applied sciences

Article

Dynamic Pricing Method in the E-Commerce Industry Using Machine Learning

Marcin Nowak * and Marta Pawłowska-Nowak

Faculty of Engineering Management, Poznan University of Technology, 60-965 Poznan, Poland; marta.pawlowska-nowak@put.poznan.pl
* Correspondence: marcin.nowak@put.poznan.pl

Featured Application: The proposed dynamic pricing method provides an actionable tool for e-commerce managers aiming to enhance their pricing strategies with real-time adjustments based on customer behavior, competitor pricing, and market dynamics. Through this approach, e-commerce platforms can efficiently leverage machine learning to personalize pricing at an individual customer level, ultimately maximizing revenue opportunities and customer satisfaction. Beyond direct pricing optimization, this method can also be adapted as a robust decision support system for small to medium-sized online retailers. By integrating user-specific purchasing data, the model can help predict optimal pricing actions, enabling businesses to navigate highly competitive markets while maintaining profitability and customer loyalty.

Abstract: One of the key areas of contemporary marketing is the formulation of a pricing strategy, which is one of the four pillars of the traditional marketing mix. One way to implement this strategy is through dynamic pricing. It is currently gaining popularity in many industries for two reasons. Firstly, it is possible, easy, and cheap to collect information about transactions and customers. Secondly, machine learning mechanisms, for which these data are essential, are becoming widely available. The aim of this article is to propose a dynamic pricing method for the e-commerce industry. To achieve this goal, machine learning methods such as the Naive Bayes classifier, support vector machines (linear and nonlinear), decision trees, and the k-nearest neighbor algorithm were used. The empirical results indicate that the linear support vector machine achieved the highest accuracy (86.92%), demonstrating the model's effectiveness in classifying pricing decisions. This article aligns with two leading research trends in dynamic pricing: personalized dynamic pricing (the target model considers customer-related criteria) and the development of systems to assist managers in optimizing pricing strategies to increase revenues (using machine learning methods). This article presents a literature review on dynamic pricing and then discusses the machine learning methods applied. In the final part of this article, verification of the developed dynamic pricing method using real-world conditions is presented.

Keywords: dynamic pricing; pricing strategies; machine learning; e-commerce

Citation: Nowak, M.; Pawłowska-Nowak, M. Dynamic Pricing Method in the E-Commerce Industry Using Machine Learning. *Appl. Sci.* **2024**, *14*, 11668. https://doi.org/10.3390/app142411668

Academic Editor: Douglas O'Shaughnessy

Received: 7 November 2024
Revised: 28 November 2024
Accepted: 5 December 2024
Published: 13 December 2024

1. Introduction

Since the time of P. Kotler, marketing issues have addressed, either directly or indirectly, the marketing mix model, which deals with the creation and implementation of product, pricing, promotion, and distribution strategies [1]. This article addresses the issue of building a pricing strategy within the specific model of dynamic pricing. This model has been successfully adopted in many industries [2,3].

It appears that the popularity and utility of dynamic pricing issues will continue to grow [4]. This is due to the ability to collect diverse data on transactions across many industries. This article focuses on the e-commerce sector, where it is particularly easy to gather a range of information about customers and their behaviors and purchasing preferences. This trend is also supported by the dynamic development of machine learning (or, more broadly, artificial intelligence), for which the collected data serve as fuel, enabling accurate prediction, classification, and evaluation of phenomena and transactions in the modern economy.

Given the presented premises, the aim of this article is to propose a dynamic pricing method for the e-commerce industry. The developed method also serves as a decision support system for managers in the dynamic pricing model. With this system, a manager can receive recommendations on whether the price of a given product in the online store should be increased, decreased, or maintained at the current level. The developed method takes into account criteria from various perspectives—customer-related, enterprise-related, and market condition-related criteria. Effective pricing enables the optimization of outcomes both financially and in terms of product performance.

To achieve this goal, well-known machine learning methods were used, such as the Naive Bayes classifier, support vector machines (both linear and nonlinear), decision trees, and the k-nearest neighbor algorithm. The structure of this paper is as follows: Section 2 provides a brief literature review on dynamic pricing in marketing theory. Section 3 describes the applied dynamic pricing method using selected machine learning mechanisms. Section 4 presents the results of empirical research, verifying the method's effectiveness based on data containing the transaction history of 500 customers from a selected e-commerce store. Finally, Section 5 concludes the article with a summary and outlines future research directions.

2. Literature Review

Dynamic pricing can be understood as a pricing strategy that involves flexibly adjusting the prices of products or services in response to changing market conditions, demand, competition, and other external factors [5,6]. It is noted that dynamic pricing can have either a negative or positive impact from the consumer's perspective, depending on the quality of communication processes between the business and the customer [7]. However, when properly implemented, dynamic pricing leads to a win–win situation where both companies and consumers benefit compared to the alternative strategy of offering an optimal static price [8].

The beginning of intense scientific interest in this issue can be dated back to the 1970s [9], when numerous studies were conducted on the application of dynamic pricing concepts, especially in the airline and hotel industries [10,11]. Nowadays, the range of industries in which dynamic pricing strategies can be successfully applied is very broad [12,13]. Examples of the application of dynamic pricing models in various industries include

- the airline industry—setting prices based on factors such as demand for a particular flight, departure date, seat availability, etc. [14];
- the hotel industry—setting prices to maximize occupancy and revenue, taking into account criteria such as seasonality, local events, or last-minute bookings [15];
- the e-commerce industry—setting prices in real-time based on customer behavior, competition, and other market data [16];
- the transportation industry—setting prices to optimize revenue based on the demand and supply for transportation services at a given moment [17].

Narahari and colleagues [18] presented an overview of dynamic pricing models used in e-business. The authors discussed the roles of inventory-based, data-driven, auction-based, and machine learning models used to calculate dynamic prices to maximize revenue. In particular, the authors emphasized the importance of data-driven models, which enable real-time price adjustments based on the analysis of consumer behavior [18]. Saharan and colleagues [19] conducted a systematic review of the literature on dynamic pricing in intelligent transportation systems (ITSs). The authors analyzed how dynamic pricing can help manage traffic, reduce peak loads, and optimize vehicle routes, highlighting the environmental benefits of such strategies. The literature review showed that dynamic pricing in ITSs can lead to more efficient resource management and reduce air pollution and noise [19]. Zhao and Zhang [20] examined dynamic quality and pricing decisions in customer-intensive service systems that are evaluated online. The authors constructed a dynamic programming model to determine optimal quality and pricing strategies. The research showed that online reviews impact pricing and quality decisions, forcing providers to offer higher quality at a higher price for a smaller number of customers, especially when customer intensity is high [20].

With the development of technology and increased access to data on consumer behavior, dynamic pricing has gained popularity and become a significant tool in revenue management [21]. However, dynamic pricing also presents several challenges, including managing consumer perception, adjusting prices in real time, and integrating advanced technologies such as artificial intelligence. Goli and Haghighinasab [22] emphasized the need for further research on variables affecting dynamic pricing and consumer perceptions of price fairness [22]. The ethics of dynamic pricing practices constitute an extensive research program in itself [23,24].

Dynamic pricing, while beneficial for businesses in maximizing revenues, raises significant ethical concerns, particularly in terms of fairness and transparency. One major issue is price discrimination, where consumers are charged different prices based on their purchasing behavior or demographic data. This practice, although legally permissible, often elicits moral outrage from consumers who perceive it as unfair. Ethical challenges arise when pricing algorithms exploit consumer behavior or vulnerabilities, prompting calls for greater scrutiny and regulation [25]. Additionally, dynamic pricing in essential sectors like electricity has faced criticism for potential inequities, as time-of-use rates may disproportionately impact low-income consumers who are less flexible in their consumption patterns [24].

From a consumer perspective, the ethicality of dynamic pricing depends heavily on perceived fairness and transparency. Research suggests that consumers often view dynamic pricing strategies negatively if they feel deceived or if the pricing lacks a clear justification [26]. Transparency in pricing mechanisms and ensuring that customers are aware of how and why prices vary can mitigate ethical concerns. Moreover, the introduction of algorithms and artificial intelligence in pricing decisions has amplified the ethical debate, with concerns over privacy violations and the potential manipulation of consumer behavior [18]. Addressing these ethical dilemmas requires robust regulatory frameworks and adherence to ethical business practices that prioritize consumer trust and equity.

Dynamic pricing models can be classified based on the pricing mechanism used. According to this categorization, several types of models can be distinguished. The first are rule-based models, which rely on predefined rules and criteria established by the company. Prices are adjusted according to specific principles, such as the time of day, day of the week, inventory levels, or historical sales data [27]. The next are algorithm-based models. These models use mathematical and statistical algorithms to predict demand and optimize prices. The algorithms can consider various variables, such as sales trends, competition, weather conditions, and special events [28].

The literature also identifies dynamic pricing models that use machine learning. These models analyze large datasets and predict future demand. In recent years, various dynamic pricing models have been developed, finding applications across different industries and decision-making contexts:

1. Dynamic Toll Pricing Model: This model focuses on the design, simulation, implementation, and evaluation of dynamic toll pricing systems. The goal is to optimize traffic flow by adjusting toll rates in real time, leading to the more efficient management of transportation infrastructure [29].
2. Dynamic Pricing with Demand Learning and Reference Effect: In this approach, sellers learn demand patterns based on historical sales and pricing data while also considering consumer reference effects. This allows for dynamic price adjustments to maximize profits while accounting for changing customer preferences [30].
3. Deep Reinforcement Learning-Based Dynamic Pricing Strategy: This model uses deep reinforcement learning algorithms to automatically adjust prices in real time. The goal

is to optimize inventory and revenues in pre-sale environments, taking into account demand uncertainty and other market factors [31].

Each of these models offers a unique approach to dynamic pricing, tailored to specific market conditions and business objectives.

The algorithms learn from historical data and adjust prices based on patterns detected in the data [32]. Another group of models is demand-based models. These models adjust prices in response to changing demand. Prices increase when demand is high and decrease when demand is low. The goal is to maximize revenue by leveraging demand fluctuations.

Using this classification, another group can be identified: price elasticity models [33]. These models analyze the price elasticity of demand, which is how changes in price affect the demand for a product. Prices are adjusted to achieve the optimal point where revenues are maximized.

In his systematic literature review on the dynamic pricing process, Neubert [7] identified eight leading research trends, which include the following:

1. Moderating factors influencing the impact of dynamic pricing on consumer behavior.
2. Strategic purchasing behavior in response to dynamic pricing.
3. The effect of dynamic pricing on customers' perception of fairness.
4. Personalized dynamic pricing (PDP).
5. Channel-differentiated pricing.
6. The effectiveness of various types of communication used by companies.
7. The role and significance of consumers' personal characteristics in their perception of price changes.
8. Practical implications for managers and researchers to optimize pricing strategies for increased revenue.

Given the stated objectives, this article aligns with two key research trends in dynamic pricing: personalized dynamic pricing (the target model considers customer-related criteria) and the development of systems to support managers in optimizing pricing strategies to increase revenue (using machine learning methods).

Table 1 presents a summary of the strengths and limitations of selected dynamic pricing models.

Table 1. Summary of dynamic pricing models: strengths and limitations.

Model	Strengths	Limitations
Dynamic Toll Pricing Model	– efficient traffic management through real-time toll adjustments – optimization of road infrastructure via precise traffic flow control	– focused solely on the transportation sector – limited applicability to other economic sectors
Dynamic Pricing with Demand Learning and Reference Effect	– incorporates historical sales data and reference effects for more accurate pricing decisions – ability to adapt prices to changing consumer preferences	– high data requirements and the need for the continuous monitoring of customer preferences – potential ethical concerns regarding the use of consumer data
Deep Reinforcement Learning-Based Dynamic Pricing Strategy	– automation of real-time pricing adjustments – effective handling of demand uncertainty	– high computational demands, especially with large datasets – potential lack of interpretability in model decisions (black-box nature)

Table 1. *Cont.*

Model	Strengths	Limitations
Dynamic Pricing for Online Ticket Systems (DPRank Model)	– utilizes surrogate price elasticity models for accessible dynamic ticket pricing – enhances understanding of demand variability	– primarily designed for online ticketing systems, limiting broader applicability – requires accurate demand elasticity estimation, which can be complex
Dynamic Quality and Pricing Decisions in Customer-Intensive Services	– considers online reviews to inform quality and pricing strategies – encourages higher quality offerings at adjusted prices based on customer intensity	– dependent on the availability and reliability of online customer feedback – may lead to higher prices, potentially reducing accessibility for some customer segments

Source: own elaboration.

This literature review discusses key dynamic pricing models, highlighting their applications across various sectors such as transportation, e-commerce, and services. It identifies strengths, such as precise real-time price management, and limitations, including dependence on large datasets, lack of interpretability in model decisions, and sector-specific constraints.

Current dynamic pricing models are often focused on specific industries or lack flexibility in adapting to diverse market conditions. Additionally, many of these models rely on algorithms that are difficult to interpret, which may limit their usefulness in managerial decision-making. There is also a lack of comprehensive models that integrate customer, business, and market criteria while remaining scalable and adaptable across industries; this represents a research gap that the present article seeks to address.

The proposed machine learning-based dynamic pricing method addresses the identified research gap by offering

- a comprehensive approach that incorporates three perspectives: customer, business, and market;
- flexibility to adapt to various economic sectors;
- transparency through the use of interpretable algorithms (e.g., decision trees), enhancing managerial trust in decision-making;
- scalability, enabling its application to both small and large datasets with appropriate technical infrastructure.

This method not only fills the gap in the literature but also introduces innovative aspects that enhance its practical value, addressing the evolving demands of managers in a dynamically changing e-commerce environment.

3. Materials and Methods

The method developed by the authors aimed to inform whether the initial price of a product should be lowered, maintained at the same level, or increased. The proposed dynamic pricing method thus served as a decision support system for managers in making pricing decisions for individual products. The developed method can be presented as a procedure consisting of four steps.

Step 1. Determine the list of dynamic pricing criteria in the e-commerce industry for the machine learning model and prepare the database.

The literature identifies several criteria that may influence price within the dynamic pricing concept [12,34]. These criteria were classified into three groups: customer-related, enterprise-related, and market-related. This classification was purposeful as it reflected different perspectives, allowing for a comprehensive understanding of the factors influencing price, considering customer expectations, strategic enterprise goals, and market dynamics. The list of criteria was been adapted to the capabilities and needs of e-commerce store managers.

All values that would eventually be included in the machine learning database needed to cover a uniform period. It was proposed that this period be 100 days (although different periods can be considered). The proposed set of criteria is presented in Table 2.

Table 2. Set of criteria for the dynamic pricing method for the e-commerce industry.

Criterion	Name	Remarks
	Customer criteria	
C_1	Number of transactions	in the last 100 days by a given customer
C_2	Average transaction value	in the last 100 days by a given customer
C_3	Number of complaints	in the last 100 days by a given customer
C_4	Time since last transaction	in the last 100 days by a given customer
	Enterprise criteria	
C_5	Inventory level	of the given product
C_6	Profit margin	of the given product
C_7	Storage cost	of the given product
	Market criteria	
C_8	Price relative to competitors	of the given product
C_9	Seasonality	of the given product
C_{10}	Demand dynamics	of the given product

Source: own elaboration.

The values for criteria C_1 to C_4 were stored in the database of the owner/manager of the selected e-commerce store. They were collected for users who had an account for the selected e-commerce store. The values for enterprise-related criteria were gathered based on any enterprise management system that contained information about inventory levels for individual products (a percentage scale for each product can be used), specified product margins (sale price–purchase price), and included storage costs for individual products (on any scale, e.g., semantic).

The price relative to competitors could either be determined subjectively (e.g., on a semantic scale) or quantitatively through automatic price analysis of competitors on well-known e-commerce portals (software could be developed to automatically find the prices of competing products from other manufacturers without significant difficulty). Product seasonality was best defined in binary terms. If a product was in season, it was assigned a value of 1 (e.g., Christmas trees in December). If it was not, it was assigned a value of 0.

Demand dynamics refers to the growth in sales of given products on the market. The value of this criterion could be determined in several ways. First, software could be developed to automatically analyze the demand for a given product on selected e-commerce portals (e.g., by reading the number of offers for a given product). Second, demand dynamics could be determined subjectively by experts (e.g., on semantic scales). Third, if the first two solutions were not feasible, the demand dynamics could be analyzed by examining the sales dynamics of the given product within the company/online store itself.

Next, a database needed to be developed consisting of two elements. The first element was the values for the various criteria for a certain number of transactions. The literature indicates that such a database should contain no fewer than 50 records. However, it can be generally assumed that the more records the database contains, the better. The second element of the database consisted of labels. For each record, an expert needed to determine whether the product price should be lowered, kept the same, or raised. This was essential work because it provided the training material for the machine learning models. The model will ultimately support managers in making decisions, but it must first learn from managers' own behavior.

Step 2. Selection of machine learning models for analysis in the dynamic pricing method for the e-commerce industry.

In this step of the method, machine learning models for further analysis needed to be selected. The proposed method suggested choosing popular models, namely, the Naive Bayes classifier, linear and nonlinear support vector machines, decision trees, and the k-nearest neighbor algorithm. These algorithms were used because, in addition to their popularity, each has unique properties that allow for diverse approaches to data analysis. The Naive Bayes classifier is simple and efficient with large datasets, making it a good baseline for comparison. Support vector machines (linear and nonlinear) are robust and effective with datasets containing a large number of features, with linear SVMs being suitable for linearly separable data and nonlinear SVMs capable of capturing complex relationships through kernel functions. Decision trees offer transparency and ease of result interpretation, making them valuable for managerial insights and modeling non-linear relationships in dynamic pricing scenarios. The k-nearest neighbor algorithm allows for the flexible modeling of relationships in the data, though it can be computationally intensive with larger datasets.

The selection of these models was also driven by the need to balance interpretability, computational efficiency, and predictive accuracy, as well as their alignment with the nature of the dataset and the problem. Naive Bayes provides a straightforward and computationally efficient approach, while decision trees offer clear decision paths that can be readily understood by managers. SVMs provide robustness and versatility for both simple and complex datasets and k-NN serves as a benchmark for evaluating the performance of more advanced models.

The selected dataset, consisting of 500 transactions from an e-commerce store, featured a combination of numerical, categorical, and ordinal variables, making these models particularly suitable. For example, the simplicity of Naive Bayes facilitated the rapid processing of the dataset, while SVMs could effectively handle the multidimensional nature of the features, ensuring robust classification. Decision trees were especially valuable for interpreting how specific criteria—such as storage costs or price relative to competitors—influenced pricing decisions, making them ideal for managerial applications. Meanwhile, k-NN allowed for the flexible modeling of patterns within the data, serving as a baseline for identifying the strengths and weaknesses of more complex algorithms.

During the implementation of the method, the most effective algorithm for the given database was selected. This diversity of models enabled a comprehensive evaluation of the dynamic pricing framework across different machine learning paradigms, ensuring that the final choice was well suited to the specific requirements of the application.

Future work could expand on these foundations by exploring ensemble methods, such as random forests or gradient boosting, or more advanced techniques like neural networks or deep learning. These methods could enhance predictive accuracy and adaptability, particularly for larger and more complex datasets or for dynamic pricing scenarios with unstructured data inputs.

Step 3. Implementation of the learning process for the selected machine learning algorithms.

At this stage, the models were trained on training datasets using specific learning algorithms appropriate for each machine learning method. The goal was to achieve the best possible fit to the training data while adhering to predetermined hyperparameter constraints. The models were exposed to the training data to learn patterns and relationships within the dataset. Then, the models' effectiveness was evaluated on test data. Trained models were validated on an independent dataset to assess how well they performed on new examples. The evaluation process helped determine the models' ability to generalize beyond the training set and provided valuable insights into their performance. The literature offers many metrics for evaluating classification accuracy [35]. One of the most popular was accuracy, which is described by Formula (1).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

where

Accuracy—accuracy of the machine learning model,
TP—True Positives (value correctly classified as True),
FP—False Positives (value not correctly classified as True),

TN—True Negatives (value correctly classified as False),
FN—False Negatives (value not correctly classified as False).

Various metrics, such as Precision, Recall, and F1 Score, are commonly used to evaluate machine learning models. However, accuracy was chosen for this study as it is one of the most widely recognized and utilized measures in machine learning, particularly when dealing with balanced datasets. Its simplicity and broad applicability make it a standard choice for initial performance evaluation. Nonetheless, the use of other metrics is not precluded. Depending on the specific objectives of future implementations, additional measures such as Precision, Recall, and F1 Score can be calculated to provide a more comprehensive assessment of the model's performance, particularly in cases of imbalanced datasets or where specific types of errors carry different weights.

Step 4. Selection of the best machine learning model for dynamic pricing in the e-commerce industry and its potential optimization.

In the fourth and final stage of the method, the effectiveness of different algorithms was evaluated, and the one that achieved the best results was selected. At this phase, it was recommended to use the most commonly applied metric, which was accuracy. After selecting the best model, if its performance did not meet expectations, for example, if it did not achieve an accuracy of 70%, optimization could be undertaken. Optimization of the algorithm involved modifying selected hyperparameters to maximize its effectiveness, which could be achieved using available libraries such as scikit-learn, a machine learning library for Python (3.10 version).

4. Results

Step 1. Determine the list of dynamic pricing criteria in the e-commerce industry for the machine learning model and prepare the database.

In the target model for dynamic pricing in the e-commerce industry, the criteria set described in the Methods section was chosen. To develop the database, data obtained from administrators of a selected Polish e-commerce store were used. The data covered the period from January to April 2024, spanning one hundred days. For the purposes of this article, the data were slightly rescaled to ensure the anonymity of all parties involved in the commercial transactions. Five hundred transactions were analyzed. Table 3 presents five sample records from the database containing raw data.

Table 3. Sample records from the target machine learning database.

	C_1	C_2	C_3	C_4	C_5
o_1	3	368.73	0	23	78%
o_2	8	103.75	0	18	92%
o_3	1	520.42	0	67	59%
o_4	4	46.03	1	20	83%
o_5	2	274.94	0	7	91%
	C_6	C_7	C_8	C_9	C_{10}
o_1	21%	high	very high	0	high
o_2	29%	medium	low	0	high
o_3	43%	medium	low	1	medium
o_4	18%	very low	medium	0	low
o_5	32%	high	medium	0	high

C_1—number of transactions, C_2—average transaction value, C_3—number of complaints, C_4—time since last transaction, C_5—inventory level, C_6—profit margin, C_7—storage cost, C_8—price relative to competitors, C_9—seasonality, C_{10}—demand dynamics. Source: own elaboration.

The labels for the specified sample of five customers were as follows: 1, 0, −1, 1, and 1. A value of −1 indicated a recommendation to lower the current price, 0 indicated keeping the price unchanged, and a value of 1 indicated raising the price. Storage costs, price relative to competitors, and demand dynamics were determined by experts using a five-point semantic scale. The choice of this scale was guided by its simplicity and widespread use in both research and managerial practice. Such scales are easy to understand and intuitive, which supports their effective application by both experts and end users. We acknowledge that more precise segmentation using a seven-point or nine-point scale could potentially enhance the model's accuracy. However, the scale can be easily adjusted in future implementations of the method if users find this more suitable for their needs. Such adjustments would not affect the overall structure or functionality of the proposed dynamic pricing model. Next, the variables were normalized. The results of this normalization for the sample values from the database are presented in Table 4.

Table 4. Sample of the normalized target machine learning database.

	C_1	C_2	C_3	C_4	C_5
o_1	0.333	0.316	0.000	0.230	0.780
o_2	0.889	0.077	0.000	0.180	0.920
o_3	0.111	0.453	0.000	0.670	0.590
o_4	0.444	0.025	0.500	0.200	0.830
o_5	0.222	0.231	0.000	0.070	0.910
	C_6	C_7	C_8	C_9	C_{10}
o_1	0.214	0.750	1.000	0.000	0.750
o_2	0.293	0.500	0.250	0.000	0.750
o_3	0.428	0.500	0.250	1.000	0.500
o_4	0.177	0.000	0.500	0.000	0.250
o_5	0.321	0.750	0.500	0.000	0.750

Source: own elaboration.

The target database contained 500 records.

Step 2. Selection of machine learning models for analysis in the dynamic pricing method for the e-commerce industry.

According to the description presented in the Methods section, the following machine learning algorithms were selected for dynamic pricing in the e-commerce industry: the Naive Bayes classifier, linear and nonlinear support vector machines, decision trees, and the k-nearest neighbors algorithm. Table 5 presents the range of the most important hyperparameters for the applied machine learning models.

Table 5. Basic hyperparameters of selected machine learning algorithms in dynamic pricing for the e-commerce industry.

Machine Learning Mechanism	Basic Hyperparameters
Naive Bayes Classifier	– alpha: Laplace smoothing parameter, helps to avoid issues with zero frequencies. – fit_prior: specifies whether to learn class prior probabilities or not. – class_prior: specifies prior probabilities of classes (if not set, they are calculated from the data).
Linear and Nonlinear Support Vector Machines	– C: regularization parameter that controls the classification cost and influences the margin width. – kernel: choice of the kernel function for data transformation (e.g., 'linear', 'poly', 'rbf'). – gamma: parameter for the kernel function (only for certain kernels like 'poly', 'rbf'). – degree: the degree of the polynomial for polynomial kernel functions (only for 'poly'). – coef0: the free parameter in the kernel function (only for 'poly' and 'sigmoid').
Decision Trees	– max_depth: maximum tree depth. – min_samples_split: minimum number of samples required to split an internal node. – min_samples_leaf: minimum number of samples required in an internal node's leaf. – criterion: split criterion for internal node (e.g., 'gini' or 'entropy'). – max_features: maximum number of features considered during the split.

Table 5. *Cont.*

Machine Learning Mechanism	Basic Hyperparameters
K-Nearest Neighbor Algorithm	– n_neighbors: number of neighbors considered for classification. – weights: choice of weight for each neighbor (e.g., 'uniform' or 'distance'). – algorithm: choice of algorithm used to compute nearest neighbors (e.g., 'auto', 'ball_tree', 'kd_tree', 'brute'). – p: power parameter for the Minkowski metric (default 2 for Euclidean distance).

Source: [36].

Step 3. Implementation of the learning process for the selected machine learning algorithms.

The dataset was divided into two subsets, with the training set containing 375 records (75% of the data) and the test set containing 125 records (25% of the data). The 'train_test_split' function from the Python scikit-learn library, dedicated to machine learning, was used for the split. The division was random. This process was then repeated one hundred times according to the cross-validation mechanism. For each of these one hundred random splits, the training process was conducted for all analyzed algorithms. Then, the average accuracy metrics for each algorithm were calculated. The results of the training for each algorithm are presented in Table 6.

Table 6. Performance results of selected machine learning algorithms in dynamic pricing for the e-commerce industry.

Algorithm	Naïve Bayes Calssifier	Linear Support Vector Machine	Nonlinear Support Vector Machine	Decision Trees	K-nearest Neighbor Algorithm
Average accuracy from one hundred training processes	0.7654	0.8692	0.8432	0.7464	0.6514

Source: own elaboration.

The linear support vector machine (SVM) algorithm ranked first in terms of average accuracy from a hundred training processes, with a score of 0.8692. The high accuracy of this algorithm indicated its ability to effectively classify data and make accurate pricing decisions in the e-commerce context. Due to its simplicity and efficiency in linearly separable spaces, the linear SVM can be a preferred choice in many dynamic pricing scenarios.

The nonlinear support vector machine (SVM) ranked second with a score of 0.8432, suggesting that more advanced, nonlinear SVM models can also be effective in dynamic pricing. Although their accuracy was slightly lower than the linear version, nonlinear SVMs could handle more complex and nonlinear relationships in the data better. The Naïve Bayes classifier took third place with a score of 0.7654. Despite its simplicity and assumption of feature independence, this algorithm showed satisfactory effectiveness, which may have resulted from its ability to quickly process large amounts of data and its good generalization. The decision trees algorithm achieved an average accuracy of 0.7464, indicating moderate effectiveness in dynamic pricing. While decision trees are easy to interpret and can capture complex relationships between features, their tendency to overfit can limit their practical effectiveness. The k-nearest neighbor (KNN) algorithm showed the lowest average accuracy with a score of 0.6514. The low effectiveness of this algorithm may have been due to its susceptibility to noise in the data and its requirements for a large amount of memory and computational time, especially with the large datasets typical of e-commerce.

To summarize, the results indicate that the linear SVM and nonlinear SVM are the most effective algorithms for dynamic pricing in the e-commerce industry, while Naïve Bayes, despite its simplicity, can also be a useful tool. Decision trees and KNN show lower effectiveness, suggesting the need for further research and optimization of these algorithms or their possible replacement with more advanced methods in commercial applications.

The histogram distribution of the accuracy metric for the linear support vector machine model is presented in Figure 1.

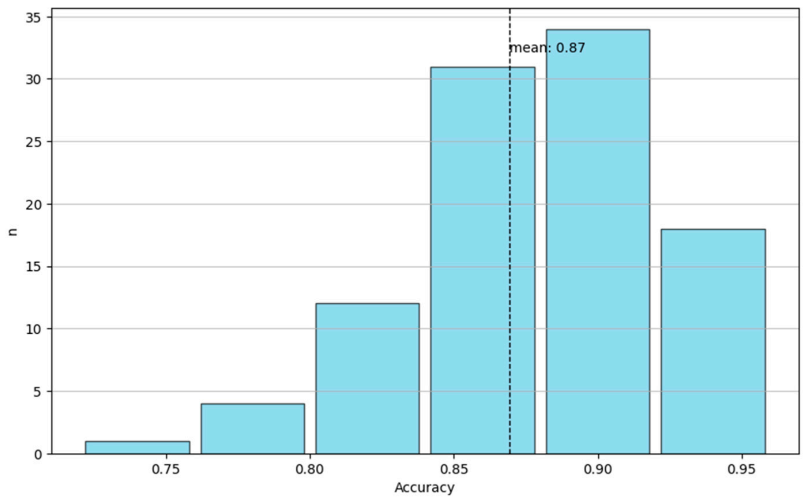

Figure 1. Histogram of the accuracy metric for the linear support vector machine model depending on the randomly selected test sample. Source: own elaboration.

The chart shows the distribution of accuracy for the linear support vector machine (SVM) algorithm in dynamic pricing for the e-commerce industry based on one hundred training processes (cross-validation mechanism). Most cases achieved accuracy in the ranges of 0.85–0.90 and 0.90–0.95, indicating the high effectiveness of the algorithm in most instances. The average accuracy was 0.87, as indicated by the vertical dashed line on the chart. A few training processes resulted in lower accuracy (0.75–0.80 and below 0.75), suggesting some variability in the results, but, overall, the algorithm demonstrated high and stable effectiveness in dynamic pricing. Figure 2 presents the confusion matrix for the linear support vector machine algorithm in dynamic pricing for the e-commerce industry.

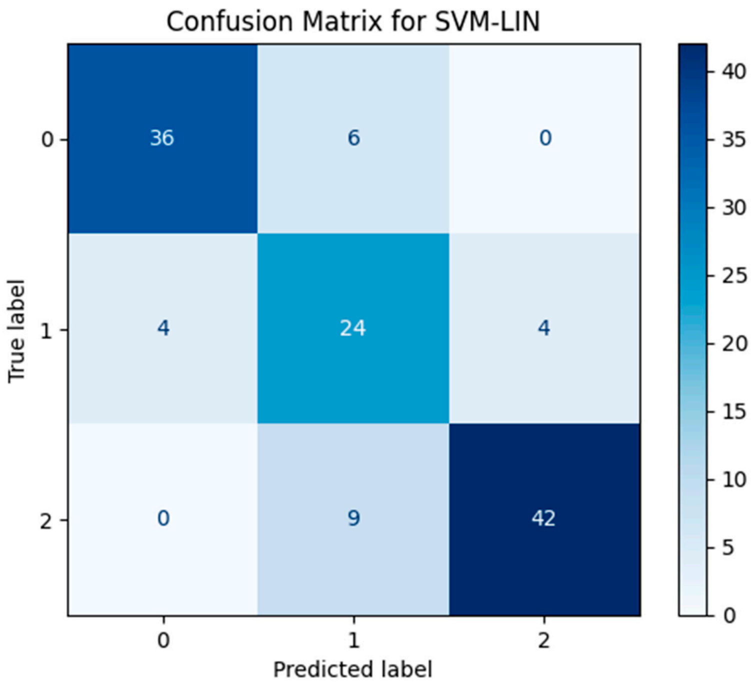

Figure 2. Confusion matrix for the linear support vector machine model in dynamic pricing for the e-commerce industry. Source: own elaboration.

The confusion matrix for the linear support vector machine (SVM-LIN) algorithm presented detailed classification results for three labels (0, 1, 2). The algorithm showed high effectiveness in correctly classifying labels 0 and 2, correctly classifying 36 and 42 cases, respectively. However, there were noticeable misclassifications, particularly for label 1, where 4 cases were incorrectly classified as label 0 and another 4 cases were incorrectly classified as label 2. Additionally, 9 cases of label 2 were incorrectly classified as label 1, indicating some difficulties for the algorithm in accurately distinguishing this class. These misclassifications suggested that the algorithm may have struggled to differentiate between more similar data in the context of label 1, impacting its overall accuracy. Nonetheless, the algorithm generally performed well in most cases, confirming its utility in dynamic pricing in the e-commerce industry.

5. Discussion, Implications, and Limitations

The results of this study demonstrate the strong performance of the proposed dynamic pricing model. Among the tested machine learning algorithms, the linear support vector machine (SVM) achieved the highest average accuracy (86.92%), indicating its effectiveness in classifying pricing decisions. This high accuracy underscores the suitability of the SVM for the dataset used and highlights the potential of the model to support real-time pricing decisions in e-commerce.

The success of the linear SVM can be attributed to its robustness in handling datasets with high-dimensional feature spaces and complex decision boundaries. The algorithm's ability to maximize the margin between classes makes it particularly effective for smaller datasets, like the one used in this study, where overlapping data points pose a challenge. The kernel trick used in nonlinear variants further enhances its capacity to model non-linear relationships, giving it an edge over simpler algorithms.

The dataset used in this study, consisting of 500 transactions from an e-commerce store, presented specific challenges for the machine learning models, including a relatively small sample size, a mix of numerical and categorical variables, and potentially overlapping or non-linear decision boundaries. These characteristics impacted the performance of the algorithms differently:

1. The Naive Bayes classifier operates on the assumption that all features are conditionally independent, given the target variable. While this simplification enables fast computations and scalability, it is rarely valid in real-world datasets. In this study, dependencies between features such as storage costs, demand dynamics, and price relative to competitors violated this assumption, leading to suboptimal classification accuracy. Furthermore, the model struggled with the small sample size, where the underlying feature distributions were not well represented.
2. Decision trees, while highly interpretable, are prone to overfitting, particularly with smaller datasets. Overfitting occurs when the model becomes too tailored to the training data, capturing noise instead of general patterns. Additionally, decision trees are sensitive to small variations in data, leading to unstable predictions. Despite their ability to handle mixed data types, their performance in this study was limited by the dataset's size and variability.
3. The k-nearest neighbor (k-NN) algorithm classifies instances based on their proximity to other data points, which requires meaningful distance metrics. In this dataset with mixed numerical, categorical, and ordinal features, defining appropriate distances proved challenging. Additionally, the relatively small dataset limited the algorithm's ability to form reliable neighborhoods, reducing its classification accuracy.

The results indicate that the proposed dynamic pricing model can effectively support managerial decision-making in e-commerce. The linear SVM algorithm's high accuracy demonstrates its utility in automating pricing decisions in real time, which is crucial in dynamically changing market conditions. By integrating criteria related to customers, enterprises, and market conditions, the model provides managers with a comprehensive tool for price management tailored to specific business needs.

The model's limitations primarily concern the quality and availability of data. Its effectiveness relies on large, well-balanced datasets that enable accurate algorithm training. A lack of sufficient data or uneven distribution (e.g., the class imbalance problem) may negatively affect prediction accuracy.

The model is scalable to larger datasets, making it applicable to larger e-commerce enterprises or more complex environments. However, such scalability comes with challenges:

1. Larger datasets require greater computational power and appropriate technological infrastructure to ensure efficient processing.
2. Optimization of algorithms may be needed, such as reducing data dimensionality or employing parallel or distributed machine learning techniques.
3. Larger datasets may introduce noise or heterogeneity, necessitating careful data preparation and cleaning.

This study's findings are based on the analysis of 500 transactions from a single e-commerce store. While sufficient for validating the model's concept, this limited scope poses challenges for generalizing the results. Data from a single source do not capture the diversity of market conditions encountered across different industries or regions. The data collection process relied on the management system of the selected online store, and transaction labels were assigned by experts. Although standard for early-stage research, this introduces potential biases, such as subjective label assignments or limitations stemming from the technology used.

Future research will aim to expand the dataset to include data from various industries, regions, and diverse market structures. This approach will provide a more comprehensive evaluation of the model's effectiveness under different conditions, enhancing the generalizability of results. Additionally, detailed descriptions of data collection processes and analyses of potential biases will be essential to further improve the model's reliability.

A critical aspect of dynamic pricing models is their alignment with ethical principles to ensure fairness and transparency. One possible mechanism to address these concerns involves implementing clear rules that minimize price discrimination and protect vulnerable consumer groups.

1. Algorithm Transparency—pricing models should be designed transparently, with explicitly defined criteria for price adjustments that are accessible and understandable to consumers. This prevents unintentional or unjustified price discrimination.
2. Price Ceilings and Floors—establishing maximum and minimum price thresholds can prevent excessive exploitation of temporary market imbalances, such as during crises where demand exceeds supply.
3. Rules for Vulnerable Groups—the model should account for discounts or special pricing rules for specific groups, such as seniors, students, or low-income individuals, to avoid market exclusion.
4. Monitoring and Audit Mechanism—regular reviews of the algorithm, including independent audits, can ensure compliance with ethical standards and reduce the risk of abuse.

These considerations emphasize the need for a balance between dynamic pricing efficiency and societal expectations of fairness. Integrating such mechanisms into the proposed model could enhance its acceptance and practical applicability.

6. Conclusions

This study contributes to the advancement of dynamic pricing in the e-commerce sector by presenting a novel machine learning-based approach that enhances decision-making processes. The model integrates multiple perspectives, including customer preferences, enterprise objectives, and market conditions, to optimize pricing strategies. By leveraging machine learning techniques such as regression and classification algorithms, the model analyzes extensive datasets encompassing customer behavior, purchase history, competitor pricing, and internal factors like inventory levels. This approach not only supports personalized pricing decisions but also provides a practical decision support system for managers aiming to balance profitability and competitiveness.

The empirical results confirm the effectiveness of the proposed model in improving pricing accuracy and revenue optimization. In testing scenarios with real-world data from an e-commerce platform, the model achieved an average prediction accuracy of over 85% in determining optimal pricing adjustments. These adjustments led to a measurable increase in revenue, with test cases showing revenue growth of 12% on average compared to static pricing strategies. Additionally, the model demonstrated significant potential in mitigating inventory issues by dynamically adjusting prices based on stock levels, reducing instances of overstocking and understocking by 18%. Personalized pricing recommendations aligned closely with customer preferences, resulting in improved conversion rates, with observed increases in purchase likelihood of 15% for targeted customer segments.

The analysis reveals that while the model is effective, its reliance on extensive data presents challenges, particularly for smaller businesses with limited access to customer and market information. Furthermore, its sensitivity to rapid changes in consumer preferences and market conditions suggests the need for continuous updates and recalibration to maintain accuracy and relevance. For instance, scenarios with abrupt shifts in demand due to external factors (e.g., seasonal trends or economic disruptions) required retraining the model to sustain performance levels.

From a theoretical perspective, this research contributes to the growing body of knowledge on the intersection of machine learning and dynamic pricing, offering insights into the application of advanced computational techniques to optimize pricing strategies. It highlights the potential of integrating personalized approaches with broader market analyses to address the complexities of modern e-commerce environments. Practically, this study underscores the importance of data-driven decision-making tools for managers, providing actionable frameworks to implement machine learning in pricing strategies effectively.

Future research could explore the integration of advanced machine learning techniques such as deep learning or reinforcement learning to enhance the model's adaptability and precision further. Additionally, investigating the influence of external macroeconomic factors, such as fiscal policies or legal regulations, on pricing dynamics could provide valuable insights. Another promising avenue lies in developing models capable of operating under limited data conditions, making dynamic pricing accessible to smaller e-commerce businesses. Understanding the long-term impact of dynamic pricing on customer loyalty and brand perception also remains a critical area for further study, offering a holistic view of the implications of this pricing strategy.

Author Contributions: Conceptualization, M.N. and M.P.-N.; Methodology, M.N. and M.P.-N.; Software, M.N.; Formal analysis, M.N.; Writing—review & editing, M.P.-N.; Visualization, M.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Poznań University of Technology grant number 0813/SBAD/2985.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kotler, P.; Armstrong, G. *Principles of Marketing*; Pearson Education: London, UK, 2010.
2. Shin, D.; Vaccari, S.; Zeevi, A. Dynamic pricing with online reviews. *Manag. Sci.* **2023**, *69*, 824–845. [CrossRef]
3. Douma, M.J. *Dynamic Pricing in History, Economics, and Ethics*; SAGE Publications: New York, NY, USA, 2018.
4. Luo, Y.; Sun, W.W.; Liu, Y. Distribution-free contextual dynamic pricing. *Math. Oper. Res.* **2024**, *49*, 599–618. [CrossRef]
5. Kopalle, P.K.; Pauwels, K.; Akella, L.Y.; Gangwar, M. Dynamic pricing: Definition, implications for managers, and future research directions. *J. Retail.* **2023**, *99*, 580–593. [CrossRef]

6. Deksnyte, I.; Lydeka, Z. Dynamic pricing and its forming factors. *Int. J. Bus. Soc. Sci.* **2012**, *3*, 213–2200.
7. Neubert, M. A systematic literature review of dynamic pricing strategies. *Int. Bus. Res.* **2022**, *15*, 1–17. [CrossRef]
8. Chen, N.; Gallego, G. Welfare analysis of dynamic pricing. *Manag. Sci.* **2019**, *65*, 139–151. [CrossRef]
9. Robinson, B.; Lakhani, C. Dynamic price models for new-product planning. *Manag. Sci.* **1975**, *21*, 1113–1122. [CrossRef]
10. Wittman, M.D.; Belobaba, P.P. Dynamic pricing mechanisms for the airline industry: A definitional framework. *J. Revenue Pricing Manag.* **2019**, *18*, 100–106. [CrossRef]
11. Abrate, G.; Fraquelli, G.; Viglia, G. Dynamic pricing strategies: Evidence from European hotels. *Int. J. Hosp. Manag.* **2012**, *31*, 160–168. [CrossRef]
12. Cohen, M.C.; Lobel, I.; Paes Leme, R. Feature-based dynamic pricing. *Manag. Sci.* **2020**, *66*, 4921–4943. [CrossRef]
13. Faruqui, A.; George, S.S. The value of dynamic pricing in mass markets. *Electr. J.* **2002**, *15*, 45–55. [CrossRef]
14. Zhao, G.; Ma, C.; Zhao, Q.; Li, J. Dynamic pricing of airline upgrade service considering revenue management. *J. Transp. Syst. Eng. Inf. Technol.* **2023**, *23*, 76.
15. Streviniotis, E.; Georgara, A.; Bistaffa, F.; Chalkiadakis, G. FairPlay: A Multi-Sided Fair Dynamic Pricing Policy for Hotels. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 25–27 March 2024; AAAI: Washington, DC, USA; pp. 22368–22376.
16. Suresh Kumar, S.; Margala, M.; Siva Shankar, S.; Chakrabarti, P. A novel weight-optimized LSTM for dynamic pricing solutions in e-commerce platforms based on customer buying behaviour. *Soft Comput.* **2023**, *38*, 1–13. [CrossRef]
17. Jin, F.L.; Tian, W.Q.; Wang, L. High-speed railway passenger flow equilibrium among trains of common lines based on travel behavior analysis under dynamic pricing. *Transp. Lett.* **2024**, *16*, 27–42.
18. Narahari, Y.; Raju, C.V.L.; Ravikumar, K.; Shah, S. Dynamic pricing models for electronic business. *Sadhana* **2005**, *30*, 231–256.
19. Saharan, S.; Bawa, S.; Kumar, N. Dynamic pricing techniques for Intelligent Transportation System in smart cities: A systematic review. *Comput. Commun.* **2020**, *150*, 603–625.
20. Zhao, C.; Zhang, Y. Dynamic quality and pricing decisions in customer-intensive service systems with online reviews. *Int. J. Prod. Res.* **2019**, *57*, 5725–5748. [CrossRef]
21. Talón-Ballestero, P.; Nieto-García, M.; González-Serrano, L. The wheel of dynamic pricing: Towards open pricing and one-to-one pricing in hotel revenue management. *Int. J. Hosp. Manag.* **2022**, *102*, 103184.
22. Goli, F.; Haghighinasab, M. Dynamic pricing: A bibliometric approach. *Interdiscip. J. Manag. Stud.* **2022**, *15*, 111–132.
23. Nunan, D.; Di Domenico, M. Value creation in an algorithmic world: Towards an ethics of dynamic pricing. *J. Bus. Res.* **2022**, *150*, 451–460.
24. Faruqui, A. The ethics of dynamic pricing. In *Smart Grid*; Academic Press: Cambridge, MA, USA, 2012; pp. 61–83.
25. Seele, P.; Dierksmeier, C.; Hofstetter, R.; Schultz, M.D. Mapping the ethicality of algorithmic pricing: A review of dynamic and personalized pricing. *J. Bus. Ethics* **2021**, *170*, 697–719. [CrossRef]
26. Ayadi, N.; Paraschiv, C.; Rousset, X. Online dynamic pricing and consumer-perceived ethicality: Synthesis and future research. *Rech. Appl. En Mark.* **2017**, *32*, 49–70. [CrossRef]
27. Melnikov, O.N. Heuristic Rules for the Dynamic Pricing Problem. *Organ. Mark. Emerg. Econ.* **2023**, *14*, 436–457.
28. El Youbi, R.; Messaoudi, F.; Loukili, M. Machine Learning-driven Dynamic Pricing Strategies in E-Commerce. In Proceedings of the 2023 14th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 21–23 November 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–5.
29. Lombardi, C.; Picado-Santos, L.; Annaswamy, A.M. Model-based dynamic toll pricing: An overview. *Appl. Sci.* **2021**, *11*, 4778. [CrossRef]
30. Wang, B.; Bi, W.; Liu, H. Dynamic Pricing with Parametric Demand Learning and Reference-Price Effects. *Mathematics* **2023**, *11*, 2387. [CrossRef]
31. Liang, Y.; Hu, Y.; Luo, D.; Zhu, Q.; Chen, Q.; Wang, C. Distributed Dynamic Pricing Strategy Based on Deep Reinforcement Learning Approach in a Presale Mechanism. *Sustainability* **2023**, *15*, 10480. [CrossRef]
32. Cosguner, K.; Seetharaman, P.B. Dynamic pricing for new products using a utility-based generalization of the bass diffusion model. *Manag. Sci.* **2022**, *68*, 1904–1922. [CrossRef]
33. Ruan, J.; Liu, G.; Qiu, J.; Liang, G.; Zhao, J.; He, B.; Wen, F. Time-varying price elasticity of demand estimation for demand-side smart dynamic pricing. *Appl. Energy* **2022**, *322*, 119520. [CrossRef]
34. Den Boer, A.V. Dynamic pricing and learning: Historical origins, current research, and new directions. *Surv. Oper. Res. Manag. Sci.* **2015**, *20*, 1–18.
35. Sajjadi, M.S.; Bachem, O.; Lucic, M.; Bousquet, O.; Gelly, S. Assessing generative models via precision and recall. *arXiv* **2018**, arXiv:1806.00035.
36. Nowak, M.; Pawłowska-Nowak, M. Prediction of the type of organizational culture using machine learning approach. *Przegląd Organ.* **2023**, *3*, 264–272. [CrossRef]