Improving Sales and Marketing Effectiveness Using Clustering and Market Basket Analysis

Shefali Jain, Akash Kashyap, Sagar Pradhan, Aditi Vatse, Matthew A. Lanham

Purdue University Krannert School of Management

jain291@purdue.edu; kashyap5@purdue.edu; pradhan4@purdue.edu; avatse@purdue.edu; lanhamm@purdue.edu

Abstract

A comprehensive portfolio of products in agriculture challenges sales and marketing teams to tailor optimal offers that combine the right products to different customer segments. In collaboration with a major US based agro science firm, we:

- Build and assess traditional choice models of the firm,
- Segment customers based on number of acres, distributors, purchase behavior and transactions using an unsupervised clustering algorithm,
- Perform market basket analysis to provide effective marketing techniques to salesforce and provide efficient product line options to customers.

The business goals are to segment the market, increase marketing and sales effectiveness and identify untapped market potential. They key objective is to identify revenue growth opportunities that can increase the firm's market share.

Introduction

The motivation for this study is to enable sales and marketing teams to tailor optimal offers that combine the right products to different customer segments. Choices of crop productivity level and crop protection management are customer and geographic specific, and in many instances affected by temporary market conditions. Therefore, understanding customer needs in this complex scenario can help the firm fulfil demand, design optimal combinations of seeds and crop protection products for distinct customer segments. Analytical solutions can help us design optimal combinations of Seeds and Crop Protection products for distinct customer segments and can be leveraged across geographies.

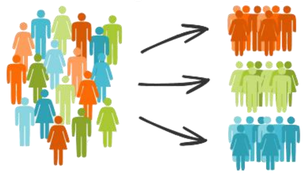

Figure 1. Customer segmentation

Figure 2. Market potential capture through product lines

Research objectives:

- Segment and profile customer groups which can be used to effectively market the right product to customers
- Provide a market basket motivated result which can be used by the marketing team to identify uncaptured market potential for different products from their product line, and identify future revenue growth

Literature Review

Once we frame the business problem in an analytics problem, we researched techniques have been used to develop similar solutions. Below is the list of research papers that used either clustering (CF, RFM) or prediction techniques (MBA, Regression). Based on the goals of our sales team, we used hierarchical clustering to profile their customers and market basket analysis to identify best subsets of product offerings.

Sr No	Paper	MBA	RFM	Reg	CF	RF
1	Prediction Approach			●	●	
2	Higher Education Recommendation			●		
3	Predicting Online Behavior			●		
4	Random Forest for Online Behavior					●
5	Clustering item preference		●			
6	Building Prediction Model	●				
7	Product Recommendation		●		●	
8	Recommendation for grocery store				●	
9	Cross category dependence	●			●	
10	Market Basket Analysis	●				
	Product Recommendation Model	●	●			

Table 1. Literature review summary by method researched for the project

Methodology

Figure 3 outlines our study design, starting from data collection, data cleaning, data pre-processing, model design and methodology selection.

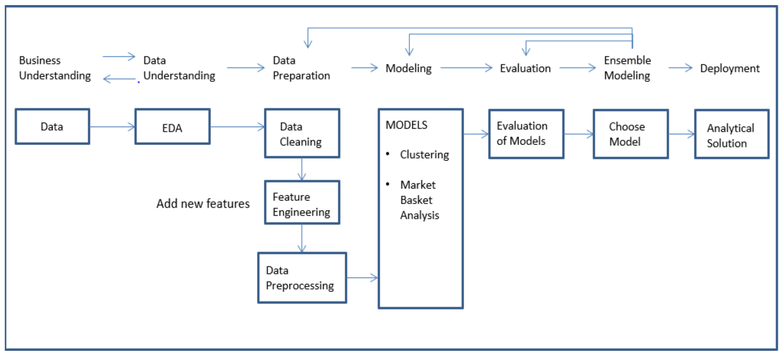

Figure 3. Study Design

Data procured from the firm

1) Farm features and size
2) Sales data of last 3 years
3) Territory alignment for sales representative

Data Cleaning & Pre-Processing

EDA was used to identify the outliers and trends. Missing values were treated and incorrect formats were standardized.

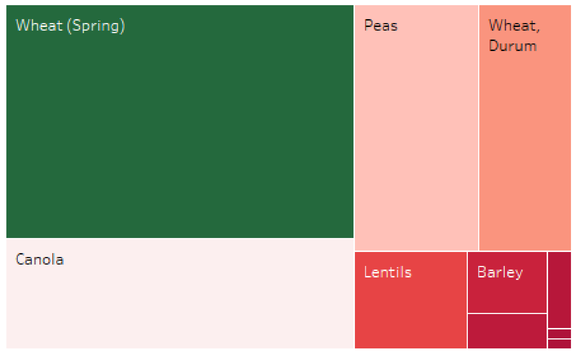

Figure 4. Acres owned by customers divided by crop

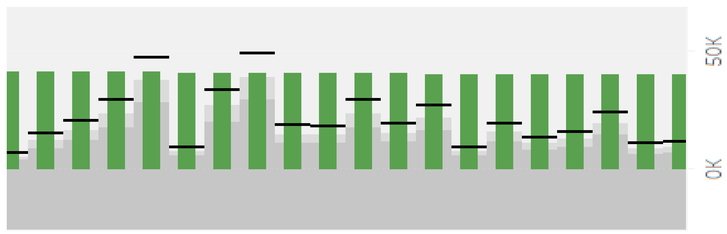

Figure 5. Sales potential (in green) vs actual sales (in gray)

Model Design and Methodology Selection

Farm size data and sales data were merged to get all features in one data source. This helped in clustering on customer demographic, geography and transactional data. It also helped to see how much of the farm area has not been captured.

Methods**: Hierarchical Clustering, Market Basket Analysis**

Clustering was performed to identify different customer profiles. Market Basket Analysis helped in determining different association rules which could later be coupled with customer profiles to create customized marketing strategies.

Model Evaluation / Statistical & Business Performance Measures

Clusters were evaluated by using Hartigan's rule and the Silhouette method. Lift ratio and confidence metric was used to determine the efficiency of the association rules.

Results

After hierarchal clustering and k-means verification, we concluded that the customers can be partitioned into 5 segments. The firm can market to each segment differently after understanding the demographic of each cluster.

The market basket analysis showed that seeds and crop protection are complementary and customers who bought one of the products are highly likely to buy the other. Selling these products together can help the firm capture more revenue and acre share. The market basket analysis provides the firm with a simple tool to judge what product to market to which customer in order to improve marketing effectiveness and increase revenue from both seed and crop protection products.

The lift in our market basket analysis model was quite high for few of the combinations in product offerings when controlled for both support and confidence. We applied threshold on lift and confidence to limit the number of recommendations per product and customer.

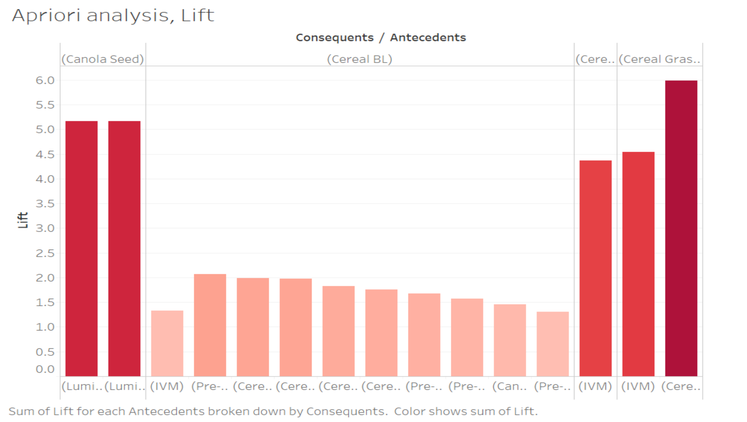

Figure 6. Market basket result for Canola seed, Cereal BL etc.

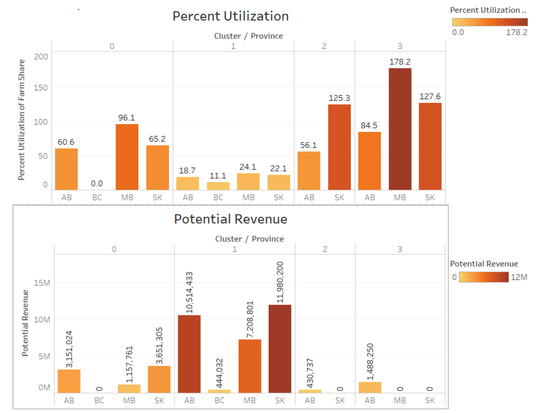

Figure 7. Potential revenue gain from Clustering

Conclusions

To sell a product to a customer, it is important to understand the customer and its needs. Mass marketing is not useful anymore. Effective marketing involves providing right options to the customers. The model we have built can help the firm achieve effectiveness.

The study not only helped to understand the customer buying behaviour but also showed where the firm was missing out with market share. The model was able to successfully identify and recommend products for different customer segments as well as individual customers.

Acknowledgements

We thank Professor Matthew Lanham for constant guidance on this project. We also thank the firm to provide us with the opportunity to work on such an interesting problem.