

A Comparative Study of various Recommendation System to Identify Key Customer

Karthikeyan S M (B.E., M.Tech1)
Asst. Prof, Dept. Of CS & E,
Adichunchanagiri Institute of Technology
Chikkamagaluru-577101, Karnataka, India

Dr.Pushpa Ravikumar (B.E., M.Tech., Ph. D2)
Professor and HOD, Dept. Of CS & E,
Adichunchanagiri Institute of Technology
Chikkamagaluru-577101, Karnataka, India

Abstract: Emergence of recommendation frameworks is software devices and methods committed to create important proposals about latest things for specific customer. In the past few year recommending frameworks comprise altered the method of correspondence among the two sites and customer. The main aim is to identify and recommend a key customer using social network analysis techniques. The loan dataset is worn to perform a analysis of key customer. Data mining is used to building a model for preprocessing, feature selection, construction of banking social network, and identifying and recommending key customer. Techniques which are used to build a social network is ego centric network for customer seeking loan and for extracting the transaction pattern of banking customer. Scoring based on behaviour and pattern of the customer is calculated using classification techniques. Classification of customers using ANN, SVM, and Regression techniques. The social network analysis technique and data mining technique is used and compared to identify key customer for recommend.

Keywords:

Data Mining, Customer Behaviour, Feature Extraction, Centrality Measure, Key-Customer Identification, Recommendation.

I. INTRODUCTION

Data mining is advancing into a deliberately significant measurement for some business associations including banking area. It is a technique for examining the information from various perspectives and summing up it into significant data. Data mining helps the banks to search for hidden pattern design in a gathering and find obscure relationship in the information. Data mining can add to taking care of business issues in banking and money by discovering examples, causalities, and relationships in business data and market costs that are not promptly clear to administrators in light of the fact that the volume information is excessively enormous or is produced excessively fast to screen by specialists [1]. Data mining methods help organizations' especially banking, telecommunication, insurance and retailing to construct precise customer profile dependent on customer behaviour. Subsequently it is turning into a need in this serious condition to break down the information from information distribution center containing many gigabytes or terabytes of information.

Data mining apparatuses predict patterns, future patterns and practices, permitting organizations to impact proactive, knowledge driven choices. The computerized, forthcoming examinations offered by information mining move past the investigation of past occasions gave by review apparatuses run of the mill of choice emotionally supportive networks [2]. The significance of gathering and breaking down information mirrors any business movement to accomplish upper hand is generally perceived in the present period of data. Modelling and researched framework and finding relations that interface factors in a database is the target of data mining [7].

Data mining utilizes various models for the formation of data about information which is known as discovery model. Data mining utilizes techniques that can filter through the information looking for every now and again happening designs, can identify patterns, produce speculations about the information, and so on. These tools can find these kinds of data with next to no direction from the customer. The principle errand for example, prediction, classification, detection of relations, explicit modelling, Clustering and Deviation Identification. Also, since the data mining process is orderly, it offers firms the capacity to find hidden patterns in their data designs that can assist them with understanding customer conduct and market patterns. [7].

There are two sorts of ways to deal with create suggestion outline [1].

1. Content Based Filtering Method [2]
2. Collaborative Filtering Method [2]

Sometimes mix of together methodologies [2] is favored through the analysts.

Content based [2] separating procedure sifting is set dependent on customer intrigued things. In content based sifting strategy, a site pages are suggested for customer rapidly starting older database. That database distinctive substance of a things be included that a customer have utilized in antiquated occasions or else potentially customer very own data and inclinations. The

customer information records can be developed by utilizing reactions to questions, thing evaluations, or the customer route data to gather the customer inclinations and interests. Utilizing this strategy, proposal should exist possible for the most part from the accessible information bottom and history experience of a site guest. The detriment of a technique is, not a occasions customer present their appraisals appropriately in favor of a site or site page [2].

In community separating move toward, [3] website page are prescribed to specific customer when further comparable sort of a customer additionally favors those page. Intended for instance, it might be a characterized because customers have comparative evaluation of site page otherwise site or customer having associated route conduct.

A community oriented sift framework gathers all data about customer enthusiasm on site from the web server/database in addition to ascertain the similitude in the middle of the customers intrigue. Customer include comparative qualities resolve to be classified in the direction of a similar gathering. The strategy had two inconveniences: i) Sparsity [4] ii) Scalability [5]. Suggestion frameworks utilizing community sifting way to deal with discover the area for the most part require outstandingly long calculation occasion that develops straightly with both a quantity of a customer and quantity of items or else website page.

II. RELATED WORK

Bamshad [6] have projected another calculation dependent on network use mining called profile aggregations. Within that calculation clustering is done on a database concerning relative sort of exchanges and furthermore online visit bunching is applied to a expect the comparative page in every exchange.

Yoon Ho Cho [7] have utilized choice hierarchy enlistment strategy, affiliation regulation mining calculations and information ware housing advancements to a take care of a issue of sparsity and adaptability in a community oriented sorting out method. As a result of that new half breed strategy has better the effectiveness of a community oriented separating come near by utilizing network utilization mining. Creators have utilized network logs as database to locate successive examples utilizing apriori calculation. Furthermore is used to group the customer creator use decision tree enlistment strategy.

Olfa Nasraoui [8] proposed fluffy estimation thinking technique on astute web suggestion framework. They have removed the client profile utilizing utilized web utilization mining and furthermore they apply bunching strategy for gathering the customer data on customer database. For grouping they have utilized various leveled unaided bunching strategy. Furthermore, for proposal they have utilized Fluffy estimate thinking

strategies.

Magdalini Eirinaki [9] have anticipated a bunching technique to a deliver better and fast suggestions to the end of a customer. Designed for grouping they utilized semantically by a intelligible bunches. Likewise designed for proposal they make use of Area metaphysics which ever depends on a catchphrases detached from a web substance.

Feng Hsu Wanga [10] have utilized bunching in addition to affiliation imperative mining utilizing network use digging for improved suggestion. Designed for bunching they have utilized various leveled bisect medioids.

Baoyao Zhou [11] proposed consecutive example digging strategy for anticipating the following site page. In the second step they used model base sifting strategy, which stores a consecutive network get to designs, in addition to furthermore helpful for a customer design coordinating and suggestion rules age.

Samadhi et al [12] have utilized Web use digging methods for deciding the enthusiasm of "comparable" customers. Creators separate it interested in two primary parts

i) Offline: This part of a creators use Information pre-processing, Example Revelation plus Example Examination strategies.

ii) Online: This part coordinate a current customer profile to a total utilization profile is finished. For this reason utilize various proposals methods. This work help for civilizing suggestion framework excellence, utilize dynamic customer search design which is correlation of various dynamic customers search design.

Haibo Liu [13] work utilized a information structures, for example, Intrigue Lattice, Customer Intrigue Framework, Class- Intrigue Network and Continuous Way Grid for proposal and personalization of sites dependent on the customer intrigue. On account of that item proposal is created based on customer intrigue.

Florent Garcin[14] have used setting ranking method for better proposal of information and story. Using framework gives improved proposal based on a customer's advantage. On account of with the purpose of improved forecast exactness and a proposal quality.

III. COLLABORATIVE FILTERING TECHNIQUE

In a proposal framework Community leaning separating procedure assumes significant job. It utilizes just the score information diagonally enormous dataset. Collaborative filtering various customer score 'N' things or contain

comparable practices as a result CF resolve rate otherwise follow up on top of different things likewise. Collaborative filtering method utilizes effectively accessible data from log servers identified with things/customer enthusiasm

to foresee things/customer enthusiasm to various dynamic customers which may like dynamic customer. [1] It deals with the increasing number of customer and things, to make a successful acclamation in a brief timeframe period and furthermore to manage different issues like virus star issue, synonymy, and information clamor, CF calculations must arrangement with exceptionally meager information. Essentially CF procedures are separated in three sections:

- ❖ Memory Based Collaborative Filtering Technique
- ❖ Model Based Collaborative Filtering Technique
- ❖ Hybrid Recommendation Technique

A. *Memory Based Collaborative Filtering Technique*

It utilizes finished dataset identified with customer thing dataset. It has portray by breeze memory based CF calculations by and huge use ranking grid to store customer thing catalog to produce suggestion. For the most part in memory base communitarian separating strategy use fellow citizen thing dataset to discover a enthusiasm of the customer, it use in future of every one evaluations by alluding to a customer otherwise things whose appraisals are like the additional customer.

B. *Model Based Collaborative Filtering Technique*

The disadvantage of this filtering is it utilize total datasets identified with customer thing datasets and on account of that this framework isn't fill in as quick as other community oriented framework and furthermore happens adaptability issue when create constant sections in suggestions framework catalog. To strike those issues, model based suggestion frameworks be presented by analysts. It proposes the frameworks utilize a little datasets call as model.

The configuration utilizing removing a few data from a enormous catalog identified with specific boundary/property and utilizations this model each time without utilizing tremendous database, due to that models speeds up and versatility of proposal framework.

The structure and progress of models grant the framework to perceive fairly complex examples dependent on the preparation information, and afterward issue proposals for the community oriented separating errands for testing information or genuine information, in light of a fitted model. This filter calculation incorporates Bayesian model, bunch base CF and relapse based technique to tackle a weaknesses of memory based calculations.

C. Hybrid recommendation technique

Presently a day's half and half community oriented separating is progressively famous on the grounds that it improves nature of site page suggestion or customer intrigue proposal. Crossover Community oriented Separating frameworks consolidate Shared Sifting with other proposal procedures, to improve expectations or suggestions of site pages to new customers [7][14]. The crossover proposal strategies are essentially separated addicted to two sections in first segment that incorporate every part of preprocessing techniques and second stage that incorporates every standard verdict. In light of with the intention of cross breed suggestion framework improves the predication versatility and quality.

IV. PROPOSED METHODOLOGY

In this section, the proposed recommendation methodology is described step by step in the figure 1. Recommendation model for banking loan is carried out by various techniques using social network analysis.

A. *Data Collection*

Data collection is a procedure of grouping and calculating information in a specific variable for establishing organized way which were unordered, saved in a divergent format have inconsistency with a insufficient and null entry. A customer loan information of various banks is considered for analysis.

B. *Preprocessing*

Data preprocessing is a very important step in data mining process. Preprocessing step is important phase in data mining techniques and in recommendation system. Bank dataset generally have dimensionality and imbalanced distribution of data. Data preprocessing helps to discover knowledge in a database process which reduce complexity of a data and gives better conditions. Different tools and techniques is used for preprocessing such as sampling which selects the representative subset from a big population of data. Transformation defines manipulates the huge data to get a single input for denoising and also removes noise from a data normalization which organize data for additional efficient to use.

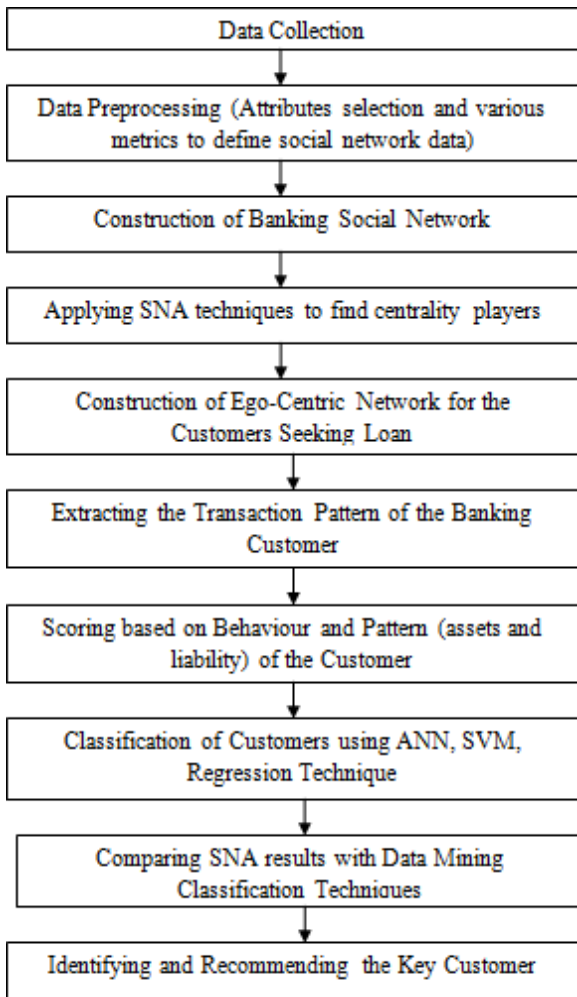


figure1: Methodology for the Proposed Work

C. Construction of Banking Social Network

As inescapable as social media life has become, there is a solid case for empowering this medium inside the banking association, too. A bank's inner customers are as significant a wellspring of criticism as outer ones, and setting up a framework that can group representative input and transform it into helpful bits of knowledge can have a large number of advantages.

Social media and network is itself dynamic with several networking sites which comes and go in a rapid speed. In past few year with a emergence of strong contenders like facebook and twitter as stabled. Compared to other consumer serving industries called retail and consumer goods, banking sector has more participants through social network commitment. Banks are essentially service focused, high customer connection and social network gives an opportunity which helps as their advantage.

D. Applying Social Network Analysis Techniques to find Centrality Players

Social network extraction permits the derivation of inferred connections and the choice of connections applicable for proposed examination methods. The examination of the subsequent informal community diagram depends on authoritative and singular investigation so as to allow a propelled customer to reach inferences with respect to the conduct of the association, in light of built up interpersonal organization chart measurements.

Long range social network locales are developing in ubiquity, and have quickly become an installation in mainstream society. Strikingly, the elements of criminal associations have for quite some time been researched by understanding the availability of key and key players in the association. On account of deducing the conduct elements of an association, just as every sensible situation, just incomplete perception of the informal organization is conceivable. Once the graph is created it is possible to perform analysis, organizational analysis. Individual level of analysis characterizes the members of organization in order to identify key customer. Many decisions based on known data are taken in bank. Extraction of transaction pattern helps the bankers to identify and authenticate a customer for loan.

E. Scoring based on the Customer Transaction

Scoring and Behavioral scoring are the procedures that assist associations with choosing whether or not to give credit to consumers who concern them Credit scoring these days depends on factual or operational exploration techniques. The measurable devices incorporate separate investigation which is basically direct relapse, a variation of this called logistic regression and classification trees, in some cases called recursive partitioning.

The Operational research methods incorporate variations of linear programming. Most scorecard developers utilize one of these strategies or a blend of the procedures Conduct scoring frameworks permit banks to settle on better choices in overseeing existing clients by estimating their future execution. The choices to be made incorporate what credit breaking point to dole out, regardless of whether to advertise new items to these specific customers, and if the record turns awful how to deal with the recuperation of the debit. The additional data in conduct scoring utilizes attributes of customer ongoing conduct to anticipate whether they are probably going to default. The technique is fundamentally the same as that of credit scoring.

An example of customers is picked so the information on their exchange execution either side of a self-assertively picked perception point is accessible.

The period before the perception time is known as the presentation or perception period and. The attributes that will be utilized in the conduct scorecard portray the customer's presentation during this time. Run of the mill factors would be normal, most extreme and least degrees of parity, credit turnover, and charge turnover.

Different attributes gauge the pattern in payments or balance during the period either by taking weighted midpoints or taking proportions of execution in the last piece of the period contrasted and that in the previous part. A portion of the qualities are markers of delinquent behavior – number of missed installments, times over overdraft or credit limit, while others reflect trouble in cash the executives, for example, the quantity of loans utilizing a Visa.

An pure behavior conduct scoring framework will just incorporate factors managing the customer presentation and the current estimations of factors from month to month credit agency reports. Other conduct frameworks incorporate individual qualities, for example, age, time with bank or private status too the unadulterated social attributes.

F. Classification of Customers using ANN, SVM and Regression Technique

Classification of the customers using data mining techniques are helps to classify the customer behaviour and their patters. The data mining techniques gives the key customer among the customer who are seeking for loan. Here the Artificial Neural Networks, Support Vector Machine and Regression techniques results are compared and accurate score Is taken for further comparison.

G. Comparing SNA results with Data Mining Classification Techniques

The result of classification technique is compare with social network analysis based on the customer behaviour and patter. After comparing both the results the highest scored customer (Key Customer) is identified and recommended for seeking loan. In this method social network analysis can identify and recommend the customer more effectively than classification techniques.

V. CONCLUSION

As of late a few suggestion frameworks have been proposed for prediction and recommendation for loan facility. The proposed model is developed for known your customer model to understand the key customer profile with an assortment and to view customer usage with various offers and benefits. Using ego centric network helps the bankers to easily visualize patterns with respect to loan recommendation. Using social network analysis it helps to develop multirelational social model with a score. The techniques which is used from

data mining is ANN, SVM and Regression tree be used for classification and comparison for better accuracy. For better result and accuracy the social network analysis results is compared with data mining classification results to identify and to recommend a key customer. In this proposed work the key customer is authenticated for loan approval and it helps to provide customized services for collection of customers.

REFERENCES

- [1] Mobasher, B., Cooley, R. and Srivastava, J. 2000a. Automatic personalization based on web usage mining. *Communications of the ACM*, 43(8), 142–151.
- [2] Shiva Nadi, Mohammad Hossein Saraee, Ayoub Bagheri. 2011. A Hybrid Recommender System for Dynamic Web Users, *International Journal Multimedia and Image Processing (IJMIP)*, 1(1).
- [3] Riecken, D. 2000. Personalized Views of Personalization. *Communications of the ACM* 43, (8) 27–28.
- [4] Claypool M., Gokhale. A., Miranda. T., Murnikov, P., Netes, D., and Sartin, M. 1999. Combining content-based and collaborative filters in an online newspaper. *ACM SIGIR '99 Workshop on Recommender Systems*, Berkely.
- [5] Sarwar. B., Karypis. G., Konstan. J., and Riedl. J. 2000. Analysis of recommendation algorithms for e-commerce. *Proceedings of ACM Ecommerce Conference*. 158–167.
- [6] Bamshad Mobasher, Honghua Dai, Tao Luo, Miki Nakagawa. 2002. Improving the Effectiveness of Collaborative Filtering on Anonymous Web Usage Data.
- [7] Yoon Ho Cho, Jae Kyeong Kim, Soung Hie Kim. 2002. A personalized recommender system based on web usage mining and decision tree induction. *Expert Systems with Applications* 23, 329–342.
- [8] Olfa Nasraoui and Chris Petenes. 2003. An Intelligent Web Recommendation Engine Based on Fuzzy Approximate Reasoning. *Proceedings of the IEEE International Conference on Fuzzy Systems - Special Track on Fuzzy Logic and the Internet*.
- [9] Magdalini Eirinaki, Charalampos Lampos, Stratos Paulakis, Michalis Vazirgiannis. 2004. Web Personalization Integrating Content Semantics and Navigational Patterns. *WIDM'04*, November 12-13. Washington, DC, USA. Copyright 2004 ACM 1-58113-978-0/04/0011.
- [10] Feng Hsu Wanga, Hsiu-Mei Shao. 2004. Effective personalized recommendation based on time-framed navigation clustering and association mining. *Expert Systems with Applications*. 27, 365–377.
- [11] Baoyao Zhou, Siu Cheung Hui and Kuiyu Chang. 2004. An Intelligent Recommender System using Sequential Web Access Patterns. *Cybernetics and Intelligent Systems*. *IEEE Conference on Cybernetics and Intelligent Systems*.
- [12] Sumathi, C., P., Padmaja Valli, R., and Santhanam, T. Automatic Recommendation of Web Pages in Web Usage Mining. *(IJCSE) International Journal on Computer Science and Engineering* 02(09),3046-3052.
- [13] Haibo Liu, Hongjie Xing, Fang Zhang. 2012. Web Personalized Recommendation Algorithm Incorporated with User Interest Change. *Journal of Computational Information Systems* 8(4), 1383-1390.
- [14] S. Jagannatha, M. Niranjanamurthy, and P. Dayananda, "Algorithm Approach: Modelling and Performance Analysis of Software System", *Journal of Computational and Theoretical Nanoscience (American Scientific publishers)*, December 2018, Volume 15, Issue 15, PP. 3389–3397
- [15] Florent Garcin, Christos Dimitrakakis, Boi Faltings, 2013. Personalized New Recommendation with Context Trees. *ACM Journal*.