**Monitoring & Governance Plan – Reducing Hospital Readmissions**

Project: Predictive Modelling – Readmissions

Model Owner: Mohamed Thalha Ahamed Ali

Validator: Independent reviewer

Last Updated: Aug 13, 2025

**Table of Contents**

[1) Purpose & Scope 3](#_Toc205940097)

[2) Key Concepts 3](#_Toc205940098)

[2.2 Risk Tiering 3](#_Toc205940099)

[2.3 Model Risk Management Framework 4](#_Toc205940100)

[2.4 Drift 4](#_Toc205940101)

[2.5 Chi-Square Test for Feature Drift 4](#_Toc205940102)

[3) Risk Tiering for the Readmission Model 4](#_Toc205940103)

[3.1 Risk Classification 4](#_Toc205940104)

[3.2 Guardrails 5](#_Toc205940105)

[3.3 Escalation Triggers 5](#_Toc205940106)

[4) Roles & Responsibilities (RACI) 5](#_Toc205940107)

[5) Model Card 6](#_Toc205940108)

[6) Monitoring Plan 7](#_Toc205940109)

[6.1 Feature Drift — Chi-square (Categorical Inputs) 8](#_Toc205940110)

[6.2 Performance Drift 12](#_Toc205940111)

[6.3 Target Rate Drift 12](#_Toc205940112)

[6.4 Data Quality & Pipeline Health 12](#_Toc205940113)

[7.0 Variables in model 13](#_Toc205940114)

[7.1 Monitored Categorical Features 13](#_Toc205940115)

[7.2 Numeric Variable Baseline & Monitoring 17](#_Toc205940116)

[8.0 Change Management 18](#_Toc205940117)

[8.1 Scope of Changes 18](#_Toc205940118)

[8.2 Change Approval Workflow 18](#_Toc205940119)

[8.3 Roles 18](#_Toc205940120)

[8.4 Documentation Requirements 19](#_Toc205940121)

[8.5 Emergency Changes 19](#_Toc205940122)

[9.0 Model Retirement 19](#_Toc205940123)

[9.1 Retirement Triggers 19](#_Toc205940124)

[9.2 Retirement Process 19](#_Toc205940125)

# 1) Purpose & Scope

This governance plan defines how we will govern, monitor, and manage the Decision Tree (Random Undersampling) model for predicting 30-day readmission risk in diabetic patients, along with the data pipeline that feeds the model.

The model is an integral decision-support tool within the hospital’s discharge planning process. Predictions are intended to be followed as part of standard workflows, with all high-risk cases reviewed by clinical staff before action.

The governance plan covers:

* The deployed Decision Tree model and any associated challenger models tested in parallel.
* The full data pipeline, from ingestion of patient encounter records to final model scoring.
* Monitoring for data drift, feature drift, target drift, performance degradation, and operational issues.
* Guardrails to ensure safe, ethical, and compliant use.
* Escalation and change management procedures when predefined thresholds are breached.

The intended audience is the project team, hospital data science group, and other stakeholders responsible for maintaining the model and pipeline in production.

# 2) Key Concepts

**2.1 Model Risk**

Model risk is the potential for incorrect or inappropriate predictions to cause adverse outcomes.  
In the context of predicting 30-day diabetic patient readmissions, model risk could mean:

* **False Negative**: A high-risk patient is predicted as low-risk, leading to inadequate follow-up and avoidable readmission.
* **False Positive**: A low-risk patient is predicted as high-risk, leading to unnecessary interventions and higher costs.  
  Model risk increases with greater model complexity, uncertainty in inputs, and the potential impact of wrong predictions on patient care and hospital resources.

## 2.2 Risk Tiering

Risk tiering is the process of classifying the potential harm of a model and deciding what controls are needed.  
We use a four-level scale: **Unacceptable, High, Limited, Minimal**.  
For this project:

* **Classification**: **Limited Risk** The model is advisory and used to support, not replace, clinical judgment.
* **Reasoning**: Predictions inform discharge planning but do not directly trigger admissions, transfers, or discharges without human review.

## 2.3 Model Risk Management Framework

This is the structured set of processes and policies for managing model risk through the entire lifecycle:

1. **Development** – Data preparation, model selection, testing.
2. **Validation** – Independent review before deployment.
3. **Approval** – Governance sign-off.
4. **Deployment** – Integration into hospital workflows.
5. **Monitoring** – Ongoing tracking of model performance and drift.
6. **Change Management** – Versioning, documentation, and safe rollout of updates.
7. **Retirement** – Controlled decommissioning when the model is no longer fit for use.

## 2.4 Drift

Drift is when the data or relationships the model learned during training change over time, reducing accuracy.  
In hospital readmissions, drift can occur due to:

* New treatment protocols.
* Seasonal or pandemic-related shifts in patient demographics.
* Coding or documentation changes in EMR systems.

We monitor three main types:

* **Feature drift** – Inputs change (e.g., discharge types distributed differently than during training).
* **Target drift** – The base rate of 30-day readmissions changes.
* **Performance drift** – Recall, precision, or AUC degrade over time.

## 2.5 Chi-Square Test for Feature Drift

Since the target variable is binary and many key inputs are categorical, we use the **Chi-square test** to detect shifts in categorical feature distributions between the training (build) dataset and ongoing production data.

* **Null Hypothesis (H₀)**: No significant difference between the distributions.
* **Threshold**: p-value < 0.05 → statistically significant drift; investigate.

# 3) Risk Tiering for the Readmission Model

## 3.1 Risk Classification

* **Category**: **Limited Risk**
* **Reasoning**:
  + The model is advisory only it provides a readmission risk score to assist clinical staff in discharge planning.
  + All model outputs must be reviewed by qualified clinical staff before patient discharge decisions are finalized.
  + Predictions are not used for fully automated decision-making.

## 3.2 Guardrails

The following safeguards ensure the model remains within the Limited Risk category:

1. **Human Oversight** – Clinical staff must review all flagged high-risk cases before action.
2. **Transparency** – Risk scores and key contributing factors (e.g., number of prior inpatient visits, discharge type) must be visible to decision-makers.
3. **Scope Control** – Model use is limited to diabetic patient readmission prediction; not to be used for other conditions without retraining and revalidation.
4. **Data Privacy** – All patient identifiers are pseudonymized; only necessary features are stored.
5. **Version Control** – Only validated model versions may be deployed; all changes require governance approval.

## 3.3 Escalation Triggers

The model’s risk classification will be escalated to High Risk if:

* Predictions are used as the sole basis for admission, transfer, or discharge without human review.
* The model is integrated into automated decision systems that bypass clinical oversight.
* It is repurposed for other high-impact medical predictions without full governance approval.
* Model performance drops below agreed thresholds (e.g., recall falls ≥ 10 percentage points from baseline for two consecutive months) and it is still used without remediation.

# 4) Roles & Responsibilities (RACI)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Activity** | **Owner** | **Data Engineer** | **Clinical Lead** | **Reviewer** | **Risk & Compliance Officer** | **IT Operations** |
| **Define business objective** | R | C | A | C | C | I |
| **Feature engineering & data preparation** | R | C | I | C | I | I |
| **Model training & documentation** | R | C | I | C | I | I |
| **Independent validation** | I | I | I | A | C | I |
| **Approval to deploy** | A | C | C | R | C | C |
| **Monthly monitoring (Chi-square drift, performance checks)** | R | C | I | C | I | C |
| **Fairness & bias review** | R | I | C | C | A | I |
| **Incident response** | A | C | C | C | C | R |
| **Change management & versioning** | A | C | I | C | C | R |

**R** – Responsible (does the work), **A** – Accountable (final authority), **C** – Consulted (gives input), **I** – Informed (kept updated)

# 5) Model Card

**Algorithm:** Decision Tree Classifier with Random Undersampling (RUS) applied to the training data for class imbalance handling.

**Training Data:**

* **Source:** *Diabetes 130-US hospitals for years 1999–2008* dataset (UCI Machine Learning Repository)
* **Records:** 66,222 patient encounters after cleaning, exclusions, grouping, and feature engineering
* **Target Variable:** Binary — 1 = Readmitted within 30 days, 0 = Not readmitted (records with >30 days excluded)
* **Features:** 45 variables including grouped admission/discharge types, diagnosis groupings, medical specialty groupings, lab results, visit history, medication status, and engineered interaction features.

**Data Split and Balancing:**

* **Split:** 70/30 stratified split on the cleaned dataset
  + Train: ~46,355 records (pre-balancing)
  + Test: ~19,867 records (untouched, imbalanced to preserve real-world distribution)
* **Balancing:** RUS applied only to the training set, producing a balanced training set of ~10,954 records (~5,477 per class).

**Intended Use:**  
To support hospital discharge planning by identifying diabetic patients at high risk of 30-day readmission. Predictions are intended to be integrated into discharge workflows, with high-risk cases reviewed by clinical staff before action.

**Initial Performance (Test Set):**

* Recall: **68%**
* ROC AUC: **0.688**
* Accuracy: **77%**
* Precision: **59%**
* F1-Score: **63%**

**Key Predictive Drivers (SHAP analysis):**

1. Number of prior inpatient visits
2. Discharge group (transfer/facility)
3. Age × Number of diagnoses – interaction Feature
4. Diabetes medication status
5. Time in hospital

**Challenger Models Evaluated:**  
Random Forest, Neural Network, Logistic Regression, XGBoost, and AdaBoost — none exceeded the Decision Tree’s recall while maintaining interpretability.

# 6) Monitoring Plan

**Reporting & Alerts:**

* Results stored in the monitoring dashboard and archived as PDF/CSV.
* Alerts sent to Owner, Reviewer, Clinical Lead, and IT Ops when thresholds are breached.

## 6.1 Feature Drift — Chi-square (Categorical Inputs)

**Purpose:** Detect statistically significant shifts in the distribution of categorical input features between the build (training) dataset and the latest scoring data.

* **Method:** Chi-square test of independence for each monitored categorical feature.
* **Threshold:** p-value < 0.05 → Flag for investigation.
* **Action:**
  + Single breach → review source data and patient cohort changes.
  + Two consecutive monthly breaches for a **critical feature** (top 10 by SHAP importance) → initiate model retraining.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Feature** | **Category** | **Expected\_Count** | **Chi2\_Stat** | **p\_value** | **dof** |
| A1Cresult | >7 | 1774 |  |  | 3 |
| A1Cresult | >8 | 3738 |  |  | 3 |
| A1Cresult | None | 38477 |  |  | 3 |
| A1Cresult | Norm | 2365 |  |  | 3 |
| acarbose | Down | 2 |  |  | 3 |
| acarbose | No | 46244 |  |  | 3 |
| acarbose | Steady | 103 |  |  | 3 |
| acarbose | Up | 5 |  |  | 3 |
| admission\_source\_grouped | Emergency | 25159 |  |  | 4 |
| admission\_source\_grouped | Newborn/Birth | 3 |  |  | 4 |
| admission\_source\_grouped | Others | 3168 |  |  | 4 |
| admission\_source\_grouped | Referral | 14503 |  |  | 4 |
| admission\_source\_grouped | Transfer | 3521 |  |  | 4 |
| admission\_type\_grouped | Elective | 9150 |  |  | 3 |
| admission\_type\_grouped | Emergency\_Urgent | 32749 |  |  | 3 |
| admission\_type\_grouped | Others | 22 |  |  | 3 |
| admission\_type\_grouped | Unknown | 4433 |  |  | 3 |
| change | Ch | 20841 |  |  | 1 |
| change | No | 25513 |  |  | 1 |
| chlorpropamide | Down | 1 |  |  | 3 |
| chlorpropamide | No | 46313 |  |  | 3 |
| chlorpropamide | Steady | 40 |  |  | 3 |
| chlorpropamide | Up | 0 |  |  | 3 |
| citoglipton | No | 46354 |  |  | 0 |
| diag\_1\_group | Blood\_Diseases | 474 |  |  | 17 |
| diag\_1\_group | Circulatory | 13654 |  |  | 17 |
| diag\_1\_group | Diabetes | 3831 |  |  | 17 |
| diag\_1\_group | Digestive | 4130 |  |  | 17 |
| diag\_1\_group | Endocrine\_Other | 1281 |  |  | 17 |
| diag\_1\_group | Genitourinary | 2415 |  |  | 17 |
| diag\_1\_group | Infectious | 1299 |  |  | 17 |
| diag\_1\_group | Injury | 3286 |  |  | 17 |
| diag\_1\_group | Mental\_Disorders | 1064 |  |  | 17 |
| diag\_1\_group | Musculoskeletal | 2446 |  |  | 17 |
| diag\_1\_group | Neoplasms | 1827 |  |  | 17 |
| diag\_1\_group | Nervous\_System | 553 |  |  | 17 |
| diag\_1\_group | Other | 1559 |  |  | 17 |
| diag\_1\_group | Other\_E | 1 |  |  | 17 |
| diag\_1\_group | Other\_V | 738 |  |  | 17 |
| diag\_1\_group | Respiratory | 4383 |  |  | 17 |
| diag\_1\_group | Symptoms | 3403 |  |  | 17 |
| diag\_1\_group | Unknown | 10 |  |  | 17 |
| diag\_2\_group | Blood\_Diseases | 1316 |  |  | 17 |
| diag\_2\_group | Circulatory | 13877 |  |  | 17 |
| diag\_2\_group | Diabetes | 6000 |  |  | 17 |
| diag\_2\_group | Digestive | 1798 |  |  | 17 |
| diag\_2\_group | Endocrine\_Other | 3813 |  |  | 17 |
| diag\_2\_group | Genitourinary | 3585 |  |  | 17 |
| diag\_2\_group | Infectious | 892 |  |  | 17 |
| diag\_2\_group | Injury | 1227 |  |  | 17 |
| diag\_2\_group | Mental\_Disorders | 1250 |  |  | 17 |
| diag\_2\_group | Musculoskeletal | 828 |  |  | 17 |
| diag\_2\_group | Neoplasms | 1253 |  |  | 17 |
| diag\_2\_group | Nervous\_System | 633 |  |  | 17 |
| diag\_2\_group | Other | 1830 |  |  | 17 |
| diag\_2\_group | Other\_E | 363 |  |  | 17 |
| diag\_2\_group | Other\_V | 847 |  |  | 17 |
| diag\_2\_group | Respiratory | 4486 |  |  | 17 |
| diag\_2\_group | Symptoms | 2152 |  |  | 17 |
| diag\_2\_group | Unknown | 204 |  |  | 17 |
| diag\_3\_group | Blood\_Diseases | 1117 |  |  | 17 |
| diag\_3\_group | Circulatory | 13417 |  |  | 17 |
| diag\_3\_group | Diabetes | 7929 |  |  | 17 |
| diag\_3\_group | Digestive | 1574 |  |  | 17 |
| diag\_3\_group | Endocrine\_Other | 4313 |  |  | 17 |
| diag\_3\_group | Genitourinary | 2772 |  |  | 17 |
| diag\_3\_group | Infectious | 852 |  |  | 17 |
| diag\_3\_group | Injury | 967 |  |  | 17 |
| diag\_3\_group | Mental\_Disorders | 1441 |  |  | 17 |
| diag\_3\_group | Musculoskeletal | 899 |  |  | 17 |
| diag\_3\_group | Neoplasms | 946 |  |  | 17 |
| diag\_3\_group | Nervous\_System | 846 |  |  | 17 |
| diag\_3\_group | Other | 1289 |  |  | 17 |
| diag\_3\_group | Other\_E | 578 |  |  | 17 |
| diag\_3\_group | Other\_V | 1640 |  |  | 17 |
| diag\_3\_group | Respiratory | 2930 |  |  | 17 |
| diag\_3\_group | Symptoms | 2085 |  |  | 17 |
| diag\_3\_group | Unknown | 759 |  |  | 17 |
| discharge\_group | Expired | 1166 |  |  | 7 |
| discharge\_group | Home | 32504 |  |  | 7 |
| discharge\_group | Hospice | 507 |  |  | 7 |
| discharge\_group | Left AMA | 272 |  |  | 7 |
| discharge\_group | Other | 12 |  |  | 7 |
| discharge\_group | Outpatient Followup | 11 |  |  | 7 |
| discharge\_group | Transfer/Facility | 9563 |  |  | 7 |
| discharge\_group | Unknown | 2319 |  |  | 7 |
| gender | Female | 24559 |  |  | 2 |
| gender | Male | 21793 |  |  | 2 |
| gender | Unknown/Invalid | 2 |  |  | 2 |
| glimepiride | Down | 94 |  |  | 3 |
| glimepiride | No | 44001 |  |  | 3 |
| glimepiride | Steady | 2102 |  |  | 3 |
| glimepiride | Up | 157 |  |  | 3 |
| glipizide | Down | 245 |  |  | 3 |
| glipizide | No | 40793 |  |  | 3 |
| glipizide | Steady | 4974 |  |  | 3 |
| glipizide | Up | 342 |  |  | 3 |
| glyburide | Down | 222 |  |  | 3 |
| glyburide | No | 41556 |  |  | 3 |
| glyburide | Steady | 4196 |  |  | 3 |
| glyburide | Up | 380 |  |  | 3 |
| insulin | Down | 5268 |  |  | 3 |
| insulin | No | 21913 |  |  | 3 |
| insulin | Steady | 14319 |  |  | 3 |
| insulin | Up | 4854 |  |  | 3 |
| max\_glu\_serum | >200 | 690 |  |  | 3 |
| max\_glu\_serum | >300 | 536 |  |  | 3 |
| max\_glu\_serum | None | 43959 |  |  | 3 |
| max\_glu\_serum | Norm | 1169 |  |  | 3 |
| medical\_specialty\_grouped | Anesthesiology/Pain | 15 |  |  | 12 |
| medical\_specialty\_grouped | Emergency/Trauma | 3188 |  |  | 12 |
| medical\_specialty\_grouped | General Practice | 10392 |  |  | 12 |
| medical\_specialty\_grouped | Medicine Subspecialty | 4373 |  |  | 12 |
| medical\_specialty\_grouped | Obstetrics & Gynecology | 449 |  |  | 12 |
| medical\_specialty\_grouped | Orthopedics | 1380 |  |  | 12 |
| medical\_specialty\_grouped | Other | 456 |  |  | 12 |
| medical\_specialty\_grouped | Pediatrics | 274 |  |  | 12 |
| medical\_specialty\_grouped | Psychiatry/Psychology | 474 |  |  | 12 |
| medical\_specialty\_grouped | Radiology/Pathology | 558 |  |  | 12 |
| medical\_specialty\_grouped | Rehabilitation | 209 |  |  | 12 |
| medical\_specialty\_grouped | Surgery | 2473 |  |  | 12 |
| medical\_specialty\_grouped | Unknown | 22113 |  |  | 12 |
| metformin | Down | 267 |  |  | 3 |
| metformin | No | 37019 |  |  | 3 |
| metformin | Steady | 8572 |  |  | 3 |
| metformin | Up | 496 |  |  | 3 |
| miglitol | Down | 1 |  |  | 3 |
| miglitol | No | 46340 |  |  | 3 |
| miglitol | Steady | 12 |  |  | 3 |
| miglitol | Up | 1 |  |  | 3 |
| nateglinide | Down | 7 |  |  | 3 |
| nateglinide | No | 46051 |  |  | 3 |
| nateglinide | Steady | 287 |  |  | 3 |
| nateglinide | Up | 9 |  |  | 3 |
| payer\_code | BC | 2334 |  |  | 17 |
| payer\_code | CH | 76 |  |  | 17 |
| payer\_code | CM | 885 |  |  | 17 |
| payer\_code | CP | 1203 |  |  | 17 |
| payer\_code | DM | 228 |  |  | 17 |
| payer\_code | FR | 0 |  |  | 17 |
| payer\_code | HM | 2762 |  |  | 17 |
| payer\_code | MC | 14343 |  |  | 17 |
| payer\_code | MD | 1578 |  |  | 17 |
| payer\_code | MP | 25 |  |  | 17 |
| payer\_code | Missing | 18595 |  |  | 17 |
| payer\_code | OG | 496 |  |  | 17 |
| payer\_code | OT | 43 |  |  | 17 |
| payer\_code | PO | 302 |  |  | 17 |
| payer\_code | SI | 28 |  |  | 17 |
| payer\_code | SP | 2173 |  |  | 17 |
| payer\_code | UN | 1209 |  |  | 17 |
| payer\_code | WC | 74 |  |  | 17 |
| pioglitazone | Down | 57 |  |  | 3 |
| pioglitazone | No | 43104 |  |  | 3 |
| pioglitazone | Steady | 3091 |  |  | 3 |
| pioglitazone | Up | 102 |  |  | 3 |
| race | AfricanAmerican | 8820 |  |  | 4 |
| race | Asian | 332 |  |  | 4 |
| race | Caucasian | 35490 |  |  | 4 |
| race | Hispanic | 972 |  |  | 4 |
| race | Other | 740 |  |  | 4 |
| repaglinide | Down | 15 |  |  | 3 |
| repaglinide | No | 45721 |  |  | 3 |
| repaglinide | Steady | 559 |  |  | 3 |
| repaglinide | Up | 59 |  |  | 3 |
| rosiglitazone | Down | 36 |  |  | 3 |
| rosiglitazone | No | 43609 |  |  | 3 |
| rosiglitazone | Steady | 2615 |  |  | 3 |
| rosiglitazone | Up | 94 |  |  | 3 |
| tolazamide | No | 46338 |  |  | 1 |
| tolazamide | Steady | 16 |  |  | 1 |
| troglitazone | No | 46353 |  |  | 1 |
| troglitazone | Steady | 1 |  |  | 1 |

## 6.2 Performance Drift

**Purpose:** Identify decline in model predictive power over time.

* **Metrics:** Recall, F1-score, ROC AUC.
* **Thresholds:**
  + Recall drops by ≥ **5 percentage points** from baseline (68%) for two consecutive months → trigger investigation.
  + ROC AUC drops by ≥ **5 percentage points** from baseline (0.688) for two consecutive months → trigger investigation.
* **Action:** Review feature drift results; if drift confirmed, retrain model with updated data.

## 6.3 Target Rate Drift

**Purpose:** Detect changes in the proportion of positive class (readmitted within 30 days).

* **Metric:** Monthly proportion of positive cases.
* **Threshold:** Change of ± 3 percentage points from baseline proportion (17.7%) for two consecutive months.
* **Action:** Investigate patient population changes, admission practices, or seasonal factors.

## 6.4 Data Quality & Pipeline Health

**Purpose**: Ensure the integrity of the incoming data stream.

**Checks**:

* Schema validation — all expected fields present with correct data types.
* Missing values in critical features:
  + For features where missing was informative during modelling (e.g., medical\_specialty\_grouped, diag\_1\_group, diag\_2\_group, diag\_3\_group), retain “Missing” or “Unknown” as a separate category.
  + For other categorical features, impute with training-mode category.
* Rare categories:
  + Categories representing <1% of the training data are merged into an “Other” bucket, unless they are clinically significant (e.g., Hospice/Expired in discharge\_group).
  + Clinically significant rare categories are retained as separate values.
* Out-of-range values: Map to “Other” or “Missing” as defined for each feature.
* Duplicate encounter records: Must be removed before scoring.
* Data timeliness: Ensure scoring data is received on the agreed schedule; stale data is flagged for review.

**Threshold**: Any critical failure halts scoring until resolved.

**Action**: IT Ops notified immediately; Owner and Data Engineer coordinate the fix.

# 7.0 Variables in model

## 7.1 Monitored Categorical Features

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Feature** | **Valid Values** | **Build % (top categories)** | **Rare-Category Rule** | **Impute** |
| A1Cresult | {None, >8, Norm, >7} | None: 83.01%, >8: 8.06%, Norm: 5.10%, >7: 3.83% | None | Mode |
| acarbose | {No, Steady, Up, Down} | No: 99.12%, Steady: 0.83%, Up: 0.04%, Down: 0.01% | Bucket <1% into 'Other' | Mode |
| admission\_source\_grouped | {Emergency, Referral, Transfer, Others, Newborn/Birth} | Emergency: 54.28%, Referral: 31.29%, Transfer: 7.60%, Others: 6.83%, Newborn/Birth: 0.01% | Bucket <1% into 'Other' | Mode |
| admission\_type\_grouped | {Emergency\_Urgent, Elective, Other\_Unknown} | Emergency\_Urgent: 70.65%, Elective: 19.74%, Other\_Unknown: 9.61% | None | Mode |
| change | {No, Ch} | No: 55.04%, Ch: 44.96% | None | Mode |
| chlorpropamide | {No, Steady, Up, Down} | No: 99.62%, Steady: 0.33%, Up: 0.04%, Down: 0.01% | Bucket <1% into 'Other' | Mode |
| citoglipton | {No, Steady} | No: 99.99%, Steady: 0.01% | Bucket <1% into 'Other' | Mode |
| diag\_1\_group | {Circulatory, Respiratory, Digestive, Diabetes, Symptoms, Injury, Musculoskeletal, Genitourinary, Neoplasms, Other, Infectious, Endocrine\_Other, Mental\_Disorders, Other\_V, Nervous\_System, Blood\_Diseases, Unknown} | Circulatory: 29.46%, Respiratory: 9.46%, Digestive: 8.91%, Diabetes: 8.26% | Bucket <1% into 'Other' | Unknown |
| diag\_2\_group | {Circulatory, Diabetes, Respiratory, Endocrine\_Other, Genitourinary, Symptoms, Other, Digestive, Blood\_Diseases, Neoplasms, Mental\_Disorders, Injury, Infectious, Other\_V, Musculoskeletal, Nervous\_System, Other\_E, Unknown} | Circulatory: 29.94%, Diabetes: 12.94%, Respiratory: 9.68%, Endocrine\_Other: 8.23% | Bucket <1% into 'Other' | Unknown |
| diag\_3\_group | {Circulatory, Diabetes, Endocrine\_Other, Respiratory, Genitourinary, Symptoms, Other\_V, Digestive, Mental\_Disorders, Other, Blood\_Diseases, Injury, Neoplasms, Musculoskeletal, Infectious, Nervous\_System, Unknown, Other\_E} | Circulatory: 28.94%, Diabetes: 17.11%, Endocrine\_Other: 9.30%, Respiratory: 6.32% | Bucket <1% into 'Other' | Unknown |
| discharge\_group | {Home, Transfer/Facility, Other/Unknown, Hospice/Expired, Left AMA, Outpatient Followup} | Home: 70.12%, Transfer/Facility: 20.48%, Other/Unknown: 5.18%, Hospice/Expired: 3.61% | Keep Hospice/Expired (<1%) as clinical exception; bucket others <1% | Mode |
| examide | {No} | No: 100.0% | None | Mode |
| gender | {Female, Male, Unknown/Invalid} | Female: 53.21%, Male: 46.77%, Unknown/Invalid: 0.02% | Bucket <1% into 'Other' | Mode |
| glimepiride | {No, Steady, Up, Down} | No: 95.02%, Steady: 4.63%, Up: 0.26%, Down: 0.09% | Bucket <1% into 'Other' | Mode |
| glipizide | {No, Steady, Up, Down} | No: 88.28%, Steady: 11.26%, Up: 0.34%, Down: 0.12% | Bucket <1% into 'Other' | Mode |
| glyburide | {No, Steady, Up, Down} | No: 95.33%, Steady: 4.25%, Up: 0.31%, Down: 0.11% | Bucket <1% into 'Other' | Mode |
| insulin | {No, Steady, Up, Down} | No: 53.80%, Steady: 43.10%, Up: 2.02%, Down: 1.08% | Bucket <1% into 'Other' | Mode |
| max\_glu\_serum | {None, Norm, >200, >300} | None: 94.83%, Norm: 2.52%, >200: 1.49%, >300: 1.16% | Bucket <1% into 'Other' | Mode |
| medical\_specialty\_grouped | {Unknown, PrimaryCare, Emergency/ICU, Surgery, Cardiology, Other, Orthopedics, Radiology, Psychiatry, Obstetrics/Gynecology, Oncology/Hematology, Endocrinology, Pediatrics} | Unknown: 47.70%, PrimaryCare: 22.38%, Emergency/ICU: 7.80%, Surgery: 5.45% | Bucket <1% into 'Other' | Missing category |
| metformin | {No, Steady, Up, Down} | No: 76.42%, Steady: 21.42%, Up: 1.45%, Down: 0.71% | Bucket <1% into 'Other' | Mode |
| miglitol | {No, Steady, Up, Down} | No: 99.64%, Steady: 0.33%, Up: 0.02%, Down: 0.01% | Bucket <1% into 'Other' | Mode |
| nateglinide | {No, Steady, Up, Down} | No: 98.50%, Steady: 1.34%, Up: 0.10%, Down: 0.06% | Bucket <1% into 'Other' | Mode |
| payer\_code | {MC, HM, BC, SP, UN, Other, Missing} | MC: 45.12%, HM: 20.94%, BC: 15.42%, SP: 6.33% | Bucket <1% into 'Other' | Missing category |
| pioglitazone | {No, Steady, Up, Down} | No: 96.37%, Steady: 3.40%, Up: 0.18%, Down: 0.05% | Bucket <1% into 'Other' | Mode |
| race | {Caucasian, AfricanAmerican, Hispanic, Asian, Other, Missing} | Caucasian: 74.23%, AfricanAmerican: 18.52%, Hispanic: 4.19%, Asian: 1.80% | Bucket <1% into 'Other' | Missing category |
| repaglinide | {No, Steady, Up, Down} | No: 97.49%, Steady: 2.29%, Up: 0.14%, Down: 0.08% | Bucket <1% into 'Other' | Mode |
| rosiglitazone | {No, Steady, Up, Down} | No: 96.75%, Steady: 3.12%, Up: 0.10%, Down: 0.03% | Bucket <1% into 'Other' | Mode |
| tolazamide | {No, Steady, Up, Down} | No: 99.62%, Steady: 0.32%, Up: 0.05%, Down: 0.01% | Bucket <1% into 'Other' | Mode |
| troglitazone | {No, Steady, Up, Down} | No: 99.95%, Steady: 0.05% | Bucket <1% into 'Other' | Mode |

## 7.2 Numeric Variable Baseline & Monitoring

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Feature** | **Mean** | **Median** | **Std** | **Min** | **Max** | **Missing %** | **Impute** |
| age | 4.3958 | 4.0 | 1.7102 | 0 | 9 | 0.0 | Median |
| num\_lab\_procedures | 43.2135 | 44.0 | 19.7282 | 1 | 132 | 0.0 | Median |
| num\_medications | 16.0299 | 15.0 | 8.0978 | 1 | 81 | 0.0 | Median |
| num\_procedures | 1.3394 | 1.0 | 1.3328 | 0 | 6 | 0.0 | Median |
| number\_diagnoses | 7.3536 | 7.0 | 1.6065 | 1 | 16 | 0.0 | Median |
| number\_emergency | 0.2283 | 0.0 | 0.7795 | 0 | 31 | 0.0 | Median |
| number\_inpatient | 0.6344 | 0.0 | 1.2514 | 0 | 21 | 0.0 | Median |
| number\_outpatient | 0.3288 | 0.0 | 0.9503 | 0 | 40 | 0.0 | Median |
| time\_in\_hospital | 4.3976 | 4.0 | 2.9355 | 1 | 14 | 0.0 | Median |

# 8.0 Change Management

**Purpose:**Ensure all modifications to the *Reducing Readmissions in Diabetic Patients* model are controlled, documented, and validated before implementation. This process maintains model integrity, compliance, and audit readiness.

## 8.1 Scope of Changes

**Changes requiring this process include:**

1. Model Modifications – algorithm change, retraining, hyperparameter tuning.
2. Feature Changes – adding, removing, or redefining input variables.
3. Data Pipeline Changes – new data sources, schema changes, or preprocessing updates.
4. Threshold/Rule Updates – changes to decision thresholds, rare-category rules, or imputation methods.
5. Monitoring Updates – adjustments to drift thresholds or risk tiering rules.
6. Infrastructure Changes – migration to new platforms or scoring environments.

## 8.2 Change Approval Workflow

1. Initiation – Owner logs a Change Request (CR) in the Change Request Log.
2. Impact Assessment – Reviewer evaluates potential effects on performance, fairness, compliance, and clinical utility.
3. Validation – Change is tested in a non-production environment using holdout data.
4. Approval – Reviewer signs off after validation and documentation review.
5. Deployment – IT/Data Engineering deploys the change to production.
6. Post-Implementation Review – Performance and monitoring metrics reviewed after 1 scoring cycle.

## 8.3 Roles

* Owner – Accountable for model performance and initiating changes.
* Reviewer – Validates the change, ensures compliance, and approves/rejects requests.
* Data Engineering – Implements technical changes and updates pipelines**.**

## 8.4 Documentation Requirements

Every approved change must update:

* Model Governance Document.
* Variable-level baselines (Sections 7.1 / 7.2).
* Monitoring & risk tiering thresholds if modified.
* Version control records (model code, configuration, parameters).

## 8.5 Emergency Changes

* Applied only if critical performance degradation or data errors threaten business/clinical safety.
* Must still be logged in CR Log after implementation.
* Reviewer must retroactively validate and approve.

# 9.0 Model Retirement

**Purpose:**

Define the criteria, process, and documentation requirements for safely decommissioning the Reducing Readmissions in Diabetic Patients model, ensuring traceability and compliance after it is no longer in use.

## 9.1 Retirement Triggers

The model will be considered for retirement when any of the following occur:

1. Performance Degradation – Persistent drop below agreed recall/ROC-AUC thresholds despite retraining.
2. Data Drift – Significant, sustained Chi-square drift in critical variables that cannot be corrected without full redevelopment.
3. Business Process Change – Hospital readmission prevention strategy changes that render the model obsolete.
4. Regulatory/Compliance Change – New healthcare data regulations requiring a fundamentally different modelling approach.
5. Model Replacement – Deployment of a newer, better-performing model after validation.

## 9.2 Retirement Process

1. Trigger Identification – Owner initiates retirement request and documents reason.
2. Impact Assessment – Reviewer evaluates operational, clinical, and compliance implications.
3. Approval – Reviewer signs off on the retirement plan.
4. Decommissioning Steps:
   * Remove the model from the production scoring environment.
   * Disable all scheduled data pipelines linked to the model.
   * Stop monitoring processes.