An investigation of how personal sensitive information impact the effectiveness of targeted online advertising

Author: Nadia Nickel
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT,
In today's advertising world, the method of targeting customers to effectively place ads to relevant customer groups is a common tool. However, when customers are targeted, this happens based on personal information that advertisers use to define their target groups. In this context, the issues of biases concerning gender, age, or ethnicity arise, which can impact the advertising in terms of privacy concerns or discrimination, which can be seen in Google Ads. This study researched how effective target online advertising can remain, when attributes like age or gender, are excluded. Therefore, these attributes are considered as protected attributes, which are used as mitigation strategy to test the effectiveness of targeting. In more detail, a dataset provided by the company IBM was used, that contained information of 1,443,140 observations with nine different variables that characterize each customer, like age, politics, or income. The analysis of the data has shown that target advertising with sensitive information is more effective than without them, however, it indicated that target advertising is still possible for marketers but gets less precise. Hence, this shows that the targeting strategies do not have to violate data protection regulations and privacy desires of customers, to conduct successful ad placements.

Graduation Committee members:
Dr. Agata Leszkiewicz
Dr. Hatice Kizgin

Keywords
Bias in online advertising, effectiveness of target advertising, sensitive information, mitigation strategy, protected attributes, privacy regulations

1. INTRODUCTION

In today's digitalized world, online advertising plays a crucial role as a marketing tool for organizations. In the last couple of years, there has been a shift in how organizations place their advertisements, coming from traditional advertising in newspapers, door-to-door sales, or paper-based brochures to social media ads, email newsletters, and digital websites (Hanlon, 2019). However, major issues that come along with online advertising are biases that can heavily influence the effectiveness of the advertising. Research has shown that there are many different angles when it comes to bias, for example, biases that are based on protected attributes like age, gender, or ethnicity (Hamman, Chen & Dutta, 2023), but also algorithmic biases that include the use of unrepresentative datasets, imprecise models or even non-suitable algorithms that bring human biases into artificial intelligence (AI) learning (van Giffen, 2022).

This research will focus on the effectiveness of advertising, which aims to help measure whether advertisements and campaigns achieved the initial goals that the marketers intended. Assija et al. present several characteristics that indicate effective algorithms, which include amongst others the precision of the targeting. This means that if ads are presented to the most relevant customer group, the algorithm is considered effective. In this context, also the presented ads need to be tailored to the customer's online behavior and interactions to match their preferences and make the ads more personalized. Furthermore, algorithmic efficiency also includes the minimum use of resources, which means that algorithms need fewer computations to define the target audience for an ad. The use of a minimum number of parameters within the data sets is also considered an indicator of algorithmic efficiency since they require less time and resources for computations and take up less memory space (Assija, Baliyan & Jain, 2018 and Hernandez & Brown, 2020).

In order to measure how effective a campaign was, Estévez and Fabrizio defined that more complex approaches that take message relevance, emotional resonance, or brand familiarity into account are best used. To make advertisement strategies from marketers more effective, it is crucial to analyze consumer preferences and behavior by using surveys, focus groups, and sales data analysis, where variables such as gender, age, religion or educational background help to target the essential groups (Estévez & Fabrizio, 2014). When considering effective advertising, in the context of digital marketing effective algorithms need also be considered.

In the context of algorithmic efficiency, the use of Artificial Intelligence (AI) becomes more and more useful for marketers, since AI can take far more variables into account when targeting customers than human analysts would do (Akter, 2022). However, AI is yet not free from biases, meaning that even if AI takes more variables into account, the outcome can still be affected because of insufficient machine learning, meaning that the training data that has been used to train the systems can have a biased outcome because the data might have been pre-selected and categorized by possibly biased humans (Akter, 2022). The use of AI technologies in business practices such as Marketing is becoming more and more frequent since nowadays AI has reached a development stage where it can carry out human-like activities like decision-making based on given datasets (van Giffen, 2022). An example of this kind of advertising is Google Ads, where marketers define criteria like gender, race, age, educational background, or criminal record in advance to specify their target group for their ads (Sweeney, 2013). The marketers provide Google with the ad texts or images which are then published either via Google AdWords or Google AdSense, where with Google AdWords advertisers define search keywords so that when users enter these keywords, the matching ads occur. On the other hand, Google AdSense matches ads based on the search content and website visits. Since Google evaluates every search and website visit, occurring ads vary with every new search entry so that one user gets different ads even when they enter and visit the same websites (Sweeney, 2013). However, this dynamic advertising approach of Google AdSense also integrates biases like discrimination. Sweeney discovered that Google ads concerning arrest records occur more often in combination with names that are black-sounding (an example is Latanya Farrel) than white-sounding names (an example is Jill Foley). Therefore, Google ads is a representative example of racial discrimination within online advertising. Furthermore, it needs to be considered what sensitive information (gender, ethnicity, or educational background) can be used for targeted advertising and which are seen as too critical to use (Hamman, Chen & Dutta, 2023 and Speicher, Ali, Venkatadri, Ribeiro, Arvanitakis, Benevenuto, Gummadi, Loiseau & Mislove, 2018).

1.1 Research Objective and Question

This thesis aims to illustrate how effective advertising can be when limiting the use of sensitive personal information to specify target audiences. The investigation will be based on datasets that include the use of that information and how effective this kind of advertising is in terms of the likelihood of a consumer being affected by the ad. Therefore, the following research question will help to address the objective:

What is the impact of the mitigation strategy of protected attributes on the effectiveness of targeted online advertising?

1.2 Academic and Practical Relevance

When considering the academic relevance of this research, literature already presents that using sensitive personal information in targeted advertising (Estévez & Fabrizio, 2014) and that personally tailored ads to customers' preferences (Assija, Baliyan & Jain, 2018) is considered effective targeted online advertising. This thesis will therefore create a strategy that aims to be implemented in Digital Marketing practices so that marketers can create ad campaigns that reach the highest effectiveness (Estévez & Fabrizio, 2014), but also take into consideration legal concepts that deal with respecting consumers' privacy or anti-discrimination laws (Wang, Yang, Chen & Zhang, 2015).

Besides the academic relevance, this thesis is also relevant to actual business practices in the field of Digital Marketing. Research papers where existing biases in advertisements are explained and analyzed (e.g., Wang, 2023, Asadi, 2023, Tran, 2020) have shown that in several sectors, stereotyping is especially a major issue with advertising for specific products (Asadi, 2023). When implementing the mitigation strategy of protected attributes in practice, this means that sensitive information about consumers is respected, which can lead to higher trustworthiness towards the product or company of the ad campaign (Speicher, Ali, Venkatadri, Ribeiro, Arvanitakis, Benevenuto, Gummadi, Loiseau & Mislove, 2018). However, investigating how effective targeted online advertising is can bring value to the advertisers to understand how important demographic factors like gender or age are, and to improve ad campaigns to reach out to a more diverse audience.

1.3 Outline of this Report

The next section of this report will focus on reviews of the existing literature that is relevant to the research question and what concepts will be used. Following this section the methodology will be discussed and the findings and results will be analyzed. Ensuing, the limitations (practical and theoretical) will be discussed, and the report will close with answering the research question and making recommendations for further research. In this report observational data from the company IBM will be used.

2. THEORETICAL FRAMEWORK

2.1 Online Behavioral Advertising

In online advertising, “advertisers are increasingly monitoring people’s online behavior and using the information collected to show people individually targeted advertisements" (Boerman, Kruikmeier & Borgesius, 2017). This concept of online behavioral advertising is essentially for organizations to reach their target group with their advertisements. In this way, all advertisements will be more personalized, meaning based on people’s online preferences, but it raises concerns about data protection and privacy as well since personal data is used to select consumer preferences and track their online behaviors according to Boerman, Kruikmeier & Borgesius. Despite other personalized and customized advertising that uses all personal data available (not only information from online preferences and behavior), Online Behavioral Advertising conducts only information about consumers from their online behavior. This information collection is based on cookies or device fingerprints, where cookies store websites or information about users search entries and preferences in the device like the phone or computer, whereas device fingerprinting stores information about the device and the browser itself (Boerman, Kruikmeier & Borgesius, 2017). Dehling et al. provide the consumer perspective of Online Behavioral Advertising, where criteria like awareness or attitudes towards this approach were investigated. This study concluded that consumers have overall limited awareness about how this approach works, express concerns about their privacy and the transparency of how and to what extent data gets stored, but also that targeted advertising is seen as disruptive (Dehling, Zhang & Sunyaev, 2019).

2.2 Mitigation Strategies

When it comes to reducing biases in advertising, mitigation strategies are approaches that can be implemented, especially when personally sensitive information is involved (Mazumder & Singh, 2022). These strategies aim to limit bias within the algorithms that are used to target a desired audience by marketers so that the outcomes of the algorithms are as fair as possible and do not violate concepts like protected attributes (van Giffen, Herhausen & Fahse, 2022). Multiple mitigation strategies address bias in advertising, such as fairness in algorithms which can be achieved by using diverse datasets for training (Ferrara, 2023), and less reliance on standardized techniques when using AI in advertising so that AI is not using its optimized capabilities and evaluates more human-like (Saeedi, Fong, Mohanty, Gupta & Carr, 2021) or that advertisers are aware of possible bias in datasets and actively correct them (Saeedi, Fong, Mohanty, Gupta & Carr, 2021). This research will, however, focus on the mitigation strategy of protected attributes (Hamman, Chen & Dutta, 2023).

2.2.1 Protected Attributes

The concept of protected attributes refers to sensitive information that exist about people, which includes gender, age, religion, or ethnicity (Hamman, Chen & Dutta, 2023). This means that personal information is protected legally and ethically, in terms of discrimination uses and are therefore also protected from being used in any kind of decision-making, including advertising. However, this kind of information is still relevant for marketers to target the desired audience, since ad campaigns aim to reach out to specific groups that are defined by gender, age, or educational levels. Therefore, when using this kind of sensitive information, advertisers need to follow laws and regulations set by the government to not use private information without the consent of the users and to make sure to target within legal possibilities (Speicher, Ali, Venkatadri, Ribeiro, Arvanitakis, Benevenuto, Gummadi, Loiseau & Mislove, 2018).

2.3 Social Identity Theory

The Social Identity Theory created by Tajfel and Turner (1986) can be used by marketers to identify target groups based on the existing group perceptions of the customers. Dimofte et al. (2014) stated that consumers categorize themselves into groups in which they identify themselves, but also put others into groups based on how they see them. Since this is a psychological approach, people develop their self-esteem and perception by seeing their social group’s positive advantages when comparing their group to characteristics from other groups. In the context of advertising, consumers can be influenced by ads in a way that they idealize a certain social group or image that belongs to this specific group (Dimofte, Goodstein & Brumbaugh 2014 and Rawal & Torres, 2017). Furthermore, this idealization of advertisements can threaten the perception of other social groups, meaning that when a specific lifestyle of a group is presented, groups that do not belong to this group can feel left out and excluded which can result in lowering the self-esteem of the groups.

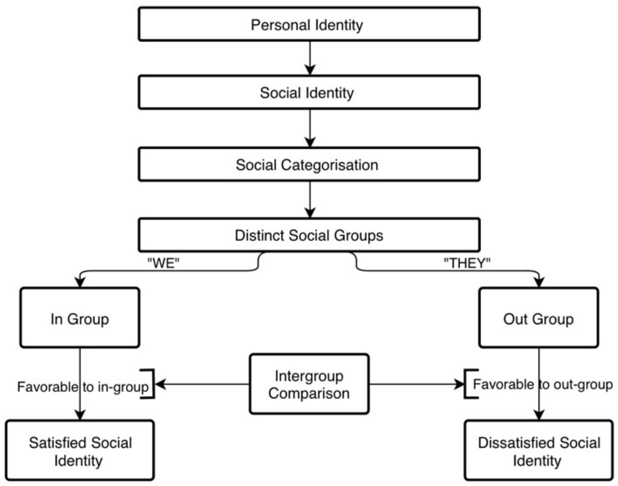

Figure 1: Social Identity Theory (Tajfel & Turner, 1986)

2.4 Privacy

Since ethical considerations concerning the use of algorithms in AI systems for online advertising, regulations concerning privacy and transparency need to be considered. AI systems use personal data from consumers to make sound recommendations for their online behavior (e.g. van Giffen, 2022, and Lee, 2018).

2.4.1 Privacy in Targeted Advertising

In targeted advertising, marketers use demographic information about consumers so that the advertisements reach the intended target group. Therefore, customer preferences and online behavior are tracked, and sensitive information is used in this online advertising technique, which is proven to increase the efficiency and effectiveness of marketing ads (Juels, 2001). However, when using sensitive consumer information, there is the issue of data and privacy protection, which can be addressed by using several approaches so that marketers can target customers but respect the privacy of their information. Wang et al. stated that information that can identify a person (such as exact names) are removed so that information is anonymized but still give relevant information to marketers to target the intended audience. Another approach is that consumers need to give their consent to advertisers before they are used. This approach gives marketers the advantage that they can freely use the provided information because consumers already gave their consent. Another privacy and data protection concern in this context is the threat of third-party violations or data breaches, which can be ensured by using cryptography techniques like encryption, where information is encoded by an algorithm and the information stays between two parties, the information provider (consumer) and the information receiver (advertiser) (Wang, Yang, Chen & Zhang, 2015).

2.4.2 GDPR in Online Advertising

The General Data Protection Regulation was introduced by the EU to overcome the privacy issue when it comes to personal data processing. This means that this regulation targets overcoming uncertainties (Rupp and von Grafenstein, 2024) when, for instance, organizations use personal data from their consumers for advertising purposes. However, this regulation sets out a common law within the EU, it is still unclear when data is considered personal consumer data and when it is not and can therefore be used without following GDPR requirements (Rupp and von Grafenstein, 2024).

2.4.3 Privacy Protection in AI

Using AI for advertising purposes can give several advantages, for instance, it can process larger data sets (Big data) that are used to identify the target audience for ads (Boerman et al. 2017). These technologies, however, expose issues in terms of data and privacy protection. The AI is trained with datasets that give the system as much information as possible so that it can develop a deep learning process (Li et al. 2023). When feeding the system with such data, especially data that is related to consumer information, these datasets might have personal data that are protected by privacy regulations. To secure data from potential malicious threats, several techniques like differential privacy add random and not real data to the dataset, so that it is unclear what data are true and confidential (Li et al. 2023), however, this might reduce the accuracy of an algorithm to detect target consumers. Another technique is multiparty computation (MPC) which cryptograms the data so that it is not visible to another party (Li et al. 2023).

2.5 Conceptual Framework

In order to address the research question of how effective targeted advertising can be when limiting the use of protected attributes as a mitigation strategy. Therefore, independent, dependent and moderator variables can be defined which means that the mitigation strategy of protected attributes is considered as an independent variable, which makes demographic characteristics (such as age, gender, or ethnicity) the moderator variable that impacts the relationship between the independent and dependent variables. The dependent variables is conclusively the effectiveness of targeted online advertising, which can be measured through click-through rate (CTR), sales revenue, or the conversion ratio (number of customers that clicked on an ad).

With this context, the following conceptual framework has been designed:

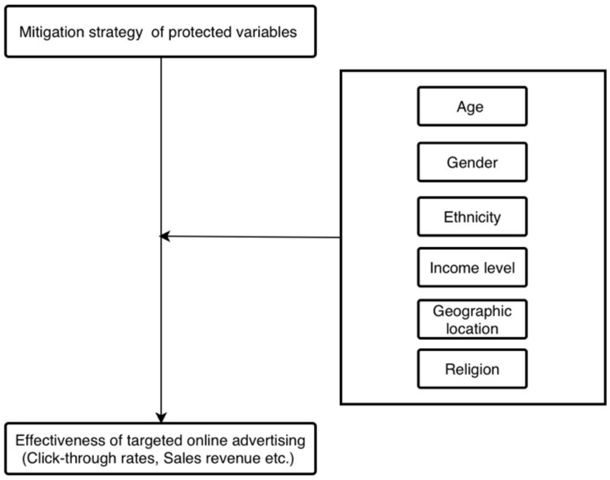

Figure 2: Conceptual framework

3. METHODOLOGY

3.1 Research Design

To answer the research question of how effective targeted advertising can be when limiting the use of protected attributes as a mitigation strategy, it is necessary to define an appropriate research method. Therefore, secondary research and data will be used, where these data are already collected by others and are openly available (Chrysochou, 2017). The data sets that contain information about bias in advertising and the effective rate of targeted advertising using personal information are provided by IBM. In this context, the key factors when using this data are effectiveness parameters such as the click-through rate and the number of protected attributes that are available that belong to the consumers.

3.2 Methods

To address the stated research question, a dataset about bias in online advertising is provided by the company IBM (https://dax-cdn.cdn.appdomain.cloud/dax-bias-in-advertising/1.0.0/bias-in-advertising.tar.gz). Within this data, in total 1,443,140 observations are stated with nine parameters from which some can be considered as protected attributes since they contain personal information about the consumer (see Table 1 and Appendix 11.1.1). Based on these parameters, a predicted probability was computed to show how probable it is a customer clicks on an ad, and then considering these probabilities it is predicted if the ad gets clicked or not. Furthermore, the actual outcome whether the ad was clicked or not was also included in the dataset, so it can be measured if there is a connection between clicking the ad and the personal information that was available about the consumer.

Table 1: Variables of the dataset

Variables	Levels	Privacy Status
college_educated	1, 0 (Yes, No)	No
parents	1, 0 (Yes, No)	No
homeowner	1, 0 (Yes, No)	No
religion	Christianity, Other, Unknown	No
politics	Communism, Conservatism, Conservative, Environmentalism, Liberal, Liberailsm, Moderate, Socialism, Unknown	No
gender	F, M, Unknown	Yes
age	18-24, 25-34, 45-54, 55-64, Unknown	Yes
income	<100K, >100K, Unknown	Yes
area	Urban, Rural, Unknown	Yes

In order to measure how effective ads are, this thesis will use the dataset to test how effective ads remain when excluding protected variables to target customers. Therefore, the classifying method of logistic regression will be used, which focuses on the click-through rate (CTR) of ads (Kumar, Naik,

Naik, Shiralli, Sunil & Husain, 2015 and Shi & Conrad 2009). This means that this method will analyze the impact of predictor variables, which are protected variables in the given dataset, on whether a customer clicks on the ad or not. Furthermore, this method will analyze the impact of protected attributes on the CTR and give an overview of what customer attributes are most important to consider target advertising as effective (Kumar et al., 2015). In the further context of this research, the method of correlation analysis will be used to determine which protected attributes included in the dataset are correlated and have an impact on the click-through rate (Shi & Conrad, 2009). This helps to investigate which attributes are considered most useful when doing targeted advertising and gives an insight into whether excluding protected attributes or at least some, will decrease the effectiveness of ads. Furthermore, by using the approach of Pearson's correlation coefficient (Shi & Conrad, 2009), it can be determined which attributes have the most or the least impact on ad effectiveness.

When performing logistic regression and correlation analysis, the used dataset will be divided into two data models, from which one contains full information about a customer and the other one can be considered privacy-friendly (Xu & Goodacre, 2018). By doing so, both methods will create models that investigate the relationship between protected variables and the click-through rate. However, the validation set will be used to evaluate which model will be the best and will then be used on the test data to perform the model and investigate how effective advertising is without protected attributes (Xu & Goodacre, 2018 and Kumar et al. 2015).

4. ANALYSIS

To analyze the dataset, it will be imported into RStudio where the data will be analyzed by using different concepts. First of all, the data will be analyzed by using descriptive statistics. This will help to summarize and understand the data and its characteristics such as the different variables (see Appendix 11.1.1). In this context, it is also possible to get a central tendency of how the data looks like and in this case to see which variables might be more crucial for further analysis in this thesis regarding the effectiveness of the target advertising. The next step will be to run a correlation analysis which will show the relationship between variables used for targeting (e.g. age, gender, income) and the key performance indicators like the conversion (Shi & Conrad, 2009. Further, the logistic regression analysis will be investigated, which means that by using this method, the attributes that have the highest impact on the click-through-rate can be defined, and then it can be tested how effective advertising is when removing those attributes that are considered as protected attributes (Hamman, Chen & Dutta, 2023).

In the first step, the variables within the dataset will be transformed into numerical data so that all, since some of the variables were categorical and could not be used for the intended research methods. Therefore, new variables were formed, for instance, the variable "age" was transformed into four variables called "age_25-34", "age_45-54", "age_55-64" and "age_Unknown", and the outcomes are binary which means that if the variable is met it is 1 and if it is not met it is 0 (see Appendix 11.1.2). The correlation analysis will be done by creating a correlation matrix in which the relationship of the variables will be analyzed, and it can be evaluated whether some variables have a higher impact on the conversion (if someone clicked on the ad or not) than others, so that variables that are crucial for target advertising can be identified. The next step will be to perform a logistic regression for two different datasets. Therefore, the existing dataset will be used to perform a logistic model that includes all variables and after that, a data model that can be considered privacy-friendly will be created that will exclude sensitive information like age, gender, income, and area (where the person's lives). By doing both models, it can be analyzed how much impact the personal information about a person has on the effectiveness of the ad and it will be possible to see how much they impact target advertising (Xu & Goodacre, 2018). When both models are performed, the results will be analyzed, and it will be concluded how effective advertising is with and without protected attributes.

5. RESULTS

The dataset was imported into RStudio and at first, an overall overview of the variables included in the dataset was made. From the dataset, it can be stated that eight independent variables were age, gender, income, politics, religion, parents, homeowner, area, and college_educated and the dependent variable is true conversion. This means that within the analysis, it was tested how much effect the independent variables have on the dependent variable (true conversion represents whether it was clicked on the ad or not). Before any method was tested, the categorical variables needed to be transformed into numerical variables, which was done by creating dummy variables, so that at the end there were 23 independent variables (e.g. the variable "age" was turned into variables "age_25-34", "age_45-54" and so on). Furthermore, it was important to clean the data before analyzing it. Since this thesis aims to find out how effective target advertising can be by eliminating personal information that are considered sensitive, the rows with unknown information about age, gender, area, and income were deleted from the dataset. After this was done, two formulas were created that contained a full set of information about a person, and the other one contained privacy-friendly information which means that the variables age, gender, area, and income were excluded. The further testing methods used both data frames so that the probabilities of whether a person clicked on the ad or not were comparable.

5.1 Descriptive Statistics

In Table 2 the outcome of the descriptive statistics is shown, for the independent and dependent variables. Here, the focus lies on investigating the dependent variable "true_conversion", which indicates whether a customer clicked on the ad or not. Some of the independent variables are considered "protected variables", which means that they contain information about a customer that are not considered privacy-friendly (Mazumder & Singh, 2022). From the table, it can be seen that in terms of the mean, there are significant differences between the variables. This means that variables such as "college_educated", "homeowner", "parents", "politics_Unknown", "age_55-64" and "area_Urban" the means are high, indicating that most of the customers in the dataset belong to that group. On the other hand, the dependent variable "true_conversion" has a mean close to 0.00, which states that there are overall only a few conversions in the dataset. Furthermore, the independent variables that show the political orientation are also close to 0.00, which means that these orientations are not represented in the dataset. Moreover, the variable "income_>100K" indicates that there is nearly an even split of people who earn more or less than 100,000. Also, it can be seen that there are more female customers in the dataset than male ("gender_M" has a mean of 0.34), and for most of the customers from whom the religion is unknown (mean of 0.7). When looking at the standard deviation, it can be said that the variables "college_educated" (0.23) and "parents" (0.23) have a low standard deviation, while variables "politics_Conservative" (0.02), "politics_Environmentalism" (0.07) and "politics_Moderate" (0.03) have a high standard deviation and the rest of the variables can be considered as moderate. When analyzing the kurtosis, it can be stated that there are several variables (which are "poltics_Conservative", "politics_Environmentalism", "politics_Moderate", and "true_conversion") with a high kurtosis, indicating that there are more data points around the mean and that it has extreme outliers (Chissom, 1970). On the other hand, the low kurtosis variables (with values of 0.00 or NaN) indicate that data points are more spread around the mean and do not have that many heavy outliers. The rest can be considered moderate kurtosis, which means that the distribution comes close to a normal distribution (Chissom, 1970).

Table 2: Descriptive Statistics

	Mean	SD	Kurtosis
college_educated	0.94	0.23	12.65
parents	0.95	0.22	14.53
homeowner	0.86	0.35	2.37
religion_Other	0.15	0.35	2.03
religion_Unknown	0.70	0.46	-1.20
politics_Conservatism	0.00	0.00	NaN
politics_Conservatisve	0.00	0.02	3266.00
politics_Environmentalism	0.00	0.07	199.44
politics_Liberal	0.00	0.00	NaN
politics_Liberalism	0.00	0.00	NaN
politics_Moderate	0.00	0.03	812.75
politics_Socialism	0.00	0.00	NaN
politics_Unknown	0.99	0.08	143.69
gender_M	0.34	0.48	-1.57
gender_Unknown	0.00	0.00	NaN
age_25-34	0.01	0.12	63.16
age_45-54	0.02	0.14	46.13
age_55-64	0.96	0.19	22.76
age_Unknown	0.00	0.00	NaN
income_>100K	0.49	0.50	-2.00
income_Unknown	0.00	0.00	NaN
area_Unknown	0.00	0.00	NaN
area_Urban	0.88	0.33	3.18
true_conversion	0.00	0.05	358.45

5.2 Correlation Analysis

In Table 3 the correlation matrix is shown that investigates the correlation of the independent variables on the dependent variable "true_conversion" (Shi & Conrad, 2009), and shows which variables have a higher effect than others. When looking at the outcomes presented in Table 3, the variables with the greatest relation are "college_educated" (0.0129), "parents" (-0.0142), and "homeowner" (-0.0296), which indicates that there is a positive relationship of "college_educated" with "true_conversion" and a negative relation between "homeowner" as well as "parents" and "true_conversion". The table shows that some variables have no correlation with "true_conversion" (presented as "NA", but were removed from the table), and the rest can be considered weak correlations (either positive or negative).

As an interpretation of these results, it can be stated that variables that have a positive impact on "true_conversion" imply that if the predicted probabilities probability of the independent variable increases, it is more likely that "true_conversion" is occurring (e.g. when the number of "college_educated" customer increases, the likelihood of them clicking the ad increases) (Gogtay & Thatte, 2017). On the other hand, variables with a negative correlation indicate that if their appearance increases, it will lower the probability of having "true_conversion" (e.g. if the number of "homeowner" increases, there will be fewer people clicking the ad) (Gogtay & Thatte, 2017). Furthermore, variables that do not correlate do conclusively not influence whether the ad is clicked on or not. The variables that only have a weak correlation (positive or negative) with the dependent variable imply that they have an influence on whether "true_conversion" occurs or not, but do not explain variations in the dependent variable very well (Gogtay & Thatte, 2017, and Shi & Conrad, 2009).

Table 3: Correlation Matrix on "true_conversion"

	true_conversion
college_educated	0.01288945
parents	-0.01422583
homeowner	-0.0295672
religion_Other	-0.005153189
religion_Unknown	-0.01710821
politics_Conservatisve	-0.0009179933
politics_Environmentalism	-0.003680419
politics_Moderate	-0.001836829
politics_Unknown	0.004319655
gender_M	0.02342043
age_25-34	-0.006406221
age_45-54	-0.007415686
age_55-64	-0.02114228
income_>100K	0.01877641
area_Urban	0.01978815

5.3 Logistic Models

In the analysis part of the thesis, it was described that the method of logistic regression will be used to investigate how effective target advertising can be when leaving out protected attributes such as age or gender (Kumar, Naik, Naik, Shiralli, Sunil & Husain, 2015). As previously described, a dataset that excluded sensitive information (in this case variables about age, gender, area, and income) were created, so that the outcomes of the privacy-friendly data could be compared to the original dataset (including protected attributes). From the outcome of the logistic model that contained full information, it can be stated that the variables that impacted the likelihood for the occurrence of "true_conversion" the most were "college_educated", "income_>100K" and "area_Urban" (positive coefficients). On the other hand, the variables "homeowner" and "age_55-64" have a negative coefficient which indicates that the likelihood of customers clicking on the ad that belongs to these groups is very low (see table in Appendix 11.2). Furthermore, the logistic model shows that based on the residual deviance of 106.17, the model fits relatively well (Peng, Lee & Ingersoll, 2002).

The second dataset that was analyzed excluded variables that could raise privacy issues and is therefore considered privacy-friendly (Wang, Yang, Chen & Zhang, 2015 and Hamman, Chen & Dutta, 2023). From this logistic regression model, it is indicated that also the variable "college_educated" has a similar effect on "true_conversion", however since variables that are considered as sensitive information are excluded, the only variable that shows a similar trend as in the full dataset if "homeowner" (see table in Appendix 11.3). The residual deviance is slightly higher than from the full dataset model (120.41), which indicates that the model is slightly worse fitting to the dataset (Peng, Lee & Ingersoll, 2002).

5.3.1 Confusion Matrix

Based on the predictions about the likelihood of customers clicking the ad, the following matrix shows the effectiveness of the targeted advertising for the provided dataset. The confusion matrix (see Table 5) generated by the full dataset can be considered as similar to the one generated with the privacy-friendly dataset (see Table 4), however, it can be stated that in the privacy model, there are no predicted positives that are also a true yes. Since there are no true positives and false positives, the outcomes can be used to compute effectiveness measures, which are accuracy (0.9973), recall (0), and specificity (1), however, the precision and F1 Score were not possible to compute (see calculations in Appendix 11.4.1). This indicates that the probabilities to predict the correct clicking behavior of customers are low and close to zero, which means that the privacy-friendly model does not give reliable information whether a customer belongs to the intended target group or not. Furthermore, it can be seen that the model does not have any true conversions, which means that within the privacy model, no customer clicked on the ad.

	No
True No	3264
True Yes	9

Table 4: Confusion Matrix Privacy Model

The matrix of the full model shows that overall, 3,247 are negatives, from which 17 are false positives (FP), which means that the model predicted "Yes" but it was a "No" (Maria Navin & Pankaja, 2016). Furthermore, there were a total of nine positives, however, eight of them were false negatives (FN) which means that they were actually "Yes", but the model predicted "No" (Maria Navin & Pankaja, 2016).

	Predicted No	Predicted Yes
True No	3247	17
True Yes	8	1

Table 5: Confusion Matrix Full Model

The outcomes of the confusion matrix are then used to measure the effectiveness of target advertising by computing the accuracy (which is 0.9924), precision (positive predictive value, which is 0.0556), recall (true positive rate, which is 0.1111), specificity (true negative rate, which is 0.9952) and the F1 score (which is 0.0741) (calculations can be found in the Appendix 11.4.2).

$$\text{Accuracy} = \frac{TrueNegatives + TruePositives}{TrueNegatives + FalsePositives + FalseNagatives + TruePositives}$$

$$\text{Precision} = \frac{TruePositives}{TruePositives + FalsePositives}$$

$$\text{Recall} = \frac{TruePositives}{TruePositives + FalseNegatives}$$

$$\text{Specificity} = \frac{TrueNegatives}{TrueNegatives + FalsePositives}$$

$$\text{F1 Score} = 2x \frac{(PrecisionxRecall)}{(Precision + Recall)}$$

5.3.2 Histograms Predicted Probabilities

To analyze the variable "predicted_probability" for the models that include sensitive information and also for the model that is considered privacy-friendly, histograms are plotted. Those histograms show the ability of the two models to correctly predict whether an ad was clicked or not. The histograms were plotted with a binwidth of 0.001 so that it shows the data distribution in more detail. From the histograms, it becomes clear that the prediction variation in the privacy-friendly model is lower, which means that the probabilities lie in the range between -0.005 and 0.015 (see Figure 3).

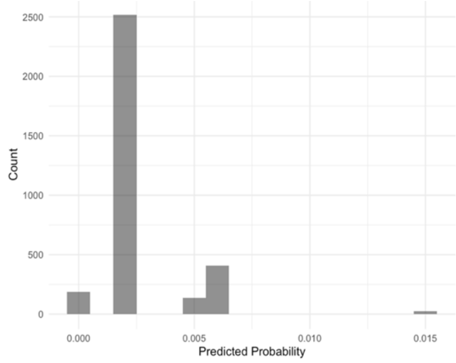

Figure 3: Histogram Privacy Model

The other model including sensitive information shows a wider range with slightly higher probabilities, those lie between 0.00 and 0.383 (see Figure 4). Furthermore, it can be investigated that the histogram of the privacy model does not show a very detailed distribution of the probabilities, even though the binwidth is low (Carlotto, 1987). These results of probability values are very close to zero and do not have a great variation.

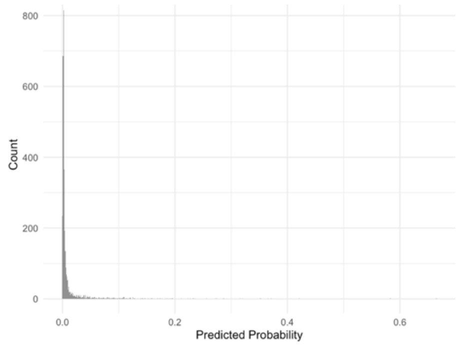

Figure 4: Histogram Full Model

5.4 Summary of the Results

In the first step of the analysis, the descriptive statistics results show that variables with high means such as "college_educated" or "homeowner" are relatively often represented in the dataset, which can indicate that they have a certain relevance when predicting conversions. These variables had a low standard deviation, while variables that had a low mean had a high standard deviation which can be already considered as not that relevant for targeting. The correlation analysis indicated that "college_educated" had the strongest positive relation, meaning that when a customer went to college, he/she was more likely to click on the ad. On the other hand, variables such as "parents" (having parents) or "homeowner" (having an own home) had a negative relation, meaning that if the customer has these attributes, it is less likely that the ad is clicked. Other variables had only weak correlations on the conversions so there is only a low influence on the clicking behavior of the customer. When looking at the logistic model, the results present that the variables "college_educated", "income_>100K" and "area_Urban" had the highest impact on the clicking of customers, however, the results of the privacy-friendly model excluded two of the most influential variables, which makes the model less accurate in predicting the conversion. These results are supported by the outcomes of the confusion matrix, which means that the privacy model is less precise (more false negatives and low true positives). Furthermore, the histograms of the predicted probabilities indicate the same assumption, since the privacy-friendly model has a lower probability range (around zero, between -0.005 and 0.015) than the full model which suggests predictions about conversions are less accurate and marketers can target less effective.

6. DISCUSSION

The goal of this thesis was to investigate how effective target advertising can be when excluding variables that can be defined as protected attributes. In order to analyze the effectiveness of sensitive information in advertising, a logistic regression analysis was performed to evaluate predictions and conversions of a privacy-friendly model (Hamman, Chen & Dutta, 2023).

When performing the logistic regression analysis, the findings support the assumptions that protected attributes do affect target advertising (Speicher et al., 2018). The key results from the two models indicated that variables concerning college education, income, or the living area of a customer do affect the clicking on the ad positively, however, other variables affect it negatively ("homeowner" and "age_55-64"). Furthermore, the deviance of the full model suggested that the model has a relatively good fit, but the privacy-friendly model was slightly worse. These results show that target advertising can be considered more effective when including protected attributes such as the demographic information of a customer (Boerman, Kruikmeier & Borgesius, 2017). In this context, the full model indicates that customers who are college educated ("college_educated"), have a higher income than 100,000 per year ("income_>100K"), and live in an urban area ("area_Urban") are more likely to click on an ad, so they can be considered as the target group for this ad (Estévez & Fabrizio, 2014). However, in the privacy model that excluded sensitive information like age, gender, demographic information, and income, the model's predictions decreased, which means that target advertising is still possible but can be considered less effective and provides less information about customers that can support targeting. Moreover, the lower deviance in this model indicates a poorer fit of the overall privacy model, which means that the model predicts the conversions less accurately so that marketers are less successful when targeting potential customers (Manning, 2007).

Another analysis that showed significance in evaluate the effect of protected attributes on target advertising, was the histogram analysis of the probabilities that predict the conversions in the two models. The findings present a higher range of probabilities in the full model (between 0.00 and 0.383) than in the privacy model (between -0.005 and 0.015). Based on these ranges it can be assumed that the privacy model loses its prediction capability when protected attributes get excluded (Carlotto, 1987). Since the privacy model does not show a wide range of probabilities, it suggests that the model does not identify differences in customer behaviors well enough to give marketers enough information to perform effective target advertising (Richardson, Dominowska & Ragno, 2007). This means that the results of the logistic regression analysis are supported by these findings since both suggest that target advertising is more precise when using protected attributes, however, it must be considered that privacy concerns arise when using them (Juels, 2001). These privacy concerns need to be considered when deciding whether to use sensitive information about customers since even though the targeted advertising is less precise and accurate, marketers can still use enough important information in to define their target groups (Cabañas, Cuevas & Cuevas, 2018).

The research has shown that based on the given dataset, differences in the prediction accuracy for target advertising can be investigated. Furthermore, it implies that marketers need to consider multiple factors when deciding whether they use privacy-friendly data or not since even though protected attributes give more accuracy, there are privacy concerns linked with the use of sensitive information in advertising (Juels, 2001 and Cabañas, Cuevas & Cuevas, 2018).

7. CONCLUSION

The research question of this thesis was "*What is the impact of the mitigation strategy of protected attributes on the effectiveness of targeted online advertising?*". After performing various analysis methods on the IBM dataset, the results support the assumption that the use of protected attributes raises the effectiveness of target advertising. This means that if the intention is to have the highest precision in targeting customers, marketers need to use sensitive information, which indicates that protected attributes have an impact on whether the customer clicks on the ad or not. However, marketers need to consider privacy regulations, which means that protected attributes need to be excluded so that those regulations are not violated and comply with ethical marketing practices. To conclude whether protected attributes impact the effectiveness of target advertising, it can be stated that they are useful to precisely target customers, but marketers need to find a balance between using the personal information of users and complying with privacy regulations. This further indicates that excluding protected variables does not make it impossible to target customers, but it gets less precise.

To conclude this thesis, the theoretical and practical implications of the conducted research will be discussed in the next section 7.1. Furthermore, limitations and suggestions for future research will be discussed in the section afterward (7.2).

7.1 Implications

In terms of theoretical implications from the literature, the results of the thesis support existing theories presented that sensitive information impact the effectiveness of target advertising and enhance the effectiveness of ad placements (Etévez & Fabrizio, 2014; Sweeney, 2013; Hamman, Chen & Dutta, 2023). However, the conducted research has shown that targeting is still possible to a certain extent, even though sensitive information gets excluded and respects privacy considerations that are already discussed in existing research (Boerman, Kruikmeier & Borgesius, 2017). This thesis contributes an extended framework to marketers for using protected attributes as a mitigation strategy, when targeting methods are developed (Mazumder & Singh, 2022). Since this thesis shows that target advertising is still possible and customers can be targeted, even though protected attributes are excluded, this research gives empirical support when theories (e.g. Social Identity Theory) aim to identify certain groups of customers (Dimofte, Goldstein & Brumbaugh, 2014; Rawal & Torres, 2017). However, this research has academic relevance in the field of bias in online advertising, since it has not yet been investigated, how the accuracy and precision of target advertising change, when excluding sensitive information. By using protected attributes as a mitigation strategy, this study gives insights into how privacy regulations can be respected and at the same time marketing goals (like reaching customers that belong to the target group) can still be achieved, which makes it not necessary to violate customers' privacy to achieve those goals.

When considering the practical implications that this thesis has, it is important to mention that by using privacy-friendly information about customers, marketers can enhance the customer's perceptions of the company, but it is still possible that marketing goals can be achieved (Sweeney, 2013). Furthermore, this study provides advertisers with an example of how the mitigation strategy of protected attributes can be implemented to reduce bias in online advertising (Sweeney, 2013, and Hamman, Chen & Dutta. 2023). As a result, companies will be able to not raise privacy concerns that are linked with their advertising methods (like Google Ads presented by Sweeney, 2013), but do not waste resources by placing ads to customers that do not belong to the target group.

7.2 Limitations and Future Research

The study performed in this thesis has some limitations that are important to consider. These limitations mostly concern the dataset that was used to research the effectiveness of target advertising without sensitive information. First of all, the dataset was provided by IBM, so it is considered secondary data, which means that no data collection has been properly made. One limitation in this context is therefore that the data is synthetic (simulated), which means that it is not known how well the data represents real customer targeting. Furthermore, it cannot be investigated to what extent the data is possibly biased. This means that when the dataset was created, it is possible that biases were introduced to prove the relevance of sensitive information for targeting. It also does not correctly reflect actual conditions, implying that real-world datasets might have a higher diversity and variety, also when it comes to outliers. Furthermore, this dataset might not make accurate assumptions in terms of the importance of sensitive information, which means that this dataset might aim to show that targeting is possible but less precise when using sensitive information, however, a non-simulated dataset might provide different assumptions. In addition to this, it is not known what industry, product, or ad campaign the dataset represents, which makes it harder to draw an overall conclusion about target ad effectiveness, without focusing on one specific industry (targeting might vary across different industries, which makes possibly a difference in concluding what role protected attributes have in target advertising).

For future research, it is recommended that this study will be done with real-world data. This means that data from a target ad campaign will be used to validate the findings from this research. Furthermore, it can be helpful to analyze different ad campaigns from different industries, so that marketers get useful suggestions in terms of targeting methods for their specific industry. Another recommendation is, that the used data should be validated, meaning the source of that data generation should be transparent so that it is possible to understand where the data comes from (which industry, which company, which product, which ad campaign, etc.). Moreover, it is important to address biases within the datasets, which can be addressed by ensuring diverse data but also by using bias mitigation algorithms that aim to make datasets fairer. The recommendation of diverse data in this study context means that the customers presented in the dataset should not all belong to the same group (for instance same age, gender, or demographic area), but include also people that do not belong to the target audience, which makes the accuracy of predicting potential customers more valid.

8. ACKNOWLEDGEMENT

I would like to thank my supervisor Dr. Agata Leszkiewicz for the guidance during the process of writing my thesis. Especially I would like to express my gratitude for her critical comments and helpful suggestions when performing the analysis for this thesis. Finally, I would like to thank my family and close friends who supported me during the past couple of weeks, and especially my parents who closely supported me during my whole studies.

9. AI USE DISCLAIMER

The ideas, concepts, and arguments of this thesis are not generated by AI models and are therefore human-generated. However, AI tools were used to support finding scientific papers that are related to the topic or are related to papers that are considered useful (researchrabbit.ai). Furthermore, AI was used to improve academic phrasing and check for grammar mistakes (ChatGPT and Grammarly) and to help quickly summarize relevant papers.

10. REFERENCES

Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R. J., Michael, K. (2022, February 08). *Algorithmic bias in machine learning-based marketing models.* Journal of Business Research 144 (2022) 201-216. https://doi.org/10.1016/j.jbusres.2022.01.083

Asadi, L. K., Shah, A. A. (2023, March 09). *Gender bias in antidepressant direct-to-consumer pharmaceutical advertising.* Comprehensive Psychiatry Vol. 123 (2023) 152384. https://doi.org/10.1016/j.comppsych.2023.152384

Assija, V., Baliyan, A., & Jain, V. (2018). Effective and Efficient Digital Advertisement Algorithms (pp. 83–91). https://doi.org/10.1007/978-981-10-6602-3_9

Bagchi, R., Ham, S. H., He, C. (2020, March 17). *Strategic Implications of Confirmation Bias-Inducing Advertising.* Production and Operations Management (2020). https://doi.org/10.1111/poms.13176

Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2017). *Online Behavioral Advertising: A Literature Review and Research Agenda.* Journal of Advertising, 46(3), 363–376. https://doi.org/10.1080/00913367.2017.1339368

Cabañas, J. G., Cuevas, Á., & Cuevas, R. (2018). *Facebook use of sensitive data for advertising in Europe*. arXiv preprint arXiv:1802.05030. https://doi.org/10.48550/arXiv.1802.05030

Carlotto, M. J. (1987). *Histogram analysis using a scale-space approach.* IEEE Transactions on Pattern Analysis and Machine Intelligence, (1), 121-129.

Chissom, B. S. (1970). *Interpretation of the kurtosis statistic.* The American Statistician, 24(4), 19-22.

Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin. EUR. (n.d.). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32000L0043

Crane, A. & Matten, D. (2016). *Business Ethics* (4th ed.). Oxford University Press.

Dehling, T., Zhang, Y., & Sunyaev, A. (2019). *Consumer Perceptions of Online Behavioral Advertising.* 2019 IEEE 21st Conference on Business Informatics (CBI), 345–354. https://doi.org/10.1109/CBI.2019.00046

Dimofte, C. v., Goodstein, R. C., & Brumbaugh, A. M. (2015). A social identity perspective on aspirational advertising: Implicit threats to collective self-esteem and strategies to overcome them. Journal of Consumer Psychology, 25(3), 416–430. https://doi.org/10.1016/j.jcps.2014.12.001

Estévez, M., & Fabrizio, D. (2014). *Advertising Effectiveness: An Approach Based on What Consumers Perceive and*

What Advertisers Need. Open Journal of Business and Management, 02(03). https://doi.org/10.4236/ojbm.2014.23022

Ferrara, E. (2023). *Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies.* Sci2024, 6, 3. https://doi.org/10.3390/sci6010003

Gogtay, N. J. & Thatte, U. M. (2017). *Principles of correlation analysis.* Journal of the Association of Physicians of India, 65(3), 78-81.

Hamman, F., Chen., J. & Dutta, S. (2023, June). *Can querying for bias leak protected attributes? Achieving privacy with smooth sensitivity.* Proceedings of the 2023 ACM Conference on Fairness, Accountability and Transparency. 2023

Hanlon, A. (2019). *Digital Marketing: Strategic Planning & Integration.* SAGE Publications Ltd.

Juels, A. (2001). *Targeted Advertising ... and Privacy Too* (pp. 408–424). https://doi.org/10.1007/3-540-45353-9_30

Kim, T., Barasz, K., & John, L. K. (2019). *Why Am I Seeing This Ad? The Effect of Ad Transparency on Ad Effectiveness.* Journal of Consumer Research, 45(5), 906–932. https://doi.org/10.1093/jcr/ucy039

Kumar, R., Naik, S. M., Naik, V. D., Shiralli, S., Sunil, V.G & Husain, M. (2015). *Predicting clicks: CTR estimation of advertisements using Logistic Regression classifier.* 2015 IEEE International Advance Computing Conference (IACC), Banglore, India, 2015, pp. 1134-1138, doi: 10.1109/IADCC.2015.7154880.

Laboratory Experiments in the Social Sciences. (2014). Elsevier. https://doi.org/10.1016/C2011-0-07562-2

Li, Z., Kong, D., Niu, Y., Peng, H., Li, X., & Li, W. (2023). *An overview of AI and blockchain integration for privacy-preserving*. arXiv preprint arXiv:2305.03928. https://doi.org/10.48550/arXiv.2305.03928

Manning, C. (2007). *Logistic regression (with R).* Changes, 4, 1-15.

Maria Navin, J. R. & Pankaja, R. (2016, December). *Performance Analysis of Text Classification Algorithms using Confusion Matrix.* International Journal of Engineering and Technical Research (IJETR). ISSN: 2321-0869 (O) 2454-4698 (P), Volume 6, Issue 4.

Mazumder, P., & Singh, P. (2022). *Protected attribute guided representation learning for bias mitigation in limited data.* Knowledge-Based Systems, 244, 108449. https://doi.org/10.1016/j.knosys.2022.108449

Meda, K. (2023, October 27). *Biased advertising: Identifying & addressing the problem*. KORTX. https://kortx.io/news/biased-advertising/#:~:text=Bias%20in%20advertising%20can%20influence,concern%20for%20ethical%20advertising%20practices.

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). *An introduction to logistic regression analysis and reporting.* The journal of educational research, 96(1), 3-14.

Pilgrim, C., Sanborn, A., Malthouse, E., Hills, T. T. (2023, January 19). *Confirmation bias emerges from an approximation to Bayesian reasoning.* Cognition 245 (2024) 105693. https://doi.org/10.1016/j.cognition.2023.105693

Rawal, M., & Saavedra Torres, J. L. (2017). *Empathy for Emotional Advertisements on Social Networking Sites: The Role of Social Identity.* Marketing Management Journal, 27(2).

Richardson, M., Dominowska, E., & Ragno, R. (2007, May). *Predicting clicks: estimating the click-through rate for new ads*. In Proceedings of the 16th International conference on World Wide Web (pp. 521-530).

Saeedi, S., Fong, A. C., Mohanty, S. P., Gupta, A. K., & Carr, S. (2021). *Consumer artificial intelligence mishaps and mitigation strategies*. *IEEE Consumer Electronics Magazine, 11(3)*, 13-24.

Shah, N., Engineer, S., Bhagat, N., Chauhan, H., Shah, M. (2020, November 25). *Research Trends on the Usage of Machine Learning and Artificial Intelligence in Advertising.* Augmented Human Research (2020) 5:19. https://doi.org/10.1007/s41133-020-00038-8

Shi, R., & Conrad, S. A. (2009). *Correlation and regression analysis*. Ann Allergy Asthma Immunol, 103(4), S34-S41.

Sierra, J. J., Hyman, M. R., & Torres, I. M. (2009). *Using a Model's Apparent Ethnicity to Influence Viewer Responses to Print Ads: A Social Identity Theory Perspective.* Journal of Current Issues & Research in Advertising, 31(2), 41–66. https://doi.org/10.1080/10641734.2009.10505265

Speicher, T., Ali, M., Venkatadri, G., Ribeiro, F. N., Arvanitakis, G., Benevenuto, F., ... & Mislove, A. (2018, January). *Potential for discrimination in online targeted advertising.* In Conference on fairness, accountability and transparency (pp. 5-19). PMLR.

Sweeney, L. (2013). *Discrimination in Online Ad Delivery.* Queue, 11(3), 10–29. https://doi.org/10.1145/2460276.2460278

Taggart, G., Zenor, J. (2022, July 14). *Evaluation as a moral practice: The case of virtue ethics.* Evaluation and Program Planning, Vol. 94 (2022) 102140. https://doi.org/10.1016/j.evalprogplan.2022.102140

Title VII of the Civil Rights Act of 1964. US EEOC. (n.d.). https://www.eeoc.gov/statutes/title-vii-civil-rights-act-1964

Van Giffen, B., Herhausen, D., Fahse, T. (2022, February 07). *Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods.* Journal of Business Research 144 (2022) 93-106. https://doi.org/10.1016/j.jbusres.2022.01.076

Wang, W., Yang, L., Chen, Y., & Zhang, Q. (2015). *A privacy-aware framework for targeted advertising.* Computer Networks, 79, 17–29. https://doi.org/10.1016/j.comnet.2014.12.017

Wang, Y., Yin, P., Tao, Z., Venkatesan, H., Lai, J., Fang, Y., Xiao, P. (2023, August). *An empirical study of selection bias in Pinterest ads retrieval.* KDD '23, August 6-10, 2023. https://doi.org/10.1145/3580305.3599771

Xu, Y. & Goodacre, R. (2018). *On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning.* Journal of analysis and testing, 2(3), 249-262.

11. APPENDIX

11.1 Variables in Dataset from IBM

11.1.1 Original Variables in the Dataset

	religion	politics	college_educated	parents	homeowner	gender	age	income	area	true_conversion	predicted_conversion	predicted_probability
1	Unknown	Unknown	1	1	1	Unknown	55-64	Unknown	Unknown	0	0	1.350511e-03
2	Other	Unknown	1	1	1	Unknown	55-64	Unknown	Urban	0	0	2.238162e-03
3	Unknown	Unknown	1	1	1	F	55-64	Unknown	Unknown	0	0	2.704057e-03
4	Unknown	Unknown	1	1	1	F	55-64	Unknown	Unknown	0	0	1.967166e-03
5	Unknown	Unknown	1	1	1	F	55-64	Unknown	Urban	0	0	1.681378e-03
6	Unknown	Unknown	1	1	1	F	55-64	Unknown	Unknown	0	0	5.596169e-03
7	Unknown	Unknown	1	1	1	M	55-64	Unknown	Unknown	0	0	6.281301e-04
8	Unknown	Unknown	1	1	0	M	55-64	<100K	Unknown	0	0	1.927780e-03
9	Christianity	Unknown	1	1	1	F	55-64	Unknown	Unknown	0	0	2.649384e-03
10	Unknown	Unknown	1	1	1	F	55-64	Unknown	Unknown	0	0	1.031408e-03
11	Christianity	Unknown	1	1	1	M	55-64	Unknown	Unknown	0	0	3.593597e-03
12	Unknown	Unknown	1	1	1	F	25-34	<100K	Unknown	0	0	5.218610e-03
13	Christianity	Unknown	1	1	1	Unknown	55-64	Unknown	Unknown	0	0	3.154423e-02
14	Unknown	Unknown	1	1	1	Unknown	55-64	Unknown	Unknown	0	0	5.820587e-04
15	Other	Unknown	1	1	1	F	55-64	Unknown	Unknown	0	0	1.840694e-03
16	Unknown	Unknown	1	1	0	F	55-64	<100K	Unknown	0	0	2.248923e-02
17	Unknown	Unknown	1	1	1	M	55-64	Unknown	Unknown	0	0	2.032652e-03
18	Unknown	Unknown	1	1	1	Unknown	55-64	Unknown	Unknown	0	0	1.950345e-03
19	Other	Unknown	1	1	1	Unknown	55-64	Unknown	Unknown	0	0	7.055193e-03
20	Unknown	Unknown	1	1	1	Unknown	55-64	Unknown	Unknown	0	0	3.891051e-03

11.1.2 Variables as Dummies (numerical variables)

```
> variables_names <- names(ad_campaign_dummies_clean)
> print(variables_names)
 [1] "college_educated"              "parents"                       "homeowner"
 [4] "true_conversion"               "predicted_conversion"          "predicted_probability"
 [7] "religion_Other"                "religion_Unknown"              "politics_Conservatism"
[10] "politics_Conservative"         "politics_Environmentalism"     "politics_Liberal"
[13] "politics_Liberalism"           "politics_Moderate"             "politics_Socialism"
[16] "politics_Unknown"              "gender_M"                      "gender_Unknown"
[19] "age_25-34"                     "age_45-54"                     "age_55-64"
[22] "age_Unknown"                   "income_>100K"                  "income_Unknown"
[25] "area_Unknown"                  "area_Urban"                    "predicted_probability_full"
[28] "predicted_probability_priv"    "predicted_conversion_full"     "predicted_conversion_priv"
```

11.2 Logistic Regression Full Model

Deviance Residuals	Min	1Q	Median	3Q	Max
	-0.9827	-0.0824	-0.0583	-0.0408	3.5674

Coefficients: (6 not defined because of singularities)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-49.1617	48355.8748	-0.001	0.9992
college_educated	16.8340	3225.1941	0.005	0.9958
parents	-0.9974	1.0790	-0.924	0.3553
homeowner	-1.3190	0.7278	-1.812	0.0699
religion_Other	-0.9567	1.1749	-0.814	0.4155
religion_Unknown	-1.0436	0.7497	-1.392	0.1639
politics_Conservatism	NA	NA	NA	NA
politics_Conservatisve	-0.7290	68159.6384	0.000	1.000
politics_Environmentalism	-1.3492	49541.2012	0.000	1.000
politics_Liberalism	NA	NA	NA	NA
politics_Moderate	-0.3300	53396.1953	0.000	1.000
politics_Socialism	NA	NA	NA	NA
politics_Unknown	15.4757	48196.1427	0.000	0.9997
gender_M	0.7290	0.6960	1.047	0.2949
gender_Unknown	NA	NA	NA	NA
age_25-34	-20.1651	6492.6818	-0.003	0.9975
age_45-54	-20.2213	5476.5929	-0.004	0.9971
age_55-64	-3.8529	1.2844	-3.000	0.0027
income_>100K	0.7123	0.7338	0.971	0.3317
income_Unknown	NA	NA	NA	NA
area_Unknown	NA	NA	NA	NA
area_Urban	16.9746	2240.6664	0.008	0.9940

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)

Null deviance: 124.11 on 3772 degrees of freedom
Residual deviance: 106.17 on 3257 degrees of freedom
AIC: 138.17

Number of Fisher Scoring iterations: 21

11.3 Logistic Regression Private Model

Deviance Residuals:

Min	1Q	Median	3Q	Max
-0.1725	-0.0652	-0.0652	-0.0652	3.5089

Coefficients:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-18.8568	1282.9738	-0.015	0.988
college_educated	14.6566	1282.9733	0.011	0.991
parents	-0.8310	1.0648	-0.780	0.435
homeowner	-1.1227	0.7092	-1.583	0.113

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 124.11 on 3772 degrees of freedom
Residual deviance: 120.41 on 3269 degrees of freedom
AIC: 128.41

Number of Fisher Scoring iterations: 19

11.4 Confusion Matrix Calculations

11.4.1 Privacy Model

Accuracy = (TN + TP) / (TN + FP + FN +TP)
Accuracy = (3264 + 0) / (0 + 3264 + 0 + 9) = 0.9973 (99.73%)

Precision = TP / (TP + FP)
Precision = 0 / (0 + 0) (undefined, since TP + FP = 0)

Recall = TP / (TP + FN)
Recall = 0 / (0 + 9) = 0

Specificity = TN / (TN + FP)
Specificity = 3264 / (3264 + 0) = 1

F1 Score = 2 * (Precision * Recall) / (Precision + Recall)
F1 Score = 2 * (0 * 0) / (0 + 0) (undefined, since Precision and Recall are 0)

11.4.2 Full Model

Accuracy = (TN + TP) / (TN + FP + FN +TP)
Accuracy = (3247 + 1) / (3247 + 17 + 8 + 1) = 0.9924 (99.24%)

Precision = TP / (TP + FP)
Precision = 1 / (1 + 17) = 0.0556 (5.56%)

Recall = TP / (TP + FN)
Recall = 1 / (1 + 8) = 0.1111 (11.11%)

Specificity = TN / (TN + FP)
Specificity = 3247 / (3247 + 17) = 0.9952 (99.52%)

F1 Score = 2 * (Precision * Recall) / (Precision + Recall)
F1 Score = 2 *(0.0556 * 0.1111) / (0.0556 + 0.1111) = 0.0741 (7.41%)