REVA
UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

REVA Academy for Corporate Excellence (RACE)

Minimizing Losses on Trials of "Strategy Builder" Tool using Business Analytics

Tushar Nigam
SRN: R18MBA57

Date: 27/08/2022

MBA in Business Analytics

Capstone Project Presentation
Year: II

race.reva.edu.in

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

Agenda

Introduction

Studies mention that 95% of new products fail. Not enough exposure to the customers could be one of the major reasons for this as they are not familiar with the capability and potential of the product. Therefore, a **Trial** becomes very important for customers to evaluate the product and make purchase decisions.

Considering this, **RLP Software Pvt. Ltd***, a software development company, after launching its new product **Strategy Builder*,** extended the tool's trial version to its existing customers. But the problem here is the cost associated with every trial. The resources which contribute to the cost are Dev-Ops, Training, Cloud Storage and Computation, Licensing, Technical Support, etc. If the customer does not purchase the product after the trial, this amount gets wasted.

The **conversion rate** after the trial is just **44%,** leading to a loss of **$8.5 M** on unsuccessful trials. Hence, the company is looking for ways to minimize these losses.

* Names changed for confidentiality

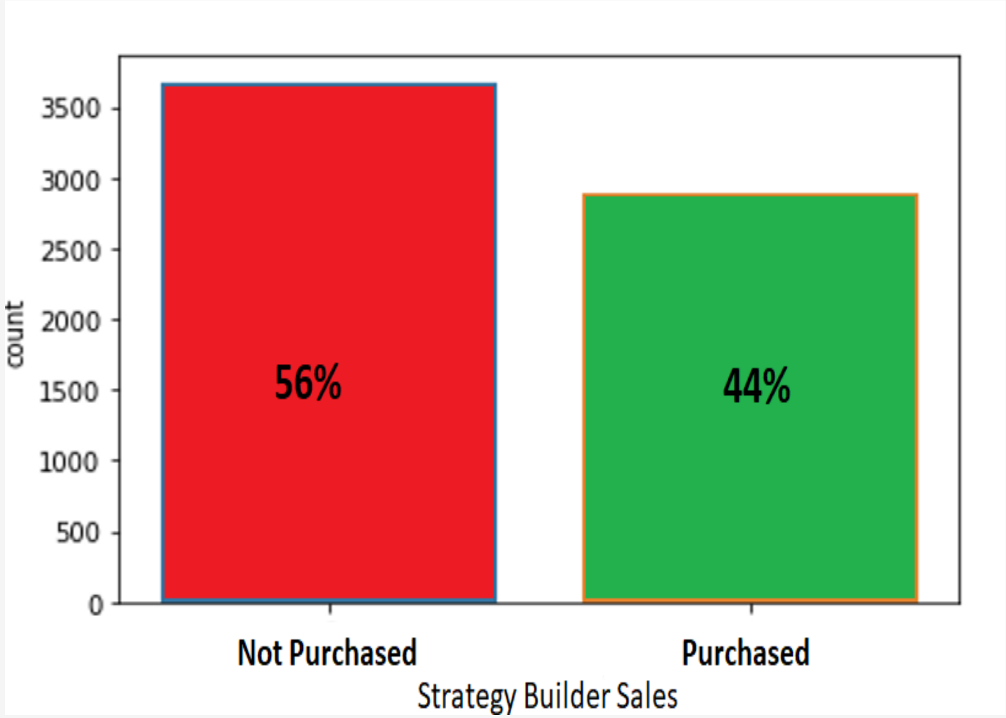
Literature Review

Title of the Paper	Author	Source	Major Insight
Who are the Likeliest Customers: Direct Mail Optimization with Data Mining	Bole, Uroš Papa, Gregor - 2011	Contemporary Engineering Sciences	This paper paved the path for this study and gives an overview of the data mining work performed with the aim to build a prediction model, which could help save costs for companies that engage in direct marketing activities. Binary Classification and Clustering were the techniques used.
One-class versus binary classification: Which and when?	Bellinger, Colin Sharma, Shiven Japkowicz, Nathalie - 2012	Proceedings - 2012 11th International Conference on Machine Learning and Applications, ICMLA 2012	One-class classification is useful when there is an overabundance of data for a particular class. In such imbalanced cases, binary classifiers may not perform very well. The paper investigates the performance of binary and one-class classifiers as the level of imbalance increases.
Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining	Kaur, Manpreet Kang, Shivani - 2016	Procedia Computer Science	The main aim to provide the information to the retailer to understand the purchase behavior of the buyer for correct decision making. The existing algorithms work on static data, and they do not capture changes in data with time. But proposed algorithm not only mine static data but also provides a new way to consider changes happening in data

Literature Review

Title of the Paper	Author	Source	Major Insight
Market Basket Analysis Algorithm with Map/Reduce of Cloud Computing	Woo, Jongwook Xu, Yuhang - 2011	The 2011 International Conference on Parallel Processing	The paper presents the Market Basket Analysis algorithm with Map/Reduce. The results show that it increases the performance by adding more nodes but at a certain point, there is a bottleneck that does not allow the performance gain.
An overview of clustering methods	Omran, Mahamed G.H. Engelbrecht, Andries P. Salman, Ayed - 2007	Intelligent Data Analysis	Clustering is not unique, and it strongly depends upon the analyst's choices. The paper described how it is possible to combine different results in order to obtain stable clusters, not depending too much on the criteria selected to analyze data. Clustering always provides groups, even if there is no group structure.
Review on determining of cluster in K-means	Kodinariya, Trupti M Makwana, Prashant R - 2013	International Journal of Advance Research in Computer Science and Management Studies	The paper explores six different approaches to determine the right number of clusters in a dataset Keywords: Akaike's information criterion, Bayesian inference criterion, Clustering, Cross-validation, Elbow Method, Jump Method, Number of Cluster, Silhouette.

Problem Statement



Cost of each trial: $2300

Customers who did not purchase the product: 3674

Loss = (3674 * 2300) = **$8,450,200**

The company loses out on Revenue because of this. Hence, the company is looking out for ways to **"Minimize the Losses on Trials of Strategy Builder Tool".**

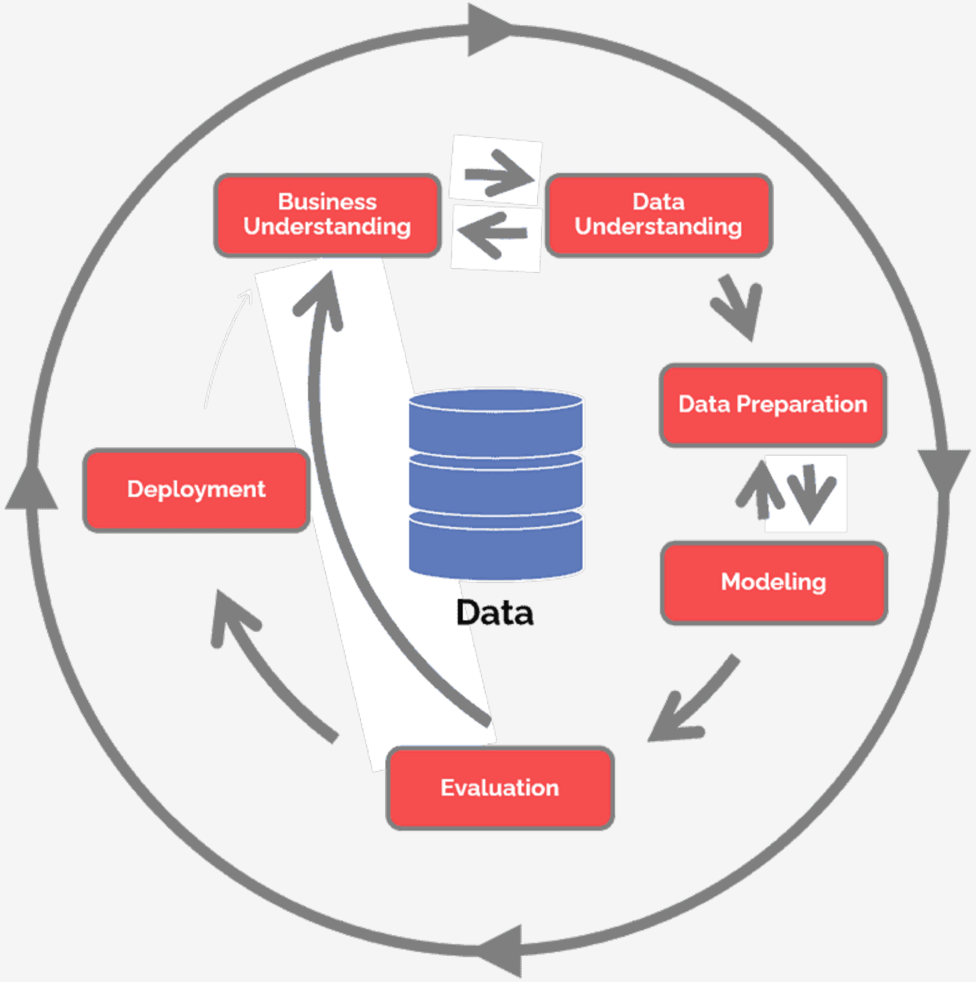
Project Objectives

There is no direct solution to this problem, hence the problem was divided into sub-problems, and accordingly below objectives were found to proceed with the project:

1. **Identify the most probable customers to whom this product can be sold.**
2. **Identify the set of products with which the new product gets sold frequently.**
3. **Identify the set of similar customers who are least responsive in buying the new product.**

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Project Methodology



CRISP-DM: The Cross Industry Standard Process for Data Mining is adopted for this study. It has six sequential phases:

1. Business Understanding

2. Data Understanding

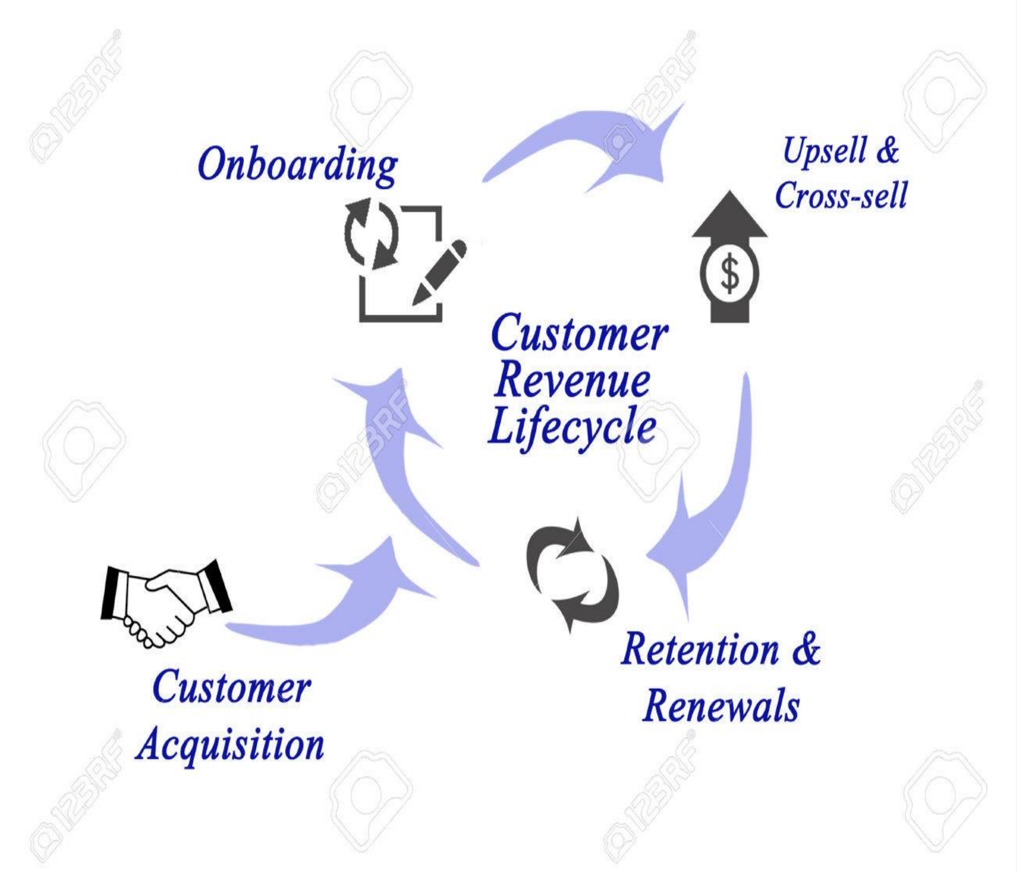
3. Data Preparation

4. Modeling

5. Evaluation

6. Deployment

Business Understanding



RLP Software Pvt Ltd. headquartered in San Mateo, California is a **B2B** company developing business process automation solutions.

It offers the Commercial Operations Suite, which enables businesses to optimize quotes and digital commerce, manage contracts and documents, as well as automate revenue management. The company caters to the energy, financial services, healthcare, media, retail and other sectors.

Well-known products used by **11000+ customers** of the company are :

1. CPQ (Configure, Price, Quote)
2. CLM (Contract Lifecycle Management)
3. IWA (Intelligent Workflow Approvals), etc.

The New Product

Strategy Builder: A new product of RLP launched in Q1 2021 which focuses on Incentive Management. This tool aims to give indefinite ease and flexibility to create rules/schemes for a business to incentivize their customers for their orders/purchases.

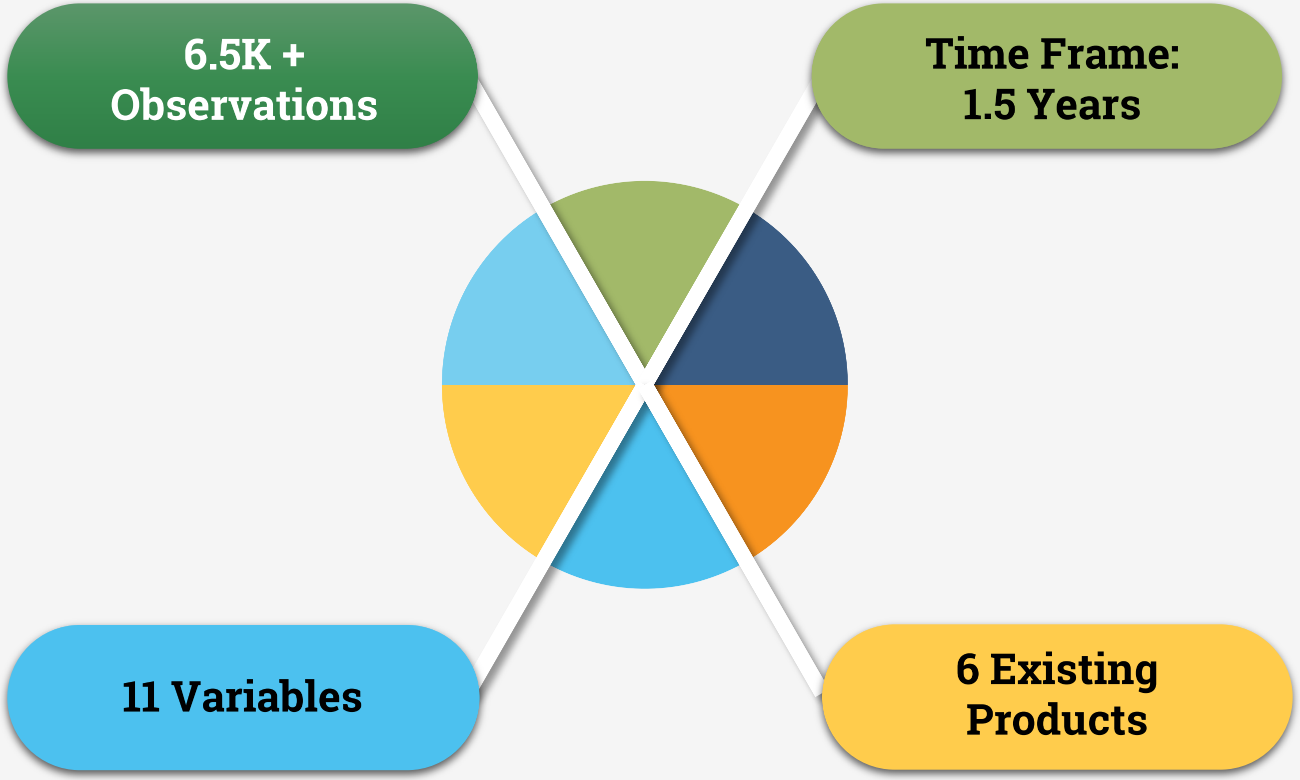
Billing is an important part of the Revenue Lifecycle, and this is where this tool comes into the picture.

Significant Use Cases of tool (but not limited to):

1. Buy X Get X – Discounted prices on the product if a certain amount quantity is purchased.
2. Buy X Get Y – Discounted price on a different product if another product is bought of a certain quantity.
3. Package Savings - Discounts if a set of products are bought together as a package.
4. Subscription – Discounted price if the product is purchased on a regular basis

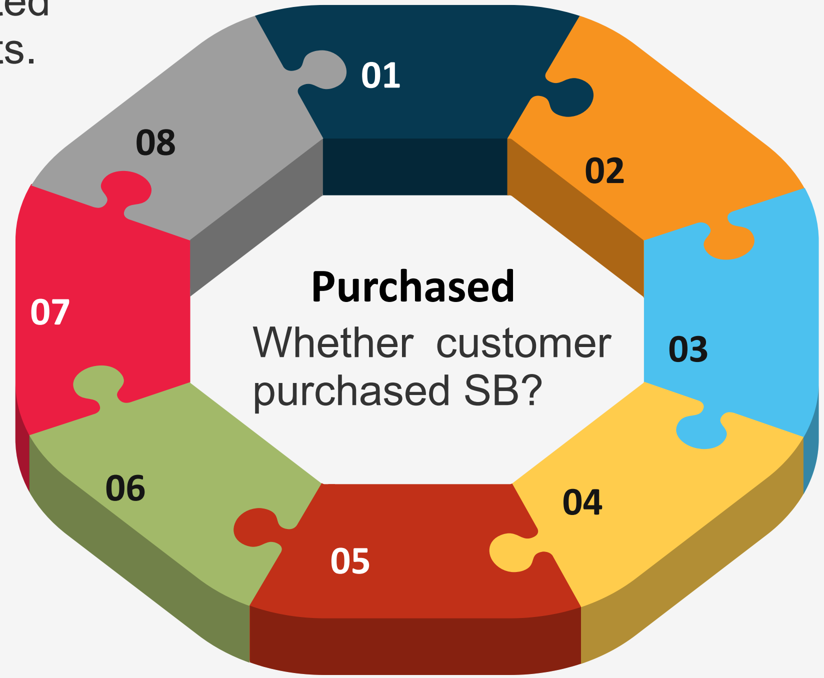
Data Understanding

Data Collection | Variables



REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

Variables



Domain

The functional area of the customer company.

Region

The continent in which the company majorly operates.

Products already purchased

Count of all the products that the customer has already purchased.

Using Similar Product

Whether the company is using a similar product?

Renewals

Count of Renewals already done for other products.

Product's Sales Data

Customer purchase history for 3 products.

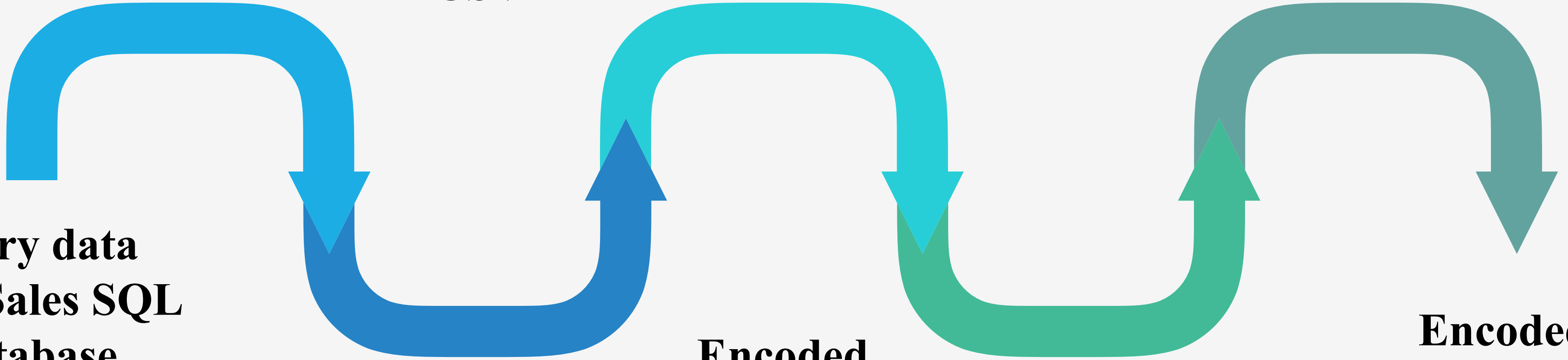
Years of Association

No. of years the customer is associated with RLP.

Bugs

No. of bugs reported from existing products.

Data Preparation



Query data from Sales SQL Database (Select + Join Queries)

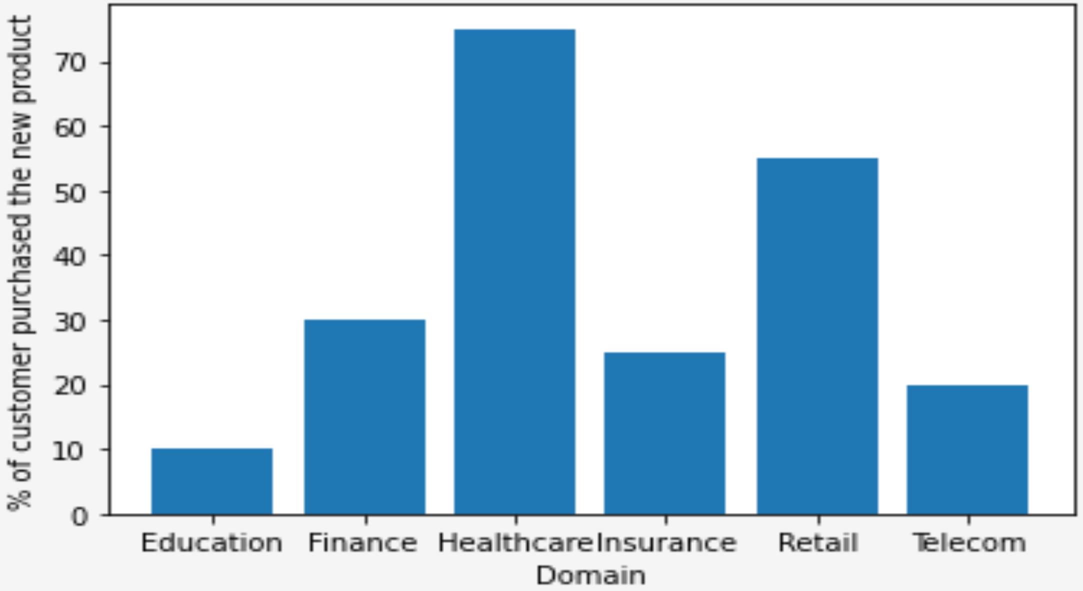
Converted data to a single view and exported to CSV

Encoded Nominal Variables

Binned Discrete Variables

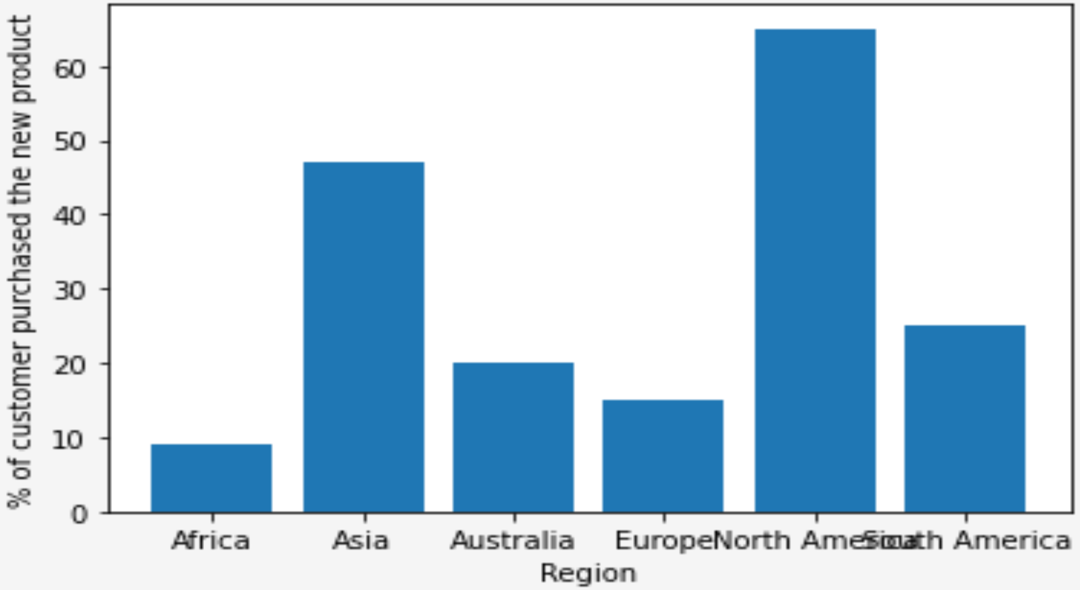
Encoded Ordinal Variables

Sales Data



Sales data based on Domain

Highest Sales: Healthcare and Retail
Lowest Sales: Education and Telecom

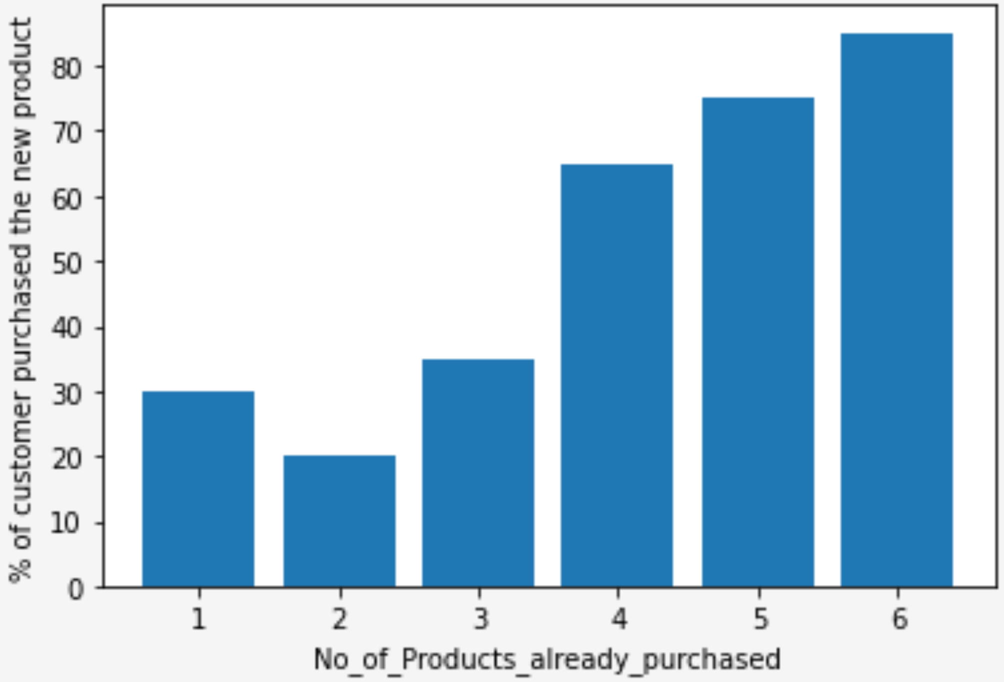


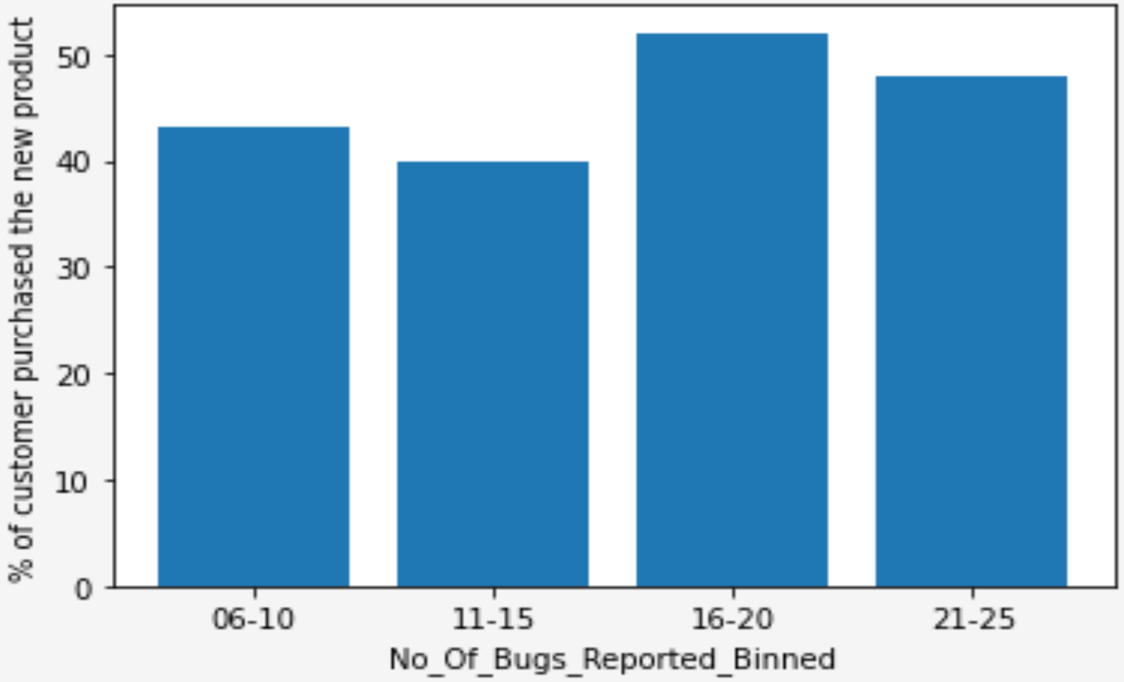
Sales data based on Region

Highest Sales: Asia and North America
Lowest Sales: Africa and Europe

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956, Approved by AICTE, New Delhi

Sales Data





Sales data based on No. of Products already purchased

Customers who already purchased higher no. of products are purchasing the new product more.

Sales data based on No. of Bugs Reported

No significant effect of Bugs on Sales.

Binary Classification

Objective: Identify the most probable customer for the new product.

Method: Binary Classification

Dependent Variable: Purchased

Independent Variables: 10 (All columns - Purchased)

Result: Random Forest Model performed the best among other models.

Modeling Technique	Accuracy
Decision Trees	77.9%
Logistic Regression	70.6%
Naïve Bayes	72.8%
Random Forest	**83%**
XGBoost	82%

Model Deployment

Model as API: It is a business logic deployed on the server with a contract of input and output. It is platform-independent and can be called from any tech stack provided the contract is followed.

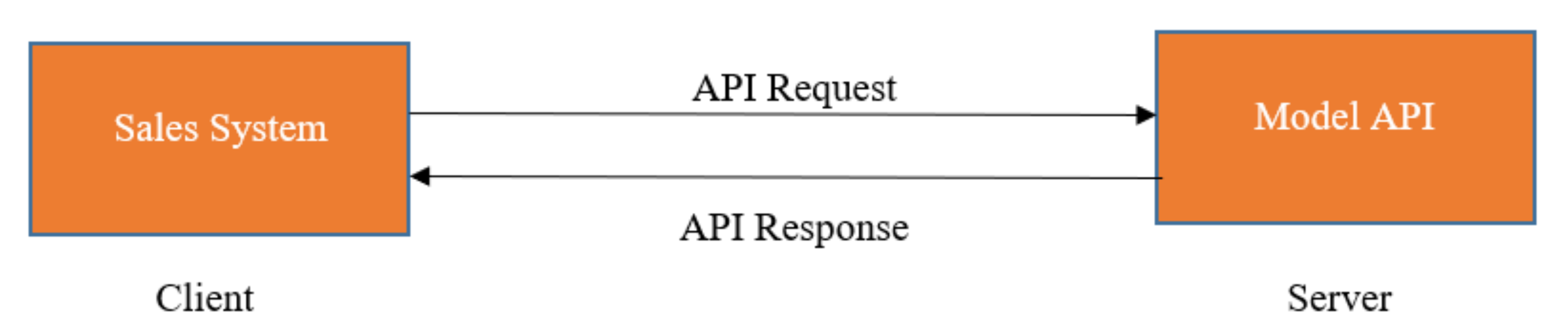
Tentative Framework: Flask

Sample Endpoint for Caller: https://{routing}/is-probable-customer

Verb: Get

Body: Customer data object

Response: Boolean (True or False)



Reality Check

The model gave correct response for 3 trial results that came after the project was done. The prediction was checked manually as the model is not deployed yet.

Market Basket Analysis

Objective: Identify the affinity of Strategy Builder with the existing products of RLP.

Technique: Apriori (Association Rules)

Observation: The customers who have purchased CLM and CPQ both are more likely to purchase the new product, followed by the customers who just bought the CPQ product.

Clustering

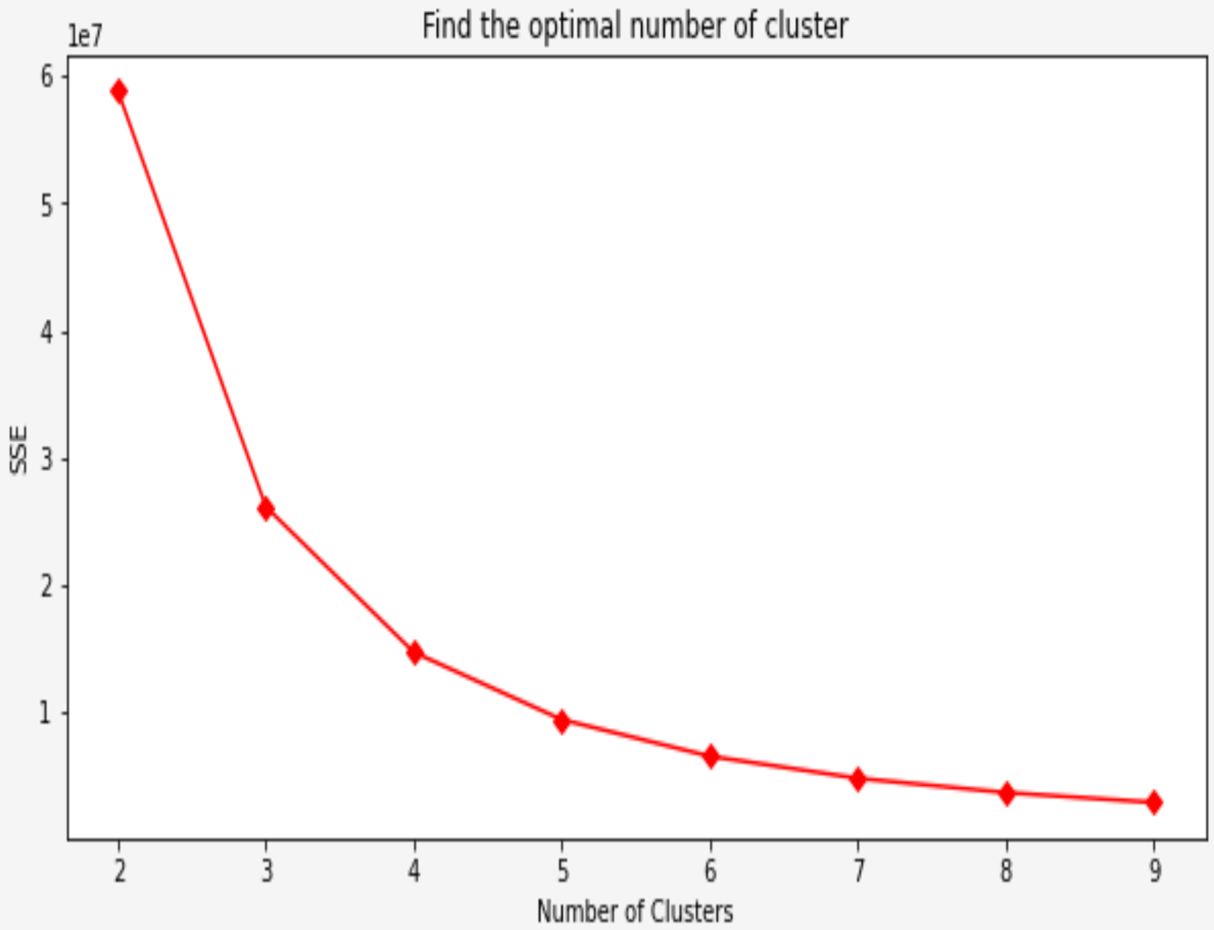
Objective: Cluster the customers. Identify similar customers who are buying/not buying the product.

Variables: All

Technique: k-means clustering

Optimal number of Clusters found: 4

Observation: First cluster had the highest number of Strategy Builder customers while the second one had lowest number of customers.



Conclusions and Recommendations

Based on the above study, below are the conclusions and recommendations for RLP software:

- Measures should be taken to attract customers from Education and Insurance domains as these are the lowest performing domains. It is strongly recommended to get feedback from them and add some domain-specific features to the tool to attract these customers.
- Africa and Europe are the lowest performing regions for this product, marketing and promotion should be reviewed and strategized to acquire the market share in these regions. Competition should be analyzed and measures to be taken accordingly.
- Random Forest Model API to be integrated with RLP's Sales system and before giving a trial version to any customer, the probability must be checked.

Conclusions and Recommendations

- For the customers who are less likely to buy the new product as per the model's response, they should be given a demo of the product (in premise) instead of giving a separate trial version, this is an attempt to introduce our product at a negligible cost. (After all, they are customers too!!)
- The new product should be considered for cross-selling with CPQ and CLM products. This recommendation is for existing as well as future customers.
- Customers in cluster 2 should be targeted for onboarding by applying similar strategies in terms of marketing, promotions, discounts, etc. as they have shown similar characteristics while clustering.

Scope For Future Work

1. Discover insights from the lowest and highest performing cluster.

2. Identify ways to move customers from the lowest performing cluster to the higher performing one.

3. Improvise the model based on fresh/future data.

References

[1] "95 Percent of New Products Fail. Here Are 6 Steps to Make Sure Yours Don't | Inc.com." https://www.inc.com/marc-emmer/95-percent-of-new-products-fail-here-are-6-steps-to-make-sure-yours-dont.html (accessed Aug. 10, 2022).

[2] K. Sun, M. Zuo, and D. Kong, "What Can Product Trial Offer?: The Influence of Product Trial on Chinese Consumers' Attitude towards IT Product," *https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/IJABIM.2017010102*, vol. 8, no. 1, pp. 24–37, Jan. 1AD, doi: 10.4018/IJABIM.2017010102.

[3] "What is purchase intent? - Definition from WhatIs.com." https://www.techtarget.com/whatis/definition/purchase-intent (accessed Aug. 10, 2022).

[4] W. Mai, F. Wu, F. Li, W. Luo, and X. Mai, "A data mining system for potential customers based on one-class support vector machine," *J. Phys. Conf. Ser.*, vol. 2031, no. 1, 2021, doi: 10.1088/1742-6596/2031/1/012066.

[5] "Purchase Intention - Meaning, Importance, Factors & Example | MBA Skool." https://www.mbaskool.com/business-concepts/marketing-and-strategy-terms/10976-purchase-intention.html (accessed Aug. 10, 2022).

[6] V. Mirabi, H. Akbariyeh, and H. Tahmasebifard, "A Study of Factors Affecting on Customers Purchase Intention Case Study : the Agencies of Bono Brand Tile in Tehran," *J. Multidiscip. Eng. Sci. Technol.*, vol. 2, no. 1, pp. 267–273, 2015.

[7] A. Fathy, S. Younus, F. Rasheed, and A. Zia, "Identifying the Factors Affecting Customer Purchase Intention," *Glob. J. Manag. Bus. Res. Adm. Manag.*, vol. 15, no. 2, pp. 1–6, 2015.

[8] "Machine Learning Models - Javatpoint." https://www.javatpoint.com/machine-learning-models (accessed Aug. 11, 2022).

[9] U. Bole and G. Papa, "Who are the Likeliest Customers: Direct Mail Optimization with Data Mining," *Contemp. Eng. Sci.*, vol. 4, no. 6, pp. 259–268, 2011.

[10] C. Bellinger, S. Sharma, and N. Japkowicz, "One-class versus binary classification: Which and when?," *Proc. - 2012 11th Int. Conf. Mach. Learn. Appl. ICMLA 2012*, vol. 2, pp. 102–106, 2012, doi: 10.1109/ICMLA.2012.212.

References

[11] M. Kaur and S. Kang, "Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining," *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 78–85, 2016, doi: 10.1016/j.procs.2016.05.180.

[12] S. Gupta and R. Mamtora, "A Survey on Association Rule Mining in Market Basket Analysis," *Int. J. Inf. Comput. Technol.*, vol. 4, no. 4, pp. 409–414, 2014, [Online]. Available: http://www.irphouse.com/ijict.htm

[13] J. Woo and Y. Xu, "Market Basket Analysis Algorithm with Map/Reduce of Cloud Computing," *2011 Int. Conf. Parallel …*, no. April 2012, 2011, [Online]. Available: http://www.lidi.info.unlp.edu.ar/WorldComp2011-Mirror/PDP4494.pdf

[14] M. G. H. Omran, A. P. Engelbrecht, and A. Salman, "An overview of clustering methods," *Intell. Data Anal.*, vol. 11, no. 6, pp. 583–605, 2007, doi: 10.3233/ida-2007-11602.

[15] T. M. Kodinariya and P. R. Makwana, "Review on determining of cluster in K-means," *Int. J. Adv. Res. Comput. Sci. Manag. Stud.*, vol. 1, no. 6, pp. 90–95, 2013, [Online]. Available: https://www.researchgate.net/publication/313554124

[16] "Understanding K-means Clustering in Machine Learning | by Education Ecosystem (LEDU) | Towards Data Science." https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1 (accessed Aug. 11, 2022).

[17] "What is CRISP DM? - Data Science Process Alliance." https://www.datascience-pm.com/crisp-dm-2/ (accessed Aug. 10, 2022).

[18] "Understanding Random Forest. How the Algorithm Works and Why it Is… | by Tony Yiu | Towards Data Science." https://towardsdatascience.com/understanding-random-forest-58381e0602d2 (accessed Aug. 10, 2022).

[19] "What is XGBoost? | Data Science | NVIDIA Glossary." https://www.nvidia.com/en-us/glossary/data-science/xgboost/ (accessed Aug. 10, 2022).

[20] "A Gentle Introduction on Market Basket Analysis — Association Rules | by Susan Li | Towards Data Science." https://towardsdatascience.com/a-gentle-introduction-on-market-basket-analysis-association-rules-fa4b986a40ce (accessed Aug. 10, 2022).

Plagiarism Report

Minimizing Losses on Trials of Strategy Builder Tool

ORIGINALITY REPORT

8%	7%	4%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

REVA
UNIVERSITY
Bengaluru, India

Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Thank you!