2024

CASE STUDY: ALGORITHMIC RISK AND IMPACT ASSESSMENT

DATA ETHICS AND SECURITY

Table of Contents

Risk Assessment and Mitigation Plan for Loan Request Acceptance Predictive Model

Introduction

Developing a predictive model to determine loan request acceptance carries significant implications, particularly concerning fairness, bias, and legal compliance. This risk assessment aims to identify potential risks at various project stages, propose mitigation strategies, and emphasize the importance of adherence to data ethics frameworks and legal regulations.

Project Scope

The project's primary goal, which is of utmost importance, is to develop a predictive model that assists decision-making regarding loan request acceptance. However, the nature of the dataset and the algorithms employed introduce risks related to bias, discrimination, and lack of fairness.

Dataset Preprocessing

Risk 1: The dataset consists of loan applicants who are white, not Hispanic, and male, resulting in a positive acceptance ratio. Conversely, black, Hispanic, or female requests face slight rejection. This imbalance, which is a potential risk, leads to algorithmic bias and discrimination.

- Risk Type: Data Quality and Algorithmic Bias
- Analysis: The dataset's reliance on race and ethnicity introduces inherent biases, affecting the model's fairness and accuracy.
- Mitigation: Utilizing diversified datasets that represent various demographic groups adequately can mitigate bias. Furthermore, adherence to data ethics frameworks, such as promoting just and equitable outcomes, is crucial (Data Ethics Framework, 2020).
- Risk Impacts: High, as biased outcomes can perpetuate discriminatory practices and undermine trust in the model's decisions.

Risk 2: Applicants from areas with high minority populations tend to face higher rejection rates, regardless of their income levels.

- Risk Type: Algorithmic Bias and Social risks
- Analysis: The prediction model's ratio of acceptance and rejection is based on ethnicity and race, which shows that the model design does not benefit the public but only specific ethnicities, like whites.
- Mitigation: The prediction model should follow the data ethics policy, i.e., show fairness and not determine the acceptance ratio based on ethnic background.
- Risk Impacts: High

Risk 3: several columns in the dataset contain missing values, potentially impacting model accuracy.

- Risk Type: Technical
- Analysis: missing values compromise the model's effectiveness and fairness.
- Mitigation: Implementing proper data anonymisation, privacy protection measures, and transparency in model development can mitigate these risks.
- Risk Impacts: medium, as they affect the model's ability to provide fair and accurate predictions.

Model Development

Risk 1: Lack of Accountability and Algorithmic Bias

The model's development stage is susceptible to algorithmic bias and lack of accountability, primarily due to the dataset's reliance on race and ethnicity.

- Analysis: Failure to follow data ethics frameworks and legal regulations can result in biased outcomes and discriminatory practices.
- Mitigation: Utilizing diversified datasets and aligning the project with data ethics frameworks and relevant laws, such as the Equality Acts 2018 and GDPR, is crucial. Secondly, diverse perspectives should be ensured by the development team to identify and address biases.
- Risk Impacts: High, as it undermines the model's fairness, accountability, and compliance with legal standards.

Risk 2: Lack of Compliance with Legal Regulations

Furthermore, there is a risk of non-compliance with legal regulations, such as the Equality Acts 2018 and GDPR, which may result in legal repercussions.

- Analysis: Ignoring legal regulations can lead to discriminatory practices and legal consequences for the organisation.

- Mitigation: Ensuring the model's design aligns with relevant laws and codes of practice is essential to mitigate legal risks. Document and disclose the model's decision-making process to stakeholders to achieve this.
- Risk Impacts: Non-compliance can lead to legal liabilities and reputational damage.

Model Evaluation

Risk 1: Data Bias in Model Evaluation

During model evaluation, there is a risk of bias persisting in areas with high percentages of minorities, resulting in the rejection of even high-income applicants.

- Analysis: Dataset bias based on race and ethnicity can lead to unfair outcomes, undermining the model's credibility.
- Mitigation: Using diversified datasets and ensuring transparency in model evaluation can mitigate these risks.
- Risk Impacts: High, as it perpetuates discrimination and undermines trust in the model's decisions.

Risk 2: Missing Values in Model Evaluation

Additionally, missing values in many records pose a challenge during model evaluation.

- Analysis: Missing values compromise the reliability and accuracy of model predictions.
- Mitigation: These risks can be mitigated by employing appropriate techniques to handle missing values and ensuring transparency in the evaluation process.
- Risk Impacts: High and medium, respectively, as they affect the model's reliability and transparency.

Risk 3: Model Complexity and Overfitting

Moreover, due to many categories in the "lender" variable, model complexity increases the risk of overfitting.

- Analysis: Overfitting reduces the model's generalisation ability and reliability.
- Mitigation: Simplifying the "lender" variable or employing regularisation techniques can mitigate the risk of overfitting.
- Risk Impacts: High, as overfitting undermines the model's reliability and generalisation ability.

Risk 4: Technical issue

Clipping values above certain thresholds or percentiles and replacing outliers based on interquartile range (IQR)

- Analysis: Clipping values above certain thresholds or percentiles and replacing outliers based on interquartile range (IQR) may lead to loss of valuable information or distortion of the data distribution, primarily if outliers represent genuine data points.
- Mitigation: Before implementing any outlier treatment method, thoroughly evaluate its impact on model performance using cross-validation.
- Risk Impacts: Medium

Maintenance

Over time, the absence of long-term evaluation and maintenance structures poses a significant risk to the model's performance and reliability.

- Analysis: Without proactive maintenance measures, the model may degrade over time due to shifts in data distribution and external factors.
- Mitigation: Implementing proactive maintenance measures aligned with organisational goals is essential to optimise the model's performance and reliability.
- Risk Impacts: High, as it jeopardises the model's effectiveness and reliability over time.

Conclusion

In conclusion, developing a predictive model for loan request acceptance entails various risks related to bias, discrimination, legal compliance, and maintenance. Adherence to data ethics frameworks, legal regulations, transparency, and proactive maintenance measures is crucial to mitigate these risks and ensure the model's fairness, accountability, and reliability.

References

Data Ethics Framework - GOV.UK (www.gov.uk) (accessed 06/04/2024)

General Data Protection Regulation (GDPR) – Official Legal Text (gdpr-info.eu) (accessed 06/04/2024)

Predicting mortgage approvals: Data analysis and prediction with Azure ML Studio | by Eduardo Muñoz | Analytics Vidhya | MediumPredicting mortgage approvals: Data analysis and prediction with Azure ML Studio | by Eduardo Muñoz | Analytics Vidhya | Medium (accessed 08/04/2024)

Predicting mortgage approvals: Data analysis and prediction with Azure ML Studio(Part 2) | by Eduardo Muñoz | Analytics Vidhya | Medium (accessed 08/04/2024)