



## A Simplified Data Mining Framework to Plan Effective Direct Marketing Campaigns

**Bharat Gupta\***  
Amity University  
India

**Shelly Gupta**  
Banasthali University  
India

**Sheetal Sharma**  
Amity University  
India

---

**Abstract**— *The advancement in technology, competitive market environment and changing behaviour of customers has generated an increased interest in direct marketing method for marketing. Data mining technology can predict and detect the changes in customer behaviour which facilitates the effective planning of direct marketing campaigns. The objective of this paper is to generate a simplified data mining framework to facilitate the marketers who have little knowledge in data mining to effectively carry out the direct marketing campaigns. The framework is built using the literature of data mining concepts and direct marketing concepts.*

**Keywords**— *Data Mining, Knowledge Discovery, Direct Marketing, Classification Techniques, Customer Relationship Management*

---

### I. INTRODUCTION

In marketing, there are two opposed approaches to communication: mass marketing and direct marketing. In mass marketing, a communication message is broadcast to all customers through media. In spite of large investments in market research and media planning, it is hard to know the benefits of mass marketing. This has led to an increased popularity of direct marketing [1]. The final goal of direct marketing is cost-effective, two-way, one-to-one communication with each and every customer. Hence, it is essential to learn present and predict future customer preferences. Due to high demanding nature of customers' and increased competitiveness in market, it has become necessary to apply customer-focused approach for effective direct marketing. A customer-focused approach will first discover what are the customers' preferences, needs, and behavioural characteristics on recent trends. This approach improves the chances of a particular goal in being successful.

Large amounts of data are generated every day in many organizations. These data can be used to establish and maintain direct relationship with the customers in order to target them individually for specific offers. To extract hidden predictive information from large amounts of data, Organizations are now realizing the importance of data mining in their strategic planning and successful application of data mining methods that can generate useful knowledge to organizations [4]. Using data mining concepts, we propose a simplified framework of data mining for direct marketing which divide the total market and perform a sampling of best segments for effective marketing.

The rest of paper is organized as section 2 provides an overview of knowledge discovery and data mining. In section 3, we explain classification in data mining. In section 4, we describe the CRM model and in section 5, we describe the data mining framework for direct marketing and highlight the main phases and functions involved in each phase and finally, we present a conclusion and references.

### II. KNOWLEDGE DISCOVERY AND DATA MINING

In this information age, information leads to power and success. We have been collecting tremendous amounts of information. Initially, with the advent of computers and means for mass digital storage, we started collecting and storing all sorts of data. Unfortunately, these massive collections of data stored on disparate structures very rapidly became overwhelming. This initial chaos has led to the creation of structured databases and database management systems (DBMS). The efficient database management systems have been very important assets for management of a large corpus of data and especially for effective and efficient retrieval of particular information from a large collection whenever needed. The proliferation of database management systems has also contributed to recent massive gathering of all sorts of information. Today, we have far more information than we can handle: from business transactions and scientific data, to satellite pictures, text reports and military intelligence. Information retrieval is simply not enough anymore for decision-making. Confronted with huge collections of data, we have now created new needs to help us make better managerial choices. These needs are automatic summarization of data, extraction of the "essence" of information stored, and the discovery of patterns in raw data [6].

Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods. Consequently, data mining consists of more than collecting and managing data, it also includes analysis and prediction [16].

Data Mining, also popularly known as Knowledge Discovery in Databases (KDD), refers to the nontrivial extraction of implicit, previously unknown and potentially useful information from data in databases. While data mining and knowledge discovery in databases (or KDD) are frequently treated as synonyms, data mining is actually part of the knowledge discovery process. The following figure shows data mining as a step in an iterative knowledge discovery process.

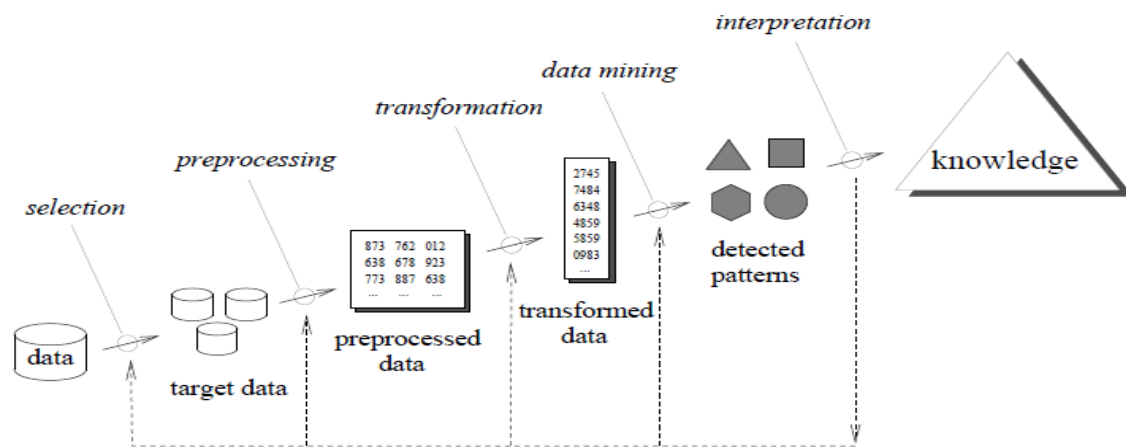


Fig.1. Steps in KDD Process

The Knowledge Discovery in Databases process comprises of a few steps leading from raw data collections to some form of new knowledge. The iterative process consists of the following steps [6]:

- **Data cleaning:** also known as data cleansing, it is a phase in which noise data and irrelevant data are removed from the collection.
- **Data integration:** at this stage, multiple data sources, often heterogeneous, may be combined in a common source.
- **Data selection:** at this step, the data relevant to the analysis is decided on and retrieved from the data collection
- **Data transformation:** also known as data consolidation, it is a phase in which the selected data is transformed into forms appropriate for the mining procedure.
- **Data mining:** it is the crucial step in which clever techniques are applied to extract patterns potentially useful.
- **Pattern evaluation:** in this step, strictly interesting patterns representing knowledge are identified based on given measures.
- **Knowledge representation:** is the final phase in which the discovered knowledge is visually represented to the user. This essential step uses visualization techniques to help users understand and interpret the data mining results.

In the KDD process, the data mining methods are for extracting patterns from data. The patterns that can be discovered depend upon the data mining tasks applied. Generally, there are two types of data mining tasks: descriptive data mining tasks that describe the general properties of the existing data, and predictive data mining tasks that attempt to do predictions based on available data. Data mining can be done on data which are in quantitative, textual, or multimedia forms.

Data mining applications can use different kind of parameters to examine the data. They include association (patterns where one event is connected to another event), sequence or path analysis (patterns where one event leads to another event), classification (identification of new patterns with predefined targets) and clustering (grouping of identical or similar objects). Data mining involves some of the following key steps-

- **Problem definition:** The first step is to identify goals. Based on the defined goal, the correct series of tools can be applied to the data to build the corresponding behavioural model.
- **Data exploration:** If the quality of data is not suitable for an accurate model then recommendations on future data collection and storage strategies can be made at this. For analysis, all data needs to be consolidated so that it can be treated consistently.
- **Data preparation:** The purpose of this step is to clean and transform the data so that missing and invalid values are treated and all known valid values are made consistent for more robust analysis.
- **Modelling:** Based on the data and the desired outcomes, a data mining algorithm or combination of algorithms is selected for analysis. These algorithms include classical techniques such as statistics, neighbourhoods and clustering but also next generation techniques such as decision trees, networks and rule based algorithms. The specific algorithm is selected based on the particular objective to be achieved and the quality of the data to be analysed.
- **Evaluation and Deployment:** Based on the results of the data mining algorithms, an analysis is conducted to determine key conclusions from the analysis and create a series of recommendations for consideration.

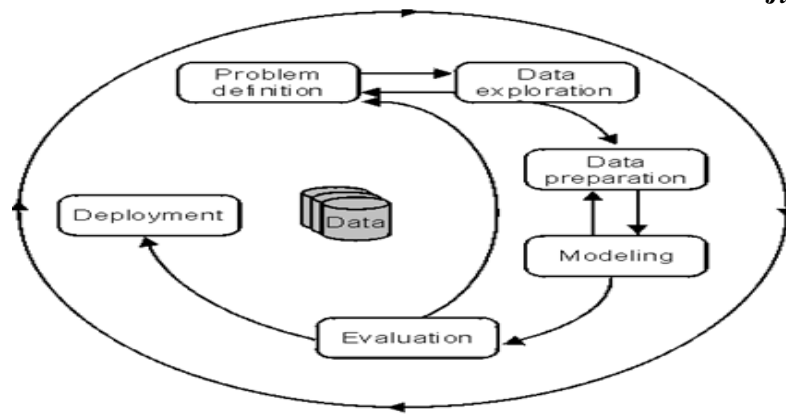


Fig.2. Data Mining Process Representation

### III. CLASSIFICATION IN DATA MINING

The Classification maps the data into predefined targets (supervised learning) [18]. The aim of the classification is to build a classifier based on some cases with some attributes to describe the objects or one attribute to describe the group of the objects. Then, the classifier is used to predict the group attributes of new cases from the domain based on the values of other attributes [6]. There are different classification algorithms used to extract relevant relationship in the data such as—

- **Decision tree induction** is the learning of decision tree from class-labeled training data set. In this method training set is broken down into smaller and smaller subsets while at the same time an associated decision tree get incrementally developed. At the end of the learning process, a decision tree covering the training set is returned.
- A **Bayes classifier** is a simple probabilistic classifier based on applying Bayes' theorem (from Bayesian statistics) with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be "independent feature model" i.e. a naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature.
- **SVM** separates a given set of binary labeled training data with a HYPERPLANE that is maximally distant from them (Maximal Margin Hyperplane MMH). When no linear separation is possible they can work in combination with the technique of KERNELS that automatically realizes a non linear mapping to a feature space.
- **KNN classification** is an example of lazy learner, which memorizes the entire training data & performs the classification when the attributes of a test tuple matches relatively similar with k nearest tuples in training data.

### IV. CUSTOMER RELATIONSHIP MANAGEMENT (CRM) MODEL

CRM involves tailoring products and services based on customers' preferences rather than some general characteristics. It is certainly critical to have a good understanding of customers' needs and preferences in order to achieve an effective CRM application. In addition, the marketing function is very important for CRM effectiveness, as it is the way companies interact with their customers [19]. There are four categories of CRM dimensions and their main purpose is to provide a deeper understanding of customers in order to capitalize on their value for the organization.

- **Customer Identification** involves target customer analysis and customer segmentation, which lead to the discovery of the population who are most likely to become customers.
- **Customer Attraction** consists of looking into customer characteristics and carrying out direct marketing.
- **Customer Retention** dimension involves direct marketing, loyalty programs, and complaints management. The main objective of this dimension is to keep customers satisfied.
- **Customer Development** entails lifetime value analysis, up/cross selling, and market basket analysis. This dimension's primary aim is to intensify transaction value and individual customer profitability.

From the above CRM description, one can conclude that direct marketing is part of CRM.

### V. DATA MINING FRAMEWORK FOR DIRECT MARKETING

The various phases of proposed data mining framework for direct marketing are outlined below:

1. **Direct Marketing Planning:** The initial step of any data mining project is to identify the marketing goals and requirements of the marketing. The next step is to translate the data mining problem into a plan in order to attain the project objective. In this step the problem and selection of data mining tools to be used is identified. For the above the activities involved in this phase are primary research, trend analysis, customer progression, customer profiling, project goals, anomalies detection and deviation analysis.
- **Customer Progression:** Lead to the discovery of the people who are likely to become customers. In other words we can say, marketing moves individuals or businesses through the progression from eyeballs to addicts.
- **Customer Profiling:** Mostly measures the outcome of a self-fulfilling prophecy. A customer profile is a model of the customer based on which the marketer decides on the right strategies and tactics to meet the needs of the customer [19].

- **Primary Research:** It includes surveys and focus groups. Surveys asking a stratified and prequalified sample of individuals a well structured set of questions to provide information that cannot be obtained through third party sources; examples of survey purposes include gauging customer loyalty, determining future intents to buy or understanding why a customer stopped purchasing. Whereas, focus groups means gathering a small group of individuals together to gauge a reaction to new creative, products or services to be marketed.
- **Trend Analysis:** Market trends are the upward or downward movement of a market, during a period of time. The market size is more difficult to estimate if one is starting with something completely new. In this case you will have to derive the figures from the number of potential customers or customer segments.( means identifying the target marketing)[22].
- **Deviation Analysis:** Deviations are useful for the discovery of anomaly and changes. Anomalies are things that are different from the normal. Anomalies can be detected by the analysis of the means, standard deviations and volatility measures from the data.
- **Project Goals:** For a campaign can be increasing awareness on brand, enhancing sales by some minimum percentage, or changing the outlook of a product. The objective is simply to attract people and recognize their needs for the product [21].

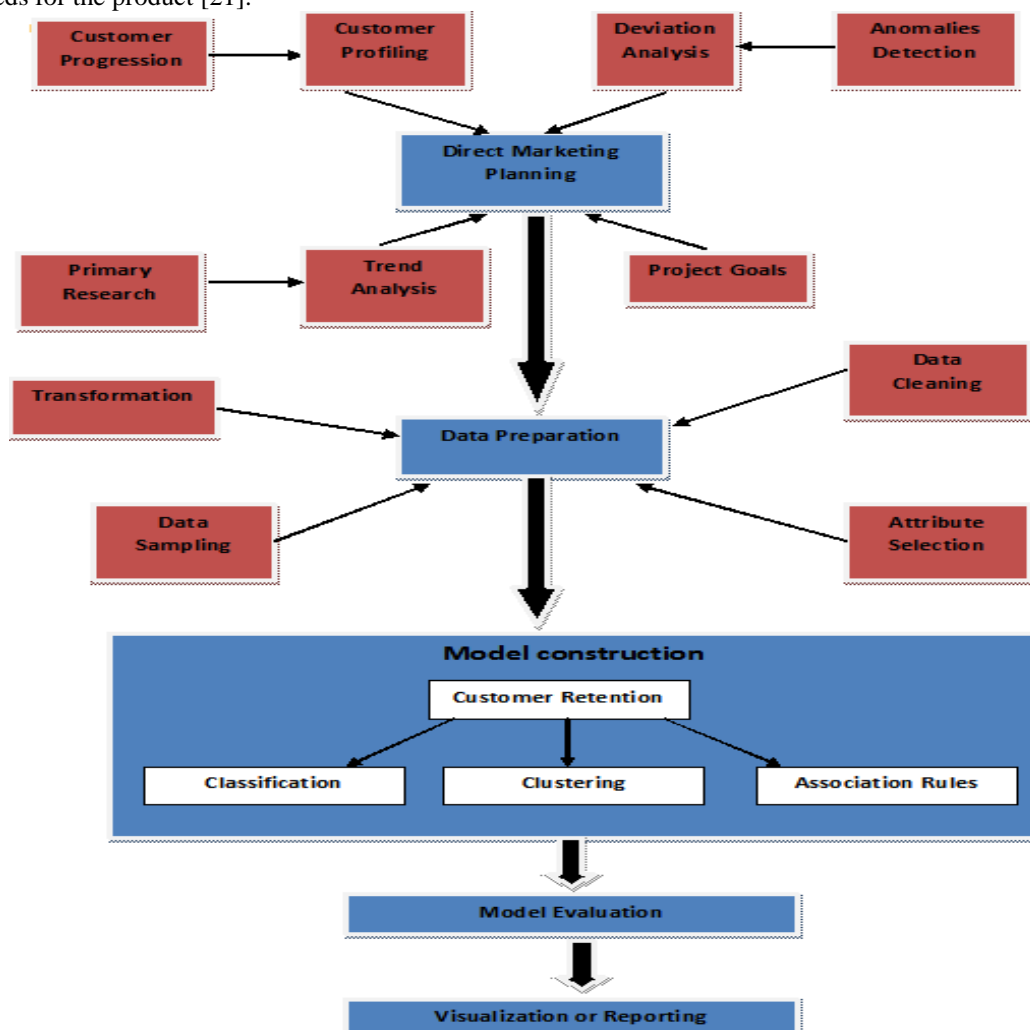


Fig.3. A Data Mining Framework for Direct Marketing

2. **Data Preparation:** The data preparation phase includes data sampling, record and attribute selection, cleaning of data and transformation.
  3. **Model construction:** When the data get prepared, the next step is to construct a model by applying the various data mining techniques. The main purpose of creating a data mining model is to identify the behaviour of customers on the marketing strategy. As the target is already predefined that either the purchase will occur or not that is why the classification techniques will be used for model construction. On the base of performance and presentation the decision tree classifier are proved the best and popular models for classification as these models generate rules that can be translated into natural language very well.
- **Customer Retention:** This is the central concern for CRM. Customer satisfaction which refers to the comparison of customers' expectations with his or her perception of being satisfied, is the essential condition for retaining customers. As such elements of customer retention include one to one i.e. direct marketing which are supported by analysing detecting and predicting changes in customer behaviour [18].

4. **Model Evaluation:** Marketers can use predictive accuracy for model analysis and evaluation.
5. **Visualization or reporting:** This stage focus to simplify the clearness about the marketing model and hence the retrieval of knowledge [21].

## VI. CONCLUSIONS

Data mining can help marketers and managers to plan effective direct marketing campaigns. Data mining technology can predict and detect the changes in customer behaviour which facilitates the effective planning of direct marketing campaigns. Here it is demonstrated that data mining is an effective tool for direct marketing which can improve direct marketing campaigns. In this paper we have outlined a simplified theoretical data mining framework for marketers and managers to carry out effectively the direct marketing campaigns using the literature concepts of data mining and marketing. This framework will help them to focus their advertising and promotions in order to reduce time and cost by targeting the interestingness of customers and their needs.

## REFERENCES

- [1] Sing L. et. al. "Data Mining Framework for Direct Marketing: A Case Study of Bank Marketing", IJCSI ,vol.10,issue 2, No.2, 2013.
- [2] Hany A. Elsalamony , Alaa. M. Elsayad "Bank Direct Marketing Based on Neural Network", International Journal of Engineering and Advanced Technology (IJEAT), vol. 2, Issue 6, August 2013.
- [3] Cote. R, Direct Marketing Trends for 2012, February 2012.
- [4] Eniafe Festus Ayetiran, Adesesan Barnabas Adeyemo "A Data Mining-Based Response Model for Target Selection in Direct Marketing", I.J. Information Technology and Computer Science ,vol. 1, pp. 9-18, February 2012.
- [5] <http://archive.ics.uci.edu/ml/datasets/Bank+Marketing>
- [6] Han J. and Kamber M., Data Mining: Concepts and Techniques, 2nd edn., San Francisco, Morgan Kauffmann Publishers, 2006.
- [7] Derrig, Richard A., and Louise A. Francis, "Distinguishing the Forest from the TREES: A Comparison of Tree-Based Data Mining Methods," Variance 2:2, pp. 184-208, 2008.
- [8] A. Flores., A. Birlutiu. "Decision Tree Models for Developing Molecular Classifiers for Cancer Diagnosis", IEEE World Congress on Computational Intelligence, pp. 10-15, June 2012.
- [9] Su-lin Pang, Ji-zhang Gong, "C5.0 Classification Algorithm and Application on Individual Credit Evaluation of Banks", Systems Engineering - Theory & Practice, vol. 29, issue 12, pp. 94-104, December 2009.
- [10] Flici A., Kevin L. Fearn A. 2011 "A Framework for Direct Marketing with Business Intelligence: An illustrative case study in Retailing", Informatics Engineering and Information Science · Communications in Computer and Information Science, vol. 252, pp. 545-558, 2011.
- [11] Bose, I., Chen, X., "Quantitative models for direct marketing: A review from systems perspective", European Journal of Operational Research vol. 195(1), pp. 1-16, 2009.
- [12] Berry, M.J.A., Linoff, G.S.: Data Mining techniques: For Marketing, Sales, and Customer Relationship Management, 2nd edn. Wiley Publishing, Inc., Indianapolis 2004.
- [13] Chen, M., Chiu, A., Chang, H., "Mining changes in customer behavior in retail marketing", Expert Systems with Applications , vol.28(4), pp. 773-781, 2005.
- [14] Page, C. and Luding, Y., "Bank manager's direct marketing dilemmas – customer's attitudes and purchase intention", International Journal of Bank Marketing, vol. 21, no.3, pp. 147-163, 2003.
- [15] Boone & Kurtz; Principles of Marketing, Thomson, 2004.
- [16] Witten, I. and Frank, E.; Data Mining – Pratical Machine Learning Tools and Techniques , 2nd edn., Elsevier, 2005.
- [17] Ackerberg Daniel A., "Advertising, learning and consumer choice experience good markets: an empirical examination.", International Economic Review, vol. 44, Issue 3, pp. 1007-1040, August 2003.
- [18] E.W.T. Ngai et. al., "Application of data mining techniques in customer relationship management:A literature review and classification", Experts Systems with Applications, vol. 36, pp. 2592-2602, 2009.
- [19] Michael J. Shaw et. al., "Knowledge management and data mining for marketing", Decision Support Systems, vol. 31, pp. 127-137, 2001.
- [20] M.Govindarajan, "A Hybrid Framework using RBF and SVM for Direct Marketing", International Journal of Advanced Computer Science and Applications, vol. 4, no. 4, 2013.
- [21] Bharat Gupta ,Shelly Gupta, Sheetal Sharma , "A Survey on Bank Direct Marketing Campaigns and Advertisement Activities to Achieve a Successful Marketing Management using Data Concepts ", International Journal of Advanced Technology in Engineering and Science, vol. 2, Issue 5, pp. 172-178, 2014.
- [22] [http://en.wikipedia.org/wiki/Market\\_analysis#Market\\_trend](http://en.wikipedia.org/wiki/Market_analysis#Market_trend)