
Huimin Jiang, Farzad Sabetzadeh and Chen Zhang

Abstract: In previous research on the development of the relationships between product attributes and customer satisfaction, the models did not adequately consider nonlinearity and the fuzzy emotions of customers in online reviews. Also, stable customer satisfaction was considered. However, customer satisfaction is changing with time rapidly, and a time-series analysis for customer satisfaction has not been conducted previously. To address these challenges, this study designed a novel methodology using adaptive neuro-fuzzy inference systems (ANFIS) in conjunction with Bi-objective particle swarm optimization (BOPSO) and sentiment analysis techniques. Sentiment analysis is employed to extract time-series customer satisfaction data from online reviews. Then, an ANFIS with the BOPSO method is proposed for the establishment of customer satisfaction models. In previous studies, ANFIS is an effective method to model customer satisfaction which can handle fuzziness and nonlinearity. However, when dealing with a large number of inputs, the modeling process may fail due to the complexity of the structure and the lengthy computational time required. Incorporating the BOPSO algorithm into ANFIS can identify the optimal inputs in ANFIS and effectively mitigate the inherent limitations of ANFIS. Using mobile phones as a case study, a comparison was performed between the proposed approach and another four approaches in modeling time-series customer satisfaction.

Keywords: time-series customer satisfaction; ANFIS; sentiment analysis; bi-objective particle swarm optimization

1. Introduction

Customer satisfaction surveys are a valuable tool for businesses to understand the needs and desires of their target audience. However, these surveys often lack the flexibility to capture the depth of emotional responses, as the respondents only can answer the questions that are provided in advance. Another limitation of surveys is the difficulty in obtaining time-series values of customer satisfaction. To assess changes in customer satisfaction over time, it is essential to conduct a series of surveys with respondents, which is a time-consuming process as it requires multiple iterations. The lengthy period to complete such surveys may lead to outdated or inaccurate insights due to evolving customer satisfaction. With the ubiquity of the internet, e-commerce has become the primary way for customers to shop, with online comments emerging as a potent form of word-of-mouth marketing. These reviews, born from real post-purchase experiences, not only provide a more accessible avenue for information dissemination, but also offer rich insights for analyzing customer satisfaction, the factors influencing satisfaction decisions, and the underlying mechanisms. Dividing online reviews into different periods is an effective method to collect time-series data for customer satisfaction. However, the research on modeling time-series customer satisfaction based on online comments is not found in previous studies. Previous research assumed that the customer satisfaction remains static in the modeling. Failing to consider the dynamic changes in customer satisfaction can potentially impact the accuracy of customer satisfaction predictions in the modeling process.
Therefore, it is imperative to incorporate techniques that capture and analyze the evolving trends of customer satisfaction to ensure more accurate and reliable results. Also, online reviews often involve ambiguity in the expression of customer sentiment, and the relationship between product attributes and customer satisfaction is nonlinear. Therefore, the modeling of customer satisfaction requires us to solve three critical issues, namely, dynamic customer satisfaction, fuzziness, and nonlinearity. However, the majority of existing approaches in modeling customer satisfaction are capable of resolving only one or two of the aforementioned issues, and cannot solve all three challenges in one modeling framework. In previous research, adaptive neuro-fuzzy inference system (ANFIS) is a popular method for modeling customer satisfaction. It integrates the strengths of artificial neural networks, which can address the nonlinearity, and fuzzy inference systems, which can capture the fuzziness in the modeling [1]. However, the modeling process of ANFIS can encounter challenges when dealing with a large number of inputs. The issues of the complexity of the neural network’s structure and the extensive computation time can pose a failure in the modeling process. To address the above challenges, this study proposes a new methodology that combines sentiment analysis and ANFIS with Bi-objective particle swarm optimization (BOPSO) for analyzing and modeling customer satisfaction based on online reviews. The method employs sentiment analysis on online reviews, thereby gathering time-series quantitative data on customer satisfaction. BOPSO is employed in ANFIS to search for the optimal inputs and simplify the model structure by removing unimportant inputs. PSO exhibits high stability and fast convergence speed, making it particularly well-suited for searching for the optimal solution [2]. Beiranvand and Mobasher-Kashani [3] discovered that the multi-objective PSO algorithm provides distinct advantages in generating association rules compared to other methods such as the multi-objective genetic algorithm, genetic association rules, and the multi-objective differential evolution algorithm.

This paper covers several parts, as follows: Section 2 describes related research; Section 3 details the proposed methodology for customer satisfaction modeling; Section 4 applies the new modeling method to an actual case analysis, evaluating its effectiveness and practicality; Section 5 analyzes the results of validation tests; and finally, Section 6 provides a comprehensive summary of the paper.

2. Related Works

The following parts describe the main sentiment analysis approaches and the previous studies on customer satisfaction modeling for product design.

2.1. Online Reviews in Product Design

Online reviews provide valuable insights into the strengths and weaknesses of a product from the perspective of real users. They offer a channel for customers to express their opinions, user experiences, and overall satisfaction with a product. By analyzing these reviews, product designers can gain a deeper understanding of customer satisfaction, needs, and pain points [4]. A data-driven product design methodology was to extract product features based on large-scale texts [5]. Kang and Zhou [6] proposed an unsupervised rule-based method to mine objective and subjective features from online comments to enable personalized recommendations. Jiang et al. [7] adopted fuzzy inference and fuzzy time-series methods for the determination and prediction of the importance weights of product features using the sentiment scores from customer’s comments. A Kansei texting mining method was used to extract and summarize customer affective responses and features from online product reviews [8]. Chiu and Lin [9] proposed a case-based approach to analyze online comments on a particular consumer product and mine consumer preferences by combining text mining and Kansei engineering. A target feature selection model was designed to determine the to-be-improved product features with the consideration of redesign lead time, engineering cost, and technical risk based on online reviews [10]. Rintyarna et al. [11] took into account the semantics of words in the supervised method of sentiment analysis to exact product features. To understand the
user experience in product design, Yang et al. [12] established a user experience base from online reviews to support customer-centered design. To analyze customer needs within a product ecosystem, a machine-learning approach was proposed by examining a substantial volume of online user-generated product reviews [13]. Joung and Kim [14] conducted the importance performance analysis based on the methods of latent Dirichlet allocation, aspect-based sentiment analysis, and a deep neural network for product attributes from online reviews. Cai et al. [15] introduced a product and user-oriented approach to elicit and prioritize product attributes using online comments and to detect their differences between users. Jin et al. [16] revealed affective customer needs by mining online reviews based on a Kansei-integrated Kano model. A linguistic pattern mining method was proposed to derive product functions from online product comments which were used to identify customer needs for improving product design [17]. Zhang et al. [18] combined the initial and supplementary reviews to improve customer requirement identification and product development based on text mining. An explainable neural network-based method was designed to extract comprehensive design strategies based on online user-generated data and to predict customer choices [19]. Wu et al. [20] proposed a decision support model through online reviews to analyze customer preferences and discuss product rankings in different regions.

2.2. Approaches for Modeling Customer Satisfaction

Customer satisfaction is a key indicator that can help measure the extent to which a product or service meets or exceeds customer expectations. If customers are pleased with their purchasing experience and feel satisfied with the value of the product, they are more likely to become loyal patrons, and the chance of that product being successful in the market will be higher. On the contrary, dissatisfied customers are not only prone to switching to rival offerings, but also may share negative experiences with other buyers, potentially damaging a company’s reputation. Understanding and modeling customer satisfaction can help businesses to identify the key factors and attributes affecting customer satisfaction and tailor or improve their products or service to increase the satisfaction level. It has a great impact on customer loyalty and the growth of businesses, and can enhance the companies’ competitiveness among competitors. Many methods exist for constructing the models that are utilized to establish the connection between product attributes and customer satisfaction. Yang et al. [21] adopted a belief rule-based method to represent the nonlinear relationships between product attributes and consumer satisfaction. However, the techniques mentioned above fail to capture the inherent uncertainty and fuzziness associated with customer satisfaction.

For the issue of fuzziness in the modeling of customer satisfaction, Kwong et al. [22] considered the fuzziness while modeling the relationships in quality function deployment, and introduced a generalized fuzzy least-squares regression for developing models based on fuzzy observations. Sener and Karsak [23] employed fuzzy linear regression to estimate the parameters of the functional relationship for translating customer needs into product attributes. Shirouh and Keramati [24] proposed an algorithm based on fuzzy regression and data envelopment analysis to model customer satisfaction for use in marketing and new product development. Kang et al. [25] employed the fuzzy weighted association rule mining method for developing the connection between product features and customer satisfaction to reduce the risk of product failure and improve customer satisfaction. Wang et al. [26] proposed an integrated rough-set theory, continuous fuzzy kano model, and fuzzy weighted-association rule-mining method to extract the relationship between customer needs and product morphological features. In previous studies, the approaches combining polynomial structure with fuzzy regression are used to deal with the nonlinear and fuzzy aspects of customer satisfaction modeling. Chan et al. [27] developed an intelligent fuzzy regression method to capture nonlinearity and fuzziness for affective customer satisfaction modeling. A method-based fuzzy regression and chaos optimization algorithm was proposed to generate a polynomial fuzzy customer satisfaction model for mobile phone product design [28]. A fuzzy stepwise regression method was proposed for
establishing customer satisfaction models in a tea-maker design [29]. The above existing research approaches typically utilize traditional data surveys to model customer satisfaction, assuming it to be a static concept.

At present, there are few studies combining the development of relationship between customer satisfaction and product attributes with online comments. Bi et al. [30] used an ensemble neural network and effect-based Kano model to model customer satisfaction based on online comments. Chan et al. [31] introduced a hybrid ensemble genetic programming algorithm to illustrate the relationship between the design attributes and dimensions of customer satisfaction based on online reviews. Yakubu et al. [32] designed a multigene genetic programming-based fuzzy regression method to generate explicit customer satisfaction models using online reviews. A probabilistic linguistic group decision-FlowSort methodology was proposed to model customer satisfaction through online comments [33]. However, customer satisfaction is a dynamic concept that can change over time due to various factors such as product updates, market trends, or changing customer expectations. In the above approaches, there is a limitation in considering dynamic customer satisfaction.

Based on the above illustration, three issues should be considered in modeling customer satisfaction, which are time-series customer satisfaction, fuzziness, and nonlinearity. However, no previous studies have been found to solve all the three problems in a single process of modeling. Therefore, research that deals with all of the above problems using information from online comments needs to be explored.

3. Proposed Methodology

The introduction of the proposed methodology is provided in the following parts, including the collecting of online reviews, sentiment analysis, and construction of the time-series customer satisfaction models based on the ANFIS with BOPSO method. A flowchart of the process of the proposed methodology is shown in Figure 1.

![Figure 1. A flowchart of the proposed methodology.](image)
3.1. Sentiment Analysis in Online Comments

Firstly, the appropriate sample products are chosen. Octopus collector is utilized to obtain online comments for the sample products from different sources and platforms. Octopus is a web crawler tool that can help crawl websites, capture review data, and save them to Excel files. According to the time duration of the crawled online reviews, the time interval for the time-series modeling is determined, and the collected online interviews are divided into different periods, which are the time-series online reviews. The process of performing sentiment analysis in online reviews involves the following steps:

In the first step, pre-processing is conducted on the text extracted from the reviews to remove noise and extraneous information, such as HTML tags, stop-words, and punctuation marks.

In the second step, the topic words are extracted using the MiniTagCloud tool, which can intelligently perform word segmentation, generate word clouds, and conduct sentiment analysis based on large texts. Based on the analysis results from this tool, the key phrases and words with high frequency of occurrence are selected for the determination of the dimensions of customer satisfaction. The ones with low frequency are removed.

In the third step, the selected phrases or words with similar meanings are summarized into one category using the K-means clustering approach. Using the phrases or words for mobile phone products as an example, “operation”, “lag”, “smoothness”, “system optimization”, “upgrade”, and “program running” can be summarized into one group, as they all relate to the “system performance”, which can be defined as one dimension of customer satisfaction.

In the fourth step, based on the above-determined categories, online reviews from different time intervals are imported into the MiniTagCloud tool for sentiment analysis. For each period, the polarity of each dimension of customer satisfaction (positive, negative, or neutral) and their emotional scores are determined, which are the time-series datasets representing the changing trend in customer satisfaction.

3.2. Constructing Customer Satisfaction Model Using an ANFIS with BOPSO

Customer satisfaction models are generated by utilizing the proposed ANFIS with the BOPSO approach based on the obtained time-series emotional scores and the product attributes settings. The BOPSO is employed to solve the bi-objective optimization problem, aiming to minimize the mean absolute percentage error (MAPE) and variance of errors (VoE) of the models. By obtaining the Pareto optimal solution, the optimal inputs of ANFIS are identified for generating the models.

3.2.1. Bi-Objective PSO Method

In the BOPSO algorithm, a “particle” refers to a potential optimal solution to the given optimization problem. The particle continuously adapts its flight state by dynamically adjusting its speed based on both its own flight experience and the flight experience of the swarm. Based on the objective functions, each particle is designated to a fitness set. A particle’s own current best position, \( p_i \), where \( 1 \leq i \leq m \), and \( m \) is the size of the particle swarm, is defined as the position possessing the best fitness set, \( PF_{\text{best}} \), which is its own flight experience. Simultaneously, each particle also references the globally best position, \( p_{G} \), which is identified to be the best value, \( GF_{\text{best}} \), in \( PF_{\text{best}} \). It is the global flying experience of the particles. The \( i^{\text{th}} \) particle’s speed is \( v_i = (v_{i1}, v_{i2}, \ldots, v_{id}) \), where \( 1 \leq d \leq D \) and \( D \) represents the total number of inputs in the datasets. The position of the \( i^{\text{th}} \) particle is denoted as \( x_i = (x_{i1}, x_{i2}, \ldots, x_{id}) \).

In this study, the output of ANFIS is denoted as \( y(t) \), which is the predicted value of customer satisfaction at the current period, \( t \). The inputs of ANFIS include the product attributes settings and the values of customer satisfaction at the historical periods that are represented as \( y(t-T), \ldots, y(t-2), y(t-1) \), and \( T \) is the number of the previous time intervals.

In the BOPSO, each particle’s position corresponds to an input arrangement for ANFIS, which represents a different structure for modeling customer satisfaction. Table 1 shows the designed structure of the particle. In ANFIS, inputs are identified based on the values...
of the elements in $x_i$. If the value is 1, the corresponding product attribute or customer satisfaction in previous periods is chosen as an input. Otherwise, it is disregarded.

Table 1. The designed particle structure.

<table>
<thead>
<tr>
<th>Product Attributes</th>
<th>Customer Satisfaction in the Historical Periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>...</td>
</tr>
<tr>
<td>0 or 1</td>
<td>0 or 1</td>
</tr>
</tbody>
</table>

For example, if the total number of inputs in the modeling is six, which involves three product attributes and customer satisfaction in the last three time periods, then the best particle position value is obtained as $x_i = (1, 1, 0, 1, 0, 1)$. The first three elements in $x_i$ denote the selection of product attributes, and the last three elements indicate the choice of customer satisfaction in the last three time periods, denoted as $y_{(t-3)}, y_{(t-2)}$ and $y_{(t-1)}$. Therefore, the first and second product attributes, $y_{(t-3)}$ and $y_{(t-1)}$, are determined to be the optimal inputs, while the third product attribute and $y_{(t-2)}$ are removed. The historical best position for the $i$th particle, $p_i = (p_{i1}, p_{i2}, \ldots, p_{id})$, is determined by the position that possesses the best values of the fitness set. The best position of the entire swarm is $p_g = (P_1, P_2, \ldots, P_g)$, $g \in \{1, 2, \ldots, m\}$. The value of $p_g$ indicates the optimal selection for the inputs of ANFIS. Based on the concept of PSO, the particle’s speed and position can be updated in each iteration [34]:

$$v_{id}^{k+1} = \omega v_{id}^{k} + c_1 r_1 (p_{id}^{k} - x_{id}^{k}) + c_2 r_2 (p_{gd}^{k} - x_{id}^{k})$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1}$$

where $k$ is the predefined iteration number in BOPSO; at the $k$th iteration, $v_{id}^{k}$ and $x_{id}^{k}$ are the $i$th particle’s speed vector and position vector, respectively; $\omega$ indicates the inertia weight that balances the exploitation and development abilities of the particles within the search space; $c_1$ and $c_2$ are equal to 2 as the learning factors; and $r_1$ and $r_2$ are randomly selected from the range $[0, 1]$.

To minimize the modeling errors, the two objective functions, MAPE and VoE, are used to form a bi-objective optimization model in BOPSO.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100$$

(3)

$$VoE = \frac{1}{n-1} \sum_{i=1}^{n} \left( \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100 - MAPE \right)^2$$

(4)

where $n$ represents the datasets number, and $y_i$ and $\hat{y}_i$ are the $i$th actual and predictive emotional scores of customer satisfaction, respectively.

Based on the values of $x_{id}^{k+1}$ and datasets, the fitness set of each particle, $OF = \{f_1, f_2\}$, where $f_1 = MAPE$ and $f_2 = VoE$, is obtained by calculating the two objective functions using Equations (3) and (4). For the Bi-objective optimization problem, the Pareto dominant theory is performed to identify the optimal solution. In the above minimum optimization issue, the following conditions, (5) and (6), should be satisfied to prove that one solution, $s_1$, dominates another solution, $s_2$. The solution $s_1$ is not worse than the solution $s_2$ in both $f_1$ and $f_2$, and $s_1$ is strictly better than $s_2$ in no less than one objective function.

$$f_i(s_1) \leq f_i(s_2), \text{ for all } i \in \{1, 2\}$$

(5)

$$f_j(s_1) < f_j(s_2), \text{ for some } j \in \{1, 2\}$$

(6)

If the optimization problem is about a maximization issue, the solution $s_2$ is dominated by $s_1$ if $f_i(s_1) \geq f_i(s_2)$, for all $f \in F$ and $f_j(s_1) > f_j(s_2)$, for at least one $f \in F$. Among all the
solutions, the one not dominated by any other solutions is referred to as the Pareto optimal solution. In each iteration, every particle compares its current position with its personal best position. If the solution at the current position is better than the personal best position, the personal best position, \( p_i \), is updated to be the current position. Next, the individual solutions in \( p_i \) are compared with one another. Based on the above two requirements, the best one is determined as the global best position, \( p_g \).

3.2.2. ANFIS Structure

As a result of \( p_g \), the optimal inputs of ANFIS are identified for modeling customer satisfaction. Figure 2 illustrates a typical ANFIS structure, consisting of one output and two inputs. Additionally, each input comprises two linguistic descriptions, namely, two membership functions.

![Figure 2: An example of an ANFIS structure.](image)

In the first layer, the membership function for the \( i \)th linguistic description of \( x_1 \) is denoted as \( \mu_i(x_1) \), the membership function of the \( j \)th linguistic description of \( x_2 \) is \( \lambda_j(x_2) \), where \( i = 1, 2 \) and \( j = 1, 2 \). The membership function in the triangular shape is applied in this study and is shown in (7), where \( (a_i, b_i, c_i) \) and \( (s_j, t_j, u_j) \) are the triangular fuzzy numbers.

\[
\mu_i(x_1) = \begin{cases} 
\frac{x_1 - a_i}{b_i - a_i} & \text{if } a_i \leq x_1 \leq b_i \\
\frac{c_i - x_1}{c_i - b_i} & \text{if } b_i \leq x_1 \leq c_i \\
0 & \text{otherwise} 
\end{cases} \quad \lambda_j(x_2) = \begin{cases} 
\frac{x_2 - s_j}{t_j - s_j} & \text{if } s_j \leq x_2 \leq t_j \\
\frac{u_j - x_2}{u_j - t_j} & \text{if } t_j \leq x_2 \leq u_j \\
0 & \text{otherwise} 
\end{cases}
\] (7)

In the second layer, \( \mu_i(x_1) \) with \( \lambda_j(x_2) \) are combined, and the total number of nodes is four. The outcome of this layer, \( w_{ij} \), is called the firing strength of each combination.

\[
w_{ij} = \mu_i(x_1)\lambda_j(x_2) \quad (\forall i = 1, 2; j = 1, 2)
\] (8)

For the third layer, the outputs are the normalized firing strength \( \tilde{w}_{ij} \).

\[
\tilde{w}_{ij} = \frac{w_{ij}}{W} \quad W = \sum_i \sum_j w_{ij} \quad (\forall i = 1, 2; j = 1, 2)
\] (9)

In the fourth layer, the fuzzy rules are described as follows:

\[
R_{ij} : \text{IF } x_1 \text{ is } \mu_i \text{ AND } x_2 \text{ is } \lambda_j, \text{ THEN } f_{ij} = p_{ij}x_1 + q_{ij}x_2 + r_{ij}
\] (10)

where \( p_{ij} \) and \( q_{ij} \) with \( r_{ij} \) are the parameters of the internal model \( f_{ij} \) of fuzzy rules \( R_{ij} \).
In the fifth layer, there is one node, being the output of ANFIS, and it is computed by summing all of the input signals together.

\[
\hat{y} = \sum_{i=1}^{2} \sum_{j=1}^{2} O_{ij} = \sum_{i=1}^{2} \sum_{j=1}^{2} \bar{w}_{ij} f_{ij} = \sum_{i=1}^{2} \sum_{j=1}^{2} \bar{w}_{ij} (p_{ij} x_1 + q_{ij} x_2 + r_{ij})
\] (11)

where \(\hat{y}\) is the predicted value of customer satisfaction in the current period.

### 3.3. Process of the Proposed Methodology

Based on the above descriptions, the procedure of the proposed methodology is arranged in the following steps.

The first step: Online comments of sample products are crawled, and sentiment analysis is performed as described in Section 3.1. The time-series data of customer satisfaction are obtained. The product attributes settings that are relevant to customer satisfaction are gathered. Therefore, the data sets for generating times series customer satisfaction models are ready.

The second step: For the proposed ANFIS with BOPSO approach, each particle’s position and speed are initialized in the appropriate ranges based on Table 1. Also, the initial settings for iteration number, search space dimension, swarm size, inertia weight, and learning factor are determined.

The third step: In the first iteration, ANFIS is utilized to model customer satisfaction according to Section 3.2.2. Then, based on (3) and (4), the values of the fitness set of each particle, including MAPE and VoE, are calculated and are defined as the initial individual best fitness set, \(PF_{best}\). Among the solutions in the \(PF_{best}\), the global best fitness set, \(GF_{best}\), is found using the Pareto dominant theory, as described in Section 3.2.1, and the index of the best particle is recorded. Its position is determined as the initial global optimal position, \(p_g\).

The fourth step: In the following iterations by \(k\) to \(k + 1\), the particle’s speed, \(v_{id}^{k+1}\), and position, \(x_{id}^{k+1}\), are adjusted based on Equations (1) and (2). Using the process in the third step, the fitness set, \(OF_{i}^{k+1}\), of the \(i\)th particle at the \((k + 1)\)th iteration is obtained based on the upgraded particles. According to the Pareto dominance theory, the \(PF_{best}\) of the \(i\)th particle and \(OF_{i}^{k+1}\) are compared. When \(PF_{best}\) is dominated by \(OF_{i}^{k+1}\), the value of \(PF_{best}\) is replaced by the value of \(OF_{i}^{k+1}\). Also, the individual optimal position of the \(i\)th particle will be updated to be \(p_i = x_{id}^{k+1}\). Among \(PF_{best}\), the Pareto dominance is again performed to define the global optimal fitness set, \(GF_{best}\), and record the index of the particle. Then, its position is used as the new global best position, \(p_{g}\).

The fifth step: When the preset iteration condition is reached, the algorithm is stopped. In the results, the values of \(p_{g}\) indicate the optimal inputs of ANFIS for modeling customer satisfaction, and \(GF_{best}\) are the values of MAPE and VoE corresponding to the developed models.

### 4. Implementation

For an in-depth study of the proposed methodology, this paper adopts a case study on mobile phone products. The selected samples involve ten popular mobile phone products currently available in the market, denoted A to J. The brands or models of the mobile phones are not restricted, but it is imperative to ensure that each sample has adequate number of online reviews spanning at least a 2-year timeline to guarantee that sufficient information about time-series customer satisfaction can be extracted. The timeline of all the ten samples is consistent. Through an extensive collection of information on social media platforms such as TikTok and Weibo, 81,176 published customer reviews about these products were accumulated. To facilitate the analysis, these reviews were divided according to the timeline, which is 2 years, with every 6 months as a phase, forming a total of four time periods. For example, in time interval sequence 1, the number of comments and opinions for products A to J were 1181, 1964, 1405, 2426, 4818, 1663, 1483, 2858, 1422, and 643, respectively. Subsequently, the Excel tool is used to store the collated review data.
For each period, the sentiment analysis is performed on ten mobile phone products. Following the steps described in Section 3.1, the pre-processing is first conducted to get the clean text using the MiniTagCloud tool. The topic words are then obtained, and the ones with the high frequency of occurrence are selected, such as “program running”, “smoothness”, “appearance design”, “color”, “durable”, “picture”, and so on. The synonymous words and phrases are categorized using a K-means method. Consider the following example: words or phrases such as “screen display”, “pixel”, “night shot”, “portrait”, “photo taking”, and “dual shot” are categorized into the photo quality category, which is one important customer satisfaction metric for mobile phone products. In total, four major customer satisfaction metrics, system performance, appearance design, battery, and photo quality, are identified. Finally, the emotional scores of the four customer satisfaction in the four time intervals are obtained through sentiment analysis. Table 2 provides some examples of reviews for categories of system performance and photo quality, as well as their sentiment polarity and scores.

Using the customer satisfaction metric photo quality as an example for the modeling, Table 3 shows the obtained sentiment scores in the four time periods. Five product attributes that are related to photo quality are identified, which are main screen size (inch), maximum brightness (nits), rear camera pixel quantity (megapixel), front camera pixel quantity (megapixel), and ultra-wide camera pixel quantity (megapixel). Their setting information is collected from the official website of the mobile phone products, and is provided in Table 4.

<table>
<thead>
<tr>
<th>Examples of Online Comments</th>
<th>Sentiment Polarity</th>
<th>Sentiment Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>“The system of this mobile phone is really smooth, it doesn’t lag at all when you use it, and it operates smoothly, really satisfied!” “After the mobile phone system update, I feel that the functions are richer and the interface is more beautiful, I really like it.” “The optimisation of the mobile phone system is really well done, the power consumption is much slower, the range is greatly enhanced, it’s really great!”</td>
<td>Positive</td>
<td>1.660</td>
</tr>
<tr>
<td>The mobile phone system feels okay, there are no too prominent advantages, but there are no obvious shortcomings, it’s kind of middle of the road.” “For this mobile phone system, I think it’s okay, the functions are all complete, it’s just that the interface design is a bit ordinary, I hope that next time it can be improved.”</td>
<td>Neutral</td>
<td>0.980</td>
</tr>
<tr>
<td>“The mobile phone system is really laggy, often lagging, very unsmooth to use, really disappointed.” “The system update of this mobile phone is too frequent, each update takes up a lot of time and traffic, and there are some new features I don’t need at all.” “The battery optimisation of the phone system is not well done, it consumes too much power, and it’s really inconvenient to charge it several times a day.”</td>
<td>Negative</td>
<td>−0.037</td>
</tr>
</tbody>
</table>
Table 2. Cont.

<table>
<thead>
<tr>
<th>Examples of Online Comments</th>
<th>Sentiment Polarity</th>
<th>Sentiment Score</th>
</tr>
</thead>
</table>
| "Photo quality

"The camera on this phone is just brilliant, the pictures are so clear and the colours are so vibrant that I get a satisfactory picture every time I take a picture." "The night mode on the phone’s camera is really great, it takes beautiful photos even in bad lighting, I really like it." "The camera is very powerful with a variety of shooting modes to choose from, and it focuses quickly, the photo experience is great."

"The mobile phone camera is not bad, the photo effect is average, not too stunning feeling, but there is no obvious shortcomings." "The camera function is relatively complete, but I feel that there is no special outstanding place, use it moderately." "The photo effect is okay, is sometimes a little slow to take pictures, need to wait a while to take pictures."

"The phone camera is really bad at taking pictures, the photos are blurry and the colours are distorted, it’s very disappointing to use." "The camera has many features, but many of them are not practical, and there is often lag when taking pictures, which is really annoying." "The focus speed of the mobile phone camera is so slow that it often fails to capture the desired moment, and the photo taking effect in low light environment is also very poor." |
| Positive | 0.633 |
| Neutral | 0.873 |
| Negative | −0.096 |

Table 3. Photo quality: The sentiment scores in the four periods.

<table>
<thead>
<tr>
<th>Sentiment Scores in the Four Time Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile Phone</td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
<tr>
<td>F</td>
</tr>
<tr>
<td>G</td>
</tr>
<tr>
<td>H</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>J</td>
</tr>
</tbody>
</table>

Based on the above datasets, the proposed approach, ANFIS with BOPSO, is utilized to construct relationships between customer satisfaction in the current period, \(y_{(t)}\), with product attributes \(x_1 \sim x_5\), and the customer satisfaction in the historical time intervals \(y_{(t-3)}\), \(y_{(t-2)}\), and \(y_{(t-1)}\). In BOPSO, the search space dimension is 8, which is the sum of the five product attributes and the three previous periods of customer satisfaction. Several operations aiming at the least program running time and high prediction accuracy were performed, and the parameter settings of the iteration number and the swarm size are 20 and 5. Using validation 1 as an example, mobile phones A and B are used as the
validation data, while products C to J are used as training data. In ANFIS with BOPSO, the number of membership functions of each input was set to three, and the training epoch number was set to five. The inertia weight $\omega$ is equal to 0.9. $c_1$ and $c_2$ are set as 2. $r_1$ and $r_2$ are random values from the interval of [0, 1]. The searching ranges of speed and position of particles were [0, 0.5] and [0, 1], respectively. Time-series customer satisfaction models were established by utilizing MATLAB 2023a software. The optimal solutions obtained for the dimension of photo quality, along with the respective values of MAPE and VoE using the training data, are provided in Table 5. Among all the solutions, the seventh solution has the smallest training MAPE and VoE, which is the Pareto optimal solution. Therefore, in validation 1, $x_1$, $y(t-3)$, and $y(t-1)$ were determined as the inputs of ANFIS for modeling customer satisfaction photo quality and predicting the sentiment score, $y(t)$.

Table 4. The settings of the five product attributes.

<table>
<thead>
<tr>
<th>Mobile Phone</th>
<th>Main Screen Size $x_1$</th>
<th>Maximum Brightness $x_2$</th>
<th>Rear Camera Pixel $x_3$</th>
<th>Front Camera Pixel $x_4$</th>
<th>Ultra-Wide Camera Pixel $x_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6.82</td>
<td>2000</td>
<td>5000</td>
<td>3200</td>
<td>5000</td>
</tr>
<tr>
<td>B</td>
<td>6.78</td>
<td>1200</td>
<td>10,800</td>
<td>3200</td>
<td>800</td>
</tr>
<tr>
<td>C</td>
<td>6.10</td>
<td>2000</td>
<td>4800</td>
<td>1800</td>
<td>1200</td>
</tr>
<tr>
<td>D</td>
<td>6.80</td>
<td>1600</td>
<td>5000</td>
<td>3200</td>
<td>5000</td>
</tr>
<tr>
<td>E</td>
<td>6.36</td>
<td>1900</td>
<td>5400</td>
<td>3200</td>
<td>1200</td>
</tr>
<tr>
<td>F</td>
<td>6.67</td>
<td>2600</td>
<td>5000</td>
<td>2000</td>
<td>2000</td>
</tr>
<tr>
<td>G</td>
<td>6.78</td>
<td>1300</td>
<td>5000</td>
<td>3200</td>
<td>1200</td>
</tr>
<tr>
<td>H</td>
<td>6.74</td>
<td>1600</td>
<td>5000</td>
<td>1600</td>
<td>800</td>
</tr>
<tr>
<td>I</td>
<td>6.78</td>
<td>1800</td>
<td>5000</td>
<td>1600</td>
<td>800</td>
</tr>
<tr>
<td>J</td>
<td>7.60</td>
<td>1750</td>
<td>5000</td>
<td>1000</td>
<td>5000</td>
</tr>
</tbody>
</table>

Table 5. The optimal solutions in the validation 1.

<table>
<thead>
<tr>
<th>Optimal Solutions</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
<th>$y(t-3)$</th>
<th>$y(t-2)$</th>
<th>$y(t-1)$</th>
<th>MAPE</th>
<th>VoE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0275</td>
<td>0.0007</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0117</td>
<td>0.0002</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.0001</td>
<td>1.6427 $\times 10^{-8}$</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.2468 $\times 10^{-5}$</td>
<td>6.4924 $\times 10^{-9}$</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1.3362 $\times 10^{-6}$</td>
<td>2.6186 $\times 10^{-12}$</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>7.1471 $\times 10^{-7}$</td>
<td>3.2273 $\times 10^{-13}$</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>6.1730 $\times 10^{-8}$</td>
<td>2.3573 $\times 10^{-15}$</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4.6991 $\times 10^{-6}$</td>
<td>1.1399 $\times 10^{-11}$</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.0001</td>
<td>1.7467 $\times 10^{-8}$</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6.5646 $\times 10^{-5}$</td>
<td>1.3417 $\times 10^{-9}$</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.5684 $\times 10^{-5}$</td>
<td>5.2629 $\times 10^{-10}$</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2.8770 $\times 10^{-7}$</td>
<td>7.9218 $\times 10^{-14}$</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6.4812 $\times 10^{-8}$</td>
<td>2.2775 $\times 10^{-15}$</td>
</tr>
</tbody>
</table>
Three inputs in ANFIS were identified, and each input had three membership functions. Therefore, 27 fuzzy rules were generated by using (10), and 6 examples are given as follows:

\[
R_{111} : \text{If } x_1 \text{ is in1mf1, } y_{(t-3)} \text{ is in2mf1 AND } y_{(t-1)} \text{ is in3mf1,} \\
\text{Then } f_{111} = 0.0406x_1 + 0.0091y_{(t-3)} + 0.0074y_{(t-1)} + 0.0063
\]

\[
R_{112} : \text{If } x_1 \text{ is in1mf1, } y_{(t-3)} \text{ is in2mf1 AND } y_{(t-1)} \text{ is in3mf2,} \\
\text{Then } f_{112} = 0.0339x_1 + 0.0070y_{(t-3)} + 0.0072y_{(t-1)} + 0.0051
\]

\[
R_{113} : \text{If } x_1 \text{ is in1mf1, } y_{(t-3)} \text{ is in2mf1 AND } y_{(t-1)} \text{ is in3mf3,} \\
\text{Then } f_{113} = 0.0974x_1 + 0.0232y_{(t-3)} + 0.0407y_{(t-1)} + 0.0160
\]

\[
R_{121} : \text{If } x_1 \text{ is in1mf1, } y_{(t-3)} \text{ is in2mf2 AND } y_{(t-1)} \text{ is in3mf1,} \\
\text{Then } f_{121} = 0.1457x_1 + 0.0374y_{(t-3)} + 0.0244y_{(t-1)} + 0.0226
\]

\[
R_{122} : \text{If } x_1 \text{ is in1mf1, } y_{(t-3)} \text{ is in2mf2 AND } y_{(t-1)} \text{ is in3mf2,} \\
\text{Then } f_{122} = 0.0611x_1 + 0.0141y_{(t-3)} + 0.0125y_{(t-1)} + 0.0093
\]

\[
R_{123} : \text{If } x_1 \text{ is in1mf1, } y_{(t-3)} \text{ is in2mf2 AND } y_{(t-1)} \text{ is in3mf3,} \\
\text{Then } f_{123} = 0.1956x_1 + 0.0466y_{(t-3)} + 0.0817y_{(t-1)} + 0.0321
\]

5. Validation

In order to comprehensively assess the proposed method, this study compares and evaluates the modeling results obtained through the proposed method with those obtained from fuzzy regression (FR), fuzzy least square regression (FLSR), ANFIS, and K-means-based ANFIS. In the last method, the K-means algorithm is utilized to identify the membership function of each input within the ANFIS framework. Five validation tests are arranged. In validations 1 to 5, the validation data are selected as the datasets of mobile phones A and B, C and D, E and F, G and H, as well as I and J, respectively. In each test, the remaining eight mobile phones are applied in the training process. When building the time-series customer satisfaction model for photo quality using ANFIS, it was found that it cannot be realized because eight inputs cause a complicated structure and excessive computational time. Another three approaches, including FR, FLSR, and K-means-based ANFIS, can model time-series customer satisfaction successfully. The parameter \( h \) in FR and FLSR was set to 0.1 and 0.99, respectively, as these settings lead to the smallest modeling errors after a number of experiments by using different values in the range of [0, 1]. ANFIS was set up with three clusters using the K-means algorithm. In Section 4 of this paper, the parameter settings for the proposed method are described. In the proposed approach, the determined optimal inputs for ANFIS in tests 1 and 4 were \( x_1, y_{(t-3)}, \) and \( y_{(t-1)}, \) while the inputs in test 2 were \( y_{(t-3)} \) and \( y_{(t-2)}, \) and the inputs in tests 3 and 5 were \( x_1, y_{(t-2)}, \) and \( y_{(t-1)}). \) In Tables 6 and 7, the values of MAPE and VoE derived from the validation datasets in the five validation tests by adopting four approaches, and their mean values, are described.

<table>
<thead>
<tr>
<th>Validation</th>
<th>FR</th>
<th>FLSR</th>
<th>K-Means-Based ANFIS</th>
<th>The Proposed Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.1753</td>
<td>1.4437</td>
<td>0.7861</td>
<td>0.5468</td>
</tr>
<tr>
<td>2</td>
<td>2.2153</td>
<td>1.3491</td>
<td>0.3605</td>
<td>0.0723</td>
</tr>
<tr>
<td>3</td>
<td>5.6298</td>
<td>2.2849</td>
<td>0.0448</td>
<td>0.0132</td>
</tr>
<tr>
<td>4</td>
<td>0.4539</td>
<td>0.4421</td>
<td>0.2641</td>
<td>0.1297</td>
</tr>
<tr>
<td>5</td>
<td>0.4601</td>
<td>0.4349</td>
<td>0.4079</td>
<td>0.2333</td>
</tr>
<tr>
<td>Mean MAPE</td>
<td>2.9869</td>
<td>1.1910</td>
<td>0.3727</td>
<td>0.1991</td>
</tr>
</tbody>
</table>

Based on the modeling results, it can be concluded that all four approaches can effectively address the inherent fuzziness of the models. The models based on the K-means-based-ANFIS and the proposed approach demonstrate their capabilities in modeling nonlinearity. Furthermore, the values of MAPE and VoE of the validation datasets
Table 7. The values of VoE in the five validation tests.

<table>
<thead>
<tr>
<th>Validation</th>
<th>FR</th>
<th>FLSR</th>
<th>K-Means-Based ANFIS</th>
<th>The Proposed Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40.1228</td>
<td>2.1718</td>
<td>0.7342</td>
<td>0.0036</td>
</tr>
<tr>
<td>2</td>
<td>0.0259</td>
<td>0.4747</td>
<td>0.1375</td>
<td>0.0022</td>
</tr>
<tr>
<td>3</td>
<td>4.5814</td>
<td>0.8213</td>
<td>0.0001</td>
<td>7.0616 × 10⁻⁵</td>
</tr>
<tr>
<td>4</td>
<td>0.0346</td>
<td>0.0578</td>
<td>0.0074</td>
<td>0.0005</td>
</tr>
<tr>
<td>5</td>
<td>3.9425 × 10⁻⁵</td>
<td>0.0008</td>
<td>0.0283</td>
<td>3.2072 × 10⁻⁵</td>
</tr>
<tr>
<td>Mean VoE</td>
<td>8.9529</td>
<td>0.7053</td>
<td>0.1815</td>
<td>0.0013</td>
</tr>
</tbody>
</table>

6. Conclusions

Based on traditional surveys, gathering time-series data on customer satisfaction is quite challenging. The respondents for the questionnaires are required to express their satisfaction with the same products in different periods. The process of data collection is hard to conduct and takes a long time. Also, questionnaires are pre-set with specific questions, limiting customers to expressing opinions based on those available options. As a result, the data collected from traditional surveys lack the ability to capture significant emotional expressions from customers. In comparison, online comments have the advantage of containing a multitude of sentimental opinions from customers, involving feedback from the user experience or suggestions for improving the product design. By conducting sentiment analysis on online comments, it is possible to segment them into different periods and obtain time-series data on customer satisfaction. This allows for tracking the changing sentiments and perceptions of customers over time. To address the fuzziness and nonlinearity in the modeling, as well as to solve the modeling problem of ANFIS, which will fail because of the complexity issue with a large number of inputs, a new methodology is introduced, involving sentiment analysis on the divided periods of comments and an ANFIS with a BOPSO approach to develop customer satisfaction models. This study carried out the application case analysis on mobile phone products. The sentiment analysis provided valuable insights into the emotional polarity and scores associated with different customer satisfaction. Furthermore, the division of the comments into different time intervals enabled us to observe changes in customer satisfaction over time. A comparison is made between the proposed modeling approach with FR, FLSR, ANFIS, and K-means-based ANFIS approaches. The ANFIS with BOPSO approach offers an effective solution for addressing the challenges of fuzziness and nonlinearity in the models. The validation results demonstrate the superior performance of the proposed approach in modeling time-series customer satisfaction, as evidenced by lower MAPE and VoE values compared to other methods. In this study, the use of time-series customer satisfaction modeling in the design of the mobile phone products is described. In fact, the proposed methodology can also be applied in various other consumer products, which meet the requirement of having a substantial number of online reviews spanning an appropriate lengthy timeline, reflecting changes in the customer satisfaction and obtaining the sentiment scores of customer satisfaction in continuous different time periods. The length of each time period can be determined by the specific requirements and objectives of the companies, or by the experience of the product designers.

In contrast to existing approaches in the literature, the proposed ANFIS with the BOPSO approach incorporates time-series customer satisfaction into the modeling process and enhances prediction accuracy. The developed models enable product development companies to predict future sentiment scores of customer satisfaction by inputting the
product attribute settings and the updated historical satisfaction scores. The evolving trends in consumer satisfaction for existing products in the market can be obtained, which helps companies make timely product adjustments. Additionally, the derived fuzzy rules provide direct insights into the effects of product attributes on customer satisfaction. Based on the developed models, product designers can leverage optimization algorithms to identify alternative product attributes that align with consumer requirements, which helps in designing new and more appealing products with an increased satisfaction level. The proposed methodology fills a gap in the existing literature by providing a method for modeling and predicting changing trends in customer satisfaction over time using online comments, and offers a new direction for research in the fields of new product development and customer behaviors. Researchers can explore different methods and techniques to enhance the accuracy and reliability of the models in predicting future consumer satisfaction.

In the proposed methodology, there exists some limitations. The setting of parameters in the proposed approach, as described in Section 3.3, is essential for achieving good prediction accuracy. However, the determination of the parameter settings currently relies on a trial-and-error method, which is a time-consuming process and needs the experience of researchers. Future research could employ advanced optimization algorithms to help determine the settings adaptively to enhance the overall performance of the proposed methodology. On the other hand, in this study, the product attributes primarily focus on the functional aspects of mobile phone design. However, the design elements related to the affective aspect, such as the shape and layout of the cameras in the phones, also have impact on the changes in customer satisfaction. In future studies, the integration of quality function deployment with affective design can be considered in the proposed methodology to provide more comprehensive suggestions for product improvement. Furthermore, an investigation of the intelligent optimization of time-series customer satisfaction models would also be conducted to systematically search for the best combination of attributes that maximizes customer satisfaction.

**Author Contributions:** Conceptualization, H.J.; Data curation, H.J. and C.Z.; Funding acquisition, H.J.; Methodology, H.J.; Project administration, F.S.; Validation, H.J.; Writing—original draft, H.J. and F.S.; Writing—review and editing, H.J. and F.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Faculty Research Grants (FRG) from Macau University of Science and Technology under Grant FRG-23-045-MSB.

**Data Availability Statement:** All datasets generated during the current study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**References**

2. Pang, J.; Li, X.; Han, S. PSO with Mixed Strategy for Global Optimization. *Complexity* **2023**, *1*, 7111548. [CrossRef]
34. Huang, W.; Liu, Y.; Zhang, X. Hybrid Particle Swarm Optimization Algorithm Based on the Theory of Reinforcement Learning in Psychology. Systems 2023, 11, 83. [CrossRef]