptions are assumptions made based on domain knowledge, experience, and intuition of data scientists or industry experts.** For example, ‘the labels provided by issuers/acquires/merchants are trustworthy and adequately represent the objective reality’, or ‘the pattern of transactions identified as fraudulent in the past will be somewhat similar in the future’.

**Analytical Assumptions are assumptions made about the structure or properties of the model itself or the underlying data.** For example, the assumptions about the distribution of data, the independence of features, or the form of the relationship between input variables and the target variable.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| # | Description | Testing Evidence | Controls | Analytical or Expert |
|  | *Briefly describe the assumptions of the model* | *Provide the descriptions of all tests conducted for supporting the assumptions and the results/conclusions of the tests.* | *Describe the relevant controls in place.* |  |

* List all limitations (shortcomings that cannot reasonably be remediated by the Model Owner and Model Developer) of the methodology in the table below including a clear description of the limitation.
* Where applicable, link this to mitigants and controls in place

|  |  |  |
| --- | --- | --- |
| # | Description | Remediation of actions |
|  | *Briefly describe the limitations of the model* | *Describe any remediation or control plans* |
|  |  |  |
|  |  |  |
|  |  |  |

* 1. Expert Consultation and Industry Research (if applicable)
* *Describe supporting insights from experts regarding the reasonability of the proposed model approach/design.*
* *Describe any relevant industry literature supporting and evaluating the reasonability of the proposed model approach/design.*
* *Include any meeting minutes form discussions with business experts and research publications in the Appendix.* 
  1. Model Methodology Testing

*Provide the results of the tests conducted for the final model.*

**Assumption Testing**

*For each assumption listed in section 4.5, conclude whether the assumption is reasonable and how well this is evidenced. If the assumption is an analytical assumption, describe any tests performed on these assumptions.*

**Model Components Testing**

*Detail any testing results conducted for each model components (e.g., performance, stability, or seasonality). Model components may encompass aspects such as, preprocessing models for determining optimal segmentations or sample weights, encoding/embedding models employed during the feature engineering process, and seasonality models utilized for determine seasonal factors.*

**Model Performance Testing**

Detail the final model performance by providing the values for all select metrics for model performance evaluation.

Description of the procedure for hyperparameter tuning process and provide evidence (this can be logs in the development code provided) that the final selected set of hyperparameters produces the optimal model performance, if applicable.

List the final set of hyperparameters/architectures of the model.

**Alternative Model Testing**

Provide results for performance evaluation for all or selected alternative models as evidence that the final champion model outperforms all its alternatives.

Describe the process and rationale of selecting the final champion model/methodology from all candidate/alternative models.

* 1. Model Specification

*Describe the functional structure of the model along with identifying essential features and their associated parameters.*

*If the model's structure and parameters cannot be distinctly outlined, alternate interpretation methods such as Partial Dependence Plots, SHAP, or LIME can be deployed. When dealing with a large feature set, focus on the key features. Similarly, when the data volume is substantial, it is acceptable to use a representative sample for analysis.*

* 1. Model Development and Testing Code

Provide a link to the centralized location where model development code is saved. Code should be clearly labeled and documented to support reproduction by someone other than the model developer. Please include the following:

* Link to code.
* Official documented scoring code for ongoing use [Required to avoid version control issues]
* Final development & scoring code used in validation [Required for reproducing testing results]
* Final model testing code [Required for reproducing testing results]
* Any other relevant documentation

# Model Output

* 1. Sensitivity, Stability and Robustness Testing

Detail results for model performance testing on the development data out-of-time, as well as segmentation analysis to assess model performance stability across various segments (i.e., region, funding source, card presented or not, etc.).

* 1. Calibration and Overlay Testing

Describe the methodology and parameters for any post-model calibrations and/or overlays. Provide rationales for transforming the raw model output to the final ready-in-production output. Details the evaluation of potential impact on any post-model calibrations and/or overlays.

Describe ongoing change plans for determining calibration and/or overlay parameters, if applicable. Include step-by-step process for any adjustments to raw model output (i.e., by Model Owner, business unit representative, etc.).

* 1. Benchmark Testing

Detail the results for benchmark testing. Provide evidence that the final champion model exhibits superior or equivalent performance when compared to any benchmark models, upstream models or current production model.

For updated versions of models that are currently in production, a benchmark comparison with the refitted version (keeping everything identical but refitting with more recent data) of the existing model is strongly recommended, unless an adequate justification is presented (e.g. the updated model involves a scope expansion or production data is not retrievable).

* 1. Bias Testing

Describe any bias testing conducted. Detail the process of evaluating if the final model is unfairly favoring certain categories or groups over others. This could be based on attributes such as gender, race, income, or any other characteristic, if related testing is conducted.

# Model Implementation

* 1. Model Implementation Processes and Procedures

Describe in detail the model’s application and configuration specifications including – technical platform specifications, sequencing, modules, reporting, etc.

Describe the process for putting the model into production, platform etc. For example, if the model is implemented in an incremental manner, describe any analysis conducted to ensure there are no major issues as a result of the new model upload. Describe where the production fields are stored.

Include formal process / procedures for maintaining and updating key modeling parameters, assumptions, data sources, computer code and parties responsible for approving the changes including individuals allowed to request a change, documentation requirements (e.g. change log), approvals, implementing and testing a change etc.

* 1. Production Input Data

*Describe source, control and management plan for data feeding into the model in production.*

|  |  |  |  |
| --- | --- | --- | --- |
| Production Data | | | |
| Source | Description | Location | Different from Development? |
| ***XXX*** | *XXX* | *XXX* | *Y/N* |
| ***XXX*** | *XXX* | *XXX* | *Y/N* |

* 1. Model Implementation Testing

Describe the implementation testing results to demonstrate that the model is developed and validated is the same model that is implemented. Implementation testing should include **both input and output testing results.** Provide evidence as well as results for implementation testing as follows.

Detail the results of input implementation testing below. If the list for input features is long, list a group of key features’ testing results.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Implementation Summary for Input** | | | | | |
| **Input Feature Name** | **Number of Cases Tested** | **Threshold for Matching** | **% of Matching** | **Max Absolute Difference** | **Max Relative Difference** |
| *XXX* | *XXX* | *XXX* | *XXX* | *XXX* | *XXX* |
| *XXX* | *XXX* | *XXX* | *XXX* | *XXX* | *XXX* |

Detail the results of output implementation testing below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Implementation Summary for Output** | | | | |
| **Number of Cases Tested** | **Threshold for Matching** | **% of Matching** | **Max Absolute Difference** | **Max Relative Difference** |
| *XXX* | *XXX* | *XXX* | *XXX* | *XXX* |

* 1. Updating Production Model

Describe system performance testing conducted as well as any plans in place to ensure production continuity in the event of system failures or implementation failure.

# Model Ongoing Monitoring

* 1. Ongoing Monitoring

Detail the ongoing performance monitoring process including the measurement metrics, acceptable ranges, and remediation actions when score distributions are not aligned. Also include reference documents in appendix, if any. Example for monitoring plan is as follows:

|  |  |
| --- | --- |
| Monitoring Plan Item | Description of requirement |
| **Model ID** | *The ID number of the model starting with ‘MD\_’ in model inventory* |
| **Model Owner Names** | *Model Owner name in the model inventory* |
| **Start Date** | *The starting date of this monitoring plan* |
| **Frequency of Monitoring** | *The frequency of the ongoing monitoring being conducted* |
| **Monitoring for Model Input(s)** | *e.g., Characteristics Stability Index, mean, median, Standard deviation* |
| **Monitoring for Model Output(s)** | *e.g., Population Stability Index, AUC, FPR, Override rate, Alert rate* |
| **Declare justification for chosen metrics and tolerance thresholds** | *List the description and rationale for selected metrics and their thresholds for ongoing monitoring.* |
| **Overlays Monitoring** | *Describe any post-model adjustments, if applicable.* |
| **Escalation protocols to Senior mgmt. for any breaches of thresholds** | |  |  |  | | --- | --- | --- | | Escalation Framework | | | | **Final Tier** | **Metrics and Thresholds** | **Action** | | **Green** | Proposed metrics and their thresholds for green | Proposed action for green | | **Yellow** | Proposed metrics and their thresholds for Yellow | Proposed action for Yellow | | **Red** | Otherwise | Proposed action for Red | |
| **Model Limitations Monitoring** | *If there are any model limitations mentioned in Section 4.5 that require ongoing monitoring, please describe the plan to monitor these limitations.* |

* 1. Business Continuity

Describe any plans in place to ensure business continuity in the event of model failure.

* 1. Model Access Management

Describe how access to the production is managed.

# Appendix

The Appendix should include, but is not limited to, the following:

* Detailed data integrity results.
* Complete list of and description of input variables considered.
* Additional supporting exhibits and detailed model development results.
* Meeting minutes from discussions with business experts.
* Implementation testing results for input.
* Model code location.
  1. Supplemental Documents
  2. Other

1. For classification problem, comparing the distributions of the values of a feature segment by different labels of the target variable. [↑](#footnote-ref-2)