International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

Data Mining Techniques for Customer Lifecycle Management

A. V. Murali

Software Consultant

Abstract: ***Businesses in every industry strive to increase customer value and achieve a high level of customer satisfaction. Sales teams focus on opportunities for cross-sell and up-sell while customer care focusses on certain key metrics such as first call resolution, quick resolution to customer's issues, high service levels and quality scores. Huge amounts of data are generated by various teams in an organization as a result of interactions with their customers and prospects. Data analysts perform different types of analysis on this data to meet various objectives such as finding out the root cause of issues, predicting some trend or suggesting plans and schedule. Employing data mining techniques not only gives a better insight into the problem areas but also reveals unknown associations between variables. This paper elaborates on the data mining techniques such as association rule mining, anomaly detection, classification, clustering and regression and how businesses can take advantage of these techniques to gain better insight into customer lifecycle and build better customer relationships.***

Keywords: Data mining, Association rule mining, Anomaly detection, Classification, Clustering, Regression, Customer lifecycle.

1. Introduction

Businesses move towards the goal of understanding their customers better and use that understanding to facilitate ease of doing business with them**.** Such efforts further helps them offer their customers better value than their competitors. Information about customers is their key asset because it confers on them the competitive advantage.

1.1 Customer lifecycle

Customer relationship is evolutionary in nature. Irrespective of the nature of industry, customers of every business go through an established lifecycle. The five life-cycle phases that customers go through are prospects, responders, new customers, established customers and churned customers. Forward-looking businesses develop well-thought-out strategies pertaining to every one of these five stages. For every business, one of the prime activities that goes into each of these phases is the identification and collection of relevant data. For example, to run a campaign a business would need not only the names of prospects and their contact details but also many additional information about the individuals and their demographic profiles to evolve an effective response model within the available budget so that the campaign does not fail. In the case of established customers, a business would want to increase the customer‘s value and to do so may resort to cross-selling, up-selling and usage stimulation. These activities demand a right model that would target the right customers and increase the campaign effectiveness.

Data may be available in various forms. Service businesses generally have data about a variety of activities, namely, web site visits of prospects as well as customers, the initial product and plan chosen by the customer, changes in the product or plan, their responses to campaigns, service usage patterns, bill payment pattern and in the case of churned customers the circumstances that lead to churn, either voluntarily or involuntarily. Data mining can play a key role in extracting patterns out of these volumes of data and allow us to perform various types of analysis on these data to detect flaws and weaknesses in the system at different levels.

1.2 Customer care analytics scenarios

Customer care generates volumes of data everyday as a result of inbound and outbound contacts with its customers on multiple media. Various types of analysis can be carried out on these data. Some of the analysis efforts that are carried out within customer care are listed below:

Mouse clicks per call: A customer service representative (CSR) who is attending to a call from a customer may have to search through the knowledge portal for the information that the customer is looking for. Depending on the design of the knowledge portal, this may require certain number of mouse clicks. Storing the places or links, where she clicked the mouse while searching for the required information, would enable us to measure the user-friendliness of the GUI and improve it. The same applies to CRM client application that is accessed on every call. Fewer mouse clicks per call for a given interaction type can be a useful measure to consider.

Mouse clicks on multiple client applications back and forth is another useful information that can throw light on the integration requirements of the client applications.

Interaction reasons: Summarization of interaction data gives us the most frequently encountered interaction reasons (say, Top 10) in customer care. This is commonly tracked in customer care to reduce such calls by either transferring that information to IVR or by proactively sending out information using a dialer or by SMS or by Email. Alternatively, we may try to reduce the time required to handle such calls.

Self-care click streams: Tracking the clickstreams on the self-care portal of an organization and analyzing this information can help us improve the navigational features of the website.

IVR traversals: An Interactive Voice Response (IVR) system greets the caller who contacts customer care and offers information using a hierarchical menu structure. Analyzing the traversal sequence of the menu items, the scripts that were listened to either fully or partially, the time spent there, the repeat hits on a menu item and the number of menu items that were accessed in a call can give insight into the effectiveness of IVR and the areas for improvement.

Analyzing the IVR traversal data in conjunction with clicks on CRM and the knowledge portal at the agent desk for the same customer on the same call can help us look at improvements in IVR and transferring of information between them to reduce costs and ensure better services to customer.

Repeat call data: Customers make repeat calls when they do not get resolution in time. There can be other reasons for this customer behavior, like lack of clarity of information given by CSR. A detailed analysis of such calls can help us identify areas for improvement.

CSR interaction data: CSRs note down the reasons for interaction and a description of the specific problem faced by the customer in text form. Analyzing this text is challenging but can yield new information.

2. Data mining methods

Data mining is the computational process of exploration and analysis of large amounts of data with the objective of discovering meaningful patterns [1]. Its goal is the extraction of previously unknown patterns and knowledge from large amounts of data. Data mining helps when an organization realizes that informed decisions are preferable over uninformed decisions. Data mining techniques must be applied to the right problems using the right data.

Data mining employs metrics for interestingness and inference mechanisms along with specialized algorithms to discover interesting patterns in data. Patterns that may be of interest include dependencies or co-occurrences among items (or events), groups of data records and unusual records. Identification of such patterns in the data are studied using methods of association rule mining, cluster analysis and anomaly detection respectively. Classification and regression are other two analytical tasks in data mining [2]. These data mining functions can be implemented using a variety of technologies that fall under database-oriented techniques, machine learning and statistical techniques. Data mining is an integral part of the more comprehensive process of Knowledge Discovery in Databases (KDD). CRSIP-DM [3] [4] and SEMMA [5] are the two popular variations of KDD that are followed by the data mining community.

The following section gives a brief description of the five types of data mining tasks that we commonly encounter in analytics scenarios. The variations in these techniques and their relative merits are discussed.

2.1 Association rule mining

An association rule is a statement that says when an event A occurs, an event B will also occur with certain probability. It can be represented as ―If A Then B" or as ―A → B". Association rule mining is used to find out frequent co-occurring associations among a collection of items that are seemingly unrelated. They are expressed in the form of if-then rules. An example of an association rule would be "If a customer visits a shopping mall, he is 90% likely to have some food at one of the outlets there. In a large dataset there can be too many rules to be of interest and only a few of them would be of significance. To arrive at a small but significant set of rules, certain minimum thresholds on ‗support' and ‗confidence' are imposed. Support of an item set with respect to a set of transactions in a database is the proportion of the transactions that contains that item set. Confidence value of a rule like A → B with respect to a set of transactions is the proportion of transactions that contains A which also contains B. Association rules are derived by looking for frequent if-then patterns in the data using the support and confidence criteria to filter out less likely relationships. In short the goal is to find all rules that satisfy the user-specified minimum support and minimum confidence. More details on this can be found in [6]. In addition to support and confidence, many other measures of interestingness can be found in the literature. They are, collective strength [7], leverage [8], conviction [9], lift [10] and all-confidence [11].

There are several variations to the idea of association rules including sequential rules that consider time ordering of events and sequential pattern mining. Higher order pattern discovery is about events or patterns that are polythetic in nature.

Several algorithms have been proposed for mining association rules. Some of the first algorithms for generating frequent item sets include SETM, AIS, A priori, Eclat (Equivalence Class Transformation) and FP-Growth (FP stands for frequent pattern). Other algorithms that have been proposed include AprioriDP [12], Node-set-based algorithms including FIN [13], PrePost [14] and PPV [15], Context Based Association Rule Mining Algorithm (CBPNARM) [16], GUHA procedure ASSOC [17] and OPUS search [18].

Apriori was among the earliest of the algorithms for association rule mining. Its performance is better than that of SETM and AIS. It works best for ‗closed item sets'. An itemset is closed if there exists no superset of it that has the same support count as this itemset. While it generates less candidate sets it consumes a lot of memory. A variation of it namely AprioriTID works better than Apriori in later passes on the dataset. Apriori Hybrid is another algorithm that makes use of the advantages of Apriori and APrioriTID. Eclat, which is a depth first search algorithm using set intersection, fares better with free item sets and utilizes less memory. FP-Growth does not utilize candidate set generation unlike Apriori but uses tree structure that adds to complexity. Its execution time is much smaller than Apriori in the lower ranges for support factor. Yet another algorithm called Matrix Apriori works faster than FP-Growth when finding itemsets although it takes more time to build a matrix structure than it takes to build a tree structure.

The PPV algorithm is better than APriori, Eclat and FP-growth and it derives its performance from the use of node-lists, a compact vertical structure and the use of intersection node-lists for counting support, which reduces complexity. AprioriDP uses dynamic programming and it eliminates candidate generation just like FP-growth but stores support count in a specialized data structure instead of a tree. CBPNARM, which uses context for mining association rules, uses context variable that decides changes to the support of an itemset which in turn decides population of rules to the rule set. FIN and PrePost are based on node sets just like PPV. They use FP-tree nodes to represent itemsets and employ depth-first search to discover frequent itemsets. ASSOC is a GUHA procedure that uses fast bitstring operations for mining association rules. Association rules mined by GUHA are more general than APriori in the sense that both conjunction and disjunction can be used for linking items. Moreover, the relation between antecedent and consequent of the rule is not restricted to setting minimum support and confidence. OPUS search does not require either monotone or anti-monotone constraints such as minimum support unlike most alternative algorithms. It can be used to find rules with any item as a consequent.

2.2 Cluster Analysis

Cluster analysis aims at grouping of similar or closely related items together. There are many cluster models and accordingly there are many algorithms for clustering [19]. Some of the common clustering models are distribution models, density models, centroid models, connectivity models, graph-based models, subspace models and group models. Each of these models has associated algorithms that may not work for other models. In the case of distribution models, clusters are modelled using statistical distributions like multivariate normal distribution. Density models define clusters as connected dense regions in the data space; DBSCAN and OPTICS are examples of this. k-means algorithm is an example of centroid models; it represents each cluster by a single mean vector. Given a set of n inputs, k-means clustering groups them into k clusters in such a way that each input belongs to the cluster with the closest mean. Connectivity models use distance connectivity as in the case of hierarchical clustering. In the case of graph-based models, connectivity between every pair of nodes in a subset of a graph is used to represent a cluster. In subspace models, clusters are modelled with cluster members as well as the relevant attributes, which is a form of bi-clustering. Group models provide grouping information but do not refine it further.

2.3 Anomaly Detection

Anomaly detection is the identification of items or events which do not conform to an overall observed pattern that other items or events belong to. In real world scenarios, anomalies are usually of the nature of some defect or an undesirable event like a fraud or hacker attack. Some of the popular anomaly detection techniques are ensemble techniques using feature bagging, score normalization, cluster analysis-based outlier detection, replicator neural networks, one class support vector machines, subspace and correlation-based outlier detection and density based techniques including local outlier factor and k-nearest neighbour [20].

2.4 Classification

Classification is the task of identifying the class or category that an observation belongs to on the basis of observations that belong to known categories. In machine learning, classification belongs to supervised learning [21] whereas clustering that we saw earlier is an unsupervised form of learning. Both of them belong to the more general class of pattern recognition problems. Some of the popular classification algorithms are linear classifiers, quadratic classifiers, decision trees, neural networks, support vector machines and learning vector quantization (LVQ) [22], [23].

A linear classifier identifies the class that an object belongs to, using the value of the linear combination of the object‘s characteristics, also called feature values that are fed into the classifier in the form of a feature vector. In quadratic classifiers, the separation between the classes is achieved using a quadric surface. It works better than linear classifier in scenarios requiring a more complex surface for separation. Decision trees resemble a flowchart. Every node of a decision tree represents a test or an if-condition on an attribute, the branch coming out of the node represents an outcome of the test and the leaf nodes represent a class. While decision trees are easy to understand and interpret, they are not suitable for handling problems where there is uncertainty in values or when outcomes are linked.

Neural networks are part of machine learning paradigm and they are particularly suited for applications where the input set is large and complex enough to estimate the function from observations. They can be used in both supervised (classification) and unsupervised (clustering) learning scenarios. Support vector machines are supervised learning models having an associated learning algorithm that builds a model assigning new items into one of the existing classes. An SVM model represents example items as points in space that are mapped in such a way that there is a clear separation between items belonging to different classes. SVMs can perform both linear and non-linear classifications. LVQ is a prototype based supervised classification algorithm used widely in classification of text documents. It uses a distance measure to find the prototype that is closest to a given input and by further adapting the position of the prototype based on the correctness of classification.

2.5 Regression

Regression is a statistical measure that determines the relationship between a dependent variable and one or more independent variables. There are several regression models [24], including linear, polynomial, logistic, probit, principal components and Poisson, to mention a few. These methods can be used in forecasting and prediction. Their domain of applicability varies with the nature of distribution of the response variable and the nature of data. While Poisson distribution is suited for count data, probit and logistic regression are used when the dependent variable is categorical and binary in nature. Probit uses cumulative normal distribution while logistic regression uses a logistic function. Principal component regression uses linear regression model for regressing the dependent variable on a set of explanatory variables but uses principal component analysis for estimating the unknown regression coefficients. Principal component regression not only helps in dimension reduction in cases where there are many explanatory variables, it also overcomes the problem of multi-collinearity that arises when some explanatory variables are close to being collinear, by excluding some of the low-variance principal components in the regression step.

3. Data mining in customer lifecycle

3.1 The broad steps

Data mining techniques are being applied by forward-looking organizations in a variety of scenarios that arise during their customer life cycle. Before embarking on a data mining project, the business goal should be clear to the data miner. Data miner should first understand the business goal and translate that into data mining tasks. Only after the tasks are identified he should go to the next step of choosing the right data mining techniques. The outcome of the data mining technique should be interpreted correctly to take right business actions.

3.2 Data mining types

Data mining can be broadly divided into two types, namely directed data mining where the focus is on one or more target variables and undirected data mining where we look for overall patterns in the data with no special roles assigned to any variable. Taking the case of finding fraud customers, a directed data mining approach would try to match records with those of known fraudulent cases while in the case of undirected approach we look for records that are unusual or differ from other records in some way.

3.3 Information types

Coming to the types of data for analysis, every record in a transactional database is a result of a certain event. It could be a service event or a sale event. Mining of co-occurrence of such events can be carried out by considering all forms of service interactions either separately or in conjunction with sales transaction data. Whether it is to establish a hypothesis or to discover something unknown, it may be useful to treat all such events homogeneously irrespective of whether it is a service event or a sale event.

3.4 Data preparation

In large organizations, data of relevance to data mining typically occur in various systems in raw form. Raw data needs to be cleaned. Missing data and outliers may need to be filled with appropriate values or discarded. Categorical data needs to be transformed to numeric values as some data mining techniques can only handle numeric values. In some cases, calculations may need to be performed on different fields in the raw data to arrive at derived data that is more suited for the mining tasks. In the case of data mining using neural networks, it is required to standardize different variables so that they all fall within a small range. Where there are too many variables, it may be required to reduce the number of variables to a few using techniques such as principal components.

3.5 Data Availability

Information capture and availability of right data is crucial to data mining. Understanding prospects may require data about existing customers. For example, it would be possible to evolve a model that relates acquisition-time data of existing customers with future outcomes of interest. Knowledge about who responded to a campaign and who did not can be a useful component of future response models. Such models should additionally include information about the campaign such as the channel used to contact customer, the channel used by the customer for response, the campaign message, the timing of the message and the timing of response from the customer. Such information can help in matching campaigns to different categories of prospects to achieve campaign effectiveness.

3.6 Lifecycle Scenarios

The following paragraphs furnish instances of application of specific data mining techniques to scenarios that arise in customer life cycle management.

Analysis of customer interactions and complaints employ all the data mining techniques that we discussed. Classification is one technique that is very common among all business scenarios. Telecom operators classify customers in to different credit classes based on several variables that together can be referred as customer signature. Credit scores are assigned to customers which in turn are used to classify them. Decision trees are used to decide on campaign strategies. A comparison of customer profile with values in decision rules can provide the right campaign strategy across demographics. Classification problems can also be handled using logistic regression, support vector machines and nearest neighbor models like memory based reasoning.

In the case of established customers, cross-selling, up-selling and usage stimulation can be done to increase customer value. A right model should be evolved that finds the right time for an offer or making recommendations. Association rule mining helps in this task by looking up past customer behavior and identifying patterns in the events that have taken place at different phases of customer lifecycle. Customers with similar purchase behavior form clusters. Association rule mining is widely used in determining affinity groups among products, i.e., those products that sell together.

We have seen a variety of clustering techniques. Distance and similarity measures are used to identify clusters in the dataset. Clusters are identified based on what its constituents have in common. It is also possible that some clusters have a characteristic that the rest of the entire dataset does not have. Clustering algorithms can be used in detecting outliers and has been used as a means to identify overlooked customer segments. In deciding which customers to include for a direct marketing campaign, it would be required to group customers into clusters by analyzing the customer signatures. Binary response models are built for the products to be marketed to estimate the customer propensity scores for those products. This can be done using logistic regression or a decision tree. Propensity scores can be arrived at using clustering techniques like k-means clustering, wherein the dominant product in each cluster is used to assign propensity score. RFM which stands for Recency-Frequency-Monetary is a widely used lookup model that is used in estimating customer response to a direct marketing campaign. The three parameters of RFM form a cube which is divided into cubic cells. Customers are assigned to these cells and these assignments can change over time based on customer behavior. Although table lookup models like RFM are simple, they suffer from the problem of fewer data samples with increase in number of input variables. Naïve Bayesian models can be used in such situations.

Different clustering algorithms are used for different scenarios. k-means clustering creates clusters whose boundaries may extend into the neighboring clusters. This can be problematic when there are outliers which are far away. k-medians is less sensitive to outliers as we try to determine the central item in the set of items instead of the average value, but it is computationally more expensive than k-means. A problem with both k-means and k-medians is that they allow the centroid to fall where a cluster member cannot be placed. k-medoids clustering overcomes this problem by choosing one item as the most representative for the cluster. k-means works well when clusters are clearly separated. But in real life business scenarios this is hardly the case. Situations where the cluster membership is fuzzy in nature employ soft clusters, wherein an item can be associated to more than one cluster. Gaussian mixture model (GMM), also referred as Expectation Maximization (EM) clustering technique is used in such scenarios where we would like to know the likelihood of each item being in every cluster.

Attrition is one area where data mining is widely employed. Attrition modelling is an example of directed data mining where we look for patterns that explain the target values. Attrition models typically produce scores for the likelihood of customers leaving within a certain time horizon. Businesses develop attrition models to ensure that retention campaigns are successful and also cost-effective. The model should make an accurate prediction of at-risk customers who would probably require an extra incentive without which they might leave. Inaccurate predictions can lead to wrong treatments that can alienate otherwise loyal customers. While building attrition models, it would be better to consider all types of attrition, namely voluntary attrition, involuntary attrition and expected attrition. Attrition models can be built using decision trees or logistic regression. Neural networks can also be useful in cases where the underlying relationships between the explanatory variables and the target variable are not precisely understood.

While linear regression is commonly used in modelling target values that are unbounded, modelling binary outcomes like yes or no is common in several business scenarios. Logistic regression can be used in these scenarios.

Anomaly detection techniques are useful in outlier detection which can mean detecting undesirable occurrences such as customer fraud or product defects. Excess usage of a service by a post-paid customer is one such example. Complaints about a specific model of a phone, network intrusion, errors in a specific media campaign, and adverse comments on social media about a specific business outlet are all examples of outliers that need immediate detection and rectification. Anomaly detection techniques mentioned in the previous section can be applied to these situations.

Decision trees can be used in many ways. It is a useful technique in prediction, classification and estimation. Decision trees are employed even in the selection of variables for the chosen model. Once a decision tree is built from a list of input variables and a target variable, it can be used to assign scores to a new subject, which can be a customer. Decision trees can be used to generate rankings among customers as in the case of campaigns. They can be used to classify customers using class probabilities in scenarios that require knowing the probabilities of class membership. They are also used in cluster assignment wherein a set of rules is derived describing the clusters. Decision trees are useful in learning about churn in cases where the effect of different variables on retention is available.

Text mining techniques are useful in analyzing textual data that appears in the form of CSR interaction with customers, social media posts, emails, forum posts, blogs, tweets, emails and product reviews. Text mining can be done using natural language processing techniques, statistical and probabilistic techniques and machine learning algorithms. Text mining techniques are used for mining sequential patterns, frequent itemsets, co-occurring terms and several others which in turn help in doing such things as sentiment analysis, frequent problem identification and churn analytics.

At the heart of all these data mining efforts is the business goal of building customer relationships. While data mining can provide new business insights, it is up to the data miner to make judicious choices and carry out every step of the data mining project meticulously so that the outputs are not only meaningful and actionable but equally importantly contribute to achieving the business goal.

4. Conclusion

We have seen different data mining algorithms and their applications to different phases of customer lifecycle management. While some techniques can be used in more than one business scenario, relative merits and demerits have to be considered in choosing the right data mining technique for the problem at hand. To reap the benefits of data mining, the data miner should have a clear idea about the business goal so that he can translate it into data mining tasks which in turn can be accomplished using the right techniques. Advanced data mining techniques can provide answers to many of the questions that cannot be answered using classical data analysis methods alone. Data mining helps in building customer relationships by providing a variety of insights about customers. This in turn helps in enhancing customer value and building customer loyalty over time.

References

[1] "Data Mining Curriculum". ACM SIGKDD. 2006-04-30. Retrieved 2014-01-27.
[2] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic (1996). "From Data Mining to Knowledge Discovery in Databases" (pdf).
[3] Shearer C., (2005). The CRISP-DM model: The New Blueprint for Data Mining, J Data Warehousing. 5:13—22.
[4] Gregory Piatetsky-Shapiro (2014), KDnuggets Methodology Poll.
[5] Azevedo, A. and Santos, M. F. (2008); KDD, SEMMA and CRISP-DM: A Parallel Overview. In Proceedings of the IADIS European Conference on Data Mining, pp. 182-185.
[6] Agrawal, R.; Imieliński, T.; Swami, A. (1993). "Mining Association Rules Between Sets of Items in Large Databases". Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data - SIGMOD '93. p. 207.
[7] Aggarwal, Charu C.; and Yu, Philip S.(1998); A New Framework for Itemset Generation, in PODS 98, Symposium on Principles of Database Systems, Seattle, WA, USA, pages 18-24.
[8] Piatetsky-Shapiro, Gregory. (1991). Discovery, Analysis, and Presentation of Strong Rules, Knowledge Discovery in Databases, pp. 229-248.
[9] Brin, Sergey; Motwani, Rajeev; Ullman, Jeffrey D.; and Tsur, Shalom; Dynamic itemset counting and implication rules for market basket data, in SIGMOD 1997, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD 1997), Tucson, Arizona, USA, May 1997, pp. 255-264.
[10] Brin, Sergey; Motwani, Rajeev; Ullman, Jeffrey D.(1997); and Tsur, Shalom; Dynamic Itemset Counting and Implication Rules for Market Basket Data, in SIGMOD 1997, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD 1997), Tucson, Arizona, USA, pp. 265-276.
[11] Omiecinski, Edward R. (2003); Alternative Interest Measures for Mining Associations in Databases, IEEE Transactions on Knowledge and Data Engineering, 15(1):57-69.
[12] D. Bhalodiya, K. M. Patel and C. Patel. (2013). An Efficient Way to Find Frequent Pattern with Dynamic Programming Approach. Nirma Univ. Intl. Conf. on Engineering, pp. 28-30 Nov.
[13] Z. H. Deng and S. L. Lv. (2014). Fast Mining Frequent Itemsets Using Nodesets. Expert Systems with Applications, 41(10): 4505–4512.
[14] Z. H. Deng, Z. Wang, and J. Jiang. (2012). A New Algorithm for Fast Mining Frequent Itemsets Using N-Lists [3]. SCIENCE CHINA Information Sciences, 55 (9): 2008 – 2030.
[15] Z. H. Deng and Z. Wang. (2010) A New Fast Vertical Method for Mining Frequent Patterns [4]. International Journal of Computational Intelligence Systems, 3(6): 733 - 744.
[16] Shaheen, M; Shahbaz, M; and Guergachi (2013). A; Context Based Positive and Negative Spatio Temporal Association Rule Mining, Elsevier Knowledge-Based Systems, pp. 261-273.
[17] Hájek, Petr; Havránek, Tomáš (1978). Mechanizing Hypothesis Formation: Mathematical Foundations for a General Theory. Springer-Verlag.
[18] Webb, Geoffrey I. (1995); OPUS: An Efficient Admissible Algorithm for Unordered Search, Journal of Artificial Intelligence Research 3, Menlo Park, CA: AAAI Press, pp. 431-465.
[19] Everitt, Brian (2011). Cluster Analysis. Chichester, West Sussex, U.K: Wiley.
[20] Chandola, V.; Banerjee, A.; Kumar, V. (2009). "Anomaly detection: A Survey". ACM Computing Surveys 41 (3): 1–58.
[21] Alpaydin, Ethem (2010). Introduction to Machine Learning. MIT Press.
[22] Rao, C.R. (1952) Advanced Statistical Methods in Multivariate Analysis, Wiley.
[23] Alpaydin, Ethem (2010). Introduction to Machine Learning. MIT Press.
[24] A. Sen, M. Srivastava. (2011). Regression Analysis - Theory, Methods, and Applications, Springer-Verlag, Berlin.

Author Profile

A.V. Murali received his B.Tech and M.Tech degrees in Metallurgical Engineering from IIT Madras. He has got an extensive work experience of over 25 years in diverse software application domains. His areas of work are primarily centered on applied research and systems implementation. His areas of research interest include software engineering, data mining and text mining, mathematical problem solving, gamification and game based learning.