Model Governance Plan

Airline Satisfaction Analysis

Jagjeet Singh - 301263211

Prof: David Parent

Date: 15th August 2023

Table of Contents

[1.0 Introduction: 3](#_Toc143165718)

[2.0 Variable Level Monitoring: 3](#_Toc143165719)

[2.1 Acceptable Ranges 4](#_Toc143165720)

[2.2 Missing Values 4](#_Toc143165721)

[2.3 Variable Drift Monitoring Tolerance 5](#_Toc143165722)

[3.0 Model Monitoring, Health & Stability 6](#_Toc143165723)

[3.1 Initial Model Fit Statistics 6](#_Toc143165724)

[3.2 Parameter 1: Accuracy 7](#_Toc143165725)

[3.3 Parameter 2 : Feature Importance Shift 8](#_Toc143165726)

[3.4 Parameter 3 : False positives and False negatives 8](#_Toc143165727)

[4.0 Risk Tiering : 9](#_Toc143165728)

[5.0 References 9](#_Toc143165729)

1.0 Introduction:  
In the ever-evolving landscape of predictive analytics and machine learning, the concept of Model Governance emerges as a cornerstone for ensuring the accuracy, fairness, and ethical soundness of models deployed in decision-making processes. This report explores the integral role of Model Governance within the context of the Airline Satisfaction Analysis capstone project. Model Governance represents a comprehensive framework encompassing processes, controls, and guidelines that are strategically established to oversee the development, deployment, monitoring, and maintenance of predictive models or machine learning algorithms. Its significance within this project cannot be overstated, as it addresses critical dimensions such as accuracy, ethics, compliance, data privacy, transparency, stakeholder collaboration, and long-term value. (GOVERNANCE MODELS AND FRAMEWORKS, n.d.)

The analysis of airline satisfaction is a multifaceted endeavor, deeply intertwined with the diverse facets of the airline industry. The insights derived from this analysis hold the potential to guide strategic decisions, optimize passenger experiences, and influence broader industry trends. As such, the application of robust Model Governance principles takes on a paramount role, safeguarding the quality and credibility of findings while ensuring alignment with ethical, legal, and operational considerations.

In the following sections, we will delve into the intricacies of Model Governance and its far-reaching implications for the Airline Satisfaction Analysis capstone project. From accuracy and fairness to compliance and stakeholder collaboration, each facet of Model Governance will be examined within the context of our project's objectives and outcomes. Through this exploration, we will unveil the pivotal role that Model Governance plays in fortifying the integrity and enduring value of our analysis within the dynamic realm of airline operations and passenger satisfaction.

# 2.0 Variable Level Monitoring:

Variable level monitoring is an ongoing process that involves tracking and analyzing individual variables within a dataset. The main objective is to ensure the quality, integrity, and consistency of these variables. By closely observing different aspects of each variable's behavior, distribution, and characteristics over time, we can ensure the accuracy and relevance of the data used in our analysis. (Sultan, 2021)

A screenshot of a computer

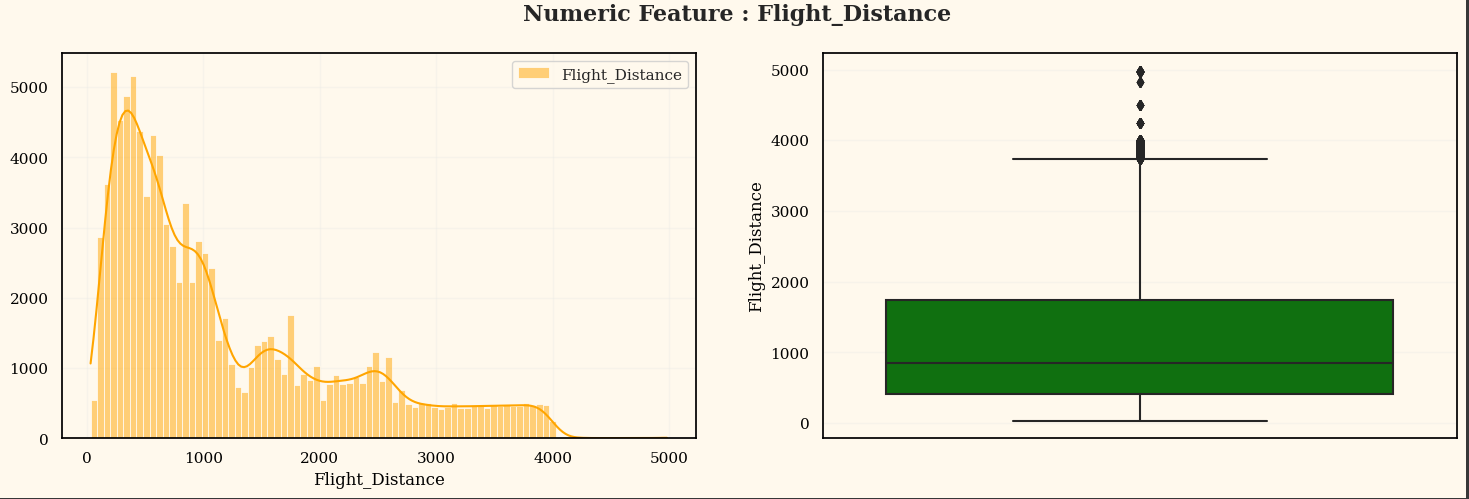
Description automatically generated

## 2.1 Acceptable Ranges

Considering the nature of the variables in our dataset and the exploratory data analysis (EDA) conducted, we have established a specific approach to variable level monitoring. In this case, we're dealing with a classification problem, and our data is derived from surveys. To maintain the authenticity of our model and ensure that the scale remains unchanged, we need to adhere to a set criterion: the maximum value should be 5, and the minimum value should be 0. It's crucial to follow this criterion throughout the survey development and data input processes. This prevents any alteration in the meaning of the data that was initially collected for building the models.

## 2.2 Missing Values

When it comes to monitoring the variables, I've devised a practical approach to tackle missing values. Given that we're dealing with a survey dataset, I've chosen to impute the missing values with the median. This decision aligns with the survey nature, where using the median proves more meaningful and appropriate compared to the mean. This thoughtful consideration ensures that our dataset's integrity remains intact. Additionally, addressing outliers is a crucial aspect of variable monitoring. In this scenario, I've opted for a strategy that involves including the outliers and not treating them with any appropriate method, because from the histogram of data we can see that we have outliers in our **“FLIGHT\_DISTANCE”** variable, and we would like to know as to what the features or important variables in long distance flight are. Because then we can devise different strategies for different flight distances.



As we transition into the modeling phase, I've implemented a strategic conversion for categorical variables. Through the process of label encoding, all categorical data is transformed into a numeric data type. This transformation significantly streamlines the imputation process and enhances the modeling efficiency.

## 2.3 Variable Drift Monitoring Tolerance

The developed models provide invaluable insights into customer satisfaction parameters, shedding light on the factors that significantly influence customer contentment. Hence, diligent monitoring of these parameters is of paramount importance to ensure the models' continued relevance to our business objectives and the accuracy of parameter determination.

Our meticulous analysis underscores the prominence of Online Boarding and Inflight Services as pivotal variables within our models. These findings form the cornerstone of our insights, underscoring the necessity to uphold the stability of these variables against drift.

To maintain the integrity of our models, we have established precise tolerance levels. Specifically, we've set a tolerance of 1 standard deviation for Online Boarding and 2 standard deviations for Inflight Services. This strategic approach allows us to detect and address deviations promptly, thus preserving the models' reliability.

From a feature importance perspective, changes in relative importance values exceeding 10% between model updates could signal a substantial shift in the influence of features on customer satisfaction prediction. Hence, we've defined an importance shift threshold of 10%.

By adhering to these rigorous monitoring and tolerance measures, we ensure the enduring accuracy and efficacy of our models in providing actionable insights for optimizing customer satisfaction within the airline industry.

# 3.0 Model Monitoring, Health & Stability

In the dynamic landscape of data analytics and predictive modeling, the ongoing monitoring, health assessment, and stability of models are essential pillars that ensure the continued relevance, reliability, and accuracy of insights generated. Model monitoring encompasses a comprehensive set of practices and methodologies that collectively safeguard the performance and utility of predictive models throughout their lifecycle.

**Importance of Model Monitoring:**

Effective model monitoring is indispensable in maintaining the quality of predictions and insights derived from models. It serves as a proactive mechanism to detect deviations, shifts, or anomalies that might arise due to changes in data distribution, external factors, or inherent model deterioration. Without vigilant monitoring, even the most meticulously crafted models could experience degradation in performance over time.

**Ensuring Model Stability:**

Model stability refers to the capacity of a model to produce consistent and reliable predictions over time. Ensuring model stability involves both continuous monitoring and the implementation of corrective measures when deviations occur. Strategies for maintaining stability include:

1. **Regular Retraining**: Periodically retraining models using fresh data helps them adapt to changing patterns and prevents them from becoming obsolete.
2. **Feedback Loops:** Incorporating feedback from users or domain experts helps in refining models and addressing discrepancies between predicted outcomes and real-world observations.
3. **Drift Correction:** When data drift is detected, taking corrective action such as retraining models on the updated data distribution is crucial to maintaining accuracy.
4. **Retrospective Analysis**: Regularly analyzing model performance over time provides insights into long-term trends, aiding in understanding the stability of predictions.

## 3.1 Initial Model Fit Statistics

One of the standout achievements of this analysis is the development of a Random Forest model, a powerful ensemble of decision trees that harnesses collective intelligence to predict customer satisfaction levels. Through meticulous exploration and a rigorous hyperparameter tuning process employing GridSearchCV, an optimal configuration has been identified:

Number of Trees: 200

Maximum Depth of Trees: Unrestricted

Minimum Samples per Leaf: 2

Minimum Samples per Split: 10

This model demonstrates exceptional predictive capabilities, achieving an accuracy of approximately 96.3%. Accuracy, a foundational performance metric, reflects the proportion of correctly predicted customer satisfaction outcomes. The model's proficiency is further illustrated through a comprehensive Confusion Matrix analysis:

|  |  |  |
| --- | --- | --- |
|  | Predicted Not Satisfied | Predicted Satisfied |
| Actual Not Satisfied | 14279 | 294 |
| Actual Satisfied | 666 | 10737 |

Interpreting the Confusion Matrix reveals the model's ability to accurately identify satisfied and unsatisfied customers, with a notably low number of false positives and false negatives.

Furthermore, the analysis has unveiled crucial insights into feature importance, underscoring the significance of Online Boarding and Inflight Services as pivotal variables impacting customer satisfaction. This knowledge empowers strategic decision-making, enabling tailored services and strategies that align precisely with customer preferences.

## 3.2 Parameter 1: Accuracy

Parameter: Accuracy is a crucial performance metric that provides insight into how effectively our predictive model is performing its task. It quantifies the proportion of correctly predicted outcomes in relation to the total number of predictions made. In the context of our Airline Satisfaction Analysis project, accuracy signifies the ability of the model to correctly classify whether a customer is satisfied or not based on the input data and features.

Tolerance or Shift: By consistently assessing this metric, we can promptly detect any potential changes or anomalies that might arise due to shifts in data distribution, model degradation, or external factors affecting the predictive process. We've set a specific threshold to trigger deeper investigation: if the accuracy value decreases by 5% or more compared to the baseline, it serves as a red flag. Such a decrease could suggest underlying issues that warrant further examination. This might involve evaluating the data quality, assessing the model's response to evolving patterns, or considering if any external factors have influenced customer satisfaction. When the accuracy value crosses this predefined threshold, it serves as a signal to delve deeper into the model's behavior.

## 3.3 Parameter 2: Feature Importance Shift

**Parameter:** Feature importance is a critical aspect of our predictive modeling process. It grants us the ability to discern which attributes or variables exert the most significant influence on our model's predictions. By understanding the relative importance of these features, we gain insights into what factors are driving customer satisfaction levels within the airline industry. This knowledge serves as a compass, guiding our strategic decisions and actions towards areas that require attention or improvement.

**Tolerance or shift**: Maintaining the stability of feature importance is pivotal for the ongoing validity of our models. Since these insights underpin our decision-making, it's essential to ensure that the importance of features remains consistent over time, reflecting the underlying dynamics of customer satisfaction.

To achieve this, we've established a predetermined threshold for importance shift at 10%. If the relative importance values of features undergo a change that exceeds this threshold, it acts as a trigger for further investigation. Such a shift might indicate alterations in customer behaviors, shifts in preferences, or changes in the competitive landscape of the airline industry.

## 3.4 Parameter 3: False positives and False negatives

Parameter: False positives and false negatives are critical elements in evaluating the performance of our predictive model. They provide insights into where our model might be making incorrect predictions and illuminate potential areas for optimization. These metrics are particularly important in understanding the model's behavior and identifying scenarios where it may need refinement.

Tolerance: Monitoring false positives and false negatives is key to identifying any shifts in the model's performance that might be affecting its accuracy. Both metrics carry distinct implications:

* **False Positives:** These occur when the model predicts a positive outcome (e.g., customer satisfaction) when the actual outcome is negative (e.g., customer dissatisfaction). In our context, false positives might represent customers incorrectly classified as satisfied when they're not, potentially leading to missed opportunities for improvement.
* **False Negatives**: These arise when the model predicts a negative outcome when the actual outcome is positive. In the airline satisfaction domain, false negatives could signify instances where customers are predicted to be dissatisfied even though they are satisfied, leading to missed chances for recognizing and nurturing positive experiences.

To ensure the model's continued optimization and accuracy, we've defined a threshold for change in these metrics. Specifically, if there is a sudden increase of 5% or more in either false positives or false negatives compared to the baseline, it serves as an alert.

# 4.0 Risk Tiering:

This table outlines the different parameters, the shifts or changes that might occur, the established tolerances or thresholds, the associated risk tiers (ranging from no action required to rebuild), and recommended actions for each scenario. Monitoring these aspects and taking appropriate actions based on the risk tier ensures the ongoing accuracy, stability, and effectiveness of your predictive models for airline satisfaction analysis.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter / Variable | Shift / Change | Tolerance / Threshold | Risk Tier | Action |
| Accuracy | Decrease > 5% |  | High (Red) | Investigate model degradation, potential retraining |
| Feature Importance | > 10% shift |  | Medium (Yellow) | Investigate reasons behind shift, adapt model if needed |
| False Positives | Increase > 5% |  | Medium (Yellow) | Investigate factors contributing to increase, refine model |
| False Negatives | Increase > 5% |  | Medium (Yellow) | Investigate factors contributing to increase, refine model |
| Variable Drift - Online Boarding | Deviation beyond ±1 std. deviation | ±1 std. deviation | Low (Green) | Monitor closely, evaluate reasons for drift |
| Variable Drift - Inflight Services | Deviation beyond ±2 std. deviations | ±2 std. deviations | Low (Green) | Monitor closely, evaluate reasons for drift |
| Hyperparameter Changes | Significant shifts | Range explored during tuning | Medium (Yellow) | Assess impact on model performance, consider re-tuning |
| Feature Importance Stability | > 10% shift | 10% | Medium (Yellow) | Investigate changes in data dynamics, adjust if needed |
| Model Stability | Fluctuations |  | Medium (Yellow) | Investigate reasons for instability, retraining if necessary |

# 5.0 References

*GOVERNANCE MODELS AND FRAMEWORKS*. (n.d.). Retrieved from James Cook University: https://jcu.pressbooks.pub/pmgovasset/chapter/module-7-governance-models-and-frameworks/#:~:text=A%20governance%20model%20or%20framework,or%20provided%20by%20the%20PMO.

Sultan, H. (2021, 03 18). *How Can You Create Ethical and Explainable AI*. Retrieved from Metrostar: https://blog.metrostar.com/ethical-and-explainable-ai