

Data Mining: Perception of Customers towards Privacy Invasion

D Kumar^{1*}, Ankita²

¹Amity Institute of Information Technology, Amity University, India

²Mewar University, Rajasthan, India

*Corresponding Author: deepakgupta_du@rediffmail.com

Available online at: www.ijcseonline.org

Abstract - Nowadays ample amount of data including personal data is being collected by government and private sector. This data in turn helps in running countries and their economy effectively and efficiently. In order to achieve this, private as well as government institutions collect clients or consumers data. To be more specific, in retail sector, we can say, the more data they gather, the better will be the results. But, the more personal data which is collected and analyzed, the more individuals are at risk of crimes. One must be aware of privacy invasions and its rules. As coin has two sides, same goes with data mining; it too has positive and negative aspects. Nowadays, technology is advancing and so more of the awareness is required from the customers' side. In this paper I look at the perception of customers towards privacy invasion with reference to data mining and how Customer tracking technology work in offline Retail Stores.

Keywords—Data mining, retail, customer privacy, Personal Data, micro data, information

I. INTRODUCTION

The arrival of information technology in different fields of human life has prompt the expansive volumes of data stockpiling in different configurations like records, reports, pictures, sound recordings, logical information, and numerous new data configurations. The data gathered from various applications need appropriate way of deriving knowledge from big warehouse for improved predictions. From the previous two decades information mining and knowledge discovery applications have a rich concentration because of its noteworthiness in basic leadership and it has turned into a necessary element in different institutions. The field of information mining have been flourished and postured into new regions of human existence with different mixes and headways like in the area of Statistics, Databases, Machine Learning, Pattern Reorganization, Artificial Intelligence and Computation capabilities etc. Some of the application areas of data mining are Life Sciences (LS), Customer Relationship Management (CRM), Web Applications, Manufacturing, Competitive Intelligence, Retail/Finance/Banking, computer /Network/Security, Monitoring/Surveillance, Teaching Support, Climate modeling, Astronomy, and Behavioral Ecology etc. Practically every field of human life has progressed toward becoming information concentrated, which made the information mining as a fundamental part. Information mining includes the utilization of a suite of strategies that plan to initiate from information, models that meet specific targets. DM calculations are based on a scope of strategies, including data hypothesis, insights, straight and non-direct models, AI, meta-heuristics. Inside the setting of information investigation techniques, information mining can be thought to be an exploratory learning disclosure way to deal with is appeared differently in relation to a corroborative approach in which a theory is indicated and the legitimacy of the speculation is tried against the information. DM models can be intended to discover new helpful examples inside the information that can be abused. For instance, an insurance agency will as of now know that young fellows in quick autos represent a

high hazard (as do more seasoned ladies). Such an example in the information will be of no enthusiasm to them. In any case, they might be intrigued if they somehow happened to locate that young fellows in quick, great autos are substantially less of a hazard. They could then create valuing and advertising techniques to abuse this newfound example. With regards to contemporary associations, the significant spurring factors for the enthusiasm for, and use of, information mining strategies can be sorted as: changed business condition (eg weight on customary promoting methods, shorter time to showcase, shorter item life cycles, and expanded rivalry and business dangers), drivers (eg client, rivalry, and information resource), and empowering agents (eg information surge in associations, new data innovation arrangements in information mining, information warehousing, and equipment). Numerous associations routinely gather a lot of information on their customers, representatives, and providers. Industrial facilities gather information on their apparatus and efficiency. Doctor's facilities gather information on patients, specialists, solutions, and techniques. These information are frequently chaotic, deficient, and here and there mistaken but, in the event that utilized legitimately, it can be an important resource for administration and compelling associations can discover approaches to abuse it. Foreseeing false conduct in a client database, for instance, will by and large require profoundly exact models regardless of whether they must be connected to a little extent of the database. Then again, if the association is arranging a focused on advertising effort and wishes to anticipate which clients may purchase a specific item, the exactness of the model can be much lower.

Data mining is the extraction of interesting patterns or knowledge from large amount of data. In the past years, with the massive growth in Internet, data storage and data processing technologies, privacy preservation has been one of the greater concerns in data mining. [10]

II. DATA MINING IN RETAIL INDUSTRY

In an undeniably aggressive market space, retailers need to know all that they can about their clients: their identity, what they purchase, when they purchase, why they purchase? What's more, on account of the measure of information flying around about client purchasing practices retailers can answer every one of those inquiries and that's only the tip of the iceberg, if they have the innovation expected to gather, arrange, clean, and dissect that data. With information mining as a component of a business knowledge activity, retailers can have genuine responses to genuine inquiries progressively.

Retailers have been gathering gigantic measures of information consistently, much the same as the banking industry, and now have the tool expected to deal with this information and fetch valuable information.

For retailers, data mining is a way to give information on product sales trends, customer buying habits and preferences, supplier lead times and delivery performance, seasonal variations, customer peak traffic periods, and similar predictive data for making proactive decisions. Retailers must think about data mining advantages which they can utilize in the following manner:

Conduct shopping cart analysis.

Shopping basket (or market bin) examination is usually utilized by retailers to better comprehend client buying inclinations. Information followed with shopping basket examination can be utilized to grow more directed activities composed toward strategically pitch items that are much of the time acquired together, up-offer clients at registration, and combine items with higher overall revenues close by items sought after. Organizations can use the bits of knowledge picked up from a shopping basket investigation to upgrade collection arranging and approve advancements, as well. For example, when an online retailer offers a free delivering bargain does that create more movement than a 10% off coupon? After some time, retailers can utilize the information mined from shopping basket examination to manage their promoting and deals exercises to hit the correct clients with the correct offers at the correct circumstances.

Must know your best clients

Nothing is more imperative to a retailer than extremely knowing their clients. What's more, with information mining programming they can learn precisely who their best clients are, what push them to shop, how regularly they purchase, the amount they spend per request and the sky is the limit from there. With this sort of business knowledge, retailers can without much of a stretch partition clients into high-spend, medium-spend and low-spend client fragments. They can utilize the information to better comprehend the spending designs, correspondence inclinations, and promoting inclinations of their clients, and would then be able to amass them by these qualities and spotlight on one fragment at any given moment.

Understand marketing drive efficiency.

Retailers should likewise guarantee their showcasing efforts are contacting the correct groups of onlookers at the ideal time with the correct offers that provoke activity. With information mining abilities, they can track all their different advertising efforts or advancements to see which ones have the greatest return.

Utilizing information mining and business insight programming, retailers can expand benefit and income by understanding which clients and items truly drive their business.

As retail industry is growing day by day, so are the customers. Retail industry holds huge amount of sales data, customer shopping history, as well as their personal information.

Personal information is something that is very critical for the organization as well as for the customer. Here comes the concept of privacy invasion, which is very likely to be done, if the customer's data is not handled properly.

III. CUSTOMERS ARE KEEN IN THE SAFEGUARD OF THEIR PRIVACY

Consumer privacy, also known as customer privacy, includes the taking care of and insurance of delicate individual data that people give throughout regular exchanges.

Organizations are progressively gathering and utilizing information on their present and potential clients to enhance their CRM, deals, and administration adequacy. In the meantime, they are looked with a developing hesitance of clients to uncover their own data or to permit following of their practices on account of security concerns.

An information supplier possesses information from which profitable data can be fetched. Basically 2 types of data providers exist:

1. information supplier who gives information to information authority, and
2. Information gatherer who gives information to information mineworker.

Data reporting information about an individual are often referred to as "microdata" [1]. If a customer announces her microdata to the data collector, her privacy could be on stake because of the unwanted data infringement. Subsequently, the protection worry of a client is whether he can take control over what sort of and how much data other individuals can get from his information. To examine the measures that the information supplier can receive to secure protection, we think about the accompanying three circumstances

1. In the event that the customer views his information as extremely delicate, that is, the information may uncover some data that he doesn't need any other individual to know, the customer can simply decline to give such information. Powerful access-control measures are wanted by the customer, with the goal that he can keep his delicate information from being stolen by the information gatherer.
2. Understanding that his information are important to the information gatherer (and also the information mineworker), the information supplier might will to hand over a portion of his private information in return for certain advantage, for example, better administrations or money related prizes. The information supplier has to know how to consult with the information

authority, so he will get enough remuneration for any conceivable misfortune in security.

3. In the event that the information supplier can neither avoid the reach to his crucial information nor strike a deal with the information authority, the information supplier can twist his information that will be collected by the information gatherer, with the goal that his actual data can't be effectively revealed.

So, we can say that individual data, when abused or not properly secured, can bring about wholesale fraud, budgetary extortion and different issues that altogether cost individuals, organizations and governments a huge number of dollars every year.

IV. HOW ORGANIZATIONS GATHER DETAILS OF INDIVIDUALS FROM THEIR SHOPPING INTERESTS

Organizations have an assortment of approaches to gather data on their clients. A great part of the data is openly given by shoppers who take an interest in dedication programs, round out studies, answer messages and play sweepstakes. Data ranges from what you purchase and whether you utilize coupons to which magazines you frequently read. Individual data they gather incorporates your name, address, occupation, number of individuals in your family unit and how much cash you make. In the event that you take an interest in investment funds projects or utilize the Internet, your own data is out there for organizations to assemble. Some of the ways through which customer's data is collected:

- **Subscription Loyalty Programs**

A standout amongst the most evident spots where stores gather purchasing data about you is through the loyalty card you agree to accept to get in-store deals. Once your card is swiped, every one of the information about what you purchased, when you made the buy and what number of coupons you utilized is quickly recorded in your distinguishing proof record. Stores arrange the information to make sense of your examples and which items you routinely buy. They can send you coupons for comparative items, arrange records they at that point can exchange and make advertisement crusades focusing on you and your companions by investigating the data you unreservedly give the store by agreeing to accept the program.

- **Consumer Credit Lists**

Buyer credit organizations gather generous measures of data from the numerous credit applications you round out and the buys you make. They know when you take out another home loan, back a bike for your youngster and add another ward to your wage government form. Organizations buy this data and make information records they thus pitch to retailers. Through the credit organization records, organizations know when you got separated, moved or accepted another position, all occasions that could trigger changes in your ways of managing money, making you ready for an alluring promoting effort.

- **Public Record Databases**

Everybody, it appears, is in the matter of offering information. Department of motor vehicles offices gather immense measures of data they frequently pitch to information authorities who aggregate the rundowns

purchased by retailers. Open records are inspected by information authorities who offer arrangements of the as of just wedded, separated and dead. Voting records are available to people in general and scoured over to find out about your leanings for political publicizing purposes and also retailers focusing on particular politically dynamic customers. The main records ensured by law, as per the Federal Trade Commission, are your medicinal services records and the data gave to organizations by kids younger than 13. Something else, it's up to customers to know about the kinds of organizations that gather data and to precisely read security proclamations gave by merchants.

V. CONSUMER PRIVACY LITERATURE REVIEW

The study is done on the basis of various articles printed between 1991 -2006 related to customer's privacy. This study is divided into 2 broad categories 1) conceptualizing customer's privacy, 2) customer's pertinent privacy concerns

Conceptualizing customer's privacy

In early days, privacy was limited to two forms only: 1) control of presence of others in the marketing environment and 2) control of transactional information [5]. These early definitions of privacy have been enlarged to encompass customer's knowledge, or the extent to which customers are aware about and consider a firm's information practices and privacy policies [4, 7]. Clients immovably trust that they have a privilege to protection when it comes to marketing.

VI. CUSTOMER'S RELATED PRIVACY ISSUES

There are large number of factors that have a straight impact on customer's privacy concerns which includes alertness, information management, information awareness, closeness with the entity etc.. Statistic factors, for example, sex, age, and pay additionally influence customer's perspectives of different protection issues[2,9]. A portion of the ways that customers self-deal with their security concerns incorporate going through privacy laws, staying away from accumulation of data, taking part in name expulsion, practicing their lawful rights, and dealing with their online characters [6,11]. Research recommends that if concerns about customer's security are not reassures, they may have unfavorable outcomes on fundamental authority, purchasing, and firm/advertise trust. [3,8].

VII. DEFINING RESEARCH MODEL

This research paper is interested only in consumer perception towards privacy concern i.e. it becomes our dependent variable. Other factors that are related to this research are consumer's life style, attitudes and interests etc. These are independent variables.

Now, the privacy concern comes under in a variable group one and all other factors come under in second variable group.

VIII. DEFINING RESEARCH OBJECT & HYPOTHESIS

To discover the impact of V1 on consumer's shopping interest;

H1: There is a considerable impact of V1 on consumer's shopping interest

To discover the impact of V2 on consumer’s shopping interest;

H2: There is a considerable impact of V2 on consumer’s shopping interest

THE SAMPLE SELECTION PLAN:

This paper mainly targets the attitude of distinct consumers towards their privacy concerns. Consumers from distinct educational, history, environment, culture etc have participated. Sample size for this study is 40. The sampling technique adopted is random sampling.

IX.DATA COLLECTION:

In this research paper, data has been collected from two sources, broadly classified as primary and secondary sources.

Primary Sources:

The primary source for this research is the questionnaire. Data analysis has been done on the gathered data, for finding consumer’s interest towards privacy concern.

Secondary Sources:

The secondary sources for this research are the journals & published books and access information from the internet, whereas some data has been collected from the specific related websites.

Tools for data collection:

Questionnaire method was followed for collecting the data from the respondents. The questionnaire consisted of questions on the variables that measure the attitude of consumers towards privacy while visiting/ shopping in a retail store. Data collected has been shown in a tabular form and on that basis further analysis has been carried out.

X.DATA ANALYSIS

Table 1.1- Showing factors related to the first variable group (V1) i.e privacy concern while visiting/ shopping in a retail store

Table 1: Showing factors related to the first variable group (V1) i.e privacy concern while visiting/ shopping in a retail store						
		Strongly Disagree	Disagree	Neutral	Agree	Strongly agree
1	As long as data doesn’t get abused, privacy is less of an issue these days	3.20%	48.40%	16.10%	29.00%	3.20%
2	Individual data will be sold to outsiders without authorization.	4.20%	35.40%	25.80%	32.30%	2.30%
3	The exchange of individual data is fundamental for the smooth running of present day society	0	22.60%	48.40%	29.00%	0.00%
4	I believe my information will not be misplaced.	0	37.30%	20.10%	36.10%	6.50%
5	I am aware of data protection regulations	0	16.10%	35.50%	45.20%	3.20%
6	Through your data, retailers can keep an eye on your shopping habits.	0	6.50%	12.90%	54.80%	25.80%
7	Email addresses and other individual data will be caught and utilized for advertising or different purposes without consent	0	15.70%	23.70%	57.10%	3.50%
8	I don’t find anything wrong in using wifi while shopping	0	3.20%	35.20%	58.30%	3.30%

Interpretation:

The above table showing factors related to first variable group (V1) i.e. privacy concern. Customers are genuinely concerned with hiding their personal data but they don’t know how to protect their data from big eyes of retailers.

Figure 1. showing the extent of factors related to the first variable group (V1) while visiting / shopping from a retail store.

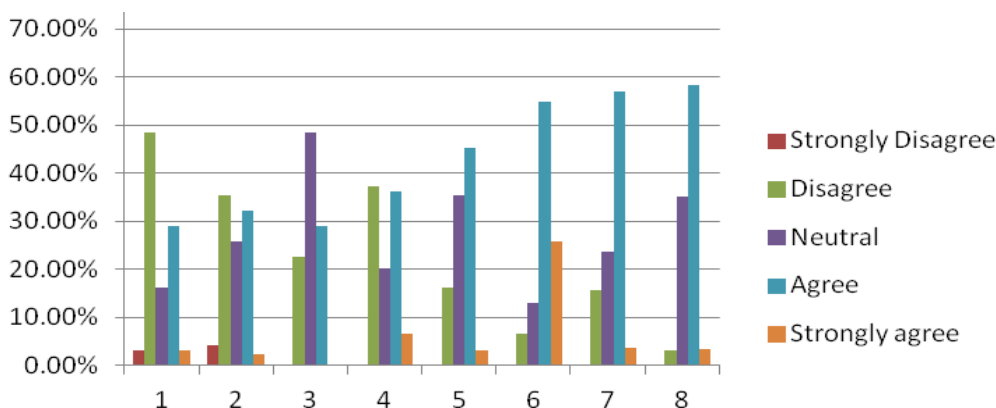


Figure 1. showing the extent of factors related to the first variable group (V1) while doing mobile-commerce

Table 2 showing the extent of factors related to the first variable group (V1) while visiting / shopping from a retail store.

Table 2: Factors related to the second variable group while visiting / shopping from a retail store.				
		Yes	No	Maybe
1.	Do you carry mobile while visiting a retail store?	98%	1%	1%
2.	Do you use wi-fi of the retail store (if available freely)	51.6%	48.4%	
3.	Do you fill the forms given by the store officials to collect your details	48.4%	19.4%	32.3%
4.	Are you aware of the purpose behind filling those forms?	58.1%	41.9%	
5.	Do you feel that your personal info may get into wrong hands	38.7%	3.2%	58.1%

Interpretation:

The above table showing factors related to second variable group (V2) i.e. consumer’s attitude, behaviour. Interests etc

Figure 2 showing the extent of factors related to the second variable group (V2) while visiting / shopping from a retail store.

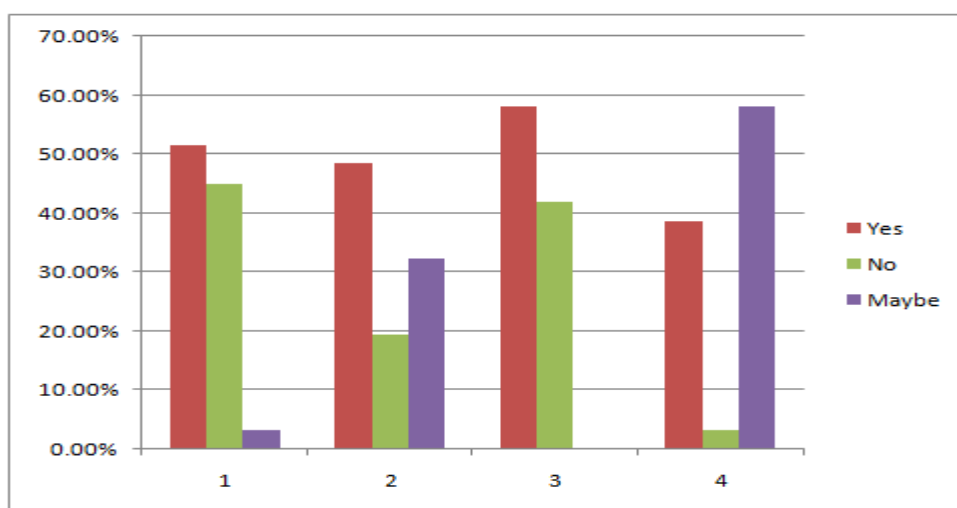


Figure 2 showing the extent of factors related to the second variable group (V2) while visiting/shopping in a retail store

XI.CONCLUSION

Based on the findings of the research the study reveals that both the variable plays a significant role for understanding consumer’s perception towards privacy invasion while visiting/ shopping from retail store. As the retail sector is growing, data mining technology is advancing too. Lacs of customers visit retail stores on daily basis. Each day, their personal information is on stake. Even though some of them are aware of privacy rules and regulations, still they are falling prey to the retailer’s hands. So, privacy concern must be handled carefully by employing privacy preserving techniques on the customer’s data, so that they feel secure and safeand their information does not go in wrong hands.

REFERENCES

[1] V. Ciriani, S. De Capitani di Vimercati, S. Foresti, P. Samarati, "Microdata protection", *Secure Data Management in Decentralized Systems*, pp. 291-321, 2007.
 [2] Culnan, Mary J. and Pamela K. Armstrong. 1999. "Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation." *Organization Science* 10 (1): 104-115.
 [3] Eastlick, Mary Ann, Sherry L. Lotz, and Patricia Warrington. 2006. "Understanding Online B-to-C Relationships: An

Integrated Model of Privacy Concerns, Trust, and Commitment." *Journal of Business Research* 59 (8): 877-886.
 [4] Foxman, Ellen R. and Paula Kilcoyne. (1993). "Information Technology, Marketing Practice, and Consumer Privacy: Ethical Issues." *Journal of Public Policy & Marketing* 12 (1): 106-119.
 [5] Goodwin, Cathy. (1991). "Privacy: Recognition of a Consumer Right." *Journal of Public Policy & Marketing* 10 (1): 149-166.
 [6] Milne, George R. and Andrew J. Rohm. 2000. "Consumer Privacy and Name Removal Across Direct Marketing Channels: Exploring Opt-In and Opt-Out Alternatives." *Journal of Public Policy & Marketing* 19 (2): 238- 249.
 [7] Nowak, Glen J. and Joseph Phelps. (1997). "Direct Marketing and the Use of Individual-Level Consumer Information: Determining How and When 'Privacy' Matters." *Journal of Direct Marketing* 11 (4): 94-108.
 [8] Phelps, Joseph E., Giles D'Souza, and Glen J. Nowak. 2001. "Antecedents and consequences of consumer privacy concerns: An empirical investigation." *Journal of Interactive Marketing* 15 (4): 2-17.
 [9] Sheehan, Kim Bartel and Mariea Grubbs Hoy. (2000). "Dimensions of Privacy Concern Among Online Consumers." *Journal of Public Policy & Marketing* 19 (1): 62-73.
 [10] WANG,P. (2010) Survey on Privacy Preserving Data Mining, *International Journal of Digital Content Technology and its Applications* Volume 4, Number 9, December
 [11] Zwick, Detlev and Nikhilesh Dholakia. 2004. "Whose Identity Is It Anyway? Consumer Representation in the Age of Database Marketing." *Journal of Macromarketing* 24 (1): 31-43.