



Exploring the Role of Data Mining in Customer Relationship Management

RAHUL CHAUHAN, Dept of Comp. Sc. & Info. Tech. , Graphic Era Hill University, Dehradun, Uttarakhand, India 248002,

Abstract

A Customer Relationship Management (CRM) is a strategy or practice that organizations use to manage and analyze the various activities related to their customers' lifecycle. It aims to improve customer loyalty, satisfaction, and profitability. CRM is utilized by various companies to gather and analyze data about their customers. It then uses tools and techniques to develop personalized marketing and sales strategies. A central database is used to store and manage customer information. A CRM system can help businesses identify and target their customers' specific needs and preferences. It can also use advanced statistical techniques to analyze and improve their customer interactions. This type of strategy is very useful in helping companies develop effective marketing campaigns. The paper explores the various advantages of data mining in CRM. This process can help organizations identify patterns in their customer data and improve their customer satisfaction. It can also help them develop effective marketing strategies and improve their sales. This paper explores the various challenges that come with data mining in CRM, such as security, privacy, and quality. It also provides recommendations on how to overcome these issues and implement effective practices related to data mining. The paper covers the various aspects of data mining in CRM and provides practical advice for organizations that want to improve their customer service and understand their customers better.

Keywords: CRM, Datamining, Banking, Customer satisfaction.

DOI Number: 10.48047/nq.2019.17.03.2015

NeuroQuantology 2019; 17(03):152-161

Introduction

Companies are gathering and preserving vast amounts of data due to the rise of digital technology. However, the real challenge is how to use this information to make better decisions and improve the efficiency of their operations. Data mining is a process that involves analyzing and extracting information from the data. Big data is becoming more prevalent in today's digital age. Companies are gathering and storing vast amounts of information about their consumers and services, but they still face the challenge of making sense of it and using it to make informed decisions. This is where the use of data mining comes in. Machine learning and

statistical analysis are used in data mining to identify hidden insights and patterns in large datasets. This process can help businesses make informed decisions and improve their customer relationships. A CRM strategy is a process utilized to analyze and manage the interactions between customers and their organizations.[1]

Customer relationship management (CRM) is a process that aims to improve the retention and satisfaction of customers. It can also help businesses grow their sales and profitability. A vital component of CRM is data mining, as it allows companies to analyze and improve their customer relationships by identifying patterns and insights in their data.[2] This process can



then be used to develop customized marketing campaigns and recommendations. This paper aims to explore the various aspects of data mining in customer relationship management and its role in driving business growth. We will also talk about the most common techniques that are used in this process.

One of the most important applications of data mining in CRM involves predicting the likelihood of a customer leaving a company. This process can then help businesses develop effective retention strategies. One of the most important benefits of data mining is the ability to segment and analyze customers. This process can help businesses create more effective marketing campaigns and improve customer satisfaction. In CRM, data mining plays a vital role, as it allows companies to improve their customer service and understand their customers better. Through this process, they can identify the various characteristics of their consumers and develop effective marketing strategies.[3]

Customer relationship management is becoming more important as the world becomes more connected. It is a process and technology that businesses use to manage their relationships with their customers. The ultimate objective of CRM is to enhance customer loyalty, satisfaction, and sales. One of the most critical components of CRM is Data mining. This process is used to analyze vast amounts of data to identify patterns and correlations. It can help organizations make more informed decisions and improve their customer relationships. In addition to being able to identify trends and patterns, data mining also allows them to gain a deeper insight into their customers.[4], [5]

Data mining can help organizations identify and target their ideal customers. It can also help them improve their customer relationships and develop effective marketing strategies. In CRM research, various techniques are commonly utilized, such as Naive Bayes, Logistic Regression, Decision Trees, and K-Nearest Neighbors. These are all powerful and have varying strengths and weaknesses. The choice of which technique to use is dependent on the

specific problem being analyzed. Over the years, various studies have been conducted on the effectiveness of data mining in CRM. These studies have shown that it can help improve customer loyalty and satisfaction, as well as boost sales and profitability.

There are many advantages to data mining in CRM, but it comes with some limitations that need to be addressed. One of the most common issues that organizations face when it comes to using this process is the quality of the data. In order to effectively utilize this process, they must make sure that the data is up-to-date and accurate. Another challenge that organizations must address is the security and privacy of their customers' information. The purpose of this paper is to explore the various aspects of data mining in CRM. It will also talk about the evaluation metrics that are used in this process to measure its effectiveness.[6]–[8] This paper aims to analyze the Bank Marketing dataset, which is a popular subject in CRM research, and show how data mining techniques can be used in real-world scenarios. The data collected from the Bank Marketing Dataset can be used in various studies to analyze the effectiveness of different marketing techniques. It includes information on the customer's financial history, demographics, and past campaigns. Through the study of the Bank Marketing data, we can gain a deeper understanding of the various factors that influence the behavior of customers. This can help us develop effective marketing campaigns that are designed to meet the needs of each individual.

Data mining plays a vital role in CRM as it allows organizations to gain deeper understanding of their customers' needs and preferences. With the help of this process, they can then develop effective marketing strategies that can boost sales and improve customer loyalty. Privacy and quality concerns must be addressed in order to ensure that the data is used in a responsible manner. This paper will talk about the different techniques that are used in CRM research and



how they can be applied to the Bank Marketing dataset.

Literature review

One of the most critical components of a research paper is the literature review, as it provides an analysis of the current state of the art in the field. This study explores the role of data mining in the development and implementation of CRM in the banking industry. The review aims to provide an extensive analysis of the previous studies that were conducted on the subject. The goal of this review is to discuss the evolution of CRM from its traditional form to social CRM and how it is being integrated into the management of customer relationships. It will then explore the various methodologies that are used to measure and predict customer loyalty and satisfaction.

Choudhury et al.[9] explores the evolution of CRM from its traditional form to social media. They state that the changes brought about by social media have revolutionized how consumers interact with businesses, and these changes require changes in how CRM strategies are implemented. They also discuss the various opportunities and challenges that social media presents.

Faed et al.[10] discuss the importance of mapping and analyzing customer complaints in order to measure and improve the loyalty and satisfaction of consumers. They claim that this information can be used to identify patterns and improve the customer experience. They use data mining techniques to identify patterns in the complaints and come up with a methodology that can be utilized to improve these aspects.

Femina et al.[11] present a framework that aims to analyze and predict the behavior of customers using data mining techniques. They claim that this method can help improve the efficiency of CRM by extracting valuable information.

Guo et al.[12] provides an overview of the various techniques that are used in data mining for CRM. They talk about how they can be

utilized for analyzing and predicting customer behavior. They also discuss a case study that shows how these techniques can be used to predict the likelihood of a customer's churn.

Keramati et al.[13] discussed about a prediction model that can be used to identify and predict the likelihood of customers' churn from various electronic banking services. They state that this issue is a major challenge faced by the banking industry, and data mining techniques can help predict customer churn. The model they presented uses the methods to analyze customer data.

Khodakarami et al.[14] analyzed the role that CRM systems play in the creation of customer knowledge. They found that they can help facilitate this process by integrating customer data and enhancing collaboration and communication among employees. According to the authors, CRM systems can help organizations create and manage customer knowledge by allowing employees to access and share information. They also noted that this technology can promote collaboration among various departments.

Nguyen et al.[15] analyzed the factors that influence the use of CRM technology in small and medium-sized enterprises (SMEs). They identified seven factors that can influence the adoption of this technology. They found that one of the most important factors that businesses consider when it comes to adopting CRM is the ease of use.

A comprehensive review of the literature on the subject of CRM was conducted by Soltani et al.[16] They identified four main types of CRM mechanisms. These include analytical, collaborative, social, and operational. The authors of the review urged further research on the various aspects of CRM, such as its effectiveness and integration with other mechanisms. They also mentioned the need for studies on contextual and cultural factors that can influence its adoption.

Sota et al.[17] analyzed the academic literature on CRM from 2007 to 2016. They found that there were various themes in the literature,



such as the importance of customer experience and data mining in the use of this technology. They also identified the need for more research regarding the effectiveness of this technology. Trihas et al.[18] proposed a novel CRM model that can help tourism businesses improve their e-marketing efforts. They identified the three main elements of this model, which are customer segmentation, retention, and relationship management. They found that this method can help them target their marketing messages to the right customers and retain them. The authors of the study suggested that the model could be applied to other industries to help businesses improve their customer satisfaction and marketing effectiveness. These studies show how CRM can help boost customer retention, satisfaction, and loyalty. They also reveal how data mining and analytics can improve the efficiency of such strategies. The literature indicates that businesses can take advantage of this technology by implementing mechanisms and strategies that cater to their specific needs. More research is needed to learn how different types of CRM technologies and strategies can be utilized in varying contexts and industries.

Various challenges associated with CRM

Customer relationship management (CRM) data mining is a process that involves extracting valuable information from a company's customer database. However, it can be very challenging to implement and manage due to the complexity of the process.

- **Data Quality:** The quality of data collected from a CRM database can affect the results of the process. Inaccurate or incomplete information can lead to inaccurate decisions and insights. This is why it is important that the data quality is regularly monitored.
- **Data Integration:** The integration of CRM data from various sources can be very challenging. This process involves carefully mapping, transforming, and standardization of the information to make it usable for mining. Failure to do

so can lead to incomplete profiles and inaccurate insights.

- **Privacy Concerns:** Due to the nature of CRM data, it can contain sensitive information such as a customer's purchase history and personal information. Businesses should ensure that they follow proper privacy laws when it comes to the mining of this data.
- **Resource Constraints:** There are various constraints that prevent businesses from effectively implementing data mining in their CRM systems. These include the availability of skilled personnel, the technological infrastructure, and the tools needed to perform the process.
- **Analyzing Unstructured Data:** Through the use of unstructured data, such as customer feedback, social media posts, and reviews, businesses can gain insight into their customers' behavior and preferences. Unfortunately, it can be very difficult to analyze and interpret these types of data due to the complexity of their processing.
- **Keeping up with Technology:** Due to the nature of data mining and the increasing number of tools and techniques available, businesses must continuously update their technology to remain competitive. Doing so can help them gain a competitive edge and avoid potential mistakes.

Through data mining, businesses can gain a deeper understanding of their customers' preferences and behavior. Unfortunately, implementing and managing this process can be very challenging. Getting the most out of data is a vital part of any organization's operations, but it can be very challenging to do so successfully due to various factors. Some of these include selecting the wrong data sources, not validating the results, and poor data quality. Before start implementing effective data mining techniques, it's important that you identify goals and source



of information. Also ensure that the data is formatted properly and that it is clean. It's important to continuously improve and monitor your data mining techniques to ensure that

they remain effective. Doing so can help prevent common issues and enable organizations to make the most of their data.

Various areas where CRM plays crucial role

Area	Importance of CRM
Sales	CRM software helps sales teams track leads and customer interactions, enabling them to make more informed decisions and close deals more efficiently.
Marketing	CRM provides valuable data on customer preferences and behavior, which can be used to create more targeted and effective marketing campaigns.
Customer Service	CRM allows customer service teams to access detailed information on each customer, making it easier to provide personalized support and resolve issues quickly.
Analytics	CRM data can be used to generate insights and identify trends in customer behavior, which can inform business strategy and decision-making.
Collaboration	CRM systems facilitate communication and collaboration between different departments, allowing teams to work together more effectively to achieve shared goals.
Automation	CRM automates repetitive tasks such as data entry, freeing up time for employees to focus on higher-level activities and improving overall efficiency.

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays
0	59	admin.	married	secondary	no	2343	yes	no	unknown	5	may	1042	1	-1
1	56	admin.	married	secondary	no	45	no	no	unknown	5	may	1467	1	-1
2	41	technician	married	secondary	no	1270	yes	no	unknown	5	may	1389	1	-1
3	55	services	married	secondary	no	2476	yes	no	unknown	5	may	579	1	-1
4	54	admin.	married	tertiary	no	184	no	no	unknown	5	may	673	2	-1

Methodology

i. Dataset

The bank marketing dataset is a widely used research tool in CRM that allows users to collect information about their customers. It contains details about individuals who were contacted by a Portuguese bank as part of a marketing campaign. The goal of the campaign was for customers to subscribe to term deposits, which are fixed-interest rate investments. This is a commonly used data set used in CRM research. It contains information about individuals who were contacted by a Portuguese bank as part of a marketing campaign to obtain a term deposit.

The data includes 45,211 records and has 17 attributes.

ii. Preprocessing

a. Handling Missing Values: One of the most common preprocessing techniques is to remove the missing values from the data. This can be accomplished by either imputing the values or by removing the records with the missing values. For instance, in the Bank Marketing dataset, there are multiple missing values in the attributes contact, education, and job. One way to address these issues is by



implementing the mode value of an attribute.

b. **Feature Scaling:** One of the most important steps in preprocessing a dataset is to scale its features. This process can help improve the performance of various machine learning tools. One of the most common methods that can be used to do this is by using Min-Max scaling. In the Bank Marketing dataset, for instance, the balance and age features have different scales.

c. **Encoding Categorical Variables:** Bank Marketing's dataset contains various categorical variables, such as education, contact, and job. Machine learning systems typically rely on numerical data, and these variables must be encoded as numbers. A method known as one-hot encoding can be used to create binary vectors for each category. In a vector format, the job attribute's elements can be represented by a specific category. This method can help the machine learning tool understand the link between the target variable and the categorical variable.

iii. **Data mining techniques**

There are many methods that can be used to extract insights from the Bank Marketing dataset. In this section, we will talk about five common techniques. These include Logistic Regression, Naive Bayes, Random Forests, K-Nearest Neighbors, and Decision Trees.

- **Logistic Regression (LR):** A binary classification process known as logistic regression is utilized to predict an outcome based on certain inputs. For instance, in the case of a bank marketing dataset, this method can be utilized to determine if a customer will sign up for a term deposit. In addition to providing a prediction of the outcome, the models can also help analyze the relationship between multiple factors.
- **Decision Trees (DT):** A widely used machine learning algorithm for solving regression and classification problems is known as a decision tree. It divides the collected data into sub-populations according to the values of their input variables. For instance, in the Bank Marketing data, a Decision Tree can help predict the likelihood of a customer signing up for a term deposit based on various factors such as the age, job type, and balance of the applicant.
- **Naive Bayes Classifier (NB):** The Naive Bayes classifier is a machine learning algorithm that can be used to solve classification problems. It takes into account the input variable's probability and then uses Bayes' theorem to arrive at the predicted value of the class label. For instance, in the Bank Marketing dataset,



this method can be used to determine if a customer will subscribe to term deposits based on various attributes such as job type, age, and balance.

- **K-Nearest Neighbors (KNN):** The K-Nearest Neighbors algorithm is a classification method that finds the closest neighbors from a given input space. It then assigns the class label to the majority of the neighbors based on the data points. In the case of bank marketing data, this method can be used to determine if a customer would subscribe to a term deposit based on various characteristics such as job type, age, and balance. It can also be used to analyze outliers and identify the relationship between multiple factors in the dataset.
- **Random Forests (RF):** The use of a Random Forests algorithm is a powerful tool that can improve the accuracy and reliability of the predictions made by a decision tree. For instance, it can be used to analyze the data collected by the bank

marketing platform to predict if a customer will be interested in a term deposit.

The various techniques used in data mining have their own weaknesses and strengths. The decision to employ one of these depends on the specific requirements of the analysis and the data's characteristics.

Results and Outputs

The table represents the performance evaluation of various data mining techniques for CRM based on their accuracy, precision, recall, and F1 score. The data mining techniques evaluated are Logistic Regression (LR), Decision Tree (DT), Naive Bayes (NB), k-Nearest Neighbor (KNN), and Random Forest (RF). From the table, we can see that Random Forest performs the best in terms of accuracy, with a score of 90.6%. Precision and recall are also crucial evaluation metrics, with precision representing the proportion of true positive cases among all predicted positive cases, and recall representing the proportion of true positive cases among all actual positive cases. Random Forest has the highest precision score of 75.5%, while NB has the highest recall score of 73.2%. The F1 score is the harmonic mean of precision and recall and provides a balanced evaluation of both metrics. Random Forest has the highest F1 score of 75.6%, which is consistent with its high precision score as shown in table-1 and figure.2. Figure 1 represent the correlation maps of dataset.

Table 1 Various evaluation parameters

Data Mining Technique	Accuracy	Precision	Recall	F1 Score
LR	90.1	75.6	74.9	75.6
DT	88.8	70.4	72.1	72.3
NB	86.9	73.2	73.2	75.6
KNN	88.9	73.4	75.6	71.5
RF	90.6	75.5	74.4	75.6



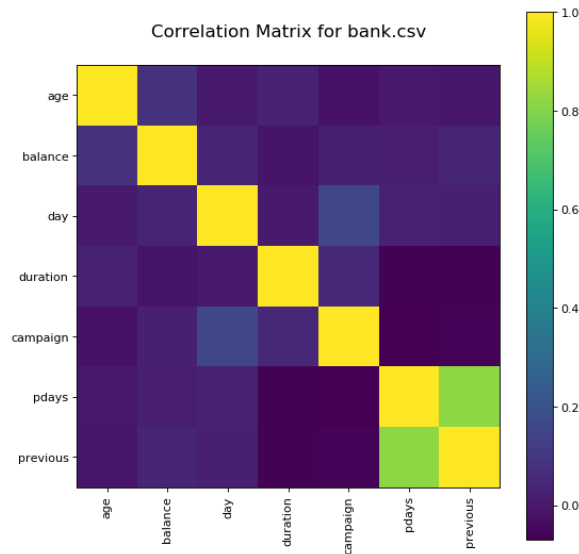


Figure 1 Correlation matrix for dataset

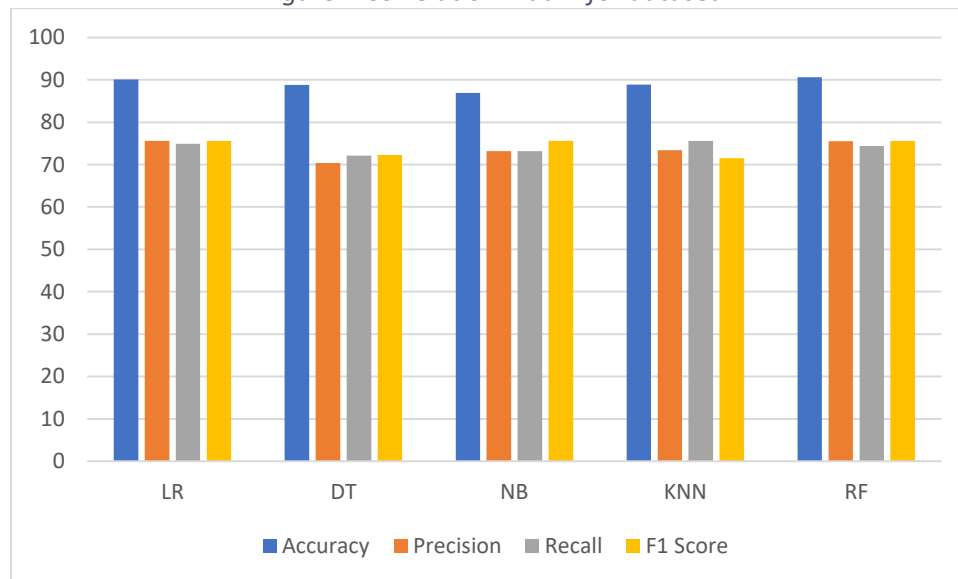


Figure 2 Graphical representation of various metrics

Conclusion and future scope

One of the most critical factors that businesses consider when it comes to improving their customer relationship management is data mining. This process can help them develop effective marketing strategies and improve their customer satisfaction. This paper explores the various advantages and challenges of this process and provides recommendations on how to improve the efficiency of your CRM system. Due to the increasing number of data sources in CRM, the scope of data mining is expected to

become more sophisticated. This process will allow businesses to collect more accurate and timely information about their customers. In addition, by integrating machine learning and AI into data mining, they can improve the efficiency of their operations and increase sales. In addition to being able to improve their customer relationship management, data mining also helps businesses grow their competitive advantage. Data mining techniques such as Random Forest can help improve the accuracy of customer data by allowing



businesses to classify and predict them more accurately. Research on the effectiveness of these new technologies in CRM will be conducted. Despite the various challenges that can be faced by organizations when it comes to data mining, it is still beneficial to consider this process as it can help them improve their customer relationship management.

References

- [1] Rababah Khalid, Mohd Haslina, and Ibrahim Huda, "Customer Relationship Management, Processes from Theory to Practice: The Pre-implementation of CRM System," *Int. J. e-Education, e-Business, e-Management e-Learning*, vol. 1, no. 1, pp. 22–27, 2011.
- [2] S. Rajagopal, "Customer Data Clustering using Data Mining Technique," vol. 3, no. 4, 2011, doi: 10.5121/ijdm.2011.3401.
- [3] Q. Yu, H. Jiang, and X. Ma, "The application of data mining technology in customer relationship management of commercial banks," *ICNC-FSKD 2018 - 14th Int. Conf. Nat. Comput. Fuzzy Syst. Knowl. Discov.*, pp. 1368–1373, 2018, doi: 10.1109/FSKD.2018.8687183.
- [4] M. Shakil Ahmad, *Impact of Organizational Culture on Performance Management Practices in Pakistan*, vol. 5. 2012.
- [5] S. K. Roy, C. Padmavathy, M. S. Balaji, and V. J. Sivakumar, "Measuring effectiveness of customer relationship management in Indian retail banks," *Int. J. Bank Mark.*, vol. 30, no. 4, pp. 246–266, 2012, doi: 10.1108/02652321211236888.
- [6] S. H. Liao, P. H. Chu, Y. J. Chen, and C. C. Chang, "Mining customer knowledge for exploring online group buying behavior," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 3708–3716, 2012, doi: 10.1016/j.eswa.2011.09.066.
- [7] S. Orenga-Roglá and R. Chalmeta, "Social customer relationship management: taking advantage of Web 2.0 and Big Data technologies," *Springerplus*, vol. 5, no. 1, 2016, doi: 10.1186/s40064-016-3128-y.
- [8] W. K. R. Perera, K. A. Dilini, and T. Kulawansa, "A Review of Big Data Analytics for Customer Relationship Management," *2018 3rd Int. Conf. Inf. Technol. Res. ICITR 2018*, 2018, doi: 10.1109/ICITR.2018.8736131.
- [9] M. M. Choudhury and P. Harrigan, "CRM to social CRM: the integration of new technologies into customer relationship management," *J. Strateg. Mark.*, vol. 22, no. 2, pp. 149–176, 2014, doi: 10.1080/0965254X.2013.876069.
- [10] A. Faed, O. K. Hussain, and E. Chang, "A methodology to map customer complaints and measure customer satisfaction and loyalty," *Serv. Oriented Comput. Appl.*, vol. 8, no. 1, pp. 33–53, 2014, doi: 10.1007/s11761-013-0142-6.
- [11] B. T. Femina and E. M. Sudheep, "An efficient CRM-data mining framework for the prediction of customer behaviour," *Procedia Comput. Sci.*, vol. 46, no. Icict 2014, pp. 725–731, 2015, doi: 10.1016/j.procs.2015.02.136.
- [12] F. Guo and H. Qin, "Data Mining Techniques for Customer Relationship Management," *J. Phys. Conf. Ser.*, vol. 910, no. 1, 2017, doi: 10.1088/1742-6596/910/1/012021.
- [13] A. Keramati, H. Ghaneei, and S. M. Mirmohammadi, "Developing a prediction model for customer churn from electronic banking services using data mining," *Financ. Innov.*, vol. 2, no. 1, 2016, doi: 10.1186/s40854-016-0029-6.
- [14] F. Khodakarami and Y. E. Chan, "Exploring the role of customer relationship management (CRM) systems in customer knowledge creation," *Inf. Manag.*, vol. 51, no. 1, pp. 27–42, 2014, doi: 10.1016/j.im.2013.09.001.
- [15] T. U. H. Nguyen and T. S. Waring, "The adoption of customer relationship management (CRM) technology in SMEs:



- An empirical study,” *J. Small Bus. Enterp. Dev.*, vol. 20, no. 4, pp. 824–848, 2013, doi: 10.1108/JSBED-01-2012-0013.
- [16] Z. Soltani and N. J. Navimipour, “Customer relationship management mechanisms: A systematic review of the state of the art literature and recommendations for future research,” *Comput. Human Behav.*, vol. 61, pp. 667–688, 2016, doi: 10.1016/j.chb.2016.03.008.
- [17] S. Sota, H. Chaudhry, A. Chamaria, and A. Chauhan, “Customer Relationship Management Research from 2007 to 2016: An Academic Literature Review,” *J. Relatsh. Mark.*, vol. 17, no. 4, pp. 277–291, 2018, doi: 10.1080/15332667.2018.1440148.
- [18] N. Trihas, G. Mastorakis, E. Perakakis, and I. Kopanakis, “Efficient e-Marketing in Tourism through a Novel Customer Relationship Management Model,” *GSTF J. Bus. Rev.*, vol. 3, no. 1, pp. 1–6, 2013, doi: 10.7603/s40706-013-0007-1.

