IE 492 PROJECT REPORT

Personalized Dynamic Pricing in Agricultural Loan Applications

Ferhat Turhan
Şeyma Çakır
Kaan Aslan
Serdar Yıldız

Project Instructor: Mustafa Gökçe Baydoğan

ABSTRACT

Agricultural loan applications have a wide range of customer types and profit maximization is a very common topic in this area. Personalized dynamic pricing is one of the approaches to use in pricing strategies with the improvement in the data driven models. In this paper, we propose different machine learning methods such as logistic regression, clustering methods, and model based trees to learn the demand curves of different types of customers and segment the customers according to their price sensitivities. Then, the outputs are used to make a dynamic pricing strategy in order to reach the target of collected amount of money with high profitability. Our methodology dynamically prices the loans according to how close the target is. Finally, the models and the pricing strategy is tested by simulation.

Tarım kredi uygulamalarının geniş çaplı bir müşteri kitlesi vardır ve kar maksimizasyonu bu alanda sıkça karşılaşılan konulardan biridir. Kişiye özel dinamik fiyatlandırma, veriye bağlı modellerin gelişmesi ile bu alanda kullanılan bir strateji türüdür. Bu makalede; lojistik regresyon, segmentasyon metodları ve model bazlı karar ağacı makine öğrenmesi algoritmalarının yardımı ile müşteri talep eğrilerinin öğrenimi ve müşterileri fiyat hassasiyetine göre gruplandırma yaklaşımını uyguladık. Sonrasında, bu modellerden elde ettiğimiz çıktıları dinamik fiyatlandırma stratejisinde, hedeflenen satış miktarına en karlı şekilde ulaşmak için kullandık. Uygulamamız hedefe olan uzaklığa göre müşteriye özel dinamik bir fiyatlandırma sunar. Son olarak, makalemizde bahsedilen modeller ve stratejileri simulasyon ile test ettik.

Key Words: Agricultural Loan, Price Sensitivity, Dynamic Pricing with Milestones, Price Segmentation, Logistic Regression

Anahtar Kelimeler: Tarım Kredisi, Fiyat Duyarlılığı, Kilometre Taşlarıyla Dinamik Fiyatlandırma, Fiyat Segmentasyonu, Lojistik Regresyon

1. Introduction

The aim of this project is to offer the optimal interest rate to the customers such that the customers should accept the offered interest rate and at the same time the loan value collected from the customers reaches the predetermined target value. It is a fact that lenders offer varied interest rates to different loan applications based on pricing segments of customers. This variance in price is mostly related to the customer's risk profile since variable costs occur across various price segments. Since profit maximization is one of the primary corporate objectives in this scenario, companies mostly offer high prices for risky customers. However, there is a trade-off between price and customer loss in a competitive market. In other words, the pricing of competitors should also be considered. The risk assessment is not a part of the project but the metrics which are also credibility scored are used for finding customers' price sensitivity.

The project focuses on agricultural loans. Agricultural loans differ from other loans such as mortgages because farmers borrow it to plant their land and, since farmers earn their income by planting their lands, there is a difference in willingness to accept the offers of customers by comparing other types of loans. Moreover, the applicant's price that they are willing to pay may be estimated depending on the loan requirement's specifics. For instance, if a farmer requires a loan to buy fertilizers to grow wheat or a tractor to harvest, the priority and urgency of buying the fertilizer will be higher and the customer may be willing to pay more. Therefore, customized pricing particular to each farmer application is required. This project involves determining the price sensitivity of applicants and pricing loans accordingly to the price sensitivity.

The first part of the project is data analysis. The data is provided by a company that offers customers loans indirectly. It means that the company purchases products for a customer and sends them instead of giving money to the customer and the customer pays back the money to the company with interest. The data includes product information that customers purchased, the crop information which is customer landing if the customer is a farmer, demographic data of the customer, customers' financial statistics, interest rates offered to the customer, and information on whether customers accept the offer or not.

The financial statistics of the customers are very correlated, hence, many columns are dropped. The product and crop information were very detailed so they are grouped to increase the accuracy of the data. There are too many null values in the data. Since the data is scaled for the protection of personal data, it is hard to estimate appropriate values for null values. Therefore, the columns with a high amount of null values were either dropped or categorized. The main problems in data analysis are that the data is difficult to interpret and it is biased because it is the result of the company's pricing strategy and risk analysis decisions. The ratio of the offered interest rates to the consumer interest rates at the application date is used as the factor that determines the price sensitivity of the customer since competitive prices are not available. This ratio is called TLREF. After data analysis, the project is divided into two phases which are determining price sensitivity by using data prepared in data analysis and determining the pricing strategy by using price sensitivity curves in the first phase.

The first phase of the project is to determine customers' price sensitivity based on historical data. As a base approach, the logistic regression curve is used to find the probability of a customer accepting an offered price with particular characteristics. This approach gives only one price sensitivity curve for all customers but the base probability of a customer acceptance is determined by other characteristics of the customer. The second approach is dividing customers into different segments by excluding interest rate offers. After customer segmentation, the sensitivity curve for each segment is obtained by using offered interest rates. By this approach, it is expected to show that the customer's price sensitivity differs by particular characteristics. The last approach is dividing customer segments based on sensitivity curve accuracies. This approach has an objective that divides customers such that the sensitivity curve of each segment gives the best curve to explain the probability of customer acceptance in offered prices.

The second phase of the project is to determine the best price strategy to maximize revenue. The main approach is dynamic pricing with milestones (Besbes, Maglaras, 2012) which is an algorithm taking a target value and trying to reach the target value by adjusting price offers. In each iteration, the algorithm calculates the expected money required from a customer to reach the target and returns the probability required for the customer to accept the offer. By using sensitivity curves the price offer is determined and if the customer accepts the offer the required amount is decreased by the gain of the customer. As a result, if the company is close to the target, it can take risks and offer high prices to increase profitability or if the company is far from the target, it does not take risks and offers lower prices and increases the sales. The second approach is the aggressive model. This model offers customers high prices by comparing the market, aiming for high profitability with fewer customers. The third model, in contrast to the second model, offers customers lower prices they are probably willing to pay and tries to increase the total revenue with more customers but low profitability. The last model is the controlled variance pricing model which offers prices in an interval that is determined by the historically accepted prices.

In the end, it is observed that as expected, model-based trees have better results compared to the logistic regression model when learning the demand curve of the customers. For the second phase, dynamic pricing with milestones offering is a better model compared to others in that it can achieve the determined target better due to the nature of its algorithm. In this paper, different kinds of pricing strategies are also examined and they can be used as well for different objectives.

The following segments of this paper contain detailed explanations of the above-mentioned information. In the second part, the problem is defined, requirements, and limitations are explained which is followed by the literature review, assumptions, and method selections in part three. In part 4, detailed explanations of the selected methods for both learning price sensitivities and dynamic pricing strategies can be found. Validations and comparisons of these methods are shown with simulation results in part 5. Suggestions for implementation of the results are discussed in part 6 and it is followed by the discussion section in part 7. At the end, references and appendix are attached.

2. Problem Definition, Requirements, and Limitations

2.1 Problem Definition

It has been very important to offer optimal prices to the customers for profit maximization of any kind in monetary system products. Especially when it comes to the banks and companies who offer loans, different variable costs occur in different loan types due to the customer's risk of paying back. Since there is a competitive environment, a trade-off between highly pricing the loans and losing the risky customers is highly concerning.

The focus of this paper is specifically on agricultural loans which differ from other loans in some aspects. Borrowers of this loan type are mainly farmers and farmers take these loans in order to plant their land, which in the end has a yield. This yield can be affected by the environmental factors over which borrowers don’t have any control. In a nutshell, the overall prices of agricultural loans depend not only on the risk segment of the customer but also the environmental factors such as the farming season type, the location of the farming land, and the plant type.

While building their pricing strategies, firms may have different objectives such as maximizing the profit, achieving a target sale, and minimizing the number of customers lost. The company providing data for this study wants to have a pricing strategy which guarantees to achieve a target sale in a certain period of time. The current procedure of the loan applications for this company is shown in Figure 1.

At first, customers, mostly farmers, want to purchase some products being used in the agricultural process. If customers can not afford the price of products, they apply for a loan, however, the application is not carried out directly through the company. The application is received by the dealer and forwarded to the company. The company uses customer data and credit scores to decide whether a customer is eligible for a loan, and offers an interest rate and an installment day for eligible customers. If the customer accepts the offer, the company makes the sale through the dealer, and the products are delivered to the customer. Customers do not directly borrow money.

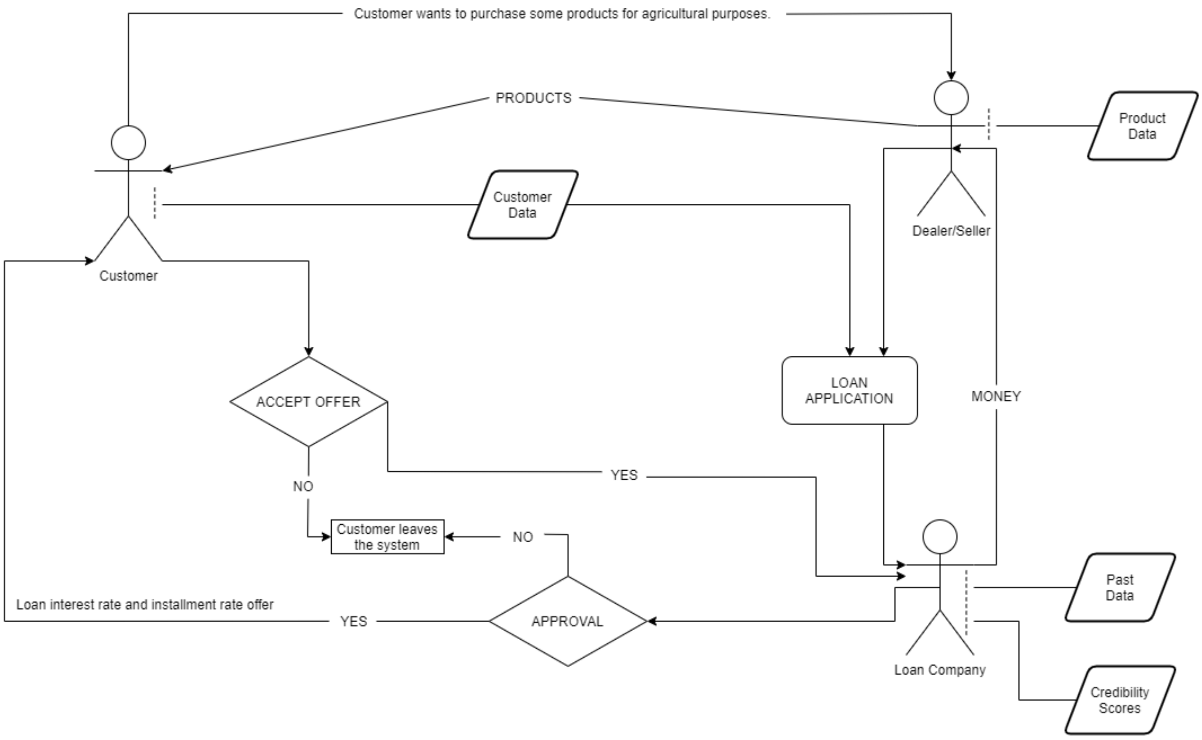

Figure 1: Flow Diagram

2.2 Requirements

In the current flow mechanism, loans' interest rates are decided based on the credibility scores of the customers. The main problem here is that there is no attempt either to achieve the targeted sale or maximize the profit in the pricing strategies. In order to have a satisfactory pricing strategy, the credibility risk of the customer, market risk, and competitor risk should be taken into consideration.

Credibility risk can be quantified by the details of customer profile and loan requirements. Payability of the previous loans, the characteristics of the products that are going to be purchased with the loan, potential price after the harvest, geographic and climatic conditions are some of the details that can be taken into account.

Market risk is defined as the possibility that investments experience losses due to external factors such as exchange rate, decrease or increase in the price of agricultural products, inflation, etc. Therefore, it is important to measure or predict the market risk in agricultural areas since it directly affects the profit.

Competitor risk is another significant element. Knowing the offers from other lenders is a key concept and should be considered so that possible customers won't end up as failures. There are two possibilities after a lender makes an offer: Success and Failure. To make a profit in the first place, possible customers must be convinced to accept the offer and resulting in success.

2.3 Limitations

As mentioned before, agricultural loans, by their nature, are loans that are dependent on external factors. When farmers take out the loan, they think that they will pay it back with the money they will get from the harvest. If the harvest is unproductive due to environmental factors such as weather conditions, they will have a hard time finding the money to pay it back. Likewise, in unstable economies like Turkey, it is difficult to predict product value at the due date. Even if the borrowers had a productive harvest, they may not have the money to pay back when the market drops. Or, the loan may lose its value more than the given interest rate due to reasons such as increasing inflation. Due to these factors that could be hard to predict sometimes, having the above-mentioned requirements in a perfectly shaped manner is quite challenging. Besides, most customers manage their money through more traditional methods, such as spending their own assets or borrowing money from people they know. This makes it difficult to draw a customer profile for unbank customers from their credit scores.

Competitor prices also play an important role in price determination. Most of the customers choose the best price after doing market research. However, it is not possible to reach the individual offers offered to customers. Therefore, it is never certain that the customer accepts the offered price because there might be a competitor giving a lower price, which we can’t know.

Above all, the provided data brings its own limitations to the study as well. First of all, all of the values in the data except the yearly rates are scaled. This makes it difficult to comprehend the meanings behind the variables. Secondly, there are lots of null values. Since data is scaled and the meanings of some variables are blurred, handling the null values is another challenge. In addition, the fact that the same customer has made a very small amount of transactions in the data makes it difficult to customize customer details. Lastly, the yearly interest rate is biased because it is the output of the pricing strategy of the company providing the data.

3. Analysis for Solution/ Design Methodology

3.1 Literature Review

There are many studies in the literature about the pricing strategies since this is a hot topic and most of the companies try to find the optimum pricing strategies for their products and services. Before studying this project, previous studies about the topic are looked up in order to get away from similar mistakes that might have been made in the past. In addition, it is a very useful practice to examine what has been done before and maybe get different perspectives and new ideas from there in terms of shedding light on the work. For these reasons, several studies have been reviewed.

In the article, Besbes and Zeevi (2015) considered a multiperiod single product pricing problem with an unknown demand curve under the stationary demand environment assumption. The seller's goal is finding the optimal price to maximize cumulative expected revenue by adjusting prices in each period. Although it is a dynamic optimization problem, the demand curve is not known by the seller. Besbes and Zeevi argued that the demand curve is modeled with a linear model. The demand curve could be represented with parameters according to the past pricing data. Then, the optimal price could be found, this strategy is known as learning and earning. Ban and Keskin (2020), investigated personalized dynamic pricing with learning as well. According to them, a personalized demand model could be constructed, and the relation between the product demand and customer features could be learned through previous sales observations. The seller should observe the high dimensional feature vector for each customer and learn the product demand according to these features. Then, the seller could employ this information to make personalized pricing for each customer and try to maximize revenue. The difference between the two articles is that Ban and Keskin (2020) used another approach to learn the demand curve.

Rothschild (1974) clearly revealed in his study that in most cases, what the demand model actually concerns can never be learned by the seller even if she/he sticks to the optimal pricing policy. This concept is later called one of incomplete learning by McLennan (1984). McLennan shows that optimal learning can be a seller's optimal choice in a model with discounting whereas, Harrison, Keskin, and Zeevi (2012) prove that this occurs generally due to myopic decision making, which chooses at each decision point the price that maximizes expected profit from the next sales opportunity, given the current (posterior) distribution over demand parameters, updating the posterior distribution as new response data accumulate.

Besbes and Zeevi (2009) approached this problem by using a classical statistics framework instead of leaning on dynamic programming. On the other hand, Levina et al. (2009) present a machine learning approach to dynamic pricing with model uncertainty, analyzing the case of strategic consumers. There are some extensions of the previously discussed models. Keller and Rady (1999) investigated pricing when the demand model changes over time. Bolton and Harris (1999) studied the strategic experimentation including multiple firms instead of a monopolist.

Zhijun Chen, Chongwoo Choe, Noriaki Matsushima (2020) claim in their study that customers who are loyal to a brand Y will buy brand X if and only if the price of firm Y is higher than the price of company X in a market which each firm has a target group and consumers information is available for both firms. The study focuses on market profit analysis of price offers for active and passive customers. They concluded that, in both cases, using more information about customers could make personalized pricing easier to implement and benefit consumers by increasing competition. A firm tends to ensure that its personalized price targeted consumers can compete with the rival's poaching price and its uniform prices will not be chosen by their active-targeted consumers.

Another approach is suggested by Arnoud V. den Boer and Bert Zwart (2014) which is updating the price with a determined interval around the mean of the previously generated prices. The idea is called controlled variance pricing. The price of the next iteration is determined by a function of the previous price and the demand against the selling price. At the end of the paper, the authors show that the parameters of the function approaches the true values, which makes the price converge to the optimal.

Besbes and Maglaras (2012), studied the above-mentioned pricing problem assuming that demand is known and stochastic. They set financial milestones to achieve at the end of the prespecified time period. A discrete review policy that updates prices at discrete points in time, and at each such point adjusts the price upward or downward according to a feedback signal that depends on how well the observed revenue and sales trajectories are tracking their milestones is suggested.

By the literature review, it is concluded that in personalized dynamic pricing there are some key points that should be considered. First of all, it is important to decide how the demand curve affects the offered price and the relationship between past pricing data and optimal price in demand should be investigated. Besides the demand curve in the market, the personalized demand curve is another important point. The personalized demand curve can be predicted by customer statistics. Moreover, it is a fact that there are predictions in customers' preferences; however, it is not possible to claim that they reflect customers’ exact preferences due to the uncertainty. Finally, the pricing strategy should include the competitors’ prices effect since there is a possibility of customer losses.

3.2 Assumptions and selected approaches

In this study, the preferences of customers throughout time are assumed to be stable. To learn the demand curve of the customers, logistic regression, model-based tree, unsupervised, and semi-supervised clustering methods are used. According to some parameters that will be later in this study, the logistic regression method is selected as the base method and the model-based tree is selected as the main method. After probabilistic price sensitivities are found, assuming that they give the correct results, a dynamic pricing method with financial milestones (Besbes and Maglaras, 2012) is used. Here a predetermined target sale is selected as the milestone, which was the aim of the company providing data. In addition to that controlled variance pricing, aggressive and defensive pricing methods are also used. Also in the simulation part, instead of estimating the number of customers that will come in the determined period of time, values of the test data is directly used due to the difficulties in estimation.

To be able to carry out this study, background in probabilistic and statistics, data mining and machine learning concepts are highly essential. Simulation knowledge for the dynamic pricing and evaluation parts is also needed.

4. Development of Alternative Solutions

4.1. Price Sensitivity Curves

The pricing strategy will be built according to the Besbes and Maglaras (2012). The pricing strategy will be personalized since outputs will be collected from data-driven models. In addition, price sensitivity for all the customers should be extracted. To build predictive models 2 different strategies will be followed: a base model with personalized but constant price sensitivity and models with personalized and customer-segmented according to price sensitivity.

4.1.1. Logistic Regression Base Model

Logistic Regression is the first approach that will be used in price sensitivity. This model gives outputs according to the features of the customers; therefore, it is personalized. However, customers are not segmented according to the price sensitivity curves, which means all customers are assumed to have the same price constant.

Logistic regression is a statistical model that models the probability event using the linear combination of independent variables. In this study, the information of whether a customer accepted the price or not is used as the target variable, and other features used as input variables. Thus, logistic regression estimates the coefficients in all the features. Among these parameters there is also a coefficient for the "price" variable, therefore it is used as price sensitivity parameter, and it is constant for all the customers.

Logistic regression is implemented on Python programming language using the "sklearn" package, more information can be found here. In the package, there are 3 parameters to tune: "penalty" which is the norm of the penalty, "C" inverse regularization strength and "solver" which is the algorithm to solve the optimization problem. Parameter tuning is performed with a 5-fold cross-validation process and the best parameters are chosen according to the Area Under the Curve classification performance measure. Best penalty, C, and solver parameters are "l2","0.1", and "liblinear" respectively.

4.1.2. Models with Price Sensitivity Segmentation

In this section, different ensembles of models are used. Primarily, customers are segmented according to their different sets of features, then logistic regression models are built on every segment of customers, by using only the "price" variable as input. This segmentation is made unsupervised, semi-supervised, and automatically within a model.

4.1.2.1. Unsupervised Customer Segmentation

In the unsupervised customer segmentation, k-means clustering algorithm is used and applied to all features of the customers except “price” and “completed”. Then, logistic regression models are constructed for all different segments of customers.

K-means clustering is an unsupervised clustering algorithm that partitions the unlabeled data to “k” different segments. It makes the clustering as an iterative process. The clustering could be applied with different “k” values; however, the best “k” value is chosen according to the clusters’ “inertia” score.

This ensemble method is implanted in the Python programming language using the “sklearn” package. Firstly, all the features except “price” and “completed”, are scaled between 0 and 1 to make the clustering algorithm work accurately. Then, 10 iterations k-means algorithm are applied with the “k” values between 2 and 21, and inertia scores of different “k” values are calculated. The best “k” value is chosen according to the elbow method which depends on the inertia graph of “k” values. The Elbow method is a rule of thumb which chooses the “k” value where the graph fractures as an elbow. The elbow shape of inertias of the Model is shown in figure 2.

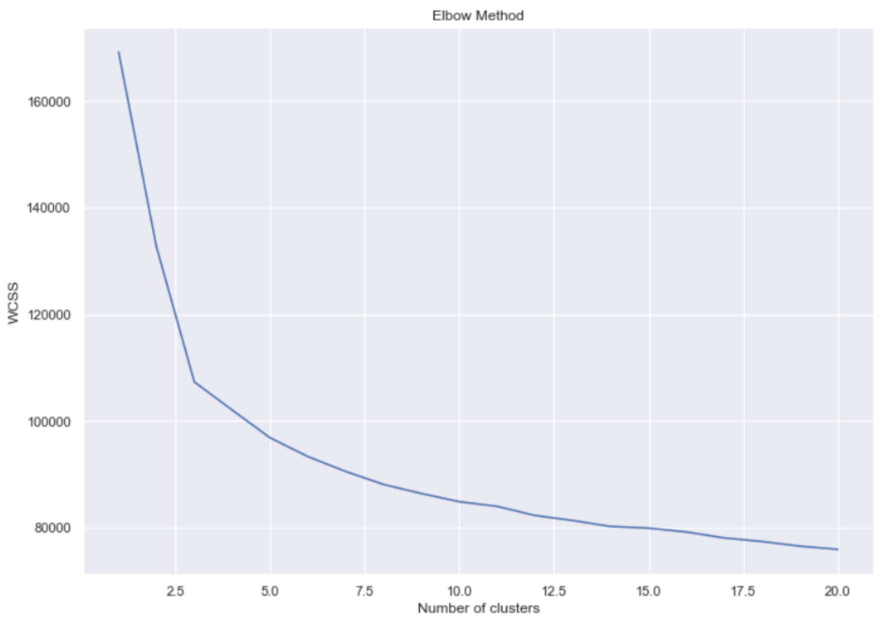

Figure 2: Inertia Scores of Unsupervised Customer Segmentation over K-values

The best “k” value is selected as 5 according to the graph shown in figure 2 above. Then logistic regression is implemented for all the segments from again the “sklearn” package, using only the “price” as the input, therefore 5 different price coefficients are obtained. Price parameters are -1.32, -0.36, -0.80, -0.07 and 1.38. Having a positive price coefficient means the probability of the acceptance is increasing while the price is increasing, therefore it can’t be accurate in real life and this segment is assumed to be not price sensitive at all. Finally, the overall model is used for the prediction of the test data, and the area under the curve score is calculated.

4.1.2.2. Semi-Supervised Customer Segmentation

In this section, the same approach is used as in section 4.1.2.1, the only difference is clustering is applied to only the important features of the customers. The important features are extracted with a supervised model which is called “LightGBM”, then “k-means” algorithm is applied on important features to segment the customers, finally logistic regression models are built for every segment of customers.

In this ensemble method, Python programming language is used with “sklearn” and “lightgbm” packages. LightGBM is a tree-based gradient boosting model which is implemented efficiently, more information about LightGBM can be found here. Although it requires a fine tuning, the model is used only for the feature importance extraction, so a LightGBM classifier is applied to the customer features without a “price” column, then features with more than 5% importance values are selected and shown in Table 1 below.

Table 1: Feature Importance of LightGBM

feature	**importance**
farmer_paid_amount	1.000
installment_day	0.304
com_total_limit	0.234
ind_acik_kredi_limit_kullanim_orani_all_23_26	0.187
ind_acik_kredi_limit_sum_23_26	0.182
month	0.147
farmer_payment_count	0.085
farmer_applications_unique_retailer_count	0.081
ind_acik_kredi_bakiye_sum_02_03_tk	0.076
birthyear	0.076
ind_last_account_creation_date_month_today_diff	0.075
ind_acik_kredi_bakiye_sum_23_26	0.073
com_nakdilimit_sum	0.065
ind_account_per_bank	0.062
farmer_open_accounts_total_risk_cash	0.057
com_nakdirisk_sum	0.055

The selected features are scaled, and the “k-means” algorithm is applied with k values between 2 and 21. Again, the elbow method, shown in Figure 3, is used to select the best “k” value.

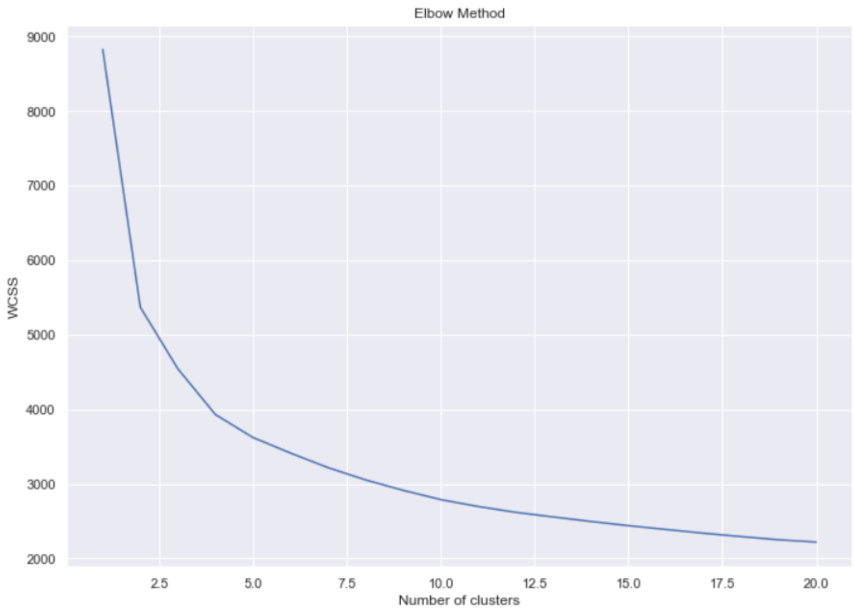

Figure 3: Inertia Scores of Semi-Supervised Customer Segmentation over K-values

The best “k” value is selected as 5. For every segment, logistic regression models are built with only the “price” feature as input. Price coefficients of the segments are: -1.34, -0.93, -0.20, 0.02, and 1.02. In this method, 2 non-sensitive segments are obtained. The overall predictive method is used to predict the test date and the area under the curve performance is recorded.

4.1.3. Model-Based Tree Model

Model-Based Tree models is another tree-based ensemble method, and it is very suitable for a price segmentation application. It partitions the with a decision tree, then it automatically builds logistic regression models. The decision tree algorithm it uses, makes the splits according to the performances of the logistic regression models.

In this section, a model-based tree model is implemented by using the R programming language with the “partykit” package. For this model, there are 2 parameters to tune: “minsize” which is the minimum size of the terminal nodes in the final tree, and “maxdepth” which is the maximum depth of the decision tree. These parameters are tuned with 5-fold cross validation and best “minsize”, and “maxdepth” parameters are 500 and 8 respectively according to the area under the curve score. Model based tree partitioned the data with all customers features except the “price” and built logistic regression models using “price” as the only feature. Customers are partitioned to 13 different segments and their price coefficients are -0.70, -2.20, -2.54, -2.27, 1.59, 0.12, -0.61, -1.09, -1.47, -2.14, -0.80, -1.32, -1.71.

4.1.4. Evaluation and Model Selection

All the models and methods explained in the previous section are compared based on the area under the curve performance score. The base logistic regression model has the 0.69 area under the curve score and it is selected because it is the only model without price segmentation. The models with the price sensitivity segmentation are unsupervised clustering, semi-supervised clustering and model-based tree models and their area under the curve scores are 0.52, 0.54, and 0.65 respectively. Thus, a model-based tree is chosen as the model with segmentation. In the following sections, base logistic regression model and the model-based tree model are compared with different pricing simulations.

4.2. Pricing Strategies

The last three weeks of the data is used as the test data when deciding optimal pricing strategy. It is assumed that the number of customers and the average flow rate of customers are known. The models use the sensitivity curves obtained by MOB and logistic regression models. It is also assumed that the sensitivity curve reflects the real demand curve. However, the probability of accepting offered price is bounded, since the results of sensitivity curves lose their meanings at extremes. Moreover, the offered price is also bounded to protect competitive advantage and profitability. The upper bound of price is determined as 200% of the consumer interest rates at the application date and the lower bound is 80% of the consumer interest rate at the application date. Four different pricing strategies are tried to find optimal pricing strategies which are Dynamic Pricing with Milestones, Aggressive Pricing, Defensive Pricing, and Controlled Variance Pricing. Logistic regression and MOB results are used for each of them. Moreover, three different target values are used in the Dynamic Pricing with Milestones approach. As a result, 12 different pricing strategy models are compared in terms of net money, profit, and total revenue.

4.2.1. Dynamic Pricing with Milestones

The figure 4 below shows the algorithm of dynamic pricing with milestones. At first, the target value which is the net money collected from the customers, the expected number of customers, and the average flow rate is given as an input to the model. When a customer arrives the required flow rate is calculated and equalized to the expected flow rate from the customer which is calculated by multiplying the probability of the customer accepting the offer and the average flow rate. This calculation returns a probability since the average flow rate per customer is constant. The probability is used in an inverse logit function by using sensitivity curves obtained before from MOB and Logistic Regression to find how much a customer is willing to pay. Then, a random number between 0 and 1 is generated. If the random number is smaller than the probability of acceptance, it is assumed that the customer accepts the offer and the actual net money paid by the customer is decreased from the target value. The Required Flow rate is calculated by dividing the remaining amount to achieve the target by the number of remaining customers.

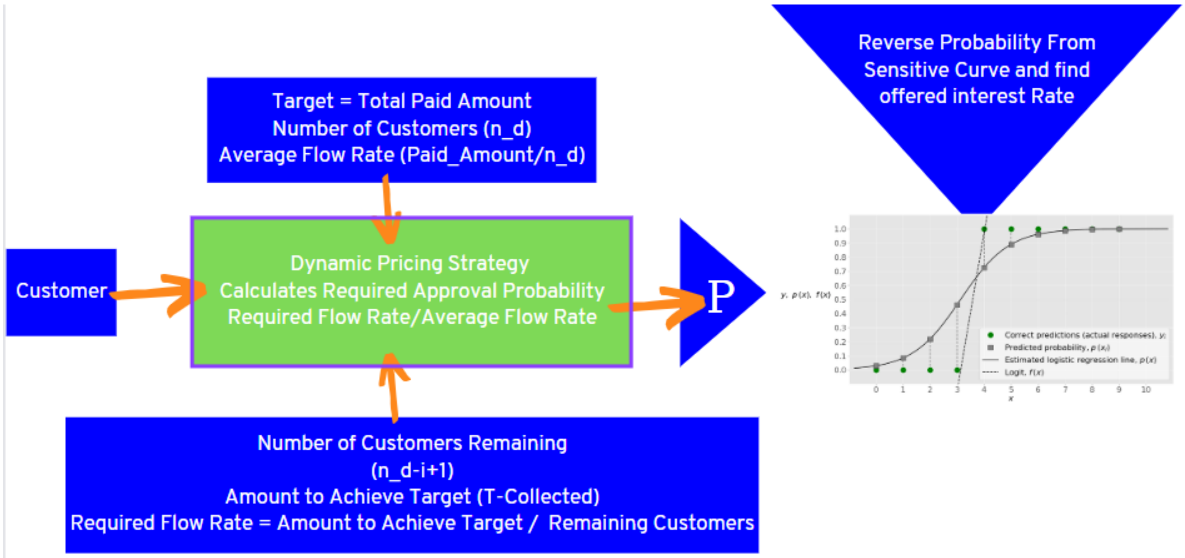

Figure 4: Algorithm of Dynamic Price Strategy with Milestones

There are some circumstances that should be considered when the models are built. As mentioned above, the sensitivity curve can return unreasonable price offers in extremes therefore if the required probability is higher than 0.95, it is assumed to be 0.95 or if the required probability is lower than 0.05, it is assumed to be 0.05. This method also prevents using negative probability or probability higher than 1. Another point is that the MOB model returns some customer segments with a positive coefficient which means that customer acceptance level is increased by price. Since it is not a realistic case, the price offered to these types of customers is equalized real offered price by the company to obtain more accurate results in comparison.

In dynamic pricing with a milestones approach, the offered price is mostly dependent on the target. If the target is higher than expected net money which is actual money collected by the company in that case, It is expected that the model offers lower prices with lower profitability but higher net collected money and if the target is lower than expected net money which is actual money collected by the company, It is expected that model should offer higher prices and increase profitability resulting in a low rate of acceptance.

4.2.2. Aggressive Pricing

The aggressive pricing strategy in this study refers to offering a price such that customers are willing to pay with the probability of 0.05. It is expected that customers' acceptance level will be lower but the profitability will increase. However, the upper bound of price is still valid for this model since the maximum relative accepted rate is 0.9813821 and there is competitors' risk.

4.2.3. Defensive Pricing

The defensive pricing strategy in this study refers to offering a price such that customers are willing to pay with the probability of 0.95. It is expected that customers' acceptance level will be higher but the profitability will decrease. However, the lower bound of price is still valid for this model to protect the company's profitability level.

4.2.4. Controlled Variance Pricing

The Controlled variance pricing, in this study, refers to offering a random interest rate to customers over a range calculated from past accepted offers. The mean of the accepted offered rates in training data is 0.4930381 and the standard deviation is 0.1003401. The interval is determined as [0.442868, 0.5432081]. The lower bound is determined as 0.5 standard deviation below the mean and the upper bound is determined as 0.5 standard deviation above the mean.

5. Comparison of Alternatives and Recommendations

5.1. Validation of Models

The Model-Based Tree and Logistic Regression give the best ROC values and they are selected to be used in pricing strategies. However, it is necessary to see that models give similar results to real life. Therefore, the simulation is prepared to see if they are valid or not.

In the simulation, the real offered prices are given to models and the predicted probability of customers accepting the price is calculated. Then, a random number is generated, if the random number is smaller than the predicted probability of customers accepting the price it is assumed that the customer accepts the offer. The simulation was repeated ten times and average acceptance rates, average net collected amount, average profit values, and average revenue were calculated. Table 2 and Table 3 show the detailed results of the simulations, while Table 4 shows the mean values and actual values of the simulations.

Table 2: Validation of the Model-Based Tree

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	82%	0.56	4650.63	2092.87	6743.50	45%
2	81%	0.56	4629.95	2081.79	6711.74	45%
3	81%	0.56	4631.86	2084.16	6716.02	45%
4	82%	0.56	4724.54	2124.80	6849.34	45%
5	81%	0.56	4649.34	2091.76	6741.10	45%
6	81%	0.56	4651.49	2091.31	6742.79	45%
7	80%	0.56	4601.42	2069.20	6670.62	45%
8	82%	0.56	4680.66	2105.41	6786.07	45%
9	81%	0.56	4664.81	2098.50	6763.31	45%
10	81%	0.56	4611.42	2075.07	6686.49	45%
Average	81%	0.56	4649.61	2091.49	6741.10	45%
Std	1%	0.00	35.43	15.88	51.29	0%

Table 3: Validation of the Logistic Regression

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	80%	0.56	4600.55	2069.39	6669.94	45%
2	81%	0.56	4598.14	2067.33	6665.47	45%
3	80%	0.56	4604.73	2071.64	6676.37	45%
4	82%	0.56	4695.82	2112.44	6808.25	45%
5	80%	0.56	4620.60	2079.30	6699.90	45%
6	81%	0.56	4640.20	2085.74	6725.94	45%
7	80%	0.56	4584.23	2061.85	6646.08	45%
8	81%	0.56	4647.04	2090.53	6737.57	45%
9	81%	0.56	4655.66	2094.71	6750.37	45%
10	80%	0.56	4572.43	2056.55	6628.98	45%
Average	81%	0.56	4621.94	2078.95	6700.89	45%
Std	1%	0.00	37.64	17.08	54.71	0%

Table 4: Validation Results

<table>
<tr><th colspan="2"></th><th>Acceptance Rate</th><th>Average Accepted Price</th><th>Collected Net Amount</th><th>Profit</th><th>Revenue</th><th>Profitability</th></tr>
<tr><td colspan="2">Actual Results</td><td>85%</td><td>0.45</td><td>4788.49</td><td>2153.50</td><td>6941.99</td><td>45%</td></tr>
<tr><td rowspan="2">Model-Based Tree</td><td>Average</td><td>81%</td><td>0.56</td><td>4649.61</td><td>2091.49</td><td>6741.10</td><td>45%</td></tr>
<tr><td>Std</td><td>1%</td><td>0.00</td><td>35.43</td><td>15.88</td><td>51.29</td><td>0%</td></tr>
<tr><td rowspan="2">Logistic Regression</td><td>Average</td><td>81%</td><td>0.56</td><td>4621.94</td><td>2078.95</td><td>6700.89</td><td>45%</td></tr>
<tr><td>Std</td><td>1%</td><td>0.00</td><td>37.64</td><td>17.08</td><td>54.71</td><td>0%</td></tr>
</table>

As shown in Table 4 above, models give very close the acceptance rate, collected net amount, and profit values on average, however, the average accepted price is close but not as much as expected. It can be said that models overestimate the accepted prices. When the overall results are considered, both models are valid for the pricing strategies.

5.2. Simulation Results of Pricing Strategies

By using the model-based tree and the logistic regression sensitivity curves, each pricing strategy is simulated ten times. The detailed results of each simulation are given below.

Table 5: Dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve:

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	86%	0.35	4759.07	1917.52	6676.59	40%
2	86%	0.36	4771.53	1953.34	6724.86	41%
3	85%	0.36	4771.38	1950.25	6721.63	41%
4	85%	0.37	4784.04	2039.99	6824.03	43%
5	85%	0.37	4782.33	2024.27	6806.60	42%
6	85%	0.36	4766.65	1943.19	6709.84	41%
7	84%	0.36	4747.60	1926.24	6673.84	41%
8	86%	0.36	4777.68	1949.48	6727.15	41%
9	86%	0.37	4783.07	2016.50	6799.57	42%
10	86%	0.35	4750.54	1880.86	6631.40	40%
Average	85%	0.36	4769.39	1960.16	6729.55	41%
Std	0%	0.01	13.29	51.06	63.06	1%

Table 5 shows the simulation results of dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve which gets the target value as the actual net amount of money collected by the company. The dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve results in an 85% average acceptance rate, and an average 41% profit. The average total revenue is 6729.55. The dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve manages to collect 99.6% of the target value on average if the target value equals the actual net amount of money collected by the company.

Table 6: Aggressive Pricing by Model-Based Tree Sensitivity Curve

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	72%	0.57	4171.72	2391.04	6562.77	57%
2	74%	0.57	4273.00	2449.97	6722.97	57%
3	74%	0.57	4272.65	2450.46	6723.11	57%
4	74%	0.57	4232.15	2427.67	6659.82	57%
5	73%	0.57	4178.67	2395.20	6573.88	57%
6	74%	0.57	4282.90	2455.32	6738.23	57%
7	73%	0.57	4213.96	2415.54	6629.49	57%
8	72%	0.57	4170.61	2392.52	6563.13	57%
9	74%	0.57	4232.92	2427.75	6660.67	57%
10	74%	0.57	4257.66	2441.31	6698.96	57%
Average	73%	0.57	4228.62	2424.68	6653.30	57%
Std	1%	0.00	43.51	25.07	68.57	0%

Table 6 shows the result of Aggressive Pricing by Model-Based Tree Sensitivity Curve. Aggressive Pricing by Model-Based Tree Sensitivity Curve results in a 73% average acceptance rate and an average 57% profit, and the average total revenue is 6653.3.

Table 7: Defensive Pricing by Model-Based Tree Sensitivity Curve

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	88%	0.27	5021.08	1498.40	6519.48	30%
2	88%	0.27	4980.28	1482.75	6463.03	30%
3	89%	0.27	5067.62	1507.90	6575.52	30%
4	88%	0.27	4990.37	1486.60	6476.97	30%
5	88%	0.27	4957.83	1477.48	6435.31	30%
6	88%	0.27	5023.86	1500.20	6524.06	30%
7	88%	0.27	4952.05	1471.73	6423.78	30%
8	87%	0.27	4957.34	1479.44	6436.79	30%
9	88%	0.27	4950.68	1474.11	6424.79	30%
10	88%	0.27	4965.95	1478.72	6444.68	30%
Average	88%	0.27	4986.71	1485.73	6472.44	30%
Std	1%	0.00	39.14	12.29	51.28	0%

Table 7 shows the result of Defensive Pricing by Model-Based Tree Sensitivity Curve. Defensive Pricing by Model-Based Tree Sensitivity Curve results in an 88% average acceptance rate and an average 30% profit, and the average total revenue is 6472.44.

Table 8: Controlled Variance Pricing by Model-Based Tree Sensitivity Curve:

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	74%	0.57	4272.39	2449.69	6722.09	57%
2	82%	0.43	4701.92	2038.68	6740.60	43%
3	81%	0.43	4690.15	2032.79	6722.94	43%
4	82%	0.43	4687.98	2031.80	6719.78	43%
5	82%	0.43	4699.59	2036.44	6736.04	43%
6	82%	0.43	4703.16	2037.51	6740.67	43%
7	83%	0.43	4732.43	2051.68	6784.10	43%
8	82%	0.43	4697.84	2037.00	6734.84	43%
9	83%	0.43	4735.27	2051.37	6786.64	43%
10	82%	0.43	4662.28	2018.66	6680.94	43%
Average	81%	0.45	4658.30	2078.56	6736.86	45%
Std	3%	0.04	137.21	130.75	30.86	4%

Table 8 shows the result of Controlled Variance Pricing by Model-Based Tree Sensitivity Curve. Controlled Variance Pricing by Model-Based Tree Sensitivity Curve results in an 81% average acceptance rate and an average 45% profit, and the average total revenue is 6736.86.

Table 9: Dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve when Target Increased by 5%

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	88%	0.31	4945.39	1694.95	6640.34	34%
2	87%	0.30	4900.93	1633.66	6534.59	33%
3	87%	0.31	4913.52	1692.07	6605.58	34%
4	87%	0.30	4880.14	1611.02	6491.16	33%
5	88%	0.31	4936.23	1701.19	6637.42	34%
6	88%	0.30	4924.17	1671.74	6595.91	34%
7	88%	0.31	4963.86	1726.11	6689.97	35%
8	88%	0.31	4940.02	1698.06	6638.08	34%
9	88%	0.31	4922.64	1701.59	6624.24	35%
10	87%	0.30	4867.03	1626.02	6493.05	33%
Average	87%	0.30	4919.39	1675.64	6595.03	34%
Std	1%	0.00	29.91	38.63	67.10	1%

Table 9 shows the simulation results of Dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve which the target is determined as 5% more than the actual net amount of money collected by the company. The dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve results in an 87% average acceptance rate, and an average 34% profit. The average total revenue is 6595.03. The dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve achieves to collect 97.84% of the target value on average if the target value equals 105% of the actual net amount of money collected by the company.

Table 10: Dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve when Target Decreased by 5%

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	81%	0.45	4549.65	2238.47	6788.12	49%
2	82%	0.42	4549.76	2122.96	6672.72	47%
3	81%	0.42	4548.52	2143.58	6692.10	47%
4	82%	0.45	4549.61	2210.53	6760.14	49%
5	82%	0.43	4549.73	2177.59	6727.31	48%
6	82%	0.43	4549.82	2173.39	6723.21	48%
7	82%	0.43	4551.55	2168.80	6720.34	48%
8	82%	0.45	4549.29	2218.31	6767.60	49%
9	82%	0.43	4550.41	2167.97	6718.38	48%
10	82%	0.41	4551.75	2111.36	6663.11	46%
Average	82%	0.43	4550.01	2173.29	6723.30	48%
Std	0%	0.01	0.99	40.80	40.45	1%

Table 10 shows the simulation results of Dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve which the target is determined as 5% less than the actual net amount of money collected by the company. The dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve results in an 82% average acceptance rate, and an average 48% profit. The average total revenue is 6723. The dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve achieves collecting 100% of the target value on average if the target value equals 95% of the actual net amount of money collected by the company.

Table 11: Dynamic Pricing with Milestones by the Logistic Regression Sensitivity Curve:

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	83%	0.33	4697.09	1811.71	6508.79	39%
2	83%	0.33	4706.39	1795.85	6502.24	38%
3	83%	0.33	4721.26	1827.82	6549.08	39%
4	84%	0.35	4769.44	1928.83	6698.27	40%
5	83%	0.33	4714.51	1804.03	6518.54	38%
6	83%	0.34	4723.72	1872.30	6596.03	40%
7	82%	0.33	4691.61	1787.65	6479.26	38%
8	83%	0.34	4747.71	1857.09	6604.80	39%
9	83%	0.34	4746.78	1880.09	6626.88	40%
10	82%	0.31	4671.51	1711.17	6382.68	37%
Average	83%	0.33	4719.00	1827.65	6546.66	39%
Std	1%	0.01	29.50	60.40	88.74	1%

Table 11 shows the simulation results of Dynamic Pricing with Milestones by the Logistic Regression Sensitivity Curve which gets the target value as the actual net amount of money collected by the company. Dynamic Pricing with Milestones by the Logistic Regression Sensitivity Curve results in an 83% average acceptance rate, and an average 39% profit. The average total revenue is 6546.66. The dynamic Pricing with Milestones by Model-Based Tree Sensitivity Curve achieves collecting 98.55% of the target value on average if the target value equals the actual net amount of money collected by the company.

Table 12: Aggressive Pricing by the Logistic Regression Sensitivity Curve

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	78%	0.58	4488.97	2596.93	7085.90	58%
2	79%	0.58	4503.09	2605.14	7108.23	58%
3	79%	0.58	4516.75	2613.05	7129.80	58%
4	80%	0.58	4595.65	2658.74	7254.39	58%
5	78%	0.58	4535.38	2623.86	7159.24	58%
6	79%	0.58	4539.78	2626.34	7166.12	58%
7	77%	0.58	4484.50	2594.05	7078.54	58%
8	79%	0.58	4567.11	2642.20	7209.31	58%
9	78%	0.58	4538.79	2625.77	7164.57	58%
10	78%	0.58	4484.90	2594.65	7079.55	58%
Average	79%	0.58	4525.49	2618.07	7143.56	58%
Std	1%	0.00	37.08	21.51	58.58	0%

Table 12 shows the result of Aggressive Pricing by the Logistic Regression Sensitivity Curve. Aggressive Pricing by the Logistic Regression Sensitivity Curve results in a 79% average acceptance rate and an average 58% profit, and the average total revenue is 7143.56.

Table 13: Defensive Pricing by the Logistic Regression Sensitivity Curve

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	84%	0.26	4749.60	1396.74	6146.34	29%
2	84%	0.26	4738.55	1394.68	6133.23	29%
3	84%	0.26	4760.94	1397.61	6158.56	29%
4	84%	0.26	4808.71	1411.47	6220.18	29%
5	83%	0.26	4740.78	1400.35	6141.13	30%
6	84%	0.26	4777.20	1401.74	6178.94	29%
7	83%	0.26	4727.94	1396.88	6124.82	30%
8	84%	0.26	4775.24	1408.08	6183.32	29%
9	84%	0.26	4803.89	1412.55	6216.44	29%
10	83%	0.26	4703.66	1387.71	6091.36	30%
Average	84%	0.26	4758.65	1400.78	6159.43	29%
Std	1%	0.00	33.25	7.87	40.71	0%

Table 13 shows the result of Defensive Pricing by the Logistic Regression Sensitivity Curve. Defensive Pricing by the Logistic Regression Sensitivity Curve results in an 84% average acceptance rate and an average 29% profit, and the average total revenue is 6159.43.

Table 14: Controlled Variance Pricing by the Logistic Regression Sensitivity Curve

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	81%	0.44	4676.08	2133.31	6809.39	46%
2	81%	0.44	4676.48	2132.91	6809.39	46%
3	81%	0.44	4672.94	2130.52	6803.46	46%
4	80%	0.44	4612.15	2103.73	6715.88	46%
5	82%	0.44	4662.74	2126.52	6789.25	46%
6	81%	0.44	4582.68	2090.50	6673.18	46%
7	79%	0.44	4575.40	2089.49	6664.89	46%
8	82%	0.44	4681.80	2135.02	6816.81	46%
9	81%	0.44	4637.86	2114.79	6752.65	46%
10	81%	0.44	4639.19	2117.00	6756.19	46%
Average	81%	0.44	4641.73	2117.38	6759.11	46%
Std	1%	0.00	39.71	17.50	57.20	0%

Table 15 shows the result of Controlled Variance Pricing by the Logistic Regression Sensitivity Curve. Controlled Variance Pricing by the Logistic Regression Sensitivity Curve results in an 81% average acceptance rate and an average 46% profit, and the average total revenue is 6759.11.

Table 16: Dynamic Pricing with Milestones by the Logistic Regression Sensitivity Curve when Target increased by 5%

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	84%	0.20	4779.82	1196.20	5976.02	25%
2	84%	0.20	4772.71	1224.96	5997.67	26%
3	84%	0.21	4798.75	1243.40	6042.15	26%
4	83%	0.20	4738.34	1172.07	5910.41	25%
5	85%	0.21	4799.42	1238.74	6038.16	26%
6	85%	0.20	4796.34	1212.54	6008.88	25%
7	85%	0.20	4828.03	1238.01	6066.04	26%
8	85%	0.21	4829.24	1256.01	6085.25	26%
9	85%	0.20	4801.38	1234.97	6036.35	26%
10	83%	0.20	4733.57	1176.43	5910.00	25%
Average	84%	0.20	4787.76	1219.33	6007.09	25%
Std	1%	0.00	32.54	29.02	60.06	0%

Table 16 shows the simulation results of Dynamic Pricing with Milestones by the Logistic Regression Sensitivity Curve which the target is determined as 5% more than the actual net amount of money collected by the company. The dynamic Pricing with Milestones by Logistic Regression Sensitivity Curve results in an 84% average acceptance rate, and an average 34% profit. The average total revenue is 6007.09. The dynamic Pricing with Milestones by Logistic Regression Sensitivity Curve achieves to collect 95.22% of the target value on average if the target value equals 105% of the actual net amount of money collected by the company.

Table 17: Dynamic Pricing with Milestones by the Logistic Regression Sensitivity Curve when Target decreased by 5%

Replication	Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
1	84%	0.33	4209.70	1657.63	5867.33	39%
2	82%	0.32	4092.85	1555.44	5648.29	38%
3	83%	0.32	4114.14	1570.86	5685.00	38%
4	83%	0.33	4145.24	1599.13	5744.36	39%
5	84%	0.33	4189.27	1626.15	5815.42	39%
6	82%	0.32	4140.87	1581.38	5722.25	38%
7	82%	0.33	4122.59	1578.22	5700.82	38%
8	84%	0.33	4161.20	1603.28	5764.48	39%
9	84%	0.33	4196.05	1620.73	5816.77	39%
10	82%	0.32	4130.56	1568.09	5698.65	38%
Average	83%	0.33	4150.25	1596.09	5746.34	38%
Std	1%	0.00	38.21	31.57	69.13	0%

Table 17 shows the simulation results of the dynamic pricing with milestones by the logistic regression sensitivity curve in which the target is determined as 5% less than the actual net amount of money collected by the company. The dynamic pricing with milestones by the logistic regression sensitivity curve results in an 83% average acceptance rate, and an average 38% profit. The average total revenue is 5746.34. The dynamic pricing with milestones by the logistic regression sensitivity curve achieves collecting 91.23% of the target value on average if the target value equals 95% of the actual net amount of money collected by the company.

The overall results of the simulations are given in Table 18 below. As seen in the table Aggressive Pricing strategy gives the best results in profit and revenue maximization. As mentioned above, models overestimate accepted prices. In the real world, it is not expected that 70% of customers accept such a high price offer in a competitive market. Besides, it also can be seen that the upper bound of price forces the give the same offer for everyone which is double the consumer interest rates at the application date. In that case, an aggressive pricing policy is not a suitable choice by considering a personalized price offer and a realistic perspective.

It is seen that a situation similar to aggressive pricing has emerged in defensive pricing as well. The lower bound of the price prevented offering lower prices and it lost personal pricing characteristics. In addition, although it is expected to collect a high amount with a low-profit rate in defensive pricing, it lagged behind other pricing strategies in terms of amount.

The Controlled Variance Pricing strategy is the second approach that gives the best results in revenue maximization. However, this approach does not offer a personalized price, and the range of the offered price in this approach is the result of the company's pricing strategy. Moreover, the controlled variance pricing does not offer any improvement to the existing pricing strategy and its success is very dependent on the correlation between historical data and future data. If the market circumstances change, it takes time for the controlled variance pricing to adjust itself.

As seen in the table, the Dynamic pricing strategy does not offer a constant price which means that it uses proper sensitivity curves and chooses the appropriate price level for customers with a 85% acceptance rate. Besides, the aim of the dynamic pricing with milestones is reaching a target sale level and as seen in the results dynamic pricing model manages to get a very close target level with 41% profitability. By considering acceptance rate, collected net amount, profitability, and the realistic results of dynamic pricing, further analysis is carried out with dynamic pricing with milestone strategy.

Table 18: Overall Results of Pricing Strategies

			Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
	Actual Results		85%	0.45	4788.49	2153.50	6941.99	45%
Model-Based Tree	Dynamic Pricing with Milestones	Average	85%	0.36	4769.39	1960.16	6729.55	41%
		Std	0%	0.01	13.29	51.06	63.06	1%
	Aggressive Pricing	Average	73%	0.57	4228.62	2424.68	6653.30	57%
		Std	1%	0.00	43.51	25.07	68.57	0%
	Defensive Pricing	Average	88%	0.27	4986.71	1485.73	6472.44	30%
		Std	1%	0.00	39.14	12.29	51.28	0%
	Controlled Variance Pricing	Average	81%	0.45	4658.30	2078.56	6736.86	45%
		Std	3%	0.04	137.21	130.75	30.86	4%
Logistic Regression	Dynamic Pricing with Milestones	Average	83%	0.33	4719.00	1827.65	6546.66	39%
		Std	1%	0.01	29.50	60.40	88.74	1%
	Aggressive Pricing	Average	79%	0.58	4525.49	2618.07	7143.56	58%
		Std	1%	0.00	37.08	21.51	58.58	0%
	Defensive Pricing	Average	84%	0.26	4758.65	1400.78	6159.43	29%
		Std	1%	0.00	33.25	7.87	40.71	0%
	Controlled Variance Pricing	Average	81%	0.44	4641.73	2117.38	6759.11	46%
		Std	1%	0.00	39.71	17.50	57.20	0%

Figure 5 shows the cumulative collected net amount behavior with customer arrivals, Figure 6 shows the profit over customer arrivals, and Figure 7 shows the cumulative revenue over customer arrivals to show the behavior of pricing strategies.

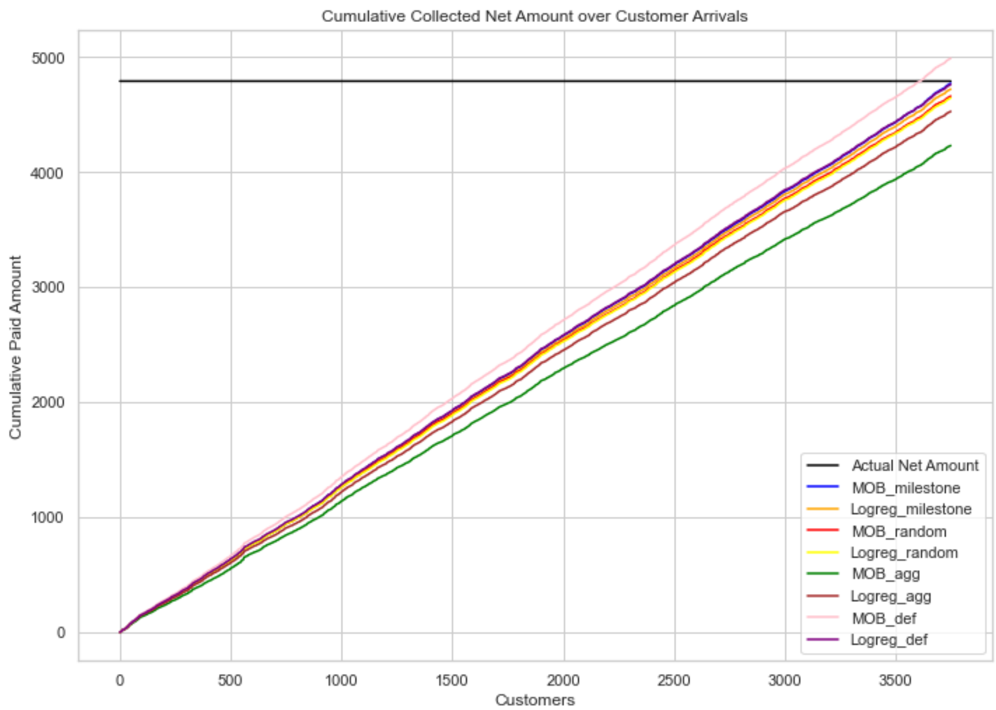

Figure 5: The Cumulative Collected Amount over Customers Arrivals with Different Price Strategies

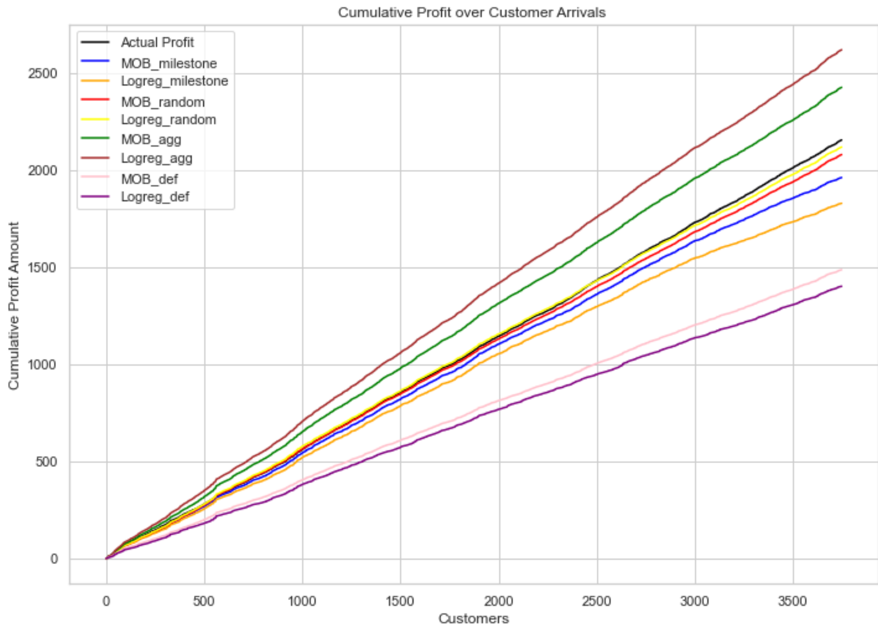

Figure 6: The Cumulative Profit over Customers Arrivals with Different Price Strategies

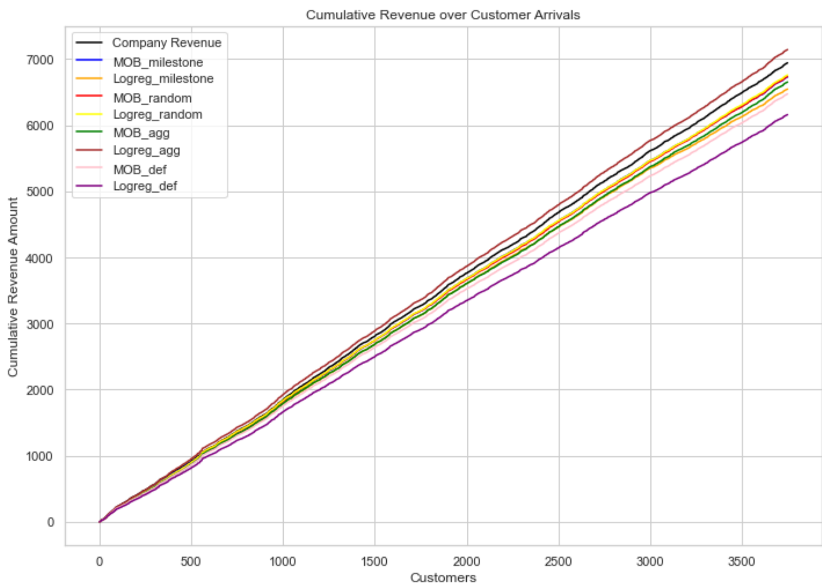

Figure 7: The Cumulative Revenue over Customers Arrivals with Different Price Strategies

Table 19 shows the average results of dynamic pricing with milestones strategy by using different target values. As seen in Table 19, dynamic pricing with milestones adjusts profitability by using a sensitivity curve to achieve a target. It is an important factor since the market is a very dynamic environment.

The dynamic pricing with milestones strategy gives better results by using model-based sensitivity curves as expected. Logistic regression gives one sensitivity curve for all customers and as seen below the range of the meaningful price is narrow, therefore it could not adjust itself as well as the algorithm using the model-based tree.

Table 19: Dynamic Pricing with Milestones in different Target Values

			Acceptance Rate	Average Accepted Price	Collected Net Amount	Profit	Revenue	Profitability
	Actual Results		85%	0.45	4788.49	2153.50	6941.99	45%
Dynamic Pricing with Milestones by using Model-Based Tree	100%	Average	85%	0.36	4769.39	1960.16	6729.55	41%
		Std	0%	0.01	13.29	51.06	63.06	1%
	105%	Average	87%	0.30	4919.39	1675.64	6595.03	34%
		Std	1%	0.00	29.91	38.63	67.10	1%
	95%	Average	82%	0.43	4550.01	2173.29	6723.30	48%
		Std	0%	0.01	0.99	40.80	40.45	1%
Dynamic Pricing with Milestones by using Logistic Regresssion	100%	Average	83%	0.33	4719.00	1827.65	6546.66	39%
		Std	1%	0.01	29.50	60.40	88.74	1%
	105%	Average	84%	0.20	4787.76	1219.33	6007.09	25%
		Std	1%	0.00	32.54	29.02	60.06	0%
	95%	Average	83%	0.33	4150.25	1596.09	5746.34	38%
		Std	1%	0.00	38.21	31.57	69.13	0%

Dynamic pricing with milestones can be implemented into the company's pricing strategies since it is a very common practice in the business world to determine targets and try to achieve them. Besides, budget planning can be simplified using this approach. Therefore, it is easy to implement in the business world.

There are many restrictions and limitations in the data which is mentioned above, however, the company has the opportunity to use unscaled, meaningful, and unbiased data. A further sensitivity analysis is needed to implement this project with more accurate data. If the sensitivity curve is better, the model will be more successful.

The expected customer number and the average flow rate were pre-determined by the examined test data since the predictions of future customers are not in the scope of the project. However, it is important to have accurate demand predictions to improve and implement this method.

The model has a capacity of adjusting itself to changes since in each iteration requirements are re-calculated and the milestones are given in a determined time horizon. At the end of the period, the results can be interpreted and in the long-run model works better.

Dynamic pricing model constructed with this study is robust and sustainable for the company providing data since it can be adjusted to the changes in the objectives very easily. As shown above, it can reach the target value at the end in all of the different situations.

6. Suggestions for a Successful Implementation

When a potential customer comes to the firm, the payment history with some personal information of the customer is examined in the current system. The agents of the firm make a risk assessment. The process for customers with a risk score below the threshold continues with the pricing part. The professionals from the firm consider the farmer's case and determine the price of the loan according to the firm's pricing policy. At this stage, the pricing model mentioned in this paper could be used to find the optimal price for a loan.

The farmer could be given the price of the loan according to her/his price sensitivity and how close the firm is to the target. As a result, the firm approaches its target by giving its customers a personalized price. The acceptance of this offer affects the total amount of money collected and the system continues with new customers waiting for their loans to be priced.

At the end of each period, the firm should check whether the target money is collected or not. If the target is reached, the profitability level of the period could be analyzed. If the firm is not satisfied with the obtained profit for three periods in a row, the price sensitivities of the customers can be revised. The price sensitivity coefficients of the farmers may be changed over time due to the economic developments in the country. The models can be retrained and the process can be progressed with new sensitivity curves.

7. Conclusion and Discussion

In this project, several IE tools, techniques, and methods are integrated to achieve results in a satisfactory level. For handling the null values, data mining tools, probabilistic and statistics have been used. Data preparation was quite challenging, and these tools are required to have a clean data set before putting it into models. Correlations were highly significant to have the data as much unbiased as possible. Since all the values were scaled and yearly interest rates were biased to the pricing strategy of the company providing data, statistics knowledge gathered through statistics-related courses helped a lot in detecting the correlations.

Machine learning concepts such as model-based trees, logistic regression, lightgbm, and clustering are used in determining the price sensitivities of the customers. Learning the true demand curve of the customers was important and necessary for later in the pricing part. Results from the logistic regression and model-based tree approaches are used for the pricing part. Former gives the demand curve in general, it doesn't segment the customers according to their profile and features, hence it is used as the base model. The latter, on the other hand, segments the customers according to their profile, thus giving a more reasonable demand curve for the different types of customers Hence it is used as the main model.

Systems Simulation knowledge is needed for the dynamic pricing and evaluation parts. Simulation is used also for the validation of the demand curves. The acceptance of the offered prices is simulated in a python code chunk using random generators. With the help of simulation, reaching the targeted value, collected profits, interest rate acceptance rate and more are evaluated. Jupyter Notebook with python and R languages is used in all of the steps throughout the project.

The output of this project has economic, environmental, ethical, and societal impacts. If we think about the economic effects, on the one hand, the profitability of the company increases with the suggested model and on the other hand, customers have economic benefits such that they are offered according to their profile so that the prices are affordable. When environmental effects are considered, policies of the company providing data have strong impacts. Since customers want to borrow products for agricultural purposes, the company suggests more productive and environmental products for sustainability. When ethical and societal impacts are considered, with the affordable customized prices, people borrow from reliable places instead of no-corporate sources of money.

8. References

Ban, G. Y., & Keskin, N. B. (2021). Personalized dynamic pricing with machine learning: High-dimensional features and heterogeneous elasticity. *Management Science*, *67*(9), 5549-5568.

Besbes, O., & Maglaras, C. (2012). Dynamic pricing with financial milestones: Feedback-form policies. *Management Science*, *58*(9), 1715-1731.

Besbes, O., & Zeevi, A. (2009). Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research*, *57*(6), 1407-1420.

Besbes, O., & Zeevi, A. (2015). On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. *Management Science*, *61*(4), 723-739.

Bolton, P., & Harris, C. (1999). Strategic experimentation. *Econometrica*, *67*(2), 349-374.

Chen, Z., Choe, C., & Matsushima, N. (2020). Competitive personalized pricing. *Management Science*, *66*(9), 4003-4023.

den Boer, A. V., & Zwart, B. (2014). Simultaneously learning and optimizing using controlled variance pricing. *Management science*, *60*(3), 770-783.

Harrison, J. M., Keskin, N. B., & Zeevi, A. (2012). Bayesian dynamic pricing policies: Learning and earning under a binary prior distribution. *Management Science*, *58*(3), 570-586.

Keller, G., & Rady, S. (1999). Optimal experimentation in a changing environment. *The review of economic studies*, *66*(3), 475-507.

Levina, T., Levin, Y., McGill, J., & Nediak, M. (2009). Dynamic pricing with online learning and strategic consumers: an application of the aggregating algorithm. *Operations research*, *57*(2), 327-341.

McLennan, A. (1984). Price dispersion and incomplete learning in the long run. *Journal of Economic dynamics and control*, *7*(3), 331-347.

Rothschild, M. (1974). A two-armed bandit theory of market pricing. *Journal of Economic Theory*, *9*(2), 185-202.

Sklearn.linear_model.logisticregression. scikit. (n.d.). Retrieved June 6, 2022, from https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html

Sklearn.linear_model.logisticregression. (n.d.). Retrieved June 6, 2022, from https://lightgbm.readthedocs.io/en/latest/