Proposal for Acme Innovations

Executive Summary:

We at Data Catalyst Consultancy are a team of professionals who have witnessed firsthand how revolutionary the correct data can be for accelerating business growth. We understand the very real difficulties that emerge when consumer preferences start to change, and competition heats up since we have worked in the trenches with businesses in a variety of industries. We can comprehend Acme Innovations' critical predicament because of that hard-won experience.

Your well-earned market position is in jeopardy because of the dropping customer retention rates that have affected your iconic brand, which generations of consumers have loved and trusted. We understand that losing devoted clients is a gut punch that no company wants to experience, particularly after investing years and years of effort into developing those strong client connections. But for exactly this reason, our team is excited to work with you.

We'll bring our battle-tested experience and roll up our sleeves alongside your team to study the data, truly understand what's behind this trend, and map out a smart path forward to reinvigorate those customer bonds.

Our strategy is to unlock the rich insights hidden within your customer data and use that knowledge to rekindle client loyalty and accelerate growth. We'll use advanced data mining techniques such as customer segmentation models and product recommendation engines to gain a thorough understanding of your consumer base.

We'll find your most valuable client segments and customize strategies to their specific needs by evaluating purchase behaviors, demographics, and levels of involvement. Our product recommendation system will make highly relevant recommendations based on actual purchase patterns, resulting in enhanced satisfaction and cross-selling opportunities.

What distinguishes us is our demonstrated ability to generate measurable results. We're more than simply numbers of crunchers; we're strategic partners who want to help you navigate today's changing economy. With Data Catalyst on your side, we are convinced Acme Innovations can restore its proper role as an industry leader for future generations.

Our Approach:

Data Exploration and Cleaning:

We have two datasets, “customer_data” containing 50,000 customer details such as “customer ID”, “age”, “salary”, “spending score”, “time as customer”, “loyalty member”, “state”, “gender”

and “customer_purchase_history_final” contains the products purchased by customers from ACME Innovations.

The first row of “customer_purchase_history_final” are the products purchased by customer ID 1 from “customer_data” and so on.

We observed that “customer_purchase_history_final” has more than 50,000 rows which doesn’t match the 50,000 customers from our “customer_data” file.

To link both datasets, we removed the rows past 50,000.

We have checked both datasets for null values but didn’t any.

We cleaned the “customer_purchase_history_final” dataset by removing prefixes starting with "X" or "X." from product names, replacing dots with spaces within product names to correct formatting like "washing.machine" to "washing machine", and finally trimming any leading or trailing whitespace from all cell values.

Model Selection:

Customer Segmentation:

We have employed the K-Nearest Neighbors (KNN) clustering technique for customer segmentation.

K-Means clustering is highly effective for your dataset because features like salary, spending score, and time as a customer naturally group into distinct clusters. This algorithm excels in identifying these clusters based on spending behavior and customer engagement, which are crucial for segmenting customers into meaningful groups.

Having these well-defined customer segments will be extremely valuable. It will enable you to tailor your marketing messages and offers to better align with the specific needs and preferences of each group. It can also help optimize how you deliver customer service and support.

We’ve used salary, and spending_score for clustering the customers.

Product Recommendation:

For the product recommendation, we chose the Apriori algorithm, a well-known technique for extracting association rules from transactional data. The Apriori algorithm is particularly well-suited for this task because of its capacity to detect frequent co-occurring patterns in purchase histories, allowing the identification of often purchased products.

The Apriori algorithm enables us to use Acme's purchase data to identify relevant product relationships and patterns. This allows the development of a recommendation system that can suggest highly relevant product offerings to customers based on their previous purchases and recognized relationships.

Model Training & Evaluation:

KNN Clustering:

The KNN clustering algorithm groups the customers based on their feature similarities (salary and spending score in this case).

We scaled the feature data to ensure equal contribution from both salary and spending scores and applied the elbow method to evaluate the within-cluster sum of squares (WCSS) for different numbers of clusters.

WCSS is the sum of the squared distances between each data point in a cluster and the centroid of that cluster.

Apriori Recommendation Algorithm:

We organized the customer_purchase_history_final data into a format that the Apriori algorithm can understand, where each customer's purchases are treated as a single transaction or group of items.

Next, we set two key parameters: support and confidence. Support determines how popular a product combination needs to be in the data to be considered frequent. Confidence measures how likely it is for one product to be bought if another product is also bought.

The Apriori algorithm then analyzes the transaction data. It first identifies the most common individual products based on the support we set. It then looks for combinations of two products that are frequently bought together, then three products, and so on. At each step, it eliminates any product combinations that don't meet the minimum support level.

Finally, from the frequent product combinations, the algorithm generates rules in the form of "If product A is bought, then product B is also likely to be bought." These rules must meet the confidence threshold we specified earlier.

Metrics:

For the product recommendation, we used the Apriori algorithm to discover patterns and relationships between products that customers frequently purchase together.

Three key metrics from the Apriori algorithm are:

Support: This measures how frequently a set of products appear together in customer purchases. A higher support value indicates that the product combination is more popular and commonly purchased. We want to identify rules with reasonably high support, as they represent recurring purchasing patterns across many customers.

Confidence: This tells us how likely it is for a customer to buy Product B if they have already bought Product A. A confidence value close to 100% means that whenever Product A is purchased, Product B is almost always purchased along with it. Rules with high confidence represent strong associations between products that can enable effective recommendations.

Lift: This metric compares the confidence of a rule to what we would expect if the products were purchased independently by chance. Higher lift values suggest more meaningful and actionable relationships between products.

We may discover the most significant and dependable product affinities by reviewing the Apriori algorithm's generated rules and focusing on those with high support, confidence, and lift. These data will enable us to create a recommendation system that can accurately propose complementary products to clients based on previous purchases, resulting in enhanced customer satisfaction and potential cross-selling opportunities.

Recommendations:

From the customer segmentation, we saw that some customers with high salaries could spend more but they aren’t. We can focus more on those customers to increase the revenue of ACME Innovation by upselling or downselling based on the product recommendation. We can also offer membership offers to the customers who have good spending scores at ACME which can benefit them as well as the recurring revenue of the company.

From the product recommendation, we can focus on the products with more support and confidence, as they are mostly likely to get sold if other products are getting sold. It is good to go through all the product recommendations from the model and recommend the products to the customer. We can also offer discounts based on these recommendations and order value to increase sales.