Major Project Report On

Analysis of Customer Lifetime Value Modeling Techniques.

Submitted By:

Pawan Raj Gopal (2K19/BMBA/12)

Tushar Sharma (2K19/BMBA/20)

Under the Guidance of:

Dr. Kamal Gulati

Associate Professor

UNIVERSITY SCHOOL OF MANAGEMENT & ENTREPRENEURSHIP

Delhi Technological University

May 2021

CERTIFICATE

This is to certify that Pawan Raj Gopal (2K19/BMBA/12) and Tushar Sharma (2K19/BMBA/20) are bona fide students of University School of Management and Entrepreneurship, Delhi, and have successfully completed the project work as prescribed by the Delhi Technological University in the partial fulfillment of the requirement of Master Of Business Administration (MBA), Business Analytics Program for the academic year 2019-2021.

The Project Work titled “ Analysis of Customer Lifetime Value modeling techniques.”.

Project Guide

Dr. Kamal Gulati

Assistant Professor

DECLARATION

We, the undersigned, hereby declare that the project report entitled, “Analysis of Customer Lifetime Value modeling techniques.” submitted by us to the Delhi Technological University, in partial fulfillment of the requirement for the award of the degree of Master of Business Administration (MBA), Business Analytics under the guidance of Dr. Kamal Gulati, is our original work and the conclusions drawn therein are based on the material collected by ourselves.

The Report submitted is our own work and has not been duplicated from any other source. We shall be responsible for any unpleasant moment/situation.

Place: New Delhi

Date: 31st May, 2021

Pawan Raj Gopal (2K19/BMBA/12)

Tushar Sharma (2K19/BMBA/20)

ACKNOWLEDGEMENT

We are using this opportunity to express our gratitude to our family who supported us throughout the course of this project. We are thankful for their aspiring guidance, invaluably constructive criticism, and friendly advice during the project work. We are sincerely grateful to them for sharing their truthful and illuminating views on several issues related to the project.

We express our warm thanks to Dr. Kamal Gulati for his support and guidance during the project. Our thanks and appreciation go to our family and also our colleagues at USME, Delhi Technological University in developing the project and people who have willingly helped us out with their abilities.

ABSTRACT

Customer lifetime Value has become an important parameter in order to find and reach out to the customers who tend to contribute more heavily and frequently. This parameter therefore depends on the marketing industry or even we can say they are interdependent. It is important to know about any customer's purchase value and continuously monitor his transaction frequency and value for accurately determining CLV. One of the most important aspects or parameters that relates marketing and CLV is CAC (Customer Acquisition Cost) . It is definitely important to know how potentially a customer will contribute to the company and is it really worth spending money on an aspect to boost his purchases.

In this project, we aim to study the different CLV Modelling techniques to predict the customer lifetime value, analyse them and find the best suitable model for our dataset. We have analysed in detail about both the Probabilistic models and Machine Learning models, Our study includes the detailed analysis of the two most common and widely used probabilistic model i.e. BG/NBD and in Machine Learning, we used DNN to develop the model and to predict the future value of both existing customers with more transaction history and also for the new customers with very few purchases.

Keywords: Probabilistic Models, BG/NBD, CLV Modelling techniques **,** Machine Learning Models.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Industry Profile

Analytics is the process of identifying, interpreting, and visualising relevant patterns in data. India is becoming a major analytics centre. By 2020, the country's analytics market is expected to double in size, with big data accounting for nearly a quarter of that growth. According to the 'Analytics India Industry Study 2017' published by Analytics India magazine and AnalytixLabs, over 60% of analytics revenue in India originates from exports to the United States. Domestic revenue amounts for only 4% of total analytics revenue in the United States. The analytics sector's revenue is generated through a variety of industry segments. The finance and banking industry accounts for about 37% of the total analytics market, or $756 million in sales, making it the highest revenue-generating sector. With 26 percent, marketing and advertising come in second, followed by e-commerce with 15 percent.

In India, the Delhi/NCR area is the leading player in the analytics business. The capital region accounts for around 28% of overall revenue, or $565 million USD. Bengaluru, in the southern state of Karnataka, comes in second with roughly 27% of the vote. The rise of IT centres and the ever-increasing population of high net worth individuals (HNWIs) across the country are both contributing to this trend. Data analytics employees are generally paid a high salary because of their niche technical expertise. Data analysts in the e-commerce sector get paid about 1,3 million Indian rupees per annum, making it the most attractive area for budding analysts, while the analysts in the retail and FMCG sectors come a close second with about 1.2 million per annum.

Market Overview

The Market was valued at $2.13 billion in 2020, and is predicted to grow at a Compounded Annual Growth Rate of 14% to $4.68 billion by 2026. (2021 - 2026). As the struggle for client retention has become a need for firms, the benefits of analytics have become increasingly clear. Companies nowadays employ a variety of methods to keep customers informed and connected. They can use an analytics system to track the effects of their activities. As a consequence, these solutions have been integrated with an existing ERP system, which is useful in terms of using the created and available data.

A Brief History of Marketing

Most marketers are aware of the digital revolution's significant industry advancements, but what was marketing like before that? How did marketers throughout history use the resources and tactics at their disposal to deliver value and engage audiences? What's more, how did marketers generate metrics to aid in the optimization of their efforts prior to the digital revolution? Even thousands of years ago, humans had a natural tendency to brand things, as evidenced by mosaic advertising found in a wealthy businessman's home in Pompeii. As merchants travelled to frequently frequented "market towns" and understood they needed to distinguish out among merchants with identical items, commerce grew steadily over time. As international trade became more widespread, the demand for marketing services grew.

The printing press transformed our capacity to communicate in the 1400s, and from then on the strategy has been to explore innovative ways to connect with people with the priorities they need. The power of this new technology along with its ability to reach a huge number of people in a very short span of time did not go unnoticed. The first known print advertising was created in the 1450s, little than a decade after the printing press was invented. Following this, the usage of print increased dramatically, as did the sophistication of ads. The quantity of print advertisements increased dramatically throughout this period, when initially advertising began in Benjamin’s popular Magazine in 1740’s via the then introduced method of direct mail promotion in 1860’s. It didn't take long for radio advertising to overtake print advertising when it first entered the marketing landscape. Similarly, the debut of television commercials in 1942 increased the number of people who saw them.

Marketing Shifts to Customer

When such a time came when organisations started using mass communication to popularise their products, no sooner after they used this strategy they realised the importance of understanding the impacts of their advertising and how the media mix might be improved. When it comes to real life scenarios things are different, the University of Pennsylvania for the first time offered this course related to marketing, it was started in the early 1900’s named "The Marketing of Products,". When media first started to livestream events first, this was the first situation ever faced by marketing experts in the marketing industry. Because all the leading organisations were investing very large amounts on media advertising as they wanted to experiment with the trend.

The Introduction of Marketing Performance Measurement

Soon after broadcast television was introduced, the marketing environment was flooded with commercials from dozens of firms fighting for customers' attention and revenue. As a result, marketers had to devise techniques that went beyond the typical "spray and pray" approach to reaching as many people as possible. When this happened, in due course the industry for the very first time faced a tough situation which eventually led them to a new concept called "marketing mix." This term for the first time came into practice in the 1950's and eventually became familiarised to the industry in the 1960's for which Borden was an important reason. This concept of Marketing mix gave a brief on all the basic elements that are required to get an idea for different product strategies that deliver different aspects in each form of marketing medium which made a significant change to the approach of the experts as in to align product goals with their marketing strategy.

The Digital Revolution and a New Era of Marketing

The invention of the first wireless phone in 1972 ushered in a new era of rising technology. In 1975, the first personal computer joined the new digital landscape, following in its footsteps. Digital technology — and the advertising that followed to profit on those capabilities — proliferated from there. Everyone got their own personal computer soon after, and they were all connected via a new kind of communication: the contemporary internet. The first search engines appeared in the early 1990s to assist users in navigating the web, bringing with them early versions of search engine optimization and advertising. Smartphones and tablets have also enabled sophisticated on-the-go browsing and participation.

T*he Emergence and Evolution of Digital Attribution Models*

Digital technology has opened up new avenues for detailed marketing measurement. Prior to this, the predominant measuring approach for print and broadcast media was media mix modelling (MMM). It did, however, rely on long-term assessments and gave comprehensive insights into marketing initiatives' overall efficacy. Marketers wanted marketing analytics that could keep up with the speed of digital encounters. Consumers had access to a greater array of digital platforms as technology advanced, both at home and on the road. As a result, marketers found themselves interacting with customers across numerous channels, implying that marketing measurement efforts were necessary.

The shift towards unified marketing management

While digital attribution models are useful for assessing the digital marketing landscape, consumers continue to engage with traditional print and broadcast media. Today, the digital revolution has made it feasible to use successful marketing performance tools that go well beyond digital attribution in terms of marketing measurement.

Marketers can now correctly assess their efforts throughout the marketing mix as part of a unified marketing attribution approach that gives them a 360o perspective of their online and offline marketing activities. Marketing metrics, because of the digital revolution, can now assist marketers in providing fluid, personalised services.

Final Thoughts

The evolution of this industry has been seen as a significant change which had a great influence to the people as well as the marketing experts as the society has noticed the change from promotions through just messages to communicating through mass media with simultaneous reach of thousands of people which helped in reaching out to the customer more easily and also understand and prioritize their needs. The marketing mix model was established as part of this endeavour, and it remained a fundamental measurement that marketers could rely on until the digital invasion.

When digital technology initially came out, it drastically transformed how consumers and businesses interacted. Furthermore, as the number of channels increased, the importance of marketing metrics grew. Marketers may now acquire previously inconceivable insights about the convention's behaviour with the correct marketing insights and optimization software.

Customer Lifetime Value (CLTV)

As a marketer, it feels like it is important to keep track of a new measure very often. There are many parameters that need to be constantly or to say frequently monitored , beginning from moderate order value to the repetition of buying rate.If there were any relatively measurable metrics experts would have strongly believed that customer lifetime value will definitely become a significant metric in a few years to understand a customer more.

"The total amount of money a customer will spend with your firm during their time as a paying customer is referred to as customer lifetime value." CLTV informs you how much a client is worth to your company at a glance and provides a clear view of their whole worth by analysing each of their aspects. So, these analyses eventually give an idea of how much you should invest in customer retention. Not only that, but customer lifetime value indicates whether or not a customer is likely to return. If their customer lifetime value is high, they're likely to be brand loyalists who will buy more of your items in the future. If they don't, they're probably simply a one-time buyer who will require extra work to re-engage.

Calculation

Definitely, this only considers customers in terms of gross income and overlooks operating expenditures such as the cost of manufacturing your product, the cost of advertising, and the cost of running your firm. When estimating Customer Lifetime Value, these data must be considered.

Figure 1. Factors Involved in CLV Calculation

Source: https://survicate.com/

Why does CLV matter?

CLV allows you to categorise your clients into Profitable, Highly Profitable, and Not Profitable buckets based on their acquisition cost and income. You may now uncover the commonalities among many of your highly benefited clients and optimise the acquisition expenses for the highest possible price rather than the minimum cost option.

This enables you to concentrate your ad expenditure, email campaigns, and reward programmes. With the correct consumer group and proven message, you can get the best results.

Maximization of CLTV

As a result, the most important aspects that you can control are: Number of Repeat Sales, Average Transaction Value and Customer Retention Rate on Average. So, here are some strategies for boosting the customer lifetime value:

SURVEYS

Long-term connections can be fostered by developing effective techniques for increasing consumer loyalty. Customers are more likely to stick with brands they are familiar with because of their need and use to that particular brand.

Feedback survey of the Customer is an intriguing approach to engage with your clients and build loyalty. According to experts, website based surveys will definitely help us to get a view on a customer's aspect of reviewing a product or any service and also a clear idea of their intentions and expectation or even the changes to be made in any particular product. For example

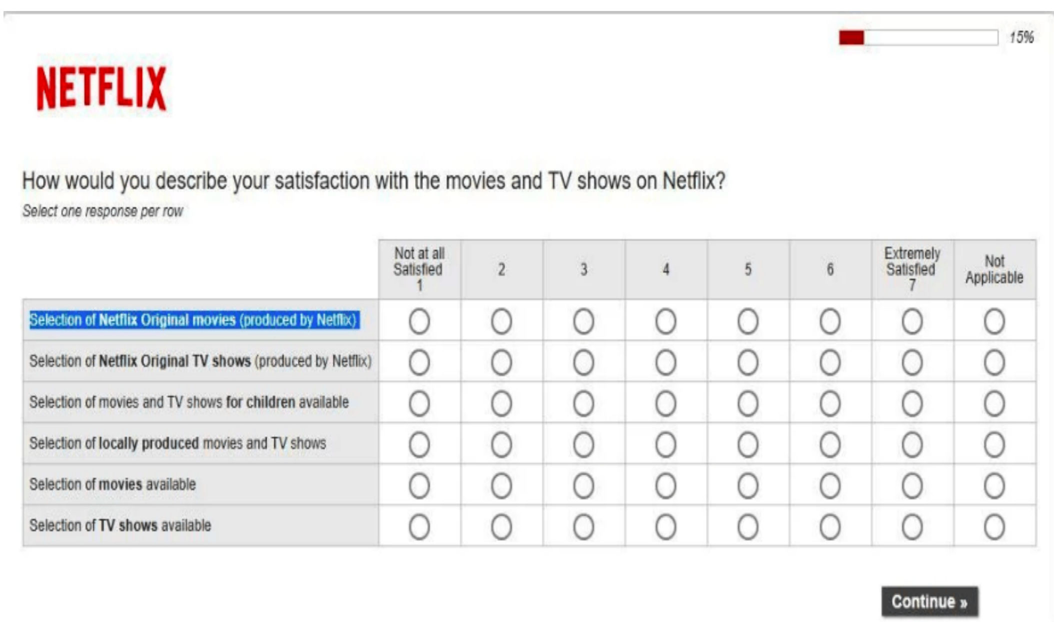

Figure 2. Sample of Netflix Survey

Source: https://survicate.com/

Staying Relevant

Another effective strategy to increase the amount of repeat transactions you receive is to stay at the top of your customers' minds. Clients will surely be considering your brand when there is a need for them if your brand is tagged as relevant and are topmost in the suggestion list.

Figure 3. Role of Influencers in Social Strategy

Source: https://survicate.com/

An influencer is someone who, among other things, will promote your goods on their own social media networks.

Cross-Selling and Up-Selling

Overloading the cart with irrelevant and fantasy products is a terrific method to increase sales while also increasing client loyalty. It also has a negative impact on your basket abandonment rate. Amazon has been particularly foresighted in this regard, providing superior shipping services, unique pricing, and access to a vast movie library through its Prime membership.Essentially, they're keeping their brand in front of their clients' minds by providing them with discounted additional things.

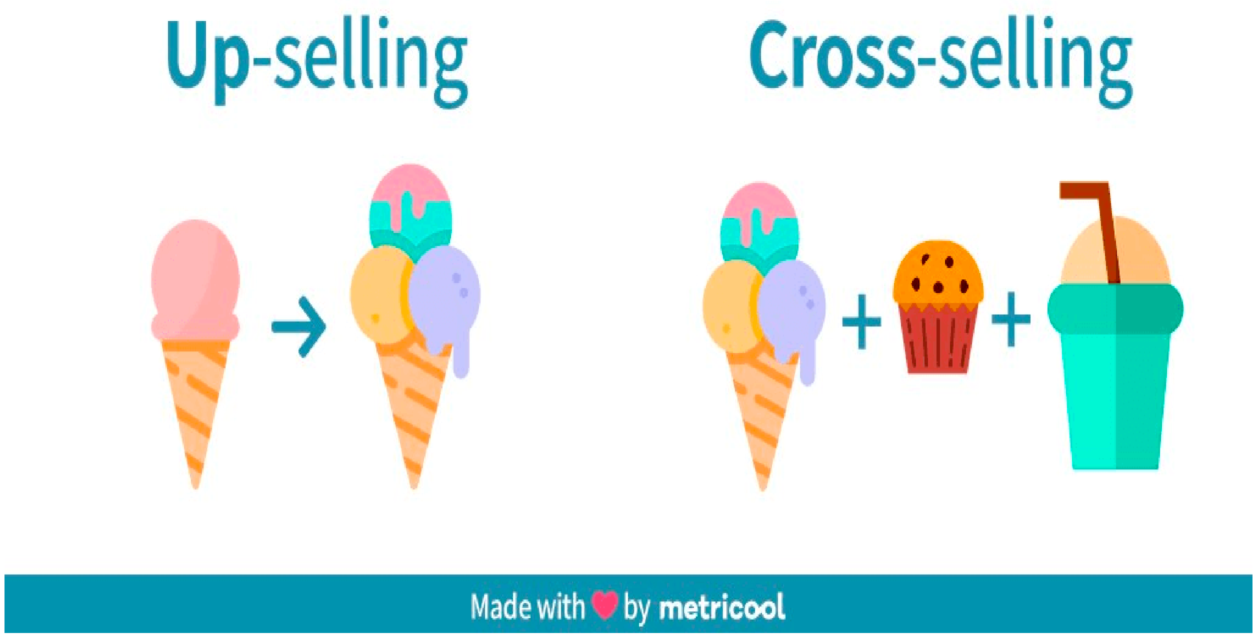

Figure 4. Example of Up-Selling and Cross-Selling

Source: https://survicate.com/

Boosting Customer Service

This aspect focuses on additional requirements of the customers who have already made sufficient purchases. The cornerstone to your company's success is customer happiness. These are frequently viewed as services that convert a one-time customer into a repeat consumer.

Loyalty Programs

Customers' lifetime values can be increased through loyalty programmes. This strategy compensates existing clients for staying in the company partnership for a longer period of time. It can also be used to figure out who your most valuable customers are.
Each of these strategies will help you increase the lifetime value of your customers while also increasing your long-term success rate. Perhaps outstanding buyer feedback or excellent customer service are the most beneficial aspects of your business. Some firms may get the greatest bang for their buck by implementing a successful loyalty programme. The final plan will be fully determined by what works best for you. Continue to test new things and track your progress until you find your ideal CLTV recipe.

1.2 Objective

The significance of CLV in marketing and advertising has been increasing in recent years and it has become a necessity for an organisation to have a record of their customers and to check their lifetime value to know about the business prospects. So, Our objective in this project is to study in detail about the various CLV Modelling techniques and to use the best suitable technique to predict the CLV for the dataset that we use.

1.3 Organisation of the report

This section includes the Glimpse of all the further sections that are included in this report. In Section 2 all the past work that has been done in this similar area has been discussed i.e. the algorithms used, various modelling techniques used in predicting the customer lifetime value and also different metrics useful for benchmarking the results from prediction. Section 3 is the Research Methodology which describes all the steps followed in completing this project, starting from data collection to analysing different modelling techniques and to find the suitable modelling technique for the dataset and finally predicting the customer lifetime value. Next comes Section 4 in which all the results or performance measures were summarised and analysed in detail. Then in Section 5 all the findings in these projects were summarised and then further possible work that could be done in this particular area is also mentioned i.e. possible algorithms that can be applied to increase the accuracy, better prediction models that could be developed etc.

Next in Section 6 we have written the limitations of the study i.e. the difficulties faced in implementing various steps in this Project, starting from collecting the data, developing a suitable model to predict CLV etc.

CHAPTER 2

LITERATURE REVIEW

Modelling Customer lifetime Value by Sunil Gupta et al.

This paper provides reviews about several CLV models which are useful for market segmentation and the allocation of marketing resources for acquisition, retention and cross-selling. There are two categories of models found. One category of models consists of those that attempt to find the impact of marketing programs on customer acquisition, retention and/or expansion (or cross-selling). The other category of models deals with the relationship between various components of CLV.

RFM model gives scores to the customers based on their recency, frequency and monetary value of their purchases. The customers are then segmented based on the scores allotted and are targeted with different marketing strategies. **Probability models** assume that the consumer's behaviour varies across the population based on some probability distribution. **Econometric models** combine customer acquisition, retention and expansion by modelling them. **Persistence models** are used to study the impact of advertising, discounting and product quality on consumer equity and to examine differences in CLV resulting from different customer acquisition methods. Many **computer science models** are being used like generalised additive models (GAM), multivariate adaptive regression lines (MARS), classification and regression trees (CART) and support vector machines (SVM) for their predictive ability. **Diffusion/Growth models** can be used to forecast the acquisition of future customers.

There are scopes of future work which are suggested by the authors like working on data which suggests customers' attitudes and their share of wallet, can focus on the portfolio of a customer rather than focusing on one customer's CLV, estimation of costs per customer. Understanding the limits CLV possesses, understanding the limits of theory-based models.

Customer Lifetime Value Measurement by Sharad Borle et al

Sharad Borle et al. used a hierarchical bayes approach to model a customer's lifetime value. They have modeled the purchase timing, purchase amount and risk of defection from the firm for each customer. They have taken the data from a membership-based direct marketing company where the number of times each customer joined and terminated the membership are known. This model is then compared with other models on a separate dataset. The other models include extended NBD-Pareto model, RFM model two models nested in their model, a heuristic model that takes the average customer lifetime, the average interpurchase time, and the average dollar purchase amount and uses them to predict the present value of future customer revenues at each purchase occasion. They have proved through the results that the hierarchical bayes approach model performs better than the other models in predicting CLV and in targeting valuable customers.

Customer-Base Analysis with Discrete-Time Transaction Data by Fader, Hardie and Berger

Fader et al. proposed a model to predict future purchase patterns of customers in discrete-time, which means transactions occur at some intervals. The model proposed was "beta-geometric/beta-binomial" (BG/BB) which act as a discrete time analog. Any customer purchase history in discrete time can be represented as a binary string where 1 represents a purchase and 0 represents not a purchase. Given this string, the model tries to find a probability that the customer is still and what are the expectations of future purchasing. They have applied this model to a dataset of cruise-line transactions for 6094 customers over a period of five years. They have observed that the customers who took a cruise in 1997 have the same probability of being alive in 1998. Customers who have taken a cruise in each of the last 4 years will have a higher probability than others. Similarly, there have been estimations made about all the categories of users based on their recency and frequency.

Customer Lifetime Value: Marketing models and applications by Paul D. Berger and Nada I. Nasr

Berger and Nasr have presented a series of mathematical models of customer lifetime value and managerial applications of these models. The mathematical models are presented in 5 different cases of different assumptions of the following: (1) The number of times sales take place in a year. (2) spending to retain customers and the change in customer retention rate. (3) difference in revenues received per customer. These models are useful to decide how much a company should spend on promotional campaigns and to check different profitability among different market segments. The CLV determination can help us find out the effect of a marketing strategy with its acquisition and retention rates. These models help to determine in various situations like the effect of price skimming strategy on acquisition rate. They can also help to decide how much to spend on the acquisition and how much to spend on retention of customers.

A model to determine customer lifetime value in retail banking context by Haenlein, M et al.

Haenlein, M et al. have proposed a model to find out CLV of retail banking customers which is based on a combination of first-order Markov Chain model and CART (classification and regression tree) analysis. They used profitability driver's age, demographics or lifestyle, type and intensity of product ownership and activity level as independent variables, and carried out age-dependent CART analyses to split customers of similar age into same sub-groups to find out the target variable contribution margin. These sub-groups were then used as states, among which customers are allowed to flow, of the first-order Markov model. The transition probabilities were estimated by transition frequencies. The CLV for each customer was determined as the discounted sum of state-dependent contribution margins which were weighted by their corresponding transition probabilities.

CHAPTER 3

RESEARCH METHODOLOGY

In this section we will be discussing methodology followed by us throughout this project. Section 3.1 discusses the problem statement of our project in detail. Section 3.2 describes the detailed sequence of steps followed , starting from analysing the past work done, finding suitable data to tabulation of results and evaluations.The section 3.3 includes the details of our experimental setup.

3.1 Problem Statement

We have seen from the introduction of this study, the importance of finding out the Customer Lifetime Value. In this study, we aim to find out the different modeling techniques which can be used for determining CLV. We also aim to implement 2 of those techniques and to have a comparative study between them.

3.2 Proposed Framework

As discussed above this section we have discussed the sequence of steps followed in completing our entire project. The below diagram just represents the framework of our entire project.

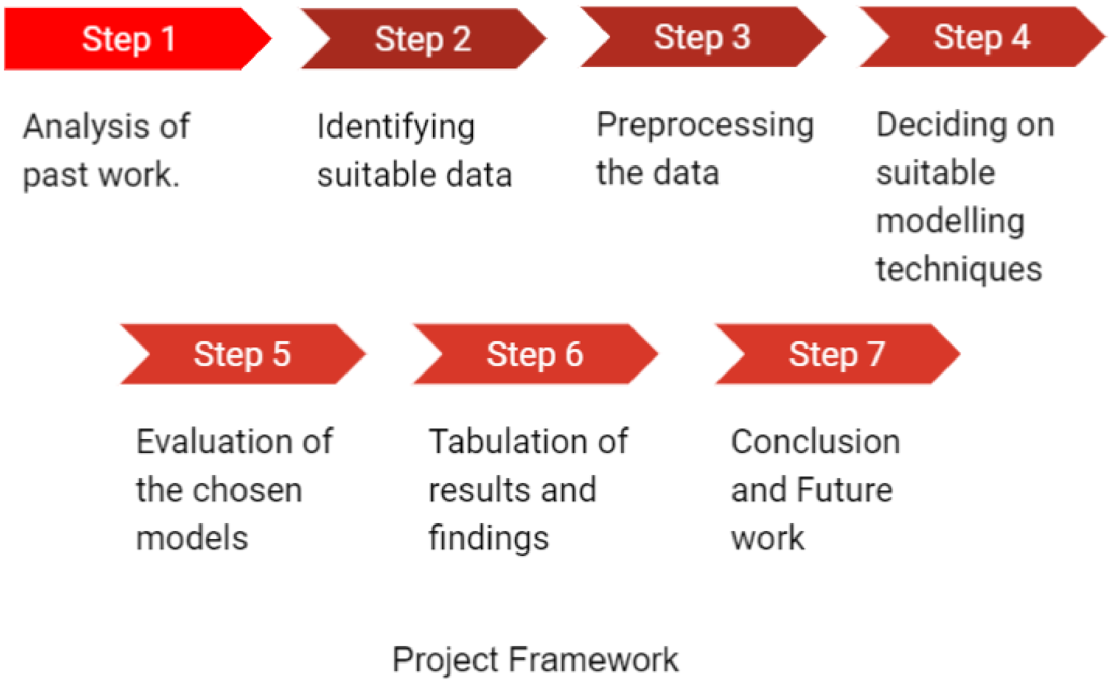

Figure 5.

Step 1 : Analysing the past work done.

The first step for us in this project was to have a clear idea on what are the CLV modelling techniques that have already been in use by most researchers and how efficient or successful were the researchers in achieving the the predictions more accurately so that we can narrow down our objective and decide on what is to be done in our project.

Step 2 : Identifying the suitable dataset.

After analysing the past work it was important for us to decide on the dataset with all the requirements which is suitable for applying the CLV modelling techniques. There are certain attributes which are very significant in predicting the CLV and this was the primary constraint for us in selecting the suitable dataset.

Step 3 : Preprocessing of the dataset.

It is very important to apply the necessary preprocessing techniques, because datasets might have missing values, outliers and duplicate entries which might definitely affect the results of our predictions. We have also applied the appropriate preprocessing techniques before actually using it for Prediction.

Step 4 : Deciding on suitable Modelling techniques.

Once the dataset was preprocessed we needed to decide on the suitable modelling techniques that would give the best result for this particular dataset. After a detailed study and analysis we decided to apply BG which is a probabilistic modelling technique and DNN which is ML modelling technique.

Step 5 : Evaluation of the chosen models.

The chosen models were then applied on this dataset to find the customer lifetime value and also the models were evaluated on the basis of various possible metrics.

Step 6 : Tabulation of results and findings.

The results for these two models were tabulated and interpreted and were analysed in the aspect of successfulness in prediction of CLV to a greater extent and with good accuracy.

Step 7 : Conclusion and Future work.

We have finally concluded our work by mentioning the findings and interpretation of our results and also have discussed the possible work that researchers in this field can work upon and have scope.

3.3 Experimental setup

3.3.1 Dataset:

The dataset that we have used for our project is of the type transactional data which contains the information of all transactions processed in an UK-based and registered online retail store for a duration of nearly one year i.e. from December 1st 2010 to December 9th 2011. The dataset consists of the attributes like Invoice Number, Product code, Product Description (name) , Quantity of Purchase, Invoice date, Unit price, Customer ID and finally Country Name.

Below is the view of the sample dataset that has been described above .

	A	B	C	D	E	F	G	H
1	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerI	Country
2	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	01/12/2010 08:26	2.55	17850	United Kingdom
3	536365	71053	WHITE METAL LANTERN	6	01/12/2010 08:26	3.39	17850	United Kingdom
4	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	01/12/2010 08:26	2.75	17850	United Kingdom
5	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	01/12/2010 08:26	3.39	17850	United Kingdom
6	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	01/12/2010 08:26	3.39	17850	United Kingdom
7	536365	22752	SET 7 BABUSHKA NESTING BOXES	2	01/12/2010 08:26	7.65	17850	United Kingdom
8	536365	21730	GLASS STAR FROSTED T-LIGHT HOLDER	6	01/12/2010 08:26	4.25	17850	United Kingdom
9	536366	22633	HAND WARMER UNION JACK	6	01/12/2010 08:28	1.85	17850	United Kingdom
10	536366	22632	HAND WARMER RED POLKA DOT	6	01/12/2010 08:28	1.85	17850	United Kingdom

Figure 6. Sample of the Dataset

The Time range of transactions is: 2010-01-12 to 2011-12-10

Total number of unique customers: 4338

Total Quantity Sold: 5167812

Total Sales for the period: 8911407.904

3.3.2 Data Preprocessing

In this section we have mentioned various preprocessing techniques applied on our dataset before proceeding for model evaluation.

Data Cleaning

This dataset had the transaction details of various customers over a period of one year which in itself had many irrelevant symbols, values and text which needed to be replaced with blank spaces as they are in no way relatable to the intended objective. The dataset was highly noisy.

Removal of Duplicate Entries

This transactional dataset also consisted of certain duplicate rows and the repetition of records might produce biased results. It is always preferable that a model learns from a single unique record, so we decided to remove all duplicate data before proceeding to the next step.

Missing Values in Dataset

Whenever there is a dataset with more missing values , it is of no use in applying modelling techniques as the results tend to be more biased. Even this dataset had many missing values which was then filled with the most appropriate method for that particular attribute.

Feature Extraction

In this dataset there were some attributes that were not required for predicting CLV. So, feature extraction was an important part of our preprocessing and finally our dataset consisted of customer ID, Invoice number , Date, Quantity and Unit price. All the other attributes were extracted.

3.3.3 About the models

We have studied different models from various research papers and have listed them out here. The fundamental concept of calculating customer lifetime value remains the same. The CLV is defined as the total profit which we expect from a customer over its relationship with the firm. The formula is simple:

$$CLV = \frac{(Average\ Sales \times Purchase\ Frequency) \times Profit\ Margin}{Churn\ rate}$$

The different models studied are:

Econometric models:

These types of models share the similar underlying principles as probabilistic models. These models combine the customer acquisition, customer retention and expansion (or cross-selling) by individually modeling them. (Sunil Gupta et al.)

Persistence models:

These models are used to study the impact of advertising, discounting and product quality on consumer equity and to examine differences in CLV resulting from different customer acquisition methods. The persistence modeling consists of three different steps (Sunil Gupta et al.) :

1. The first step is to examine the evolution of each variable of the system, over time. This step identifies temporary and permanent movements in that variable. For instance, determining how the firm's retention rates are behaving overtime.
2. The next step is to estimate the vector-autoregressive model (VAR) with the least square methods.
3. The last step is to derive the impulse response functions. The estimated parameters of the VAR models are not always directly interpreted. They are used

to obtain the estimates of short and long-run impact of a single shock in a variable on the system.

Diffusion/Growth models:

These models are used to predict the number of customers a firm will be acquiring in the future (Sunil Gupta et al.). Gupta, Lehmann, and Stuart (2004) have suggested a model for forecasting the number of new customers at time t:

$$nt = \frac{alpha*gamma*exp(-beta-gamma*t)}{[1+exp(-beta-gamma*t)\ ]2}$$

Where alpha, beta, gamma are the parameters of the growth curve of the customers.

Computer science models or machine learning models:

There are many machine learning models which are used for their predictive ability in determining CLV. Some of them are generalised additive models (GAM), multivariate adaptive regression lines (MARS), classification and regression trees (CART), support vector machines (SVM), deep neural networks (DNN) and other regression models.

For this study we have implemented DNN as one of our models. Deep neural networks are basically neural networks with multiple hidden layers. In our work, we have used 3 hidden layers. The machine learning models have two major phases, training the model and testing the model. To implement DNN in python we have used Keras API to Tensorflow. The benefits of using a machine learning model over a probabilistic model is that in machine learning models we can use multiple parameters which we consider important. Along with recency, frequency and monetary value we have considered average basket value, average basket size as the features to the model.

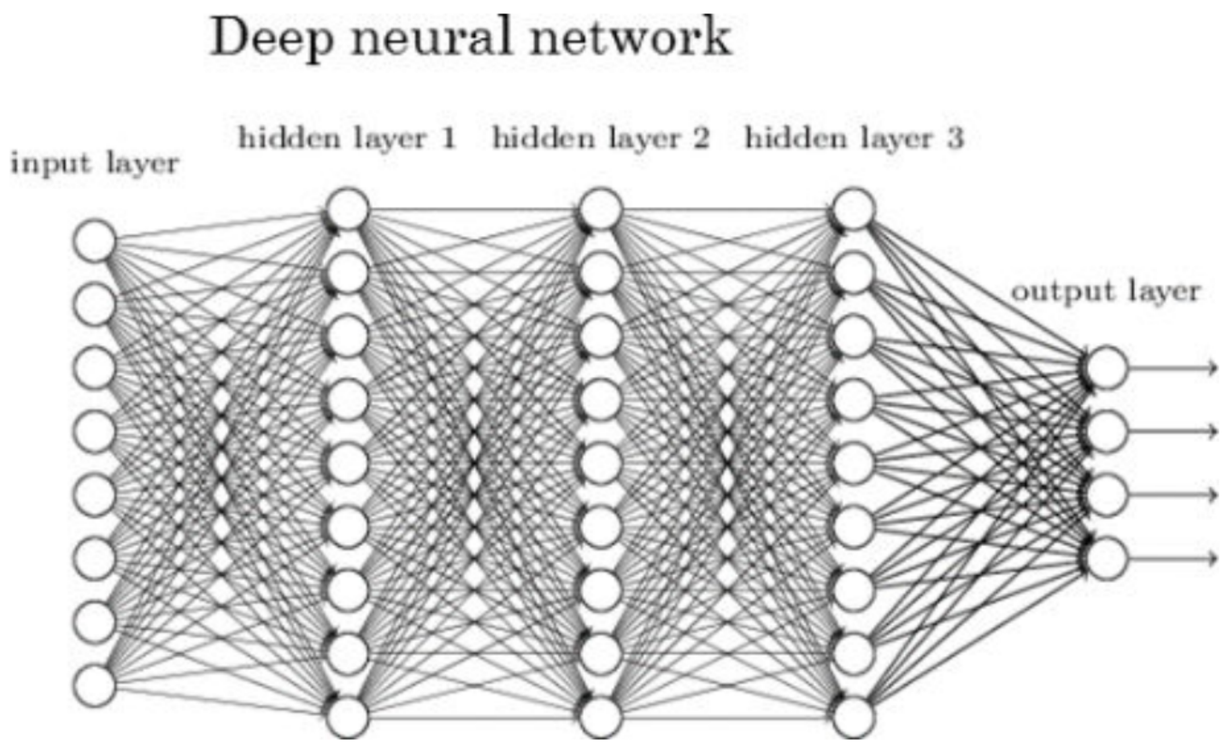

Figure7. Deep Neural Network. Source:
https://www.kdnuggets.com/2020/02/deep-neural-networks.html

Modeling through historical approach:

These models consider the past data to determine CLV. There are two major models in this category- Aggregate model and Cohort model. Aggregate model calculates the CLV with the help of average revenue per customer in past data. Cohort models group the data into cohorts based on features like transaction date and then it determines average revenue per cohort. This is used to determine CLV of each cohort.

Probabilistic models.

These models fit a probability distribution to the data given and estimate the number of transactions a customer will make in the future and the amount of money they are likely to spend. There are various probabilistic models to determine CLV. We have to take this into account that not all the variables are determined from a single model. The transaction variables like the purchase frequency and churn rate are determined separately from monetary variables like the amount the customer will spend. We have shown this through the Figure . To determine future transactions we have two popular models - **Pareto/NBD** (Pareto Negative Binomial Distribution) model and **BG/NBD** (Beta Geometric Negative Binomial Distribution) model. The steps for using both the models are the same. In this study, we have used the BG/NBD model as one of our models for the comparative study because we have to take into consideration the discrete time analogue. We have combined **Gamma-Gamma** with the BG/NBD model to determine the monetary benefit a customer will likely give. These models are implemented in python using the **lifetimes** library.

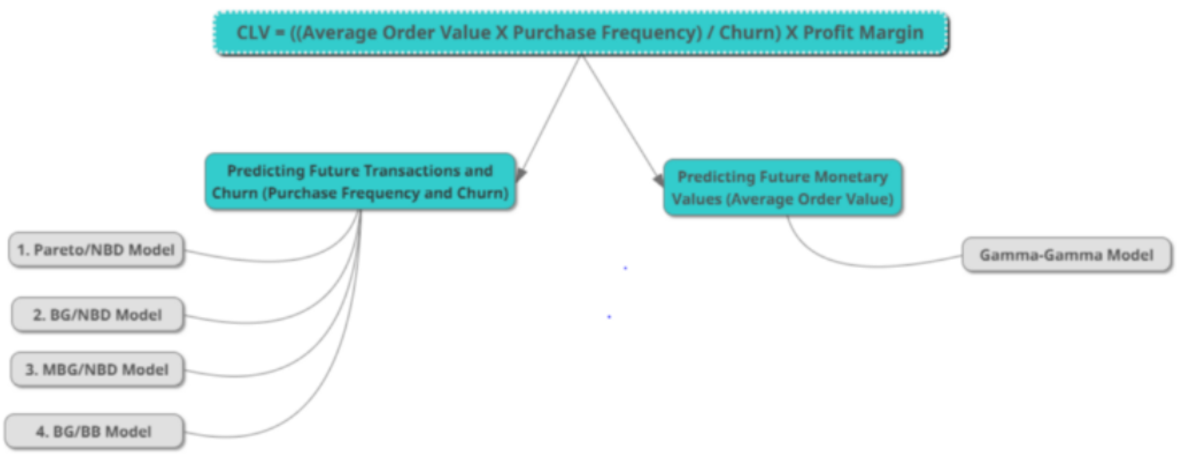

Figure 8. Types of Probabilistic models

Source: https://www.analyticsvidhya.com/

3.3.4 Evaluation Metric

In this study, we have used one evaluation metric to compare the models -Root Mean Square Error (RMSE). It is determined by taking the square root of the mean of the square of all the errors.

CHAPTER 4

RESULTS

For this study, we have implemented two models - one probabilistic model and one machine learning model. We have used BG/NBD along with Gamma-Gamma model for the probabilistic model and DNN for the machine learning model.

For implementing the probabilistic models we have used the lifetimes library of python. For that, we determined recency, frequency, T and monetary value to feed in the model. Table 1 shows the recency, frequency, T and monetary value calculated.

Table 1. Recency, Frequency, T and Monetary Value

	Customer ID	Frequency	Recency	T	Monetary Value
0	12346.0	0.0	0.0	326.0	0.000000
1	12347.0	6.0	476.0	516.0	599.701667
2	12348.0	3.0	283.0	359.0	359.0
3	12349.0	0.0	0.0	19.0	0.000000
4	12350.0	0.0	0.0	311.0	0.000000

Frequency is the number of times a customer makes purchases. **Recency** here is the time difference between the first and the last transaction. **T,** given in the table is the time difference between a customer's first purchase to the end of the transaction period. **Monetary Value** is the average value of sales of a customer.

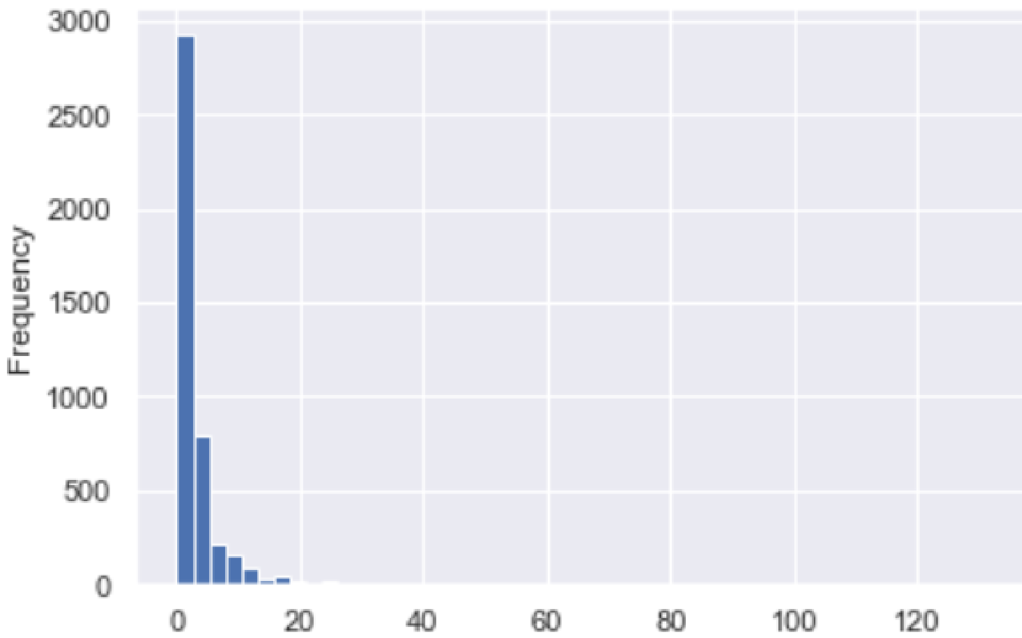

Figure 9. Plot showing the frequency of each purchase. X-axis is the frequency and the graph shows the number of customers.

Table 2. shows the model summary of the BG/NBD model. It shows the estimated distribution parameter values.

Table 2. Summary of BG/NBD model in python

	coef	**se(coef)**	**lower 95% bound**	**upper 95% bound**
r	0.949337	0.03203509	0.886547849	1.012125401
alpha	96.88384	3.660641212	89.7089882	104.0587018
a	6.42E-07	0.003256946	-0.006382971	0.006384256
b	10354.74	48676056.15	-95394715.32	95415424.81

The table 3. shows the probability a customer is alive given its recency, frequency and T.

Table 3. Probability that a customer is alive

	CustomerID	frequency	recency	T	monetary_value	probability_ alive
0	12346	0	0	326	0	1
1	12347	6	476	516	599.7017	1
2	12348	3	283	359	301.48	1
3	12349	0	0	19	0	1
4	12350	0	0	311	0	1
5	12352	6	268	341	253.565	1
6	12353	0	0	205	0	1
7	12354	0	0	233	0	1
8	12355	0	0	96	0	1
9	12356	2	303	326	269.905	1

Now that the probability of being alive is dependent on frequency and recency of a customer we can derive the following by looking at the figure 8. :

If the customer has bought many times but the customer is not recent, the probability of being alive is less. However, if a customer is less frequent but he recently purchased something (recency is high) then the probability of being alive is high.

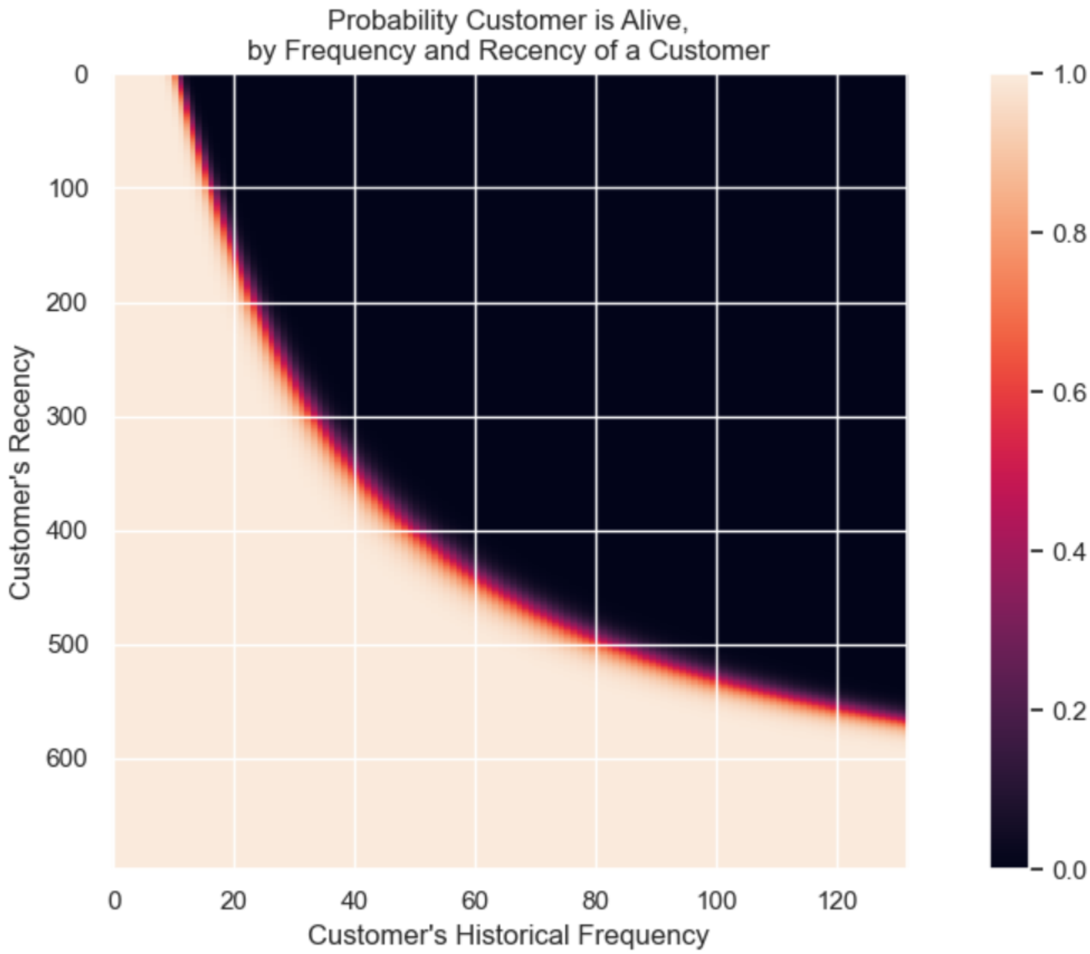

Figure 10. Plot showing the probability that a customer is alive

After calculating the probability, we have shown the likely future transactions of each customer. Table 4 shows the predicted number of transactions in the next 10 days. To verify if the predicted number of transactions in table 4 is correct or not we have performed a calculation manually for a customer's future transaction.

Let us check for customer ID - 12748, In the given 372 days, he has purchased items 112 times. Therefore, he will purchase 112/372 = 0.301 times in one day. And,in 10 days he will purchase 3.01 times.

Our predicted result came out to be 4.27. This result is actually close to the manually calculated value. The reason that there is a difference is because of the various assumptions made about the customers, e.g., dropout rate, customers lifetime distribution, etc.

Table 4. Predicted number of transactions in next 10 days

	index	Customer ID	frequency	recency	T	Monetary value	Probability of alive	Predicted number of txn
0	1879	14911	131	697	697	1093.662	1	4.99
1	326	12748	112	692	697	301.0248	1	4.27
2	4010	17841	111	697	697	364.4522	1	4.23
3	2176	15311	89	697	697	677.7294	1	3.4
4	1661	14606	88	692	697	135.8901	1	3.36
5	1689	14646	44	353	355	6366.706	1	2.98
6	562	13089	65	571	577	893.7143	1	2.94
7	481	12971	70	657	666	159.2113	1	2.79
8	2083	15189	37	329	335	429.8768	1	2.64
9	3992	17811	37	340	341	208.5197	1	2.6

Now, we have predicted the expected average sales for each customer transaction and Customer Lifetime Value using the Gamma Gamma model. This model takes care of the monetary aspect of the problem.

Table 5. Average expected sales

Cust ID	frequency	recency	T	Monetary value	Probability alive	pred_num_txn	exp_avg_sales
12347	6	476	516	599.7017	1	0.34	602.9928
12348	3	283	359	301.48	1	0.26	306.7138
12352	6	268	341	253.565	1	0.48	256.0666
12356	2	303	326	269.905	1	0.21	277.5566

12358	1	117	120	484.86	1	0.27	503.2447

Table 6. Predicted Customer Lifetime Value of sales of customers

CustID	frequency	recency	T	Monetary value	Probability alive	pred_num_txn	Exp_avg sales	Predicted CLV
12347	6	476	516	599.7017	1	0.34	602.9928	203.0847
12348	3	283	359	301.48	1	0.26	306.7138	78.92291
12352	6	268	341	253.565	1	0.48	256.0666	120.7083
12356	2	303	326	269.905	1	0.21	277.5566	57.49826
12358	1	117	120	484.86	1	0.27	503.2447	134.3503

Using the model we have obtained the Predicted CLV as shown in Table 6. Now, let us check and compare it with the manually predicted CLV. We can see that both the values are actually very close in Table 7. We have calculated the CLV for the next 30 days.

Table 7. Predicted CLV versus Manually predicted CLV

Cust ID	probability_alive	pred_num_txn	exp_avg_sales	predicted_clv	manual_predicted_clv
12347	1	0.34	602.9928	203.0847	205.0175
12348	1	0.26	306.7138	78.92291	79.74558
12352	1	0.48	256.0666	120.7083	122.912
12356	1	0.21	277.5566	57.49826	58.28689
12358	1	0.27	503.2447	134.3503	135.8761

Now, we need to observe here that the CLV calculated above is based on the total sales and not the actual profit per customer. To obtain the net profit per customer, we multiply the predicted CLV with the net profit margin of 5%.

Table 8. Predicted CLV (net profit margin) of each customer

Cust ID	probability_alive	pred_num_txn	exp_avg_sales	predicted_clv	manual_predicted_clv	CLV
12347	1	0.34	602.9928	203.0847	205.0175	**10.15423**
12348	1	0.26	306.7138	78.92291	79.74558	**3.946146**
12352	1	0.48	256.0666	120.7083	122.912	**6.035417**
12356	1	0.21	277.5566	57.49826	58.28689	**2.874913**
12358	1	0.27	503.2447	134.3503	135.8761	**6.717516**

In Table 8 we have predicted the final CLVs of customers for the next 30 days. Now, looking at these results, the marketing team will know which customer to target to increase their sales and hence profit.

Implementing Machine Learning Model : DNN

Implementing the DNN model, R-square comes out to be 0.7 which seems good enough to say that the model will make meaningful predictions. Our DNN model underpredicts the total number of sales however this is because of the large outliers that the model is unable to predict. The error can be reduced by dealing with the outliers. We can see the scatter plot between the actual CLVs and the predicted CLVs. in Figure 9. We can observe that for greater values of CLV, predicted values are also greater.

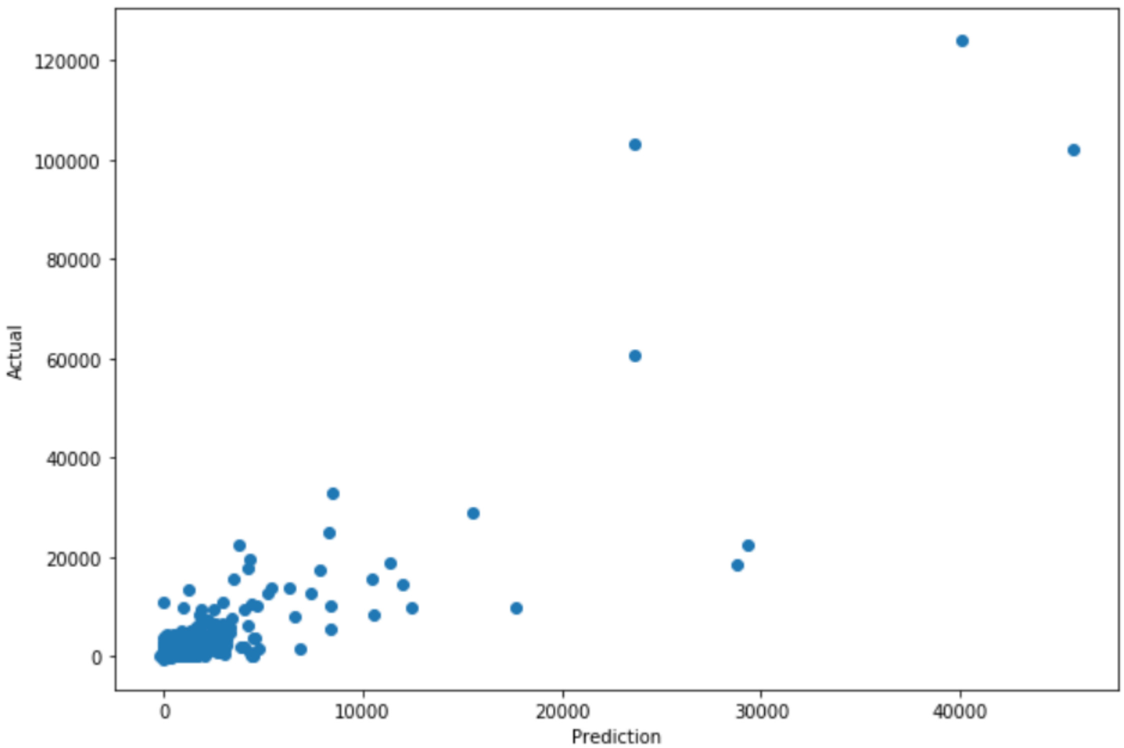

Figure 11. Scatter plot of actual vs. predicted CLV

It was necessary to compare the two models implemented. One is the probabilistic model BG/NBD+Gamma Gamma and the other is the machine learning model DNN. In Figure we can see the probability density function of actual CLV versus DNN predicted CLV. We can see that it is very fairly predicting the lifetime value of the customers. In Figure we have shown. In Figure . we have shown the probability function of actual versus predicted CLV from the BG/NBD+Gamma Gamma model. We can observe that both the models are good measures of revenue as both of them show the long skewed tail. However it is clear from the figures that the DNN model fits better than the BG/NBD+Gamma Gamma model because the DNN model does not have the second spike as we can observe in the other models.

We also made comparisons in the RMSE value of both the models which we can see in the Table . Although we can see the difference in the RMSE values of the models, we can fairly say that both the models perform similarly.

Table 9. RMSE of the models

Model	RMSE
BG/NBD	1595
DNN	836

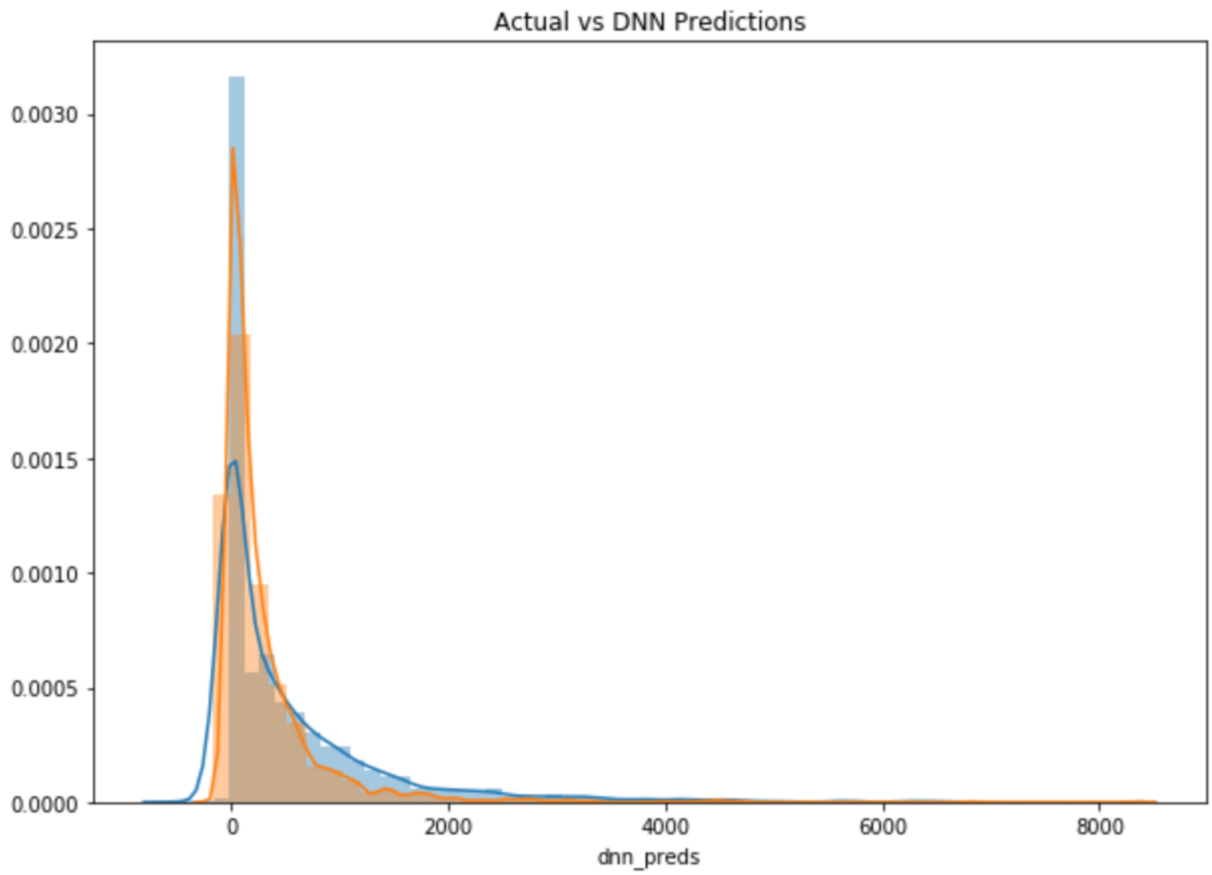

Figure 12.

Probability density representation of Actual VS. DNN Predictions

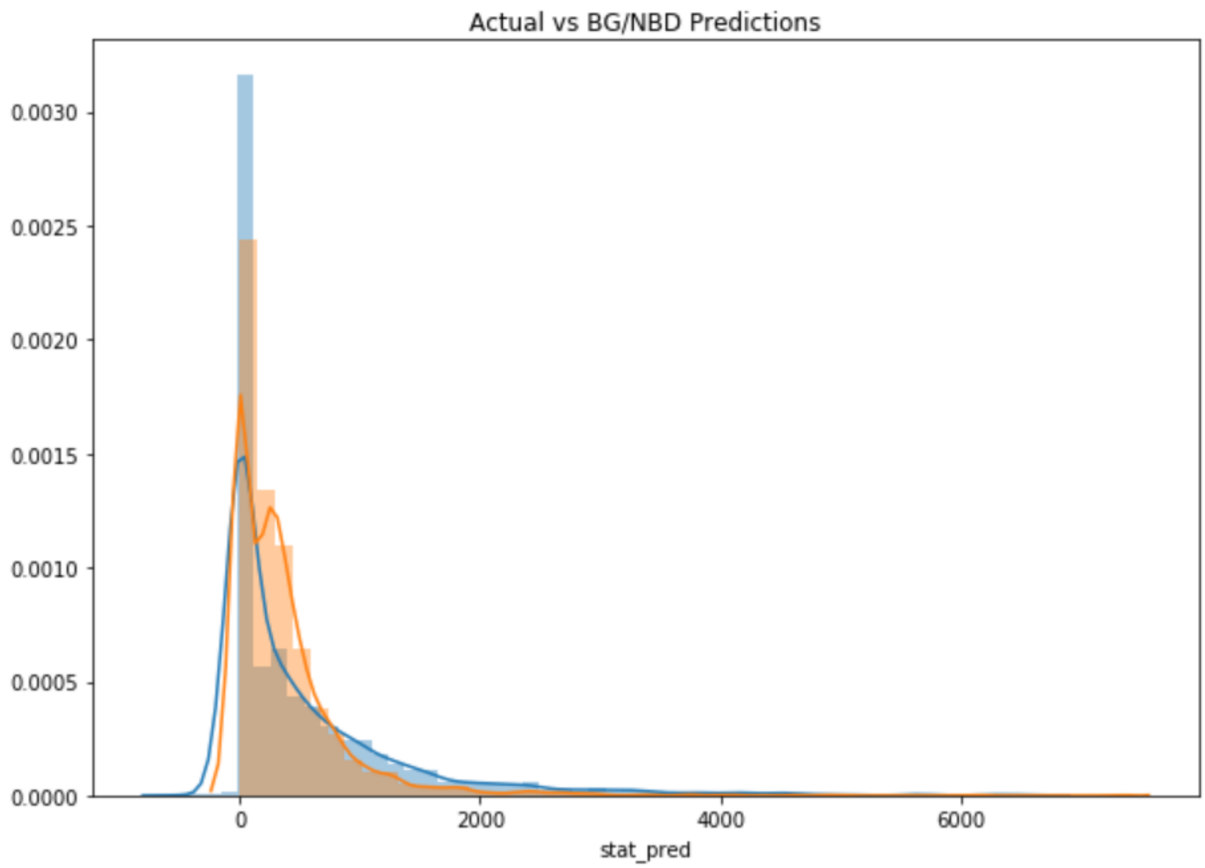

Figure 13.

Probability density representation of Actual VS. BG/NBD predictions of CLV

CHAPTER 5

LIMITATIONS OF THE STUDY

In this study we have targeted each customer to find out their CLV and help the marketing team to focus on each customer. However, in reality it is hard to target each customer individually. Customers should be targeted in groups which show similarity.

To form the groups we need to have access to the demographics of the customers. If we have the demographics data we can create the customer segments and predict the CLVs of each of the segments. This information about the segment can then be used for personalised targeting. This study could have also formed segments with the RFM of the customers when there is unavailability to demographics data of customers.

To use the DNN model effectively we could have accessed a bigger dataset as it will perform better on that.

CHAPTER 6

RECOMMENDATIONS

In this project, we have worked on finding the modeling techniques of customer lifetime value and we have implemented two of those techniques. For future work, we can explore and implement more modeling techniques to arrive at a better result. We should also use multiple evaluation metrics to have a better understanding of which model performs better. We can also work on implementing the different models for different situations and assumptions. Different models will perform better with different datasets also.

For future work, we also suggest that along with the models, the proper usage and application of the model should be listed so that it can be used in real time. Further, we can also determine the costs to be spent on various marketing activities to keep in check the costs with the maximum profit.

CHAPTER 7

REFERENCES

Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., ... & Sriram, S. (2006). Modeling customer lifetime value. *Journal of service research*, *9*(2), 139-155.

Borle, S., Singh, S. S., & Jain, D. C. (2008). Customer lifetime value measurement. *Management science*, *54*(1), 100-112.

Fader, P., Hardie, B., & Berger, P. D. (2004). Customer-base analysis with discrete-time transaction data. *Available at SSRN 596801*.

Berger, P. D., & Nasr, N. I. (1998). Customer lifetime value: Marketing models and applications. *Journal of interactive marketing*, *12*(1), 17-30.

Haenlein, M., Kaplan, A. M., & Beeser, A. J. (2007). A model to determine customer lifetime value in a retail banking context. *European Management Journal*, *25*(3), 221-234.

Gupta, Sunil and Donald R. Lehmann and Jennifer Ames Stuart (2004), “Valuing Customers,” Journal of Marketing Research, 41 (1), 7-18.

www.kdnuggets.com/2020/02/deep-neural-networks.html

https://www.statista.com/topics/4677/analytics-market-in-india/#dossierSummary

https://www.dataiku.com/learn/samples/customer-lifetime-value/

https://github.com/

http://archive.ics.uci.edu/ml/datasets/Online+Retail

https://www.kaggle.com/

https://stackoverflow.com/

Dekimpe, Marnik and Dominique M. Hanssens (2004), “Persistence Modeling for Assessing Marketing Strategy Performance,” in Assessing Marketing Strategy Performance, C. Moorman and D. Lehmann, eds. Cambridge, MA: Marketing Science Institute.

Cooper, Lee G. and Giovanni Giuffrida (2000), “Turning Data Mining into a Management Science Tool: New Algorithms and Empirical Results,” Management Science, 46 (2), 249-64.

TY - JOUR AU - Jain, Dipak AU - Singh, Siddharth PY - 2002/03/01 SP - 34 EP - 46 T1 - Customer lifetime value research in marketing: A review and future directions VL - 16 DO - 10.1002/dir.10032 JO - Journal of Interactive Marketing ER -

Wu, S. I., Li, P. C.: The relationships between CRM, RQ, and CLV based on different hotel preferences. International Journal of Hospitality Management, 30(2), 262-271 (2011)

CHAPTER 8

ANNEXURE - PLAGIARISM REPORT

turnitin

Major Project Report- Pawan(12)_Tushar(20).pdf
May 31, 2021
8044 words / 41028 characters

Major Project Report- Pawan(12)_Tushar(20).pdf

Sources Overview

12%
OVERALL SIMILARITY

1	www.statista.com INTERNET	2%
2	docplayer.net INTERNET	1%
3	www.researchgate.net INTERNET	1%
4	pdfs.semanticscholar.org INTERNET	1%
5	Michael Haenlein, Andreas M. Kaplan, Anemone J. Beeser. "A Model to Determine Customer Lifetime Value in a Retail Banking Context... CROSSREF	<1%
6	National Institute of Technology, Kurukshetra on 2014-11-08 SUBMITTED WORKS	<1%
7	Symbiosis International University on 2017-06-13 SUBMITTED WORKS	<1%
8	citeseerx.ist.psu.edu INTERNET	<1%
9	Bocconi University on 2015-04-19 SUBMITTED WORKS	<1%
10	lifetimes.readthedocs.io INTERNET	<1%
11	www.michaelhaenlein.eu INTERNET	<1%
12	Indian Institute of Management, Bangalore on 2021-03-07 SUBMITTED WORKS	<1%
13	University of Newcastle upon Tyne on 2021-05-20 SUBMITTED WORKS	<1%
14	hdl.handle.net INTERNET	<1%
15	Informatics Education Limited on 2014-01-26 SUBMITTED WORKS	<1%
16	mansci.journal.informs.org INTERNET	<1%
17	www.indiaclass.com INTERNET	<1%

18	Heriot-Watt University on 2020-08-02 SUBMITTED WORKS	<1%
19	University of Hertfordshire on 2020-04-01 SUBMITTED WORKS	<1%
20	Indian School of Business on 2018-02-21 SUBMITTED WORKS	<1%
21	docshare04.docshare.tips INTERNET	<1%
22	repository-tnmgrmu.ac.in INTERNET	<1%
23	nccur.lib.nccu.edu.tw INTERNET	<1%

Excluded search repositories:

- None

Excluded from Similarity Report:

- Bibliography
- Quotes
- Small Matches (less than 10 words).

Excluded sources:

- None