Bias in Health-Auto Insurance Premiums

Ankit Tripathi[1]
ankittri@usc.edu
4612676999
[1]University of Southern California

Kshitij Harish Parab[1]
kparab@usc.edu
9459566432
[1]University of Southern California

Abstract*— *This research investigates potential biases in using auto insurance data to predict individuals' likelihood of purchasing health insurance premiums. The study aims to conduct demographic analysis on gender and age to identify biases, employ statistical metrics to assess model bias, and develop a predictive model utilizing health insurance data. The dataset contains around 450,000 samples from a health insurance company with features like age, gender, prior insurance, vehicle age, annual premiums, and customer duration. Exploratory analysis reveals representation bias in age and gender. Techniques like balanced sampling (SMOTE), algorithm selection, and excluding protected groups are used for bias mitigation. Three models (Logistic Regression, Random Forest, and XGBoost) are evaluated with and without SMOTE, and with/without excluding gender and age. Results show excluding sensitive features reduces bias metrics like statistical parity (as low as 0.0489 for Logistic Regression) and equal opportunity (up to 0.1796 for Logistic Regression), but with some accuracy loss. Random Forest achieves highest accuracy (0.8754) but more bias , while Logistic Regression shows best bias mitigation and fairness when excluding protected groups. XGBoost balances accuracy and fairness/bias reduction. The study provides insights into detecting, analyzing and mitigating biases when using personal data for predictive modeling in insurance analytics. It contributes towards ethical and socially responsible insurance practices by proposing principled approaches to eliminate biases while leveraging personal data for developing fair predictive systems.

I. Introduction

The realms of motor vehicle insurance and health insurance are extensive but intricately interlinked areas within the broad field of insurance. With this kind of data in their hands, insurers can begin to look at the relationship between health profiles and likelihood to purchase auto insurance or any other similar product. Nevertheless, such integration requires dealing with major issues regarding responsible use of personal data, elimination of biases as well as creating dependable predictive systems. Our comprehensive proposal is an attempt to lead the way in breaking new ground in insurance analytics by using existing health insurance data strategically to illuminate predictive patterns for the acquisition of car coverage. The initiative will help determine if certain age-groups and gender-health premium combinations imply one's intention on purchasing automobile insurances.

However, this goes beyond mere predictive modeling. It will give a chance to rigorously examine many angles on potential biases that might underlie system predictions about interest in auto insurance.We want to ensure fairness, transparency as well as absence of bias while employing personal health information to predict developments in car insurance. Our request is a principled data driven approach to insurance analytics which anticipates and eliminates sources of bias. The implications are significant and could contribute to advancing insurance analytics towards a more ethical and socially responsible direction. This could lead to predictive systems guiding insurers in serving diverse populations more fairly and without undue bias.

In short, through this initiative it is possible to come up with new rules surrounding responsible and non-discriminatory use of AI in the insurance industry. These findings would be crucial not only for insurers but other stakeholders like regulators, customers as well as society in general.

II. The Project Objectives

1) **Demographic Analysis**: Conduct an in-depth demographic analysis focusing on Gender And Age. The goal is to identify potential biases in the data and whether there is a change in performance metrics of the subsequent predictive model.

2) **Bias Assessment** : Employ statistical metrics and fairness indicators to evaluate the presence of biases in the predictive model. Address any identified biases to ensure the model's outputs are fair and unbiased.
3) **Predictive Modeling**: Develop an advanced predictive model utilizing health insurance data to estimate the likelihood of policyholders expressing interest in auto insurance.

III. Dataset

The dataset employed in this project has been sourced from Kaggle, containing approximately 450,000 samples. It encompasses various parameters from the Health Insurance Company's records, including the individual's age, location region, gender, prior insurance ownership status, vehicle age, annual health insurance premium, and the duration of the person's association with the insurance company. These parameters will undergo a detailed assessment to determine the presence of bias within the data. Subsequently, specific predictive models will be employed to derive insights and predictions from the dataset.

IV. Related Works

The research by Graminha and Pedro Brandao [1] explains how biases like framing, anchoring and certainty effects guide auto insurance buyers. For this reason, the researchers tested for their existence through an experiment involving 163 respondents who participated in an experiment to investigate these effects on insurance purchase. From a different perspective, the conclusion is that younger individuals, bachelors or single ladies and males have higher proclivities towards risks while payment of deductibles increased respondents' risk aversion. Finally, it probes other cognitive imperfections such as biases and heuristics that result non-optimal insurance buying decisions. Thus in order to construct a suitable choice architecture and to help achieve market efficiency as well as consumer protection for insurers, consumers, or regulators; it is important to comprehend consumer behavior and demand determinants.

To sum up, this paper provides insights into the effect of biases and heuristics on auto insurance purchase decisions, showing how individual characteristics together with cognitive bias shapes consumer choices. This study also has implications for insurance companies which highlight the importance of understanding customer behavior so as to improve efficiency of markets within the sector offering protection to them.

The research paper on checking personal bias in automobile claims settlement [2] is about the personal demographic biases and state liability rules of claims adjusters when assessing fault on bodily injury liability case for automobile accidents. Research findings show that, there is tendency to be biased towards gender and age in the assessment of fault. The research found out that young women insured defendants are rated higher levels of fault relative to other things. The study also unearthed rule-of-thumb decision making practices during the assessment of fault; hence, male insured were assigned lower percentage fault as compared to their female counterparts. Furthermore, the study established a direct relationship between severity of driving violations and assignment of fault whereby insured defendant when charged with a violation results into high rates while low rates occur when claimant is charged with a violation.

Implications of the study results can be useful to insurers trying to improve their customer service, consumer groups supporting fair claim settlement practices, and regulators. The study evidence indicates that insurer claims handling practices need to be investigated in order to assess the efficiency and fairness of victim compensation. Additionally, there is a need for more robust data on insurance companies' claims management strategies that incorporate demographic variables into insurer practice analysis. Also, in addition, when doing business online as shown by this research paper it can be seen that such systems are not dependent upon individual decision makers thereby decreasing personal bias in business decisions.

In conclusion, this research provides valuable insights on potential biases in claims adjustment practices and highlights the importance of further investigation regarding the relationship between comparative negligence rules and fault assignment. More importantly though, this evidence suggests that both demographic biases and legal rules may affect an adjuster's determination of fault with consequences for the fairness or efficiency of claims settlement processes.

In the paper [3], the researcher discusses the topic of Bias and fairness in Auto insurance premiums where biases from an actuarial perspective are discussed and industry wide discrepancies. This study employs advanced statistical models such as GLMs and the three-way analysis with data from the motor vehicle statistical plan to show that there is a significant gap between loss costs and premiums especially due to higher loadings on urban drivers. Delving into the intricacies of bias and fairness within auto insurance rates, the study illuminates the challenges of traditional actuarial fairness rooted in company-level data, aiming to minimize mean square errors across sub-populations. The subjective nature of fairness unfolds, revealing differing opinions between insured individuals and insurance companies, encapsulated by a shift from collective risk-based to individual risk-based fairness.

The essence of regulatory rules cannot be overlooked, because even though there are no excessive premiums at the industry level, traditional actuarial fairness was far from reality due to a shift from collective risk- based to individual risk- based fairness. In order for well-informed decisions on insurance prices this research makes suggestions on how policy makers can base their judgement on auto insurance equity by using a robust statistical framework such as three way analysis. Noteworthy findings underscore a substantial gap between loss costs and premiums, with urban drivers prominently exhibiting higher loadings.

V. METHODOLOGY

Our approach to detect, analyse and mitigate bias involves key steps like- Thorough analysis of the columns of the dataset will be done to recognise potential bias inducing features. Various potential biases can be present in the dataset such as Gender Bias, Age bias, Demographic bias, Selection Bias. Initially, utilization of performance metrics like Statistical Parity and Equal Opportunity to evaluate bias in the potential bias-inducing variable. Methods like Resampling, Bootstrapping can be used to address these biases. Once the bias inducing variable is identified it will be removed from the model training so that the model do not pickup the bias in its prediction. In this way, we can mitigate the level of bias in the prediction of model. To tackle the Gender bias, we aim to analyse whether a specific gender are treated inappropriately (for eg:- Female have higher insurance premiums). On the Other hand we will also analyse the Age demographic among the dataset where whether a certain age group are charged with higher insurance premiums or vice versa. This will help us understand the Potential Biased features and help build more fairer models.

To assess the model's capability in detecting biases, particularly concerning demographic groups, we'll use the following metrics:

1) Statistical Parity: Comparing the proportion of positive outcomes between different demographic groups.Formulated as the absolute difference between the probabilities of positive outcomes for each group being less than or equal to a predefined threshold.
2) Equalized Opportunity: Assessed by comparing the true positive rates (sensitivity) between different demographic groups. This means ensuring that individuals from various demographics have an equal likelihood of being correctly classified as positive by the model.

The type of bias we are hoping to encounter and mitigate are the Gender Bias, Representation Bias, Demographic Bias.

VI. IMPLEMENTATION

A. Pre-Processing

The 'Vehicle_Age' column contains categorical data representing the age of the vehicle. To incorporate this information into the model, one-hot encoding was performed, which converts categorical variables into binary vectors. Each category of 'Vehicle_Age' was represented as a binary feature. The 'Annual_Premium' column contains the annual premium paid by the policyholder. To address outliers and extreme values, we computed the average Annual Premium and transformed continuous values into categorical ones by assigning a value of 1 to those greater than the mean and 0 to those lesser than the mean. The 'Vehicle_Damage' column indicates whether the vehicle has been damaged in the past (Yes/No). To convert this categorical feature into a numerical one, 'Yes' was mapped to 1 and 'No' to 0.

B. Exploratory Data Analysis & Bias Assessment

Gender can affect car insurance rates due to statistical data showing that male drivers tend to be involved in more accidents and receive more traffic violations than female drivers. Insurance companies use this data to determine risk and set premiums accordingly. Following is the Representation of Sex in the dataset-

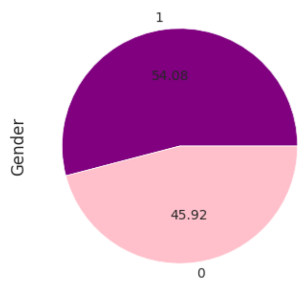

Fig. 1. Gender Representation

As the Sex Ratio was 54.08% in the favor of Male and 45.92% in the favor of Female indicates that there is little to None Gender bias in the dataset.

Age is a significant factor that affects insurance rates for vehicles. Younger drivers, especially those under the age of 25 , typically have higher insurance rates because they are considered more risky to insure. Older drivers, on the other hand, may be eligible for lower insurance rates due to their experience and history of safe driving. Following is the Presentation of Age in the dataset-

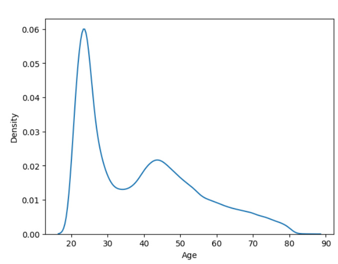

Fig. 2. Age Representation

There is significant Representation Bias in the training dataset where the people in the age between 20-30 comprises of about 60% of data which is not a true representation of the population.

Age of the Vehicle can be used to auto predict the insurance premiums of the vehicle upon how old the car is. Following is the representation of the age of vehicles and there tendency to buy the auto insurance premiums.

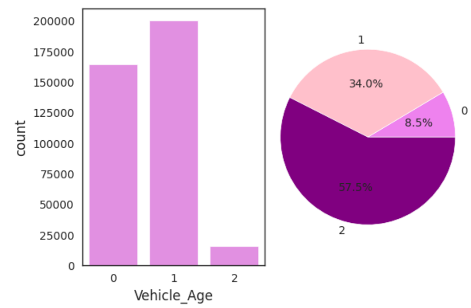

Fig. 3. Age of Vehicle Representation

The subgroup comprising vehicles aged over two years appears notably underrepresented in the dataset, serving as an illustration of representation bias within the dataset. The analysis of vehicle age emerged as a noteworthy aspect within the dataset, showcasing a relatively low frequency yet a significantly heightened propensity for auto insurance purchases.

The dataset exhibits a high level of imbalance, with approximately 87.7% of individuals showing disinterest in auto insurance. To mitigate this issue, counter-balancing techniques such as Over Sampling were used to balance the dataset. To address the data imbalance among groups, various data balancing methods such as Under Sampling, Over-Sampling, and SMOTE technique were employed. However, the SMOTE technique yielded the most favorable results. Following is the dataset class representation before and after the SMOTE technique was used-

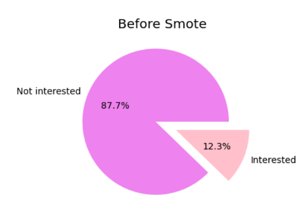

Fig. 4. Class labels Before SMOTE vs After SMOTE

C. Bias Mitigation

Following are the Mitigation Bias techniques used to address the issues-

- Balanced Sampling: Adjust the dataset to ensure a balanced representation of different classes, particularly in cases of class imbalance like Over-Sampling, SMOTE technique.
- Algorithm Selection: Try to explore and use algorithms which are less bias or better suited to handle imbalanced data and which works best on the balance dataset as-well.
- Implementing the models after converting the numerical values into categorical values for better adaptability for the models.
- Compute bias detection metrics such as Statistical Parity and Equalized Opportunity on the model excluding the protected groups, namely Age and Gender, and assess its impact on bias.

VII. Results

Following are the Tables of the Performance of 3 models Logistic Regression, Random Forest and XgBoost without SMOTE and With SMOTE on the Dataset for 3 parameters:

- Entire dataset without any removal of Protected groups
- Without the Protected Group Age
- Without the Protected Group Gender

TABLE I
Logistic Regression Results

Method	Accuracy	Statistical Parity	Equalized Opportunity
1. Logistic Regression without smote	0.7844	0.2312	0.2732
2. Logistic Regression without smote no gender	0.7722	0.0489	0.1796
3. Logistic Regression without smote no age	0.7760	-0.1210	-0.1718
4. Logistic Regression with smote	0.7842	0.2283	0.2699
5. Logistic Regression with smote no gender	0.7714	0.0520	0.1753
6. Logistic Regression with smote no age	0.7756	-0.1202	-0.1737

For Logistic Regression, SMOTE had minimal impact on accuracy (0.7844 without vs. 0.7842 with). Without SMOTE, bias was slightly higher (mean difference 0.2312 vs. 0.2283) and fairness was slightly lower (disparate impact 0.2732 vs. 0.2699). Excluding gender slightly reduced accuracy (0.7722 vs. 0.7760 without age) but significantly lowered bias (mean difference 0.0489 vs. -0.1210) and improved fairness (disparate impact 0.1796 vs. -0.1718).

For Random Forest, without SMOTE accuracy was marginally higher (0.8732 vs. 0.8739) but bias was slightly higher (mean difference 0.1528 vs. 0.1521) and fairness was slightly better (disparate impact 0.2190 vs. 0.2174). Excluding gender led to a slight accuracy improvement (0.8754 vs. 0.8739 without age) while significantly reducing bias (mean difference 0.1048 vs. -0.1596) and improving fairness (disparate impact 0.1980 vs. -0.1841).

TABLE II
Random Forest Results

Method	Accuracy	Statistical Parity	Equalized Opportunity
1. Random Forest without smote	0.8732	0.1528	0.219
2. Random Forest without smote no gender	0.8754	0.1048	0.1980
3. Random Forest without smote no age	0.8739	-0.1596	-0.1841
4. Random Forest with smote	0.8739	0.1521	0.2174
5. Random Forest with smote no gender	0.8751	0.1044	0.1969
6. Random Forest with smote no age	0.8737	-0.1588	-0.1862

TABLE III
XGBoost Results

Method	Accuracy	Statistical Parity	Equalized Opportunity
1. XGBoost without smote	0.8434	0.1807	0.2348
2. XGBoost without smote no gender	0.8416	0.0893	0.1944
3. XGBoost without smote no age	0.8406	-0.1805	-0.1920
4. XGBoost with smote	0.8414	0.1766	0.2316
5. XGBoost with smote no gender	0.8388	0.0853	0.1917
6. XGBoost with smote no age	0.8397	-0.1845	-0.1965

For XGBoost, without SMOTE accuracy was marginally higher (0.8434 vs. 0.8414) but bias was slightly higher (mean difference 0.1807 vs. 0.1766) and fairness was slightly better (disparate impact 0.2348 vs. 0.2316). Excluding gender slightly improved accuracy (0.8416 vs. 0.8406 without age) while significantly reducing bias (mean difference 0.0893 vs. -0.1805) and improving fairness (disparate impact 0.1944 vs. -0.1920).

Random Forest models achieved the highest accuracy among the three models, but they exhibited the highest bias and lowest fairness scores. Conversely, Logistic Regression models demonstrated the best bias mitigation and fairness, especially when gender and age features were excluded, but at the cost of lower accuracy compared to Random Forest and XGBoost.

The inclusion of sensitive features like gender and age consistently amplified bias and reduced fairness across all three models, even though they may have contributed to higher accuracy. Excluding these features led to a noticeable reduction in bias and improved fairness metrics, particularly for Logistic Regression and XGBoost models.

XGBoost models struck a balance between accuracy and fairness/bias mitigation, with performance metrics falling between those of Random Forest and Logistic Regression. This makes XGBoost a potential compromise when both accuracy and fairness/bias mitigation are important considerations.

The use of SMOTE (Synthetic Minority Oversampling Technique) had a minimal impact on the accuracy, bias, and fairness metrics for all three models, suggesting that it may not be effective in significantly improving performance or mitigating bias for the given dataset.

Positive values of the mean difference and disparate impact metrics indicated potential bias in favor of the majority or privileged group, and vice versa. For Logistic Regression and XGBoost models without age features, the negative values of these metrics indicated potential bias against the minority or underprivileged group based on age.

The choice of the most appropriate model depends on the specific requirements and priorities of the problem at hand. If accuracy is the primary concern, Random Forest models may be preferred. However, if fairness and bias mitigation are critical, Logistic Regression models without sensitive features like gender and age may be the better choice, even at the cost of slightly lower accuracy.

VIII. Conclusion and Future Work

This study successfully developed and evaluated predictive models to estimate health insurance purchase likelihood using personal auto data. Key biases related to gender, age, and representation were identified and mitigated through techniques like balanced sampling and excluding protected attributes. Results showed trade-offs between accuracy, bias mitigation, and fairness across models. Logistic Regression achieved best fairness and bias reduction when excluding sensitive features like gender/age, though with some accuracy loss compared to Random Forest and XGBoost. XGBoost provided a balance between accuracy and fairness considerations. The research highlights the importance of detecting and mitigating biases for ethical, responsible use of personal data in insurance predictive modeling. Future opportunities include exploring advanced bias mitigation techniques, such as adversarial debiasing, causal modeling, or constrained optimization approaches, incorporating domain knowledge, expanding bias assessment scope, real-world model deployment/validation, and investigating regulatory/policy implications - all towards advancing fairness, transparency, and ethical practices in leveraging data and modeling to better serve diverse populations.

IX. Github Link

Code Link

References

[1] Graminha, Pedro Brandão, and Luís Eduardo Afonso. "Behavioral economics and auto insurance: The role of biases and heuristics." Revista de Administração Contemporânea 26 (2022): e200421.

[2] Doerpinghaus, Helen, Joan Schmit, and Jason Jia-Hsing Yeh. "Personal bias in automobile claims settlement." Journal of Risk and Insurance 70.2 (2003): 185-205.

[3] Xie, Shengkun, Rebecca Luo, and Yuanshun Li. "Exploring Industry-Level Fairness of Auto Insurance Premiums by Statistical Modeling of Automobile Rate and Classification Data." Risks 10.10 (2022): 194..

[4] Lanes, Stephan F., and Cynthia de Luise. "Bias due to false-positive diagnoses in an automated health insurance claims database." Drug safety 29 (2006): 1069-1075.

[5] Meyers, Gert, and Ine Van Hoyweghen. "Enacting actuarial fairness in insurance: From fair discrimination to behaviour-based fairness." Science as Culture 27.4 (2018): 413-438.